OPERATIONS RESEARCH, Vol. 24, No. 1, January–February 1976

# A Model for Allocating Retail Outlet Building Resources across Market Areas

GARY L. LILIEN

*Massachusetts Institute of Technology, Cambridge, Massachusetts*

AMBAR G. RAO

*New York University, New York, New York*

(Received original October 24, 1974; final, April 2, 1975)

Many factors affect retail outlet profitability, including market potential, distribution and product costs, market pricing levels, cost (and availability) of land or space, and the relation between share of outlets and share of market. This paper presents a model that was used to plan building decisions for outlets for a consumer product across time and market areas. The model has been in use for a number of years and has provided important input for budgeting and planning decisions. The implementation process for this model is also discussed. The model and its use provide an example of what we believe to be a 'successful' management science application—the characteristics of and reasons for this success are discussed.

IN A NUMBER of industries, products or services are offered to consumers through company-controlled retail outlets; each outlet offers only the products or services of the company controlling it. Examples of such industries are retail banking, gasoline, and fast foods, where retail outlets are branch banks, service stations, and franchised restaurants, respectively. Companies in such industries grow by constructing or acquiring new outlets, and one of the most important decisions faced by marketing management is the development of a plan for such construction.

The authors participated in a project to develop a systematic, model-based approach to this planning decision. The approach was to provide guidelines on how many outlets should be built in each geographical market in each of the next 5–10 years. Traditionally, each year district managers had submitted requests for construction of outlets on a number of sites that met company requirements in terms of anticipated profitability. The requests were screened and then met, subject to the availability of funds. The long-term impact of construction on company profitability was never explicitly considered. The development of a model-based approach was motivated by top management's desire to invest larger sums of money in outlet construction than it had in the past and by their recognition that the payback for such investments would occur over an extended time period. Thus the historical approach was considered inadequate.

The profitability of a given site depends on, among other factors, its sales volume. Sales volume is affected by a number of site characteristics, such as traffic flow and neighborhood population. Since, when developing long-range plans, managers rarely have a list of specific sites available, an 'average' volume figure is assumed for each potential site. Only sites that satisfy this assumption are then selected during implementation. More important for planning is the impact of the number of sites constructed on average volume per outlet, and thus, on market share. Total market demand in the product classes considered is rather inelastic—new outlets divide essentially the same 'pie.' Thus, if a very large number of outets were to be constructed in a single market, the average sales per outlet would be substantially depressed. Marketing management believed that a relation did exist between the share of outlets $s$ and the (volume) share of market $m$ enjoyed by a company, and that, other things being equal, outlets tended to have larger volumes in markets where $s$ was large than where $s$ was small. The only quantitative work on the relation between $s$ and $m$ that had been reported in the literature supported this belief. Hartung and Fisher[1] showed that for $0<s<0.2$, $dm/ds$ and $d^2m/ds^2$ were both positive. Thus other things being equal, it is preferable to build in markets where $s$ is high than where it is low. Hartung and Fisher do not consider the impact of saturation alluded to above, and their model has other more serious shortcomings, but their work formed a starting point for this analysis.

In this paper we present a model for the relation between $s$ and $m$ and then show how the relation was used to develop a model for the outlet construction decision. The output of this model is a specification of the number of outlets to be constructed each year in each market. We consider constraints on the total budget for construction and the annual availability of sites in each market. The constraints are really estimates: thus, the initial output of the model is a demand for refinement of these estimates. That is, once the model determines that $n_{it}$ outlets should be constructed in market $i$, year $t$, a search is conducted for such sites. If an adequate number cannot be found, then the constraint is revised and put into the model, and a new solution is obtained. We developed an approximate procedure for solving the model consistent with the mixture of 'hard' and 'soft' data. The procedure provides optimal solutions in most cases and always gives solutions close to optimal.

## 1. MODEL HYPOTHESES

Hartung and Fisher[1] model the sequence of purchases by a customer as a 2-state Markov chain. The states are 'purchase company brand' and 'purchase some other brand.' The probability that a customer will buy the company's brand on the $t$th occasion, given that he bought it at $t-1$, is assumed to be $k_1s$; and the probability that the customer buys the brand at $t$, given that he bought some other brand at $t-1$, is $k_2s$, where $k_1$ and $k_2$ are constants. This model implies that

$$m=k_2s/[(1-s)+(1+k_2-k_1)s].$$

The values of $k_1$ and $k_2$ are estimated from aggregate data and found to be 4.44 and 0.64, respectively. Although this model provides a good fit in the range of data available to Hartung and Fisher, it breaks down for $s$ greater than about 0.20; for $s=1/k_1$, $m=1$.

In more recent work Naert and Bultez[3] question the robustness of the Hartung-Fisher model and suggest several alternative model structures. These structures are based on empirical evidence, not on fundamental behavioral hypotheses. For example, they do not question the Markovian basis for the Hartung-Fisher model although they promise to explore it in later work.

Our initial attempts to fit the Hartung-Fisher model to the data available were unsuccessful, even in the range $0<s\leqq 0.20$; a new approach seemed desirable. The data indicated that for small values of $s$, the second derivative of the $m-s$ relation was positive (as required by the Hartung-Fisher model), but as $s$ increased the rate of growth of $m$ slowed down, very likely because of the saturation effect referred to earlier. This observation led us to hypothesize an $S$-shaped relation between $s$ and $m$. Such a relation would be consistent with the Hartung-Fisher results for small values of $s$. For larger values of $s$ the second derivative of the $m-s$ relation would be negative, leading to $m=s-1$ as an end point—see Fig. 1.

In addition to this basic $S$-shaped $m-s$ relation, we also hypothesized that in any particular market the share enjoyed by a brand would depend upon the age of its outlets as compared to competition. To illustrate this hypothesis, consider a two-brand market. Suppose that the outlets of Brand 1 have been constructed more recently. Thus, it is likely that they are better located compared to older outlets. In the industry under study, the average life of an outlet is 20 years; substantial changes in traffic patterns and neighborhoods occur during this time. Thus, we can hypothesize that if Brands 1 and 2 had the same number of outlets, Brand 1 would have a larger market share than Brand 2, because its outlets would be convenient to more people.

Combining the above hypotheses, one can specify a family of $S$-shaped curves relating market share to outlet share. The dashed curve in Fig. 1 represents a case where a brand's outlets are newer than in the case of the solid curve.

Thus, we have identified another factor impacting profitability—the age distribution of a company's outlets compared to those of competitors.

This factor reinforces the importance of developing a long-range building plan rather than relying on the traditional 'bottom up' approach to outlet construction described earlier.

## 2. EMPIRICAL VALIDATION

Initially, the company provided two types of data sources. The first was a retail competitive survey which was conducted annually by company

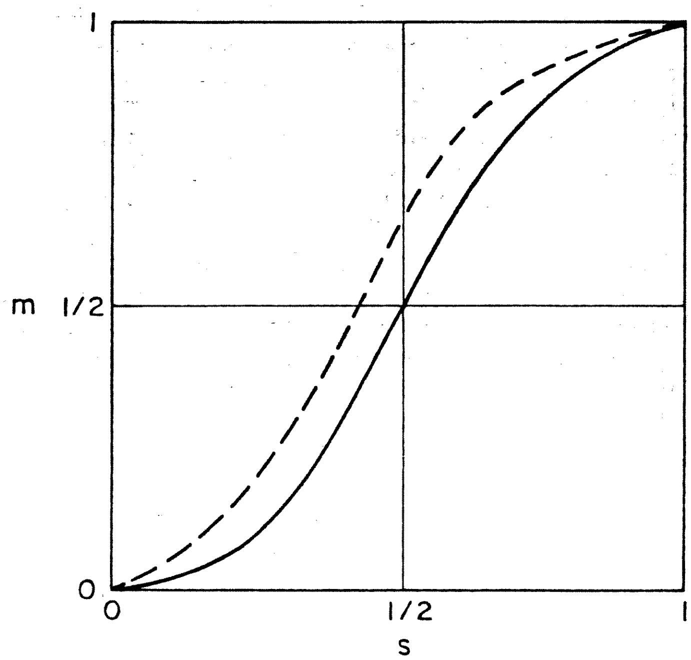


**Figure 1**

salesmen. This survey provided information on outlet numbers and estimated sales volumes by brand. Company management considered this data source to be much more reliable than commercially available data of the same type. The second data source was the new outlets openings report, a record of all new outlet openings for the last ten years. Data from both sources were initially available for thirty markets and were used for estimation. Outlets that were less than five years old were classified as 'recently built.' This admittedly arbitrary classification provided the best fit and also agreed with the intuition of marketing management. A variable called "aggressiveness" was defined as

$$a = \frac{\dfrac{\text{No. of recently built company outlets}}{\text{Total company outlets}}}{\dfrac{\text{No. of recently built industry outlets}}{\text{Total industry outlets}}}$$

and a function $m = g(a, s)$ was fitted to the data. Figure 2 shows contours of the fitted function for $a = 1.25$ and $a = 0.85$. Also shown is the fitted Hartung-Fisher model for the same data set. While in the range $0 < s \leqq 0.14$ there is not too much difference between this model and Hartung-Fisher's, beyond $s = 0.14$ substantial differences occur. A high $R^2$ ($>0.8$) was obtained and the impact of building rate found to be highly significant. The proprietary nature of the data precludes a fuller discussion of the estimates, procedure or presentation of those data.

It should be noted that the results presented were initial ones. In practice the curves are re-estimated each year to reflect the most recent data available. The most recent curves differ somewhat from those shown in Fig. 2, but their general character is as illustrated.

## 3. ALLOCATION PROCEDURE

As mentioned before, the model was designed to aid in a planning problem. The output of the model was to help construct a building plan—how many outlets should the company expect to build in each of a large number of market areas during a planning period (usually a 5-year period). The first year results become budget items—building funds are allocated in accordance with plan 'year 1.' The following year results are used to prepare profit plan projections and to help allocate outlet-site procurement funds (in anticipation of building).

The nature of the managerial decision is such that a near-optimal solution to the mathematical formulation of the problem is quite adequate. All the planned outlets cannot or are not always built, because of changing local building codes, construction difficulties, lack of sites, etc. And if an extra 'choicc' site becomes available in a desirable area, an outlet will be constructed on it immediately, even if no money was originally allocated. Management is concerned with whether it should acquire five sites or twenty sites in an area; the difference between five sites and six often washes out during implementation.

It has been demonstrated that the firm's market share $m$ is related to the aggressiveness $a$ and share of outlets $s$ by a relation

$$m = g(a, s). \tag{1}$$

In general $m$, $a$, $s$ as well as $g$ will be known for a particular market.

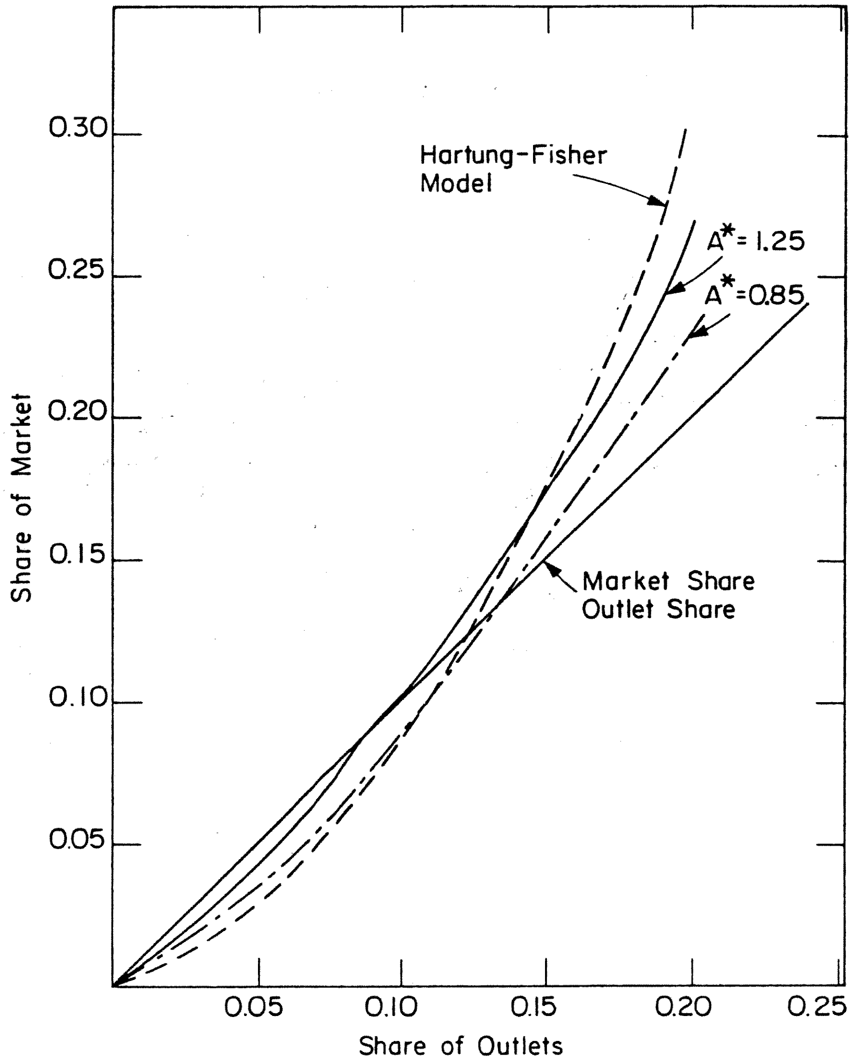


**Figure 2**
($A^*$ = Aggressiveness)

Thus, for consistency the following is assumed:
*A-1*: In equation (1) market share ($m$), aggressiveness ($a$) and the function ($g$) are known with certainty, while outlet share ($s$) is to be determined from the equation.

*A-1* gives an operational definition of outlet share that may seem to be different from the one observed. This discrepancy could be due to (a) differences in the size and effectiveness of outlets in the market (as discussed in Section 1, (b) marketing factors, (c) random, or other factors. The reason for the differences need be of no concern in general; specific, significant differences should be brought to the attention of management for purposes of control.

Given this starting point ($m=m_0$, $s=s_0$, $a=a_0$) and an assumption about nonfirm building rate, one can now calculate the annual expected market share for a given building plan for each year of a planning horizon. A host of other data (growth rates, discount rates, cost factors, margins, etc.) are then needed to choose an economically optimal building plan for a particular market. The details of the economic evaluation will vary from application to application. The highlights of one such application are sketched here.

The problem of determining an optimal building plan was originally formulated as a dynamic programming problem. The procedure was cumbersome, computationally inefficient, and unable to handle several of the constraints. An empirical market-by-market analysis of the relation between cumulative NPV and building investment indicated that most such curves were nearly concave. Thus we scrapped the dynamic programming approach and developed the following algorithm.

The objective of the algorithm is to maximize the total net present value (NPV) of a $Y$-year building program subject to restrictions on the total number of outlets that can be built (a) within a market, (b) across all markets in a given year, and (c) during the $Y$-years, where NPV is defined as

$$\text{NPV} = \sum_{j=1}^{j=J} \sum_{i=1}^{i=\tau} CF_{ij}/(1+R)^{i-1}, \qquad (2)$$

where $CF_{ij}$ = cash flow associated with market area $j$ in year $i$, $R$ = discount rate, $J$ = market areas considered in the plan, and $\tau$ = planning horizon ($\tau > Y$). To do this maximization, the procedure selects the group of outlets in the market that has the highest average NPV per outlet. It then selects the next highest NPV group and so on until all allowable outlets have been allocated.

It will be assumed that (for a particular market), if one knows: the firm's building/investment plan, the firm's current market share, market growth rate, discount rate, margin, competitive building/investment plans, current age distribution of firm/industry outlets, other financial information: land costs, improvement and equipment costs, depreciation methods, working capital needed, etc. then it will be straightforward, to calculate cash flows and, hence, the NPV associated with any particular building plan. The following assumptions have been used in practice in making such NPV calculations; though they are somewhat arbitrary, we trust they seem reasonable.

*A-2*: 'New' outlets, used in the definition of aggressiveness, are defined as those four years old or newer. In year 3 of the building plan, outlets built in years $-1$ (last year), 0 (this year), 1, and 2 are included in the definition of aggressiveness.

The building plan is designed for $Y$ years (where $Y$ usually equals 5); the planning horizon is set for $\tau$ (generally 20) years. Because of the nonlinear relation between outlet share and market share, if one assumed no building after year $Y$ the model could seriously understate the profitability of the building plan. On the other hand, it would be a mistake to assume the continuation (and reap the model-profits) of an aggressive building plan in years after $Y$ (with no capital outlay). As a compromise:

*A-3*: The model assumes, after $Y$ years, that the firm will build enough outlets to maintain its market share: $m_k = m_Y$, $k = Y+1, \cdots$. Thus aggressiveness is assumed equal to 1: $a_k = 1, k = Y+1 \cdots$.

Note that were we considering an infinite planning horizon and an infinite building horizon, *A-3* would not be necessary. There are also some minor end-off problems (which can be taken account of by properly defining salvage values) that this finite-horizon approach entails. However, since the model was developed as an operational tool for managers, it conforms to the planning practices used. The inconveniences encountered in such modelling are rather minor, and the implementation benefits are considerable.

An allocation algorithm for a single building plan can now be developed. We will then extend it to $Y$ years and give theoretical justification for why the procedure is, at least, near optimal.

Let $X_i$ = number of outlets built in market $i$, $n_i$ = market building constraint, $T$ = overall building constraint, $V_{ik}$ = incremental net present value (NPV) of the $k$th station in market $i$, $\sum_{k=i}^{k=j} V_{ik}$ = cumulative NPV of the first $j$ stations in market $i, j = 1, 2, \cdots n_i$,

$$W_{ij} = \begin{cases} \sum_{k=1}^{k=j} V_{ik}/j = \text{average NPV of the first } j \text{ stations in market } i, j = 1, \cdots, n_i, \\ -B \quad = j > n_i, B \text{ is large positive number,} \end{cases}$$

$M$ = Number of markets, and $N = \max_i [n_i]$.

The single-year problem is

$$\max Z = \sum_{i=1}^{i=M} \sum_{k=1}^{k=X_i} V_{ik},$$
$$0 \leqq X_i \leqq n_i, \qquad X_i \text{ integer}, \qquad i = 1, \cdots, M, \qquad (3)$$
$$\sum_{i=1}^{i=M} X_i \leqq T.$$

The solution algorithm is described in Fig. 3.

The rationale behind the algorithm is simple:

THEOREM. *If, in every market,* NPV *is a concave function of the number of*

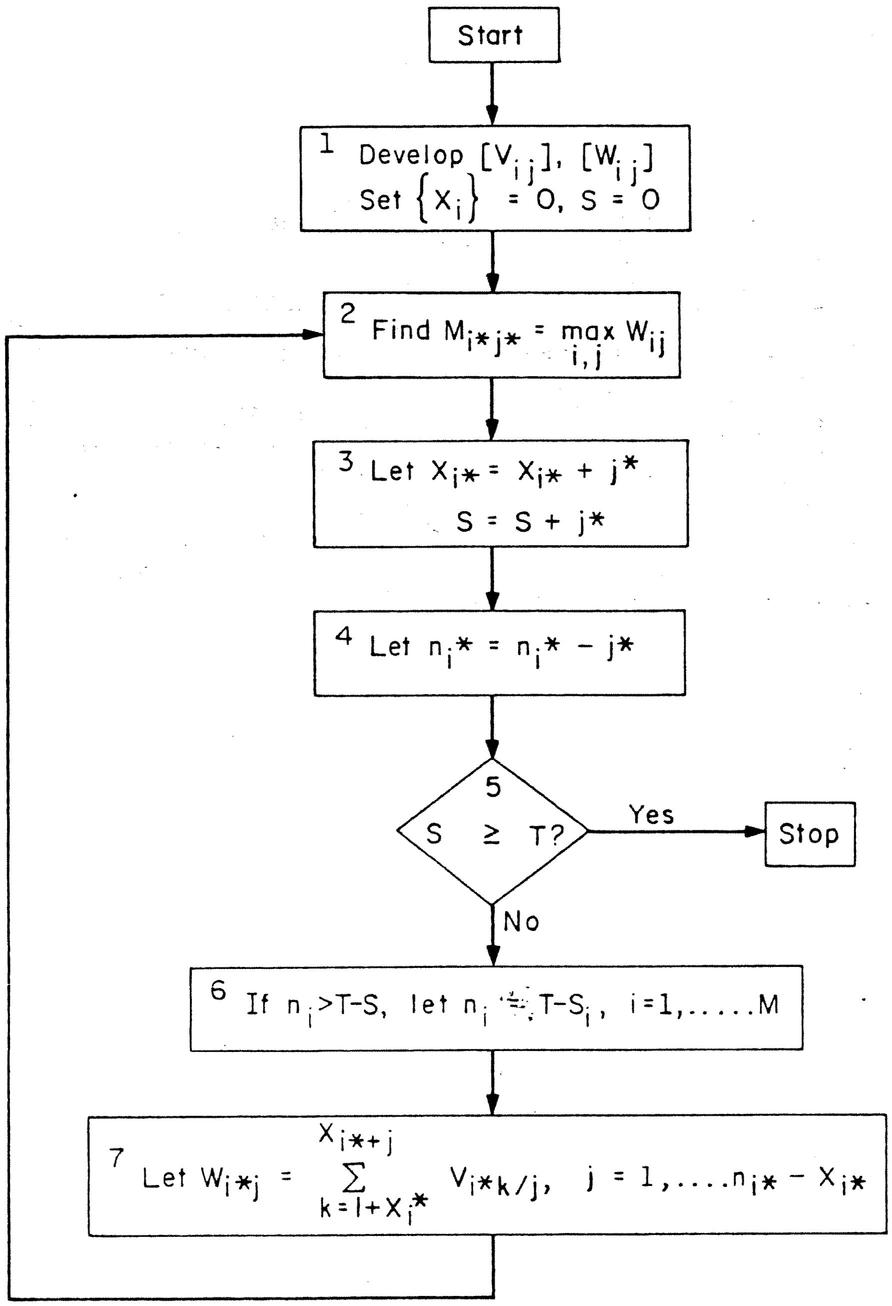


**Fig. 3.** Allocation algorithm.

*outlets built, then a simple allocation according to incremental* NPV *yields an optimal building plan.*

*Proof.* Let $g_j(X_j)$ = cumulative NPV associated with building $X_j$ outlets in market $j$, $j = 1, \cdots, J$, and assume $g_j$ is concave for all $j$. A building plan can be considered a vector $\bar{X} = (X_1, X_2, \cdots, X_J)$, and the NPV of that building plan is simply $G(\bar{X}) = \sum_{j=1}^{j=J} g_j(X_j)$.

Assuming that management uses all building resources, we have the resource constraint $\sum_{j=1}^{j=J} X_j = K$. We then form the Lagrangian: $L(\bar{X}, \lambda)$

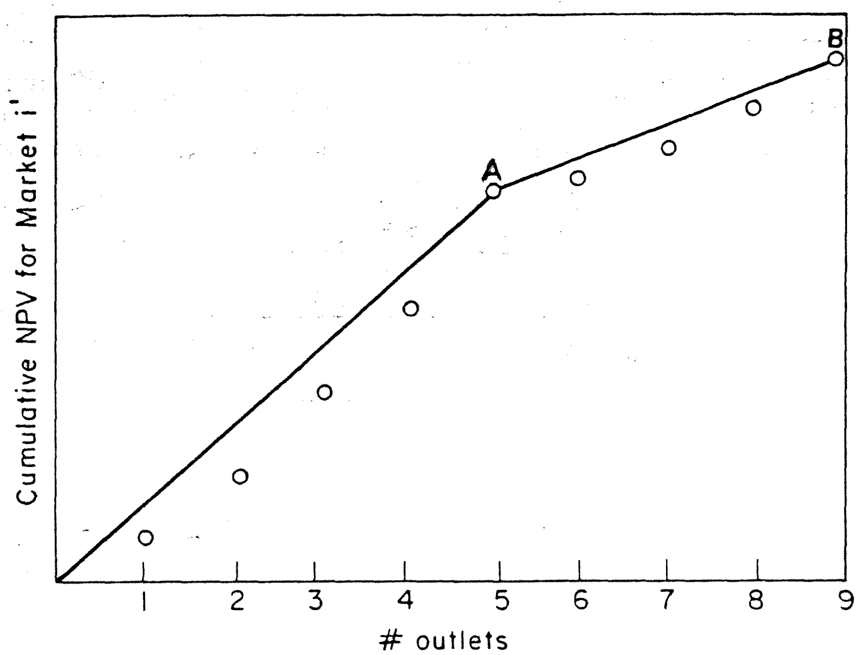


**Figure 4**

$= G(\bar{X}) + \lambda(K - \sum_{j=1}^{j=J} X_j)$. Setting the partial derivatives of $L$ to 0, we obtain $\partial L/\partial X_j = 0 = g_j'(x_j) - \lambda$, $\forall_j$. Thus the NPV maximizing solution has equal incremental NPV's for each market. The solution is a global optimum since $L(\bar{X}, \lambda)$, a sum of concave functions, is concave.

In general, the cumulative NPV curves may not be concave. Thus the $W_{ij}$ matrix is constructed and only entries of maximal size are allocated. This procedure yields a concave envelope for the cumulative NPV curves (transforming them into concave functions). Then maximal $W_{ij}$ entries, the only ones chosen for allocation, always correspond to a feasible point. As an example consider Fig. 4, with the solid line indicating the concave envelope.

Point A in Fig. 4 is a maximal entry for market $i'$. Thus five stations would be built in market $i'$ (assuming it had the highest current $W_{ij}$ entry); and then Step 7 in the algorithm would move the origin 0 to point A, where the algorithm is repeated. The next set of stations picked in this market will correspond to point B, i.e., nine stations (or four additional). Note that if the slope from 0 to B(0B/9) were greater than that from 0 to A (0A/5), either the entire set of nine outlets would be included in the building plan or none would be at all (i.e., $W_{i,9}$ would be the largest entry in the $i'$th row). A numerical example is included in the next section to illustrate the procedure.

Assume the process continues until $T$ outlets are selected (and ignore Step 6 for the moment). Two events are possible: (a) $S$, the running total of outlets, $=T$; (b) $S>T$.

If (a) occurs, the resulting $X_i$ is optimal by the theorem. If (b) occurs, an optimal solution has been found for problem (3), with $S$ replacing $T$. This is not feasible for the original (3), but $S$ is usually close enough to $T$ to be acceptable for planning purposes.

An alternative is to insert a set of steps, (2a), in the algorithm:

(2a) If $S+j^*\leqq T$, continue to 3. If $S+j^*>T$, find $M_{i'j}=\max_i W_{i'j'}$ such that $j'=T-S$. Then let $i'$ replace $i^*$ and go to 3.

This set of steps may lead to a less than optimal solution and is, in essence, an algorithm 'end effect.' The end-effect problem has not proved nearly important enough in practice to justify the alternative dynamic programming solution, which would guarantee theoretical optimality.

Now consider a building program that can span several years ($Y>1$). Define $X_{it}$ = number of outlets built in market $i$ in year $t$, $V_{ijt}$ = incremental NPV of the $j$th outlet built in market $i$, given it is built in year $t$, and $T_t$ = cumulative number of outlets that can be built up through year $t$; $t=1, \cdots, Y$.

All other quantities are altered by adding a subscript $t$ to the prior symbol. The problem becomes

$$\max Z=\sum_{i=1}^{i=M}\sum_{t=1}^{t=Y}\sum_{j=1}^{j=X_{it}} V_{ijt},$$

$$0\leqq X_{it}\leqq n_{it},\qquad X_{it}\text{ integer},\qquad i=1,\cdots,M,\qquad t=1,\cdots,Y,\qquad (4)$$

$$\sum_{k=1}^{k=t}\sum_{i=1}^{i=M} X_{ik}\leqq T_t,\qquad t=1,\cdots,Y.$$

The multi-year problem is slightly more complicated than the single-year problem. Two assumptions make the problem more tractable, however. Assume (A4) $V_{ijt}$ is independent of the *time* at which outlets $j-1$ were built. (A5) $V_{ijt}>V_{ij(t+1)}$—the earlier an outlet is built, the greater its NPV.

Then the algorithm for the multi-year case is very similar to that for the one-year case. The main difference is that the cumulative NPV matrix is formed from a three-dimensional NPV matrix $[V_{ijt}]_{MXNXT}$, where $N=\max_{i,t}[n_{it}]$.

A problem that seems to arise here (the reason for assumption A5) is that even though $V_{ijt}$ is independent of the time at which other outlets are built, the cumulative value of the first $j$ stations does depend on the time at which the first $(j-1)$ outlets are built (because of the aggressiveness definition, among other things).

Since we have assumed that $V_{ijt} > V_{ij(t+1)}$, the cumulative value is greatest when outlets are built as fast as constraints allow. Thus, the algorithm will always assume stations are built as soon as possible, and there is no ambiguity in calculating NPV's.

We still have the 'end-effect' problem mentioned above, and the comments made earlier apply here as well. In addition, another problem lies in the assumption that $V_{ijt} > V_{ij(t+1)}$. This cannot always be assumed in advance, although a large discount rate (internal rate of return) will almost always ensure it. Large market growth rates or profit growth rates could lead to this assumption's being violated.

TABLE I

| Number of outlets | Cumulative NPV Market A | Cumulative NPV Market B |
|---|---|---|
| 1 | 5 | 4 |
| 2 | 8 | 9 |
| 3 | 12 | 16 |
| 4 | 14 | 20.5 |
| 5 | 15 | 22 |

Experience with the procedure has indicated that management generally concedes that the assumptions are reasonable, if debatable. Violations of the assumptions seem to be rare and, when they occur, are slight and have little effect on allocation. As stressed earlier, the type of planning decision that the procedure is designed to support will not be grossly affected by small variations from optimal solutions.

## 4. NUMERICAL EXAMPLE AND COMPUTATIONAL EXPERIENCE

A small, two-market example indicates how the algorithm works. The data are given in Table I.

Assume $n_1 = n_2 = 5$ and $T=5$.
Initially,

*Step 0.* $W_{ij} = \begin{bmatrix} 5 & 4 & 4 & 3.5 & 3 \\ 4 & 4.5 & 5.3 & 5.1 & 4.4 \end{bmatrix} \quad \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad S=0.$

The maximum entry is 5.3 for Market B, 3 outlets. Three outlets are added to $X_2$ and $S$ and $W_{ij}$ are updated:

$$\textit{Step 1.}\quad W_{ij}=\begin{bmatrix}5 & 4 & 4 & 3.5 & 3\\ 4.5 & 3 & -B & -B & -B\end{bmatrix}\quad \begin{bmatrix}X_1\\ X_2\end{bmatrix}=\begin{bmatrix}0\\ 3\end{bmatrix},\qquad S=3.$$

The maximum entry here is 5 for Market A, 1 outlet. $\bar{X}$, $W_{ij}$, $S$ are updated.

$$\textit{Step 2.}\quad W_{ij}=\begin{bmatrix}3 & 3.5 & 3 & 2.5 & -B\\ 4.5 & 3 & -B & -B & -B\end{bmatrix}\bar{X}=\begin{bmatrix}1\\ 3\end{bmatrix},\qquad S=4.$$

The maximum entry is 4.5, Market B, 1 outlet. $S$, $X$ are updated and $S=T$. Hence the procedure stops with the allocation $\bar{X}=1$ outlet in Market A, 4 in Market B.

If this procedure were for a one-year plan, the above would be complete. For a multi-year problem, an additional check has to be made after each allocation to be sure no single-year constraint is violated. Otherwise the procedure is identical.

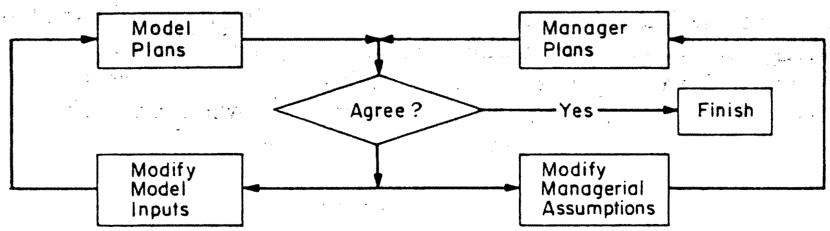


**Fig. 5.** The 'implementation' process.

The algorithm is simple and efficient. Including set-up calculation of NPV's, we have run a 170-market, 5-year problem allocating 3000 outlets in under five minutes on an IBM 360-75. The bulk of that time is $I/O$ and NPV calculation; the allocation procedure itself took less than one minute. This makes update runs and sensitivity analysis quite inexpensive.

## 5. IMPLEMENTATION

This system has been used as an aid in outlet building planning at a major U.S. corporation since 1969. For planning purposes the company breaks the U.S. down into seven operating regions, with each regional manager providing a five-year 'building proposal' for markets in his region. (A region might contain as many as 35 markets.) These proposals are then considered at a building-plan meeting, presided over by the marketing vice president. Invariably the individual proposals add up to more building requests than company annual constraints allow. Prior to the development of the model, political considerations and pseudo-quantitative arguments preceded an executive decision that left little room to reconsideration.

After the model was developed, the regional managers still produced manual proposals; but the model results, produced in parallel, became an additional input at the building-plan meetings. Initial runs were rarely close to the proposals made by the regional managers—input items were changed for further runs and building proposals were updated. After several iterations, model output and regional proposals were close enough so that the few differences could be resolved by hand. This process is schematically represented in Fig. 5.

There are a number of things to be learned from this implementation process. This model-interaction and input revision classifies the model as a 'decision-calculus' type (see Little[2]). The model does not replace or transcend the manager here; rather, the interaction process provides more meaningful model-inputs and leads to more useful outputs. The managers are involved at every stage of shaping the final results; managers trust the model because they control it.

The authors consider the process outlined in Fig. 5 to be indicative of successful model implementation. As should be the case, model results are rarely used as they are. They are one input into the decision-making process, and the assumptions behind the model should be screened and adapted until they seem reasonable. During the screening and updating process, managers learn a great deal about their own decision-situation and we have found, become more secure in their decisions.

## 6. CONCLUSION

A model was developed to help plan retail outlet building. An $S$-shaped, outlet share-market share relation was hypothesized and estimated satisfactorily from company data. This relation was then one input in a resource allocation algorithm that efficiently produced optimal or near-optimal plans.

The results of the study were 'implemented' in the sense that they had an important influence on the decision-making process. By using the procedure, management became more comfortable with it, and the procedure became an integral part of the planning process.

## REFERENCES

1. P. H. Hartung and J. L. Fisher, "Brand Switching and Mathematical Programming in Market Expansion," *Management Sci.* **11,** B231–243 (1965).
2. J. D. C. Little, "Models and Managers: The Concept of a Decision Calculus," *Management Sci.* **16,** B446–485 (1970).
3. P. A. Naert and A. V. Bultez, "A Model of Aggregate Distribution Network Performance," *Management Sci*, (forthcoming).