

UNCLASSIFIED

AD 414521

DEFENSE DOCUMENTATION CENTER

FOR

SCIENTIFIC AND TECHNICAL INFORMATION

CAMERON STATION, ALEXANDRIA, VIRGINIA



UNCLASSIFIED

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

414521

CATALOGED BY DDC

AS AD No. ~~414521~~

THE

TECHNOLOGICAL
INSTITUTE



NORTHWESTERN UNIVERSITY

EVANSTON, ILLINOIS

O. N. R. R E S E A R C H M E M O R A N D U M No. 79

The Technological Institute
Northwestern University

The College of Liberal Arts

Northwestern University

DEMON:

DECISION MAPPING VIA OPTIMUM
GO-NO NETWORKS

A Model for New Products Marketing

by

A. Charnes, W. W. Cooper,* J. K. DeVoe**
and D. B. Learner**

* Carnegie Institute of Technology

**Batten, Barton, Durstine and Osborn, Inc.

July 1963

Part of the research underlying this paper was undertaken for Office of Naval Research and National Institutes of Health projects at Northwestern University and Carnegie Institute of Technology. Reproduction of this paper in whole or in part is permitted for any purpose of the United States Government. Contract Nonr-1228(10), Project NR 047-021, Contract Nonr-760(01), Project NR 047011, N.I.H. project WP 00019-04, and N.I.H. project EF 00355-01.

SYSTEMS RESEARCH GROUP

A. Charnes, Director

DEMON
Decision Mapping Via Optimum Go-No-Go Networks;
A Model For New Products Marketing

By

A. Charnes Northwestern University
W. W. Cooper Carnegie Institute of Technology
J. K. De Voe Batten, Barton, Durstine & Osborn, Inc.
D. B. Learner Batten, Barton, Durstine & Osborn, Inc.

July 1963

This paper is part of a report which is being prepared for presentation at the Tenth International Meeting of The Institute of Management Sciences to be held at the Palace Hotel, Tokyo, Japan on August 24, 1963. The authors wish to acknowledge the considerable help they received from others on the staff at BBDO and especially the guidance, counsel and encouragement they received from Clark Wilson, Tom Dillon and E. E. Norris during all phases of the research.

I. Introduction:

This report covers certain research which has been directed to analyzing and systematizing historically independent market research steps into a rational marketing program for introducing a new product. In the general case, this system is applicable to many situations where decisions under uncertainty must be made in the light of incomplete and not necessarily comparable data.

Historically, "we view the task of introducing a new product as one of taking the product over the hurdles. The hurdles, in this case are the research projects that we conduct to determine its market potential. The deficiency in this approach is that the marketing decisions at each hurdle are not made on the basis of accumulated information - they are based upon the results of the most recent test." ^{1/}

We have, in the following, developed a formal method of integrating all that we know about a product at the time we are ready to make the next marketing planning decision so that an optimal tactical decision related to an optimal strategic objective will result.

The system's inter-relationship of tasks and decisions can best be illustrated and formalized by means of a network. The links of the network represent research tasks and the nodes represent evaluations or decisions undertaken in the light of all prior information. These alternatives together with their associated costs and risks are used to determine an optimal course of action for acquiring further information to achieve an optimum goal.

Figure I is a simplified illustration of such a network. If a company decides to use this network system for analyzing new product potential they must proceed to at least the first node labelled Evaluation I. At this and subsequent evaluation nodes, one of three choices is made in terms of previously stated objectives and policies:

1/ See (11)

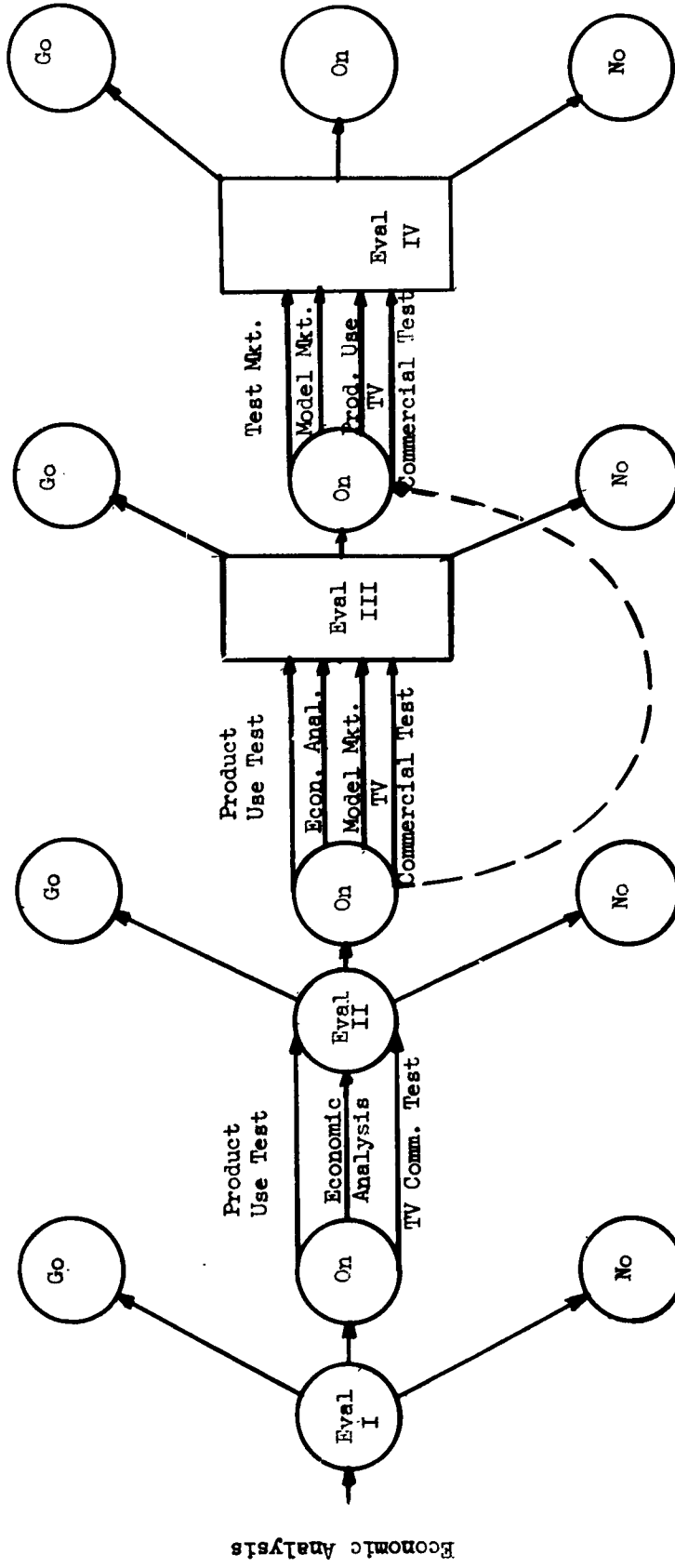
1. "Go" -- This is an abbreviation for "go national" or, more precisely, "begin national marketing of the product" as accumulated evidence indicates that stated company objectives and requirements will all be met.
2. "No" -- This is an abbreviation for "do not market and discontinue testing" since evidence is sufficient to indicate that the company requirements, as initially stated, will not be satisfied.
3. "On" -- This is an abbreviation for "continue testing" since evidence to this point is not sufficient to warrant either a "Go" or a "No" decision.

We note, that any "On" decision must be made in the light of (a) already accumulated information and (b) every alternative route that remains for traversing the network. A conditional choice then, is to be made that optimally designates one of a subset of branches for use in effecting a transit to the next evaluation node. Notice, for instance, that "Product Use Test" might be used in preference to "Economic Analysis" or "TV Commercial Test" in proceeding to "Evaluation II" node in Figure 1. Suppose, for instance, that "Product Use Test" is indicated at this stage. This does not end the matter however, as allowance may be made for a repeat of "Product Use Test" after "On" has been designated following "Evaluation II".

Usually, the data available for such decisions will be of a probabilistic nature and hence the over-riding company policy requirements, objectives, etc., must be stated with this in mind. This suggests a chance-constrained programming formulation with its related apparatus of decision rules, probabilistic constraints, risk evaluation potentials etc.^{1/}

^{1/} See (5) and (3)

Figure 1
NEW PRODUCT NETWORK



As illustrated in Figure 1, the network can (and should) be modified to permit simultaneous choices, such as when a "product use plus economic analysis" branch is introduced as an alternative because there may be a cost saving by initiating both studies at the same decision point. Provision may also be made for "leapfrogging" or bypassing various stages when, accumulated information indicates that intermediate steps are not required, (an arrow connects Evaluation II node directly to a "Product Use Test" which enters into "Evaluation IV", and so on.)

We have implied that an initial decision must be made whether to enter the network at all. Some preliminary analyses will be required at this pre-entry to clarify the relationships between new product decisions and other aspects of total company operations. A company may want to penetrate certain markets because it may be profitable in its own right, help to protect existing product lines, or enhance the company's reputation as a product innovator. It is critical to emphasize that only through new products can a company realize such corporate goals.

This model permits any product to be viewed generically as an information packet containing advertising, packaging, promotion, distribution, etc., attributes along with the usual view of a product in terms of its chemical and physical attributes. The model will distinguish between "study-intelligence" decisions, and "actual marketing implementation decisions." Study-intelligence decisions occur in the process of network traversal whereas marketing implementation occurs after a "Go" decision has been made. Both stages are to be regarded, however, as involving optimal sets of decisions. Specifically, the study-intelligence stage -- which forms the focus of the present analysis -- is regarded in terms of optimal sequence of studies, inquiries, etc., which will yield a "Go" or "No" decision in the most economical fashion. In essence, this phase is the development of an optimal marketing plan, the basis for implementing a sound marketing policy.

A best possible "Go" decision is to be effected and other, less satisfactory, "Go" nodes are to be bypassed if the study data indicate this. All alternatives are evaluated on the assumption that this best possible marketing strategy will be implemented when any Go (or No) is designated. Provision is made for evaluation of such outcomes so that the effects of altering initially stated company requirements and goals can be conveniently examined at every evaluation node.

2. Objectives and Policy Constraints:

The model can accommodate a variety of objectives and requirements. These can be given a number of forms but, we shall assume that the following three attributes are of interest: (1) amount or rate of potential profit (2) length of payout period (breakeven point) and (3) minimum acceptable confidence level for achieving profit and breakeven point. For the sake of concreteness we shall assume that the first attribute is represented by an expected value and, precisely, we shall assume that "maximize expected maximum profit" is a correct statement of company objectives. The payout period length is regarded as a policy constraint with respect to the maximum time allowable for recouping the new product marketing investment. The confidence level expresses the desired odds of achieving the potential profit and payout period.

Any probabilistic statement of objectives such as "maximize expected maximum profit," should be reflected probabilistically in the constraints. Thus the payout period will be stated as a chance constraint so that a "Go" node is designated only if a specified payout period can be achieved at the stipulated minimum level of confidence. This formulation can be varied as the problem may require and, if desired, the payout period can be moved into the function ^{1/} with a restatement of objectives to "maximize

^{1/} See, e.g., (8)

the probability of attaining the specified payout period" instead of "maximize expected maximum profit."

Generally the budget figure for a research study will be known with certainty in the sense of a precise upper limit of expenditure. This is not crucial, however, since the model is developed so that this research budget can be evaluated relative to the consequences of altering an initially budgeted amount. Thus, the study budget is itself subject to probabilistic evaluation relative to profit, payout period, and the estimated costs of pursuing the alternative study paths that the budget limit will admit.

3. Some Aspects of a Strategy for Model Synthesis in New Product Applications:

We have distinguished between the study-planning phase of this new product model and the marketing implementation phase. We have also emphasized that each phase is optimal relative to all alternatives admitted by the model. The study phase must be oriented so that at each step the studies and analyses are designed to yield test estimates (tests of hypotheses, etc.) relative to the objective (maximize expected maximum profit) and the constraints (study budget, payout period and confidence constraints). Other constraints may be introduced in particular applications, of course, and the network relations, as in Figure 1, must be considered. These additional constraint possibilities must then be considered in the light of all accumulated results at every evaluation node.

Some of the ways in which flexibility as well as experience can be built into this model have been suggested. For instance, a branch might be used to bypass Evaluation III and leapfrog over to "On" and Evaluation IV. This mode can be extended and applied in a variety of ways. Corrective action can be introduced when subsequent data suggest it, e.g. when another evaluation node is introduced into Figure 1 to

permit a reversion to "Product Use Test" even after a "Test Market" has been conducted.

Of course these alternatives should generally be built into the model, even at the expense of some redundancy, if there is any uncertainty with respect to the study outcomes. It may be possible to eliminate some redundancies and follow an alternative as checks where previous experience does not provide a wholly reliable guide.

4. Network Data and Objectives:

For this new-product model it is assumed that estimates of national demand are to be obtained in a context which permits us to study ways in which this demand can be influenced. We assume that estimates are to be obtained for national demand which is manifested only as a random variable with unknown mean, μ , and we further allow this mean to be a function of various inputs such as advertising copy, price, distribution, promotion, etc.

For the study phase, it is natural to center on predicting the level of μ while taking account of its behavior as a function.

of these inputs and their associated costs. The predictions for μ will, in any event, be secured through certain chance variates d_{ij} ($i=1,2,\dots,m; j=1,2,\dots,n$;) on which observations are obtainable by undertaking the studies and incurring the costs associated with designated branches over a network of study alternatives. Thus, in general, these d_{ij} variates will form a vector of random variables which can be used to supply information on μ .

For convenience in our subsequent development we now introduce

$$(1) \quad f_{ij}(d_{ij}, \mu)$$

as the probability--or probability density--of observing d_{ij} if μ is the true mean national demand. Next we transfer from the network of Figure 1 into the somewhat more abstract Figure 2. Here, again, we assume that at least a certain portion of the network must be traversed and, starting with node 1, each branch or link is assigned a certain traversal cost and also a certain observational variable d_{ij} . Finally, we adopt an artifact and induce a flow over this network by assigning a unit input for entry into node 1 so that via this artifact,^{1/} all possible traversals can be represented by means of the relevant Kirchhoff conditions. That is, in the symbols shown in Figure 2, we have

$$\begin{aligned} x_{10} + x_{1,n+1} + \sum_{j=1}^n x_{1j} &= 1 \\ (2) \quad x_{i0} + x_{i,n+1} + \sum_{j=1}^n x_{ij} - \sum_{k=1}^n x_{i-1,k} &= 0 \\ x_{ij} &\geq 0 \text{ and integer} \\ i &= 1, 2, \dots, m \text{ nodes} \end{aligned}$$

^{1/} See [4].

which are the Kirchhoff conservation conditions on the nodes for this network.^{1/}

As is readily seen, each x_{ij} will be either unity or zero because of (a) the integer requirement and (b) the fact that only one unit of input is to flow over the system. There will then generally be associated with each link a triple (x_{ij}, c_{ij}, d_{ij}) where $x_{ij} = 0, 1$ designates whether the j^{th} link out of node i is to be used while c_{ij} designates the cost of using this link and d_{ij} represents the information which can be acquired if this link is used. Of course the d_{ij} are variates to be interpreted in accordance with (1), above.

The c_{i0} , which are all on Go nodes, do not include any costs of manufacture since the latter are represented by different symbols which will be introduced subsequently. Also, for $j=0$ or $j=n+1$ we shall set $d_{i0}, d_{i,n+1} = 0$ to denote that these alternatives provide no additional information on μ for the study intelligence aspects of this new product introduction model. For notational convenience we also assume

$$(3) \quad f_{i0}(0, \mu) \equiv 1$$

in conformity with our assumption that no additional study information is generated when a Go node is designated. On the other hand, as already observed, there can be a c_{i0} cost associated with any Go node and hence the total study cost, $\sum_{i,j} c_{ij} x_{ij}$, is to be determined relative to all i and j so that we must have

^{1/} If leapfrogging over nodes is to be considered then we must add terms $-\sum_{r=1}^{i-1} \sum_{k=1}^n y_{rk}$ to the second system of equations, where $y_{rk} \geq 0$ and integer valued.

$$(4) \quad \sum_{i=1}^m \sum_{j=0}^{n+1} c_{ij} x_{ij} \leq c$$

if c is the budgeted limit for the study.

The information secured via the path associated with (4) may be represented as

$$(5) \quad \prod_{\{(i,j) | x_{ij}=1\}} f_{ij}(d_{ij}, \mu),$$

when independence is assumed for the estimates which can be secured. But then, by the definition of a logarithm, we have, explicitly,^{1/}

$$(6) \quad \prod_{\{(i,j) | x_{ij}=1\}} f_{ij}(d_{ij}, \mu) = e^{\sum_i \sum_j x_{ij} \ln f_{ij}(d_{ij}, \mu)}$$

which permits us to show how the estimate \hat{d} of μ depends on the path. Thus, for instance, if the estimator \hat{d} is linear in the d_{ij} with known weights k_{ij} then

$$(7) \quad \hat{d} = \sum_i \sum_j k_{ij} d_{ij} x_{ij}$$

where the k_{ij} are given by factors like stratification or regional proportionality weights, etc.

In actuality the mean national demand, μ , as well as the estimator \hat{d} is a function of the links which are traversed. I.e., both \hat{d} and μ can be altered in accordance with the data findings indicated by these links so that the d_{ij} -- as part of a triple--will vary stochastically in accordance with such things as

^{1/} See [10]. Note also the relation to the likelihood function.

alterations in copy writing, changes in product properties, changes in emphases on selected product properties, etc.

These stochastic relations will generally assume a form like

$$(8) \quad d_{ij} = d_i(1 + s_{ij})$$

or, in some cases,

$$(9) \quad \text{or} \quad d_{ij} = d_i + s_{ij}$$

$$d_{ij} = d_i s_{ij}$$

where the d_i are true values and the s_{ij} are random components. These s_{ij} , and hence d_{ij} , will be associated with the performance of certain functions of a "creative" or "analytical" variety--in accordance with the links to which they are assigned--as well as with resulting improvements in the d_{ij} estimates of d_i as functions of the links that are traversed.

Hence, in general, these d_{ij} estimates may be subjected to additional risk constraints, if desired, relative to possibilities that might attend any alterations in the marketing strategy that will finally be designated.

It should perhaps be emphasized now that each parallel link going from stage i to stage $i+1$ ^{1/} represents a potential total configuration of activities and the corresponding confidence levels must be formed accordingly. See, e.g., (6). Moreover, it may be desirable to provide for over-riding some estimates by other (e.g., later) ones. Thus the functional form for the estimates d_i of μ at stage i , say, need not be linear or involve fixed coefficients, as in (7),

^{1/} These index designations must, of course, be expanded when leapfrogging is used.

and, indeed, the functional forms need not be the same at all stages.

Bearing in mind, now, that a new product's demand is, in general, partly controllable--as indicated by the network path utilized--we can now close this section by stating our objective as

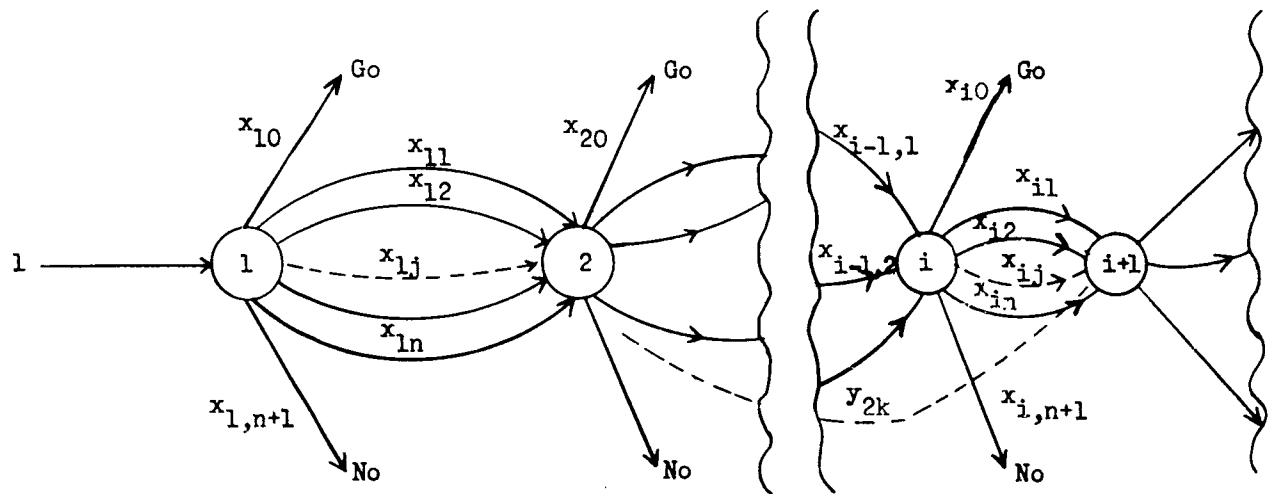
$$\begin{aligned}
 (10.1) \quad & \max E [\pi T d(\pi, x_{ij}) - \sum_i \sum_j c_{ij} x_{ij}] = \\
 & = \max E[\pi T d(\pi, x_{ij})] - \sum_i \sum_j c_{ij} P(x_{ij} = 1),
 \end{aligned}$$

where "E" means "expected value" "P" means "probability" while

$$\begin{aligned}
 (10.2) \quad & \pi = \text{profit rate} \\
 & T = \text{period in which product will} \\
 & \quad \text{continue to be marketed.}
 \end{aligned}$$

and the other terms are as previously defined.

FIGURE 2



5. Probabilistic Constraints and Preemptions:

At each evaluation stage we will have the network constraints to be considered relative to the x_{ij} possibilities. The latter are also conditional on the already obtained d_{ij} values which, in turn, also determine the best estimate of the probability function for d at this stage.

Since we are dealing with new product studies we shall need to have considerable flexibility in this model. One way to obtain this is via the use of certain "preemptions" which we shall shortly examine. First, however, we introduce

(11) $g_i(d)$ = the "current" estimate of the probability density function for d when stage i has been achieved.

We have introduced this definition so we can also write

(12) $g_i(x) = g_i(d_i, x)$

to mean that this estimate depends, through d_i , on the path which has been followed to this point^{1/} so that $g_i(x)$ is a "proximate probability density function" resulting from the path which has been followed. Finally, we let

(13) $G_i(x)$ = cumulative probability function for $g_i(x)$

Now we consider what we shall call a "Go preemption" for purposes of conducting explorations and evaluations at various stages of a study. Formally, such a preemption is designated whenever the following three

^{1/} See, e.g., (2)ff.

conditions are all satisfied:

- (i) Proximate probability of meeting payout requirements are satisfied.
- (ii) E_i (maximum profit) $\geq \pi_i T$ where π_i represents a specified profit rate and T represents a satisfactory interval of time.
- (iii) Confidence in each constraint satisfied $\geq \alpha_i$.
- (iv) $\sum_{r=1}^i \sum_{j=0}^{n+1} c_{rj} x_{rj} \leq c$ so that, at this stage, i, the study budget constraint is still satisfied.

Note, now, that this Go preemption represents a new type of additional constraint. We shall shortly also introduce a "No preemption" but first we want to observe how the above conditions may be used for network explorations. Evidently we can generally raise either the level of α_i , the value of $\pi_i T$, or the suitable length of the payout period until the indicated Go preemption becomes inoperative. This will give a reading on relationships between the α_i , $\pi_i T$, and payout period values at the same time that it enables us to secure further information on the additional network branches that will be utilized when these alterations are effected.

Generally we want to be sure that Go and No options are not taken without further study and evaluation. The meaning of such a Go preemption is that the model is to be parametrically explored and subsequent decisions --e.g., reentry into the model with new objectives and constraints from the current state, or to go national--are to be made on the basis of further study. Hence, the real effect of the model, when used in this fashion, is to obtain a response surface

map of the variation of Maximum Expected Maximum Profit, see (10), with confidence of risk levels α_i and payout period variations as well.

The "No preemptions" are developed in an analogous fashion ^{1/} with a similar response surface resulting from suitably arranged parametric variations. Thus, in general, we put down conditions for "Go-" or "No-preemptions" which involve the x_{ij} and further preemptive variables which are functions of the x_{ij} and d_{ij} up to certain points in the network. As can now be seen we are then able, in this fashion, to explore sections of the network explicitly and in depth. That is, the analytic working of the preemptions in the model is to make the values of immediate and later x_{ij} equal to zero. For instance, if a Go preemption is used then all succeeding x_{ij} for No and On choices are peremptorily equated to zero.

6. Generalized Inverses and Network Path Designations:

The system of network flow equations--see (2)--and the uni-directionality in time of flows may be written in matrix notation as

$$(15) \quad \begin{aligned} A X &= b \\ X &\geq 0 \end{aligned}$$

where A represents the coefficients matrix for the variables x_{ij} and

^{1/} "No preemptions" differ in that any one constraint violation causes a "no" condition.

$$\begin{aligned}
 \mathbf{X}^* &= (x_{10}, \dots, x_{1, n+1}; \dots; x_{m0}, \dots, x_{m, n+1}) \\
 (16) \quad \mathbf{b}^* &= (1, 0, \dots, 0) \\
 \mathbf{C}^* &= (c_{10}, \dots, c_{1n+1}, \dots, c_{m0}, \dots, c_{mn+1})
 \end{aligned}$$

For any matrix A we have uniquely associated with it, a generalized inverse A^\dagger , which is defined as the necessarily unique solution of the four equations ^{1/}

$$\begin{aligned}
 (17) \quad & A A^\dagger A = A \\
 & A^\dagger A A^\dagger = A^\dagger \\
 & (A A^\dagger)^* = A A^\dagger \\
 & (A^\dagger A)^* = A^\dagger A .
 \end{aligned}$$

For our system of network flows--including the leapfrogging requirements--it can be shown that A^\dagger is actually a right inverse ^{2/} of A and

$$(18) \quad A^\dagger = A^* (A A^*)^{-1}$$

where A^* designates the transpose of A . ^{3/} Indeed we find it convenient to

1/ See, e.g., Penrose [12].

2/ I.e., for our network we have $A A^\dagger = I$ where I is the appropriate identity matrix.

3/ We use this symbolism in conformity with now common usages for generalized inverses--which extend to matrices with complex numbers as entries. It is believed that this will not introduce any confusion because of our later use of starred values for optimal program variables.

so still further and use a right inverse which we continue to designate as A^\dagger but which will not have the last two (symmetry) properties of (17). It will, however, be easy to obtain and be better suited to the integrality conditions which are also of concern in this new product network representation.

Since there is no issue of existence for the solutions to (15) we can write

$$(19) \quad X = A^\dagger b + (I - A^\dagger A)Y$$

where Y is, in general, arbitrary. In our case,

$$(20) \quad X = A^\dagger b + (I - A^\dagger A)Y(d_{ij})$$

where $Y(d_{ij})$ denotes the class of conditional (stochastic) dependencies allowed for X . In other words, (20) denotes the class of decision rules which we are permitted to consider. Alternatively, the values for X secured from (20) precisely specify all possible paths we are permitted to consider in the network. The remaining conditions, such as $X \geq 0$ plus preemptions, then become chance constraints.

One might decide to satisfy $X \geq 0$ with probability one (certainty), for example, in which event we would have

$$(21) \quad X = X^+ = \frac{|A^\dagger b + P Y(d_{ij})| + A^\dagger b + P Y(d_{ij})}{2}$$

where $P = (I - A^\dagger A)$ and the absolute value sign is regarded as applying to each component of the vector which it embraces. In this case, then, the functional

would become

$$\begin{aligned} & E[\pi T d(d_{ij}, x_{ij}, \mu)] - EC^* X^+ \\ (22) \quad & = E(\pi T d) - \frac{1}{2} EC^* [|A^+ b + PY(d_{ij})| + A^+ b + PY(d_{ij})] \\ & = E(\pi T d) - \frac{1}{2} C^* A^+ b - \frac{1}{2} C^* PEY - \frac{1}{2} [C^* E |A^+ b + PY|] \end{aligned}$$

We now draw upon preceding work ^{1/} in order to assert that the above development will yield a specific set of "certainty equivalent relations." (Note: $X \geq 0$ means $A^+ b + PY \geq 0$,) which means the solution is in terms of "constrained generalized medians."

7. Illustrative Development of Model Details:

To make these ideas more concrete we now refer to Figure 1 again and represent the network conditions by

$$\begin{aligned}
 & x_{10} + x_{1N} + x_1 & = 1 \\
 & -x_1 + x_{21} + x_{22} + x_{23} & = 0 \\
 & -x_{21} - x_{22} - x_{23} + x_{30} + x_{3N} + x_3 & = 0 \\
 & -x_3 + x_{41} + x_{42} + x_{43} + x_{44} + \hat{y}_{36} & = 0 \\
 & -x_{41} - x_{42} - x_{43} - x_{44} + x_5 + x_{50} + x_{5N} & = 0 \\
 & -\hat{y}_{36} - x_5 + x_{62} + x_{63} + x_{64} + x_{65} & = 0 \\
 & -x_{62} - x_{63} - x_{64} - x_{65} + x_{70} + x_{7N} + x_7 = 0 & = 0
 \end{aligned}
 \tag{23}$$

where the single subscript variables refer to movement from an "Evaluation" to an "On" node and \hat{y}_{36} is the variable which enables us to leapfrog from node 3 (Evaluation II) to node 6 (the On node preceding Evaluation IV). In the variables with double subscript the second index is

coded as follows:

- 0 = Go
- N = No
- 1 = Economic Analysis
- 2 = Product Use Test
- 3 = Channel One Test
- 4 = Model Market
- 5 = Test Market

(24)

Thus $A^\dagger A$ which is 24×24 and $PY = (I - A^\dagger A)Y$, which is 24×1 , become

$$(27) \begin{array}{l} \text{row} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \\ 11 \\ 12 \\ 13 \\ 14 \\ 15 \\ 16 \\ 17 \\ 18 \\ 19 \\ 20 \\ 21 \\ 22 \\ 23 \\ 24 \end{array} \left[\begin{array}{l} +++ \\ \\ -+++ \\ \\ -+++ \\ ---+++ \\ -++++ \\ \\ -++++ \\ ----+++ \\ --++++ \\ \\ --++++ \\ \\ ---+++ \end{array} \right] = A^\dagger A, PY = \left[\begin{array}{l} -y_2 -y_3 \\ y_2 \\ y_3 \\ y_3 -y_5 -y_6 \\ y_5 \\ y_6 \\ y_3 -y_4 -y_5 -y_6 +y_7 \\ y_4 +y_5 +y_6 -y_7 -y_9 \\ y_9 \\ y_9 -y_{11} -y_{12} -y_{13} -y_{14} \\ y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{15} \\ y_9 -y_{10} -y_{11} -y_{12} -y_{13} -y_{14} +y_{16} \\ y_{10} +y_{11} +y_{12} +y_{13} -y_{15} -y_{16} \\ y_{14} +y_{15} -y_{19} -y_{20} -y_{21} \\ y_{19} \\ y_{20} \\ y_{21} \\ y_{14} +y_{15} -y_{18} -y_{19} -y_{20} -y_{21} +y_{22} \\ y_{19} +y_{20} +y_{21} -y_{22} -y_{24} \\ y_{24} \end{array} \right]$$

We observe that X differs from PY only by the addition of a 1 for the first entry in PY . Thus, the condition $X \geq 0$ goes over into

$$y_2, y_3, y_5, y_6, y_9, y_{11}, y_{12}, y_{13}, y_{14}, y_{15}, y_{19}, y_{20}, y_{21}, y_{23}, y_{24} \geq 0$$

and

Thus $A^\dagger A$ which is 24×24 and $PY = (I - A^\dagger A)Y$, which is 24×1 , become

$$(27) \begin{matrix} \text{row} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \\ 11 \\ 12 \\ 13 \\ 14 \\ 15 \\ 16 \\ 17 \\ 18 \\ 19 \\ 20 \\ 21 \\ 22 \\ 23 \\ 24 \end{matrix} \begin{bmatrix} +++ \\ \\ -+++ \\ \\ -+++ \\ ----+++ \\ -+++++ \\ \\ -+++++ \\ ----+++ \\ --++++ \\ \\ --++++ \\ ----+++ \end{bmatrix} = A^\dagger A, PY = \begin{bmatrix} -y_2 - y_3 \\ y_2 \\ y_3 \\ y_3 - y_5 - y_6 \\ y_5 \\ y_6 \\ y_3 - y_4 - y_5 - y_6 + y_7 \\ y_4 + y_5 + y_6 - y_7 - y_9 \\ y_9 \\ y_9 - y_{11} - y_{12} - y_{13} - y_{14} \\ y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{15} \\ y_9 - y_{10} - y_{11} - y_{12} - y_{13} - y_{14} + y_{16} \\ y_{10} + y_{11} + y_{12} + y_{13} - y_{15} - y_{16} \\ y_{14} + y_{15} - y_{19} - y_{20} - y_{21} \\ y_{19} \\ y_{20} \\ y_{21} \\ y_{14} + y_{15} - y_{18} - y_{19} - y_{20} - y_{21} + y_{22} \\ y_{19} + y_{20} + y_{21} - y_{22} - y_{24} \\ y_{24} \end{bmatrix}$$

We observe that X differs from PY only by the addition of a 1 for the first entry in PY . Thus, the condition $X \geq 0$ goes over into

$$y_2, y_3, y_5, y_6, y_9, y_{11}, y_{12}, y_{13}, y_{14}, y_{15}, y_{19}, y_{20}, y_{21}, y_{23}, y_{24} \geq 0$$

and

$$\begin{aligned}
 & 1 - y_2 - y_3 \geq 0 \\
 & y_3 \geq y_5 + y_6 \\
 (28) \quad & y_3 \geq y_4 - y_7 + y_5 + y_6 \geq y_9 \geq y_{11} + y_{12} + y_{13} + y_{14} \\
 & y_9 \geq y_{10} - y_{16} + y_{11} + y_{12} + y_{13} + y_{14} \\
 & y_{10} - y_{16} + y_{11} + y_{12} + y_{13} \geq y_{15} \\
 & y_{14} + y_{15} \geq y_{19} + y_{20} + y_{21} \geq y_{22} + y_{24} \quad , \\
 & y_{14} + y_{15} \geq y_{19} + y_{20} + y_{21} + y_{18} - y_{22}
 \end{aligned}$$

where these variables are the components of Y and hence are not to be confused with the leapfrog variables which were earlier designated by \hat{y} . Note y_4 , y_7 and y_{10} , y_{16} occur only in the form $y_4 - y_7$ and $y_{10} - y_{16}$ in PY so that these combinations can be replaced by y_4 and y_{10} , respectively, without loss of generality. Doing this then the zero-one character of X goes over into the zero-one character of the y_j .

To develop the study budget constraint we observe that $C^*X = C^*A^+b + C^*PY$. Since the first term on the right is constant we can move directly to PY as represented in (27) and write

$$\begin{aligned}
 & c_{10}(1 - y_2 - y_3) + c_{1N}y_2 + c_1y_3 + c_{21}(y_3 - y_5 - y_6) + c_{22}y_5 + c_{23}y_6 + \\
 & + c_{30}(y_3 - y_4 - y_5 - y_6) + c_{3N}(y_4 + y_5 + y_6 - y_9) + c_{41}(y_9 - y_{11} - y_{12} - y_{13} - y_{14}) + \\
 & + c_{42}y_{11} + c_{43}y_{12} + c_{44}y_{13} + \hat{c}_{36}y_{14} + c_5y_{15} + c_{50}(y_9 - y_{10} - y_{11} - y_{12} - y_{13} - y_{14}) + \\
 & + c_{5N}(y_{10} + y_{11} + y_{12} + y_{13} - y_{15}) + c_{62}(y_{14} + y_{15} - y_{19} - y_{20} - y_{21}) + c_{63}y_{19} + \\
 & + c_{64}y_{20} + c_{65}y_{21} + c_{70}(y_{19} + y_{20} + y_{21} - y_{23} - y_{24}) + c_{7N}y_{23} + c_7y_{24} \quad ,
 \end{aligned}$$

or, collecting terms,

$$\begin{aligned}
 & c_{10} + y_2(c_{1N} - c_{10}) + y_3(c_1 - c_{10} + c_{21} + c_{30}) + y_4(c_{3N} - c_{30}) + y_5(c_{3N} - c_{30} + c_{22} - c_{21}) + \\
 & + y_6(c_{23} - c_{21} + c_{3N} - c_{30}) + y_9(c_3 - c_{3N} + c_{41} + c_{50}) + y_{10}(c_{5N} - c_{50}) + \\
 (30) \quad & + y_{11}(c_{42} - c_{41} + c_{5N} - c_{50}) + y_{12}(c_{43} - c_{41} + c_{5N} - c_{50}) + y_{13}(c_{44} - c_{41} + c_{5N} - c_{50}) + \\
 & + y_{14}(\hat{c}_{36} - c_{41} + c_{62} - c_{50}) + y_{15}(c_5 - c_{5N} + c_{62}) + y_{19}(c_{36} - c_{62} + c_{70}) + \\
 & + y_{20}(c_{64} - c_{62} + c_{70}) + y_{21}(c_{65} - c_{62} + c_{70}) + y_{23}(c_{7N} - c_{70}) + y_{24}(c_7 - c_{70}) .
 \end{aligned}$$

Thus, if B_D is a known limit on the permissible expenditures for this new product study we can write

$$(31) \quad \sum k_i y_i \leq B_D - c_{10}$$

as our budget constant where the k_i are obtained from the coefficients of the y_i in (30).^{1/}

Next we develop the constraints on payout period, etc., in a way that will lead into the "Go" and "No" preemptions for Figure I. At Evaluation I we have the d_1 estimate of μ and $F_1(d_1)$. Thus, in terms of the proximate probability and the desired confidence α_{10} for "Go" we have

$$(32) \quad P_1 \{ \pi_1 T_1 d_1 \geq c_{10} \} = \alpha_{10} + u_{10}^+ - u_{10}^- .$$

^{1/} Only the y_i which appear explicitly in (30) are in this sum.

or,

$$(33) \quad P_1 \left\{ d_1 \geq \frac{c_{10}}{\pi_1 T_1} \right\} = \alpha_{10} + u_{10}^+ - u_{10}^- .$$

This may also be expressed in terms of $F_1(d_1)$ by

$$(34) \quad u_{10}^+ - u_{10}^- = 1 - F_1 \left(\frac{c_{10}}{\pi_1 T_1} \right) - \alpha_{10} .$$

Employing a similar course of development for $u_{1N}^+ - u_{1N}^-$, which differs from the right side of (34) in replacement of α_{10} by α_{1N} --the α confidence coefficient sufficiently lower than α_{10} to cause a "No" preemption--we can write

$$(35) \quad u_{10}^+ - u_{10}^- - u_{1N}^+ + u_{1N}^- = \alpha_{1N} - \alpha_{10} < 0 .$$

Next we require "Go" and "No" preemptive variables which are related to the proximate EMP (Expected Maximum Profit). For this we write

$$(36) \quad (EMP)_1 = \pi_1 T_1 E_1(d_1) - c_{10} = W_{10} + w_{10}^+ - w_{10}^-$$

and, similarly for W_{1N} and $w_{1N}^+ - w_{1N}^-$ where the second subscript designates, as usual, association with "Go" or "No." As before, an equivalent definition is provided by

$$(37) \quad w_{10}^+ - w_{10}^- - w_{1N}^+ + w_{1N}^- = W_{1N} - W_{10} < 0 .$$

Having these definitions and relations at hand we can now write the "Go" preemption as

$$(38) \quad x_{1N} + x_1 \leq U(u_{10}^- + w_{10}^-)$$

where U is a suitably large constant. ^{1/} This means that both the α_{10} confidence level and the W_{10} payout level must be equalled or exceeded to cause u_{10}^- and w_{10}^- to be zero and thus $x_{1N} = x_1 = 0$, also. The "No" preemption requires the simultaneous inequalities

$$(39) \quad \begin{aligned} x_{10} + x_1 &\leq Uu_{1N}^+ \\ x_{10} + x_1 &\leq Uw_{1N}^- \end{aligned}$$

to be satisfied. Thus, if either the confidence level is below α_{1N} or the payout is below W_{1N} then $x_{10} = x_1 = 0$ and a "No" decision $x_{1N} = 1$ is preempted.

At Evaluation II we have a d_2 estimate of μ and $F_2(d_2)$. Now, however, we must also be mindful of the stochastic character of the x_{2j} 's (as embodied in the eventual decision rules) in terms of d_1 , the only prior information available.

Thus

$$(40) \quad P_2 \{ \pi_2 T_2 d_2 \geq c_{30} + c_1 x_1 + \sum_{j=1}^3 c_{2j} x_{2j} \} = \alpha_{20} + u_{20}^+ - u_{20}^-$$

^{1/} See, e.g., Chapter VIII in [9].

or, noting $c_1 = 0$,

$$(41) \quad P_2 \left\{ d_2 - \frac{1}{\pi_2 T_2} \sum_{j=1}^3 c_{2j} x_{2j} \geq \frac{c_{30}}{\pi_2 T_2} \right\} = \alpha_{20} + u_{20}^+ - u_{20}^- .$$

Alternatively, as in the Evaluation I case,

$$(42) \quad 1 - \hat{F}_2 \left(\frac{c_{30}}{\pi_2 T_2} \right) = \alpha_{20} + u_{20}^+ - u_{20}^-$$

where \hat{F}_2 denotes the distribution associated with $d_2 - \frac{1}{\pi_2 T_2} \sum_{j=1}^3 c_{2j} x_{2j}$.

Correspondingly,

$$(43) \quad u_{20}^+ - u_{20}^- - u_{2N}^+ + u_{2N}^- = \alpha_{2N} - \alpha_{20} < 0 .$$

The (EMP)₂ expression is

$$(44) \quad \pi_2 T_2 E_2 (d_2 - c_1 x_1 - \sum_{j=1}^3 c_{2j} x_{2j}) - c_{30} = W_{20} + w_{20}^+ - w_{20}^-$$

and the "No" variables relation is

$$(45) \quad w_{20}^+ - w_{20}^- - w_{2N}^+ + w_{2N}^- = W_{2N} - W_{20} < 0 .$$

Since leapfrogging is now a possibility and since the "Go" and "No" preemptions must become redundant constraints if we do not arrive at "Eval II," the variable \hat{y}_{36} must appear on the left-hand side and the quantity $1 - x_1$ must appear on the

right-hand side of the corresponding inequalities as follows:

$$(46) \quad x_{3N} + \hat{y}_{36} + x_3 \leq U(1 - x_1 + u_{20}^- + w_{20}^-)$$

and

$$(47) \quad \begin{aligned} x_{30} + \hat{y}_{36} + x_3 &\leq U(1 - x_1 + u_{2N}^+) \\ x_{30} + \hat{y}_{36} + x_3 &\leq U(1 - x_1 + w_{2N}^+) \end{aligned}$$

The method of forming these constraints having now been illustrated we next record the Eval III conditions as

$$(48) \quad \begin{aligned} u_{30}^+ - u_{30}^- &= 1 - \hat{F}_3 \left(\frac{c_{50}}{\pi_3 T_3} \right) - \alpha_{30} \\ u_{30}^+ - u_{30}^- - u_{3N}^+ + u_{3N}^- &= \alpha_{3N} - \alpha_{30} < 0 \end{aligned}$$

where \hat{F}_3 is associated with the random variable $d_3 - \frac{1}{\pi_3 T_3} \sum_{i=2}^3 \sum_{j=0}^N c_{ij} x_{ij}$.

Analogously,

$$(49) \quad w_{30}^+ - w_{30}^- = \pi_3 T_3 E_3 \left(d_3 - \sum_{i=2}^3 \sum_{j=0, N} c_{ij} x_{ij} \right) - c_{50} - W_{30}$$

and

$$(50) \quad w_{30}^+ - w_{30}^- - w_{3N}^+ + w_{3N}^- = W_{3N} - W_{30} < 0$$

The preemptions are then

$$(51) \quad x_{5N} + x_5 \leq U(1 - x_3 + u_{30}^- + w_{30}^-)$$

and

$$(52) \quad \begin{aligned} x_{50} + x_5 &\leq U(1 - x_3 + u_{3N}^+) \\ x_{50} + x_5 &\leq U(1 - x_3 + w_{3N}^-) \end{aligned}$$

At "Eval IV" the analogous relations are

$$(53) \quad u_{40}^+ - u_{40}^- = 1 - \hat{F}_4 \left(\frac{c_{50}}{\pi_4 T_4} \right) - \alpha_{40}$$

and

$$(54) \quad w_{40}^+ - w_{40}^- = \pi_4 T_4 E_4 (d_4 - \sum_{i=2}^4 \sum_{j=0,N} c_{ij} x_{ij}) - c_{70} - W_{40}$$

where the random variable associated with \hat{F}_4 is $d_4 - \frac{1}{\pi_4 T_4} (\hat{c}_{36} \hat{y}_{36} + \sum_{i=2}^4 \sum_{j=0,N} c_{ij} x_{ij})$.

Since "Eval IV" is a terminal node the terminal "Go" preemption is

$$(55) \quad x_{7N} + x_7 \leq U(1 - \hat{y}_{36} - x_5 + u_{40}^- + w_{40}^-).$$

Since the only other alternative at this node is a "No" it can be assured, where relevant, by the network conditions originally set down together with a very high penalty cost assignment to x_7 . ^{1/}

^{1/} Alternatively one can merely drop the link associated with x_7 .

In the preceding we have, for clarity, first exhibited the use of the generalized inverse both to insure satisfaction of network flow conditions and to indicate the remaining arbitrariness (hence permissible stochastic decision rules) via Y . Second, we have developed the preemptive conditions, again for clarity, in terms of X and an abbreviated notation for the d_i and $X(d_i)$ where relevant. This has been done mainly to highlight the reasons and logic underlying this approach. If, instead, our emphasis were to shift to a more immediately operational focus then we would convert the preemptive inequalities involving the x_{ij} , u , w into equalities by the addition of slack variables and then by appending network conditions and budget constraint (similarly reduced to an equality) we would derive a generalized inverse for the enlarged system to encompass the restrictions and admissible flexibility for stochastic decision rules involving the x_{i0} , u , w .

Thus to conclude in this same spirit we need only indicate how the d_i are formed and thereby, also, how the F_i and \hat{F}_i formations may be considered simultaneously. To fix the ideas we write

$$(56) \quad d_2 = d_1 + (d_{21} - d_1)x_{21} + (d_{22} - d_1)x_{22} + (d_{23} - d_1)x_{23}$$

wherein d_{21} is the known combination of d_1 and the information obtained from traversing the link associated with x_{21} , etc. In particular, if none of the x_{2j} links are traversed--so that $x_{2j} = 0$, $j=1, 2, 3$ --then $d_2 = d_1$; alternatively, if $x_{21} = 1$ then $x_{22} = x_{23} = 0$ and $d_2 = d_{21}$; etc.

It is interesting to examine the expected value expression, $E d_2$, and the variance expression $V(d_2)$. Evidently

$$(57) \quad E d_2 = E d_1 (1 - x_{21} - x_{22} - x_{23}) + \sum_{j=1}^3 E(d_{2j} x_{2j}) .$$

Thus, if the x_{2j} involve, say, linear decision rules in d_1 then $E d_2$ will involve both the mean and variance of d_1 . Furthermore,

$$\begin{aligned} V(d_2) &= E d_2^2 - (E d_2)^2 \\ &= E [d_1^2 (1 - x_{21} - x_{22} - x_{23})^2 + \sum_{j=1}^3 (d_{2j} x_{2j})^2 + \\ &\quad + 2 \sum_{j=1}^3 d_1 d_{2j} x_{2j} - 2 \sum_{j=1}^3 \sum_{k=1}^3 d_1 x_{2j} d_{2k} x_{2k} + \sum_{r \neq s} d_{2r} d_{2s} x_{2r} x_{2s}] - (E d_2)^2 . \end{aligned}$$

But $(1 - x_{21} - x_{22} - x_{23})^2 = (1 - x_{21} - x_{22} - x_{23})$ since these expressions can be, respectively, only one or zero. Also products $x_{2r} x_{2s} = 0$ for $r \neq s$ and $x_{2j}^2 = x_{2j}$.

Thus,

$$\begin{aligned} (59) \quad V(d_2) &= E d_1^2 (1 - x_{21} - x_{22} - x_{23}) + \sum_{j=1}^3 E d_{2j}^2 x_{2j} + \sum_{j=1}^3 E d_1 d_{2j} x_{2j} - \\ &\quad - 2 \sum_{j=1}^3 E d_1 d_{2j} x_{2j} - (E d_2)^2 \\ &= E d_1^2 (1 - \sum_{j=1}^3 x_{2j}) + \sum_{j=1}^3 E(d_{2j}^2 x_{2j}) - (E d_2)^2 . \end{aligned}$$

In the event of linear decision rules for the x_{2j} we can note that $V(d_2)$ depends on both the variance and the skewness--via the 3^d moment--of d_1 .

From this illustration it is also clear that the general method of formation is

$$(60) \quad d_{i+1} = d_i + \sum_j (d_{i+1,j} - d_i) x_{i+1,j}$$

and

$$(61) \quad E d_{i+1} = E d_i (1 - \sum_j x_{i+1,j}) + \sum_j E (d_{i+1,j} x_{i+1,j})$$

while

$$(62) \quad V(d_{i+1}) = E d_i^2 (1 - \sum_j x_{i+1,j}) + \sum_j E (d_{i+1,j}^2 x_{i+1,j}) - (E d_{i+1})^2 .$$

We note that the above development does not assume statistical independence. Rather, the network x_{ij} variables have eliminated certain types of "covariance" terms. On the other hand, certain higher moments--e.g., those related to skewness--are coming in naturally as a result of the statistical compounding which leads to new distributions as the network decisions are being effected.

8. Additional Remarks and Conclusion:

In subsequent reports we shall develop this model through suitable numerical illustrations which will clarify the nature of the decision rules. This expansion will involve techniques of approximating the distributions of various random variables with mixtures of normal distributions. These mixtures have convenient properties (for example, closure under the operations of linear combinations of such variables) and they offer explicit expressions of the parameters of the distributions of these linear combinations. Such techniques will enable us to handle explicitly nonlinear relations of the unknown parameters in the decision rules. This network approach provides a systematic way of (a) mapping an initial set of alternatives and (b) evaluating the consequences of various policies and constraints. We can direct the model towards an overall objective subject to chance (and other) constraints while retaining the flexibility needed to explore alternatives that might be discovered in the course of gathering data. Thus, we might study risk and quality features, in the pay-out period constraints, to enable us to portray clearly the effect on profit expectations that results from tightening or loosening these constraints. Alternatively we can move to the profit objective itself and examine, the effects of varying the time parameter T to ascertain how this might effect the study strategies and costs. Finally, we observe that the network is itself subject to manipulation in ways that can be employed to advantage prior to forming a study strategy. Here, in particular, an opportunity is provided for utilizing any experience or insight that may be available from persons with expert knowledge of particular products in a systematic and flexible fashion.

BIBLIOGRAPHY

- (1) Alderson and Shapiro, Marketing and the Computer, (Englewood Cliffs, New Jersey, Prentice Hall Inc., December 1962)
- (2) Boomer, P., "Test Marketing: Boon or Bust?" Printers' Ink 282, No. 2, April 5, 1963, pp. 23-32
- (3) _____, "Test Marketing: The New Horizons," Printers' Ink 282, No. 2, April 12, 1963, pp. 23-31
- (4) Charnes, A. and W. W. Cooper, "A Network Interpretation and a Directed Sub-Dual Algorithm for Critical Path Scheduling;" Journal of Industrial Engineering XIII, No. 4, July-August 1962, pp. 213-218
- (5) Charnes, A. and W. W. Cooper, "Chance Constrained Programming," Management Science, October, 1959
- (6) Charnes, A., W. W. Cooper and Thompson, "Chance Constrained Studies for Linear Programming under Uncertainty," Parts I and II. ONR Research Memos, Northwestern University, December 1962
- (7) _____, _____ and _____, "Critical Path Analysis via Chance Constrained and Stochastic Programming," ONR Research Memo No. 101, Northwestern University - Carnegie Institute of Technology, July, 1962
- (8) Charnes, A. and W. W. Cooper, "Deterministic Equivalents for Optimizing and Satisficing under Chance Constraints" Operation Research 11, No. 1, January, February 1963 pp. 18-39
- (9) Charnes, A. and W. W. Cooper, Management Models and Industrial Applications of Linear Programming (New York: John Wiley & Sons Dec., 1961)
- (10) Charnes, A. and W. W. Cooper "Optimizing Engineering Designs under Inequality Constraints," ONR Research Memo No. 64 Northwestern University - Carnegie Institute of Technology, August, 1962
- (11) Learner, D. B. "Innovation, Progress -- or a Random Walk," Proceedings of the 46th National Conference, American Marketing Association (forthcoming)
- (12) Penrose, R., "The Generalized Inverse for Matrices" Proceedings of the Cambridge Philosophical Society, 51, 1955, pp. 406-413
- (13) Rao, C. R., "The Generalized Inverse of a Matrix," Journal of the Royal Statistical Society, Series A, February, 1962
- (14) Wilson, C. L., "Use of Linear Programming to Optimize Media Schedules in Advertising" Proceedings of the 46th Annual Conference, American Marketing Association (forthcoming)