THE INSTITUTE OF STATISTICS

THE UNIVERSITY OF NORTH CAROLINA

THE DEFINITION OF PARAMETERS IN GENERAL LINEAR MODELS

by

Ronald W. Helms

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1320
December 1980

DEPARTMENT OF BIOSTATISTICS
Chapel Hill, North Carolina

1. INTRODUCTION

In recent years a number of authors have addressed the question, "What hypothesis is being tested by such-and-such general linear model program (or procedure)?" These questions have arisen largely because of the use of traditional, less-than-full-rank (LTFR), ANOVA linear models by automatic or semi-automatic programs such as GLM in SAS (Goodnight, 1976). Although one can avoid these questions by using full rank models in conjunction with non-automatic software such as LINMOD (Helms, Christiansen, and Hosking, 1979), the inconvenience of non-automatic software, especially for large problems, is simply too great. One can expect substantial advances in automatic linear model software in the next few years.

A number of authors have addressed a variety of aspects of the "What hypothesis is being tested" issue, and related issues, over the past decade. See, for example, Urquhart, Weeks and Henderson (1973), Hocking and Speed (1975), Speed and Hocking (1976), Seeley (1977), and the recent entire special issue of Communications in Statistics (Vol. A9, No. 2, 1980). (The papers are listed in the List of References.)

The issue addressed in this paper is a fundamental one which will be of special importance to producers of automated linear model software, viz., what are the definitions of the various parameters in a linear model? The point of view adopted here (taken from Scheffe (1959), Chapter 3), is that the fundamental quantities in a linear model $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$, are $\underset{\sim}{X}$ and the elements of $E(\underset{\sim}{Y})$. The parameter vector $\underset{\sim}{\beta}$ may be, in many cases, defined as a vector-valued linear function of $\underset{\sim}{X}$ and $E(\underset{\sim}{Y})$. In such a case the appropriate interpretation of $\underset{\sim}{\beta}$ arises from the definition of $\underset{\sim}{\beta}$. For illustration, Scheffe (1959, section 3.1 defined a model for a one-way design by

$$Y_{ij} = \beta_i + e_{ij},\ i=1,2,\ldots,I;\ j=1,2,\ldots,J_i$$

$$\{e_{ij}\} \text{ are independently } N(0\ \sigma^2).$$

Here the definition of each element of $\underset{\sim}{\beta}$ is $\beta_i = E(y_{ij})$. β_i has a simple, direct interpretation. The interpretation of an estimate of β_i follows from this definition: $\hat{\beta}_i$ is an estimate of the population mean of observations from the i-th group. If the elements of $\underset{\sim}{\beta}$ can be defined as linear combinations of the elements of $\underset{\sim}{X}$ and $E(\underset{\sim}{Y})$ then one can reasonably and directly interpret the elements of $\underset{\sim}{\beta}$ as known as linear combinations of (unknown) expected values--means of well defined populations. The interpretation may, or may not, be simple, depending on the complexity of the linear combination.

If $\underset{\sim}{\beta}$ has a direct interpretation then an estimate of $\underset{\sim}{\beta}$ also has a direct interpretation.

In the usual ANOVA model for the one-way design, without side restrictions, $E(y_{ij}) = \mu+\alpha_i$, the parameter μ does not have a definition in the sense described above. Moreover, there is no reasonable interpretation of an estimate of μ (e.g., a least squares estimate), for one can obtain another least squares estimate of μ by adding an arbitrary constant. Thus, definability of a parameter is an important part of the interpretation of the parameter and its estimate.

The parameter $\underset{\sim}{\beta}$ in the model $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$ is called a primary parameter; a secondary parameter is a linear combination of the form $\underset{\sim}{\theta} = \underset{\sim}{C}\underset{\sim}{\beta}-\underset{\sim}{g}$ [or $\underset{\sim}{\theta} = \underset{\sim}{H}E(\underset{\sim}{Y}) - \underset{\sim}{g}$]. A primary parameter may (or may not) be defined as a linear combination of $\underset{\sim}{X}$ and $E(\underset{\sim}{Y})$. Similarly, a secondary parameter may (or may not) be defined as a linear combination of $E(\underset{\sim}{Y})$ and fixed, known matrices ($\underset{\sim}{X}$, $\underset{\sim}{C}$, etc.). If $\underset{\sim}{\theta}$ is so defined, the interpretation of the elements of $\underset{\sim}{\theta}$, and the interpretation of an estimator of $\underset{\sim}{\theta}$, stem directly from the definition.

The importance of a definition is the close link between the definition and interpretation of a parameter. The interpretation of a well defined parameter (and its estimator) is straightforward, though possibly complex. As illustrated by the one-way μ above, an undefinable parameter, and its estimate, probably cannot be interpreted.

Section 2 of this paper contains a definition of primary parameter definitions in the general linear model and a simple illustration. The issue of alternative, but equivalent, definitions is addressed, and a canonical definition is defined. Section 3 addresses similar issues for secondary parameters.

The results presented in this paper may be used by the developer of automatic or semi-automatic linear model software. When a model is constructed for a user the program can display the canonical definitions of $\underset{\sim}{\beta}$ and/or any definable secondary parameters, together with equivalent, more easily interpreted definitions. Given such definitions the user will be able, in a straightforward manner, to answer the questions, "What do these parameters mean?" and "What hypothesis is being tested?"

2. DEFINITIONS OF LINEAR MODEL PRIMARY PARAMETERS

2.1 Notation and Terminology

We shall assume, as a basis for discussion, the general linear univariate model with an Nx1 vector of random observations, $\underset{\sim}{Y}$, an Nxq "design matrix", $\underset{\sim}{X}$, with $r = \text{Rank}(\underset{\sim}{X}) \leq q < N$, such that

(2.1) $E(\underset{\sim}{Y}) = \underset{\sim}{X}\,\underset{\sim}{\beta}$

(2.2) $V(\underset{\sim}{Y}) = \sigma^2 \underset{\sim}{I}$.

(Generalizations of (2.2) will be discussed below.) The qx1 vector $\underset{\sim}{\beta}$ is a nonstochastic vector of primary parameters whose values are unknown; presumably $\underset{\sim}{\beta}$ can take on any value in E^q, q-dimensional Euclidean space. The model equation (2.1) is assumed to be consistent, i.e., whatever the value of $E(\underset{\sim}{Y})$ there exists a $\underset{\sim}{\beta} \epsilon E^q$ such that (2.1) is satisfied. Thus,

$$E(\underset{\sim}{Y}) \epsilon M(\underset{\sim}{X}) = \text{linear manifold of columns of } \underset{\sim}{X}$$
$$= \{\underset{\sim}{t} \epsilon E^n\text{: for some } \underset{\sim}{\beta} \epsilon E^q \quad t = \underset{\sim}{X}\underset{\sim}{\beta}\}.$$

Thus $\underset{\sim}{\beta}$ is an exact solution of (2.1). This contrasts with the inconsistent data equation, $\underset{\sim}{Y} \cong \underset{\sim}{X}\underset{\sim}{b}$ which does not have an exact solution. [The least squares estimator, $\underset{\sim}{b}$, for $\underset{\sim}{\beta}$, produces an approximate solution of the data equation and an exact solution of the consistant normal equations, $(\underset{\sim}{X}'\underset{\sim}{X})\underset{\sim}{b} = \underset{\sim}{X}'\underset{\sim}{Y}$.]

We are interested in studying specifications and definitions of the primary parameter $\underset{\sim}{\beta}$.

DEFINITION 2.1. A specification of a parameter is a collection of one or more conditions and matrix equations which involve: (1) the parameter, (2) $E(\underset{\sim}{Y})$, and (3) known, constant matrices (such as $\underset{\sim}{X}$).

DEFINITION 2.2. A definition of a parameter is a specification which uniquely determines the parameter as a vector-valued function of $E(\underset{\sim}{Y})$ and known, constant matrices (such as $\underset{\sim}{X}$), i.e., for each permissible, fixed value of $E(\underset{\sim}{Y})$ and the known constant matrices there is one unique value of the parameter which satisfies the specification.

A specification may not result in a parameter being well-defined in the mathematical sense; a definition results in a mathematically well-defined parameter (whose value, like that of $E(\underset{\sim}{Y})$ is unknown).

For example, the model equation $E(Y) = \underset{\sim}{X}\underset{\sim}{\beta}$ is a specification of $\underset{\sim}{\beta}$; it is a definition iff for each $E(\underset{\sim}{Y}) \epsilon M(\underset{\sim}{X})$ there exists a unique value of $\underset{\sim}{\beta}$ satisfying the equation. The following matrix theorem [adapted from Searle (1973), Theorem 4, page 12] is useful in the present context:

THEOREM 2.1. The number of linearly independent solutions to the consistent model equation $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$ is $q+1-\text{Rank}(\underset{\sim}{X})$ where $\underset{\sim}{X}$ is $N \times q$, $q<N$.

The following results are immediate consequences.

COROLLARY 2.1.1. The consistent model equation $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$ has a unique solution for $\underset{\sim}{\beta}$ and is, therefore, a definition of $\underset{\sim}{\beta}$, iff Rank $(\underset{\sim}{X}) = q$, i.e., iff $\underset{\sim}{X}$ is Full Rank (FR), in which case the unique solution is given by

$$\underset{\sim}{\beta} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'E(\underset{\sim}{Y}). \tag{2.3}$$

Let S be a specification of a parameter $\underset{\sim}{\beta}$ (as, for example, the model equation $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$). We shall say that $\underset{\sim}{\beta}$ is well defined (or "definable") if and only if S is a definition.

The primary parameter, $\underset{\sim}{\beta}$, in the usual Less Than Full Rank (LTFR) ANOVA model without side restrictions is not well defined. Some workers leave the parameters unrestricted but place restrictions on the estimates to determine a solution to the normal equations. This procedure leaves the parameters undefined. If one places appropriate restrictions on the parameters of the LTFR model there are two consequences: (1) The model is different; imposing restrictions changes the model itself. (2) The model, with the additional restrictions, admits a unique solution and the specification becomes a definition. The well defined, restricted parameters are different from the undefined, unrestricted parameters.

DEFINABILITY and ESTIMABILITY. Obviously, a primary parameter is well defined if and only if it is estimable; in a sense, then, the two concepts are equivalent. However, as will be seen, the study of definitions and their properties is quite different from the study of estimators and their properties. Even though definability and estimability are equivalent concepts in this context, the concepts stem from different origins and it is useful to use different names for the two concepts.

2.2 Canonical, Alternative, and Equivalent Primary Parameter Definitions

In this section we consider unconstrained, full rank models in which the primary parameter is well defined. If we let the model equation $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$ specify $\underset{\sim}{\beta}$ then the following is a reasonable consequence of Corollary 2.1.1.

DEFINITION 2.3. Let $\underset{\sim}{\beta}$ be the primary parameter specified by the full rank model equation $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$; then the canonical definition of $\underset{\sim}{\beta}$ is

(2.3) $$\underset{\sim}{\beta} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'E(\underset{\sim}{Y}).$$

There are other definitions of a primary parameter and it is often the case that alternative definitions are more useful for human interpretation than the canonical definition. The following simple illustration will help make this point.

Illustration. Consider a hypothetical experiment with three observations, $\underset{\sim}{Y}' = (y_1, y_2, y_3)$, in which one knows from experimental considerations that $E(y_1) = E(y_2)$, but $E(y_3)$ may be different. Consider the following intuitive definition of a FR model and its parameters. The definition is labelled D_2; D_1 will be the label for the canonical definition. The parameter vector defined by D_i will be denoted $\underset{\sim}{\beta}^{(i)} = (\beta_1^{(i)}, \beta_2^{(i)})'$.

(2.4) $$D_2: \beta_1^{(2)} = E[y_1] = E[y_2]$$
$$\beta_2^{(2)} = E[y_3] - E[y_1]$$

or, in matrices,

(2.5) $$D_2: \underset{\sim}{\beta}^{(2)} = \underset{\sim}{A}_2E(\underset{\sim}{Y}) = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} E[\underset{\sim}{Y}]$$

Solving (2.4) for $E(Y_3) = \beta_1^{(2)}+\beta_2^{(2)}$ leads to the model equation, in matrices:

(2.6) $$E[\underset{\sim}{Y}] = \underset{\sim}{X}\beta^{(1)} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \underset{\sim}{\beta}^{(1)}$$

where $\underset{\sim}{\beta}^{(1)}$ is defined by the canonical definition:

(2.7) $$D_1: \underset{\sim}{\beta}^{(1)} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'E(\underset{\sim}{Y})$$
$$= (1/2)\begin{bmatrix} 1 & 1 & 0 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} E(Y_1) \\ E(Y_2) \\ E(Y_3) \end{bmatrix}$$
$$= \underset{\sim}{A}_1E(\underset{\sim}{Y})$$

or, equivalently

(2.8) $$D_1: \beta_1^{(1)} = (1/2)\ [E(y_1)+E(y_2)]$$
$$\beta_2^{(1)} = (-1/2)\ [(y_1)+E(y_2)]+E(y_3)$$

There are at least two reasons for considering alternative definitions. First, if one constructs a model by first defining readily interpretable parameters (as with (2.4)) and then constructing an appropriate $\underset{\sim}{X}$ matrix the "original", or "intuitive", definition will usually not be the same as the canonical definition. For an "ANOVA setting" the intuitive definition matrix, $\underset{\sim}{A}_1$, will usually contain mostly 0's, 1's and -1's, which make the intuitive definition more readily interpretable than the canonical definition.

Secondly, if one first constructs $\underset{\sim}{X}$ and then determines the canonical definition, the $\underset{\sim}{A}_1$ matrix may be difficult to interpret. As will be illustrated later, it is possible to construct an equivalent definition, $\underset{\sim}{\beta}^{(2)} = \underset{\sim}{A}_2E(\underset{\sim}{Y})$, which is much more easily interpretable, especially for incomplete and/or unbalanced designs. This "example" illustrates the fact that a primary parameter may have two or more definitions, which may or may not be "equivalent".

DEFINITION 2.4. Equivalent Definitions of FR Model Primary Parameters. Let $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$ specify a full rank linear model and consider two alternative definitions of the primary parameter, $\underset{\sim}{\beta}$, of this model:

$$D_i: \ \underset{\sim}{\beta}^{(i)} = \underset{\sim}{A}_iE(\underset{\sim}{Y}) \ , \ i = 1,2,$$

where $\underset{\sim}{A}_1 \neq \underset{\sim}{A}_2$. Then D_1 and D_2 are equivalent definitions if and only if they define the same parameter, that is, if and only if for all $E(\underset{\sim}{Y}) \epsilon M(\underset{\sim}{X})$,

(2.9) $\underset{\sim}{A}_1E(\underset{\sim}{Y}) = \underset{\sim}{A}_2E(\underset{\sim}{Y})$.

COROLLARY 2.4.1. In the context of Definition 2.4, definitions D_1 and D_2 are equivalent if and only if

(2.10) $\underset{\sim}{A}_1\underset{\sim}{X} = \underset{\sim}{A}_2\underset{\sim}{X}$

which is equivalent to

(2.11) $\underset{\sim}{A}_1\underset{\sim}{X}\underset{\sim}{d} = \underset{\sim}{A}_2\underset{\sim}{X}\underset{\sim}{d}$ for all $\underset{\sim}{d} \epsilon E^q$.

Proof. The equivalence of (2.10) and (2.11) is a well-known theorem from matrix theory. Substituting $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$ in (2.9) and changing $\underset{\sim}{\beta}$ to $\underset{\sim}{d}$ produces (2.11), i.e., (2.9) holds iff (2.11) holds. QED.

The most important case is the equivalence of an alternative, or "intuitive", definition to the canonical definition. This issue is addressed in the following corollaries.

COROLLARY 2.4.2. In the context of Definition 2.4, let D_1 be the canonical definition of β, i.e., $A_1 = (X'X)^{-1}X'$; then D_2 is equivalent to the canonical definition if and only if $A_2X = I_q$. (Proof: Equivalence <==> $A_2X = A_1X = (X'X)^{-1}X'X = I_q$. QED).

COROLLARY 2.4.3. Class of Definitions Equivalent to the Canonical Definition. In the context of Definition 2.4 let D_1 be the canonical definition $[A_1 = (X'X)^{-1}X']$. Then definition D_2 is equivalent to the canonical definition, D_1 if and only if there exists a qxN matrix K such that

(2.12a) $\quad A_2 = (X'X)^{-1}X'-K$, and

(2.12b) $\quad KX = 0$.

That is, the class of all definitions equivalent to the canonical definition is generated by matrices A_2 of the form (2.12a) satisfying (2.12b).

Proof (1). Let A_2 be of the form (2.12a) and let $KX = 0$. Note $A_1X = I$. Then

$$A_2X = [(X'X)^{-1}X'-K]X = I \Rightarrow D_1,\ D_2 \text{ equivalent. QED(1).}$$

Proof (2). Assume D_2 is equivalent to D_1 so that $A_2X = I$. A_2 is fixed. Let $K = (X'X)^{-1}X'-A_2$ and note that, solving for A_2, (2.12a) is satisfied. Also, $KX = (X'X)^{-1}X'X-A_2X = 0$. QED(2).

In the simple illustration given above, since D_1 is the canonical definition one can easily demonstrate equivalence by verifying $A_2X = I$. The class of definitions equivalent to the canonical definition can be found with a bit of arithmetic. The condition $KX = 0$ (2.12b) requires that K be of the form

$$K = \begin{bmatrix} a & -a & 0 \\ b & -b & 0 \end{bmatrix}$$

for arbitrary scalars a,b. Further, one can easily verify that the A_2 in the illustration is generated by a = -1/2, b = 1/2.

Non-equivalent definitions are easily constructed. In fact, after defining the parameters, if one makes an error in constructing X, the intuitive and canonical definitions will generally not be equivalent. To continue the illustration consider a third definition

$$D_3:\ \beta_1^{(3)} = E(y_1) = E(y_2)$$

$$\beta_2^{(3)} = E(y_3)$$

or

$$\beta^{(3)} = A_3E(Y) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} E(Y)$$

A simple calculation shows $A_3X \neq I$, as expected, because $\beta_2^{(2)}$ is defined as a "cell mean" rather than a difference of cell means (as $\beta_2^{(2)}$). D_3 defines a "cell mean" model which would have an X matrix different from the one for definitions D_1 and D_2 .

This illustrates the point that the definition of β is taken in the context of the model equation $E(Y) = X\beta$. The model equation may contain information not contained in the definition of β alone. In equation (2.5) for example, A_2 does not specify that $E(y_1) = E(y_2)$.

2.3 Computing "Simpler" Definitions from X

Although Corollary 2.4.3 defines the class of primary parameter definitions equivalent to the canonical definition, no algorithm is provided for computing such a definition. The following algorithm will produce a canonical-equivalent definition which is often easily interpreted.

The use of Gaussian elimination (row operations) to "reduce" the matrix $[X,I]$ to Hermite normal form can be written in matrices as

$$\begin{bmatrix} A_2 \\ R \end{bmatrix} \begin{bmatrix} X, & I_n \end{bmatrix} = \begin{bmatrix} I_q & A_2 \\ 0 & R \end{bmatrix} \tag{2.13}$$

where A_2 is $q \times N$, R is $(N-q) \times N$. That is, if one uses row operations to reduce $[X, I_n]$ to Hermite normal form, the result of the computations will be as shown in the right side of (2.13).

The matrix A_2 forms an alternative definition of β which is equivalent to the canonical definition. The matrix R satisfies $RX = 0$ and represents model restrictions on $E(Y)$:

$$RE(Y) = RX\beta = 0.$$

R will contain information such as $E(y_1)-E(y_2) = 0$ in the illustration. R will also contain restrictions such as an assumption, in a factorial design, that certain interactions are zero.

To continue the illustration, the algorithm is applied to the matrix $[X,I]$ to produce:

$$[X,I] \longrightarrow \begin{bmatrix} I_2 & A_2 \\ 0 & R \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \longrightarrow \left[\begin{array}{cc|ccc} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 \\ \hline 0 & 0 & -1 & 1 & 0 \end{array}\right]$$

We see that $0 = \underset{\sim}{R}E(\underset{\sim}{Y}) = -E(y_1) + E(y_2)$, the restriction that $E(y_1) = E(y_2)$. The $\underset{\sim}{A}_2$ is the same matrix originally used in D_2, a coincidence, not unexpected in this simple illustration.

2.4 Definitions from Cell Means and "First Observations"

In ANOVA-like models with no concommitant variables one can frequently reduce the effort required to form a definition. Consider, for example, an AxB factorial treatment design run as a completely random experimental design, with y_{abk} denoting the k-th observation from the a-th level of factor A, b-th level of factor b. With no concommitant variables we may define, for each non-empty cell:

$$\mu_{ab} = E\,[y_{abk}]\ ,\ k = 1,2,\ldots,N_{ab}$$

and, in particular:

(2.14) $$\mu_{ab} = E[y_{ab1}].$$

If we extract from the model equation $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$ those rows corresponding to the first observation from each cell (y_{ab1}), we have the "essence" model, say $E(\underset{\sim}{Y}_E) = \underset{\sim}{X}_E\underset{\sim}{\beta}$. Note that only $\underset{\sim}{Y}$ and $\underset{\sim}{X}$ are affected, not $\underset{\sim}{\beta}$. But, from above, $\underset{\sim}{\mu} = E(\underset{\sim}{Y}_E)$, where $\underset{\sim}{\mu}$ is the set of cell means, $\underset{\sim}{\mu}_{ab}$, arranged in a vector, in the order of the elements $E(y_{ab1})$ in $E(\underset{\sim}{Y}_E)$.

Letting NC denote the number of (non-empty) cells in the design, $\underset{\sim}{\mu}$, $E(\underset{\sim}{Y}_E)$, and $\underset{\sim}{X}_E$ all have NC rows, $NC \leq N$. Often NC is substantially less than N.

Definitions of $\underset{\sim}{\beta}$ may be constructed in this "essence" model. An essence-model canonical definition of $\underset{\sim}{\beta}$ is:

(2.15) $$D_{E1}:\ \underset{\sim}{\beta} = (\underset{\sim}{X}_E'\underset{\sim}{X}_E)^{-1}\underset{\sim}{X}_E'\underset{\sim}{\mu} = \underset{\sim}{A}_{E1}\underset{\sim}{\mu}$$

Applying the algorithm described above to $[\underset{\sim}{X}_E, \underset{\sim}{I}_{NC}]$ produces an alternative, possibly more intuitively appealing definition of $\underset{\sim}{\beta}$ in the context of the "essence" model:

(2.16) $$[\underset{\sim}{X}_E, \underset{\sim}{I}_{NC}] \xrightarrow{\text{Alg.}} \begin{bmatrix} \underset{\sim}{I}_q & \underset{\sim}{A}_{E2} \\ \underset{\sim}{0} & \underset{\sim}{R}_E \end{bmatrix}$$

where $\underset{\sim}{A}_{E2}$ is qxNC,. $\underset{\sim}{R}_E$ is (NC-q)xNC and

$$D_{E2}:\ \underset{\sim}{\beta} = \underset{\sim}{A}_{E2}\underset{\sim}{\mu},$$

which is, by construction, equivalent to D_{E1} in the context of the "essence" model.

If the original model imposes constraints on the estimation space (e.g., AxB interactions are zero), then the number of parameters in $\underset{\sim}{\beta}$, i.e., q, is less than the number of cell means: $q < NC$, and the restrictions are defined by $\underset{\sim}{R}_E\underset{\sim}{\mu} = \underset{\sim}{0}$. If there are no constraints, q = NC and the $\underset{\sim}{R}_E$ matrix is null (0x0).

The extraction of the "essence" model may be performed as a matrix multiplication, with

$$[E(Y_E),\ X_E] = U[E(Y),\ X]$$
$$\text{NCx1} \quad \text{NCxq} \quad \text{NCxN} \quad \text{Nx1} \quad \text{Nxq}$$

where U is a matrix of 1's and 0's which "picks off" the rows of $E(Y)$ and X corresponding to the first observation in each cell. The nonzero columns of U collectively form I_{NC}.

"Essence" model definitions are easily expanded to full model definitions using the matrix U:

$$\beta = A_E E(Y_E) = A_E U E(Y) = A E(Y)$$

with $A = A_E U$. (The result of this operation is essentially the same as inserting one column of zeros into A_E to correspond to each non-first observation (y_{abk}, $k>1$) in each cell.) The correspondence between "essence" model definitions and full model definitions is established in the following theorem:

<u>NOTATION</u>. Let a full rank linear model be specified by the model equation $E(Y) = X\beta$ where $E(Y)$ is Nx1, X is Nxq, and let the canonical definition of β be

$$D_1:\ \beta = (X'X)^{-1}X'E(Y) = A_1 E(Y).$$

Let NC be the number of distinct rows in X; $NC \le N$. Let U (NCxN) satisfy:

(1) Each column of U is zero or a column of I_{NC};

(2) $U'U = I_{NC}$ (which insures that U contains every column of I_{NC} exactly once.)

Let

$$X_E = UX,\ Y_E = UY \text{ so that } E(Y_E) = UE(Y).$$

Let the model $E(Y_E) = X_E$ be called the <u>essence model</u>. By definition the essence model canonical definition of β is

$$D_{E1}:\ \beta = A_{E1}E(Y_E) = (X_E'\ X_E)^{-1}X_E'\ E(Y_E).$$

Obviously $A_{E1}X_E = I_q$.

<u>THEOREM 2.2</u>. With the notation given above, when definition D_{E1} is extended to the full model the corresponding definition is equivalent to the canonical definition. That is, if the extension of D_{E1} to the full model is

$$D_{EE1}:\ \beta = (A_{E1}U)\ E(Y) = A_{EE1}E(Y)$$

then

$$A_{EE1}X = I, \text{ i.e., } D_{EE1} \Leftrightarrow D_1:\ \beta = (X'X)^{-1}X'E(Y).$$

Proof. As noted above, $\underset{\sim}{A}_{E1}\underset{\sim}{X}_E = \underset{\sim}{I}$; substituting:

$$\underset{\sim}{A}_{EE1}\underset{\sim}{X} = \underset{\sim}{A}_{E1}\underset{\sim}{U}\underset{\sim}{X} = \underset{\sim}{A}_{E1}\underset{\sim}{X}_E = \underset{\sim}{I}.\quad \text{QED.}$$

2.5 An Example

The incomplete, unbalanced part of a 3x4 factorial, with four missing cells, shown in Figure 2.1, is taken from Searle (1971, page 287). Unlike Searle, we shall assume the interactions are zero; they are omitted from the model.

The "reference cell" technique will be used to construct a full rank model with directly interpretable parameters. For convenience, let the first level of each factor be the reference level. (The models constructed by the SAS GLM procedure are, in effect, essentially equivalent to reference cell models, but the GLM algorithm uses the last level of each factor as the reference level.)

In reference cell models "main effects" are taken at the reference levels of the other factors, not averaged over the levels of the other factors as in other methods. Instead of a "general mean", one uses the mean of the reference cell as a measure of the "general level of response". Using Figure 2.1, the following reference cell parameters are defined:

$$\mu_{11} = E[Y_{11k}],\ \text{reference cell mean}$$
$$\alpha_2 = E[Y_{21k}]-\mu_{11} = \mu_{21}-\mu_{11}$$
$$\beta_3 = E[Y_{13k}]-\mu_{11} = \mu_{13}-\mu_{11}$$
$$\beta_4 = E[Y_{14k}]-\mu_{11} = \mu_{14}-\mu_{11}.$$

The missing cells do not permit the direct definitions of α_3 and β_2 at the reference (first) levels of the other factors, but by taking note of the no-interaction assumption one can define

$$\alpha_3 = E[Y_{33k}]-E[Y_{13k}] = \mu_{33}-\mu_{13}$$
$$\beta_2 = E[Y_{22k}]-E[Y_{21k'}] = \mu_{22}-\mu_{21}.$$

Expressing these results in matrix form produces the matrix $\underset{\sim}{A}_2$, shown in Table 2.1, which specifies the "intuitive definition" of $\underset{\sim}{\beta}$. Note that $\underset{\sim}{A}_2$ uses only the first observation from each cell.

The design matrix, $\underset{\sim}{X}$, and the $\underset{\sim}{A}_1$ specifying the canonical definition of $\underset{\sim}{\beta}$ are shown in Tables 2.2 and 2.3. Clearly, $\underset{\sim}{A}_1$ and $\underset{\sim}{A}_2$ are different; for interpretive purposes the definitions given above are generally preferable to the canonical definitions. An easy calculation shows $\underset{\sim}{A}_2\underset{\sim}{X} = \underset{\sim}{I}$, which proves the definition specified by $\underset{\sim}{A}_2$ is equivalent to the canonical definition specified by $\underset{\sim}{A}_1$. Note that one of the differences between $\underset{\sim}{A}_1$

and $\underset{\sim}{A}_2$ is that the canonical definition assigns equal weight to the expected values of observations in one cell, i.e., to all observations with the same expectation under the model. This would generally not be the case if the model contained concommitant variables.

The design matrix, $\underset{\sim}{X}_E$, for the "essence" model (for the first observation from each cell) is constructed by selecting the rows of $\underset{\sim}{X}$ marked with an asterisk (*) in Table 2.2. The $\underset{\sim}{X}_E$ matrix was used to compute $\underset{\sim}{A}_{E1}$, (see equation 2.15) the matrix which is the basis of the canonical definition of $\underset{\sim}{\beta}$ in the essence model. The transpose of $\underset{\sim}{A}_{E1}$ is shown in Table 2.4. $\underset{\sim}{A}_{E1}$ was expanded to $\underset{\sim}{A}_{EE1}$ by inserting columns of zeroes to correspond to non-first observations, i.e., $\underset{\sim}{A}_{E1} = \underset{\sim}{A}_{E1}\underset{\sim}{U}$ as in Theorem 2.2. One can readily verify $\underset{\sim}{A}_E\underset{\sim}{X}_E = \underset{\sim}{I}_6$ and $\underset{\sim}{A}_{EE1}\underset{\sim}{X} = \underset{\sim}{I}_6$. In spite of the equivalence of definitions, note the substantial difference between $\underset{\sim}{A}_{EE1}$ (obtained by inserting rows of zeroes into the matrix in Table 2.4) and $\underset{\sim}{A}_1'$ in Table 2.3.

Finally, the matrix $[\underset{\sim}{X}_E, \underset{\sim}{I}_8]$ was transformed to Hermite normal form; the results are shown in Table 2.5. The resulting definition matrix, $\underset{\sim}{A}_{E3}$, when expanded to the full model, defines all the parameters, except α_3, just as they are defined by $\underset{\sim}{A}_2$ (Table 2.1).

The restrictions $\underset{\sim}{R}_E\underset{\sim}{\mu} = \underset{\sim}{0}$ are:

$$0 = \mu_{11}-\mu_{13}-\mu_{21}+\mu_{22}-\mu_{32}+\mu_{33}$$

$$0 = \mu_{11}-\mu_{14}-\mu_{21}+\mu_{22}-\mu_{32}+\mu_{34}$$

An examination of these linear combinations reveals that these are combinations of interactions.

The matrices $\underset{\sim}{A}_{E3}$ and $\underset{\sim}{R}_E$ are not uniquely defined by the algorithm. In fact, since the linear combinations above are zero, one can modify the definition of a parameter by adding zero (in the form of one of the rows of $\underset{\sim}{R}_E\underset{\sim}{\mu}$) to the definition of the parameter. To illustrate, consider the $\underset{\sim}{A}_{E3}$ definition of α_3 from Table 2.5, panel 2, row 3:

$$\alpha_3 = \mu_{11}+\mu_{21}-\mu_{22}+\mu_{32}$$

The first restriction above is:

$$0 = \mu_{11}- \mu_{13}-\mu_{21}+\mu_{22}-\mu_{32}+\mu_{33}$$

Thus, adding the expressions for α_3 and 0 yields

$$\alpha_3+0 = \mu_{33}-\mu_{13}$$

which is the same as the $\underset{\sim}{A}_2$ definition in Table 2.1. Using the second row of $\underset{\sim}{R}_E\underset{\sim}{\mu}$ instead of the first yields $\alpha_3 = \mu_{34}-\mu_{14}$. Since the interactions are zero, the equivalence of all these definitions of α_3 is obvious. The example illustrates one way in which different, but equivalent, definitions of a parameter may arise from one model.

As a result of this discussion one can observe that $\underset{\sim}{A}_{E3}$ may be modified by adding any linear combination of the rows of $\underset{\sim}{R}_E$ to any row(s) of $\underset{\sim}{A}_{E3}$; the resulting definition would be equivalent to the one from the original $\underset{\sim}{A}_{E3}$ and to the canonical definition. Computationally, this device may be used to "simplify" a computed "intuitively pleasing" definition.

The example also shows that a computed definition, as α_3 in $\underset{\sim}{A}_{E3}$ in Table 2.5, may contain more non-zero terms than, and may not be as easily interpreted as, other equivalent definitions.

2.6 Weighted Models: The Case $V(\underset{\sim}{Y}) = \sigma^2\underset{\sim}{V}$

To this point we have assumed $V(\underset{\sim}{Y}) = \sigma^2\underset{\sim}{I}$. If one assumes instead, $V(\underset{\sim}{Y}) = \sigma^2\underset{\sim}{V}$, where $\underset{\sim}{V}$ is a known, positive definite symmetric matrix then one can transform from the original model, $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$, to:

$$(2.17) \qquad \begin{aligned} E(\underset{\sim}{L}^{-1}\underset{\sim}{Y}) &= (\underset{\sim}{L}^{-1}\underset{\sim}{X})\underset{\sim}{\beta} \\ V(\underset{\sim}{L}^{-1}\underset{\sim}{Y}) &= \sigma^2\underset{\sim}{I}_N \end{aligned}$$

where

$$\underset{\sim}{V} = \underset{\sim}{L}\,\underset{\sim}{L}', \quad \underset{\sim}{V}^{-1} = (\underset{\sim}{L}^{-1})'\underset{\sim}{L}^{-1}.$$

We shall refer to (2.17) as the "transformed model". All the preceeding results apply to this model directly, with $(\underset{\sim}{L}^{-1}\underset{\sim}{Y})$ in place of $\underset{\sim}{Y}$ and $(\underset{\sim}{L}^{-1}\underset{\sim}{X})$ in place of $\underset{\sim}{X}$.

The canonical definition of $\underset{\sim}{\beta}$ in model (2.17) is, after some simplification:

$$\begin{aligned} D_{T1}\colon\ \underset{\sim}{\beta} &= (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{X}'(\underset{\sim}{L}^{-1})' \quad E(\underset{\sim}{L}^{-1}\underset{\sim}{Y}) \\ &= \underset{\sim}{A}_{T1}\,E(\underset{\sim}{L}^{-1}\underset{\sim}{Y}). \end{aligned}$$

Now consider the canonical definition of $\underset{\sim}{\beta}$ from the original model, but in the context of the transformed model:

$$\begin{aligned} D_{T2}\colon\ \underset{\sim}{\beta} &= (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'E(\underset{\sim}{Y}) \\ &= (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{L}\ E(\underset{\sim}{L}^{-1}\underset{\sim}{Y}) \\ &= \underset{\sim}{A}_{T2}E(\underset{\sim}{L}^{-1}\underset{\sim}{Y}) \end{aligned}$$

A quick application of Corollary 2.4.1 to D_{T2} in the contet of model (2.17), $\underset{\sim}{A}_{T2}\,(\underset{\sim}{L}^{-1}\underset{\sim}{X}) = \underset{\sim}{I}$, demonstrates that D_{T2} and D_{T1} are equivalent. Note that both definitions can be written in the context of the original model:

$$\begin{aligned} D_{01} = D_{T1}\colon\ \underset{\sim}{\beta} = \underset{\sim}{A}_{01}E(\underset{\sim}{Y}),\ \underset{\sim}{A}_{01} &= (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{V}^{-1} \\ D_{02} = D_{T2}\colon\ \underset{\sim}{\beta} = \underset{\sim}{A}_{02}E(\underset{\sim}{Y}),\ \underset{\sim}{A}_{02} &= (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}' \end{aligned}$$

Another application of Corollary 2.4.1 shows D_{01} and D_{02}, hence D_{T1} and D_{T2}, are equivalent in the context of the original model.

The parameter $\underset{\sim}{\beta}$ is identically the same, whether defined in the original model or in a transformed model such as (2.17). Thus: (1) whether or not one is using a weighted (transformed) model, and (2) whether or not one <u>should be</u> using a weighted model, the parameter defined by D_{02} in the original model is equivalent to the corresponding parameter defined by D_{T1} in the transformed model. Any D_{02}-equivalent alternative definitions in the original model are also equivalent to D_{T1} and its equivalents in the transformed model. Thus interpretation of $\underset{\sim}{\beta}$ may be made in the simplest, most intuitively appealing setting, and the interpretation applied directly to the canonical definition in the transformed model.

Figure 2.1 Incomplete, Unbalanced Part of a 3x4 Factorial Design (Searle)

Numbers of Observations, N_{ab}

Levels of Factor A		Levels of Factor B			
	b =	1	2	3	4
	a = 1	3	0	1	2
	2	2	2	0	0
	3	0	2	2	4

Cell Means: $\mu_{ab} = E[Y_{abk}]$

Levels of Factor A		Levels of Factor B			
	b =	1	2	3	4
	a = 1	μ_{11}	--	μ_{13}	μ_{14}
	2	μ_{21}	μ_{22}	--	--
	3	--	μ_{32}	μ_{33}	μ_{34}

Table 2.1 $\underset{\sim}{A}_2$, The Matrix of the Intuitive Definition of $\underset{\sim}{\beta}$

PARAMETER	OBSERVATION SUBSCRIPT																	
	111	112	113	131	141	142	211	212	221	222	321	322	331	332	341	342	343	344
	*			*	*		*		*		*		*		*			
μ_{11}	1																	
α_2	-1						1											
α_3				-1									1					
β_2							-1		1									
β_3	-1			1														
β_4	-1				1													

*Notes: 1. Each row marked with an asterisk is also a row of A_{E2}.
2. Zero values have been omitted.

Table 2.2 $\underset{\sim}{X}$, The Reference Cell Model Design Matrix for the Example

OBSERVATION SUBSCRIPT	μ_{11}	α_2	α_3	β_2	β_3	β_4
111 *	1					
112	1					
113	1					
131 *	1				1	
141	1					1
142	1					1
211 *	1	1				
212	1	1				
221 *	1	1		1		
222	1	1		1		
321 *	1		1	1		
322	1		1	1		
331 *	1		1		1	
332	1		1		1	
341 *	1		1		1	1
342	1		1		1	1
343	1		1		1	1
344	1		1		1	1

*Notes: 1. Rows marked with an asterisk are the rows of $\underset{\sim}{X}_E$.
2. Zero values have been omitted.

Table 2.3 $\underset{\sim}{A}'_{E1}$, The Transpose of the Matrix of the Essence Model Canonical Definition

OBSERVATION SUBSCRIPT	μ_{11}	α_2	α_3	β_2	β_3	β_4
111	0.8	-0.6	-0.2	-0.4	-0.7	-0.7
131	0.1	-0.2	-0.4	0.2	0.6	0.1
141	0.1	-0.2	-0.4	0.2	0.1	0.6
211	0.2	0.6	0.2	-0.6	-0.3	-0.3
221	-0.2	0.4	-0.2	0.6	0.3	0.3
321	0.2	-0.4	0.2	0.4	-0.3	-0.3
331	-0.1	0.2	0.4	-0.2	0.4	-0.1
341	-0.1	0.2	0.4	-0.2	-0.1	0.4

Table 2.4 A_{E1}, The Essence Model Canonical Definition for the Example

PARAMETER	OBSERVATION SUBSCRIPT 111	131	141	211	221	321	331	341
μ_{11}	0.8	0.1	0.1	0.2	-0.2	0.2	0.1	0.1
α_2	-0.6	-0.2	-0.2	0.6	0.4	-0.4	0.2	0.2
α_3	-0.2	-0.4	-0.4	0.2	-0.2	0.2	0.4	0.4
β_2	-0.4	0.2	0.2	-0.6	0.6	0.4	-0.2	-0.2
β_3	-0.7	0.6	0.1	-0.3	0.3	-0.3	0.4	0.1
β_4	-0.7	0.1	0.6	-0.3	0.3	-0.3	-0.1	0.4

Table 2.5 Matrices Used in Computing Definitions of Parameters for the Example

q = Number of parameters = Rank ($\underset{\sim}{X}$) = 6

NC = Number of non-empty cells = 8

1. $[\underset{\sim}{X}_E, \underset{\sim}{I}_{NC}]$*

CELL	PARAMETER μ_{11}	α_2	α_3	β_2	β_3	β_4	CELL ID 11	13	14	21	22	32	33	34
11	1						1							
13	1				1			1						
14	1					1			1					
21	1	1								1				
22	1	1		1							1			
32	1		1	1								1		
33	1		1		1								1	
34	1		1			1								1

2. After Reduction to Hermite Normal Form: $\begin{bmatrix} \underset{\sim}{I} & \underset{\sim}{A}_{E3} \\ \underset{\sim}{0} & \underset{\sim}{R} \end{bmatrix}$

PARAMETER	PARAMETER* μ_{11}	α_2	α_3	β_2	β_3	β_4	CELL ID 11	13	14	21	22	32	33	34
μ_{11}	1						1							
α_2		1					-1			1				
α_3			1				-1			1	-1	1		
β_2				1			0			-1	1			
β_3					1		-1	1						
β_4						1	-1		1					
"Zero-1"	0	0	0	0	0	0	1	-1	0	-1	1	-1	1	0
"Zero-2"	0	0	0	0	0	0	1	0	-1	-1	1	-1	0	1

*Note: Most zero values have been omitted.

3. DEFINITIONS OF SECONDARY PARAMETERS

3.1 Notation and Definitions; Definability

Much of the work of a general linear model analysis is performed via the estimation of secondary parameters and testing hypotheses about these parameters. Given the general linear model, $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$, either Full Rank (FR) or Less Than Full Rank (LTFR), secondary parameters are commonly specified in terms of the primary parameters,

$$\underset{\sim}{\theta}_1 = \underset{\sim}{C}\underset{\sim}{\beta}-\underset{\sim}{g}_1 \tag{3.1}$$

or in terms of $E(\underset{\sim}{Y})$,

$$\underset{\sim}{\theta}_2 = \underset{\sim}{H}_2 E(\underset{\sim}{Y})-\underset{\sim}{g}_2 \tag{3.2}$$

where $\underset{\sim}{C}$ and $\underset{\sim}{H}_2$ are specified matrices of constants and $\underset{\sim}{g}_1$ and $\underset{\sim}{g}_2$ are specified vectors of constants.

As with primary parameters, three basic issues arise: (1) Is a particular secondary parameter well defined? (That is, is the specification of $\underset{\sim}{\theta}$ a definition?) (2) What is a canonical definition for a definable secondary parameter? (3) Given two different definitions for a secondary parameter, are the two definitions equivalent? The following results address these issues.

DEFINITION 3.1 Well-Defined ("Definable") Secondary Parameter. In the context of the linear model $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$, a secondary parameter, $\underset{\sim}{\theta}$, specified by (3.1) or (3.2) is said to be well-defined (or "definable") if and only if for an arbitrary specified value of $E(\underset{\sim}{Y}) \in M(\underset{\sim}{X})$ there exists a unique value of $\underset{\sim}{\theta}$ which satisfied the specification of $\underset{\sim}{\theta}$. In particular:

(1) A secondary parameter, $\underset{\sim}{\theta}_2$ of the form (3.2) is well-defined and the specification (3.2) is a definition of $\underset{\sim}{\theta}_2$.

(2) A secondary parameter, $\underset{\sim}{\theta}_1$ specified by (3.1) is well-defined if and only if for arbitrary specified values of $\underset{\sim}{X}$ and $E(\underset{\sim}{Y}) \in M(\underset{\sim}{X})$ there is a unique value of $\underset{\sim}{\theta}_1$ satisfying (3.1), invariant to choice of $\underset{\sim}{\beta}$ satisfying $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$. If $\underset{\sim}{\theta}_1$ is well-defined, (3.1) is a definition of $\underset{\sim}{\theta}_1$.

In a LTFR model without side restrictions the primary parameter $\underset{\sim}{\beta}$ is not well-defined but the parameter specified by (3.1) may or may not be well-defined. Clearly, in a FR model all secondary parameters specified by (3.1) or (3.2) are well-defined.

COROLLARY 3.1.1. Let $\underset{\sim}{X}$ have full column rank in the linear model $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$ and let $\underset{\sim}{\theta}_1$, $\underset{\sim}{\theta}_2$ be specified by (3.1) and (3.2), respectively. Then (3.1) is a definition of $\underset{\sim}{\theta}_1$, (3.2) is a definition of θ_2, and both θ_1 and θ_2 are well-defined.

Proof. First, by Definition 3.1, θ_2 is well-defined and (3.2) is a definition. If X has full rank then by Corollary 2.1.1 β is well-defined: for any $E(Y) \in M(X)$ there is a unique value of β which satisfies $E(Y) = X\beta$. Therefore, since all other matrices in (3.1) are fixed constants, for any $E(Y) \in M(X)$ there is a unique value of θ_1 given by (3.1), which satisfies (3.1). QED.

The following theorem gives necessary and sufficient conditions for a LTFR model secondary parameter to be well-defined.

THEOREM 3.1. A secondary parameter, θ_1, specified by (3.1) is well-defined if and only if

(3.3a) $$C = CX^{-}X$$

or equivalently,

(3.3b) $$C \in M(X')$$

for any generalized inverse, X^{-}, if X, and where $M(A)$ denotes the space spanned by the columns of A. If θ_1 is definable any specification of the form

(3.4) $$\theta_1 = CX^{-}E(Y)-g_1$$

is a definition of θ_1.

Proof. First, θ_1 is unique iff θ_1-g_1 is unique. Now, by Rao and Mitra (1971), Theorem 2.3.1(c), $\theta_1-g_1 = C\beta$ has a unique value for each solution, β, of $E(Y) = X\beta$, iff the equivalent conditions (3.3a) and (3.3b) hold. From (3.3a) $\theta_1-g_1 = C\beta = CX^{-}X\beta = CX^{-}E(Y)$, which implies (3.4). QED.

As a result of this theorem one can observe that a secondary parameter is well-defined if and only if it is estimable, just as is the case for primary parameters. As before, "definability" stems from a different conceptual base than does estimability. Although the concepts are "tangent" at this point, the study of alternative definitions and equivalence of definitions is different from the study of estimators and their properties. As illustrated by teh proof of Theorem 3.1, however, some theory tools are useful in both studies.

3.2 Equivalence of Alternative Definitions of a Secondary Parameter

As with primary parameters it may be possible to have different, but equivalent, definitions of a secondary parameter. Since only well defined secondary parameters need be considered, there is no need to distinguish between LTFR and FR models.

The basic idea of equivalence of definitions of secondary parameters is straightforward: two definitions are equivalent iff they define the same parameter. Some additional notation will facilitate manipulation and comparison of definitions.

A secondary parameter θ specified by (3.1) or (3.2) is a vector-valued linear function of X, which is fixed, and $E(Y) \in M(X)$. Since $M(X) = \{X\beta: \beta \in E^q\}$, it is convenient to regard θ as a vector-valued linear function of β, either directly as in (3.1) or indirectly via $E(Y) = X\beta$ in (3.2). We shall use the notation $\theta = \theta(\beta)$ to denote this relationship.

Let $\theta_1(\beta)$ and $\theta_2(\beta)$ be two possibly different, well-defined secondary parameters. By varying $E(Y)$ over all of $M(X)$, or equivalently, by varying β over E^q, one generates all possible values of $\theta_1(\beta)$ and $\theta_2(\beta)$. Therefore, $\theta_1(\beta)$ and $\theta_2(\beta)$ are identical iff $\theta_1(\beta) = \theta_2(\beta)$ for all $\beta \in E^q$.

With this notation it is easy to define equivalence of secondary parameter definitions.

DEFINITION 3.2. Equivalence of Secondary Parameter Definitions. Let D_1 and D_2 be definitions of the secondary parameters $\theta_1 = \theta_1(\beta)$ and $\theta_2 = \theta_2(\beta)$ in the linear model $E(Y) = X\beta$. The definitions D_1 and D_2 are said to be equivalent, and θ_1 and θ_2 are said to be equivalently defined if and only if $\theta_1(\beta) = \theta_2(\beta)$ for all $\beta \in E^q$, i.e., iff θ_1 and θ_2 are identical.

The following theorem gives straightforward necessary and sufficient conditions for the equivalence of secondary parameter definitions.

THEOREM 3.2. Necessary and Sufficient Conditions for Equivalence of Secondary Parameter Definitions. Under the linear model $E(Y) = X\beta$ let four secondary parameters be defined by:

(3.5) D_{1j}: $\theta_{1j} = \theta_{1j}(\beta) = C_j\beta - g_1$, $j = 1,2$;

(3.6) D_{2j}: $\theta_{2j} = \theta_{2j}(\beta) = H_{2j}E(Y) - g_2$, $j = 1,2$.

Then:

(i) D_{11} and D_{12} are equivalent iff $C_1 = C_2$.

(ii) D_{21} and D_{22} are equivalent iff $H_{21}X = H_{22}X$.

(iii) D_{11} and D_{21} are equivalent iff

(a) $g_1 = g_2$, and (b) $C_1 = H_{21}X$.

Proof (i). D_{11} and D_{12} are equivalent iff

$$\underset{\sim}{0} = \underset{\sim}{\theta}_{11}(\underset{\sim}{\beta}) - \underset{\sim}{\theta}_{12}(\underset{\sim}{\beta}) \quad \text{for all } \underset{\sim}{\beta} \in E^q$$
$$= (\underset{\sim}{C}_1 - \underset{\sim}{C}_2)\underset{\sim}{\beta} \quad \text{for all } \underset{\sim}{\beta} \in E^q$$
$$\Leftrightarrow \underset{\sim}{C}_1 = \underset{\sim}{C}_2. \quad \text{QED(i).}$$

Proof (ii). D_{21} and D_{22} are equivalent iff

$$\underset{\sim}{0} = \underset{\sim}{\theta}_{21}(\underset{\sim}{\beta}) - \underset{\sim}{\theta}_{22}(\underset{\sim}{\beta}) \quad \text{for all } \underset{\sim}{\beta} \in E^q$$
$$= (\underset{\sim}{H}_{21} - \underset{\sim}{H}_{22})\underset{\sim}{X}\underset{\sim}{\beta} \quad \text{for all } \underset{\sim}{\beta} \in E^q$$
$$\Leftrightarrow \underset{\sim}{H}_{21}\underset{\sim}{X} = \underset{\sim}{H}_{22}\underset{\sim}{X}. \quad \text{QED(ii).}$$

Proof (iii). D_{11} and D_{21} are equivalent iff

$$\underset{\sim}{0} = \underset{\sim}{\theta}_{11}(\underset{\sim}{\beta}) - \underset{\sim}{\theta}_{21}(\underset{\sim}{\beta}) \quad \text{for all } \underset{\sim}{\beta} \in E^q$$
$$= (\underset{\sim}{C}_1 - \underset{\sim}{H}_{21}\underset{\sim}{X})\underset{\sim}{\beta} + (\underset{\sim}{g}_1 - \underset{\sim}{g}_2) \quad \text{for all } \underset{\sim}{\beta} \in E^q$$

Letting $\underset{\sim}{\beta} = \underset{\sim}{0}$ establishes $\underset{\sim}{g}_1 = \underset{\sim}{g}_2$; $(\underset{\sim}{C}_1 - \underset{\sim}{H}_{21}\underset{\sim}{X})\underset{\sim}{\beta} = \underset{\sim}{0}$ for all $\underset{\sim}{\beta} \in E^q \Leftrightarrow \underset{\sim}{C}_1 = \underset{\sim}{H}_{21}\underset{\sim}{X}$. QED(iii).

3.3 Canonical Definitions of Secondary Parameters

The objective of constructing a canonical definition is to provide one definition which is easily and uniquely determined from any member of a class of equivalent definitions. The canonical definition is thus a unique representative of the class of equivalent definitions.

The canonical definition of $\underset{\sim}{\beta}$ in the full rank model is easily justified: $\underset{\sim}{\beta} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'E(\underset{\sim}{Y})$ is the unique solution of the model equation $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$. Alternative, equivalent defintions of $\underset{\sim}{\beta}$ necessarily involve another matrix, $\underset{\sim}{K}$, as shown in Corollary 2.4.3. Thus, the canonical definition of $\underset{\sim}{\beta}$ is not only the unique solution of the model equation but is also, in one sense, the most concise definition in the class of equivalent definitions.

The justification of the canonical definition of a secondary parameter lies primarily in the fact that it is a unique definition in the sense that, beginning with any equivalent alternative definition, application of a simple formula always results in the same canonical definition. This statement is proven in the following theorem.

THEOREM 3.3. Let four equivalent definitions of a secondary parameter $\underset{\sim}{\theta}_2$ be given by:

$$D_{1_j}: \ \underset{\sim}{\theta} = \underset{\sim}{C}_j\underset{\sim}{\beta} - \underset{\sim}{g}, \quad j = 1,2$$
$$D_{2_j}: \ \underset{\sim}{\theta} = \underset{\sim}{H}_j E(\underset{\sim}{Y}) - \underset{\sim}{g}, \quad j = 1,2.$$

Then,

$$(3.7)\qquad \underset{\sim}{H} = \underset{\sim}{H}_1\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}' = \underset{\sim}{H}_2\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}' = \underset{\sim}{C}_1\,(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}' = \underset{\sim}{C}_2\,(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}'$$

for any generalized inverse, $(\underset{\sim}{X}'\underset{\sim}{X})^-$ of $\underset{\sim}{X}'\underset{\sim}{X}$.

Proof. By equivalence and Theorem 3.2,

$$\underset{\sim}{H}_1\underset{\sim}{X} = \underset{\sim}{H}_2\underset{\sim}{X} = \underset{\sim}{C}_1 = \underset{\sim}{C}_2.$$

Postmultiplying each of these terms by $(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}'$ produces $\underset{\sim}{H}$. Note that $\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}'$ is invariant to choice of $(\underset{\sim}{X}'\underset{\sim}{X})^-$. Also, by Theorem 3.1, $\underset{\sim}{C}_1 = \underset{\sim}{C}_2 \in M(\underset{\sim}{X}') \Rightarrow \underset{\sim}{C}_j = \underset{\sim}{A}\underset{\sim}{X}$, so $\underset{\sim}{C}_j(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}' = \underset{\sim}{A}\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}'$ is also invariant to choice of $(\underset{\sim}{X}'\underset{\sim}{X})^-$. QED.

Theorem 3.3 and the preceeding discussion motivate the following.

DEFINITION 3.3. Canonical Definition of a Secondary Parameter. (1) Let a secondary parameter, $\underset{\sim}{\theta}$, in the model $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$ be defined by

$$(3.8)\qquad D_1:\ \underset{\sim}{\theta} = \underset{\sim}{C}\underset{\sim}{\beta}-\underset{\sim}{g}$$

Then the canonical definition of $\underset{\sim}{\theta}$ is

$$(3.9)\qquad D_1^*:\ \underset{\sim}{\theta} = [\underset{\sim}{C}(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}']\,E(\underset{\sim}{Y})-\underset{\sim}{g}$$

(2) Let a secondary parameter, $\underset{\sim}{\theta}$, in the model $E(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}$ be defined by

$$(3.10)\qquad D_2:\ \underset{\sim}{\theta} = \underset{\sim}{H}E(\underset{\sim}{Y})-\underset{\sim}{g}$$

Then the canonical definition of $\underset{\sim}{\theta}$ is

$$(3.11)\qquad D_2^*:\ \underset{\sim}{\theta} = [\underset{\sim}{H}\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}']\,E(\underset{\sim}{Y})-\underset{\sim}{g}$$

In each case $(\underset{\sim}{X}'\underset{\sim}{X})^-$ is an arbitrary generalized inverse of $(\underset{\sim}{X}'\underset{\sim}{X})$.

COROLLARY 3.3.1. In Definition 3.3, D_1 and D_1^* are equivalent and D_2 and D_2^* are equivalent, i.e., a secondary parameter definition is equivalent to its canonical definition.

Proof. For D_1 and D_1^* note that $(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}'$ is a generalized inverse of $\underset{\sim}{X}$, i.e., $\underset{\sim}{X}^- = (\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}'$. Let $\underset{\sim}{H}^* = \underset{\sim}{C}\underset{\sim}{X}^-$ in D_1^*; using (3.3a) we have $\underset{\sim}{H}^*\underset{\sim}{X} = \underset{\sim}{C}\underset{\sim}{X}^-\underset{\sim}{X} = \underset{\sim}{C}$. Thus D_1 and D_1^* are equivalent by Theorem 3.2(iii). Now, D_2 and D_2^* are equivalent, by Theorem 3.2(ii) iff $\underset{\sim}{H}\underset{\sim}{X}$ is equal to $[\underset{\sim}{H}\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}']\underset{\sim}{X} = \underset{\sim}{H}\underset{\sim}{X}\underset{\sim}{X}^-\underset{\sim}{X} = \underset{\sim}{H}\underset{\sim}{X}$. QED.

It is interesting that although $(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}'$ is a generalized inverse for $\underset{\sim}{X}$, the expression $\underset{\sim}{C}\underset{\sim}{X}^-$ $[\underset{\sim}{C} \in M(\underset{\sim}{X}')]$ is not invariant to choice of $\underset{\sim}{X}^-$, but $\underset{\sim}{C}(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}'$ is invariant to choice of $(\underset{\sim}{X}'\underset{\sim}{X})^-$. Thus, in a sense, (3.9) represents the simplest g.i.-invariant canonical definition of $\underset{\sim}{\theta}$ in terms of $\underset{\sim}{C}$, $\underset{\sim}{X}$, and $E(\underset{\sim}{Y})$. Similarly since $\underset{\sim}{X}\underset{\sim}{X}^-$ is not invariant to choice of X^-, but $X(X'X)^-X'$ is invariant to choice of $(X'X)^-$, (3.11) is the simplest g.i.-invariant canonical definition of θ from definition D_2 (3.10).

The following corollary shows that an expression which appears to be slightly simpler than (3.9) is, in fact, identical.

COROLLARY 3.3.2. In the setting of Definition 3.2, $C(X'X)^-X' = CX^+$ where X^+ is the Moore-Penrose g.i. of X. Thus an identical expression for (3.9) is:

$$D_1^*: \ \theta = CX^+E(Y)-g.$$

Proof. Graybill (1976), Theorem 1.5.26, shows $XX^+ = X(X'X)^-X'$. Now $C \in M(X') \Rightarrow C = AX \Rightarrow C(X'X)^-X' = AX(X'X)^-X' = AXX^+ = CX^+$. QED.

COROLLARY 3.3.3. In the context of Theorem 3.2, two secondary parameter definitions of the forms (3.5) or (3.6) are equivalent if and only if they have the same canonical definition.

Proof (i). Using the notation of Theorem 3.2 and Corollary 3.3.1, the canonical definitions of θ_{1j} are

$$\theta_{1j} = H_{1j}E(Y)-g \ , \quad j = 1,2$$

where

$$H_{1j} = C_j(X'X)^-X' \ , \quad j = 1,2.$$

By Theorem 3.2, D_{11} and D_{12} are equivalent iff $C_1 = C_2$. Clearly $C_1 = C_2 \Rightarrow H_{11} = H_{12}$, i.e., θ_{11} and θ_{12} have the same canonical definition. Suppose θ_{11} and θ_{12} have the same canonical definition, i.e., $H_{11} = H_{12}$, then

$$\theta_{11}(\beta)-\theta_{12}(\beta) = (H_{11}-H_{12})X\beta = 0$$

for all $\beta \in E^q \Rightarrow \theta_{11}$ and θ_{12} are identical and D_{11} and D_{12} are equivalent. QED(i).

Proof (ii). The canonical definitions of θ_{2j} are

$$\theta_{2j} = H_{2j}^*E(Y)-g.$$

where

$$H_{2j}^* = H_{2j}(X'X)^-X' \ , \quad j = 1,2.$$

If D_{21} and D_{22} are equivalent, by original specification (3.6) and Theorem 3.2, $H_{21}X = H_{22}X$, which implies $H_{21}^* = H_{22}^*$, i.e., θ_{21} and θ_{22} have the same canonical definition. If θ_{21} and θ_{22} have the same canonical definition, i.e., $H_{21}^* = H_{22}^*$, then

$$\theta_{21}(\beta)-\theta_{22}(\beta) = (H_{2j}^*-H_{1j}^*)X\beta = 0$$

for all $\beta \in E^q \Rightarrow D_{21}$ and D_{22} are equivalent. QED(ii).

Proof (iii). The canonical definitions of θ_{11} and θ_{21} are

$$\theta_{i1} = H_{i1}^{*} E(Y)-g_i , \quad i = 1,2$$

with

$$H_{11}^{*} = C_1(X'X)^{-}X'$$
$$H_{21}^{*} = H_{21}X(X'X)^{-}X'.$$

If D_{11} and D_{21} are equivalent then (Theorem 3.2(iii)) $g_1 = g_2$ and $C_1 = H_{21}X \Rightarrow H_{11}^{*} = H_{21}^{*}$, i.e., the canonical definitions are identical. if the canonical definitions are identical, i.e., $H_{11}^{*} = H_{21}^{*}$ and $g_1 = g_2$, then $\theta_{11}(\beta)-\theta_{21}(\beta) = (H_{11}^{*}-H_{21}^{*})X\beta$ for all $\beta \Rightarrow$ D_{11} and D_{21} are equivalent. QED(iii).

The following result is useful in the next section:

COROLLARY 3.3.4. A secondary parameter defined by (3.5) has an equivalent definition in the form (3.6), and vice versa. [The canonical definition of (3.5) is of the form (3.6). For θ_{2j} of the form (3.6), let $C_j = H_{2j}X$ to obtain the form (3.5); equivalence follows from Theorem 3.2(iii).]

3.4 Definitions and Equivalence in the Essence Model

All of the definitions and results in the preceeding sections may be applied directly to the essence model, i.e., the linear model for the first observation from each cell, where all observations from a cell have the same expected value.

Parameter definitions made in the essence model may be extended to equivalent definitions in the full model. Essence model canonical definitions are equivalent to, but generally not identical to, their extensions to the full model. The following result establishes most of the possibilities.

THEOREM 3.3. In the context of the model $E(Y) = X\beta$, with X Nxq, let U be an NCxN matrix, NC<N, whose rows are a subset of the rows of I_N hence, $UU' = I_{NC}$ such that $Y_E = UY$ is the vector of first observations from the cells. Define $X_E = UX$; the essence model is $E(Y_E) = X_E\beta$. Let the following be equivalent secondary parameter definitions in the essence model:

$$D_{E1}: \ \theta = C\beta-g$$
$$D_{E2}: \ \theta = H_E E(Y_E)-g$$

It follows that $C = H_E X_E = H_E UX = HX$ with $H = H_E U$, and that the essence model canonical definition of θ may be stated in the two equivalent forms

$$D^*_{E1}: \underset{\sim}{\theta} = \underset{\sim}{C}(\underset{\sim}{X}_E{}'\underset{\sim}{X}_E)^- \underset{\sim}{X}_E\, E(\underset{\sim}{Y}_E)-\underset{\sim}{g}$$

$$D^*_{E2}: \underset{\sim}{\theta} = \underset{\sim}{H}_E\underset{\sim}{X}_E(\underset{\sim}{X}_E{}'\underset{\sim}{X}_E)^- \underset{\sim}{X}_E{}'\, E(\underset{\sim}{Y}_E)-g.$$

Define:

$$D^*_1: \underset{\sim}{\theta} = \underset{\sim}{C}(\underset{\sim}{X}'\underset{\sim}{X})^- \underset{\sim}{X}' E(\underset{\sim}{Y})-\underset{\sim}{g}$$

$$D^*_2: \underset{\sim}{\theta} = \underset{\sim}{H}\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^- \underset{\sim}{X}' E(\underset{\sim}{Y})-\underset{\sim}{g}$$

$$D_2: \underset{\sim}{\theta} = \underset{\sim}{H}E(\underset{\sim}{Y})-\underset{\sim}{g}.$$

Then all of D_{E1}, D_{E2}, D^*_{E1}, D^*_{E2}, D^*_1, D^*_2, and D_2 are equivalent definitions of $\underset{\sim}{\theta}$.

Proof. The equivalence of D_{E1} and D_{E2} is assumed; the equivalence of D_{E1} and D_{E2} to D^*_{E1} and D^*_{E2} follow from Corollary 3.2.1. Since $\underset{\sim}{g}$ is the same vector throughout it may be ignored. Now, by Theorem 3.1, $\underset{\sim}{C} = \underset{\sim}{C}\underset{\sim}{X}_E{}^-\underset{\sim}{X}_E \iff \underset{\sim}{C} \in M(\underset{\sim}{X}_E{}') \iff \exists\, \underset{\sim}{A}_E$ s.t. $\underset{\sim}{C} = \underset{\sim}{A}_E\underset{\sim}{X}_E = \underset{\sim}{A}_E\underset{\sim}{U}\underset{\sim}{X} \iff \underset{\sim}{C} \in M(\underset{\sim}{X}') \iff \underset{\sim}{C} = \underset{\sim}{C}\underset{\sim}{X}^-\underset{\sim}{X}$. Consider the difference of the parameters defined by D_{E1} and D^*_1:

$$(\underset{\sim}{C}\underset{\sim}{\beta}-\underset{\sim}{g}) - [\underset{\sim}{C}(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}'\, E(\underset{\sim}{Y}) - \underset{\sim}{g}]$$

$$= \underset{\sim}{C}\underset{\sim}{\beta}-\underset{\sim}{C}\underset{\sim}{X}^-\underset{\sim}{X}\underset{\sim}{\beta} = (\underset{\sim}{C}-\underset{\sim}{C}\underset{\sim}{X}^-\underset{\sim}{X})\underset{\sim}{\beta} = \underset{\sim}{0}.$$

The two parameters are identical for all $\underset{\sim}{\beta} \in E^q$, i.e., D_{E1} and D^*_1 are equivalent. By the equivalence of D_{E1} and D_{E2}, $\underset{\sim}{C} = \underset{\sim}{H}_E\underset{\sim}{X}_E$. Simple manipulations show $\underset{\sim}{H}_E = \underset{\sim}{H}\underset{\sim}{U}'$; thus $\underset{\sim}{C} = \underset{\sim}{H}_E\underset{\sim}{X}_E = \underset{\sim}{H}\underset{\sim}{U}'\underset{\sim}{U}\underset{\sim}{X} = \underset{\sim}{H}\underset{\sim}{X}$, which establishes the equivalence of D_{E1} and D_2 and, therefore, of D^*_2 which is the canonical definition of D_2. We have established a chain such that for all $\underset{\sim}{\beta} \in E^q$, all six definitions of $\underset{\sim}{\theta}$ produce identical values of $\underset{\sim}{\theta}$. All the definitions are therefor equivalent. QED.

Secondary parameter definitions are often more conveniently constructed in the essence model than in the full model because the vector of cell means, $\underset{\sim}{\mu} = E(\underset{\sim}{Y}_E)$ has fewer elements than $E(\underset{\sim}{Y})$. There is no difference for secondary parameters of the form $\underset{\sim}{\theta} = \underset{\sim}{C}\underset{\sim}{\beta}-\underset{\sim}{g}$, of course.

3.5 Examples

As a first example reconsider the design and model presented in section 2.5. We shall work with the essence model to keep the matrices small.

The first obvious secondary parameters to define are the A and B main effects, defined in terms of the primary parameters:

$$D_{A^1}: \underset{\sim}{\theta}_A = \underset{\sim}{C}_A\underset{\sim}{\beta} = \begin{bmatrix}\alpha_2\\ \alpha_3\end{bmatrix}$$

$$D_{B^1}: \underset{\sim}{\theta}_B = \underset{\sim}{C}_B\underset{\sim}{\beta} = \begin{bmatrix}\beta_1\\ \beta_2\\ \beta_3\end{bmatrix}$$

where

$$\begin{matrix} & \mu_{11} & \alpha_2 & \alpha_3 & \beta_2 & \beta_3 & \beta_4 \end{matrix}$$

$$\underset{\sim}{C}_A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$\underset{\sim}{C}_B = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

An obvious second set of definitions of these effects may be made in terms of cell means:

$$\mu' = (\mu_{11} \quad \mu_{13} \quad \mu_{14} \quad \mu_{21} \quad \mu_{22} \quad \mu_{32} \quad \mu_{33} \quad \mu_{34})$$

$$\underset{\sim}{H}_{A2} = \begin{bmatrix} -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\underset{\sim}{H}_{B2} = \begin{bmatrix} 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$D_{A2}: \ \underset{\sim}{\theta}_A = \underset{\sim}{H}_{A2}\underset{\sim}{\mu} = \underset{\sim}{H}_{A2}\, E(\underset{\sim}{Y}_E) = \begin{bmatrix} \mu_{21}-\mu_{11} \\ \mu_{33}-\mu_{13} \end{bmatrix}$$

$$D_{B2}: \ \underset{\sim}{\theta}_B = \underset{\sim}{H}_{B2}\underset{\sim}{\mu} = \begin{bmatrix} \mu_{22}-\mu_{21} \\ \mu_{13}-\mu_{11} \\ \mu_{14}-\mu_{11} \end{bmatrix}$$

Since definitions D_{A2} and D_{B2} coincide with original intuitive definitions of the parameters $(\alpha_2, \alpha_3)'$ and $(\beta_2, \beta_3, \beta_4)'$, one would expect equivalence between D_{A1} and D_{A2} and between D_{B1} and D_{B2}. This equivalence is easily verified by demonstrating that $\underset{\sim}{C}_A = \underset{\sim}{H}_{A2}\underset{\sim}{X}_E$ and $\underset{\sim}{C}_{B^2} = \underset{\sim}{H}_{B2}\underset{\sim}{X}_E$.

Since the model assumes zero interactions we may define, as secondary parameters, cell means for the empty cells. Some obvious definitions are:

$$\mu_{12} = \mu_{11}+\beta_2 = \mu_{11}+\mu_{22}-\mu_{21}$$
$$\mu_{23} = \mu_{11}+\alpha_2+\beta_3 = \ldots = -\mu_{11}+\mu_{13}+\mu_{21}$$
$$\mu_{24} = \mu_{11}+\alpha_2+\beta_4 = \ldots = -\mu_{11}+\mu_{14}+\mu_{21}$$
$$\mu_{31} = \mu_{11}+\alpha_3 = \mu_{11}-\mu_{13}+\mu_{33}$$

Given these secondary parameters it is interesting to define "average distance between the lines" main effects for A and B. These are defined as:

$$D_{A3}: \ \underset{\sim}{\theta}_A = 1/4 \sum_{b=1}^{4} \begin{bmatrix} \mu_{2b} - \mu_{1b} \\ \mu_{3b} - \mu_{1b} \end{bmatrix} = \underset{\sim}{H}_{A3}\,\underset{\sim}{\mu}$$

$$D_{B3}: \underset{\sim}{\theta}_B = 1/3 \sum_{a=1}^{3} \begin{bmatrix} \mu_{a2} - \mu_{a1} \\ \mu_{a3} - \mu_{a1} \\ \mu_{a4} - \mu_{a1} \end{bmatrix} = \underset{\sim}{H}_B \underset{\sim}{\mu}$$

An overall mean is also interesting:

$$D_{\mu_3}: \theta_\mu = 1/12 \sum_a \sum_b \mu_{ab} = \underset{\sim}{H}_{\mu_3}\underset{\sim}{\mu}$$

After some minor calculations, the matrices of these definitions are shown to be:

$$\begin{array}{ll} & \begin{array}{cccccccc} \mu_{11} & \mu_{13} & \mu_{14} & \mu_{21} & \mu_{22} & \mu_{32} & \mu_{33} & \mu_{34} \end{array} \\ \underset{\sim}{H}_{A3} = 1/4 & \begin{bmatrix} -4 & 0 & 0 & 4 & 0 & 0 & 0 & 0 \\ -1 & -2 & -1 & 1 & -1 & 1 & 2 & 1 \end{bmatrix} \end{array}$$

$$\underset{\sim}{H}_{A3} = 1/3 \begin{bmatrix} -1 & 1 & 0 & -2 & 2 & 1 & -1 & 0 \\ -3 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ -3 & 1 & 2 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

$$\underset{\sim}{H}_{\mu_3} = 1/12 \begin{bmatrix} 1 & 1 & 2 & 2 & 2 & 1 & 2 & 1 \end{bmatrix}$$

Simple matrix multiplication demonstrates that $\underset{\sim}{C}_A = \underset{\sim}{H}_{A3}\underset{\sim}{X}_E$ (D_{A1} and D_{A3} are equivalent) and $\underset{\sim}{C}_B = \underset{\sim}{H}_{B3}\underset{\sim}{X}_E$ (D_{B1} and D_{B3} are equivalent), as expected. One can compute $\underset{\sim}{C}_\mu$ = $\underset{\sim}{H}_{B3}\underset{\sim}{X}_E$ = 1/12 [12 4 4 3 3 3], i.e., the equivalent definition of $\underset{\sim}{\theta}_\mu$ in terms of $\underset{\sim}{\beta}$ is

$$\theta_\mu = \mu_{11} + (\alpha_2+\alpha_3)/3 + (\beta_2+\beta_3+\beta_4)/4.$$

This definition appeals to one's intuition if one recalls that, by definition, $\alpha_1 = 0$ and $\beta_1 = 0$; the terms in parentheses are averages of the α_i and β_j, respectively.

<u>A Reparameterized Model</u>. A well-known method for obtaining a full rank model in ANOVA problems is to "reparameterize" the LTFR ANOVA model for the example described above and in section 2.5; we continue to omit interactions. One device for reparameterization, based on the notation of "sum to zero" restrictions, is as follows.

Require:

$$\alpha_1+\alpha_2+\alpha_3 = 0 \Rightarrow \alpha_1 = -(\alpha_2+\alpha_3)$$

$$\beta_1+\beta_2+\beta_3+\beta_4 = 0 \Rightarrow \beta_1 = -(\beta_2+\beta_3+\beta_4)$$

One omits α_1 and β_1 from the model. When $E[Y_{abk}]$ involves α_1 or β_1, the term is replaced by the expression above. Thus for example:

LTFR Model: $E[Y_{11k}] = \mu+\alpha_1+\beta_1$

Reparameterized Model: $E[Y_{11k}] = \mu-\alpha_2-\alpha_3-\beta_2-\beta_3-\beta_4$

Application of this technique to the "essence" LTFR model produces the $\underset{\sim}{X}_E$ shown in Table 3.1.

A typical interpretation of the parameter α_2 is: α_2 is a comparison of the second level of A vs. (minus) the first level of A. Similarly, α_3 is often interpreted as comparing the third vs. the first levels of A. Similar interpretations are applied to the β_j and levels of B. The parameter μ is interpreted as an overall mean. We shall examine these interpretations.

Define the following secondary parameters (which are just subsets of the primary parameters):

$$D_{A1}: \quad \underset{\sim}{\theta}_A = \begin{bmatrix} \alpha_2 \\ \alpha_3 \end{bmatrix} = \underset{\sim}{C}_A \underset{\sim}{\beta}$$

$$D_{B1}: \quad \underset{\sim}{\theta}_B = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \underset{\sim}{C}_B \underset{\sim}{\beta}$$

$$D_{\mu}: \quad \theta_{\mu} = \mu = \underset{\sim}{C}_{\mu} \underset{\sim}{\beta}$$

The matrices $\underset{\sim}{C}_A$, $\underset{\sim}{C}_B$ and $\underset{\sim}{C}_\mu$ are shown in Table 3.1.

If the interpretations given above are correct, the following definitions, in terms of cell means, should be equivalent to the ones above:

$$D_{A2}: \quad \underset{\sim}{\theta}_A = \begin{bmatrix} \mu_{21}-\mu_{11} \\ \mu_{33}-\mu_{13} \end{bmatrix} = \underset{\sim}{H}_A \underset{\sim}{\mu}$$

$$D_{B2}: \quad \underset{\sim}{\theta}_B = \begin{bmatrix} \mu_{22}-\mu_{21} \\ \mu_{13}-\mu_{11} \\ \mu_{14}-\mu_{11} \end{bmatrix} = \underset{\sim}{H}_B \underset{\sim}{\mu}$$

$$D_{\mu 2}: \quad \underset{\sim}{\theta}_\mu = (1/12) \sum_a \sum_b \mu_{ab} = \underset{\sim}{H}_\mu \underset{\sim}{\mu}$$

where the μ_{ab} for missing cells are constructed as above. The matrices $\underset{\sim}{H}_A$ and $\underset{\sim}{H}_B$ are identical to the matrices $\underset{\sim}{H}_{A2}$, $\underset{\sim}{H}_{B2}$ presented earlier. $\underset{\sim}{H}_\mu$ is indentical to $\underset{\sim}{H}_{\mu 3}$, above.

The matrices $\underset{\sim}{C}_{A3} = \underset{\sim}{H}_A \underset{\sim}{X}_E$, $\underset{\sim}{C}_{B3} = \underset{\sim}{H}_B \underset{\sim}{X}_E$ and $\underset{\sim}{C}_{\mu 3} = \underset{\sim}{H}_\mu \underset{\sim}{X}_E$ are:

Parameter:	μ	α_2	α_3	β_2	β_3	β_4
$\underset{\sim}{C}_{A3}$ =	0	2	1	0	0	0
	0	1	2	0	0	0
$\underset{\sim}{C}_{B3}$ =	0	0	0	2	1	1
	0	0	0	1	2	1
	0	0	0	1	1	2
$\underset{\sim}{C}_{\mu 3}$ =	1	0	0	0	0	0

These matrices are compared with $\underset{\sim}{C}_A$, $\underset{\sim}{C}_B$, and $\underset{\sim}{C}_\mu$ in Table 3.1.

First, $\underset{\sim}{C}_\mu = \underset{\sim}{C}_\mu$ shows that the parameter μ is truly an overall mean in the sense described by $D_{\mu 2}$ above, where secondary parameters are used for means for empty cells.

Obviously $\underset{\sim}{C}_A \neq \underset{\sim}{C}_{A3}$ and $\underset{\sim}{C}_B \neq \underset{\sim}{C}_{B3}$. Thus for example, the parameter α_2 is not a comparison of the second level of A vs. the first level of A; that comparison is made by (from the first row of $\underset{\sim}{C}_{A3}$) $2\alpha_2+\alpha_3 = \mu_{21}-\mu_{11}$. None of the α_2 or β_b have the interpretations suggested above! Although this type of reparameterization is a useful technique, especially if one wishes only to test hypotheses, interpretation of the parameter estimates from such a model requires considerable care.

Remarks on Interpretation of Results. The canonical definition of a secondary parameter is primarily useful in concept rather than application: it is useful to have one unique definition which results from application of a simple algorithm to any member of a class of equivalent definitions. In practice the coefficients in the canonical definition matrix are often too complicated for direct interpretation; there are simpler, equivalent definitions which are much more useful for interpretation.

However, the formal equivalence of definability and estimability becomes useful, for the coefficients of the canonical definition of $\underset{\sim}{\theta}$ are the same as the coefficients of the BLUE of $\underset{\sim}{\theta}$, under the assumption $V(\underset{\sim}{Y}) = \sigma^2 \underset{\sim}{I}_N$. That is, if the canonical definition of $\underset{\sim}{\theta}$ is D: $\underset{\sim}{\theta} = \underset{\sim}{H}E(\underset{\sim}{Y})-\underset{\sim}{g}$, then $\hat{\underset{\sim}{\theta}} = \underset{\sim}{H}\underset{\sim}{Y}-\underset{\sim}{g}$ is the BLUE of $\underset{\sim}{\theta}$.

One could insist upon a data oriented interpretation, in contrast to the parameter oriented approach taken in this paper, in which one would interpret $\hat{\underset{\sim}{\theta}}$ in terms of the coefficients expressing $\hat{\underset{\sim}{\theta}}$ as a vector-valued linear function of $\underset{\sim}{Y}$. Since the same coefficients, the matrix $\underset{\sim}{H}$ above, also form the canonical definition of $\underset{\sim}{\theta}$ the two interpretative approaches are very similar at this point. An advantage of the parametric approach is that there are equivalent, simpler definitions of $\underset{\sim}{\theta}$ which often permit easy, straightforward interpretation. In contrast, while there are unbiased estimators other than the BLUE, $\hat{\underset{\sim}{\theta}}$, those estimators are not equivalent to $\hat{\underset{\sim}{\theta}}$ from an estimation point of view and it would not be appropriate, in a data oriented approach, to interpret $\hat{\underset{\sim}{\theta}}$ in terms of non-equivalent estimators.

One can usefully combine the two approaches. For example, in the reference cell model example the BLUE $\hat{\alpha}_2$ is estimating (from Table 2.3)

$$E[\hat{\alpha}_2] = E[-3/14\, y_{11\bullet} - 5/42 y_{131} - 5/42 y_{14\bullet} + 9/28 y_{21\bullet} + 5/28 y_{22} - 5/28 y_{32\bullet} + 5/84 y_{33\bullet} + 5/84 y_{34\bullet}]$$

where $y_{ab\bullet} = \sum_{k=1}^{N_{ab}} y_{abk}$. However, by adding information from alternative definitions of α_2, from H_{A2} or H_{A3} above, one can note that

$$E[\hat{\alpha}_2] = \mu_{21}-\mu_{11} = E[y_{21k}] - E[y_{11k}\cdot]$$

subject to the no-interaction assumption. Surely the latter expression for $E[\hat{\alpha}_2]$ is easier to interpret than the former. (Even the latter has pitfalls; the no interaction assumption means $\bar{y}_{21}. \neq \hat{\mu}_{21}.$)

Table 3.1 X_E for the Full Rank Reparameterized Essence Model and C Matrices for First Definitions

PARAMETER

CELL ID	μ	α_2	α_3	β_2	β_3	β_4
11	1	-1	-1	-1	-1	-1
13	1	-1	-1	0	1	0
14	1	-1	-1	0	0	1
21	1	1	0	-1	-1	-1
22	1	1	0	1	0	0
32	1	0	1	1	0	0
33	1	0	1	0	1	0
34	1	0	1	0	0	1

$= X_E$

$$C_A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$C_B = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$C_\mu = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$