Journal of Official Statistics
Vol. 3, No. 2, 1987, pp. 169–181

Classification and Properties of Rotation Sampling Designs

Ronaldo Iachan[1] *and Byron Jones*[2]

Abstract: For repeated sampling in time, a general class of rotation sampling designs is introduced. The construction and classification of such designs are followed by a detailed examination of their properties. For this purpose, an additive model is proposed that takes into account period and sampling unit effects and permits efficiency comparisons.

Key words: Rotation sampling design; efficiency; variance of period changes; construction algorithm; classification; properties.

1. Introduction

In this paper we consider the construction, classification, and properties of a class of rotation sampling designs, drawing an analogy with incomplete block designs. Efficiency comparisons are made using a linear additive model for survey responses.

Rotation sampling designs perform well when one wishes to minimize sampling and nonsampling errors. As for sampling errors, current levels are best estimated (in the sense of smallest sampling variance) when there is some overlap between successive samples. Partially overlapping samples provide a useful compromise between the complete overlap (same unit interviewed over and over), required for the most precise estimates of period-to-period changes, and the drawing of a new sample (no overlap) needed for the most precise overall (aggregate) estimates. See Cochran (1977, pp. 344–347) for related arguments and some classical developments for sampling on two or more occasions with simple random sampling. With more than two occasions, finding optimal designs becomes extremely complicated (even in a restricted and approximate sense). Patterson (1950) and Eckler (1955) have developed the basic theory for this case. Extensions to more general designs and estimators are given in Ghangurde and Rao (1969), Jain (1981), and Wolter (1979). As for non-sampling errors, the rotation aspect of the design avoids excessive respondent burden and sample attrition that lead to increased response errors and nonresponse errors.

The model-based approach of this paper permits straightforward efficiency computations when more than two occasions are considered. Recommendations as to which

[1] Research Triangle Institute, Center for Statistics Research, Research Triangle Park, North Carolina 27709–2194, U.S.A.
[2] Mathematical Institute, The University, Canterbury, Kent, CT2 7NF, England.
Acknowledgements: We are most grateful to the Editor and the referees for their helpful suggestions which have improved the presentation of the paper.

design to use in what circumstances follow from this methodology, so that design choices can be tailored to fit survey priority objectives.

In this paper we use some properties of block designs that are used in comparative experiments to derive rotation sampling designs for surveys. A block design is used to compare a number of "treatments," e.g., different drugs, different fertilizers, etc. The available experimental units, e.g., human volunteers, plots of agricultural land, etc., are divided up into "blocks" so that the units within a block are as similar as possible. The units in different blocks may then differ greatly. In the analysis of the data obtained from block experiment, the differences between the treatments are estimated from within-block comparisons. In an incomplete block design the number of units in a block is smaller than the number of treatments. This means that the estimates of the treatment differences must be adjusted for the differences between the blocks. An instructive review of the similarities between experimental design and sample survey methodologies is provided by Fienberg and Tanur (1985). They draw analogies between the design concepts in the two areas such as randomization/probability sampling, blocking/stratification and split-plots/clusters. Then they discuss the use of balanced incomplete block designs as a means of achieving a restricted randomization in sampling by reducing the "support" of a sampling plan (Chakrabarti (1963)). The support is the set of samples with positive selection probabilities. These authors then review analogies between the model-based analyses used in the two areas: while Model I or fixed effects linear models have seen several applications in sampling, Model II or random effects models are used rarely in the sampling literature (exceptions include Hartley and Rao (1978) and Fuller and Harter (1985)). Our paper is something of a synthesis in that linear models inspired by incomplete block designs are used to guide the search for optimal sampling designs. We use the framework of what are known as cyclic incomplete block designs to obtain alternatives to the symmetric designs used in sample surveys. Our designs reduce the response burden.

2. Rotation Sampling Designs

Many large-scale surveys are repeated periodically using a rotation sampling design. In such a design, a total sample is divided into b rotation groups. A rotation group can consist of a fixed number of primary sampling units (PSU's), segments within PSU's, or last stage units. Let rotation group j, where $j=1, 2, \ldots, b$, be interviewed in each of the k_j periods. The total number of periods in the design is t and $k_j \leq t$, for all j. Furthermore, let f be the constant fraction of the groups interviewed in one period that are also interviewed in the next period; the remaining fraction $(1-f)$ of the groups are replaced. The cyclical nature of our designs should be noted. In each period, r groups are interviewed, and this requires that a particular group returns to the sample.

The design used in the National Crime Survey (Fienberg (1980)), for example, has this type of overlap pattern (slightly modified to accommodate bounding of interviews). Other rotation patterns are used in the Current Population Survey carried out by the U.S. Bureau of the Census (1978) and in Statistics Canada's Labor Force Survey (Ghangurde (1982)).

A convenient representation of this design is given by the *incidence matrix* N, which has elements n_{ij}, $i=1, 2, \ldots, t$; $j=1, 2, \ldots, b$. The element $n_{ij}=1$ if group j is interviewed in period i, and is zero otherwise. The analogy with the construction of block designs, as used in comparative experiments, is that

b = the number of blocks,
t = the number of treatments,
r = the number of units which receive each particular treatment,
k_j = the number of units in block j, and
n_{ij} = the number of units in block j which receive treatment i.

The class considered here is an important subclass of the class of cyclic designs (see, e.g., John et al. (1972)).

3. Construction

A rotation design is completely specified by its incidence matrix N. The basic designs considered in this paper have an incidence matrix that is constructed as follows. Given b, r, and a "shift parameter" s:

1. Set the first r entries of the first row equal to 1 and the rest to 0. That is, set $n_{1j}=1$; $j=1, 2, \ldots, r$ and $n_{1j}=0$ for $j=r+1, \ldots, b$.
2. Construct the second row of N by shifting all entries in the first row s places to the right in a cyclical manner. That is, the $(p+s)$ th element of row 2 is set equal to the p th element of row 1, where $(p+s)$ is taken modulo b, and $p=1, 2, \ldots, b$.
3. Rows 3, 4, ... are constructed in exactly the same way from their directly preceding rows.
4. Eventually row 1 will occur in the $(t+1)$ th row. This last row is discarded to leave an incidence matrix with t rows and b columns. Then we see that the value of k_j is the number of ones in column j;$j=1, 2, \ldots, b$.

These types of designs were constructed for $b=1, 2, \ldots, 12$. For each value of b, designs were systematically obtained by varying r from 1 to b–1, and for each r by varying s from 1 to r. As an example, Table 1 gives the incidence matrix for $b=8$, $r=5$, and $s=2$, constructed as explained above. This design has $t=4$, $k_1=k_3=k_5=k_7=3$, and $k_2=k_4=k_6=k_8=2$. More realistic examples will be given in the following sections.

Table 1. Incidence Matrix for $b=8$, $r=5$ and $s=2$.

1	1	1	1	1	0	0	0
0	0	1	1	1	1	1	0
1	0	0	0	1	1	1	1
1	1	1	0	0	0	1	1

An additional parameter that is useful for describing a design is the "overlap," l, between rows. In Table 1 the overlap is $l=3$. It will be noted that $f=\frac{l}{r}$. As will be seen in the following, most designs have $k_j=k$, a constant, for all j. Designs for which the incidence matrices are the same (except for a permutation of their columns) are considered equivalent, since they correspond to a simple relabelling of the groups.

4. Classification of the Designs

In order to classify the rotation designs, we introduce a linear model for the characteristic, Y, that is measured in the survey. For $i=1, 2, \ldots, t$ and $j=1, 2, \ldots, b$, we assume

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, \tag{4.1}$$

where

Y_{ij} = the observed value of the characteristic for group j in period i,
μ = the overall mean,
α_i = the effect of period i.
β_j = the effect of group j,
and
ε_{ij} is an independent random residual with mean zero and variance σ^2.

It will be noted that the model is written in terms of the group characteristic rather than in terms of the characteristics of the individual units within each group, because in block designs it is unusual to have more than one response from each treatment (period) in a block (group). In sampling where there is usually more than one unit in each period/group combination, the above model could be applied to the mean value of the characteristic in each group in each period. This simple model is adequate for our purpose of obtaining efficient designs for comparing periods. This point and others are taken up in Section 7. We assume that the main objective of a given survey is to obtain information on contrasts, i.e., comparisons, between the period effects. We are interested in estimating $\sum_{i=1}^{t} c_i \alpha_i$, where $\sum_{i=1}^{t} c_i = 0$. Such contrasts include all period-to-period changes, but exclude the mean response in each period. The above formulation is congruous with Wolter (1979) and also with Gurney and Daly (1965).

To obtain the classification of our designs we consider comparisons of neighbouring periods. The "1st neighbour" comparisons are $\alpha_1-\alpha_2$, $\alpha_2-\alpha_3$, …, $\alpha_{t-1}-\alpha_t$, $\alpha_t-\alpha_1$, where here and in the following, the comparisons are taken cyclically. The "2nd neighbour" comparisons are $\alpha_1-\alpha_3$, $\alpha_2-\alpha_4$, …, $\alpha_{t-2}-\alpha_t$, $\alpha_{t-1}-\alpha_1$, $\alpha_t-\alpha_2$. In a similar manner, 3 rd, 4 th, …, neighbour comparisons can be defined for a suitable t. The estimates $\hat{\alpha}_i$ are taken to be the usual least squares estimates, adjusted for group effects.

The (model) variance of these comparisons is of interest and we define

$$v_n \sigma^2 = V(\hat{\alpha}_i - \hat{\alpha}_j), \tag{4.2}$$

where $\hat{\alpha}_i$ and $\hat{\alpha}_j$ are the estimated effects of two n th neighbours, $n=1, 2, \ldots, m_t$ and

$$m_t = \frac{(t-1)}{2} \text{ if } t \text{ is odd and}$$

$$m_t = \frac{t}{2} \text{ if } t \text{ is even.} \tag{4.3}$$

Some of these variances may be identical and we denote the number of distinct ones by d.

A useful class of designs has
$v_1 \leqslant v_2 \leqslant \cdots \leqslant v_{m_t}$.
This class is introduced initially, not only for its practical interest, but for its conceptual import and its simplicity. The properties of rotation designs are more clearly understood for these monotonic designs. Furthermore, these basic designs are building blocks for more complex designs, as will be described later.

These designs give more importance to comparisons between neighbours that are closer together. That is, assign the smallest variance to estimated changes between 1st neighbours, next smallest to changes between 2nd neighbours and so on. In particular, this means that period-to-period, e.g., month-to-month, changes will be estimated precisely, a desirable characteristic in practice.

Let
$v_1 < v_2 < \cdots < v_{d-1} < v_d = v_{d+1} = \cdots = v_{m_t}$.
We denote this class of designs by B_d^+. Note that there is a dual class B_d^-, say, with just the opposite property, namely,

$$v_1 > v_2 > \cdots > v_{d-1} > v_d = v_{d+1} = \cdots = v_{m_t}.$$

The latter class may be of interest, for example, if $m_t = \frac{t}{2}$ and mid-cycle changes are the main focus of attention. For instance, for annual cycles ($t=12$) semestral changes might need to be estimated precisely. It will be seen in the next sections, moreover, that general rotation designs can be constructed by "mixing" (adjoining/combining) basic designs in these simple, monotonic, classes.

Balanced designs form the class B_1^+ with one single value for v_n. It is worthwhile to remark that for a symmetric balanced incomplete block design, l corresponds to the usual design parameter λ, where λ is the number of times a pair of treatments occur together in the same block. Designs currently used in sample surveys can often be considered replicates of symmetric balanced incomplete block designs and can also be represented as cyclic designs. The rotation sampling designs, however, are cyclic designs which reduce respondent burden. It may be of interest to note that the properties of a cyclic incomplete block design can be considered in terms of a corresponding paired comparison design where each block of size k is thought of as being divided into $\frac{1}{2}k(k-1)$ blocks of size two. We do not pursue this correspondence here since not all pairwise comparisons are of interest. (See John (1966) for further details.)

Table 2 lists some of the more efficient rotation designs of the B_d^+ type, constructed using the method described earlier. In the table some designs have a serial number which is followed by a C. These designs are complements of their partners with the same serial number. A complement of a design with incidence matrix N is a design with incidence matrix $N_c=J-N$, where J is a $t \times b$ matrix of ones. Note that $d_c=d$, $r_c=b-r$, and $k_c=t-k$ relate the parameters of a design to those of its complement.

Table 2. Classification of Rotation Designs: $b=3$ to 12, $2 \leq r \leq b-1$

Serial No.	b	t	r	k^*	l	d^{**}
1	3	3	2	2	1	1*
2	4	4	2	2	1	2*R
3	4	4	3	3	2	1*
4	5	5	2	2	1	2*
4C	5	5	3	3	2	2*
5	5	5	4	4	3	1*
6	6	6	2	2	1	3*
6C	6	6	4	4	3	3*
7	6	6	3	3	2	3*R
8	6	3	4	2	2	1
9	6	6	5	5	4	1*
10	6	3	5	(2,3)	4	1
11	7	7	2	2	1	3*
11C	7	7	5	5	4	3*
12	7	7	3	3	2	3*
12C	7	7	4	4	3	3*
13	7	7	6	6	5	1*
14	8	8	2	2	1	4*
14C	8	8	6	6	5	4*
15	8	8	3	3	2	4*
15C	8	8	5	5	4	4*
16	8	8	4	4	3	4*R

(cont).

Table 2. (Cont.) Classification of Rotation Designs: b=3 to 12, 2 ⩽ r ⩾ b–1

Serial No.	*b*	*t*	*r*	*k**	*l*	*d***
17	8	4	4	2	2	2R
18	8	4	5	(2,3)	3	2
19	8	4	6	3	4	1
20	8	8	7	7	6	1*
21	8	4	7	(3,4)	6	1
22	9	9	2	2	1	4*
22C	9	9	7	7	6	4*
23	9	9	3	3	2	4*
23C	9	9	6	6	5	4*
24	9	9	4	4	3	4*
24C	9	9	5	5	4	4*
25	9	3	6	2	3	1
26	9	3	7	(2,3)	5	1
27	9	9	8	8	7	1*
28	9	3	8	(2,3)	7	1
29	10	10	2	2	1	5*
29C	10	10	8	8	7	5*
30	10	10	3	3	2	5*
30C	10	10	7	7	6	5*
31	10	10	4	4	3	5*
31C	10	10	6	6	5	5*
32	10	5	4	2	2	2
32C	10	5	6	3	4	2
33	10	10	5	5	4	5*R
34	10	5	5	(2,3)	3	2R
35	10	5	7	(3,4)	5	2
36	10	5	8	4	6	1
37	10	10	9	9	8	1*
38	10	5	9	(4,5)	8	1
39	11	11	2	2	1	5*
39C	11	11	9	9	8	5*
40	11	11	3	3	2	5*
40C	11	11	8	8	7	5*
41	11	11	4	4	3	5*
41C	11	11	7	7	6	5*
42	11	11	5	5	4	5*
42C	11	11	6	6	5	5*
43	11	11	10	10	9	1*
44	12	12	2	2	1	6*
44C	12	12	10	10	9	6*
45	12	12	3	3	2	6*
45C	12	12	9	9	8	6*

(cont).

Table 2. (Cont.)

Serial No.	b	t	r	k*	l	d**
46	12	12	4	4	3	6*
46C	12	12	8	8	7	6*
47	12	6	4	2	2	3
47C	12	6	8	4	6	3
48	12	12	5	5	4	6*
48C	12	12	7	7	6	6*
49	12	6	5	(2,3)	3	3
49C	12	6	7	(3,4)	5	3
50	12	12	6	6	5	6*R
51	12	6	6	3	4	3R
52	12	4	6	2	3	2R
53	12	4	7	(2,3)	4	2
54	12	4	8	(2,3)	5	2
55	12	3	8	2	4	1
56	12	6	9	(4,5)	7	3
57	12	4	9	3	6	1
58	12	3	9	(2,3)	6	1
59	12	6	10	5	8	1
60	12	4	10	(3,4)	8	1
61	12	3	10	(2,3)	8	1
62	12	12	11	11	10	1*
63	12	6	11	(5,6)	10	1
64	12	4	11	(3,4)	10	1
65	12	3	11	(2,3)	10	1

* Values of k_j, $j=1, 2, \ldots, b$. Either $k_j=k$, all j, or $k_j=k$ or k', $k'=k+1$.
** Value of d in B_d^+.

Also, in Table 2, some designs have values of d which are written as d^*. These designs are symmetric and are such that $b=t$, $r=k$ and N can be written as a symmetric matrix, after possibly permuting its columns. Some designs are reflexive in the sense that taking their complements does not alter the designs themselves. These designs are indicated by an R in Table 2. We further denote by P_t the symmetric balanced design with $r=k=t-1$. (Note that all symmetric designs with a given $b=t$ have $d=m_t$ as in (4.3).)

Designs which have a block size of 1, or are complements which have $r=1$, have been omitted from Table 2. Such designs are not considered useful for the estimation of contrasts, i.e., changes. Moreover, for each of the designs in Table 2 the variances v_n were obtained and are considered in Section 6.

5. Properties of the Designs

Other rotation designs can be obtained by adjoining two or more basic rotation designs, e.g., those listed in Table 2. A design D_1 is adjoined to another D_2 by taking the b_1 columns of D_1 and putting them alongside the b_2 columns of D_2. Of course, both D_1 and D_2 must have the same value of t.

Some of the designs constructed in this way had an incidence matrix which was obtained by adjoining a $t\times t$ identity matrix I_t, to an incidence matrix which had two or more 1's in each column. For example, a design for $b=8$, $t=4$, $r=3$, $k_j=2$ or 1, $j=1, 2, \ldots, b$, is obtained by adjoining I_4 to the incidence matrix for the second design in Table 2, and suitably permuting the columns. Designs with an identity matrix component have been excluded from Table 2.

Table 3 gives a list of the designs in Table 2, which are, in fact, obtained by adjoining other designs in Table 2. Clearly, other designs not in Table 2 could be constructed by adjoining designs that are in Table 2. This adjoining is, of course, subject to the final design being a rotation design.

Table 3. Designs Obtained by Adjoining Two or Three Other Designs

Serial No. in Table 2	Serial Nos. of adjoined designs
17	2,2
18	2,3
52	2,2,2
53	2,2,3
54	2,3,3
32	4,4
34	4,4C
32C	4C,4C
35	4C,5
47	6,6
49	6,7
49C	7,6C
47C	6C,6C
56	6C,9
51	7,7

Clearly, a design obtained by adjoining two balanced (B_1^+) designs is also balanced. Further, adjoining a B_2^+ and a B_1^+ design leads to another B_2^+ design. The effect of adjoining one or more B_1^+ designs to a B_2^+ design is to make the variances, v_1 and v_2, more similar in size (see Section 6).

A different operation consists of combining (vertically) designs with a common b, r, and l to obtain larger rotation designs. In fact, we note that each design of type P_t, for non-prime $t=h_1h_2$, may be formed by combining its "component" design with $t=h_1$ (h_2 times), and similarly for $t_2=h_2$ (h_1 times). Many of these components have blocks of size 1 and are not given in Table 2. Table 4 presents the components of designs P_t, for $4 \leq t \leq 10$.

Table 4. Subclass P_t of Symmetric Balanced Designs, $4 \leq t \leq 10$, and their components[1]

	b	t	r	k	l	d
$P_4 = D_1 \oplus D_1$:	4	4	3	3	2	1
Component: D_1	4	2	3	(1,2)	2	1
$P_6 = D_2 \oplus D_2$						
$= D_3 \oplus D_3 \oplus D_3$:	6	6	5	5	4	1
Components: D_2	6	3	5	(2,3)	4	1
D_3	6	2	5	(1,2)	4	1
$P_8 = D_4 \oplus D_4$						
$= D_5 \oplus D_5 \oplus D_5 \oplus D_5$:	8	8	7	7	6	1
Components: D_4	8	4	7	(3,4)	6	1
D_5	8	2	7	(1,2)	6	1
$P_9 = D_6 \oplus D_6 \oplus D_6$	9	9	8	8	7	1
Component: D_6	9	3	8	(2,3)	7	1
$P_{10} = D_7 \oplus D_7$						
$= D_8 \oplus D_8 \oplus D_8 \oplus D_8 \oplus D_8$:	10	10	9	9	8	1
Components: D_7	10	5	9	(4,5)	8	1
D_8	10	2	9	(1,2)	8	1

[1] The operation of combining (vertically) two designs, D_1 and D_2, say, is denoted by $D_1 \oplus D_2$.

Of course, in any real continuing survey t is not fixed and the designs would evolve over time by adjoining and combining designs. The effects of adjoining and combining designs, as described above, would then be of particular interest. For example, a rotation pattern similar to that used in the U.S. National Crime Survey (NCS) can be generated by successive adjoining and combining operations. By combining (6×12) matrix C = (A|B) to (O|A) one derives the (NCS) pattern displayed in Table 5. Note that the (6×12) matrix C = (A|B) is of the elementary type described in Sections 2 and 3. The pattern employed in the U.S. Current Population Survey can be generated by more complex mixtures of the elementary rotation designs considered in Table 2.

Table 5. NCS Rotation Pattern

	Rotation groups (Panels)											
	(A)						(B)					
t												
1	1	1	1	1	1	1	0	0	0	0	0	0
2	0	1	1	1	1	1	1	0	0	0	0	0
3	0	0	1	1	1	1	1	1	0	0	0	0
4	0	0	0	1	1	1	1	1	1	0	0	0
5	0	0	0	0	1	1	1	1	1	1	0	0
6	0	0	0	0	0	1	1	1	1	1	1	0
⋮	(0) 6×6						(A)					

6. Variances of Estimated Period Changes

The values of v_n, where $v_n\sigma^2 = V(\hat{\alpha}_i - \hat{\alpha}_j)$ are given in Table 6 for d=1, 2 and 3 only, to save space, the remaining ones being available from the authors. The corresponding efficiency factors for each variance are also given. The efficiency factor E_n, say, is defined as

$$E_n = \frac{V(\hat{\alpha}_i - \hat{\alpha}_j) \text{ in a saturated design}}{V(\hat{\alpha}_i - \hat{\alpha}_j) \text{ in the design under consideration}}$$

$$= \frac{2}{r\,v_n}.$$

Here, a saturated design is one in which every group is interviewed in every period. This definition is the one used in the design of comparative experiments. The saturated design of such experiments is the randomized (complete) block design, with r blocks, each of size t. Costs usually prohibit the use of a saturated design in a survey. Nevertheless the saturated design does provide a useful yardstick for measuring the efficiency of other designs. It would be possible to compare the average value of v_n with $\frac{2}{r}$, but this would conceal the true range of the efficiencies in a given design.

For comparison purposes, the total size $n = r \times t$ of the designs are also included in Table 6.

Table 6. Values of the Variances (v_n) and Efficiencies for Designs with d = 1, 2 and 3

Serial No.	b	t	r	n^1	Variances (v_n)			Efficiencies ($\frac{2}{rv_n}\times 100\%$)		
1	3	3	2	6	1.3333			75		
2	4	4	2	8	1.5000	2.0000		67	50	
3	4	4	3	12	0.7500			89		
4	5	5	2	10	1.6000	2.4000		63	42	
4C	5	5	3	15	0.7636	0.8727		87	76	
5	5	5	4	20	0.5333			94		
6	6	6	2	12	1.6667	2.6667	3.0000	60	37	33
6C	6	6	4	24	0.5359	0.5744	0.5769	93	87	87
7	6	6	3	18	0.7833	0.9333	1.0500	85	71	63
8	6	3	4	12	0.6667			75		
9	6	6	5	30	0.4167			96		
10	6	3	5	15	0.4444			90		
11	7	7	2	14	1.7143	2.8571	3.4286	58	35	29
11C	7	7	5	35	0.4174	0.4356	0.4364	96	92	92
12	7	7	3	21	0.7944	0.9826	1.1498	84	68	58
12C	7	7	4	28	0.5396	0.5881	0.6336	93	85	79
13	7	7	6	42	0.3429			97		
17	8	4	4	16	0.7500	1.0000		67	50	
18	8	4	5	20	0.4870	0.5455		82	73	
19	8	4	6	24	0.3750			89		
20	8	8	7	56	0.2917			98		
21	8	4	7	28	0.3000			95		
25	9	3	6	18	0.4444			75		
26	9	3	7	21	0.3333			86		
27	9	9	8	72	0.2540			98		
28	9	3	8	24	0.2667			94		
32	10	5	4	20	0.8000	1.2000		63	42	
32C	10	5	6	30	0.3818	0.4364		87	76	
34	10	5	5	25	0.5053	0.6316		79	63	
35	10	5	7	35	0.3126	0.3297		91	87	
36	10	5	8	40	0.2667			94		
37	10	10	9	90	0.2250			99		
38	10	5	9	45	0.2286			97		
43	11	11	10	110	0.2020			99		
47	12	6	4	24	0.8333	1.3333	1.5000	60	37	33
47C	12	6	8	48	0.2679	0.2872	0.2885	93	87	87

(cont.)

Table 6. *(Cont).*

Serial No.	b	t	r	n[1]	Variances (v_n)			Efficiencies ($\frac{2}{rv_n}\times 100\%$)		
49	12	6	5	30	0.5189	0.6838	0.7582	77	58	52
49C	12	6	7	42	0.3159	0.3515	0.3712	90	81	77
51	12	6	6	36	0.3917	0.4667	0.5250	85	71	63
52	12	4	6	24	0.5000	0.6667		67	50	
53	12	4	7	28	0.3643	0.4286		78	67	
54	12	4	8	32	0.2943	0.3158		85	79	
55	12	3	8	24	0.3333			75		
56	12	6	9	54	0.2341	0.2412	0.2414	95	92	92
57	12	4	9	36	0.2500			89		
58	12	3	9	27	0.2667			83		
59	12	6	10	60	0.2083			96		
60	12	4	10	40	0.2143			93		
61	12	3	10	30	0.2222			90		
62	12	12	11	122	0.1833			99		
63	12	6	11	66	0.1852			98		
64	12	4	11	44	0.1875			97		
65	12	3	11	33	0.1905			95		

[1] Total sample size = $n = rt$.

Adjoining x copies of a design reduces the value of v_n in that design by a factor x, but does not change E_n. For example, the design with serial number 8 can be constructed by adjoining two copies of design 1. Adjoining copies of a design of type B_2^+ to a design of type B_1^+, reduces the relative differences between the values of v_n. For example, adjoining designs 4C and 5 produces design 35. In design 4 C the ratio of the two variances is 1:1.5 but in design 35 is 1:1.05.

Since taking the complement of a design gives $r_c=b-r$, without changing b, the complement of a highly efficient design will have low efficiency. By observing the values of the efficiencies in Table 6, it is clear that, while most designs have high efficiency, there are some with particularly low efficiencies. The low efficiencies occur, as one might expect, for designs where the k_j values equal 2. In other words, all groups must be in the sample more than twice to achieve a reasonable efficiency.

The efficiencies allow a useful design to be chosen out of a number of competitors. For example, if $t=6$ and $b=12$ then design 49 is much more efficient than design 47 and only uses 6 more groups. Of course, design 47C is even more efficient than design 49 but requires 18 more groups.

7. Discussion

Some alternatives to the simple model (4.1) may be considered. First, it may be more realistic to consider the effects of individual groups being random rather than fixed. This would complicate the theory in Sections 4–6 but is not likely to substantially change the overall results.

A second alternative would be to model, individually, each unit within a group. The model for the measurement on the q *th* unit in the j th group in the i th period would then be

$$Y_{ijq} = \mu+\alpha_i+\beta_j+\gamma_{ij}+\varepsilon_{ijq}, \qquad (7.1)$$

where μ, α_i, β_j are defined as in Secton 4; γ_{ij} is the period-by-group interaction effect; and ε_{ijq} is the usual error term. This type of interaction is manifested in practice via what is called the rotation group bias (see, e.g, Bailar (1975)). This bias is related to the number of times different units have been interviewed previously. The presence of an interaction term will contaminate published period-to-period changes, i.e., estimates of the form $\hat{Y}_i - \hat{Y}_{i'}$, which are of more direct interest than contrast estimates $\hat{\alpha}_i - \hat{\alpha}_{i'}$. Clearly, $\hat{Y}_i - \hat{Y}_i'$ differences will involve the interaction parameters γ_{ij}.

Another model-based approach employs time series models to estimate the current mean response. References for this approach are Blight and Scott (1973), Scott and Smith (1974), and Jones (1980).

We have discussed in Section 4 the role of the monotonic classes B_d^+ and B_d^-; in particular, Table 2 is concerned only with B_d^+. We have also examined, however, other diverse variance patterns that are not reported here. Of particular interest is an investigation of the duals or "mirror images" in B_d^- of designs in the class B_d^+. A summary for t=5, d=2 and the designs in Table 2 is presented in Table 7.

Table 7. Mirror Images of Designs for t=5 and = d = 2

Serial No.	Image	*b*	*t*	*r*	*k*	*l*
4	4I	5	5	2	2	0
4C	4CI	5	5	3	3	1
32	32I	10	5	4	2	0
32C	32CI	10	5	6	3	2
34	34I	10	5	5	(2,3)	1
35	35I	10	5	7	(3,4)	4

It is worthwhile to remark that the approach in Sections 4–6 has completely ignored the randomization imposed by the design. In particular, all variances are considered with respect to the model distribution. Denoting such variances by V_m and expectations over the sampling design, p, say, by E_p, one may also consider the criterion E_pV_m. In this context, Bellhouse (1984) has derived optimal treatment assignments for certain subclasses of treatment contrasts. Other references where complex sample design effects have been considered include Nathan and Holt (1980) and Holt, et al. (1980).

8. References

Bailar, B.A. (1975): The Effects of Rotation Group Bias on Estimates from Panel Surveys. Journal of the American Statistical Association, 70, pp. 23–30.

Bellhouse, D.R. (1984): Optimal Randomization for Experiments in Which Autocorrelation is Present. Biometrika, 71, pp. 155–160.

Blight, B.J.N. and Scott, A.J. (1973): A Stochastic Model for Repeated Surveys. Journal of the Royal Statistical Society, B, 35, pp. 61–66.

Chakrabarti, M.C. (1963): On the Use of

Incidence Matrices of Designs in Sampling from Finite Populations. Journal of the Indian Statistical Association, 1, pp. 78–85.

Cochran, W.G. (1977): Sampling Techniques. Third Edition, Wiley, New York.

Eckler, A.R. (1955): Rotation Sampling. Annals of Mathematical Statistics, 26, pp. 664–685.

Fienberg, S.E. (1980): The Measurement of Crime Victimization: Prospectus for Panel Analysis of a Panel Survey. The Statistician, 29, pp. 313–350.

Fienberg, S.E. and Tanur, J.M. (1985): A Long and Honorable Tradition: Intertwining Concepts and Constructs in Experimental Design and Sample Surveys. Proceedings of the International Statistical Institute Meetings in Amsterdam, 45th Session, Vol. L1.

Fuller, W.A. and Harter, R. (1985): The Multivariate Components of Variance Model for Small Area Estimation. Paper presented at the Small Area Symposium, Ottawa, Canada.

Ghangurde, P.D. (1982): Rotation Group Bias in the LFS Estimates. Paper presented at the meeting of the American Statistical Association in Cincinnati.

Ghangurde, P.D. and Rao, J.N.K. (1969): Some Results on Sampling over Two Occasions. Sankhya, 31, pp. 463–472.

Gurney, M. and Daly, J.F. (1965): A Multivariate Approach to Estimation in Periodic Sample Surveys. American Statistical Association, Proceedings of the Social Statistics Section, pp. 242–257.

Hartley, H.O. and Rao, J.N.K. (1978): Estimation of Nonsampling Variance Components in Sample Surveys. In N.K. Namboodiri (Ed.), Survey Sampling and Measurement. Academic Press, New York.

Holt, D., Smith, T.M.F., and Winter, P.D. (1980): Regression Analysis of Data from Complex Surveys. Journal of the Royal Statistical Society, A, 143, pp. 474–487.

Jain, A.K. (1981): Design of a Rotation Scheme for a Stratified Multi-Stage Sample. Journal of Statistical Planning and Inference, 5, pp. 57–69.

John, J.A. (1966): Cyclic Incomplete Block Designs. Journal of the Royal Statistical Society, B, 28, pp. 345–360.

John, J.A., Wolock, F.W., and David, H.A. (1972): Cyclic Designs. National Bureau of Standards, Applied Mathematics Series, 62.

Jones, R.G. (1980): Best Linear Unbiased Estimators for Repeated Surveys. Journal of the Royal Statistical Society, B, 42, pp. 221–226.

Nathan, G. and Holt, D. (1980): The Effect of Survey Design on Regression Analysis. Journal of the Royal Statistical Society, B, 42, pp. 377–386.

Paterson, H.D. (1950): Sampling on Successive Occasions with Partial Replacements of Units. Journal of the Royal Statistical Society, B, 12, pp. 241–255.

Scott, A.J.L and Smith, T.M.F. (1974): Analysis of Repeated Surveys Using Time Series Methods. Journal of the American Statistical Association, 69, pp. 674–678.

U.S. Bureau of the Census (1978): The Current Population Survey: Design and Methodology. Technical paper 40, U.S. Government Printing Office, Washington, D.C.

Wolter, K.M. (1979): Composite Estimation in Finite Populations. Journal of the American Statistical Association, 74, pp. 604–613.

Received December 1985
Revised April 1987