NON-ADDITIVE LINEAR MODELS: ESTIMABILITY AND EFFICIENT ESTIMATIONS OF INTERACTION

Bikas K. Sinha and Rita Saharay
Stat/Math. Division
Indian Statistical Institute
Calcutta, INDIA

Anis C. Mukhopadhyay
Computer Science Unit
Indian Statistical Institute
Calcutta, INDIA

Present address:
Dept. of Statistics
University of California
Santa Barbara, CA 93106

Key Words and Phrases: Tukey's non-additive model; block designs, optimal inference on interaction parameter(s).

ABSTRACT

The simplest block design is considered and Tukey's non-additive model $y_{ij} = \mu + \beta_i + \tau_j + \theta\beta_i\tau_j + \epsilon_{ij}$ is adopted, where the parameters μ, β_i, τ_j, have their usual significance. A meaningful definition of estimability of θ, the interaction parameter, under this general set-up is considered and it is observed that the parameter is *not* always estimable.

A simple and interesting characterization of the block designs providing estimation of θ is obtained. Next some aspects of optimality for inference on

θ are discussed. Some extensions of the above model are also considered and relevant results on estimability presented.

1. INTRODUCTION

A model incorporating non–additivity in case of two–way classified data ith only one observation per cell was first suggested by Tukey (1949). The model proposed by him assumes the form

$$y_{ij} = \mu + \beta_i + \tau_j + \theta\beta_i\tau_j + \epsilon_{ij} \text{ with } \Sigma\beta_i = \Sigma\tau_j = 0 \tag{1.1}$$

where θ is the interaction parameter or non–additivity parameter, the other notations/parameters having their usual significance. Tukey provided an estimator of θ as also a test for $\text{Ho} : \theta = 0$ based on the residuals

$e_{ij} = y_{ij} - \hat{\mu} - \hat{\beta}_i - \hat{\tau}_j$ where $\hat{\mu}$, $\hat{\beta}_i$, $\hat{\tau}_j$ are ordinary least squares (LS) estimates of the parameters under the usual additive model (i.e., without the non–additivity term). Later, Milliken and Graybill (1970) considered an extension of the general linear model, hinted at earlier by Scheffe (1959). Thus, for example, in a randomized block design (RBD) instead of just one term $\theta\beta_i\tau_j$ in (1.1) to describe non–additivity, one could consider a model of the type

$$y_{ij} = \mu + \beta_i + \tau_j + \theta_1 f_1 + \theta_2 f_2 + \dots + \theta_t f_t + \epsilon_{ij} \tag{1.2}$$

where $f_1, f_2, \dots, f_t$ are any known functions of β_i and/or τ_j . Again an F–test for $\text{Ho} : \underline{\theta} = 0$ has been derived under the usual assumption on the law of distribution of ϵ_{ij}'s .

It is somewhat surprising to note that the testing problems have been formulated and solved without any reference to estimation of $\underline{\theta}$ as such. Our concern in this paper is to initiate a study in the latter direction. Specifically, we adopt Tukey's non-additive model (1.1) in the set-up of a block design and write $y_{ij} = \mu + \beta_i + \sum_h \delta_{ij.h} \tau_h + \theta\beta_i (\sum_h \delta_{ij.h} \tau_h) + \epsilon_{ij}$ with

$$\Sigma\beta_i = \Sigma\tau_h = 0 \, , \; 1 \leq j \leq k_{di} \, , \; 1 \leq i \leq b \tag{1.3}$$

where k_{di} is the ith block size for a design d and $\delta_{ij\cdot h} = 1$ if hth treatment occurs in the jth plot of the ith block; $= 0$, otherwise.

In sections 2 and 3 we point out that the non-additivity parameter θ is *not* necessarily estimable (through an analysis of the residuals) for any arbitrary choice of the design. With this observation, we provide a simple characterization of designs allowing estimation of θ . For the general model of the type (1.2), in section 4, we make an investigation on estimability of the θ_i's , assuming special forms of the f_i's . Finally, in section 5, we try to develop reasonable optimality criteria and make a relative comparison of designs for efficient estimation of θ .

2. CONCEPT OF ESTIMABILITY OF θ UNDER GENERAL LINEAR MODEL

Following Milliken and Graybill (1970) (as also Kshirsagar (1983)), we may consider a general linear model

$$\underline{Y}_{N\times 1} = X_{N\times p}\, \underline{\eta}_{p\times 1} + F_{N\times k}\, \underline{\theta}_{k\times 1} + \underline{\epsilon} \tag{2.1}$$

$$E(\underline{\epsilon}) = 0 \text{ , Var } (\underline{\epsilon}) = \sigma^2 I_N$$

where $\underline{\eta}$ = vector of unknown (additive) parameters

$\underline{\theta}$ = vector of non-additivity parameters

F = matrix associated with non-additivity parameters.

The functional forms of the elements of F, $f_{ij}(\cdot)$ say, are known and in general, they are arbitrary functions of estimable parametric functions of $\underline{\eta}$ under simple linear model

$$\underline{Y} = X\underline{\eta} + \epsilon \tag{2.2}$$

Estimation of $\underline{\theta}$ from the model (2.1) as such is formidable as basically it is non-linear in $\underline{\eta}$. One may adopt the ad-hoc procedure of premultiplying (2.1) by $I - P_x$, the orthogonal projection operator of X , to get rid of $\underline{\eta}$ and convert (2.1) to one which describes in a sense a linear model involving $\underline{\theta}$ only. Thus denoting $(I - P_x)\underline{Y}$, $(I - P_x)F$, and $(I - P_x)\underline{\epsilon}$ by $\underline{Z}$, M , and $\underline{\epsilon}^*$ respectively, from the model (2.1) using the fact that $(I - P_x)X$ is null, we get,

$$\underline{Z} = M\underline{\theta} + \underline{\epsilon}^* \tag{2.3}$$

where $E(\underline{\epsilon}^*) = 0$, Var $(\underline{\epsilon}^*) = \sigma^2 (I - P_x) = \Sigma$ say. Assume momentarily that the matrix M (even though its elements involve the unknown parameter $\underline{\eta}$ through F) is completely known. Following Rao (1965, Chapter 4, section 4i.4) we get formally,

$$\hat{\underline{\theta}} = (M'\Sigma M)^{-} M'\Sigma Z \quad \text{as} \quad \mathcal{M}(\Sigma) \supseteq \mathcal{M}(M)$$
$$= [F' (I - P_X) F]^{-} F'(I - P_X)\underline{Y} \tag{2.4}$$

where $\mathcal{M}(\Sigma)$ is the column space of Σ .

In practice, for estimation and testing purpose, we replace the functions of $X\underline{\eta}$ in F by their blue $X\hat{\underline{\eta}}$, obtained under the model (2.2). In other words, we compute

$$\hat{\hat{\underline{\theta}}} = [F'_{(X\hat{\eta})} (I - P_X)F'_{(X\hat{\underline{\eta}})}]^{-} F_{(X\hat{\eta})} (I - P_X)Y \tag{2.5}$$

(here $F_{(X\hat{\underline{\eta}})}$ denotes F with $X\eta$ replaced by $X\hat{\underline{\eta}}$). At this stage, one might wonder as to whether $\underline{\theta}$ is as such estimable or not irrespective of the choice of $\underline{\eta}$ in the relevant parameter space. To settle this, we first incorporate the following formal definition of estimability of $\underline{\theta}$. For simplicity in notation, from now onwards we write $I - P_X$ and $I - P_{X_i}$ as $I - P$ and $I - P_i$ respectively.

Definition 1. Under model (2.1), the interaction parameter vector $\underline{\theta}$ is estimable iff

$$\text{Rank } [F'(I - P)F] = k$$

i.e. $F'(I - P)F$ is nonsingular for all choices of $\underline{\eta}$ in the relevant parameter space.

Then next theorems justify some intuitive feelings regarding estimability of $\underline{\theta}$.

Theorem 2.1. Suppose $\underline{\theta}$ is estimable (in the above sense) under the model (2.1). Then $\underline{\theta}$ is also estimable under an extended set-up with additional observations (involving the same $\underline{\eta}$ parameter as in (2.1)).

Proof. Clearly, it suffices to prove the result with one additional observation. Let the extended set-up be

$$\underline{Y}_0 = X_0\underline{\eta} + F_0\underline{\theta} + \underline{\epsilon}_0 \tag{2.6}$$

where

$$\underline{Y}_0 = \left(\frac{\underline{Y}}{y_0}\right), \ X_0 = \begin{pmatrix} X \\ \underline{L}' \end{pmatrix}, \ F_0 = \begin{pmatrix} F \\ \underline{g}' \end{pmatrix}, \ \underline{\epsilon}_0 = \begin{pmatrix} \underline{\epsilon} \\ \epsilon_0 \end{pmatrix} \sim N(0,\sigma^2 I) .$$

Let Θ be the relevant parameter space of $\underline{\eta}$. Now to ensure estimability of $\underline{\theta}$ under this extended set up we distinguish between two cases: $\underline{L} \in M(X')$ and $\underline{L} \notin M(X')$ and show that

$$\text{Rank}\,[\,(I - P_0)F_0] = k \text{ for all } \underline{\eta} \in \Theta$$
$$\text{whenever Rank}\,[\,(I - P)F\,] = k \text{ for all } \underline{\eta} \in \Theta$$

Case (i): $\underline{L} \notin M(X')$.

Then

a) $(X'X + \underline{L}\,\underline{L}')^-$ is a g-inverse of $X'X$.

b) $\underline{L}'(X'X + \underline{L}\,\underline{L}')^- X' = \underline{0}$.

c) $\underline{L}'(X'X + LL)^- L = 1$.

Below we sketch a brief outline of the proofs.

The identity

$$(X'X + \underline{L}\,\underline{L}')\,(X'X + \underline{L}\,\underline{L}')^{-}X'X = X'X \quad \text{[see Rao–Mitra (1971)]}$$

implies

$$X'X[I - (X'X + \underline{L}\,\underline{L}')^{-}X'X] = \underline{L}\,\underline{L}'(X'X + \underline{L}\,\underline{L}')^{-}X'X = 0$$

As $M(\underline{L}) \cap M(X') = \{0\}$, writing $\underline{L} = \underline{L}\,\underline{L}'D_1$ and $X' = X'XD_2$ for some D_1 and D_2, (a) and (b) follow immediately. Further (c) follows as a consequence of
$(X'X + \underline{L}\,\underline{L}')\,(X'X + \underline{L}\,\underline{L}')\underline{L}\,\underline{L}' = \underline{L}\,\underline{L}'$ and through an application of the result stated in (b).

Case (ii): $\underline{L} \in M(X')$.

Then writing $\underline{L} = X'X\underline{u}$ for some $\underline{u}$, we get

$$[I - P_0] = \begin{bmatrix} I - X\underline{u} \\ 0 \quad 1 \end{bmatrix} \begin{bmatrix} \dfrac{I - P \atop - u'X'}{1 + \underline{u}'X'X\underline{u}} & \dfrac{0 \atop 1}{1 + \underline{u}'X'X\underline{u}} \end{bmatrix} \tag{2.7}$$

As $\begin{bmatrix} I - X\underline{u} \\ 0 \quad 1 \end{bmatrix}$ is non-singular, and rank $[(I - P)F] = k$ for all $\underline{\eta} \in \Theta$, it is readily seen from (2.7) that

rank $[(I - P_0)F_0] = k$ for all $\underline{\eta} \in \Theta$.

Lemma 2.2. Suppose $\underline{\theta}$ is not estimable under the model (2.1). Then $\underline{\theta}$ is also not estimable under the extended set-up (2.6) when $\underline{L} = X'\underline{e}_i$ and $g = F'\underline{e}_i$, e_i being a $N \times 1$ vector with 1 in the ith position and 0 elsewhere. In the context of block designs, an applications of the above results simplifies verification of estimability of the interaction parameter (as a scalar) in Tukey's non-additive model and also of the vector parameter $\underline{\theta}$ in the generalized set-up.

3. ESTIMABILITY OF θ UNDER TUKEY'S MODEL IN A BLOCK DESIGN SET-UP

Suppose under the model (1.3) $\beta_i = 0$ for some i . This means that for the design d , for each of the k_{di} observations in the ith block the multiplicative interaction term would vanish and hence, no information on the non-additive parameter θ would be available from this ith block. A similar consideration applies to the set of observations under hth treatment if $\tau_h = 0$ for some h . Again $\beta_i = \beta_i{}'(\tau_h = \tau_h{}')$ would mean that we have to work effectively with a model where the treatment – block incidence matrix is suitably modified in the sense that the corresponding columns (rows) of the original incidence matrix are merged to form a new column (row). Therefore, the relevant parameter space can be and will be taken as

$$\Theta = \left\{ \begin{array}{l} (\mu , \underline{\tau} , \underline{\beta}) : \ -\infty < \mu < \infty , \underline{\beta}'\underline{1} = 0 , \ \underline{\tau}'\underline{1} = 0 , \\ \beta_i \neq \beta_i{}' \neq 0 \text{ for all } i \neq i' , 1 \leq i , i' \leq b , \\ \tau_h \neq \tau_h' \neq 0 \text{ for all } h \neq h' , 1 \leq h , h' \leq v \end{array} \right\} \tag{3.1}$$

Without any loss of generality, we will be working with connected designs only.

Arranging the observations serially blockwise, the model (1.3) can be rewritten as

$$\underline{Y} = [1 : X_\beta : X_\tau] [1 , \underline{\beta}' , \underline{\tau}']' + \theta \underline{f} + \underline{\epsilon} \tag{3.2}$$

where $\underline{1}$ is $N \times 1$ vector of all 1's ,

X_β is $N \times b$ coefficient matrix corresponding to the block effects,

X is $N \times v$ coefficient matrix corresponding to the treatment effects,

X is $N \times (b+v+1)$ coefficient matrix $[1 : X_\beta : X_\tau]$,

$\underline{\eta}$ is $(b+v+1) \times 1$ parameter vector $(\mu, \underline{\beta}', \underline{\tau}')'$,

$\underline{f} = ((\beta_i \Sigma \delta_{ij \cdot h} \tau_h))$ is the $N \times 1$ column vector associated with the interaction parameter θ.

Now, the question of estimability of θ through an analysis of the residuals $\underline{Z}$ (see (2.3)) is equivalent to the following: Is the column vector $(I - P)\underline{f}$ non-null for all $\underline{\eta} \in \Theta$ for *any* choice of the design matrix X? The answer is hopelessly in the negative as we will demonstrate shortly. In this context, the following result is useful.

Property 1: The form of $X = [1 : X_\beta : X_\tau]$ above indicates $(I-P)X_\beta = 0$, $(I-P)X_\tau = 0$. That is, in each row of $(I-P)$, the sum of entries corresponding to each block and each treatment is zero.

The following lemma demonstrates that it suffices to settle estimability of θ in a binary design. Let $N_d = ((n_{dhi}))$ be the treatment x block incidence matrix under an arbitrary non-binary design $d(v, b, k_{d1}, \ldots, k_{db})$.

Let d be reduced to $\bar{d}(v, b, k_{d1}, \ldots, k_{db})$ as follows:

$$\left.\begin{aligned} n_{dhi} &= 1 \text{ if } n_{dhi} \geq 1 \\ &= 0 \text{ otherwise} \end{aligned}\right\} \quad 1 \leq h \leq v,\ 1 \leq i \leq b .$$

(In our later discussions we will refer to such $\bar{d}$ as binary reduction of d).

Lemma 3.1. θ is estimable under d iff θ is estimable under $\bar{d}$.

Proof. We first note that connectedness of d retains connectedness of $\bar{d}$.

"If part." Follows as an immediate application of Theorem 2.1

"Only if part." We can construct d starting from $\bar{d}$ by adding the observations one by one for all treatments appearing more than once in a block. Now the proof follows by using Lemma 2.2.

The following Theorem provides a simple characterization of a large class of connected block designs ensuring estimability of θ.

Theorem 3.1. Under Tukey's model, applied to the block-design set-up (1.3), a connected block design $d(v, b, k_{d1}, k_{d2}, \ldots, k_{db})$ will provide unbiased estimation of θ whenever at least one pair of treatments (h, h') say, occur in two different blocks.

Proof. The proof is by contradiction. Suppose h and h' occur together in ith and i'th blocks. Let d_1 be the sub-design of d formed by only ith and i'th blocks and hth and h'th treatments with the incidence structure

$$n_{d_1hi} = n_{d_1h'i} = n_{d_1hi'} = n_{d_1h'i'} = 1$$

Now suppose θ is not estimable in the original design d. Then

$$[I - P]f = 0 \text{ for some } \eta_1 \in \Theta \qquad \ldots (3.3)$$

Writing X and $\underline{f}$ as $X = \binom{X_1}{X_2}$, $\underline{f} = \binom{\underline{f}_1}{\underline{f}_2}$ where X_1 and $\underline{f}_1$ correspond to the sub-design d_1, (3.3) implies

$$\binom{\underline{f}_1}{\underline{f}_2} \in M\binom{X_1}{X_2} \text{ for } \eta_1 \in \Theta$$

$$\Rightarrow \quad \underline{f}_1 \in M(X_1) \text{ for } \eta_1 \in \Theta$$

$$\Rightarrow \quad (I - P_1)\underline{f}_1 = \underline{0} \text{ for } \eta_1 \in \Theta .$$

As rank $(I - P_1) = 1$ = error d.f. using Property 1 one can develop any 1×4 row vector of $(I - P_1)$, say ith row vector as $a_i(1 - 1 | -1\ 1)$ where a_i's are real numbers, not all zero's. The sets of two positions each in the above partition correspond to treatments h and h′ in order applied to the blocks i and i′ respectively. Then a typical element of $(I - P_1)\underline{f}_1$ is $a_i(\beta_i - \beta_i{}')(\tau_h - \tau_h{}')$ and thus $(I - P_1)\underline{f}_1 = \underline{0}$ for any $\eta_1 \in \Theta$ implies $(\beta_i - \beta_i{}')(\tau_h - \tau_h{}') = 0$ which can only happen if for η_1, $\beta_i = \beta_{i'}$ and/or $\tau_h = \tau_h{}'$ holds. This leads to a contradiction in the description and coverage of the parameter space Θ (see (3..1)).

Theorem 3.2. Let $d(v, b, k_{d1}, k_{d2}, \ldots, k_{db})$ be a connected block design for which each elementary treatment contrast has a unique (unbiased) estimator under the simple linear model (2.2). Then θ is *not* estimable from d.

Proof. It is evident that under the given hypothesis, from the ith block of d, we obtain exactly $k_{di} - 1$ *distinct* independent elementary treatment contrasts. Again, any estimable treatment contrast is a linear function of the within block elementary treatment contrasts. Since d is a connected design, under the given hypothesis,

$$(k_{d1} - 1) + (k_{d2} - 1) + \ldots + (k_{db} - 1) = v - 1$$

$$\text{i.e. rank } (I - P) = (\Sigma k_{di} - 1) - (b - 1) - (v - 1) = 0 \ .$$

Corollary 3.3. Let the binary reduction $\bar{d}$ of a design d be such that the condition of Theorem 3.2 holds for $\bar{d}$. Then θ is *not* estimable under d

The proof is immediate from an application of Lemma 3.1.

Remark 1. There are still many designs which neither satisfy the condition of Theorem 3.1 (ensuring estimability of θ) nor do they satisfy the condition of Corollary 3.3 (leading to non-estimability of θ . For example, for a BIBD with $\lambda = 1$, estimability of θ could *not* be settled. However, we have a strong feeling (as indicated by several examples) that the "sufficient" condition stated in Theorem 3.1 will turn out to be "necessary" also. Below we demonstrate two examples in justification of our conjecture.

Example 1. BIBD $(b = v = 3, \lambda = 1)$.

Block	Varieties
1	1, 2
2	1, 3
3	2, 3

For this design, taking a choice of $\underset{\sim}{\beta}$ and $\underset{\sim}{\tau}$ as

$$\underset{\sim}{\beta} = (1, -\frac{2}{3}, -\frac{1}{3})', \text{ and } \underset{\sim}{\tau} = (1, 2, -3)', (I - P)\underset{\sim}{f} = \underset{\sim}{0} .$$

Example 2.

Block	Varieties
1	1, 2, 3
2	2, 4, 6
3	3, 4, 5
4	1, 5, 6

For this design, one can easily verify that $(I - P)\underline{f} = \underline{0}$ for the choice of $\underline{\beta}$ and $\underline{\tau}$ as $\underline{\tau} = (1, -1, 2, -4, 4, -2)'$, $\beta = (-11, 13, 1, -3)'$. Thus θ is not estimable under these two designs.

4. ESTIMATION OF INTERACTION – PARAMETER VECTOR $\underline{\theta}$ IN A GENERAL BLOCK DESIGN

In this section we take up a model of the type (1.2) in the context of general block design set–up and show that it is not generally true that all the components of $\underline{\theta}$ in the model of the type (1.2) are estimable through an analysis of the residuals.

Lemma 4.1. Consider the model

$$y_{ij} = \mu + \beta_i + \Sigma\delta_{ij.h}\tau_h + \theta_1\beta_i \sum_h \delta_{ij.h}\tau_h + \theta_2\beta_i^2 + \theta_3\sum_h \delta_{ij.h}\tau_h^2 + \epsilon_{ij} \qquad \text{...(4.1)}$$

Then for no block design θ_2 and θ_3 are estimable.

Proof. Writing the observations blockwise in order, (4.1) can be rewritten as

$$y = [1 : X_\beta : X_\tau] \begin{bmatrix} \mu \\ \underline{\beta} \\ \underline{\tau} \end{bmatrix} + [\underline{F}_1, \underline{F}_2, \underline{F}_3] \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} + \underline{\epsilon}$$

$$= X\underline{\eta} + F\underline{\theta} + \underline{\epsilon}$$

It can be easily seen that under this model $\underline{F}_2 = X_\beta \underline{\beta}^{(2)}$ and $\underline{F}_3 = X_\tau \underline{\tau}^{(2)}$ where the elements of $\underline{\beta}^{(2)}$ and $\underline{\tau}^{(2)}$ are β_i^2's and τ_h^2's respectively. Thus referring to Property 1, we observe that both the second and third columns of $(I - P)F$ are null vectors.

Scheffe (1959) developed a different argument to ascertain a similar result in the context of RBD set-up. As a matter of fact, in view of Property 1, it turns out that whatever the block design set-up adopted in a non-additive model of the type (2.1), the interaction parameters corresponding to higher powers of β_i and τ_h *alone cannot* be estimated. What if instead we introduce more terms involving both β_i and τ_h in (4.1)? Take, for example, the model involving up to second powers of β_i's and τ_h's, i.e.

$$y_{ij} = \mu + \beta_i + \Sigma\delta_{ij.h}\tau_h + \theta_1\beta_i\Sigma\delta_{ij.h}\tau_h + \theta_2\beta_i\Sigma\delta_{ij.h}\tau_h^2 + \theta_3\beta_i^2\Sigma\delta_{ij.h}\tau_h + \theta_4\beta_i^2\Sigma\delta_{ij.h}\tau_h^2 + \epsilon_{ij} \qquad \ldots(4.2)$$

Clearly, not all designs will provide estimation of $\underline{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4)'$ as we have seen in case of single interaction parameter. Following the same arguments as given in section 3, under (4.2) also, the relevant parameter space can be taken to be the same as ϕ and similar results as stated in Lemma 3.1, Theorem 3.3 and Corollary 3.3 also hold here.

The following theorem establishes estimability of $\underline{\theta}$ under RBD, and using this result the next theorem gives a sufficient condition under which $\underline{\theta}$ is estimable.

<u>Theorem 4.1.</u> An RBD with $b \geq 3$, and $v \geq 3$, provides unbiased estimation of θ under the model (4.2).

Proof. Since $F'(I - P)F$ is invariant under any rearrangement of treatments among the plots within a block, without loss of generality, we assume that in the RBD considered, the jth plot within any block receives the jth treatment $(j = 1, 2, \ldots, v)$. Thus, writing the observations block-wise and in order, $(I - P) = (I - \frac{J}{b}) \otimes (I - \frac{J}{v})$. Under (4.2), $[(I - P)F]$ is a $N \times 4$ matrix $(N = bv)$ and let its pth row $(p = (i - 1)b + j)$ corresponding to the observation in jth plot of ith block, be denoted by $\underline{\alpha}'_{ij}$. It is not difficult to verify that this typical row takes the form

$$\underline{\alpha}'_{ij} = \Big[\beta_i\tau_j,\ \beta_i(\tau_j^2 - \frac{1}{v}\sum_{h=1}^{v}\tau_h^2),\ \tau_j(\beta_i^2 - \frac{1}{b}\sum_{t=1}^{b}\beta_t^2),$$

$$(\beta_i^2 - \frac{1}{v}\sum_{t=1}^{b}\beta_t^2)(\tau_j^2 - \frac{1}{v}\sum_{h=1}^{v}\tau_h^2)\Big] .$$

Consider the following 9×4 submatrix A of $[(I - P)F]$ formed by the rows corresponding to $i, j = 1, 2, 3$. Let

$$A = \begin{bmatrix} \beta_1 & \beta_1^2 - \frac{1}{b}\Sigma\beta_\tau^2 \\ \beta_2 & \beta_2^2 - \frac{1}{b}\Sigma\beta_\tau^2 \\ \beta_3 & \beta_3^2 - \frac{1}{b}\Sigma\beta_\tau^2 \end{bmatrix} \otimes \begin{bmatrix} \tau_1 & \tau_1^2 - \frac{1}{v}\Sigma\tau_h^2 \\ \tau_2 & \tau_2^2 - \frac{1}{v}\Sigma\tau_h^2 \\ \tau_3 & \tau_3^2 - \frac{1}{v}\Sigma\tau_h^2 \end{bmatrix}$$

Suppose

$$B = \begin{bmatrix} 1 & \beta_1 & \beta_1^2 \\ 1 & \beta_2 & \beta_2^2 \\ 1 & \beta_3 & \beta_3^2 \end{bmatrix} \otimes \begin{bmatrix} 1 & \tau_1 & \tau_1^2 \\ 1 & \tau_2 & \tau_2^2 \\ 1 & \tau_3 & \tau_3^2 \end{bmatrix},$$

Here $\otimes$ denotes the Kronecker product. It is trivially known that the column vectors of B are independent whenever $\tau_1 \neq \tau_2 \neq \tau_3, \beta_1 \neq \beta_2 \neq \beta_3$. Now expressing the column vectors of A as linear functions of the column vectors of B, one can easily verify that whenever the column vectors of B are independent, the column vectors of A are also independent. Hence rank (A) $= 4$ for all $\underline{\eta} \in \Theta$ which in turn implies that rank $(I - P)F = 4$ for all $\underline{\eta} \in \Theta$.

Theorem 4.2. A connected block design $d(v, b, k_{d1}, k_{d2}, \ldots, k_{db})$ with $b \geq 3$, $v \geq 3$ and $k_{di} \geq 3$ for at least three blocks, will provide estimation of $\underline{\theta}$ under the model (4.2) whenever at least one triplet of treatments (h, h', h'') say, occur in three different blocks.

The proof follows (through contradiction) essentially along similar line of arguments as in Theorem 3.2.

5. OPTIMAL ESTIMATION OF θ IN A BLOCK–DESIGN UNDER TUKEY'S MODEL

Now we focus our attention to the problem of efficient estimation of single non–additivity parameter θ under model (1.3). For fixed N (total number of experimental units) and v (number of treatments), let $\Omega(N, v)$ denote the class of all connected block designs (with block sizes $\leq v$) providing estimation of θ. For a design $d \in \Omega(N, v)$ let $v_d(\hat{\hat{\theta}})$ denote the variance of $\hat{\hat{\theta}}$ (see (2.5)).

Assume at this stage that $\hat{\underline{\tau}}$ and $\hat{\underline{\beta}}$ are consistent for $\underline{\tau}$ and $\underline{\beta}$. This could be achieved, for example, using the existing data set in combination with otherwise available consistent estimators for $\underline{\tau}$ and $\underline{\beta}$ from independent sources. Note that the passage from $\hat{\theta}$ in (2.4) to $\hat{\hat{\theta}}$ in (2.5) is not affected by use of such independent auxiliary information. Since we are primarily interested in making a relative comparison of various designs $\in \Omega(N, v)$, the above assumption is not unrealistic. On the other hand, this justifies the approximation

$$v_d(\hat{\hat{\theta}}) \approx v_d(\hat{\theta}) = [\underline{f}'(I-P)\underline{f}]^{-1}\,\sigma^2 \qquad \text{(see (2.4))}$$

and it depends on $\underline{\tau}$ and $\underline{\beta}$ through $\underline{f}$. We will make use of this approximate expression for $v_d(\hat{\hat{\theta}})$ while comparing the relative efficiencies of different designs. For this sort of comparison, one may certainly restrict to the effective parameter space:

$$\begin{aligned} \Theta_1 = \{(\tau,\beta) : \ & \underline{\tau}'\underline{\tau} = 1,\ \underline{\tau}'\underline{1} = 0,\ \underline{\beta}'\underline{\beta} = 1,\ \underline{\beta}'\underline{1} = 0 \\ & \tau_h \neq \tau_h{}' \neq 0 \quad h \neq h', \\ & \beta_i \neq \beta_i{}' \neq 0 \quad i \neq i'\} \end{aligned} \qquad \ldots(5.1)$$

Note that in the above we have not unnecessarily bothered to include μ in the description of effective parameter space.

Now we recall the following two definitions.

Definition 2. A design $d^* \in \Omega(N, v)$ is uniformly best if for every other $d \in \Omega(N, v)$, $v_{d^*}(\hat{\hat{\theta}}) \le v_d(\hat{\hat{\theta}})$ for all $(\underline{\tau}, \underline{\beta}) \in \Theta_1$ with strict inequality at some point.

Definition 3. A design d^* in $\Omega(N, v)$ is a minimax design if

$$\max_{\Theta_1} v_{d^*}(\hat{\hat{\theta}}) = \min_{d \in \Omega(N, v)} \max_{\Theta_1} v_d(\hat{\hat{\theta}})$$

If we confine only to binary designs in $\Omega(n, v)$, the following results may be derived.

Lemma 5.1. For an RBD $\sigma^{-2} v_d(\hat{\hat{\theta}}) = 1$ for all $(\underline{\tau}, \underline{\beta}) \in \Theta_1$.

Proof. For an RBD with parameters b and v,

$$(I - P) = (I - \tfrac{1}{v}J) \otimes (I - \tfrac{1}{v}J) \text{ and}$$
$$\sigma^{-2} v_d(\hat{\hat{\theta}}) = \frac{1}{\left[\sum_{t=1}^{b} \beta_t^2\right]\left[\sum_{h=1}^{v} \tau_h^2\right]} = 1 \text{ in } \Theta_1.$$

This result is reported in Kshirsagar (1983).

Lemma 5.2. For any binary design d, $\sigma^{-2} v_d(\hat{\hat{\theta}})$ varies between 1 and ∞.

Proof: We have for d, $\sigma^{-2} v_d(\hat{\hat{\theta}}) = (\underline{f}'(I-P)\underline{f})^{-1}$

Since for an incomplete binary block design all combinations of $\beta_i \tau_h$, i = 1, 2, ..., b and h = 1, ..., v do not appear in $\underline{f}$, we augment (I − P) to a bv × bv matrix by inserting suitable null columns and null rows corresponding to the combinations $\beta_i \tau_h$ which are missing in $\underline{f}$. With this in view, we may write

$\sigma^{-2} v_d(\hat{\hat{\theta}}) = (\underline{f}^{o\prime}(I-P)^o \underline{f}^o)^{-1}$, where $\underline{f}^0 = (\underline{\beta} \otimes \underline{\tau})$ and $(I-P)^0$ is the augmented (I − P) .

Now, $\underline{f}^{o\prime}\underline{f}^o = 1$ since $\underline{\tau}'\underline{\tau} = 1$ and $\underline{\beta}'\underline{\beta} = 1$ in Θ_1 .

Also, $\underline{f}^{o\prime}\underline{1} = 0$ since $\underline{\tau}'\underline{1} = 0$ and $\underline{\beta}'\underline{1} = 0$ both hold in Θ_1 even though one of them would have been enough.

$$\begin{aligned} \text{Let } \Theta_2 = \{(\underline{\tau}, \underline{\beta}) : \underline{f}^{o\prime}\underline{f}^o = (\underline{\beta} \otimes \underline{\tau})' (\underline{\beta} \otimes \underline{\tau}) = 1 ,\\ \underline{f}^{o\prime}\underline{1} = (\underline{\beta} \otimes \underline{\tau})'\underline{1} = 0\} \ . \end{aligned}$$

Obviously $\Theta_2 \supset \Theta_1$ and, consequently,

$$\min_{\Theta_2} \underline{f}^{o\prime}(I-P)^o\underline{f}^o \le \min_{\Theta_1} \underline{f}^{o\prime}(I-P)^o\underline{f}^o$$

and

$$\max_{\Theta_2} \underline{f}^{o\prime}(I-P)^o\underline{f}^o \ge \max_{\Theta_1} \underline{f}^{o\prime}(I-P)^o\underline{f}^o \ .$$

As (I − P) is idempotent, so also is $(I-P)^o$ and $(I-P)^o$ has the eigenvalues o (multiplicity > 1) and 1 (multiplicity ≥ 1) as trace of (I −

$P)^o < bv - b - v + 1$. So $\min_{\Theta_2} \underline{f}^{o\prime}(I - P)^o \underline{f}^o = 0$ and $\max_{\Theta_2} \underline{f}^{o\prime}(I - P)^o \underline{f}^o = 1$ and, hence,

$\sigma^{-2} v_d (\hat{\hat{\theta}}) = \{\underline{f}^{o\prime}(I - P)^o \underline{f}^o\}^{-1}$ lies between 1 and ∞ . Thus, summing up the above two lemmas, we immediately obtain the following result.

Theorem 5.1. Whenever N is a multiple of v , an RBD is the *uniformly best* among all *binary* designs in $\Omega(N , v)$.

If now non-binary designs are allowed to be considered, the RBD no longer remains uniformly best, but it turns out to be minimax. To prove this, we first establish one structural property of non-binary designs with block sizes $\leq v$.

Lemma 5.3. For any non-binary design d with $b > 2$ and $k_{di} \leq v$ for all i , there exist at least one pair of blocks say (i , i') such that considering these two blocks only, the hth treatment replication $n_{dhi} + n_{dhi'}$ is strictly less than 2 for some $h \in \{1, 2, ..., v\}$.

Moreover,

(i) if for some h, $n_{dhi} + n_{dhi'} = 0$, then there exists some $h'(\neq h) \in \{1,2, ..., v\}$ such that $n_{dh'i} + n_{dh'i'} \leq 3$;

(ii) if for some h, $n_{dhi} + n_{dhi'} = 1$, then there exists some $h'(\neq h) \in \{1, 2, ..., v\}$ such that $n_{dh'i} + n_{dh'i'} \leq 2$.

Proof. We first establish for some treatment h and for some pair of blocks (i, i'), $n_{dhi} + n_{dhi'} < 2$.

Suppose in the design d , there exists at least one block say i , such that $k_{di} < v$. Then pairing this block with any other block, say i' , we get

$$\bar{r}(i, i') = \frac{\sum_{h=1}^{v} (n_{dhi} + n_{dhi'})}{v} = \frac{k_{di} + k_{di'}}{v} < 2 \ .$$

Hence, for at least one treatment $h \epsilon \{1, 2, \ldots, v\}$, $n_{dhi} + n_{dhi'} < 2$.

If the design has constant block size $k = v$, then the proof also follows by contradiction unless $b = 2$.

Now we are ready to prove (i) and (ii).

(i) if for any pair of blocks (i, i') , $n_{dhi} + n_{dhi'} = 0$ for some h , there exists some other h' with $n_{dh'i} + n_{dh'i'} \leq 3$ as otherwise

$$k_{di} + k_{di'} = \sum_{h=1}^{v} (n_{dhi} + n_{dhi'}) \geq 4(v - 1) > 2v$$

whenever $v \geq 3$ and this is a contradiction to $k_{di} \leq v$ for all i .

Using a similar argument (ii) can be verified.

The following theorem furnishes a relative comparison of efficiencies in the class $\Omega(N, v)$ *including* non-binary designs.

Theorem 5.2. For any non-binary design $d \epsilon \Omega(N,v)$,

$$\max_{(\tau,\beta) \epsilon \Theta_1} \sigma^{-2} V_d(\hat{\theta}) \geq 1$$

Proof. We prove this theorem by contradiction. Suppose, if possible,

$$\sigma^{-2} v_d(\hat{\theta}) = (\underline{f}'(I - P)\underline{f})^{-1} < 1 \ \text{for all} \ (\underline{\tau}, \underline{\beta}) \ \epsilon \ \Theta_1 \ ,$$

then it follows that

$$\underline{f}'\underline{f} > \underline{f}'(I-P)\,\underline{f} > 1 \quad \text{for all } (\underline{\tau},\underline{\beta}) \in \Theta_1 \ .$$

Now $\underline{f}'\underline{f} = \sum_{i=1}^{b} \sum_{h=1}^{v} n_{dhi}\, \beta_i^2\, \tau_h^2$.

We will show that this leads to a contradiction .

$$\Sigma\Sigma n_{dhi} \beta_i^2\, \tau_h^2 \geq 1 \quad \text{for all } (\underline{\tau}, \underline{\beta}) \in \Theta_1 \ .$$

The lemma 5.3 above guarantees that we can always get hold of a pair of blocks (i, i′) say, in non–binary design d with $b > 2$ such that considering these two blocks only, there exist at least a pair of treatments (h, h′) say, for which the sum of four terms corresponding to replications in these two blocks $(n_{dhi} + n_{dh'i}) + (n_{dhi'} + n_{dh'i'})$ is *strictly* less than 4 . We set these two block effects viz β_i and $\beta_{i'}$ as $\pm \frac{1}{\sqrt{2}}$ and other block effects as zero. Similarly, setting these two treatment effects viz. τ_h and $\tau_h{}'$ as $\pm \frac{1}{\sqrt{2}}$ and other treatment effects as zero, we get for this particular choice of $\underline{\tau}$ and $\underline{\beta}$, $\underline{f}'\underline{f} < 1$. But this point obviously does not belong to the relevant parameter space Θ_1 . However, as $v_d(\hat{\hat{\theta}})$ is a continuous function of $\underline{\tau}$ and $\underline{\beta}$, this particular choice indicates that in the neighborhood of this point, there exists a point in Θ_1 for which $\underline{f}'\underline{f} < 1$.

Hence, for *any* non–binary design d with $b > 2$, the theorem holds.

It remains to prove the theorem for $b = 2$. For $b = 2$ the only non–trivial case corresponds to $r_{dh} = 2$ for all h .

From the very structure of the design it follows that for all points in Θ_1 (clearly with $(\beta_1, \beta_2) = \pm \frac{1}{\sqrt{2}}$ up to a permutation)

$$\underline{f}'\underline{f} = \sum_{i=1}^{b} \sum_{h=1}^{v} n_{dhi} \beta_i^2 \cdot \tau_h^2 = \frac{1}{2} \sum_{h=1}^{v} \tau_h^2 r_{dh} = \sum_{h=1}^{v} \tau_h^2 = 1 .$$

Moreover, $P\underline{f}$ can be shown to be $\neq \underline{0}$, yielding thereby $\underline{f}'(I-P)\underline{f} < f'f = 1$ for some points in Θ_1.

Thus, we get the final result:

Theorem 5.3. RBD is a minimax design within the class of designs $\Omega(N,v)$.

Remark: It is interesting to note that, there are non-binary designs, for which $v_d(\hat{\hat{\theta}})$ is strictly less than 1 for some parameter points. As for example, with b blocks, having constant block size v, take the design d_0, (1,2, ..., v denoting treatments)

$$d_0 = \begin{bmatrix} 1 & 1 & 3 & 4 & \dots & v \\ 1 & 2 & 3 & 4 & \dots & v \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 2 & 3 & 4 & \dots & v \end{bmatrix}$$

which differs from an RBD only in the first block. Now, making the choice,

$$\tau_1 = \frac{1}{\sqrt{2}}, \tau_v = -\frac{1}{\sqrt{2}}, \tau_h = 0 \ \forall\, h \neq 1, \text{ and } v$$
$$\beta_1 = \frac{1}{\sqrt{2}}, \beta_2 = -\frac{1}{\sqrt{2}}, \beta_i = 0 \ \forall\, i \neq 1, 2,$$

$v_{d_0}(\hat{\hat{\theta}})$ turns out to be strictly less than 1. Hence, by continuity of the variance function, there exists a parameter point in the neighborhood of this point for which $v_{d_0}(\hat{\hat{\theta}})$ is strictly less than 1 .

In the above we have established that RBD is the "best" (in some sense) for estimation of single interaction parameter. With this in view, as regards multiple interaction parameters as in (4.2), we first study the dispersion matrix of $\hat{\hat{\underline{\theta}}}$ under an RBD. Let the dispersion matrix be denoted by $\text{var}(\hat{\hat{\underline{\theta}}})$. Then $[\text{var}(\hat{\hat{\underline{\theta}}})]^{-1}$ is given by the matrix

$$\begin{bmatrix} \underline{\beta}'\underline{\beta}\cdot\underline{\tau}'\underline{\tau} & \underline{\beta}'\underline{\beta}^{(2)}\cdot\underline{\tau}'\underline{\tau} & \underline{\beta}'\underline{\beta}\cdot\underline{\tau}'\underline{\tau}^{(2)} & \underline{\beta}'\underline{\beta}^{(2)}\cdot\underline{\tau}'\underline{\tau}^{(2)} \\ - & \beta^{(2)\prime}(I-\frac{J}{b})\beta^{(2)}\cdot\underline{\tau}'\underline{\tau} & \underline{\beta}^{(2)\prime}\underline{\beta}\cdot\tau'\underline{\tau}^{(2)} & \underline{\beta}^{(2)\prime}(I-\frac{J}{b})\beta^{(2)}\cdot\underline{\tau}'\underline{\tau}^{(2)} \\ - & - & \underline{\beta}'\underline{\beta}\cdot\underline{\tau}^{(2)\prime}(I-\frac{J}{v})\underline{\tau}^{(2)} & \underline{\beta}'\underline{\beta}^{(2)}\cdot\underline{\tau}^{(2)\prime}(I-\frac{J}{v})\underline{\tau}^{(2)} \\ - & - & - & \underline{\beta}^{(2)\prime}(I-\frac{J}{b})\underline{\beta}^{(2)}\cdot\underline{\tau}^{(2)\prime}(I-\frac{J}{v})\underline{\tau}^{(2)} \end{bmatrix}$$

Over the restricted parameter space

$\Theta_3 = \{(\tau,\beta) : (\tau,\beta) \in \Theta_1 , \sum_{i=1}^{b} \beta_i^3 = 0, \sum_{h=1}^{v} \tau_h^3 = 0\}$ the dispersion matrix takes on the simple form, viz.,

$$[\mathrm{var}(\hat{\theta})]^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \Sigma\beta_i^4 - \frac{1}{b} & 0 & 0 \\ 0 & 0 & \Sigma\tau_h^4 - \frac{1}{v} & 0 \\ 0 & 0 & 0 & (\Sigma\beta_i^4 - \frac{1}{b})(\Sigma\tau_h^4 - \frac{1}{v}) \end{bmatrix}$$

Now $\lambda_{\min}[\mathrm{var}(\hat{\underline{\theta}})]^{-1} \leq (\Sigma\beta_i^4 - \frac{1}{b})(\Sigma\tau_h^4 - \frac{1}{v})$ and this RHS expression can be made arbitrarily small for some choice of $\underline{\tau}$ and $\underline{\beta}$ in the parameter space Θ_3. Thus, if we bring in E–optimality criterion which may be relevant to this multiparameter problem, RBD behaves very badly. Since $F'(I - P)F$ becomes intractable for an arbitrary design, it is very difficult to establish any such optimality result for multiple interaction parameter. Maybe, we have to redefine the parameter space appropriately so that the RDB achieves a better standing.

ACKNOWLEDGEMENT

The authors thank Professor P. Bimasankaram and Professor R. Mukherjee of Indian Statistical Institute for their useful comments on an earlier draft of the paper.

BIBLIOGRAPHY

Kshirsager, A.M. (1983): A Course in Linear Models. Mercel Dekker.

Milliken, G.A. and Graybill, F.A. (1970): Extensions of the General Linear Hypothesis Model. Journal of American Statistical Association 65, 797–807.

Mukhopadhyay, A.C., Sinha, B.K., and SahaRay, R. (1984): Characterization of Optimal Designs for Efficient Estimation of Non-Additive Parameter in Tukey's Non-Additive Model. Presented in 72nd Session of Indian Science Congress, January 1985.

Rao, C.R. (1965): Linear Statistical Inference and its Applications. John Wiley.

Rao, C.R. and Mitra, S.K. (1971): Generalized Inverse of Matrices and its Applications. New York, John Wiley and Sons.

Scheffe, H. (1959): The Analysis of Variance. John Wiley.

Tukey, J.W. (1949): One Degree of Freedom for Non-additivity, Biometrics 5, 232-242.

Tukey, J.W. (1953): Answer to Query 43, Biometrics, 11, 111-113.

Tukey, J.W. (1968): The Future of Data Analysis. Annals of Mathematical Statistics 33, 1-67.

Received August 1986; Revised January 1990.