COWLES FOUNDATION DISCUSSION PAPER NO. 75

Note: Cowles Foundation Discussion Papers are preliminary materials circulated privately to stimulate private discussion and critical comment. References in publications to Discussion Papers (other than mere acknowledgment by a writer that he has access to such unpublished material) should be cleared with the author to protect the tentative character of these papers.

Some Team Models of a Sales Organization*

C. B. McGuire

August 11, 1959

(Revision and extension of Cowles Foundation Discussion Paper No. 53)

* Research undertaken by the Cowles Commission for Research in Economics under Contract Nonr-358(01), NR 047-066 with the Office of Naval Research

Some Team Models of a Sales Organization*

by C. B. McGuire

1. Introduction

The models presented in this paper grew out of an attempt to apply the theory of teams of Marschak and Radner [1], [2], [3], [4] to the organization of the sales force in a typical wholesale bakery. As organizations go, the one we have chosen to analyze is an extremely simple one, even when -- as is quite obviously not the case in what follows -- it is viewed in its full complexity. At the risk of achieving results of quite limited general interest, a "simple," easily quantified subject was deliberately chosen as the best place to begin to apply a theory that pretends to prescribe optimum decisions of a team in a precise way.

We shall not here go much beyond a discussion of some mathematical models which deal only with the day-to-day problem a bakery sales force faces in attempting to "properly" supply its regular customers with a single product. Problems of advertising, of price policy, of product design, and of obtaining new customers are ignored. All the reader needs to know, to begin with, is that the sales force consists of truck-driver salesmen who daily visit each of their given customers (i.e. grocery stores) leaving, on consignment, an amount of bread to be decided by the salesman. At the end of the day the salesman returns to the plant and submits an order for the next day. For our purposes here the "organization" is characterized by the way in which these orders are jointly formed. In the last section, after the models have pin-pointed the organization problem, we shall return to a more detailed description of procedures found in practice.

A suggestion of Radner's, [3], [4] that certain organization problems can be formulated in linear programming terms, has been followed assiduously. The present examples, insofar as they are successful, hint that this approach may be quite widely useful. We shall find that severe computational problems remain, but we shall also suggest some possibilities that have so far been little exploited.

* Many of the points discussed in this paper first arose in my discussions with Jacob Marschak and Martin Beckmann. I am heavily indebted and grateful to both of them and to my other colleagues at the Cowles Foundation at Yale.

2. The Team Problem of Marschak and Radner

2.1. A General Formulation

In the Marschak-Radner formulation of the team problem, decision (or "action") functions $\alpha_i(x)$, $(i=1, \ldots, N)$, are to be determined which tell each of N team members how to act when the state of the world is x. If each member is perfectly informed about the value of x, then the payoff to the team as a whole for given x and given functions α_i is a real number $u[\alpha_1(x), \ldots, \alpha_N(x); x]$ minus whatever costs are incurred in making the value of x known to each of the members. If observation and communication are costly it will generally not be best to fully inform each member. In order to specify systems less costly in this respect let the function $\eta_i(x)$, $(i=1, \ldots, N)$, denote the information about x that is made available to the i^{th} team member. Thus if x is an M-tuple $(x_1, \ldots, x_M)$, one particularly simple example of an information function might be $\eta_i(x_1, \ldots, x_M) = x_i$. More generally η_i is some function from the set X of states of the world to the set of all subsets of X. With each of these information structures $\eta = (\eta_1, \ldots, \eta_N)$ is associated a cost $k(\eta)$. With a given structure then and a known probability distribution over X of states of the world the N-tuple of decision functions $\hat{\alpha} = (\hat{\alpha}_1, \ldots, \hat{\alpha}_N)$ is to be found which maximizes the expected payoff to the team *

(1) $$E\left\{u[\alpha_1(\eta_1(x)), \ldots, \alpha_N(\eta_N(x)); x]\right\} - k(\eta) .$$

---

* Notice that $k(\eta)$, the cost of the information structure, is here taken to be a function of η independent of the state of the world. This need not be so. Whenever costs of communication depend on the intensity of use of certain communication links -- as in the case of long-distance telephoning -- k will depend also on x and will occur under the expectation sign in (1).

Once a procedure is found for determining, for any given structure η, the corresponding optimum set of decision functions $\hat{\alpha}$, the goodness of different structures can be compared.

2.2 A More Specific Formulation

In this section a more specific, but still quite general formulation of the team problem is proposed. Let the state of the world be represented by a (2N+1)-tuple, $x = (z_1, \ldots, z_N; y_1, \ldots, y_N; w)$. The y_i are demand parameters in each of several markets. No particular real interpretation of the z_i will be adhered to; we suppose only that they are statistically related to the respective y_i and yet are very much easier to observe. That different states of the world can affect matters of production is recognized by distinguishing such states by the single variable w. (Through most of this paper w is fixed and known.)

Each of N salesmen will have an information function η_i $(i = 1, \ldots, N)$ which will be a function only of the $z_1, \ldots, z_N$, and the $y_1, \ldots, y_N$. The latter are supposed to be prohibitively expensive to observe. In this limited context a completely coordinated organization would imply an information structure η in which $\eta_i(z_1, \ldots, z_N; y_1, \ldots, y_N) = (z_1, \ldots, z_N)$, or more briefly $\eta_i(x) = z$, for each i; that is, where every salesman knows all that can "reasonably" be known about the demand parameters in all markets. The most obvious example of η in an uncoordinated system is $\eta_i(x) = z_i$ for every i. Since our investigation of the set of information structures will not in any case be exhaustive, we shall be concerned with picking out these "obvious" structures for analysis.

The action variables $a_i = \alpha_i[\eta_i(x)]$ will be interpreted as supply actions or "orders" by the salesmen: in response to the information $\eta_i(x)$, Salesman i supplies a_i units of product to his market.

The payoff to the team will be taken to be the expected profit of the firm: expected revenue from sales minus expected cost of production minus cost of information structure. More specifically the payoff is

$$(2) \quad ER(a_1, \ldots, a_N; y_1, \ldots, y_N) - EC(a_1, \ldots, a_N; w) - k(\eta)$$

where R is a revenue function and C a production cost function. It is assumed that the z_i influence profit only through their rôle as arguments in the action functions. The y_i do not influence production costs, and w does not influence demand.

For a team problem to be something more than a collection of N one-person problems there must be a source of what Marschak has called "interaction" among the actions of the N members: the effect of one man's decision of the payoff must not be quite independent of the actions of all of the other men (in continuous terminology, $\partial^2 u | \partial a_i \partial a_j$ is not always zero for all i and j). In the present instance interaction among decisions comes only from the production cost relationship. It will be supposed that production costs depend only on the sum of the individual orders; revenue from one market will be assumed to depend only on the state of the world in that market (y_i in market i) and on the supply decision in that market. The payoff function (2) can therefore be further specified:

$$(3) \quad \sum_{i=1}^{N} ER_i(a_i, y_i) - EC(\sum_{i=1}^{N} a_i, w) - k(\eta)$$

where R_i is the revenue function in market i .

Payoff function (3) represents the general situation to be investigated in this paper: one plant supplying N markets. Three analogous problems suggest themselves. We shall do no more than mention them.

The first is the "opposite" of (3): N plants supplying one market, with payoff

$$(4) \quad ER(\sum_{i=1}^{N} a_i, y) - \sum_{i=1}^{N} EC_i(a_i, w_i) - k(\eta)$$

where the state-of-the-world variables are appropriately and analogously redefined. One would expect a strong similarity between the results that analysis of (3) yields and results from (4).

The second is N plants supplying N markets with no plant impinging on another's territory:

$$(5) \quad \sum_{i=1}^{N} ER_i(a_i, y_i) - \sum_{i=1}^{N} EC_i(a_i, w_i) - k(\eta).$$

The team aspect has disappeared, leaving N one-person problems.

The third analog is N plants jointly supplying N markets:

$$(6) \quad ER(\sum_{i=1}^{N} a_i, y) - EC(\sum_{i=1}^{N} a_i, w) - k(\eta)$$

In this very simple form, where the team is indifferent as to which plant supplies a unit of product or in which market a unit is sold, it might appear that again the problem degenerates into a one-person situation. But the very multiplicity of solutions that makes the problem look easy causes trouble in the absence of coordination:* no amount

---

* The term "coordination" is used here in the sense of a once-and-for-all coordination, as opposed to the coordination supplied day-to-day by a team member (absent in our model) whose actions vary as the states of the world changes. The once-and-for-all type is exemplified by the (legally enforced!) convention among drivers to use their respective right-hand sides of a road. The example is from Marschak, who has emphasized the relation between the need for coordination and the uniqueness of $\hat{\alpha}$.

---

of information about the external state of the world is sufficient for one team member to act in a way that complements the action of his fellows. If the optimum N-tuple of action functions were unique, coordination would contribute nothing and a one-person problem would result.

3. Model I

One of the simplest interesting models with payoff function of the form (3) is the following one.

Let the random variable y_i be the price in market i , and suppose this price is unaffected by the supply action a_i of salesman i . The total revenue curve in market i will then be some straight line through the origin, the slope of which is, in general, not precisely known at the time the supply decision is made. The random variable z_i is observed by salesman i on the preceding day and is used to predict y_i . If we like, we may regard z_i as the price prevailing in the earlier period and suppose that the price series is autocorrelated. The joint distribution $p(z_1, \ldots, z_N; y_1, \ldots, y_N)$ is known. We will suppose further that the z_i and the y_i vary discretely.

Total cost of production is, in accord with (3), a function of the sum of the supply actions

$$(7) \quad c_a \sum_i a_i + (c_b - c_a) \operatorname{Max} [\sum_i a_i - w, o]$$

with $w > o$, $c_b > c_a > o$. The parameters c_a, c_b and w are assumed to be fixed and known to all participants, although in a more comprehensive study of this organization it would be necessary to include these three cost parameters in the state-of-the-world vector where they would be observed and communicated at some cost.

3.1. Centralization

If the information structure is such that every member has all obtainable information about demand in all markets then, for every i, $\eta_i(x) = z$. Given this structure, the problem is to find a set of action functions $\alpha_i(z)$, i=1, .., N, which maximizes expected profit. In order to write payoff as a linear function of the $\alpha_i(z)$, let us define a non-negative function $b(z)$ with

$$(8) \quad b(z) \geq \sum_i \alpha_i(z) - w .$$

If $b(z)$ is chosen to be as small as possible for any set of $\alpha_i(z)$, then payoff can be written

$$(9) \quad \sum_z \sum_y \left\{ \Sigma (y_i - c_a) \alpha_i(z) - (c_b - c_a) b(z) \right\} \; p(z,y)$$

where the summation signs on the left stand for $\sum_{z_1=o}^{K} \dots \sum_{z_N=o}^{K}$ and $\sum_{y_1=o}^{K} \dots \sum_{y_N=o}^{K}$ respectively, K being some high integer beyond the ranges of variation of the z_i and the y_i.

We must now select $\alpha_i(z)$ for each i and z and $b(z)$ for each z so as to maximize (9), subject only to constraint (8) on $b(z)$ and non-negativity of $b(z)$ and $\alpha_i(z)$. In case the range of y_i extends above c_b it will be necessary to put an upper bound on the $\alpha_i(z)$, say $\bar{a}$, otherwise the maximum of (9) will be unbounded. Since both the maximand and the constraints are linear in the $b(z)$ and the $\alpha_i(z)$, the necessary and sufficient conditions for the optimum values of these variables can be found by differentiating the Lagrangean expression L. Let $\lambda(z) \geq 0$ be the multiplier associated with constraint (8). Then the conditions are

$$(10) \qquad \frac{\partial L}{\partial \alpha_i(z)} = \sum_y (y_i - c_a) p(z,y) - \lambda(z) \begin{Bmatrix} \leq \\ = \\ \geq \end{Bmatrix} 0 \quad \text{if} \begin{cases} 0 = \alpha_i(z) \\ 0 < \alpha_i(z) < \bar{a} \\ \alpha_i(z) = \bar{a} \end{cases}$$

and

$$(11) \qquad \frac{\partial L}{\partial b(z)} = -(c_b - c_a) \sum_y p(z,y) + \lambda(z) \begin{Bmatrix} \leq \\ = \end{Bmatrix} 0 \quad \text{if} \quad b(z) \begin{Bmatrix} = \\ > \end{Bmatrix} 0$$

with $\lambda(z)$ vanishing if strict inequality holds in (8).

Constructing a solution will be easier if (10) and (11) are rewritten in more familiar terms:

$$(12) \qquad E(y_i | z) \begin{Bmatrix} \leq \\ = \\ \geq \end{Bmatrix} c_a + \frac{\lambda(z)}{p(z)} \qquad \text{if} \begin{cases} 0 = \alpha_i(z) \\ 0 < \alpha_i(z) < \bar{a} \\ \alpha_i(z) = \bar{a} \end{cases}$$

and

$$(13) \qquad \lambda(z) \left\{ \begin{matrix} \leq \\ = \end{matrix} \right\} (c_b - c_a)p(z) \quad \text{if} \quad b(z) \left\{ \begin{matrix} = \\ > \end{matrix} \right\} 0$$

where $p(z) \equiv \sum_y p(z,y)$.

Let us fix our attention on a particular value of z . If the conditional price expectations in all markets are below c_a , the lowest extreme of marginal production cost, then intuition and (12) both assert that $\alpha_i(z) = 0$ for all i .

If the highest conditional price expectation exceeds c_a , then, by (12), either $\alpha_i(z) = \bar{a}$ in this high market or $\lambda(z)$ is positive. But by assumption, $\bar{a}$ is a high number so, by (8), $b(z)$ must be positive and hence, by (13), $\lambda(z)/p(z) = c_b - c_a$. In either case a price expectation exceeding c_a implies a positive value of $\lambda(z)$. This in turn implies equality in constraint (8) , which means that $\sum_i \alpha_i(z) \geq w$.

If the highest expected price falls <u>between</u> c_a and c_b then (12) and $\sum_i \alpha_i(z) \geq w$ can be simultaneously satisfied only by $\alpha_i(z) = w$ in the highest* market and zero in other markets.

If one or more of the price expectations exceeds c_b then $\alpha_i(z) = \bar{a}$ in all such markets and $\alpha_i(z) = 0$ in the others since the left side of (12) can never exceed c_b .

This elliptical** argument has been presented only in an effort to motivate the following solution, the validity of which is easily checked against (8), (12), and (13):

* If there are n "highest" markets, each of the corresponding $\alpha_i(z)$ can be positive so long as $\sum_i \alpha_i(z) = w$.

** The distinction between $<$ and $\leq$ was ignored for one thing.

Let k denote the maximizer of $\operatorname*{Max}_i E(y_i|z)$ for a given z.

Then

(14) if $E(y_k|z) \leq c_a$, set $\alpha_1(z) = \ldots = \alpha_N(z) = b(z) = \lambda(z) = 0$;

if $c_a < E(y_k|z) \leq c_b$, set
$$\alpha_k(z) = w$$
$$\alpha_i(z) = 0 \quad \text{for } i \neq k$$
$$b(z) = 0$$
$$\lambda(z) = (c_b - c_a)p(z)$$

if $c_b < E(y_i|z)$ and $E(y_j|z) \leq c_b$ set
$$\alpha_i(z) = \bar{a}$$
$$\alpha_j(z) = 0$$
$$b(z) = \Sigma\, \alpha(z) - w$$
$$\lambda(z) = (c_b - c_a)p(z) .$$

The solution is not (quite) unique, the choices between $<$ and $\leq$ having been made arbitrarily in (14). For those values of z however, where the delicate issue of $<$ versus $\leq$ in applying (14) never arises, the values for the $\alpha_i(z)$ (but not necessarily $\lambda(z)$) are unique. In the case $N=2$ a graphical representation of (14) indicates the areas of non-uniqueness. In Figure 1 the optimum values of $\alpha_i(z)$ are shown over various regions of the space of conditional price expectations. Along the borders common to two regions (or the points common to three or four regions) all interpolated values are optimal. As long therefore as z is in the interior of a region, the intelligent salesman will know how to act

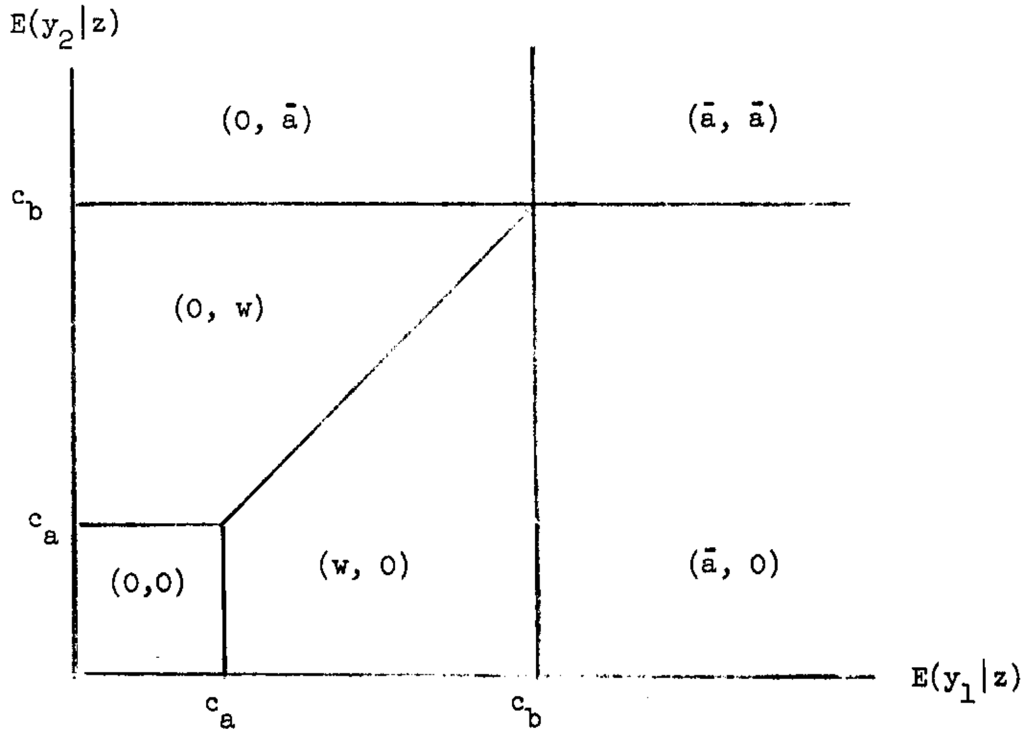

Figure 1. Optimum Supply Actions $(\alpha_1(z), \alpha_2(z))$

optimally, but if z occurs on a border he will not. Complete information about the external state of the world in these cases does not lead to automatic coordination.

3.2 Decentralization

Here we suppose the information structure η to be such that $\eta_i(x)=z_i$ for every i ; each salesman knows his own market predictor, but not those in other markets. The action functions to be determined are now $\alpha_i(z_i)$ instead of $\alpha_i(z)=\alpha_i(z_1, \ldots, z_N)$ as in the complete information case of the preceding section.

The constraint on $b(z)$ is now

$$(15) \qquad b(z) \geq \sum_i \alpha_i(z_i) - w$$

and the function to be maximized (which of course represents the payoff if $b(z)$ is properly chosen) is

$$(16) \qquad \sum_{z}\sum_{y} \left\{ \sum_i (y_i - c_a)\alpha_i(z_i) - (c_b - c_a)b(z) \right\} \quad p(z,y) \; .$$

Proceeding just as before we find conditions on the $\alpha_i(z_i)$ and $b(z)$:

$$(17) \qquad \frac{\partial L}{\partial \alpha_i(z_i)} = \sum_{\substack{z_j \\ j \neq i}} \left[\sum_y (y_i - c_a) p(z,y) - \lambda(z) \right] \begin{Bmatrix} \leq \\ = \\ \geq \end{Bmatrix} 0 \text{ if } \begin{cases} 0 = \alpha_i(z_i) \\ 0 < \alpha_i(z_i) < \bar{a} \\ \quad\; \alpha_i(z_i) = \bar{a} \end{cases}$$

and for $\frac{\partial L}{\partial b(z)}$ condition (11) again. And as before, when (17) is translated into terms of conditional expectation we have

$$(18) \qquad E(y_i | z_i) \begin{Bmatrix} \leq \\ = \\ \geq \end{Bmatrix} c_a + \sum_{\substack{z_j \\ j \neq i}} \frac{\lambda(z)}{p(z_i)} \text{ if } \begin{cases} 0 = \alpha_i(z_i) \\ 0 < \alpha_i(z_i) < \bar{a} \\ \quad\; \alpha_i(z_i) = \bar{a} \end{cases}$$

where $p(z_i)$ stands for $\sum_{\substack{z_j \\ j \neq i}} p(z)$.

For convenience, we repeat

$$(13) \qquad \lambda(z) \begin{Bmatrix} \leq \\ = \end{Bmatrix} (c_b - c_a)p(z) \qquad\qquad \text{if } \; b(z) \begin{Bmatrix} = \\ > \end{Bmatrix} 0 \; .$$

In attempting to construct a solution, we start by concentrating on a given i and given value of z_i . We assume that a solution exists in which the $\lambda(z)$ take only the values zero or $(c_b - c_a)p(z)$ (sometimes of course this assumption is violated, as will become clear in a moment - - in any case no harm is done by investigating the construction it leads to).

If $E(y_i|z_i) < c_a$ then $\alpha_i(z_i)$ must be zero from (18). If $E(y_i|z_i) \geq c_a$ then the sum on the right of (18) must be evaluated. From the above assumption about the possible values of the $\lambda(z)$ we know that when z is such that $\Sigma\alpha$ exceeds w, $(c_b-c_a)p(z)$ will enter the sum for $\lambda(z)$, and when z is such that Σa falls short of w, no contribution results. The right-hand sum in (18) therefore can be written $(c_b-c_a)p'$, where p' is the conditional probability given z_i that $\Sigma\alpha > w$. We have therefore that

$$(19) \qquad \frac{E(y_i|z_i)-c_a}{c_b-c_a} \left\{ \begin{matrix} \leq \\ = \\ \geq \end{matrix} \right\} p' \qquad \text{if} \quad \left\{ \begin{matrix} 0 = \alpha_i(z_i) \\ 0 < \alpha_i(z_i) < \bar{a} \\ \alpha_i(z_i) = \bar{a} \end{matrix} \right.$$

with the qualification that the $\lambda(z)$ must take their extreme values only.

A graphical interpretation of the uses to which (19) can be put will be given for the very special case where $N=2$ and when $E(y_i|z_i) = z_i$. Suppose one suspects that both action functions $\alpha_i(z_i)$ are non-decreasing in z_i. If this were so then $z_2' > z_2$ and $\alpha_1(z_1) + \alpha_2(z_2) > w$ would imply $\alpha_1(z_1) + \alpha_2(z_2') > w$ and similarly for $z_1' > z_1$. All the positive $\lambda(z)$ would therefore occur in the Northeast part of the z space of Figure 2. Condition (19) then suggests the definition of a function $g_1(z_1)$ by*

$$(20) \qquad \sum_{z_2=g_1(z_1)}^{K} p(z_1,z_2) \equiv \frac{z_1 - c_a}{c_b - c_a} \cdot p(z_1) \qquad (c_a \leq z_1 < c_b)$$

---

* In the present discussion we can ignore the discreteness of the z_i.

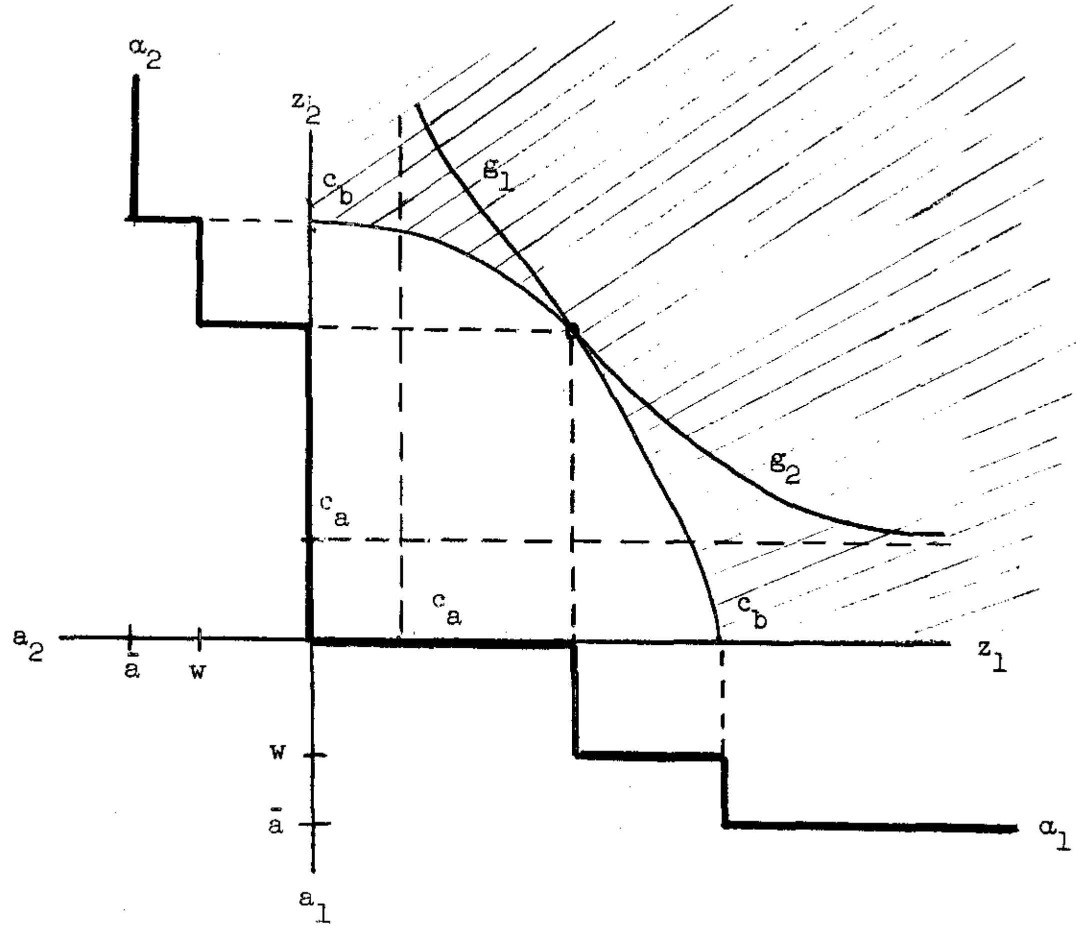

Figure 2

and a function $g_2(z_2)$ in the same fashion. When $z_1 = c_a$, g_1 must be at the "top" of the conditional z_2 distribution, and when $z_1 = c_b$, g_1 must be at the "bottom". Figure 2 shows a pair of very well behaved g functions -- well behaved in the sense that they intersect only once and are monotonic. Functions of this simple kind will result if z_1 and z_2 are independent, or if they are inversely correlated and $p(z)$ is not too misshapen.

Now set $\lambda(z) = (c_b - c_a)p(z)$ when z is in the shaded area of Figure 2 and $\lambda(z) = 0$ otherwise. Next define three-step action functions as in Figure 2 with ordinates 0, w, and $\bar{a}$, with the first steps occurring at the coordinates of g_1-g_2 intersection, and the second steps at c_b. That the three functions $\lambda(z)$, $\alpha_1(z_1)$, and $\alpha_2(z_2)$ so defined are optimum is easily established by checking against (18) and (13). Also, it should be noted, the solution is unique.

When z_1 and z_2 are positively correlated the procedure just described will quite often not work, because the g curves intersect in more than one place and/or lack the monotonic properties on which the earlier construction rested. Figure 3 portrays one such case. It will be recalled however that the definition of the g curves was prompted

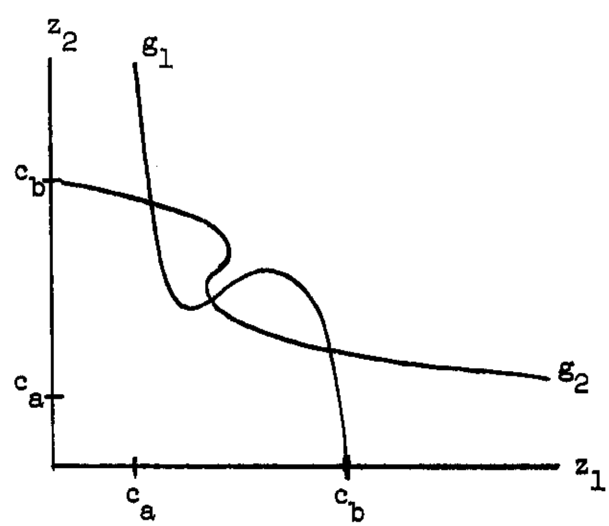

Figure 3

by a hunch that the action functions were non-decreasing. But when z_1 and z_2 are positively correlated, we would expect rather that <u>one</u> of the action functions will be non-decreasing and the other non-increasing. By "reversing" one of the g functions (i.e.,

$\sum_{z_2=0}^{g_1(z_1)} p(z_1,z_2) = \frac{z_1-c_a}{c_b-c_a} \cdot p(z_1)$) and making the obvious changes in the construction (i.e., the positive $\lambda(z)$ are put in the Northwest or Southeast corner of the z space) similarly simple step-function solutions for these positively correlated distributions can sometimes be found.

Still another class of "easy" solutions occurs with positively correlated discrete distributions which are nearly symmetric about the line $z_1=z_2$. In these cases the action functions are constant in the interval $[c_a, c_b]$ and zero and $\bar{a}$ respectively below and above this interval. The middle levels a_i need only be chosen so that $a_1+a_2 = w$; here of course the need, discussed earlier, for coordination in selecting complementary action functions is again acute.

The geometric methods suggested above for constructing solutions to the decentralized problem are obviously very limited: the changes required to replace z_i with the original $E(y_i|z_i)$ are minor, but the extension to $N > 2$ raises quite new questions. Moreover, it will have been observed that maximizing (16) subject to (15), with the restriction that the variables be non-negative, is a standard problem of linear programming; the armory of established techniques for these problems can be put to use. The point is however that the number of variables (and constraints) in our linear programming formulation is so extremely high as to make ordinary computational techniques quite impracticable for many interesting problems. The formulation is uneconomic in the sense that it makes no use of regularity properties the $p(z,y)$ distribution might have (e.g., a single mode); but rather is ready to deal with bizarre checkerboard distributions which never occur. One hopes that geometric arguments like those presented will provide a time-saving guide for machine computation, both by suggesting computational methods geared to the peculiarities of the problems, and by classifying problems.

3.3. A Mixed Case - Partial Spread of Information

Let $N=2$, $\eta_1(x)=z_1$, and $\eta_2(x)=(z_1,z_2)$. The maximand becomes

$$(21)\quad \sum_z \sum_y \left\{(y_1-c_a)\alpha_1(z_1)+(y_2-c_a)\alpha_2(z_1,z_2)-(c_b-c_a)b(z)\right\} \; p(z,y)$$

and the constraint on $b(z)$

$$(22)\qquad b(z) \geq \alpha_1(z_1) + \alpha_2(z_1,z_2) - w \; .$$

The optimum condition on $\alpha_1(z_1)$ is

$$(23)\quad \sum_{z_2} \left\{\sum_y (y_1-c_a)p(z,y) - \lambda(z) \right\} = 0 \text{ , etc.}$$

$$\text{or}\quad E(y_1|z_1)-c_a - \sum_{z_2} \frac{\lambda(z_1, z_2)}{p(z_1)} = 0 \text{ , etc.}$$

And for $\alpha_1(z_1,z_2)$

$$(24)\quad \sum_y (y_2-c_a)p(z,y) - \lambda(z) = 0\text{, etc.}$$

$$\text{or}\quad E(y_2|z_1,z_2)-c_a - \frac{\lambda(z_1,z_2)}{p(z_1,z_2)} = 0\text{, etc.}$$

If in the middle range of z we can manage to keep $\Sigma\alpha=w$ then $\lambda(z)$ can take any non-negative value not exceeding $(c_b-c_a)p(z)$, from (13). In particular it can take the value it has when equality holds in (24). Substituting this value in (23) we get

$$(24)\quad E(y_1|z_1)-c_a - \sum_{z_2} \left\{ E(y_2|z_1,z_2)-c_a \right\} \frac{p(z_1,z_2)}{p(z_1)}$$

or

$$E(y_1|z_1)-E(y_2|z_1) = 0 \text{ , etc.}$$

Thus when z_i is in the interval $[c_a, c_b]$, Man 1 supplies w whenever his conditional expectation of $y_1 - y_2$ exceeds zero. When z_2 is in the $[c_a, c_b]$ interval Man 2 supplies w when Man 1 is not acting and nothing when he is. The total then is always w in the middle z range, so the substitution for $\lambda(z)$ in (24) was permissible. As before both men order zero and $\bar{a}$ respectively below and above the $[c_a, c_b]$ interval.

An interesting question suggested by this result is whether the function $\alpha_1(z_1)$ just derived differs from the $\alpha_1(z_1)$ of the centralized case of the preceding section. In other words, does Man 1 act differently when Man 2 has more information? In the event that they are the same, the formidable computation problem of the last section is immensely simplified.

4. Model II

In the market facing a bread salesman price is constant* and the

---

* The wholesale prices of baked goods do change of course, but only as a result of relatively infrequent and ponderous decisions at the highest levels of management. To ignore these price changes is not to deny their importance to our organization. We simply wish to limit ourselves here to a smaller more manageable problem - one which almost surely can be safely "factored out" of the grander price-decision problem.

---

quantity of fresh bread demanded at this price is one of the random variables characterizing the state of the world x. Normalizing our money measure, we can let price in this model be unity. The demand in

Market i will be written y_i and whatever advance information the salesman can learn about demand z_i . Ignoring the production cost parameters, we have then $x = (z,y)=(z_1, \ldots, z_N; y_1, \ldots, y_N)$. Just as in Model I, the salesman's action $a_i = \alpha_i(\eta_i(z,y))$ is the amount of fresh bread he supplies his market. Lacking precise knowledge of the demand y_i that will materialize, he will sometimes supply too much and sometimes too little.

Throughout this section we shall assume that bread delivered fresh to the market in the morning remains saleable as "fresh" bread only to the end of the day -- i.e., its "shelf-life" is one day.* Any excess

* "Shelf-life" is of course a policy decision, not a technological parameter. Like price, we suppose it to be given.

e_i of the amount a_i supplied by the salesman over the amount y_i demanded by consumers must be picked up by the salesman at the beginning of the next day and returned to the plant where it is sold in an $(N+1)^{th}$ market at a price $1-r$, with $1-c_a < r < 1$. Demand for "stale" bread in this special market will be supposed infinitely elastic.

If Salesman i "undersupplies" his market the reaction of disappointed consumers will be supposed to bring about a diminution of the firm's future profits the present discounted value of which is directly proportional to the deficiency d_i of supply. Let such loss per unit of deficiency be the same in all markets and denote it q .

To simplify the analysis, both r and q will be regarded as given; in fact, they vary* from time to time and from place to place,

---

* Variation in r would arise if the assumption of elasticity in the demand for stale bread were dropped. Failure to meet the demand of a particularly good customer, or of an ordinary customer on an occasion particularly important to him, could give rise to changes in q .

---

but probably less than the y_i .

In these terms total revenue in Market i is

$$(25) \qquad a_i - re_i - qd_i$$

$$\text{where} \quad e_i = \max(a_i - y_i, 0)$$

$$d_i = \max(y_i - a_i, 0) .$$

In Figure 4 total revenue curves are shown for two different values of y_i .

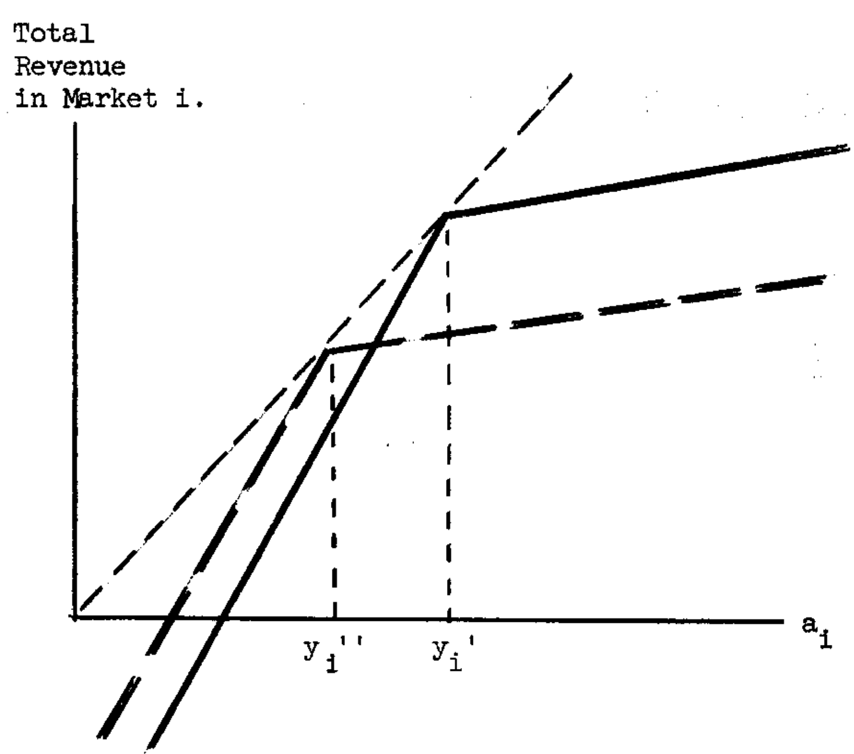

Figure 4. Total Revenue in Market i as Function of Amount Supplied

From a formal standpoint the present problem now appears more closely analogous to the Model I problem than one might have suspected at first. Once again the problem is to determine how each salesman should behave in the face of a shifting total revenue curve. To be sure, the curves are somewhat more complicated in the present case than in Model I, but the analysis based on them can proceed in the same fashion. For purposes of the analysis, the interpretation given these revenue curves is unimportant: whether we say that when the amount supplied exceeds y_i marginal revenue is lower because the **number of stales** is higher or because the "price" must be lower to clear the market makes no difference. The only new feature of general interest introduced by Model II is the dependence of marginal revenue on the amount supplied to the market; in Model I marginal revenue was unaffected by the supply decision.

4.1. Centralization

Let the information structure be specified by $\eta_i(x) = z$. The problem then is to select non-negative $\alpha_i(z)$, $d_i(z,y_i)$, $e_i(z,y_i)$ for each i , and $b(z)$ so as to maximize

$$(26) \quad \sum_z \sum_y \left\{ \sum_i [(1-c_a)\alpha_i(z) - re_i(z,y_i) - qd_i(z,y_i)] - (c_b - c_a)b(z) \right\} p(z,y)$$

subject to the constraints

$$(27) \quad d_i(z,y_i) \geq y_i - \alpha_i(z) \quad , \quad (i=1, \ldots, N)$$

$$(28) \quad e_i(z,y_i) \geq \alpha_i(z) - y_i \quad , \quad (i=1, \ldots, N)$$

$$(29) \quad b(z) \geq \sum_i \alpha_i(z) - w \quad .$$

Let $\mu_i(z,y_i)$, $\gamma_i(z,y_i)$ and $\lambda(z)$ be the Lagrange multipliers respectively associated with (27), (28), and (29) . The partial derivatives of the Lagrangean maximand with respect to $\alpha_i(z)$, $d_i(z,y_i)$, $e_i(z,y_i)$, and $b(z)$ are respectively

(30) $$(1-c_a)p(z) - \lambda(z) + \sum_{y_i} [\mu_i(z,y_i)-\gamma_i(z,y_i)] = 0 \text{ , etc.}$$

(31) $$\mu_i(z,y_i) - qp(z,y_i) = 0 \text{ , etc.}$$

(32) $$\gamma_i(z,y_i) - rp(z,y_i) = 0 \text{ , etc.}$$

(33) $$\lambda(z) - (c_b-c_a)p(z) = 0 \text{ , etc.}$$

Let us first look at (27) and (28) in relation to (31) and (32) for given i . If $y_i > \alpha_i(z)$ then $d_i(z,y_i) > 0$, so equality must hold in (31) and strict inequality in (28), in turn implying $\gamma_i(z,y_i) = 0$. If $y_i < \alpha_i(z)$ the same argument yields equality in (32) and $\mu_i(z,y_i) = 0$. If $y_i = \alpha_i(z)$ then <u>both</u> μ_i and γ_i can be positive.

With these results the third term in (30) can be written

$$\left\{ qS_i - r\ (1-S_i) \right\}$$

where S_i , or more strictly $S_i[\alpha_i(z)]$, which we shall call the "(conditional) probability of sell-out," is defined by

(34) $$\text{Prob}\ \left\{ y_i < \alpha_i(z) | z \right\} \leq 1-S_i[\alpha_i(z)] \leq \text{Prob}\ \left\{ y_i \leq \alpha_i(z) | z \right\} .$$

We can now rewrite (30) as

(35) $$1-c_a - \frac{\lambda(z)}{p(z)} + qS_i - r(1-S_i) = 0 \text{ , etc.}$$

or, in full,

(36) $$\frac{1+q - c_a - \frac{\lambda(z)}{p(z)}}{r+q} \begin{Bmatrix} \leq \\ = \end{Bmatrix} 1 - S_i \quad \text{if} \quad \alpha_i(z) \begin{Bmatrix} = \\ > \end{Bmatrix} 0 .$$

With knowledge of his own demand distribution a salesman can easily choose his $\alpha_i(z)$ so as to put S_i at any prescribed level between zero and one. Suppose the $\alpha_i(z)$ are tentatively set at values which make

$$(37) \qquad 1-S_i = \frac{1+q-c_a}{r+q}$$

For all z such that $\Sigma\, \alpha_i(z) < w$, these tentative $\alpha_i(z)$ are optimal, because in these cases $\lambda(z) = 0$ and (36) is the same as (37).

Now, for the remaining z, set the $\alpha_i(z)$ tentatively at values which make

$$(38) \qquad 1-S_i = \frac{1+q-c_b}{r+q}$$

These values will be correct for all z for which $\Sigma\, \alpha_i(z) > w$, since now (36) is the same as (38).

For the still remaining z, the $\alpha_i(z)$ must be selected so as to make $\Sigma\alpha = w$ and $S_i = S_j$, at a level intermediate between (37) and (38).

The solution is easy to see in a graph for the case $N=2$. To simplify the picture let $z_1 = E(y_i|z)$. The solid lines divide the z space into the three regions just described. In a practical case of centralization one might wish to compute a whole family of contours of the function $\lambda(z)$; the dotted line in Figure 5 is an intermediate member of this family. In the case when the z_i are not correlated, all any individual salesman need know in order to act optimally is the value of $\lambda(z)$; the particular z that gave rise to this $\lambda(z)$ is of no importance. This suggests that from the computational point of view the optimal set of actions given a z, might best be found, not by tabulating $\lambda(z)$ and

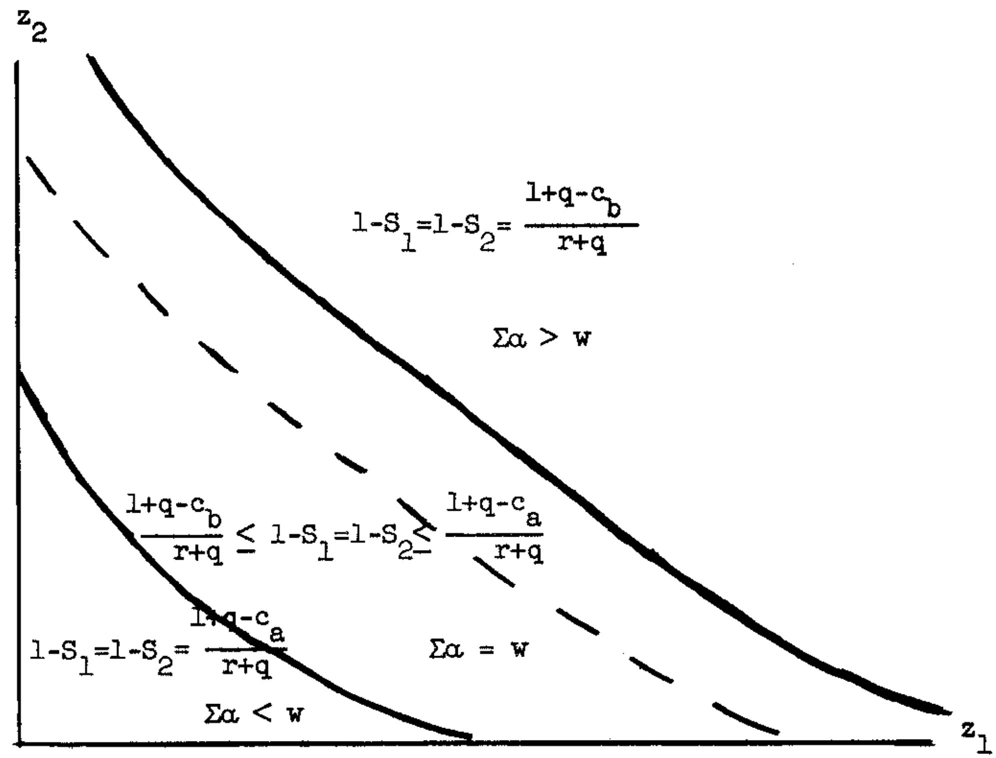

Figure 5

informing all members of z, but rather by playing a "market game" in which individuals repeatedly submit tentative orders on the basis of thier own z_i and a tentative value of λ. A "custodian" (of overtime labor) then revises λ in the light of $\Sigma\alpha$, the procedure continuing until a "centralized" solution satisfying (36) is found. Even when some correlation between markets exists, this procedure might be a reasonably good one in practice.

4.2. Decentralization

The information structure is now $\eta_i(x) = z_i$. The condition analogous to (36), namely,

$$(39) \qquad \frac{1+q-c_a - \frac{1}{p(z_i)} \sum\limits_{\substack{z_j \\ j \neq i}} \lambda(z)}{r+q} \begin{Bmatrix} \leq \\ = \end{Bmatrix} 1-S_i[\alpha_i(z_i)] \quad \text{if} \quad \alpha_i(z_i) \begin{Bmatrix} = \\ > \end{Bmatrix} 0$$

follows by exactly the same argument. In looking for a method of computation we can, as in the decentralization case of Model I, fall back on the standard techniques available for linear programming problems. But just as in Model I, this alternative cannot be regarded as much more than a last resort -- the number of variables is so high. Again we wish to make use of whatever special characteristics the problem may have that were not incorporated in the mathematical formulation.

Results -- or rather weak computational hints -- analogous to those obtained for Model I are slightly harder to derive and, once derived, are probably of less direct usefulness. Since they are suggestive, however, they follow.

Letting p and p' denote the conditional probabilities that $\Sigma\alpha < w$ and $> w$, respectively, for a given z_i, the left side of (39) can be written*

$$(40) \qquad p\left(\frac{1+q-c_a}{r+q}\right) + p'\left(\frac{1+q-c_b}{r+q}\right) .$$

* Recall (35), which allows us to ignore occasions where $\lambda(z)$ has an "intermediate" value.

An optimal action function is therefore seen to be one which results from setting $1-S_i[\alpha_i(z_i)]$ equal to a linear interpolation between the extreme values of (39), the weights being the conditional probabilities p and p'.

Designate the two hypothetical action functions that would be optimal when $w=0$ and $w=\infty$ by α_i^0 and α_i^∞, respectively (if $w=0$ then c_b always prevails; if $w=\infty$, c_a always prevails). For any given z_i, the value $\alpha_i(z_i)$ of the function we seek is a non-linear interpolation, via $S_i[\alpha_i(z_i)]$, between $\alpha_i^0(z_i)$ and $\alpha_i^\infty(z_i)$.

Figure 6, a four quadrant diagram analogous to Figure 2, shows the functions referred to for the case $N=2$, $z_i=E(y_i|z_i)$.

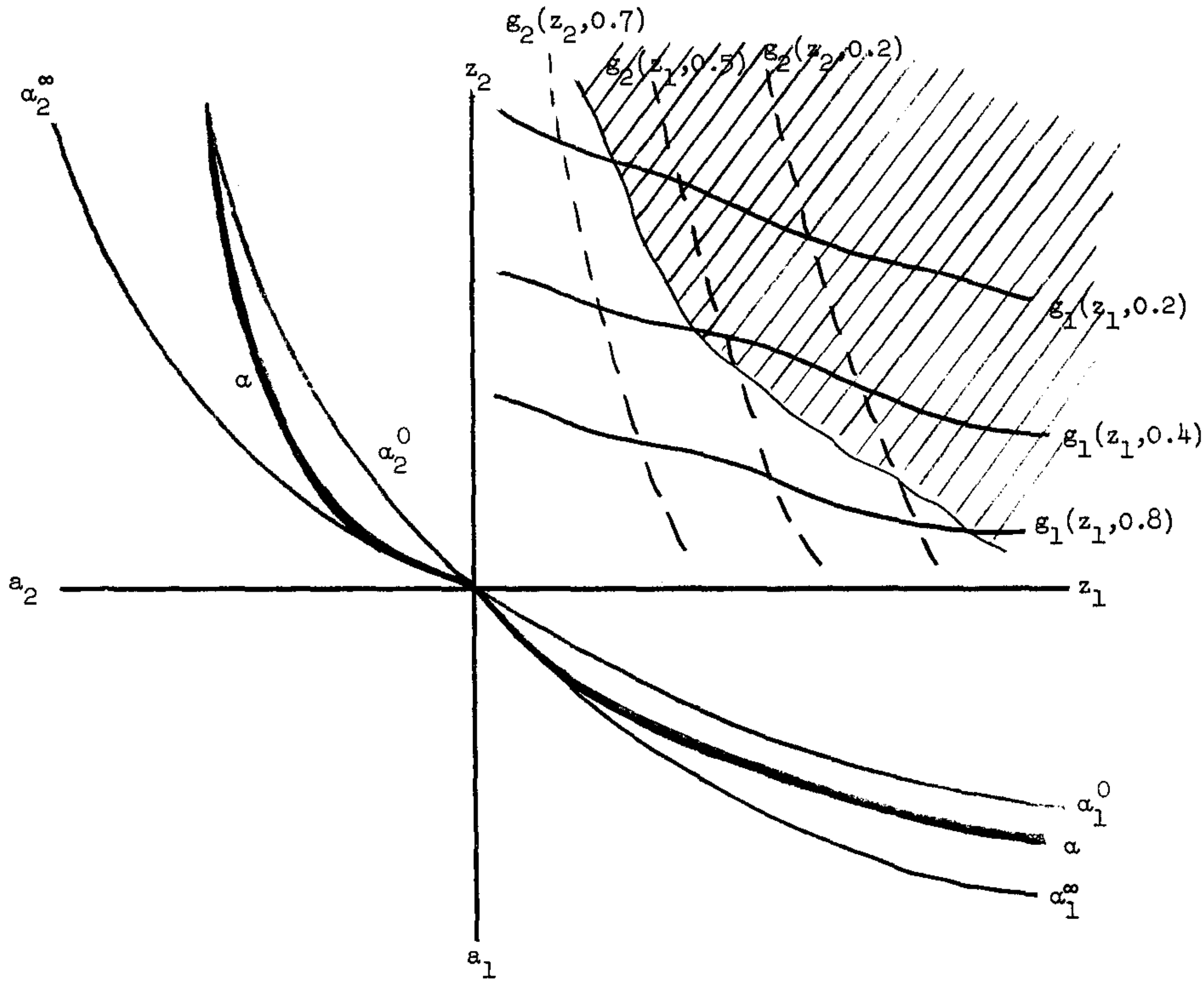

Figure 6

So long as z_1 and z_2 are not highly positively correlated, then both α_1 and α_2 are probably non-decreasing and the positive $\lambda(z)$ will occur in the Northeast region of the z space. Define a family of g_1 functions as follows:

$$\sum_{z_2 = g_1(z_1,t_1)}^{\infty} p(z_1,z_2) \doteq t_1\, p(z_1) \qquad 0 \leq t_1 \leq 1$$

and a similar family of g_2 functions. Just as in Model I, much will depend on how well behaved these g functions are. If they are monotonic as in Figure 6, and as would be expected with a "reasonable" probability distribution in which the positive correlation is not too high, then the following procedure will yield a solution without too much difficulty.

Pick a z. Observe the t_1 and t_2 associated with the g curves through z. Select tentative $\alpha_i(z_i)$ so as to make

$$1-S_i[\alpha_i(z_i)] = \frac{1+q-c_a-t_i(c_b-c_a)}{r+q} \quad .$$

If $\Sigma\alpha > w$, mark z with a "+".

If $\Sigma\alpha < w$, mark z with a "0". Next, if z is "+" then so is any $z' \geq z$ (in the vector sense); if z is "0" then so is $z' \leq z$. When either plus or zero has been associated with all z (except those dividing the two regions, which we will ignore), we set $\lambda(z) = (c_b-c_a)p(z)$ if z is "+", and $\lambda(z) = 0$ if z is "0". Set the final values of

$\alpha_i(z_i)$ by the values of t_i along the border of the "+" region (shaded in Figure 5).

That the solution so constructed satisfies (39) is easily verified.

Notice that by no means all the z need be investigated, and hence not all the g curves need be constructed. It is here that the saving of effort over a straightforward linear programming calculation is most evident. The qualifications mentioned earlier about extensions to $N > 2$ still hold, of course. Furthermore, while adapting the procedure (e.g. by making one of the g functions cumulate probability from the bottom of the z_i distributions instead of the top) to high positive correlation may be more difficult than in Model I, the need to do so arises less often because the level of positive correlation at which trouble is encountered is higher than in Model I.

5. Model III

In most wholesale bakeries the shelf-life of the leading product is set at two days rather than -- as in Model II -- one day. In this section we attempt to investigate the consequences of changing Model II in this one single respect.

No longer is one day's profit independent of another day's profit; the supply actions taken on Tuesday affect not only Tuesday's profit but also Wednesday's since some of the bread supplied fresh Tuesday morning may not be taken by consumers on Tuesday. This carryover, not yet "stale", affects the probability of sell-out on Wednesday, hence the optimum amount to be supplied fresh on that day, hence the carry-over into Thursday and the profit on that day, etc.

In order that the profit function be bounded, we must introduce either a discount factor or a beginning and end to the period considered. Since Sundays are the salesman's day off, they provide a natural break in the interrelatedness of profits on successive days, so the second alternative will be chosen here. Wherever necessary, variables will be labeled with superscript t to denote day of the week ($t=1, \ldots, T$).

On any one day three classes of bread will be distinguished: fresh, day-old, and stale, corresponding respectively to ages less than one day, between one and two, and greater than two. We shall assume that a "LIFO" rule of consumption prevails in every store: that the consumer always takes the freshest loaf of bread from the store shelf. Of course, it is true in fact that some firms attempt to force FIFO consumption either by rack arrays that invite the consumer to make what is for him the worse choice, or by actually withholding the fresh product until yesterday's carryover has disappeared. It is not known how widespread - or successful - the first policy is; the second policy requires an intimate cooperation between grocer and bread salesman that is not typical in stores open to competing suppliers. Here, in any case, we suppose the LIFO consumption rule is imposed on the firm. Some further remarks on the LIFO - FIFO comparison will be found below.

As in Model II, q will be the stock-out penalty; that is, when a demand can neither be filled from the stock of fresh or the stock of day-old bread, a loss of q is incurred. When a demand cannot be filled with a fresh loaf, but is filled with a day-old one, a loss of q' is incurred. Clearly, if the policy (called the "48-hour code") of leaving day-old bread on the store shelves is to be a profitable one, q' must not be too high. At first sight one might suppose that for the "48-hour code" to be more profitable than the "24-hour code," a day-old loaf supplied to a disappointed consumer in Model II must reduce the "good-will" loss suffered by the firm; that is, that $q' < q$. This is easily seen to be false. Let the optimal ordering policy of Model II be fixed. Now suppose one unsold loaf from Day t is carried over to Day $(t + 1)$ instead of being returned for sale in the stales market. If this loaf is not sold on Day $(t + 1)$ the firm loses r dollars on that loaf, just as it would have if the loaf had not been carried over. If the loaf is sold on Day $(t + 1)$ a loss of q' dollars occurs, but without this carryover a stockout would have occurred and a salvage loss on Day t. A sufficient condition, therefore, for the firm's preferring the "48-hour code" is that $q' < r + q$.

The definitions of "excess" and "deficiency" must be extended to day-old bread. In accordance with the LIFO assumption, demand will be first applied to the stock of fresh bread to determine e_i^t and d_i^t, the excess and deficiency of fresh bread, and d_i^t will be the second-stage "demand" that is applied to the stock of day-old bread to determine $\bar{e}_i^t$ and $\bar{d}_i^t$, the excess and deficiency of "day-old."

The state-of-the-world vector now becomes a TN-tuple of the z_i^t and the y_i^t. As the week of length T days progresses a salesman learns more and more about the true x: at the beginning of Day t he has observed $z_i^1, \ldots, z_i^t$ and $e_i^1, \ldots, e_i^{t-1}$ in his own market and, in the centralized case, the same variables in the other markets. If we follow the formulation of Model II the e_i^t are pseudo action functions, optimum values of which must be determined in the profit maximizing problem. Yet here we find they also play the rôle of information variables given by x and the - supposedly fixed - information structure η. This conflict is resolved by supposing (contrary to fact) that in making his decision for Day t Salesman i knows yesterday's demand y_i^{t-1} in place of e_i^{t-1}. We then assume that the pair (z_i^t, e_i^{t-1}) is a sufficient statistic for (z_i^t, y_i^{t-1}) relative to y_i^t, thus ensuring that no use can be made of the extra information in y_i^{t-1}. We can still regard the e_i^t, therefore, as pseudo action functions without violating the assumption of fixed η.

A convenient formalism is to regard Salesman i on Day t as a team member distinct from Salesman i on Day (t+1). For each "member," then, there is an information function η_i^t. In order to make our representation of these functions less cumbersome we shall use the notation

$$Y_i^t = (y_i^1, \cdot\cdot, y_i^t)$$

and

$$Y^t = (Y_1^t, \cdot\cdot, Y_N^t)$$

with Z_i^t and Z^t defined similarly.

In the decentralized case, then, where only local information is known, we have

(40) $$\eta_i^{\,t}\,(x) = (Y_i^{\,t-1},\ Z_i^{\,t})\ .$$

In the centralized or full-information case we have

(41) $$\eta_i^{\,t}\,(x) = (Y^{t-1},\ Z^t)$$

as our information structure.

5.1 Centralization

A useful abbreviation is to omit the information arguments from the action functions and pseudo action functions. For reference we write them out this once in full:*

(42)
$$\begin{aligned}
&\alpha_i^{\,t}\,(Y^{t-1},\ Z^t) \\
&e_i^{\,t}\,(Y^t,\ Z^t) \\
&d_i^{\,t}\,(Y^t,\ Z^t) \qquad\qquad (t=1,\ \ldots,\ T) \\
&\bar{e}_i^{\,t+1}\,(Y^{t+1},\ Z^{t+1}) \\
&\bar{d}_i^{\,t}\,(Y^t,\ Z^t) \\
&b^t\,(Y^{t-1},\ Z^t)
\end{aligned}$$

---

* From the way the four excess and deficiency variables are used in constraints (44)-(47) below, it will be seen that we have based these variables on a greater amount of information than is strictly necessary. Thus $e_i^{\,t}(Y^t,\ Z^t)$ might just as well have been replaced by $e_i^{\,t}\,(Y^{t-1},\ y_i^{\,t},\ Z^t)$, etc. As we shall see, the solution turns out to be the same in either case.

In what follows it must be remembered that, e.g., $\alpha_i^{\,t}$ is not a single-valued variable and neither is $\frac{\partial L}{\partial \alpha_i^{\,t}}$, etc.

The most natural way to begin is to follow the procedure of Model II. Thus we have

<u>Problem A</u>: Choose non-negative functions $\alpha_i^{\,t}$, $e_i^{\,t}$, $d_i^{\,t}$, $\bar{e}_i^{\,t+1}$; $\bar{d}_i^{\,t}$, and b^t for $t=1 , \cdots, T$ and $i=1 , \cdots, N$ so as to maximize

$$(43) \quad \mathcal{E} \sum_{t=1}^{T} \left\{ \sum_{i=1}^{N} [(1-c_a)\, \alpha_i^{\,t} - r\bar{e}_i^{\,t+1} - q\bar{d}_i^{\,t} - q'(d_i^{\,t} - \bar{d}_i^{\,t})] - (c_b - c_a)\, b^t \right\}$$

subject to the constraints:

$$(44) \qquad e_i^{\,t} \geq \alpha_i^{\,t} - y_i^{\,t}$$

$$(45) \qquad d_i^{\,t} \geq y_i^{\,t} - \alpha_i^{\,t}$$

$$(46) \qquad \bar{e}_i^{\,t+1} \geq e_i^{\,t} - d_i^{\,t+1}$$

$$(47) \qquad \bar{d}_i^{\,t} \geq d_i^{\,t} - e_i^{\,t-1}$$

$$(48) \qquad b^t \geq \sum_{i=1}^{N} \alpha_i^{\,t} - w$$

where $e^o = d^{t+1} = 0$.

Just as in Model II, of course, Problem A is a correct representation of the real problem we wish to solve only if in the solution the pseudo action variables take on values consistent with the interpretations of them that determine their rôles in (43). Thus in (43) $\bar{e}_i^{\,t}$ represents the excess of day-old bread on Day t . If in the solution non-zero $\bar{e}_i^t$ were found for which strict inequality held in (46), then (43) would be incorrect; extra day-old bread would have been generated mathematically. The same is true for each of the constraints (44)-(48): we want equality to hold when the right-hand side is non-negative; we want the variable on the left to be equal to zero when the right-hand side is negative.

How can one be sure that these conditions - not formally incorporated in the problem - are met? Rather than attempt to answer this question at this stage we might simply find the solution to Problem A, that is to say, characterize the solution abstractly as in Model II, and then check to see whether the desired consistency requirments are met or not. Even without carrying through this procedure, one of the results is clear: certain restrictions on the parameters are necessary for this consistency. If $q' > q > 0$, (43) is clearly unbounded, for the coefficient of $\bar{d}_i^t$ is positive and this variable is not constrained from above by (44)-(48). Yet we know that with a bounded demand distribution, profit is finite, so $q' > q > 0$ implies inconsistency. Hence we are led immediately to the very severe restriction $q' < q$. It will be recalled that $q' < r+q$ was sufficient for the "48-hour code" to be an improvement over the "24-hour code" under a LIFO consumption rule. According to our present result therefore, when $q < q' < r+q$, that is, when the parameters are such that a LIFO regime would be most advantageous to the firm, Problem A is not a correct representation of the real problem. This argument suggests - but of course does not prove - that in these circumstances the <u>real</u> problem involves maximizing a function which is not concave, and that therefore the usual price theorems do not apply.

This result is implausible. To be sure, when q' is extremely low, one would expect a FIFO rule of consumption to be more profitable for the firm than a LIFO rule. In such circumstances the optimal stocking policy of a firm faced with a LIFO market might reflect an attempt to force near-FIFO consumption by placing all the fresh bread in a few stores so as to increase sales of day-old bread in the other stores. The mathematical problem with such a solution clearly involves a maximand which is not concave. In the case at hand, however, we are concerned with high values of q'.

We are led therefore to suspect the Problem A formulation as the source of difficulty. Let us attempt a reformulation.

We first observe that nothing real is lost by using

$$e_i^t = \alpha_i^t - y_i^t + d_i^t \tag{49}$$

to eliminate e_i^t from Problem A, for if d_i^t possesses the desired consistency property, then so will the right-hand side of (49). By the same argument, to eliminate $\bar{d}_i^t$, we can use

$$\bar{d}_i^t = d_i^t - e_i^{t-1} + \bar{e}_i^t ,$$

which after substituting for e_i^{t-1}, becomes

$$\bar{d}_i^t = d_i^t - \alpha_i^{t-1} + y_i^{t-1} - d_i^{t-1} + \bar{e}_i^t \tag{50}$$

where now we must stipulate that

$$\alpha^o \doteq y_i^o \doteq d_i^o \doteq \bar{e}_i^1 \doteq d^{T+1} \doteq 0 \tag{51}$$

When (49) is substituted in the other constraint of Problem A in which it occurs, namely (46), we have

(52) $$\bar{e}_i^{\,t+1} \geq \alpha_i^{\,t} - y_i^{\,t} + d_i^{\,t} - d_i^{\,t+1} .$$

But notice that $d_i^{\,t}$ in this constraint serves no purpose; if the consistency property obtains either $d_i^{\,t} = 0$ or $d_i^{\,t} = y_i^{\,t} - \alpha_i^{\,t}$. In the last case the right side of (52) is non-positive with or without $d_i^{\,t}$. The relevant constraint can therefore be written

(53) $$\bar{e}_i^{\,t+1} \geq \alpha_i^{\,t} - y_i^{\,t} - d_i^{\,t+1} .$$

We next observe that

(54) $$\sum_{t=1}^{T} \bar{d}_i^{\,t} = \sum_{t=1}^{T} (d_i^{\,t} - \alpha_i^{\,t-1} + y_i^{\,t-1} - d_i^{\,t-1} + \bar{e}_i^{\,t})$$

$$= \sum_{t=1}^{T} (y_i^{\,t} - \alpha_i^{\,t} + \bar{e}_i^{\,t+1}) - y^T + \alpha^T - \bar{e}_i^{\,T+1} + \sum_{t=1}^{T} d_i^{\,t} - \sum_{t=1}^{T-1} d_i^{\,t}$$

$$= \sum_{t=1}^{T} (y_i^{\,t} - \alpha_i^{\,t} + \bar{e}^{t+1})$$

by using (50) in the first line, (51) in the second line, and (45) and (53) at $t=T$ in the third line.

We can now state Problem B, whose payoff function is derived from that of Problem A by using (54) to eliminate $\bar{d}_i^{\,t}$.

Problem B: Choose non-negative functions

$\alpha_i^{\,t}$, $d_i^{\,t}$, $\bar{e}_i^{\,t+1}$, and b^t for $t=1 , \cdot\cdot, T$ and $i=1 , \cdot\cdot, N$

so as to maximize

(55) $$\mathcal{E} \sum_{t=1}^{T} \left\{ \sum_{i=1}^{N} \left[(1-c_a + q - q') \alpha_i^t - (r + q - q') \bar{e}_i^{t+1} - q'd_i^t - (q - q') y_i^t \right] - (c_b - c_a) b^t \right\}$$

subject to the constraints [stated with the help of(51)]

(45) $$d_i^t \geq y_i^t - \alpha_i^t$$

(53) $$\bar{e}_i^{t+1} \geq \alpha_i^t - y_i^t - d_i^{t+1} .$$

(48) $$b^t \geq \sum_{i=1}^{N} \alpha_i^t - w$$

With constraints (45), (53), and (48) we shall associate Lagrange multipliers $\mu_i^t(Y^t, Z^t)$, $\bar{\gamma}_i^{t+1}(Y^{t+1}, Z^{t+1})$, and $\lambda^t(Y^{t-1}, Z^t)$ respectively.

The partial derivatives of the Lagrangean function L with respect to the action variables are

(56) $$\frac{\partial L}{\partial \alpha_i^t} = (1 - c_a + q - q')\, p(Y^{t-1}, Z^t) + \Sigma_{y^t} \mu_i^t - \Sigma_{y^t} \Sigma_{z^{t+1}} \Sigma_{y^{t+1}} \bar{\gamma}_i^{t+1} - \lambda^t$$

(57) $$\frac{\partial L}{\partial d_i^t} = -q'p(Y^t, Z^t) + \mu_i^t + \bar{\gamma}_i^t$$

(58) $$\frac{\partial L}{\partial \bar{e}_i^{t+1}} = -(r + q - q')\, p(Y^{t+1}, Z^{t+1}) + \bar{\gamma}_i^{t+1}$$

(59) $$\frac{\partial L}{\partial b^t} = -(c_b - c_a)\, p(Y^{t-1}, Z^t) + \lambda^t$$

To restate the necessary and sufficient conditions on the solution: there must exist non-negative μ_i^t, $\bar{\gamma}_i^{t+1}$, and λ^t such that

i) each of the derivatives (56) - (59) is respectively equal to or less than zero according as the corresponding action variable is or is not strictly positive, <u>and</u>

ii) equality must hold in any constraint whose associated Lagrange multiplier is strictly positive.

For any set of α_i^t values we might choose, there is a natural way to construct corresponding d_i^t, $\bar{e}_i^{t+1}$, and b^t; namely, give these pseudo action variables values which make them "consistent" with the given α_i^t. Now consider that set of action functions α_i^t determined by the following rule expressed in terms of the corresponding naturally defined pseudo action variables:

(60)
$$
\begin{aligned}
&-(q-q')-q' \;\text{Prob}\left\{d_i^t>0 \mid Y^{t-1}, Z^t\right\} \\
&\quad + (r+q-q')\;\text{Prob}\left\{d_i^t>0,\ \bar{e}_i^t>0 \mid Y^{t-1}, Z^t\right\} \\
&\quad + (r+q-q')\;\text{Prob}\left\{\bar{e}_i^{t+1}>0 \mid Y^{t-1}, Z^t\right\} \\
&\quad\quad \left\{\begin{matrix} = \\ \geq \end{matrix}\right\} 1 - c_a - \frac{\lambda^t}{p(Y^{t-1}, Z^t)} \qquad \text{if } \alpha_i^t \left\{\begin{matrix} > \\ = \end{matrix}\right\} 0
\end{aligned}
$$

Next,* if $b^t=0$ set $\lambda^t=0$; if $b^t>0$ set λ^t at that value that yields

---

* For brevity we have sacrificed some precision in the statement of (60) and in the following construction procedure. Strictly speaking, in place of

$\Pr\left\{d_i^t > 0\right\}$ we should have some number $\mathcal{E}$ with

$\Pr\left\{d_i^t > 0\right\} \leq \mathcal{E} \leq \Pr\left\{d_i^t \geq 0\right\}$, etc., and in place of (footnote cont'd) "if $b^t = 0$ set $\lambda^t = 0$" we should say "if $\sum_i \alpha_i^t < w$ set $\lambda^t = 0$; if $\Sigma\alpha_i^t = w$ set λ^t at any positive value not violating (59)," etc. All the statements that follow can be made precise with none other than expositional trouble. Even in the rough form in which it stands (60) could be used in practice without difficulty so long as changes of $\pm$ 1 in the variables had small effects.

---

equality in (59). Set $\bar{\gamma}_i^{t+1}$ and d_i^t in analogous fashion, in both cases using again the naturally defined pseudo action variables of (60). Does this resulting set of values constitute a solution?

Conditions (i), (ii) and the non-negativity of the multipliers are satisfied if

(61) $$c_a \leq c_b$$

(62) $$0 \leq 1/2(r + q) \leq q' \leq r + q$$

and

(63) $$1 - c_a \leq r .$$

The first restriction is derived from (59) and the second restriction from (57) and (58). Restriction (63) is necessary in order that it be always possible to find non-negative α_i^t satisfying rule (60), which is nothing more than (56) with (57)-(59) substituted. This derivation follows immediately upon observing that by choosing larger and larger values of α_i^t and α_i^{t+1} we can bring Prob $\left\{d_i^t > 0 \mid Y^{t-1}, Z^t\right\}$ arbitrarily close to zero, and Prob $\left\{\bar{e}_i^{t+1} > 0 \mid Y^{t-1}, Z^t\right\}$ arbitrarily close to one.

The first and last of these parameter restrictions have been assumed throughout and are therefore of no interest. The restriction (62) on q' is new, interesting, and not obvious. It assures us that the present analysis is applicable over a broad range of q', and it suggests the possibility that $1/2(r + q)$ marks the border between that region of the q' interval where a LIFO consumption rule is more advantageous to the firm than a FIFO rule and that region where the opposite is true.

To sum up the argument so far: if the parameters satisfy restrictions (61)-(63), then a finite solution to Problem B exists*, Rule (60) for setting

---

* Here we make use of the Kuhn-Tucker Theorem's assertion that Lagrange multipliers (with all the customary properties) exist if and only if the maximum sought is finite. By construction we have shown that the multipliers do in fact exist if (61)-(63) hold.

---

the α_i^t yields this solution, and the resulting values of the pseudo action variables are consistent.

In interpreting this result and the use of the rule it is of course quite important to realize that (60) is not an algorithm. Even in the simpler Model II no general algorithm was found (although suggestions were put forward). Rule (60) simply facilitates the salesman's recognition of the optimal ordering decision, and allows us to see what determines this decision.

The meaning of Rule (60) for setting the α_i^t is made clearer by two rephrasings, the derivations and statements of which are most easily carried out by means of a graph. Figure 7 shows the (y_i^t, y_i^{t+1}) space. The relevant joint probability distribution over this space is the conditional distribution given (Y^{t-1}, Z^t). The heavy horizontal lines, determined by α_i^t and e_i^{t-1},** divide the space into the regions $d_i^t > 0$ and $\bar{e}_i^t = 0$, $d_i^t > 0$ and $\bar{e}_i^t > 0$, and $d_i^t = 0$. The hyphenated wavy line divides the

---

** Recall (49).

---

last of these regions into two subregions, $\bar{e}_i^{t+1} > 0$ and $\bar{e}_i^{t+1} = 0$. Since $\bar{e}_i^{t+1}$ depends in part on α_i^{t+1}, which in turn is a function of Z^{t+1}, a correct subdivision of the $d_i^t = 0$ region requires more dimensions than the two of Figure 7 - hence our attempt to give this last borderline a tentative appearance. For present purposes no harm will come from this convenient graphical fiction.

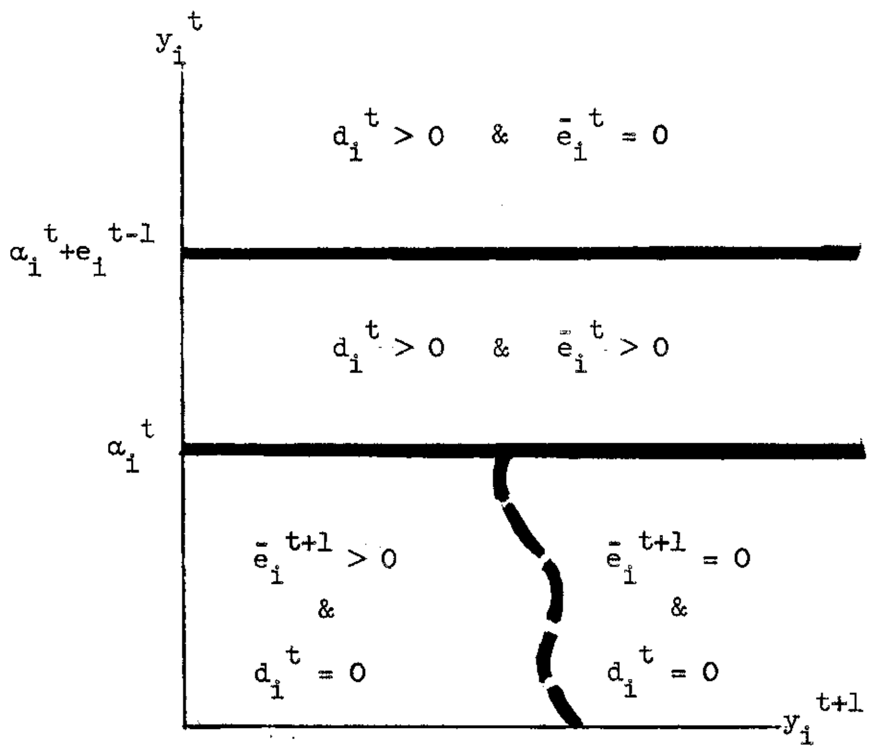

Figure 7

With an obvious shorthand, Rule (60) can now be paraphrased:*

---

* (64) is of course not a complete statement of Rule (60), but rather a mnemonic. Strict equality must hold if $\alpha_1^t > 0$, etc.; c stands for c_a or c_b or some intermediate value depending on b^t, etc.

---

(64)

$-q'$	
$-(2q'-r-q)$	
$r+q-q'$	0

$-(q-q') \geq 1-c$

where the terms in the box are the coefficients of the probabilities of the corresponding regions of Figure 7. Since the sum of the regional probabilities is unity we have the equivalent statement.

(65) $1-c+$

q	
$q'-r$	
$-r$	$q-q'$

≤ 0 .

The value of (65) is that it makes the optimal ordering rule very plausible. In this form the rule clearly says: order that amount which makes expected marginal profit equal to zero; if the amount ordered is not the optimal amount, expected marginal profit will be negative. In the case when realized demand exceeds the total amount stocked, the addition of one more fresh loaf reduces the number of depletions by one, and a loss of q is averted. When realized demand exceeds the fresh stock but not the total stock, increasing today's order by one causes (through the LIFO rule) a day-old loaf to be returned to the shelf in favor of a fresh one; a loss of q' for selling the inferior product is prevented, but since the rejected loaf cannot be sold in this market, a loss of r is incurred. When both today's demand falls short of even the fresh stock and tomorrow's demand is not high enough to cause a sell-out, then an additional fresh loaf placed in the store today is not sold either today <u>or</u> tomorrow, so a loss of r is brought about. When both today's demand is less than the fresh stock and a sell-out <u>does</u> occur the following day, then an additional fresh loaf put in today is sold tomorrow to an otherwise disappointed shopper; the loss of q is avoided, but since at the time of sale the product is not fresh a loss of q' is incurred.

Still another form of Rule (60) is useful because it has a more "workable" look. From (65) we easily obtain

$$\begin{array}{|c|c|}\hline \multicolumn{2}{|c|}{-\ (r+q-q')} \\ \hline \multicolumn{2}{|c|}{0} \\ \hline q' & 2q'-r-q \\ \hline \end{array} \geq 1 - c - r + q'$$

which can be written

$$\begin{array}{|c|}\hline -\ (r+q-q') \\ \hline 0 \\ \hline q' - (r+q-q')S^{*} \\ \hline \end{array} \geq 1 - c - r + q'$$

where S^{*} is the conditional probability of sell-out tomorrow given a positive carryover from today to tomorrow.* Compared to the rule of Model II which

---

* It is quite possible to express S^{*} directly in terms of the parameters and expected future marginal production costs, but the resulting expression is a clumsy one.

required the salesman to equalize probabilities of sell-out, the present relatively "workable" rule is still a complicated one, for it involves three separate probabilities at each store: the probability of sell-out, the probability of carryover, and the conditional probability of sell-out tomorrow given carryover from today.

5.2 Decentralization

Without much additional work we can use the preceding results to characterize the optimal decision rules for the care of decentralization. With information structure (40), the action variables written out in full become

$$\begin{aligned} &\alpha_i^t\ (Y_i^{t-1}, Z_i^t) \\ &d_i^t\ (Y_i^t, Z_i^t) \\ &\bar{e}_i^{t+1}\ (Y_i^{t+1}, Z_i^{t+1}) \\ &b^t\ (Y^{t-1}, Z^t) \end{aligned} \qquad (t=1, \cdot\cdot, T)$$

The argument which led to Problem B, and the parameter restrictions sufficient for concavity and "consistency" remain unchanged. The condition on the α_i^t analogous to (60) is

$$\begin{aligned} (67) \qquad & -(q-q') - q' \operatorname{Prob}\left\{d_i^t > 0 \mid Y_i^{t-1}, Z^t\right\} \\ & + (r+q-q') \operatorname{Prob}\left\{d_i^t > 0, \bar{e}_i^t > 0 \mid Y_i^{t-1}, Z^t\right\} \\ & + (r+q-q') \operatorname{Prob}\left\{\bar{e}_i^{t+1} > 0 \mid Y_i^{t-1}, Z^t\right\} \\ & \begin{Bmatrix} = \\ \geq \end{Bmatrix} 1 - c_a - \sum_{\substack{Y_j^{t-1} \\ j \neq i}} \sum_{Z_j^t} \frac{\lambda^t}{p(Y_i^{t-1}, Z_i^t)} \quad \text{if} \quad \alpha_i^t \begin{Bmatrix} > \\ = \end{Bmatrix} 0 . \end{aligned}$$

This rule looks somewhat less formidable when it is realized that the right-hand member can be written $1 - \mathcal{E}\left\{c \mid Y_i^{t-1}, Z_i^t\right\}$. If in (65) and (66) we replace c by $\mathcal{E}\left\{c \mid Y_i^{t-1}, Z_i^t\right\}$ and reinterpret the action variables, both rules apply here.

At just this point in our analysis of Models I and II, the discussion became more interesting from an organizational point of view in that we could begin to investigate the differences in the shapes of the order functions under centralization and decentralization. The troubles encountered in devising good algorithms for these simpler problems certainly carry over to Model III which, organizationally, differs from them very little. At the same time, whatever progress can be made at the level of the simpler problems is bound to be useful also in Model III. We shall not pursue the computation question farther here.

6. The Models and the Real Organization

Ideally, one of the first steps in assessing a given sales organization might be to attempt to set down the existing information structure. No very careful observation is required to see quickly that the structure in use is neither of the two extremes - "centralization" and "decentralization" - that we have been most concerned with in the models. To locate the structure in the wide range between these extremes, however, requires very close observation indeed, for, as must be the case in almost all live organizations, only part of the information flow between the "team" members is formal.

For an individual salesman, at least six sources of information can be distinguished: (i) his observations in his own market, (ii) his discussions with his fellow salesmen, (iii) his discussions with his route supervisor,

(iv) "bulletins" and other relatively formal communications from the sales manager, (v) his short end-of-the-day encounter with the order clerk, and (vi) inference from the actual order of goods supplied him at the beginning of the day. We shall discuss each of these sources in turn, describing the kind of information learned, its place (if any) in Model III, and the nature of the associated costs.

In discussing observation it is important to recall three respects in which Model III abstracted from reality: (i) the z_i^t , upon which the salesman forms his independent estimates of tomorrow's demands, were given no real interpretation; (ii) on Day t it was assumed that the salesman observes z_i^{t+1} and the carryover e_i^t from Day t to Day $(t+1)$, yet if he visits the store not oftener than once a day (the usual frequency) he will not in fact observe e_i^t until Day $(t+1)$; (iii) the model assumed one salesman for each store whereas in fact one salesman handles on the order of forty stores; (iv) the model dealt with only one product while the salesman may deal with thirty.*

---

* Until we return to this point we shall continue to talk about one product.

Upon visiting one of his stores on Day t , the salesman writes down in his "route book" the carryover e_i^{t-1} from yesterday, the number $\bar{e}_i^t$ of stales to be returned to the plant, and a tentative store order for tomorrow based on observations which are not themselves formally recorded in any regular way. Without specific knowledge of the salesman's art of estimating demand, it is perhaps best for us to regard z_i^t as simply an index to a family of demand distributions. If this index is one-dimensional, the tentative order can be regarded as an equivalent index which is recorded. In our theory, of course, this last index is ambiguous without a specification of the marginal production cost upon which it is based. A two-dimensional index is also used sometimes: in addition to recording a tentative order, the salesman also writes down "special" orders, i.e., sales of which he is relatively more sure. Thus we have near equivalents, for the salesman, of mean and variance.

The fact that many stores are handled by one salesman, and the fact that he does not observe e_i^{t-1} until Day t can be brought into Model III without drastic changes. The salesman can now be regarded as making two kinds of decisions: the total "route" order for Day t submitted at the end of Day (t-1); and the individual store order for Day t decided on the spot with full knowledge of e_i^t. Let $\alpha_I(Y^{t-2},Z^t)$ denote the route order for Salesman I $(I=1,\cdot\cdot,M)$. Then we have, for the case of centralization, the new constraint

$$(68) \qquad \alpha_I(Y^{t-2},Z^t) \geq \sum_{i\in I} \alpha_i(Y^{t-2},Z^t) \ ,$$

with multiplier $\Theta_I^t(Y^{t-1},Z^t)$.

The "b" constraint (48) is changed to

$$(69) \qquad b^t(Y^{t-2},Z^t) \geq \sum_{I=1}^{M} \alpha_I(Y^{t-2},Z^t) - w \ .$$

Instead of $\sum_i (1-c_a)\, \alpha_i^t$ in the maximand, we now have $\sum_i \alpha_i^t - c_a \sum_I \alpha_I^t$.

In other respects, Problem B remains the same. The optimal ordering rules can be developed without difficulty. Very roughly they prescribe the following behavior: With knowledge of Y^{t-1} and Z^t the salesman selects a "shadow" marginal production cost (Θ_I^t) which exhausts his fixed stock if he allocates according to (60). The total order for the route is decided upon the day before, with knowledge of Y^{t-2} and Z^{t}, so as to make the expected value of this shadow cost equal to actual cost c . The rule under decentralization is analogous.

Restricting attention here to only one product out of the many ordinarily handled by a salesman is perhaps not a serious shortcoming of the analysis. In practice, careful order control is exercised over not more than one or two of the largest selling items. Since this means that most of the bread is controlled and most of the sweet goods are not, there is reason to believe that substitutability between the controlled goods and the uncontrolled goods - both in demand and in production - is limited.

Discussions with fellow salesmen are completely informal. The information that is conveyed in this fashion could, with theoretical propriety, influence an individual salesman's estimate of the demand distributions facing him in his own market and the strength of his order given these demand distributions. It seems highly likely that the former does occur and that the latter does not. Salesmen do customarily believe in quite complicated demand correlations among different markets; whether the discussions exploit these correlations or whether they simply serve as a medium for contagions of irrational optimism and pessimism is an open question. Only if an ordering rule of the Model III type were being used in practice and the salesmen understood how marginal production cost (or $\Sigma\alpha$) entered this rule, would the discussions influence the strength of their orders, given their estimates of demand distributions. In any case, discussions among salesmen are a far-fetched way of communicating cost information, if only because the number of others that any one man can contact with a reasonable expenditure of time is very small. To avoid congestion, schedules are arranged so that relatively few men are present at the plant at any given moment.

The route supervisors, who may each direct five or ten salesmen, play no explicit role in our model. Much of their effort is devoted to matters not dealt with here - advertising, promotion, new customers, training salesmen, etc. In the usual absence of a hard and fast ordering rule, however, the supervisor has a strong influence on the ordering decisions of his men. He is the one man who examines these decisions, when sales results warrant it, in as much detail as the route book records allow. In addition, since he is customarily in contact with more salesmen than any individual salesman is, he serves a centralizing function. In terms of our model he is therefore a channel through which flows information on marginal production cost (or $\Sigma\alpha$). In practice, of course, his participation takes a much less passive form.

Both the supervisor and the sales manager communicate to the salesmen information relevant to demand estimation. Much of this is of a non-local type: a prospective change in a competitor's price, a weather prediction, a change in product design, packaging, or advertising. With more convenient access to past records, the supervisor is in a better position to predict the effect on demand of unusual circumstances such as a Monday Fourth of July, a special school holiday, or a strike in an important industry in the market served. In our models no non-local information except production costs affected the outcome, but there was nothing essential about this restriction.

The salesman's last duty of the day (Day t) is to decide upon his total route order for the next day in conference with the order clerk. Usually only those products over which some "order control" is exercised are dealt with in this short encounter. The salesman customarily submits to the clerk his tentative total route order, his total of "special orders," his total $\sum_{i \in I} \bar{e}_i^{t-1}$ of stales returned to the plant, and the total route carryover $\sum_{i \in I} e_i^{t-1}$ found on store shelves that morning. In an organization where no effort is made to coordinate ordering decisions, the process may just end here, in which case the tentative orders become final orders, the sum of which becomes the production order. In an only slightly more sophisticated organization, the order clerk may use his acquaintance with historical information to help the salesman estimate demand.*

---

* This sharing of the prediction job can take curious forms. In one plant the salesman submits not his order for fresh stock tomorrow but rather an "order" for the total stock of fresh and day-old bread he would like to see on the market tomorrow. The order clerk then estimates today's (as yet unobserved) demand on the basis of his very aggregate historical knowledge, subtracts this estimate from the total stock on the market this morning so as to derive an estimate of carryover from today to tomorrow. Finally he subtracts this estimated carryover from the salesman's desired market total to obtain a "fresh" order for the salesman for tomorrow. This precedure has interesting cross-checking aspects, but there is little reason to believe it leads to "good" orders.

Where effort is made - as in our models - to coordinate ordering decisions, the order clerk is the main instrument of coordination, for he alone sees every one of the salesmen each day and is in a position to encourage or discourage ordering. In practice there is no doubt that the order clerk does influence the salesmen's decisions, but the real extent and motivation of this influence are by no means as clear as in the idealized organization of Model III. Certainly our investigations have not reached a point that warrants our attaching descriptive value to the present theory.

The order clerk's interviews with the salesmen are usually sequential, but not necessarily strictly so. (Other routine duties may keep a salesman in the office long enough for the clerk to catch him once again.) This suggests as an approximation to the real structure, one of the form

(70) $$\eta_i^{\,t}(x) = (Y_1^{\,t-1}, \ldots, Y_i^{\,t-1};\ Z_1^{\,t}, \ldots, Z_i^{\,t})$$

where, to avoid introducing new notation, we again revert to the Model III assumption of one store to a salesman.* Probably not very much more can be

* For N=2, this was the structure of Section 3.3 .

said about the real information structure a priori.

If the quantity of product delivered to a salesman at the beginning of a day differs from the order agreed upon the night before by him and the order clerk, he is said to have suffered a "cut" or a "plus". These not-too-frequent small alterations by the order clerk are not really instances of truly centralized decisions, for it must be remembered that the order clerk possesses only very aggregate information.** This raises the interesting question of how

** In the fictional language of Model III the "cuts" and the "plusses" are information for the salesman. This interpretation is rather strained, however; our formulation is not well adapted to this problem.

much of this easily communicated aggregate information is required to enable the order clerk to make "perfect" decisions- "perfect" in the sense of being the same as those decisions that he could make in the centralized system of Model III. Do there exist simple statistics which, for the purpose at hand, are - loosely speaking - sufficient statistics for the whole set of observations of a salesman? The most obvious candidate is the pair of numbers consisting of a salesman's order given c_a and his order given c_b . This pair is not sufficient in the above sense because (i) demand relations among correlated markets cannot be exploited, and (ii) if the optimal solution is such that $\Sigma\alpha = w$, the total order for the plant can be decided, but the allocation of this total among the salesmen will require more information. Neither of these criticisms seems devastating. It would be interesting to know, therefore, whether the added gross profits of this near centralization more than offset the costs of having the salesmen carry out the extra calculations required for this double ordering.

References

[1] Jacob Marschak, "Towards an Economic Theory of Organization and Information," Chap. XIV in Decision Processes, Thrall, Davis, and Coombs, eds., New York, Wiley (1954).

[2] Jacob Marschak, "Elements for a Theory of Teams," Management Science, 1:2:127-37 (January 1955).

[3] Roy Radner, "The Linear Team: An Example of Linear Programming Under Uncertainty," Proc. of the Second Symposium in Linear Programming, Washington: National Bureau of Standards (1955).

[4] Roy Radner, "The Application of Linear Programming to Team Decision Problems," Management Science, 5:2:143-150 (January, 1959).