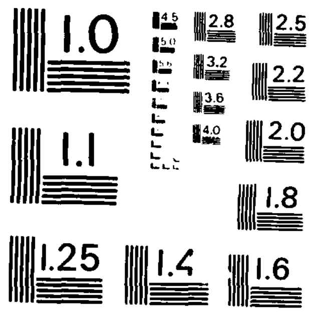
AD-A131 970 ON THE PERFORMANCE OF TWO-STAGE GROUP SCREENING EXPERIMENTS(U) DESMATICS INC STATE COLLEGE PA C A MAURO AUG 83 TR-113-9 N00014-79-C-0650 1/1

UNCLASSIFIED F/G 12/1 NL

END
DATE
FILMED
9-83
DTIC




MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

# ON THE PERFORMANCE OF TWO-STAGE GROUP SCREENING EXPERIMENTS

by

Carl A. Mauro

DESMATICS, INC.

P. O. Box 618
State College, PA. 16801
Phone: (814) 238-9621

*Applied Research in Statistics - Mathematics - Operations Research*

# ON THE PERFORMANCE OF TWO-STAGE GROUP SCREENING EXPERIMENTS

by

Carl A. Mauro

TECHNICAL REPORT NO. 113-9

August 1983

This study was supported by the Office of Naval Research
under Contract No. N00014-79-C-0650, Task No. NR-042-467

TABLE OF CONTENTS

# I. INTRODUCTION

Investigators in many fields are often confronted with research problems in which a large number of factors (i.e., independent variables) must be considered. In such cases the first step in experimentation is usually the identification of the most important factors, so that future research may be concentrated on the major factors. Accordingly we often want to conduct an efficient preliminary screening experiment aimed at determining the subset of important factors.

One resource-efficient screening strategy is two-stage group screening. In this method, introduced by Watson (1961), the individual factors (each at two levels) are partitioned into groups, forming group factors. By assigning the same level to all componenet factors within each group, the group factors are tested as if they were single factors. All factors within groups found to have significant effects are then tested individually in a second-stage experiment.

A key assumption in Watson's development of group screening is that the directions of all effects are known or can be correctly assumed, a priori. With this assumption, factor levels can be assigned so that all effects are in the same direction. Thus, there is no chance of cancellation of effects (within a group). This assumption, however, is unlikely to hold exactly in practice. Consequently one may hesitate to use a group screening design because important effects may cancel if assumed effect directions are wrong.

In a previous paper, we [Mauro and Smith (1982)] examined the extent to which cancellation affects the performance of two-stage group screening designs when the response is observed without random error (i.e., when the error standard deviation $\sigma$ is equal to zero). In the present paper we extend this work to the case $\sigma > 0$. As part of our investigation we have developed a computer-aided search routine to select an optimal (in a sense to be defined later) group screening plan. As in our earlier paper, we use the multifactorial designs of Plackett and Burman (1946) to analyze the results of the first and second stages.

## II. ASSUMPTIONS AND NOTATION

Suppose K factors are to be screened for their effects on the response. For detecting the factors having major effects, it is usually sufficient to assume a first-order model:

$$y_i = \beta_0 + \sum_{j=1}^{K} \beta_j x_{ij} + \varepsilon_i \,, \tag{2.1}$$

where $y_i$ is the $i^{\underline{th}}$ response, $\beta_0$ is a constant term common to every response, $\beta_j$ ($j \geq 1$) is the linear effect of the $j^{\underline{th}}$ factor, $x_{ij} = \pm 1$ is the level of the $j^{\underline{th}}$ factor in the $i^{\underline{th}}$ run, and $\varepsilon_i$ is the $i^{\underline{th}}$ error term. We make the following additional assumptions:

1. $k \geq 1$ ($k$ unknown) of the K factors are active (i.e., have a true effect) and $(K-k)$ are inactive,
2. all active factors have the same absolute effect, $\Delta > 0$, that is,

$$|\beta_j| = \begin{cases} \Delta, & \text{if factor } j \text{ is active} \\ 0, & \text{if factor } j \text{ is inactive,} \end{cases}$$

3. the error terms $\{\varepsilon_i\}$ are independent and normally distributed with mean zero and unknown variance $\sigma^2$.

We let $\underline{\beta}(i)$ for $i=0,1,2,\ldots,k$ denote the effects arrangement in which $i$ effects equal $-\Delta$, $(k-i)$ effects equal $+\Delta$, and $(K-k)$ effects equal 0. The $\underline{\beta}(0)$ [or $\underline{\beta}(k)$] case, therefore, corresponds to the situation where all active effects are in the same direction. Furthermore, in the version of two-stage group screening we consider, we assume that the K factors are partitioned randomly into G groups of size g; if K is not a multiple of g, we assume that the group sizes are taken as "evenly" as possible. Concerning our assumptions, we note that random grouping and equal absolute effects maximize the chance of cancellation. Thus, with regard to studying the cancellation effect, our assumptions define "worst case" conditions.

For reasons of economy and to avoid design saturation (i.e., no degrees of freedom to estimate $\sigma$), we employ at both stages of screening the smallest Plackett-Burman (PB) design that has at least one error degree of freedom. Since PB designs are only available for numbers of runs that are multiples of four, the number of first-stage runs $N_1$ required to test the G group factors will therefore be B(G+1) where

$$B(x) = x + 4 - x(\text{mod } 4). \qquad (2.2)$$

Similarly, if S denotes the number of factors that reach the second stage, then the number of second-stage runs $N_2$ will be B(S+1). Thus, the total number of runs R required by both stages of group screening will be $N_1 + N_2 = B(G+1) + B(S+1)$. We note that because S is random, so is R.

Regarding formal significance testing, the results of the first and second stages can be analyzed by the usual analysis of variance procedures for factorial experiments. We denote the significance levels of the (two-sided) t tests performed at the end of the first and second stages by $\alpha_1$ and $\alpha_2$, respectively. Our version of two-stage group screening, therefore, is completely determined by g, $\alpha_1$, and $\alpha_2$. Accordingly, we will denote such a strategy by GS(g, $\alpha_1$, $\alpha_2$). In the next section we show how the quantities g, $\alpha_1$, and $\alpha_2$ affect the performance of the GS(g, $\alpha_1$, $\alpha_2$) strategy.

## III. PERFORMANCE EVALUATION

We can define three separate measures of performance. These are:

Power. We denote by A the number of active factors that are detected correctly, and we define

$$E_A = 100E(A)/k \qquad (3.1)$$

as a percentage measure of the "power" of a GS strategy for detecting the active factors.

Type I Error. We denote by U the number of inactive factors that are declared active (important), and we define

$$E_U = 100E(U)/(K-k) \qquad (3.2)$$

as a percentage measure of Type I error (i.e., declaring active an inactive factor).

Relative Testing Cost. We define relative testing cost

$$E_R = 100E(R)/B(K+1) \qquad (3.3)$$

as the ratio, expressed as a percentage, of the expected number of runs required by a GS strategy to the number of runs required by the smallest PB design for K factors that has at least one error degree of freedom.

A larger value of $E_A$, a smaller value of $E_U$, or a smaller value of $E_R$ indicates better performance on the average, but all three measures should be considered in assessing and comparing the performance of GS(g, $\alpha_1$, $\alpha_2$) strategies. In general, selecting a suitable GS(g, $\alpha_1$, $\alpha_2$) strategy will require that trade-offs be made between $E_A$, $E_U$, and $E_R$. In many ways our problem is like the testing of a statistical hypothesis in which we want the sample size (relative testing cost) and Type I error to be small but the power to be large.

In the Appendix we derive the expected values of A, U, and R for any K, k, $\underline{\beta}(i)$, g, $\alpha_1$, $\alpha_2$, and signal-to-noise ratio $\Delta/\sigma$ . Using these results we developed a computer program that gives the performance properties of alternative GS(g, $\alpha_1$, $\alpha_2$) strategies. To illustrate its use, we applied our program to two case studies, which we will refer to as Study A and Study B. In Study A we evaluated $E_A$, $E_U$, and $E_R$ when K=60, k=8, $\Delta/\sigma$= 1,3,$\infty$ and $\underline{\beta} = \underline{\beta}(0)$, $\underline{\beta}(4)$ for g=3 and 6, $\alpha_1$=.01, .05, .10 and $\alpha_2$=.01, .05, .10 . In Study B we evaluated $E_A$, $E_U$, and $E_R$ when K=240, k=32, $\Delta/\sigma$ = 1, 3, $\infty$ and $\underline{\beta} = \underline{\beta}(0)$,$\underline{\beta}(16)$ for the same g, $\alpha_1$, $\alpha_2$ combinations considered in Study A. We chose to consider just the $\underline{\beta}(0)$ and $\underline{\beta}([k/2])$ cases because of the symmetry between the $\underline{\beta}(i)$ and $\underline{\beta}(k-i)$ arrangements. Moreover, the probability of group-factor effect cancellation is maximized in the $\underline{\beta}([k/2])$ case. The results obtained for Studies A and B are shown in Tables 1A and 1B, respectively.

In Tables 1A and 1B we give the limiting values of $E_R$, $E_A$, and $E_U$ as $\Delta/\sigma \to \infty$ ($\sigma$>0). We note that these results are not directly comparable to those given by Mauro and Smith (1982) under the assumption that $\sigma$=0. The zero and nonzero error cases are fundamentally different since when $\sigma$=0 the testing process is totally deterministic.

It is intuitive that an increase in $\alpha_1$ or $\alpha_2$ will increase both $E_A$ and $E_U$. An increase in $\alpha_1$ will also increase the expected number of runs made, thus the relative testing cost $E_R$. However, $E_R$ does not depend on $\alpha_2$, a fact that will be important to us in later discussion. The actual extent of these movements for our two case studies can be seen from Tables 1A and 1B.

$g=3$

$\Delta/\sigma = 1$

| | $E_R$ | | $\alpha_2 = .01$ | $\alpha_2 = .05$ | $\alpha_2 = .10$ |
|---|---|---|---|---|---|
| $\alpha_1 = .01$ | 60.6(58.4) | $E_A$ | 15.13(11.03) | 34.77(26.19) | 44.03(33.65) |
| | | $E_U$ | 0.13( 0.12) | 0.65( 0.61) | 1.30( 1.21) |
| $\alpha_1 = .05$ | 75.9(73.8) | $E_A$ | 33.76(28.28) | 65.77(56.37) | 78.39(67.70) |
| | | $E_U$ | 0.27( 0.26) | 1.33( 1.28) | 2.65( 2.57) |
| $\alpha_1 = .10$ | 81.5(79.4) | $E_A$ | 39.52(34.07) | 73.19(64.22) | 85.86(75.74) |
| | | $E_U$ | 0.32( 0.31) | 1.62( 1.57) | 3.23( 3.15) |

$\Delta/\sigma = 3$

| | $E_R$ | | $\alpha_2 = .01$ | $\alpha_2 = .05$ | $\alpha_2 = .10$ |
|---|---|---|---|---|---|
| $\alpha_1 = .01$ | 76.8(74.5) | $E_A$ | 76.96(67.20) | 93.86(82.22) | 99.35(87.29) |
| | | $E_U$ | 0.26( 0.25) | 1.31( 1.26) | 2.62( 2.53) |
| $\alpha_1 = .05$ | 79.2(77.0) | $E_A$ | 77.49(68.15) | 94.27(83.09) | 99.49(87.92) |
| | | $E_U$ | 0.29( 0.28) | 1.46( 1.42) | 2.92( 2.83) |
| $\alpha_1 = .10$ | 82.2(80.2) | $E_A$ | 77.91(69.17) | 94.68(84.08) | 99.59(88.63) |
| | | $E_U$ | 0.33( 0.32) | 1.65( 1.61) | 3.29( 3.21) |

$\Delta/\sigma \to \infty$

| | $E_R$ | | $\alpha_2 = .01$ | $\alpha_2 = .05$ | $\alpha_2 = .10$ |
|---|---|---|---|---|---|
| $\alpha_1 = .01$ | 76.8(74.5) | $E_A$ | 100.00(87.97) | 100.00(87.97) | 100.00(87.97) |
| | | $E_U$ | 0.26( 0.25) | 1.31( 1.26) | 2.62( 2.53) |
| $\alpha_1 = .05$ | 79.2(77.0) | $E_A$ | 100.00(88.45) | 100.00(88.45) | 100.00(88.45) |
| | | $E_U$ | 0.29( 0.28) | 1.46( 1.42) | 2.92( 2.83) |
| $\alpha_1 = .10$ | 82.2(80.2) | $E_A$ | 100.00(89.06) | 100.00(89.06) | 100.00(89.06) |
| | | $E_U$ | 0.33( 0.32) | 1.65( 1.61) | 3.29( 3.21) |

$g=6$

$\Delta/\sigma = 1$

| | $E_R$ | | $\alpha_2 = .01$ | $\alpha_2 = .05$ | $\alpha_2 = .10$ |
|---|---|---|---|---|---|
| $\alpha_1 = .01$ | 27.9(27.0) | $E_A$ | 0.74( 0.42) | 2.28( 1.34) | 3.42( 2.08) |
| | | $E_U$ | 0.03( 0.03) | 0.17( 0.14) | 0.34( 0.28) |
| $\alpha_1 = .05$ | 41.8(38.0) | $E_A$ | 7.07( 4.21) | 17.43(10.84) | 23.11(14.56) |
| | | $E_U$ | 0.17( 0.14) | 0.83( 0.69) | 1.66( 1.37) |
| $\alpha_1 = .10$ | 56.6(50.5) | $E_A$ | 16.58(10.34) | 35.15(23.22) | 44.70(29.99) |
| | | $E_U$ | 0.31( 0.26) | 1.55( 1.31) | 3.09( 2.61) |

$\Delta/\sigma = 3$

| | $E_R$ | | $\alpha_2 = .01$ | $\alpha_2 = .05$ | $\alpha_2 = .10$ |
|---|---|---|---|---|---|
| $\alpha_1 = .01$ | 34.2(31.7) | $E_A$ | 10.41( 5.93) | 15.50( 9.22) | 18.88(11.59) |
| | | $E_U$ | 0.09( 0.07) | 0.46( 0.36) | 0.92( 0.72) |
| $\alpha_1 = .05$ | 64.1(56.3) | $E_A$ | 44.80(31.11) | 66.86(46.09) | 73.81(51.68) |
| | | $E_U$ | 0.38( 0.32) | 1.88( 1.57) | 3.77( 3.15) |
| $\alpha_1 = .10$ | 79.6(72.0) | $E_A$ | 58.26(45.01) | 87.68(67.53) | 94.29(73.50) |
| | | $E_U$ | 0.54( 0.48) | 2.68( 2.37) | 5.36( 4.75) |

$\Delta/\sigma \to \infty$

| | $E_R$ | | $\alpha_2 = .01$ | $\alpha_2 = .05$ | $\alpha_2 = .10$ |
|---|---|---|---|---|---|
| $\alpha_1 = .01$ | 80.2(72.1) | $E_A$ | 100.00(78.11) | 100.00(78.11) | 100.00(78.11) |
| | | $E_U$ | 0.54( 0.47) | 2.68( 2.35) | 5.36( 4.69) |
| $\alpha_1 = .05$ | 81.8(74.0) | $E_A$ | 100.00(79.00) | 100.00(79.00) | 100.00(79.00) |
| | | $E_U$ | 0.55( 0.49) | 2.77( 2.45) | 5.54( 4.91) |
| $\alpha_1 = .10$ | 83.7(76.3) | $E_A$ | 100.00(80.10) | 100.00(80.10) | 100.00(80.10) |
| | | $E_U$ | 0.58( 0.52) | 2.89( 2.59) | 5.78( 5.18) |

Table 1A. Performance Results For Study A: K=60, k=8. Values in Parentheses Were Obtained in $\underline{\beta}(4)$ Case; Values Outside Parentheses Were Obtained in $\underline{\beta}(0)$ Case.

## g=1

$\Delta/\sigma = 1$

| | $E_R$ | | $\alpha_2 = .01$ | $\alpha_2 = .05$ | $\alpha_2 = .10$ |
|---|---|---|---|---|---|
| $\alpha_1 = .01$ | 68.8(66.5) | $E_A$ | 61.49(52.91) | 82.85(72.18) | 90.79(79.26) |
| | | $E_U$ | 0.24( 0.23) | 1.20( 1.16) | 2.41( 2.32) |
| $\alpha_1 = .05$ | 73.5(71.3) | $E_A$ | 67.16(58.89) | 88.61(78.41) | 96.76(85.75) |
| | | $E_U$ | 0.29( 0.28) | 1.44( 1.40) | 2.88( 2.79) |
| $\alpha_1 = .10$ | 76.7(74.6) | $E_A$ | 68.33(60.48) | 89.16(79.47) | 97.14(86.70) |
| | | $E_U$ | 0.33( 0.32) | 1.63( 1.59) | 3.25( 3.17) |

$\Delta/\sigma = 3$

| | $E_R$ | | $\alpha_2 = .01$ | $\alpha_2 = .05$ | $\alpha_2 = .10$ |
|---|---|---|---|---|---|
| $\alpha_1 = .01$ | 71.0(68.7) | $E_A$ | 83.57(73.66) | 99.30(87.67) | 99.99(88.42) |
| | | $E_U$ | 0.26( 0.25) | 1.29( 1.25) | 2.58( 2.49) |
| $\alpha_1 = .05$ | 73.5(71.3) | $E_A$ | 83.85(74.32) | 99.42(88.27) | 99.99(88.88) |
| | | $E_U$ | 0.29( 0.28) | 1.44( 1.40) | 2.88( 2.79) |
| $\alpha_1 = .10$ | 76.7(74.6) | $E_A$ | 84.18(75.12) | 99.54(88.99) | 99.99(89.47) |
| | | $E_U$ | 0.33( 0.32) | 1.63( 1.59) | 3.25( 3.17) |

$\Delta/\sigma \to \infty$

| | $E_R$ | | $\alpha_2 = .01$ | $\alpha_2 = .05$ | $\alpha_2 = .10$ |
|---|---|---|---|---|---|
| $\alpha_1 = .01$ | 71.0(68.7) | $E_A$ | 100.00(88.42) | 100.00(88.42) | 100.00(88.42) |
| | | $E_U$ | 0.26( 0.25) | 1.29( 1.25) | 2.58( 2.49) |
| $\alpha_1 = .05$ | 73.5(71.3) | $E_A$ | 100.00(88.88) | 100.00(88.88) | 100.00(88.88) |
| | | $E_U$ | 0.29( 0.28) | 1.44( 1.40) | 2.88( 2.79) |
| $\alpha_1 = .10$ | 76.7(74.6) | $E_A$ | 100.00(89.47) | 100.00(89.47) | 100.00(89.47) |
| | | $E_U$ | 0.33( 0.32) | 1.63( 1.59) | 3.25( 3.17) |

## g=6

$\Delta/\sigma = 1$

| | $E_R$ | | $\alpha_2 = .01$ | $\alpha_2 = .05$ | $\alpha_2 = .10$ |
|---|---|---|---|---|---|
| $\alpha_1 = .01$ | 65.5(57.2) | $E_A$ | 47.36(38.31) | 68.40(49.10) | 81.36(58.99) |
| | | $E_U$ | 0.41( 0.35) | 2.04( 1.73) | 4.09( 3.45) |
| $\alpha_1 = .05$ | 78.1(70.6) | $E_A$ | 56.72(45.07) | 82.28(65.12) | 96.32(76.97) |
| | | $E_U$ | 0.54( 0.48) | 2.67( 2.39) | 5.35( 4.77) |
| $\alpha_1 = .10$ | 80.6(73.6) | $E_A$ | 57.34(46.38) | 83.33(67.10) | 97.22(79.00) |
| | | $E_U$ | 0.56( 0.51) | 2.82( 2.55) | 5.64( 5.10) |

$\Delta/\sigma = 3$

| | $E_R$ | | $\alpha_2 = .01$ | $\alpha_2 = .05$ | $\alpha_2 = .10$ |
|---|---|---|---|---|---|
| $\alpha_1 = .01$ | 77.0(69.3) | $E_A$ | 71.41(56.19) | 99.73(79.74) | 99.99(80.12) |
| | | $E_U$ | 0.52( 0.46) | 2.60( 2.31) | 5.21( 4.61) |
| $\alpha_1 = .05$ | 78.6(71.2) | $E_A$ | 71.67(57.02) | 99.76(80.59) | 99.99(80.93) |
| | | $E_U$ | 0.54( 0.48) | 2.78( 2.41) | 5.40( 4.83) |
| $\alpha_1 = .10$ | 80.7(73.7) | $E_A$ | 71.99(58.07) | 99.79(81.64) | 99.99(81.93) |
| | | $E_U$ | 0.56( 0.51) | 2.82( 2.55) | 5.64( 5.10) |

$\Delta/\sigma \to \infty$

| | $E_R$ | | $\alpha_2 = .01$ | $\alpha_2 = .05$ | $\alpha_2 = .10$ |
|---|---|---|---|---|---|
| $\alpha_1 = .01$ | 77.0(69.3) | $E_A$ | 100.00(80.12) | 100.00(80.12) | 100.00(80.12) |
| | | $E_U$ | 0.52( 0.46) | 2.60( 2.31) | 5.21( 4.61) |
| $\alpha_1 = .05$ | 78.6(71.2) | $E_A$ | 100.00(80.93) | 100.00(80.93) | 100.00(80.93) |
| | | $E_U$ | 0.54( 0.48) | 2.78( 2.41) | 5.40( 4.83) |
| $\alpha_1 = .10$ | 80.7(73.7) | $E_A$ | 100.00(81.93) | 100.00(81.93) | 100.00(81.93) |
| | | $E_U$ | 0.56( 0.51) | 2.82( 2.55) | 5.64( 5.10) |

Table 1B. Performance Results For Study B: K=240, k=32. Values in Parentheses Were Obtained in $\underline{\beta}(16)$ Case; Values Outside Parentheses Were Obtained in $\underline{\beta}(0)$ Case.

We can further add that as $\alpha_1 \to 0$, the probability of detecting any group effect approaches zero. Thus, in this case, $E_A \to 0$, $E_U \to 0$, and $E_R \to 100\gamma$, where $\gamma = B(G+1)/B(K+1)$. As $\alpha_1 \to 1$, all groups will be found to have a significant effect with probability one, in which case the second stage simply becomes a PB experiment for all K factors in B(K+1) runs; thus, $E_U \to \alpha_2$, $E_A \to \Psi_{PB}$, and $E_R \to 100(1+\gamma)$, where $\Psi_{PB}$ is the corresponding power for detecting an effect of magnitude $\Delta$ in the associated PB design. Corresponding limits as $\alpha_2 \to 0$ or 1 are less interesting.

As noted in Tables 1A and 1B, performance values enclosed by parentheses were obtained in the $\underline{\beta}([k/2])$ case; those directly preceding the parentheses are corresponding values that were obtained in the $\underline{\beta}(0)$ case. Any differences between these values, therefore, are due to the cancellation effect. Since the chance of canc llation is zero in the $\underline{\beta}(0)$ case and at a maximum in the $\underline{\beta}([k/2])$ case, these differences specify the maximum effect of cancellation on performance.

An examination of the results shows that of the three performance measures considered, $E_A$ is the most sensitive to the cancellation effect. In some of the cases considered $E[A| \underline{\beta}([k/2])]$ is 70% of $E[A| \underline{\beta}(0)]$. Further, although we would tend to use fewer runs as the chance of cancellation increases [equivalently, as i increases from 0 to [k/2] in $\underline{\beta}(i)$], this apparent advantage is offset by the fact that we would also tend to detect fewer active factors.

In Studies A and B the proportion of factors that are active (i.e., k/K) is the same. Mauro and Smith (1982) found that in the deterministic case ($\sigma$=0) power and relative testing cost are essentially a function of k/K and $\underline{\beta}(i/k)$. Inspection of Tables 1A and 1B clearly shows that this result does not hold when $\sigma > 0$. Thus, in this case, it will be necessary to consider each (K,k) combination separately.

An important practical consideration in the use of a GS strategy is the number of error degrees of freedom (e.d.f.) for testing group effects in the first stage. For testing G groups in a PB design having B(G+1) runs, e.d.f. = B(G+1) -(G+1), and thus, $1 \leq$ e.d.f. $\leq 4$. When e.d.f. = 1 the efficiency of the PB design can be extremely poor. See, for example, the case with K=60 and g=6 in Table 1A. In such cases a more reasonable strategy may be to employ a larger PB design or to use one less group and partition the factors as "evenly" as possible. Such considerations would also apply in the second stage where e.d.f. = B(S+1) - (S+1). At this stage, however, a reasonable alternative for increasing e.d.f. may be to combine effects into a pooled error estimate.

We make one final observation. Tables 1A and 1B suggest that $E_U$ is directly proportional to $\alpha_2$ in a GS(g, $\alpha_1$, $\alpha_2$) strategy. Indeed, using (A.2) and (A.5) in the Appendix, we see that

$$E_U = 100*\alpha_2*P\ \{\text{an inactive factor reaches the second stage}\} \quad (3.4)$$

We will make use of this fact in the next section.

## IV. SEARCH ROUTINE

To select a GS strategy, the experimenter must specify the group size g and the significance levels of the first and second stage tests, $\alpha_1$ and $\alpha_2$. In general there are no obvious choices for g, $\alpha_1$, and $\alpha_2$. In order to choose the best strategy for a particular application, trade-offs will need to be made between $E_R$, $E_A$, and $E_U$. In this section we present a computer-aided search routine to help select a "good" GS strategy. The search program is written in standard FORTRAN and is available upon request from the authors.

In order to use the search routine, the experimenter must first specify a maximum tolerable relative testing cost, say $E_R^*$, and a maximum tolerable Type I error, say $E_U^*$. Subject to $E_R \leq E_R^*$ and $E_U \leq E_U^*$, the search algorithm determines, for various group sizes, the values of $\alpha_1$ and $\alpha_2$ that maximize power ($E_A$). From the program output, the group size which gives the greatest power may then be selected. The basic steps in the search algorithm are outlined in Table 2.

Step 4 of the algorithm makes use of the fact the $E_R$ is an increasing function in $\alpha_1$ and does not depend on $\alpha_2$. The $\alpha_2$ defined by Step 5 can be quickly determined from (3.4). Further, the overall logic of the algorithm is based on the premise that power increases as relative testing cost increases.

Table 3 containes sample computer printout from the search routine for the case K=60, k=8, $\Delta/\sigma$=2, $E_R^*$=50%, and $E_U^*$=5%. Asterisks appearing in the printout signify a case (i.e., group size) in which the value of $E_R$ at $\alpha_1$=0 equals or exceeds $E_R^*$. In Table 3, for instance, when g=2

| | |
|---|---|
| Step 1. | Input values for K, k, and $\Delta/\sigma$ . |
| Step 2. | Input maximum tolerable values for $E_R$ and $E_U$. |
| Step 3. | Assume $\underline{\beta}(0)$ case and g=2. |
| Step 4. | Determine $\alpha_1$ so that $E_R$ attains maximum allowable value. |
| Step 5. | For the $\alpha_1$ determined in Step 4, determine the $\alpha_2$ that maximizes $E_U$ subject to constraint specified in Step 2. |
| Step 6. | Calculate $E_A$ and $E_U$ for given $GS(g,\alpha_1,\alpha_2)$ strategy. |
| Step 7. | Repeat Steps 4, 5, and 6 as long as $g \leq \min(8, K/2)$. |
| Step 8. | Reset g=2 and repeat Steps 4 through 7 for $\underline{\beta}(\lceil k/2 \rceil)$ case. |

Table 2. Outline of Two-Stage Group Screening Search Algorithm

SEARCH RESULTS FOR B( 0) CASE:

| STRATEGY | GROUPS | KSTAR | POWER | TYPE I ERROR | RELATIVE TESTING COST |
|---|---|---|---|---|---|
| GS(2,*******,*******) | *** | **** | ****** | ****** | **** |
| GS(3,0.00031,1.00000) | 20 | 60 | 26.21 | 4.641 | 50.0 |
| GS(4,0.00015,0.49499) | 15 | 60 | 40.74 | 5.000 | 50.0 |
| GS(5,0.00110,0.30370) | 12 | 60 | 49.39 | 5.000 | 50.0 |
| GS(6,0.04072,0.22769) | 10 | 60 | 49.59 | 5.000 | 50.0 |
| GS(7,0.00714,0.19609) | 9 | 63 | 55.98 | 5.000 | 50.0 |
| GS(8,0.00016,0.22686) | 7 | 56 | 49.18 | 5.000 | 50.0 |

SEARCH RESULTS FOR B( 4) CASE:

| STRATEGY | GROUPS | KSTAR | POWER | TYPE I ERROR | RELATIVE TESTING COST |
|---|---|---|---|---|---|
| GS(2,*******,*******) | *** | **** | ****** | ****** | **** |
| GS(3,0.00043,0.95345) | 20 | 60 | 22.30 | 5.000 | 50.0 |
| GS(4,0.00027,0.44386) | 15 | 60 | 33.18 | 5.000 | 50.0 |
| GS(5,0.00198,0.27931) | 12 | 60 | 40.01 | 5.000 | 50.0 |
| GS(6,0.05444,0.19763) | 10 | 60 | 41.60 | 5.000 | 50.0 |
| GS(7,0.01312,0.18467) | 9 | 63 | 45.06 | 5.000 | 50.0 |
| GS(8,0.00075,0.20804) | 7 | 56 | 37.21 | 5.000 | 50.0 |

Table 3. Sample Computer Printout From Two-Stage Group Screening Search Routine When K=60, k=8, and $\Delta/\sigma=2$. Maximum $E_U$ Specified As 5% And Maximum $E_R$ Specified As 50% (Equivalently, E(R)=32 Runs). Note: Power, Type I Error, and Relative Testing Cost Are Expressed As Percentages.

the first-stage PB experiment requires 32 runs, thus leaving no runs available for the second-stage follow-up experiment.

The search routine is based on the performance results derived in the Appendix, which are only applicable when K is a multiple of g. Therefore, for group sizes where this restriction is not met, the program redefines K as the nearest multiple of g and denotes the new value as KSTAR. Performance results are then calculated as if there are KSTAR factors to be screened. It seems reasonable that these results should be comparable to the true performance had the K factors been partitioned as "evenly" as possible.

As a final observation, we note that if the group sizes considered in Table 3 are ranked according to their corresponding power, the ranking is the same in the $\underline{\beta}(0)$ and $\underline{\beta}(4)$ cases. We have found this phenomenon to be generally true for the different cases we have looked at. Greater power, of course, is attained in the $\underline{\beta}(0)$ case, when no cancellation can occur.

The search routine presented in this section provides guidance in using and selecting a satisfactory GS strategy. The routine supplies the user with quantitative information needed to determine whether two-stage group screening is suitable for a particular application.

## V. SUMMARY AND REMARKS

In this paper we examine the performance characteristics of two-stage group screening experiments, extending the previous work of Mauro and Smith (1982). The analysis in Section III indicates the extent to which the choice of group size and of the significance levels of the first and second stage tests affect performance. We evaluate performance as a function of a constant signal-to-noise ratio for all active (i.e., nonzero) factors, assuming random grouping and a first-order model.

In screening experiments an experimenter may hesitate to use group screening because important effects may cancel if not all effect directions are known. A key feature in our development is that we make no presumption concerning effect direction, only effect magnitude. In fact, the assumptions of random grouping and equal absolute effects define "worst case" conditions with regard to possible cancellation of effects.

In general, the efficacy of group screening in a given application must be based on trade-offs between factor classification and testing cost. To facilitate this process we have developed a computer-aided search routine which addresses this problem. The results of this paper can be used as a practical guide in decisions about the possible use and choice of a two-stage group screening strategy.

## APPENDIX

### Part 1. Introduction

Without loss of generality, let $\underline{\beta}(i)$ denote the case in which $\beta_1=\beta_2=\ldots=\beta_i= -\Delta$ , $\beta_{i+1}=\beta_{i+2}=\ldots=\beta_k= +\Delta$ , and $\beta_{k+1}=\beta_{k+2}=\ldots=\beta_K=0$. Suppose that K=gG where g denotes the group size and G denotes the number of groups. We define A and U as in (3.1) and (3.2). We now define

$$R_{1j} = \begin{cases} 1, & \text{if the } j^{\underline{th}} \text{ factor is in a group that shows a significant effect in the first stage} \\ 0, & \text{otherwise} \end{cases}$$

$$R_{2j} = \begin{cases} 1, & \text{if the } j^{\underline{th}} \text{ factor shows a significant effect in the second stage} \\ 0, & \text{otherwise.} \end{cases}$$

In a GS(g, $\alpha_1$, $\alpha_2$) strategy, the $j^{\underline{th}}$ factor is declared important only if both $R_{1j}=1$ and $R_{2j}=1$. Accordingly, we define $D_j=R_{1j}R_{2j}$ and observe that $A = \sum_{j\leq k} D_j$ and $U = \sum_{j>k} D_j$. The number of factors S that reach the second stage is given by $S = \sum_{j=1}^{K} R_{1j}$.

Because of symmetry we can write

$$\begin{aligned} E[A|\underline{\beta}(i)] &= iE[D_1|\underline{\beta}(i)] + (k-i)E[D_{i+1}|\underline{\beta}(i)] \\ &= iE[D_1|\underline{\beta}(i)] + (k-i)E[D_1|\underline{\beta}(k-i)]. \qquad (A.1) \end{aligned}$$

Similarly

$$E[U|\underline{\beta}(i)] = (K-k)E[D_{k+1}|\underline{\beta}(i)], \qquad (A.2)$$

and

$$E[S|\underline{\beta}(i)] = iE[R_{11}|\underline{\beta}(i)] + (k-i)E[R_{11}|\underline{\beta}(k-i)]$$
$$+ (K-k)E[R_{1,k+1}|\underline{\beta}(i)]. \qquad (A.3)$$

Further,

$$E[D_1|\underline{\beta}(i)] = P[R_{11}=1|\underline{\beta}(i)]P[R_{21}=1|R_{11}=1,\underline{\beta}(i)], \qquad (A.4)$$

and

$$E[D_{k+1}|\underline{\beta}(i)] = P[R_{1,k+1}=1|\underline{\beta}(i)]\alpha_2. \qquad (A.5)$$

Thus, to evaluate the expectations of A, U, and S it suffices to evaluate $P[R_{1j}=1|\underline{\beta}(i)]$ for j=1 and j=k+1 and evaluate $P[R_{2j}=1|R_{1j}=1, \underline{\beta}(i)]$ for j=1.

Regarding the expected total runs,

$$E(R) = B(G+1) + E[B(S+1)] . \qquad (A.6)$$

Using the approximation $B(x) \cong x+2.5$, we have

$$E[R|\underline{\beta}(i)] \cong B(G+1) + 3.5 + E[S|\underline{\beta}(i)]. \qquad (A.7)$$

Note that $|B(x) - (x+2.5)| \leq 1.5$.

## Part 2. Derivation of the Probability that an Active Factor Reaches Second Stage

Without loss of generality suppose that factor #1 is placed into group #1 and assume that $1\leq i\leq k$. Now, given that factor #1 ($\beta_1 = -\Delta$) is placed into group #1, there are (i-1) effects of $-\Delta$, (k-i) effects of $+\Delta$, and (K-k) zero effects left to be distributed into groups.

We define

$$\Psi(n;\delta;\alpha) = P\{|T_n(\delta)| \geq t(n;\alpha/2)\} \qquad (A.8)$$

where $T_n(\delta)$ denotes a random variable having a noncentral t distribution with n degrees of freedom and noncentrality parameter $\delta$, and where $t(n;\alpha/2)$ denotes the upper $100(1-\alpha/2)$ percentage point of "Student's" t distribution with n degrees of freedom.

It is not difficult to see that

$$P[R_{11}=1|\underline{\beta}(i)] = \sum_{j=0}^{J} \sum_{m=0}^{M} p(j,m)\Psi(f_1;\delta(j,m);\alpha_1), \qquad (A.9)$$

where $J = \min(g-1,\ i-1)$, $M = \min(g-1-j,\ k-i)$, $f_1 = N_1-G-1$, $N_1 = B(G+1)$, $\delta(j,m) = \sqrt{N_1}\ (m-j-1)\Delta/\sigma$, and

$$p(j,m) = \binom{i-1}{j}\binom{k-i}{m}\binom{K-k}{g-1-j-m} \Bigg/ \binom{K-1}{g-1}. \qquad (A.10)$$

The quantity $p(j,m)$ defined in (A.10) is the probability that j effects of $-\Delta$ and m effects of $+\Delta$ fall into group #1 along with factor #1. The $\Psi$ quantity appearing in (A.9) is the power of the t-test associated with group #1 given that $(j+1)$ effects of $-\Delta$ , m effects of $+\Delta$ , and $(g-1-j-m)$ zero effects are placed in group #1.

## Part 3. Derivation of the Probability that an Inactive Factor Reaches Second Stage

To derive $P[R_{1,k+1}=1|\underline{\beta}(i)]$ we can repeat the argument of Part 2 for the (k+1)st factor ($\beta_{k+1}=0$). Doing so, we obtain

$$P[R_{1,k+1}=1|\underline{\beta}(i)] = \sum_{j=0}^{J^*} \sum_{m=0}^{M^*} p^*(j,m)\Psi(f_1;\delta^*(j,m);\alpha_1), \qquad (A.11)$$

where $J^* = \min(g-1,i)$, $M^* = \min(g-1-j,\ k-i)$, $\delta^*(j,m) = \sqrt{N_1}(m-j)\Delta/\sigma$, and

$$p^*(j,m) = \binom{i}{j}\binom{k-i}{m}\binom{K-k-1}{g-1-j-m} \bigg/ \binom{K-1}{g-1} \qquad \text{(A.12)}$$

Part 4. Derivation of the Probability that an Active Factor that Reaches Second Stage is Declared Important

We let H denote the number of groups in the first stage that show a significant effect. Note that S=gH, so that E(H) = E(S)/g. The expected value of S can be computed per (A.3), (A.9), and (A.11).

We can write

$$P[R_{21}=1|R_{11}=1,\underline{\beta}(i)] = \sum_{h=1}^{G} P[H=h|R_{11}=1,\underline{\beta}(i)]P[R_{21}=1|H=h,\ R_{11}=1,\underline{\beta}(i)]\ . \qquad \text{(A.13)}$$

The second factor within the summation in (A.13) can be evaluated as

$$P[R_{21}=1|H=h,\ R_{11}=1,\underline{\beta}(i)] = \Psi(f_2(h);\delta_2(h);\alpha_2) \qquad \text{(A.14)}$$

where $f_2(h) = B(hg+1) - (hg+1)$ and $\delta_2(h) = -[B(hg+1)]^{\frac{1}{2}}\ \Delta/\sigma$. The conditional distribution of H given $R_{11}=1$ and $\underline{\beta}(i)$ is intractable, however. The authors have found that the conditional distribution of H required in (A.13) can be reasonably approximated as Y+1 where Y is a binomially distributed random variable with parameters (G-1) and success probability $\rho$ = E(H)/G = E(S)/K. Thus

$$P[R_{21}=1|R_{11}=1,\underline{\beta}(i)] \simeq \sum_{h=1}^{G} P(Y=h-1)\Psi(f_2(h);\delta_2(h);\alpha_2) \qquad \text{(A.15)}$$

where $Y \sim b(G-1;\rho)$.

# REFERENCES


Mauro, C.A., and Smith, D.E. (1982), "The Performance of Two-Stage Group Screening in Factor Screening Experiments," Technometrics, 24, 325-330.

Plackett, R.L., and Burman, J.P. (1946), "The Design of Optimum Multifactor Experiments," Biometrika, 33, 305-325.

Watson, G.S. (1961), "A Study of the Group Screening Method," Technometrics, 3, 371-388.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER<br>113-9 | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle)<br>ON THE PERFORMANCE OF TWO-STAGE GROUP SCREENING EXPERIMENTS | | 5. TYPE OF REPORT & PERIOD COVERED<br>Technical Report<br>6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s)<br>Carl A. Mauro | | 8. CONTRACT OR GRANT NUMBER(s)<br>N00014-79-C-0650 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Desmatics, Inc.<br>P. O. Box 618<br>State College, PA 16801 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS<br>NR 042-467 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Office of Naval Research<br>Arlington, VA 22217 | | 12. REPORT DATE<br>August 1983<br>13. NUMBER OF PAGES<br>20 |
| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) | | 15. SECURITY CLASS. (of this report)<br>Unclassified<br>15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this report is unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Factor Screening
Group Screening

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Many experimental situations require consideration of a large number of factors. For example, computer simulation studies are frequently characterized by the inclusion of an exceptionally large number of input variables. Because of resource limitations, there is often a need in such situations for an efficient method of factor screening. One possible screening strategy is two-stage group screening. This report attempts to examine the performance characteristics of two-stage group screening

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) experiments as a function of a constant signal-to-noise ratio for all important factors. The effect of group-factor cancellation on performance is also investigated.

END

DATE
FILMED

9 — 83

DTIC