In presenting the dissertation as a partial fulfillment of the requirements for an advanced degree from the Georgia Institute of Technology, I agree that the Library of the Institute shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish from, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the Dean of the Graduate Division when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

7/25/68

OPTIMAL SALES CALL ALLOCATION

A THESIS

Presented to

The Faculty of the Graduate Division

by

James Jackson Belcher

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Industrial Engineering

Georgia Institute of Technology

June, 1968

OPTIMAL SALES CALL ALLOCATION

Approved:

Chairman

Date approved by Chairman: MARCH 31, 1969

ACKNOWLEDGMENTS

The author wishes to express his appreciation to the following individuals:

Dr. Lynwood A. Johnson, thesis advisor, who contributed a great deal to the defining and clarifying of how this research should be presented.

Mr. Victor L. Harville, Division of Research and Statistics, Georgia Department of Family and Children Services, who not only managed to teach me most of whatever statistics I know, but assisted in the preparation and running of computer programs, and above all showed me through his own example and sense of humor how to develop the fortitude necessary to complete an M.S. thesis.

Mr. H. Jack Belcher, President, Belcher and Associates, Inc., who provided me with data and ideas relating to the decision variables involved with sales call allocation.

Mr. William W. Swart, who contributed his computerized "Pattern-search" program.

Mr. Brent A. Blumenstein, who wrote the excellent stepwise regression program used in this research.

Professor James M. Apple and Anang Z. Gani, in showing me the Gani algorithm for solving the traveling-salesman problem.

The other members of my thesis reading committee, whose advice is greatly appreciated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

SUMMARY

The relevant variables and data available with respect to the salesman's allocation of sales calls among customers and prospects for several product lines are examined. The objective is to determine how expected profits might be maximized.

Various regression models are used in order to mathematically describe the feasible sets of sales calls, as well as to model the relationship between these sales calls and expected sales.

Based on these derived relationships, and taking into account the travel constraints incumbent to a salesman's activities, algorithms of a directed search nature are developed in order to arrive at an optimal set of sales calls.

The results, although theoretical in nature, point to an encouraging outcome, if commercial research along similar lines is carried out.

CHAPTER I

INTRODUCTION

Five basic decisions are inherent in the job of personal selling:

1. Which customers and prospects should be visited?
2. How frequently should these parties be contacted?
3. Which route should the salesman take in contacting these parties?
4. Which products or product lines should be shown?
5. How long should each sales call last?

Notably, the salesman and the vendor for whom he works have little control over most of the relevant factors influencing the decision variables. Industrial sales especially are highly correlated to the two factors of customer needs and economic conditions. The efforts of competitors are also generally beyond the control of a vendor.

Nevertheless, sales effort will to some degree affect the share of business by product and customer that a vendor will receive. The all-important personal relationship between a company and its customers is under the firm's control; this will reflect itself in past and present sales figures; it will be an important criterion with respect to these five decisions. Assessment of past results does then constitute a major responsibility, since the five basic selling decisions will be made on the basis of these results and these decisions must inevitably reflect on a firm's sales and profits.

It would therefore be desirable for the firm to assist the salesman with answering these five basic questions. The primary objective of this research is to demonstrate how applied mathematics might be profitably used in this decision-making area.

The responsibility for making these decisions is almost always left to the salesman. His judgment, the sum of his past experience with customers, his ambitions, his desire for material prosperity, his willingness to travel, and his manner of handling his activities other than selling, form practically the sole basis for his decisions. It is important to recognize that the objectives of the salesman may not always coincide with those of the firm for which he works.

Furthermore, the marketing managers of the salesman's firm can draw upon a wider experience in promoting its products than that of any one salesman; also, they are more aware of the over-all objectives of the firm, and the needs, desires and objectives of other departments within the firm.

It is not an intended purpose to strictly set forth a regulated existence for any real or hypothetical salesman; the primary purpose is instead to show how an *optimal* set of medium to long-range sales goals could be calculated, based on statistical analysis of past data available to the firm. It will be shown that such an analysis will lead to these sales goals based on an *optimal* allocation of sales calls, this allocation being made on both a customer and product basis. It will also be shown that this expected optimal allocation of sales calls falls within the physical limitations of time and distance to be traveled by the salesman.

Literature Survey

A search of the literature indicates that considerable work has been done in several related areas of marketing research. This work falls into four general categories: (1) Sales Forecasting, (2) The Traveling-Salesman Problem, (3) Media Selection in Advertising, and (4) Research into Personal Selling.

In summary, Mercer,[1] Ackoff,[2] Lazar,[3] and others have found that most of the operations research projects carried on in the marketing field have produced little of tangible benefit to industry. Non-quantifiable aspects of marketing problems have posed severe handicaps on operations research efforts. In their attempt to circumvent this hurdle, Lazar[4] has stated that significant variables are often neglected or assumed static. Mercer[5] has noted that marketing is so integrated with production, finance, accounting and other activities that a systems approach seems the only reasonable way to achieve success in marketing studies; few researchers have attempted this, and their results may lead only to suboptimization.

Ackoff and Green[6] have made an especially comprehensive study of the research into personal selling. Their findings show that few projects have had much success; most are rather superficially conducted under a narrow range of environmental conditions; their conclusions are necessarily limited in reliability and applicability.

Sales forecasting is probably the oldest activity of market research, and a great deal of information has been written on the subject. These forecasts are of two general types: (1) forecasts of industry-wide demand, and (2) forecasts of company sales (company-wide, or subdivided by product, product line or territory). At one time, forecasting was strictly an art, and *expert opinion* formed practically the sole basis for forecasts. In recent years, mathematical techniques have gained popularity, with time series analysis and correlation techniques perhaps the most popular methods. Most standard texts on sales forecasting, such as Reichard,[7] stress these two quantitative methods, but do not neglect subjective criteria. Journal articles, such as Wagle[8] or Harrison[9] deal most generally with refinements of quantitatively oriented methods, such as exponential smoothing techniques in time series analysis or regression methods of a correlative nature. In practically all cases, sales forecasting is used to generate predictive data from descriptive data, on the assumption that resources available to the firm will be utilized in the same manner. In other words, answers are often sought to the question such as: "What will sales be next year for product X to customer Y?" but not to one such as: "What *would* sales be if we doubled the number of calls made on behalf of that product to that customer?"

The traveling-salesman problem has instigated a great amount of inquiry into the routing aspects of personal selling. The problem has been attributed to a 1934 seminar talk given at Princeton University by Hassler Whitney and has received wide attention in mathematical journals. Representative articles describing this problem and attempts at its solution are those of Heller[10] and Kuhn.[11] In essence, the problem is that of finding a permutation of distances between N points (cities) so that a route beginning and ending at a particular point x (representing the salesman's home) and having as intermediate stops this specified group of N points is of the shortest possible distance. The name of the problem comes, of course, from the analogy of a salesman traveling through a specific group of cities, who begins and ends his tour at his home.

Two basic methods of attack have been used most frequently. The first is that of mathematical programming, more specifically, linear or integer programming, such as Miller[12] presents. Methods such as these seek to establish a boundary set of feasible solutions, select an initial solution, and through a series of iterations, seek particular solutions that successively approach routes closer to the optimal tour possible. The other basic methods include *branch and bound* techniques, coupled with heuristic approaches. These seek to divide a large problem into successively smaller (and solvable) ones by process of elimination. Dantzig[13] and Little[14] use this approach, and are fairly successful with relatively large problems. While full-scale marketing problems include many more factors than the mere routing of salesmen, no result can be useful if it would not satisfy the logistics (routing) involved, and so this groundwork is both helpful and necessary for researchers in marketing research when personal selling is involved.

Another related area of endeavor by market researchers has been that of selecting the optimal media mix in advertising. Given a fixed advertising budget, the problem is to estimate how resources could be allocated to achieve optimal expected sales. This approach differs significantly from ordinary forecasting in that realignment of resources is studied, whereas most forecasting, as stated previously, assumes no change in resource allocation. This problem has been studied for several years, and linear programming has been used with some success by Bass,[15] Day[16] and others to find that combination of media, at what level of expenditure for each medium, will reach the largest audience or produce the most customers for the product.

A basic assumption in these models concerns the drawing power of repeated advertising messages to the same persons. Figure 1 illustrates how this theory operates with respect to the additional number of customers that additional messages for a product should produce.

It would seem as if these three areas of research would have been integrated into a single project to determine the optimal allocation of personal selling time. Subjective and historical data could aid in basic estimates of value versus costs for sales in future periods; territorial boundaries and customer numbers and locations constitute a set of traveling-salesmen routes that must be covered; a *mix* of calls to various customers in behalf of different products (product lines) is involved, and profits are to be optimized. However, a careful study of published literature revealed no such study, and it is likely no such study has ever been conducted. Instead, one of the above three areas is drawn upon to model the personal selling process, and the decisions involved in personal selling.

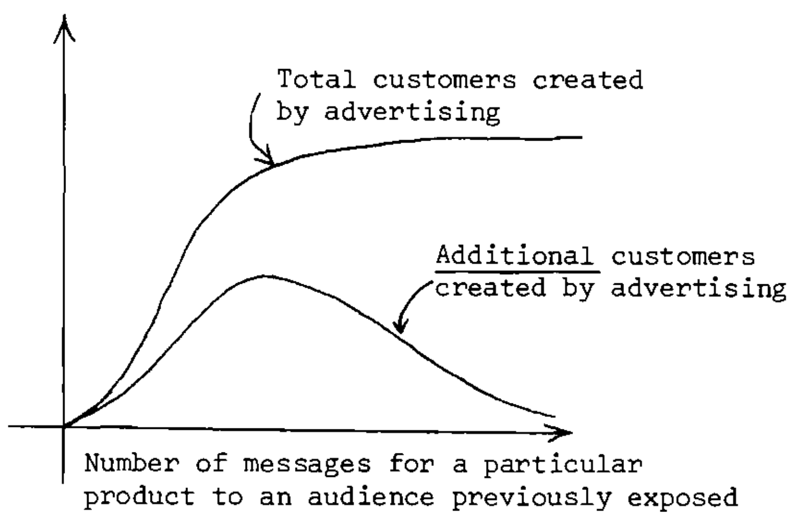


Figure 1. Theoretical Relationship of Repeated Advertising to the Same Audience

Charvat[17] discusses several decisions that can be approached on a quantitative basis, and gives several simplified examples of his techniques. Given a specified function relating Sales Volume to Sales Call duration, Charvat shows how historical data and Bayesian probabilities can be used to effect an optimal allocation of sales calls by the salesman. He uses game theory in determining competitive strategy. Linear programming is used to solve a simple traveling-salesman problem.

A Markov matrix using subjective probabilities as call values is used to show how sales potentials can be estimated.

Sevin[18] uses historical data of an accounting nature to show in general how relative measures of profitability of customers may be determined, but a general algorithm for allocation of personal selling time is not given. Brown[19] conducted consultant work along relatively similar lines, including some experimentation with one client. Without regard to routing restrictions and considering a one-product firm (a printing establishing in this instance), he was able to use experimental data and establish an *optimal* (estimated) reallocation of personal selling time, by use of an empirically derived formula.

The most ambitious, but unfortunately unsuccessful, project of this sort was carried out by Waid.[20] Various statistical manipulations were carried out on historical data in order to evaluate marginal effectiveness of sales call time with various customers, but no correlations seemed to be present between sales effort and sales volume. It was concluded that saturation levels of sales volume had been reached for General Electric lamp customers, and that calls could be reduced without lowering sales. This hypothesis proved correct.

Many of the original ideas of Brown[21] and Waid[22] will be used in this research, but the scope will cover several product lines and the results will fall within feasible routes for salesmen after verification by traveling-salesmen techniques of logistics considerations.

Mathematical Statement of the General Case

GIVEN the following data about the firm:

1. Salesmen

(1) The number of salesmen in the various sales districts.

(2) The time available to these salesmen in a time period for traveling and making sales contacts.

(3) Differences in ability of various salesmen to sell the various product lines.

2. Customers

(1) Location and distance between customers.

(2) Potential volume for the firm's products, by product (product line).

3. Prospects

(1) Location and distance between prospects.

(2) Potential volume for the firm's products by product (product line).

4. Time Period

Length and difference from T_0, the initial time period.

5. Cost Functions

(1) Travel costs per mile (average).

(2) Costs of sales calls other than travel costs, per call.

6. Product Lines

Grouping of products into related product lines.

INITIAL CONDITIONS affecting the problem:

1. Calls Made to Date

(1) By product line.

(2) By customer.

(3) By time period.

(4) By sales district.

2. Salesmen's Performance to Date (Sales Volume)

(1) By product line.

(2) By customer.

(3) By time period.

(4) By sales district.

3. Salesmen's Performance to Date (Effectiveness)

(1) Market share by product and customer.

(2) Estimated present ability to sell a given product line.

4. Gross Margin (Current) of Various Product Lines*

*For the purposes of this research, gross margin of a product will be the sales price of a product minus variable costs of production as a percentage of the sales price of that product.

DECISION VARIABLES:

1. Whom to call.**

2. When (frequency) to call.

3. In what sequence to make calls.

4. What products to show.

5. How long to stay at a given call.***

Then in order to maximize expected profits in some future time period T', the following expression should be maximized:

$$\sum_{j=1}^{M} \sum_{i=1}^{N} \alpha_i s_{ij}$$

through the obtaining of an ordered set of $\vec{C}$, subject to the physical limitations of the givens:

$$\sum_{j=1}^{M} \sum_{i=1}^{N} n_{ij}(\phi_{ij} + \theta_{ij} + \psi_{ij}) \leq T$$

**For the purposes of this research, this decision variable will be treated on a customer basis, not on an individual of the customer basis, since it is felt that quantifiable data is almost nonexistent to treat this decision variable accurately, and that nonquantifiable data pertinent to this decision variable would contain bias in the long run mostly detrimental to a constructive solution to the basic data sets sought.

***Owing primarily to lack of data, this research will not involve itself with this decision variable. Instead, the average length of time associated with calls of the various product lines will be taken as the actual call; that is, for the purposes of this research, the salesman's individual judgment will still prevail.

where:

N - the number of product lines.

$\vec{C}$ - the entire set of calls made on behalf of product line i to customer j.

α_i - the current gross margin of product line i.

s_{ij} - the sales of product line i to customer j.

ϕ_{ij} - the time spent directly in making a sales call to a customer j for product line i.

θ_{ij} - the indirect time spent in making a sales call on behalf of product line i to customer j; this would include clerical, administrative, technical and miscellaneous duties associated with a salesman's time.

ψ_{ij} - the travel and travel-related (e.g., waiting) time associated with a particular sales call.

n_{ij} - the number of calls made to customer j in behalf of product line i.

T - the total time available to the salesman in which to solicit business.

The solution to this model breaks down into three phases:

PHASE I: Determine the marginal time involved with any call x_{ij}, including the time involved with all three parts of the time constraint equation shown above.

PHASE II: Determine the relationship between s_{ij} and n_{ij} for all x_{ij}.

PHASE III: Develop an optimizing procedure for

$$\sum_i \sum_j \alpha_i s_{ij}$$

given the value of T and the results of Phases I and II.

CHAPTER II

THE CASE FOR ONE SALESMAN, ONE CUSTOMER, M PRODUCTS

In order to more fully understand the solution for the general case, first a simplified case shall be studied, one which lends itself to a logically meaningful solution without resorting to a complex model.

Consider that case in which only one customer is available, and no other firms are solicited for orders; hence, there is no variation in routing patterns. In such a situation, travel and waiting time remain at practically constant proportions of time spent working. Only two activities need be formulated: (1) actual selling time, and (2) clerical, administrative and other duties (which may be highly correlated with the sales call mix, depending upon the engineering, production or marketing aspects of the various products).

Assuming an economically rational firm, the stated goal is to establish a particular vector $\vec{C} = \{n_1, n_2, \ldots, n_m\}$, which constitutes a set of numbers n_i that represent for m products that number of calls for each product that will optimize expected total profit for the firm.

While it is apparent that direct selling expenses play a part in establishing the total cost of a product, the assumption shall be made (which can be shown to be reasonable in most practical examples) that a reallocation of sales calls will not appreciably affect average gross profit margins for individual products.

By examining past profit figures (by product) for the firm, noting profit trends as well as average profit percentages, the average expected gross profit percentages for each product can be computed. These percentages will be represented by a vector $\vec{P} = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ for the m products of the firm.

Using analogous notation for the sales (by product), we may then define the company's objective as that of maximizing expected profits per sales dollar times expected sales for each product, i.e., $\alpha_i s_i$. On a company-wide basis, the objective is therefore to maximize $\vec{S} \cdot \vec{P}$.

However, the sales for each product, s_i, are themselves very closely allied with the number of calls made in behalf of each product, e.g., $s_i = f_i(n_i)$. In fact, our problem can be stated as one of maximizing total profits $P_t = \sum_{i=1}^{m} \alpha_i s_i = \sum_{i=1}^{m} \alpha_i f_i(n_i)$. The only constraint on the n_i is that the time available to the salesman for making sales calls not be exceeded by the time required to actually make $\vec{C}$, a particular set of calls.

Assuming that average time per call is known for each product, our solution for the optimal company-wide profit, P_t, breaks down into three stages:

1. Determine the relationship between the sales call *mix* and the time not spent in sales calls or travel-related duties.

2. Determine the relationship between sales calls and sales volume for each product line.

3. After the various f_i functions have been quantified in terms of n_i, and the time constraint stated in terms of $\vec{C}$, the optimal set $\vec{C}$ must be selected from those that satisfy the time constraint.

The Relationship Between Sales Call Mix and Nontravel, Nonsales Activities

As briefly mentioned in the discussion of the three major activities that consume the time of a salesman, that segment not directly affected by particular sales calls may well be affected by the *ratio* of sales calls among product lines. For example, there are various activities, such as sales meetings or market research, that may be required for a given product line, but which cannot be logically charged to any one sales call or series of sales calls. In addition, some activities may not affect any particular product line, such as certain intra-company meetings and conferences, trade shows and conventions, personnel, clerical or administrative duties. All together, these activities make up a major segment of time, that together with travel and travel-related time (such as that spent waiting in reception rooms) must detract from that available for actual selling, and hence limit the feasible sets of n_i.

It is noteworthy that this portion of time, which we shall designate Theta (θ) is from a practical standpoint only weakly affected by the sales call mix. Only those calls that require consultation with others in the firm, or clerical attention, can really be said to be sales-related with respect to the sales call mix. The other duties are either not product-dependent or dependent solely on whether or not a product line is carried at all.

Therefore, it appears reasonable to model in the following manner:

$$\theta = \beta_0 + \beta_1 n_1 + \beta_2 n_2 + \cdots + \beta_m n_m$$

This is the general model for a linear multivariate regression analysis; from an applications viewpoint, this is precisely that technique that seems most suitable. Naturally, those techniques that show the extent of variance explained by the regression, and over-all effectiveness measures of the particular equation chosen should be simultaneously undertaken. Correlation of the regression coefficients, an F-test (ANOVA) of the regression, computation of the standard error of estimate of regression, and the coefficient of multiple correlation R are the statistical elements that effect such a rigorous test of the validity of the regression equation.

In summary, θ is developed in three steps:

1. Determine θ for each of several time periods.
2. Determine $\vec{c} = \{n_1, n_2, \cdots, n_m\}$ for those same time periods.
3. Use multiple regression analysis to relate the two sets of numbers.

The Relationship Between Sales Calls and Sales Volume

Many factors affect the sales volume of a particular product sold to a given firm. Among these are:

1. The sales effort made in behalf of this product.
2. The quality of this sales effort.

3. The need for the product.

4. The availability of substitutes, and the efforts made to sell these products.

5. The relationships between the given product and its substitutes in terms of service, reliability, quality, price, etc.

6. The personal relationships between the personnel of the supplier and the consumer.

7. General business conditions, and trends in the general economy.

8. The financial conditions and marketing performance of the customer, and trends in the company's overall condition.

Close examination of these factors reveals that all are reflected in the relationship of sales effort to sales volume. If we assume that a particular salesman's efforts to sell a given product to a given customer are accurately measured in the number of calls made in behalf of that product to that customer, we then can simplify this complex function between sales volume and the many factors that affect it into $s_i = f_i(n_i)$. Indeed, the only factors concerning the sales volume that are controllable by the marketing function deal with the allocation of efforts, and the nature of these efforts (in an attempt to upgrade them, and effect closer personal vendor-consumer relationships) directed to customers for the various products sold.

It should be expected that a graph of the number of calls for a given product to a given customer in a particular time period compared with the sales for that time period, customer and product would resemble an S-shaped curve, as shown in Figure 2.

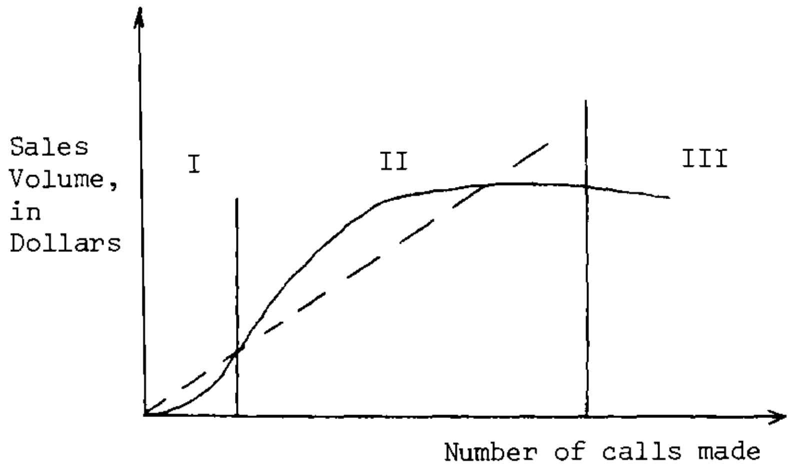


Figure 2. Theoretical Sales Call to Sales Volume Relationship

Region I marks the area where successive calls have a reinforcement value in introducing the product. It is relatively small. Region II is the area of diminishing marginal returns. It would be expected that an allocation of calls among various products would include data points almost altogether falling in this region. Region III shows where the entire available market has been absorbed, insofar as direct selling effort is concerned. Additional calls have a *nuisance* effect, and so present a negative marginal sales return.

As a matter of incidental interest, the dotted line shows how a cost-return balance would be represented, if only one product was handled. We could assume linearity of the costs of sales calls; the optimal number of calls would be that point in Region II where the value of additional sales (expected marginal gross profit) equaled the cost of additional calls.

Dealing with more than one product, however, and with a fixed time resource to allocate, the distribution of sales calls among products for an over-all optimal allocation may require n_i to fall anywhere in Region I or II for any particular product.

A practical model to reflect the expected S-curve for s_i would therefore be:

$$s_i = c + \rho_1 n_i + \rho_2 n_i^2 + \cdots + \rho_\varepsilon n_i^\varepsilon.$$

This is the model for the general polynomial regression equation with one independent and one dependent variable. Again, statistical methods can be adopted to measure the degree of variance explained by the particular equation chosen. Also, by comparing the differences in sums of squares of error variance as higher order terms of s_i are added in a stepwise procedure, the model equation selected can be limited to the power of the terms that will best balance the accuracy of the predictive equation with the difficulty of using more complex equations as parts of an objective function in our final stage of solution.

Before leaving this topic, special consideration should be given to two aspects of the s_i relationships. First, an effective call or series of sales calls might not result in immediate orders. Instead, various financial, engineering, and purchasing considerations might delay the writing of the order until a later time period. It then seems reasonable to study the s_i of later time periods with the n_i of a group of time periods to note if a consistent relationship, a time lag, does indeed exist. If this proves to be the case, this time lag should be retained in the predictive equation, unless it is relatively short in comparison to the length of the period being forecast.

Second, it must be realized that s_i is a dynamic function; it is not static with respect to time. While a given observation may reflect accurately a s_i/n_i relationship at a given time, this probably will not hold in the future. However, each observation of s_i and n_i will occur in a different time period. Theoretically, it appears that the situation reflected by Figure 3 could occur.

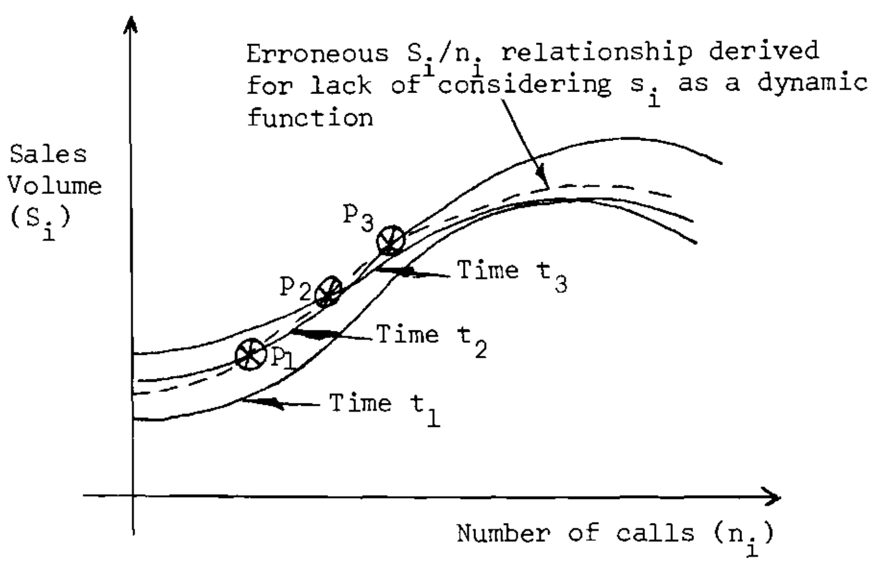


Figure 3. S_i/n_i Relationships for Three Different Time Periods

Three different S-curves, and three p_j observations show this shifting of s_i.

To overcome this difficulty, a hypothesis can be made:

H_o: No shift of $f_i(n_i)$ occurs between p_t and $p_{t'}$.

H_a: The function has changed during this interval.

We then can take a relatively lengthy period, such as a year or several months, and hypothesize that during this period, H_o holds true. Monthly or weekly observations could be made and an s_i function obtained. Then other major periods could have s_i functions developed for them. Long-term trends could be reflected in the function chosen to predict the next major time period. The differences in the s_i functions for the various major time periods could be tested for statistical significance, using an F-test such as that presented by Williams.[23] Whether or not either a time lag of s_i to n_i or a trend effect is present, each item should be investigated before a final s_i predictive equation for a given product is established with respect to future time periods.

Determination of an Optimal $\vec{C}$

At this stage of model development, both the expected sales per product line, and the time available for direct selling have been formulated in terms of a set of n_i, which has been designated $\vec{C}$.

Now notation will be introduced to represent the three different activities of the salesman's duties, and show the time constraint in terms of these three elements. First, psi (ψ) (in this case a constant, owing to our fixed routing) shall be used to represent the travel time. Second, phi (ϕ) is designated as the time spent in direct selling effort. Then, designating T_t as the total time spent working by the salesman, it is apparent that $T_t = \theta + \psi + \phi$. But $\theta = \beta_o + \beta_1 n_1 + \beta_2 n_2 + \cdots + \beta_m n_m$. Also, $\phi = t_1 n_1 + t_2 n_2 + \cdots + t_m n_m$, where t_i is the average time per call for product i. Since T_t and ψ are constants, our time constraint appears as:

$$T_t \leq \psi + \phi + \theta = \sum_{i=1}^{m} t_i n_i + \psi + a + \sum_{i=1}^{m} \beta_i n_i$$

$$\therefore T_t - \psi - a \leq \sum_{i=1}^{m} (\beta_i + t_i) n_i$$

Using notation introduced earlier, our objective function, P_t, is to be maximized:

$$P_t = \sum_{i=1}^{m} x_i f_i(n_i)$$

But

$$\sum_{i=1}^{m} \alpha_i f_i(n_i) \equiv \sum_{i=1}^{m} \alpha_i [\rho_{1i} n_i + \rho_{2i} n_i^2 + \cdots + \rho_{mi} n_i^m]$$

Therefore, the entire analysis of this case has been reduced to two principal equations, where the various α_i, β_i, and t_i are derived constants, and T_t, ψ, and β_o are also known constants.

Obviously, all $n_i \geq 0$, but the absolute upper bound of each of the n_i is also easily derived. Each of the m equations for s_i could have derivatives taken with respect to n_i, and the sum of the terms of the polynomials set to zero. This would yield a number of inflection points (equal to the power of the polynomial) for each s_i equation, one of which would be the limiting value of n_i, since it would mark the boundaries of regions II and III, as discussed previously.

We thus have defined a region in space of m + 1 dimensions, which constitutes the feasible combinations of the n_i. This would indicate that a mathematical programming program of a particular sort has been set forth: The objective function is of the highest power retained by the various polynomial terms of the functions for s_i, and the one principal constraint is of linear form.

Mathematical programming problems of this nature are solvable by several techniques, but the Hocke-and-Jeeves directed search technique seems especially suitable, due to its relative ease of computation, and its adaptability for computerized calculation.

An Example of the One-Customer Case

To illustrate the discussion of the previous section, the following problem will be solved:

A salesman working an average of 175 hours per month (which shall be considered the maximum hours he is to work) sells three different products to one customer.

The three products are quite different from a marketing standpoint. Product A, for which the gross margin is only 7%, is a supply-type item. Volume is relatively high and rather consistent. Selling embodies more of a public relations and service aspect than in other lines, and 50-65 calls per month are normally made on behalf of this product. Products B and C are of an altogether different nature, and in*

---

*Gross margin shall be defined as the average percentage difference between the selling price and variable cost of manufacture of a product.

fact are substitutes for one another. Product B is newer, less well known and accepted, technologically more sophisticated, and higher priced. It seems to require much persistence in selling efforts, as well as a great deal of coordination with the production and engineering departments for the salesman to effect sales of product B, which may come several months after initial selling efforts have started. The 14% gross margin on product B does, however, compensate somewhat for these problems in selling. (It will be assumed that the salesman's compensation is tied to the profits on sales made to his customer.) Product C is also a small equipment item for the same purpose, but better recognized and accepted, less expensive, and easier to sell. It carries a 10% gross margin.

Past studies have shown that this salesman spends about 1/2 hour per call for product A, 1 hour per call for product C, and 1-3/4 hours per call for product B. Also, the average travel time and time spent waiting to see prospects is about 1-1/2 hours per day (or 33 hours per month).

In addition, the following statistics are known with respect to sales and calls made for each month of the past two years to this customer.

Table 1. Sales and Sales Call Statistics--One Customer Example

Month	Sales Calls			Sales (In Dollars)		
	N_A	N_B	N_C	S_A	S_B	S_C
1	75	3	20	23000	3700	6800
2	60	5	20	19500	3100	6500
3	55	10	5	18000	2800	7100
4	55	5	20	18500	2300	4300
5	50	5	25	15500	3200	7000
6	60	2	25	19700	3900	8200
7	65	0	25	21800	3400	8400
8	55	0	30	18800	3000	8000
9	50	2	25	17200	1800	9200
10	40	8	20	12500	900	8300
11	35	6	25	11800	800	7200
12	45	6	33	13800	1200	8600
13	60	0	30	19000	4100	9900
14	58	2	25	19300	3300	8500
15	60	3	25	19800	3700	8500
16	85	0	15	24800	1600	6100
17	80	1	18	24900	1700	6500
18	63	4	18	22000	1600	6800

Table 1. Sales and Sales Call Statistics--One Customer Example (Continued)

Month	Sales Calls			Sales (In Dollars)		
	N_A	N_B	N_C	S_A	S_B	S_C
19	55	3	25	19800	1100	8300
20	50	3	28	16000	1100	9000
21	55	2	25	17600	2200	9000
22	50	4	21	15400	2100	7700
23	60	6	20	19300	2000	7000
24	60	9	11	19600	1700	7200

Given this information, what is the "best" mix of sales calls (from the vendor's standpoint) for the coming 12 months?

It is assumed that no significant trends are developing, and the salesman's judgment is allowed to prevail as to the length of calls, then the problem has three general parts:

1. Establish the relationship between the salesman's time and the calls he makes.

2. Establish the relationship between a salesman's calls and the gross margin on the sales made.

3. Using the results of steps 1 and 2, optimize the over-all gross margin, assuming the salesman will work 175 hours per month.

Step One

If:

$$T_{total} = \psi_{total} + \theta_{total} + \phi_{total}$$

then

$$\theta_{total} = T_{total} - \psi_{total} - \phi_{total}$$

But,

$$\phi_{total} = \phi_A + \phi_B + \phi_C$$

$$= N_A \cdot t_A + N_B \cdot t_B + N_C \cdot t_C$$

Therefore, all variables other than θ_{total} are known for all 24 observations. Tables 2 and 3 show how this is arrived at for these observations.

Table 2. Determination of ϕ_{total} for One-Customer Example

Month	ϕ_A	ϕ_B	ϕ_C	ϕ_{total}
1	37.5	5.25	20.0	62.75
2	30.0	8.75	20.0	58.75
3	27.5	17.50	5.0	50.00
4	27.5	8.75	20.0	56.25
5	25.0	8.75	25.0	58.75
6	30.0	3.50	25.0	58.50
7	32.5	0.00	25.0	57.50
8	27.5	0.00	30.0	57.50
9	25.0	3.50	25.0	53.50
10	20.0	14.00	20.0	54.00
11	17.5	10.50	25.0	53.00
12	22.5	10.50	33.0	66.00
13	30.0	0.00	30.0	60.00
14	29.0	3.50	25.0	57.50
15	30.0	5.25	25.0	60.25
16	42.5	0.00	15.0	57.50
17	40.0	1.75	18.0	59.75
18	31.5	7.00	18.0	56.50
19	27.5	5.25	25.0	57.75
20	25.0	5.25	28.0	58.25

Table 2. Determination of ϕ_{total} for One-Customer Example (Continued)

Month	ϕ_A	ϕ_B	ϕ_C	ϕ_{total}
21	27.5	3.50	25.0	56.00
22	25.0	7.00	21.0	53.00
23	30.0	10.50	20.0	60.50
24	30.0	15.75	11.0	56.75

Table 3. Determination of θ_{total} for One-Customer Example

Month	T_{total}	ϕ_{total}	ψ_{total}	θ_{total}
1	175.00	62.75	33.00	79.25
2	175.00	58.75	33.00	83.25
3	175.00	50.00	33.00	92.00
4	175.00	56.25	33.00	85.75
5	175.00	58.75	33.00	83.25
6	175.00	58.50	33.00	83.50
7	175.00	57.50	33.00	84.50
8	175.00	57.50	33.00	84.50
9	175.00	53.50	33.00	88.50
10	175.00	54.00	33.00	88.00
11	175.00	53.00	33.00	89.00
12	175.00	66.00	33.00	76.00
13	175.00	60.00	33.00	82.00
14	175.00	57.50	33.00	84.50
15	175.00	60.25	33.00	81.75
16	175.00	57.50	33.00	84.50
17	175.00	59.75	33.00	82.25
18	175.00	56.50	33.00	85.50
19	175.00	57.75	33.00	84.25
20	175.00	58.25	33.00	83.75
21	175.00	56.00	33.00	86.00
22	175.00	53.00	33.00	89.00
23	175.00	60.50	33.00	81.50
24	175.00	56.75	33.00	85.25

Now multiple regression is used to obtain a relationship between the θ_{total} for the various observations and the N_A, N_B, and N_C for those observations.

It is noteworthy that ordinary multiple regression will result in a trivial solution to answers for beta coefficients of N_A, N_B, and N_C. For the reasoning behind this statement, first examine T_{total}: By definition,

$$T_{total} \equiv \phi_{total} + \psi_{total} + \theta_{total}$$

Therefore,

$$\theta_{total} \equiv (T_{total} - \psi_{total}) + \sum_{i=1}^{M} \alpha_i N_i$$

Now if the normal multiple linear regression model is used for θ_{total}, then

$$\theta = \beta_o + \sum_{i=1}^{M} \beta_i N_i$$

and the solution will be the trivial one,

$$\beta_o = (T_{total} - \psi_{total}) \text{ -- a constant in this case.}$$

$$\text{all } \beta_i = -\alpha_i$$

This would be meaningless in this case, inasmuch as it would only confirm the above identity and not identify (as is desired) the indirect time associated with sales calls of the various product lines. A much more reasonable model would be to force θ_{total} through the origin, since this would show a salesman who did not work (made no sales calls) to spend no time on clerical, administrative duties in conjunction with those calls.

The result of such a regression analysis, with $\beta_o = 0$, is shown in Table 4.

This particular regression program produces a number of significant statistics in addition to the three required regression coefficients (β_1 = 0.74441 hrs/call; β_2 = 3.43378 hrs/call; β_3 = 1.28096 hrs/call):

1. By dividing the *UBE VARIANCE* by the *STANDARD DEVIATION* for each β_i, we can obtain a T-statistic to test the significance of each coefficient. All three β_i for this model are highly significant.

2. The *SD CONFIDENCE INTERVAL* shows the 95 per cent confidence interval of each β_i. The range is not notably large for any of these coefficients, although regrettably it is largest for the numerically greatest β_i, i.e., β_2.

3. The *RSQ#* of 0.9920182637 shows that this equation explains 99.20 per cent + of the total variance, which naturally is highly significant.

4. The *RSQ INCREMENT* shows that these particular data values indicate a high degree of interrelationship between variables. That is, much of the variation explained by one variable is explained by another

Table 4. Summary of Computer Results for Multiple Regression Model in One-Customer Example

```
    MLE, UBE PARAMETER      UBE VARIANCE       STANDARD DEVIATION        SD CONFIDENCE INTERVAL               F%1,    21@
B%  1 #@0.7444117380D 00   0.6684942113D-02   0.8176149529D-01    % 0.662650243D 00, 0.826173233D 00@  0.8289508365D 02
B%  2 #@0.3433777867D 01   0.2760098419D 00   0.5253663882D 00    % 0.290841148D 01, 0.395914426D 01@  0.4271887683D 02
B%  3 #@0.1280957474D 01   0.3816845803D-01   0.1953674948D 00    % 0.108558998D 01, 0.147632497D 01@  0.4298973903D 02

    RSQ#              0.9920182637D 00   SSE#    0.1368961244D 04   MLE VARIANCE#         0.5704005181D 02
    1.0-RSQ#          0.7981736311D-02   SSR#    0.1701427488D 06   UBE VARIANCE#         0.6518863064D 02
    F%  3,     21@#  0.8700021618D 03   DELTA#  0.6170776387D-01   STANDARD DEVIATION#  0.8073947649D 01

    STEP ANALYSIS

    STEP           RSQ               RSQ INCREMENT     NU    COND F%1, 21@       RSQ UP          NH       F%NH, 21@

1   X%  1  @0.9582685660D 00    0.9582685660D 00   23   0.2521210812D 04   0.9920182637D 00    3    0.8700021618D 03
2   X%  3  @0.9757815584D 00    0.1751299239D-01   22   0.4607679656D 02   0.3374969765D-01    2    0.4439783670D 02
3   X%  2  @0.9920182637D 00    0.1623670526D-01   21   0.4271887683D 02   0.1623670526D-01    1    0.4271887683D 02

                                            PARAMETER COVARIANCES AND CORRELATIONS

                                                1                  2                  3
                                        1  0.6684942113D-02
                                           0.1000000000D 01
                                        2 -0.1468435122D-01  0.2760098419D 00
                                          -0.3418563520D 00  0.1000000000D 01
                                        3 -0.1377070990D-01 -0.3568367152D-02  0.3816845803D-01
                                          -0.8620951391D 00 -0.3476601723D-01  0.1000000000D 01

                                                     RESIDUAL ANALYSIS

                                        ID        Y           ESTIMATE OF Y     RESIDUAL
                                        01    0.792000D 02    0.917514D 02    -0.125514D 02
                                        02    0.832000D 02    0.874527D 02    -0.425274D 01
                                        03    0.920000D 02    0.816852D 02     0.103148D 02
                                        04    0.857000D 02    0.837307D 02     0.196932D 01
                                        05    0.835000D 02    0.835562D 02    -0.561969D-01
                                        06    0.832000D 02    0.864134D 02    -0.321341D 01
                                        07    0.845000D 02    0.804107D 02     0.408930D 01
                                        08    0.845000D 02    0.793714D 02     0.512863D 01
                                        09    0.885000D 02    0.761121D 02     0.123879D 02
                                        10    0.880000D 02    0.828658D 02     0.513416D 01
                                        11    0.890000D 02    0.786810D 02     0.103190D 02
                                        12    0.760000D 02    0.963728D 02    -0.203728D 02
                                        13    0.820000D 02    0.830934D 02    -0.109343D 01
                                        14    0.845000D 02    0.820674D 02     0.243263D 01
                                        15    0.817000D 02    0.869900D 02    -0.528997D 01
                                        16    0.845000D 02    0.824894D 02     0.201064D 01
                                        17    0.822000D 02    0.860440D 02    -0.384395D 01
                                        18    0.855000D 02    0.836903D 02     0.180971D 01
                                        19    0.842000D 02    0.832679D 02     0.932084D 00
                                        20    0.837000D 02    0.833887D 02     0.311270D 00
                                        21    0.860000D 02    0.798341D 02     0.616586D 01
                                        22    0.890000D 02    0.778558D 02     0.111442D 02
                                        23    0.815000D 02    0.908865D 02    -0.938652D 01
                                        24    0.852000D 02    0.896592D 02    -0.445924D 01
```

set of N_i. We should not, however, try to oversimplify our model and *drop* variables from our model equation unless we can be confident this situation will not change.

5. The *RESIDUAL ANALYSIS* shows the accuracy of the model for each of the 24 observations; this appears to be particularly encouraging in determining the forecasting accuracy of our model.

6. The *COND F%1, 21@* shows that as additional variables were added, how significant the incremental square of the coefficient of multiple correlation was. These show all three variables contributed significantly.

7. The F%1, 21@ statistics show a very high degree of significance in each β_i value.

In light of these findings, it appears sound to consider the β_i values as good estimators of the indirect time associated with calls of the three product lines. In addition, all reflect results in line with the general description of the nature of the product lines as given in the problem description.

The time constraint (relationship of clerical, administrative and miscellaneous time in relation to the numbers of calls made in behalf of the various product lines) can now be formed. Since, in the one customer case, ψ_{total} is constant, and in this particular case is 33 hrs/month,

$$T_{total} = 33 \text{ hours/month}$$

$$+ (\alpha_1 + \beta_1)N_1 + (\alpha_2 + \beta_2)N_2 + (\alpha_3 + \beta_3)N_3$$

$$\therefore T_{total} = 33 \text{ hrs/month} + (1.49 \text{ hrs/call})(N_1)$$

$$+ (5.18 \text{ hrs/call})(N_2) + (2.28 \text{ hrs/call})(N_3)$$

and the constraint is

$$175 \text{ hrs/month} \leq 33 \text{ hrs/month} + (1.49 \text{ hrs/call})(N_1)$$
$$+ (5.18 \text{ hrs/call})(N_2) + (2.28 \text{ hrs/call})(N_3)$$

or,

$$142 \text{ hrs/month} \leq (1.49 \text{ hrs/call})(N_1) + (5.18 \text{ hrs/call})(N_2)$$

$$+ (2.28 \text{ hrs/call})(N_3)$$

This completes Step One.

Step Two

In establishing a relationship of sales of a product line to sales calls made in behalf of that product line, two points should be emphasized. First, the relationship should reflect only an s_i to N_i relationship; thus the function created through polynomial regression should be considered only in its entirety. It is important to understand that as additional terms are added to polynomial model, the determinant of the resulting correlation matrix reduces rapidly to a near-zero value, and the degree of sensitivity to change of individual terms becomes quite large; thus individual regression coefficients mean little or nothing as the power of the expression increases beyond two or three.

Second, the lag between sales and sales calls must be taken into account. If any logical aspect of the product line would indicate that sales would lag sales calls functions by as much as the time period of one observation, then polynomial regression functions should be studied for lagged sales/sales call observations. In this case, product line 2 was modeled as a 0, 1, 2, 3, and 4 time-period lag sales-to-sales-call function while product line 3 was shown with 0, 1, and 2 time period (month) time lags.

The most significant results of these polynomial regression runs are summarized in the data shown in Table 5 on page 34. On the basis of these results and the general nature of the products involved, it seems that the best relationships are obtained when product line 1's sales are lagged 0 months behind sales calls for line 1, product line 2's sales are lagged 3 months behind calls for line 2, and product line 3's sales are lagged 1 month behind sales calls for product line 3.

Third degree polynomials for these s_i/n_i relationships seemed to represent good models for all three product lines; the F-statistic for

Table 5-1. Summary of Polynomial Regression Results in One-Customer Example Product Line 1

Power of Model Equation	Time Period Lag	Approximate Percentage Variation Explained by Model
First Order	0	92.7%
Second Order	0	95.0%
*Third Order	0	96.5%
Fourth Order	0	-
First Order	1	48.8%
Second Order	1	49.6%
Third Order	1	50.1%
Fourth Order	1	-

*Model selected from this data set with this order of predictive equation.

Model Selected:

$$S_{1_T} = \$18144.92 = \$794.43n_{1_T} + \$22.03n_{1_T}^{2} - \$0.14n_{1_T}^{3}$$

Table 5-2. Summary of Polynomial Regression Results in One-Customer Example Product Line 2

Power of Model Equation	Time Period Lag	Approximate Percentage Variation Explained by Model
First Order	0	12.4%
Second Order	0	14.9%
Third Order	0	18.9%
Fourth Order	0	25.6%
First Order	1	< 1.0%
Second Order	1	< 1.0%
Third Order	1	< 1.0%
Fourth Order	1	< 1.0%
First Order	2	13.7%
Second Order	2	14.2%
Third Order	2	18.0%
Fourth Order	2	29.2%
First Order	3	88.5%
Second Order	3	89.5%
*Third Order	3	94.5%
Fourth Order	3	-
First Order	4	33.9%
Second Order	4	35.0%
Third Order	4	35.2%
Fourth Order	4	36.2%

*Model selected from this data set with this order of predictive equation.

Model Selected:

$$S_{2_T} = \$1090.04 + \$27.08n_{2_{T-3}} + \$121.71n^2_{2_{T-3}} - \$9.65\, n^3_{2_{T-3}}$$

Table 5-3. Summary of Polynomial Regression Results in One-Customer Example Product Line 3

Power of Model Equation	Time Period Lag	Approximate Percentage Variation Explained by Model
First Order	0	32.6%
Second Order	0	42.1%
Third Order	0	49.0%
Fourth Order	0	54.7%
First Order	1	73.3%
Second Order	1	73.3%
*Third Order	1	74.4%
Fourth Order	1	-
First Order	2	7.1%
Second Order	2	7.6%
Third Order	2	7.7%
Fourth Order	2	-

*Model selected from this data set with this order of predictive equation.

Model Selected:

$$S_{3_T} = \$2436.33 + \$452.47S_{3_{T-1}} - \$16.26S_{3_{T-1}}^2 + \$0.28S_{3_{T-1}}^3$$

all three showed they did explain a very large part of the variance when sales were lagged behind calls as mentioned above.

The models obtained were:

$$s_{1_T} = \$18144.922 - \$794.425\ n_{1_T} + \$22.034\ n_{1_T}^2 - \$0.139\ n_{1_T}^3$$

$$s_{2_T} = \$1090.04468 + \$27.68124\ n_{2_{T-3}} + \$121.79711\ n_{2_{T-3}}^2 - \$9.65821\ n_{2_{T-3}}^3$$

$$s_{3_T} = \$2436.33203 + \$452.47461\ n_{3_{T-1}} - \$16.26376\ n_{3_{T-1}}^2 + \$0.28443\ n_{3_{T-1}}^3$$

This completes Step Two.

Step Three

As an optimization method for this particular non-linear programming problem, the Hooke and Jeeves' directed search technique (commonly known as Patternsearch) seems especially well suited. It is easily computerized and gives accuracy to any number of decimal places desired in a relatively short period.

In addition to the time constraint developed in Step One, non-negativity constraints on the n_i, and the objective function formed by multiplying s_i function terms by gross margins and adding all s_i estimating equations together, additional constraints representing managerial decisions as to permissible ranges of n_i should also be entered at this time. For our example suppose that management has decided that sales calls for product line 1 should range somewhere between 30 and 80 calls per month, while calls for product line 2 should range between 0 and 10 per month, and between 0 and 35 for product line 3.

The results of the Patternsearch program on this data are shown in Table 6. This shows that 34, 7 and 24 calls per month of product lines 1, 2 and 3, respectively, should yield a gross margin of $2119.64 per month. Therefore, these three numbers constitute the *best* sales call *mix*.

Discussion of This Example

It should be noted that the results shown are in marked contrast to the past performance of this salesman for two of the three product lines. The average number of calls per month for the 24-month period just completed is 57.5 for line 1, 4 for line 2 and 22 for line 3. The analysis of these past statistics seems to indicate that much effort (time) is being wasted on line 1 to the detriment of line 2 and the firm as a whole, insofar as this salesman is concerned.

Specifically, gross margin will increase from about $1725 per month to $2120 per month, if the estimates established in Steps One and Two hold true. While combined estimates of accuracy and reliability would be quite difficult or impossible to obtain, the discussion and

Table 6. Summary of Patternsearch Results in One-Customer Example

I. Patternsearch Parameters:

Initial (Base) Point: $n_1 = 30$, $n_2 = 3$, $n_3 = 20$

Maximum Stepsize: 2 calls

Minimum Stepsize: 0.05 calls

II. Optimum Located After 38 Iterations

$n_1 = 34.2188380$ calls/month

$n_2 = 7.0000000$ calls/month

$n_3 = 24.0000000$ calls/month

III. Value of objective function (expected monthly gross margin) at this call level is approximately $2120.00 per month.

computer program results in each of these two steps shows that it is quite reasonable to believe they are indeed rather accurate and reliable. In other words, if this were an actual case, there is a strong likelihood of success that it has been mathematically determined how this salesman's activities could bring over 20 per cent additional profit for this firm.

Furthermore, it is likely that much greater disparities would be shown had an example of more product lines been presented. As more and more product lines are present, the decision evaluation grows geometrically more complex; the resulting judgments made mentally by the salesman would probably be far less reliable than in this three-product case.

Table 7. Summarized Statistics--Second Regression Model of One-Customer Case

Model: $\theta_T = 10$ hrs/mo $+ \beta_1 n_1 + \beta_2 n_2 + \beta_3 n_3$

β Values	UBE Parameter	UBE Variance
β_1	0.65675 hrs/call	0.00577 hrs/call
β_2	3.06864 hrs/call	0.23854 hrs/call
β_3	1.12034 hrs/call	0.03987 hrs/call

RSQ# = 0.99112

1.0 - RSQ = 0.00887

F(3,21) = 782.05417

Maximum Residual = -18.9 hrs/mo

Suppose now that the management of this firm is not satisfied with some of the criteria used in Steps One and Two. Specifically, it is said that the decision to force the regression model in Step One through the origin was in error, and also that sales trends should have been modeled into the s_i/n_i relationships.

Logically, the judgment used in deciding that β_o should be set equal to zero seemed plausible, but for the moment assume that ten hours per month of miscellaneous time are not explained by the n_i values.

That is, the same model will be run again after adjusting Y to be ten less than previously. The results of this computation are shown in Table 7.

First, notice that the amount of over-all variance explained in this model is very slightly less than before. Now notice what the effect has been on the beta coefficients:

Coefficient	Previous Model	This Model
β_1	0.74 hrs/call	0.65 hrs/call
β_2	3.43 hrs/call	3.07 hrs/call
β_3	1.28 hrs/call	1.12 hrs/call

Since lesser Y values are to be estimated with identical X values, it is likely that the beta weights would decrease but not in these proportions. The over-all decrease in $\overline{Y}$ is -11.9 per cent, in beta-1 -12.1 per cent, in beta-2 -10.5 per cent, and in beta-3 -12.5 per cent. Since these all fall in a narrow range, there is no reason to believe that this model is a better forecaster of miscellaneous time. Therefore, this model should be rejected.

Now consider the extreme case. The assumption will be made that miscellaneous time is in no way explained by the n_i values. Table 8 shows the result of using a model constructed under this assumption. Since the R^2 value is so low, some of the beta-weights are negative (meaningless with respect to time), and the various reliability measures are so unfavorable, this model should also be rejected.

Table 8. Summarized Statistics--Third Regression Model of One-Customer Case

Model:	$\theta_T = \overline{\theta_T} + \beta_1 n_1 + \beta_2 n_2 + \beta_3 n_3$	
β Values	UBE Parameter	UBE Variance
β_1	0.003783 hrs/call	0.001100 hrs/call
β_2	0.348752 hrs/call	0.045421 hrs/call
β_3	-0.076063 hrs/call	0.006281 hrs/call
	RSQ# = 0.14664	
	1.0 - RSQ# = 0.85335	
	F(3,21) = 1.20288	

Under the circumstances, it does not seem reasonable to pursue further the modeling of θ_{total}. Therefore, more models of s_i will then be examined, with special emphasis on modeling in *trends*. To appreciate the effect forecaster assumptions concerning suspected trends can have on final results, an extreme case will be studied:

1. Suppose that a forecaster assumes that real differences were present between the first and second year data.

2. Suppose he is willing to accept polynomial regression lines depicting purported s_i/n_i relationships for each *half* of the data as representing *good* estimators of s_i/n_i for those years.

3. Suppose further that this forecaster, in deciding what to forecast for next year, thinks the same difference between first and

second year data will occur between this past year's and this coming year's data.

4. Using a set of composite s_i/n_i relationships derived under such assumptions, differences in the final, as well as intermediate, results will be studied.

First, examine Table 9. This shows the same sort of calculations as were done in Step Two, although in this case only *half* (since sales are lagged behind calls in some cases, less than 24 observations are available in all cases for the three product lines) of the data for each product is taken to determine an individual s_i/n_i relationship.

In order to maintain consistency, only sales-to-call time lags of 0, 3, and 1 month, respectively, for the three products were considered in these calculations. After all, if sales of product 2 seemed to follow calls made by three months when we used all of the available data, then to search for a different time lag when using only part of the data would be illogical. Also, in line with the procedure in Step Two, polynomials were allowed to, but not required to, be of the fourth order. Now this is not entirely desirable, for with so few observations the likelihood of obtaining regression lines depicting primarily chance relationships is far more probable.

Further, the sales figures in this problem were *planned* to reflect less stability of sales of product 1 in the second year unless at least 50 calls were made, to show an upturn in sales of product 2, and a downturn and stabilizing of sales of product 3. Under such conditions it is reasonable to expect that the n_i obtained should vary at

Table 9-1. Summary of Polynomial Regression Results with Trend Allowances in One-Customer Example

Product Line 1

Power of Model Equation	Time Period Lag	Approximate Percentage Variation Explained by Model
	FIRST YEAR DATA ONLY	
First Order	0	95.5%
*Second Order	0	96.5%
Third Order	0	96.9%
Fourth Order	0	-
	SECOND YEAR DATA ONLY	
First Order	0	88.6%
*Second Order	0	93.5%
Third Order	0	-
Fourth Order	0	-

*Model selected from this data set with this order of predictive equation.

Table 9-2. Summary of Polynomial Regression Results with Trend Allowances One-Customer Example

Product Line 2

Power of Model Equation	Time Period Lag	Approximate Percentage Variation Explained by Model
	FIRST YEAR DATA ONLY	
First Order	3	89.3%
Second Order	3	95.3%
*Third Order	3	97.6%
Fourth Order	3	97.9%
	SECOND YEAR DATA ONLY	
First Order	3	84.7%
*Second Order	3	93.0%
Third Order	3	-
Fourth Order	3	-

*Model selected from this data set with this order of predictive equation.

Table 9-3. Summary of Polynomial Regression Results with Trend Allowances in One-Customer Example

Product Line 3

Power of Model Equation	Time Period Lag	Approximate Percentage Variation Explained by Model
	FIRST YEAR DATA ONLY	
First Order	1	95.7%
Second Order	1	96.4%
*Third Order	1	97.2%
Fourth Order	1	-
	SECOND YEAR DATA ONLY	
First Order	1	48.7%
Second Order	1	51.9%
*Third Order	1	52.4%
Fourth Order	1	-

*Model selected from this data set with this order of predictive equation.

Table 9-4. Summary of Polynomial Regression Results with Trend Allowances in One-Customer Example. Establishing of s_i/n_i Relationships.

PRODUCT LINE 1

Model Selected for First Year Data:

$$S_{1_T} = -\$6020.213 + (\$567.011)(n_{1_T}) - (\$2.318)(n_{1_T}^2)$$

Model Selected for Second Year Data:

$$S_{1_T} = -\$22901.010 + (\$1076.546)(n_{1_T}) - (\$6.034)(n_{1_T}^2)$$

PRODUCT LINE 2

Model Selected for First Year Data:

$$S_{2_T} = \$826.265 + (\$216.350)(n_{2_{T-3}}) + (\$80.099)(n_{2_{T-3}}^2) - (\$7.576)(n_{2_{T-3}}^3)$$

Model Selected for Second Year Data:

$$S_{2_T} = \$1303.758 + (\$24.316)(n_{2_{T-3}}) + (\$56.687)(n_{2_{T-3}}^2)$$

PRODUCT LINE 3

Model Selected for First Year Data:

$$S_{3_T} = \$5263.683 - (\$338.936)(n_{3_{T-1}}) + (\$32.153)(n_{3_{T-1}}^2) - (\$0.550)(n_{3_{T-1}}^3)$$

Model Selected for Second Year Data:

$$S_{3_T} = -\$1852.523 + (\$1146.209)(n_{3_{T-1}}) - (\$50.030)(n_{3_{T-1}}^2) + (\$0.790)(n_{3_{T-1}}^3)$$

COMPOSITE MODELS: Composite models are s_i/n_i relationships formed by adding the algebraic difference between second and first year models to the second year model.

$$S_{1_T} = -\$39781.807 + (\$1586.081)(n_{1_T}) - (\$9.750)(n_{1_T}^2)$$

$$S_{2_T} = \$1781.251 - (\$168.718)(n_{2_{T-3}}) + (\$33.275)(n_{2_{T-3}}^2) + (\$7.576)(n_{2_{T-3}}^3)$$

$$S_{3_T} = -\$8968.731 - (\$2631.354)(n_{3_{T-1}}) - (\$132.215)(n_{3_{T-1}}^2) + (\$2.130)(n_{3_{T-1}}^3)$$

least moderate amounts from those obtained with the original model (which did not adjust for sales trends).

One aspect of examining this model should be emphasized at this stage of development: without studying detailed records of sales and calls, conferring with salesmen and sales managerial personnel, and on a more or less trial-and-error basis studying various forecasting models with different trend adjustments on all available data, it cannot be said whether this model is *better* or *worse* than the original model. Sales forecasting, while appearing at times to bear exacting scientific methods, is still a combination of such *modern* methods and the *rules of thumb, horseback guesswork,* and luck of the forecaster in making these all important assumptions as to the meaning to be extrapolated from available data.

With these considerations in mind, first note that the *planned* trends in the preceding paragraph are reflected to a very large degree in the s_i/n_i relationships obtained. Notably however, the absence of any points showing calls made as less than 50 for product line 1 resulted in a relationship that seems erroneous for $N_1 < 50$ ($s_1 < 0$ for small values of n_1) for second-year data. Also, the first year data for product line 3 resulted in the producing of a model reflecting an extremely high (and unlikely) degree of stability for that product in the first year, while second year data did reflect the upturn accurately; this resulted in a grossly inaccurate composite model for s_3/n_3.

Table 10 shows how Step Three would have resulted if only first-year data were used.

Table 10. Patternsearch Results--First-Year Data Only--One-Customer Example

Initial Base Point:	$n_1 = 40$, $n_2 = 3$, $n_2 = 20$
Optimum n_i:	$n_1 = 44.2117823$
	$n_2 = 5.000000$
	$n_3 = 22.000000$

In this case, call totals of 44, 5, and 22, respectively, would have been recommended for the three product lines. From an over-all viewpoint, these figures are not really so different from those obtained with the original model. Primarily these results indicate that product line 1 rather than product line 2 was the better area in which to spend some ten hours per month.

Table 11 tells a different story.

Table 11. Patternsearch Results--Second-Year Data Only--One-Customer Example

Initial Base Point:	$n_1 = 40$, $n_2 = 3$, $n_2 = 20$
Optimum n_i:	$n_1 = 63.5993966$
	$n_2 = 8.8000000$
	$n_3 = 0.7190656$

The results indicated in this set of calculations advise the salesman to make 64 calls on behalf of product line 1, 9 for product 2 and none for product 3. Primarily this is due to the inaccurate model for product line 1, which forces the result to show an N_1 well above 50 calls per month.

Table 12 is the composite; it is constructed with the extreme assumptions mentioned earlier. Due to the errors mentioned above, it would appear to be invalid due to inaccuracies in the models for product lines 1 and 3. The results are almost identical to those for the second-year data only.

Table 12. Patternsearch Results--Composite Model with Trend Adjustments--One-Customer Example

Initial Base Point:	$n_1 = 40$, $n_2 = 3$, $n_3 = 20$
Optimum n_i:	$n_1 = 64.6929305$
	$n_2 = 8.8000000$
	$n_3 = 0.0057005$

Now we might well ask: "Do we honestly know much more than we did at the beginning of this analysis?" The four sets of *optimal* sales calls give two widely different viewpoints on two of the three product lines.

If we were to terminate discussion at this point, an unfavorable stand concerning the significance of these calculations would probably be unchallenged. However, let us now remember that:

1. Extreme assumptions concerning trend effects were made on the model used in Table 12, while no trend effects were taken into account

with the model used in Table 6.

2. Real trends were present.

3. Chance variation, a highly probable occurrence with the models in Tables 10 and 11 (because of few data points and high-order polynomials), caused erroneous models to be generated with the data of Table 11, and from this Table 12's model was similarly biased.

In essence, the lesson to be learned here is that trend effects concerning sales should be studied, and assumptions agreed upon and models tested *before* any attempt to select optimal sales call *mixes* is made. This research is extending the supposed value to be gained in studying past sales figures beyond what is normally done in sales forecasting; if this data cannot be used to give fairly accurate sales forecasts, then certainly it cannot be well used in determining an optimal sales call mix.

CHAPTER III

THE CASE FOR ONE SALESMAN, M PRODUCTS, N CUSTOMERS

In the previous case, the presence of only one firm to be solicited by the salesman precluded variations in routing patterns; naturally, the model constructed assumed that the ratio of travel time to total time available was static. If, however, more than one firm is *called on* by the salesman, the possible variations in routing patterns introduce three additional decision variables:

1. How many tours (sales trips) will be made in the given time period?

2. Which firms shall be solicited on each of these tours? For which products?

3. In what sequence shall these customers and prospects be visited?

The complicating effect of these additional decisions can best be illustrated with an example. Consider Figure 4:

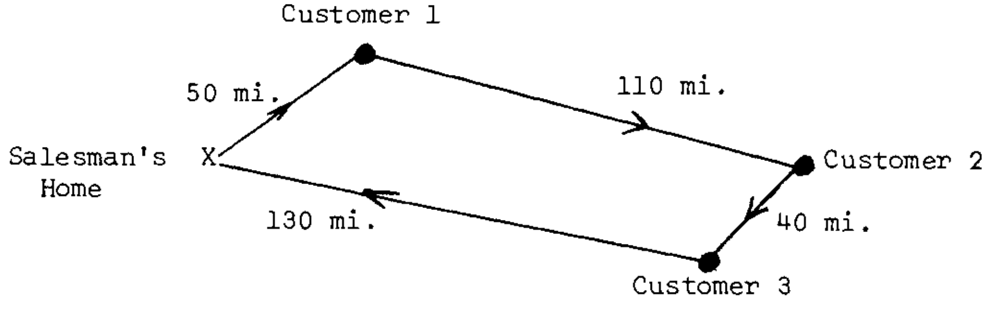


Figure 4. Hypothetical Sales Tour, Multi-Customer Case

While the figure shows a tour that could be made to three customers, it could be assumed that other customers or prospects might have been visited. A number of pertinent questions which were irrelevant in the one-customer case present themselves:

1. Based on past experience, are there some firms which should never be seen again?

2. What is the cost of a call made to a given customer?

3. What restrictions affect the time that the salesman can be away from home?

All of these are aimed, directly or indirectly, at the establishing of some *marginal* concept of the cost of a call, in order to determine what would be the logical income from expected sales that a given call might produce to offset the costs of the call. That is, money is now as important as time in restricting the activities of our hypothetical salesman, if we are to fully understand the multi-customer case.

Retaining our previous notation, now an attempt can be made to establish the costs (in dollars) associated with the ϕ, ψ, and θ time of a call (actual selling time; travel and travel-related time; and clerical and miscellaneous time).

If it can be assumed that prior studies have, for each product i and customer j, provided functions for expected sales, which we shall call s_{ij} functions, and that accounting figures can give the average fixed cost per hour associated with maintaining the salesman (for practical purposes, this would be the direct, administrative, and overhead costs of the salesman but not the travel expenses), which shall be designated γ, then two of these three components are easily calculated:

The total time associated with a calls for all products to all customers is T_{ij}. All costs must be charged to time associated with one or more of these three phases, θ_{ij}, ϕ_{ij}, or ψ_{ij}, that make up T_{ij}. But

$$\theta = \beta_o + \sum_{i,j} (\beta_{ij} \cdot n_{ij})$$

Similarly,

$$\phi = \sum_{i,j} (t_{ij} \cdot n_{ij})$$

So the costs associated with ϕ_{ij} and θ_{ij} are equal to:

$$\sum_{i,j} \gamma[\beta_{ij} + t_{ij}]n_{ij}$$

This is rather straightforward; the same is not true of the travel expense portion. First, there are certain expenses that constitute a variable expenses portion of the cost of the salesman which must be allocated to calls. Second, the time and money in actual travel must be pro-rated with the calls made.

Referring to Figure 1, 330 miles of travel comprise the entire trip. Let us say that the company pays 10¢ per mile for travel to the salesman; this amounts to $33.00 for the trip. In addition, other travel expenses might amount to $75.00, if this tour took several days to complete. And, of course, the costs of directly supporting the salesman, such as administrative expense and his salary, might run \$6.50 per hour; it will be assumed that this tour represents seven hours of actual driving. How this \$153.50 (\$33.00 + \$75.00 + (7)(\$6.50)) is pro-rated among the calls made, is certainly going to be important in determining whether or not a call is *profitable.* After all, in this rather realistic example, nontravel expenses would be \$208.50 (33 hours times \$6.50 per hour), while travel expenses are \$153.50.

Now some accounting method is needed whereby travel costs can be fairly charged to various calls. One approach, which will be called Method I, is:

1. Add up all travel expenses (in the previous example, this would amount to \$153.50).

2. Divide this amount by the number of calls made. (If a total of 25 calls were made to the three customers, then the costs of calls were \$153.50/25 for this tour.

While this is simple from a computational standpoint, such an approach does not take into account the distances needed to reach each customer. It is biased in favor of the distant customer. A more impartial approach might be the following, which can be called Method II:

1. Add all travel expenses.

2. Add up total mileage travelled during the tour.

3. Divide (1) by (2), giving a per-mile expense of travel.

4. Pro-rate according to the number of calls made to a customer, the distance to that customer, and (3). If five calls were made with customer 1, then the calls made to that customer cost (\$153.50 x 50 mi./ 330 mi.) ÷ 5 each.

With this method, it must be resolved how to allocate the costs associated with the return mileage. A simple but logical step would be to pro-rate return costs (in Figure 1, the costs of the distance from customer 3 to the salesman's home) as follows:

1. Add up all the mileage other than the mileage of the return trip.

2. Pro-rate return mileage on the basis of a customer's one-way mileage as a percentage of the total one-way mileage. For example, customer 1 would be charged with (50 mi./200 mi.) x 130 miles for its share of the return mileage, with the per-mile rate the same as that above.

While this allocation method may seem more fair, it has one especially serious drawback; it gives noticeably different results if the direction of the tour is reversed. Yet in practice the reversal of the tour direction would likely affect neither costs nor sales benefits drastically. For example, if the tour in Figure 1 were reversed, customer 1 would be charged with:

$$\frac{(110 \text{ miles}) + \dfrac{110 \text{ miles}}{(110 + 130 + 40)\text{mi.}} \times 50 \text{ miles}}{(50 + 110 + 130 + 40) \text{ miles}}$$

of the travel expense. This is 39.3 per cent of the total travel costs, as compared with only 82.5/330 (25 per cent) of the travel costs if the tour direction were reversed. This method shall not be considered further.

Now consider Method III;

1. Add all travel expenses.

2. Compute the tour length with:

   (a) a customer included and

   (b) that same customer excluded from the tour.

3. Add these distances for all customers; pro-rate total travel costs on this basis.

This is illustrated in Figure 5.

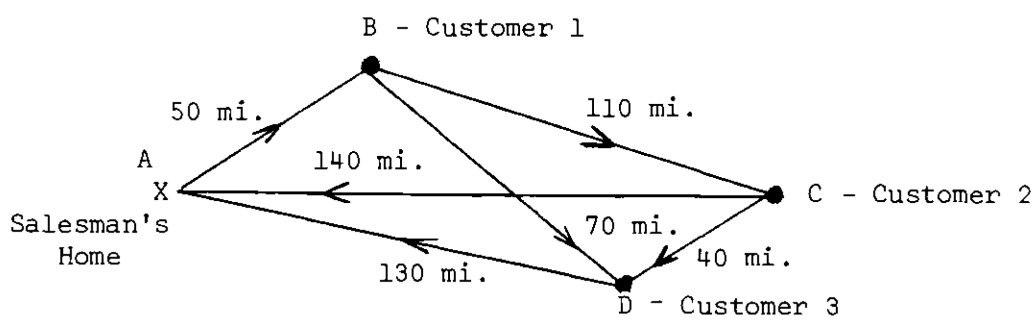


Figure 5. Hypothetical Sales Tour, with Distances for Method III Travel Expense Allocation

The present tour is ABCD. It extends 330 miles. If customer 1 were not visited, the tour would be ACD, and the difference in mileage would be AB + BC - AC, or 50 miles + 110 miles - 140 miles (20 miles). In similar fashion, customer 2 would be charged with 80 miles and customer 3, 30 miles. Thus, 20 miles/(20 + 80 + 30) miles of the travel costs (15.4 per cent) would be charged to customer 1 under this method.

While the partiality of Method I is avoided, and the unrealistic drawback of Method II is not present, there is another disadvantage to this method. Consider Figure 6:

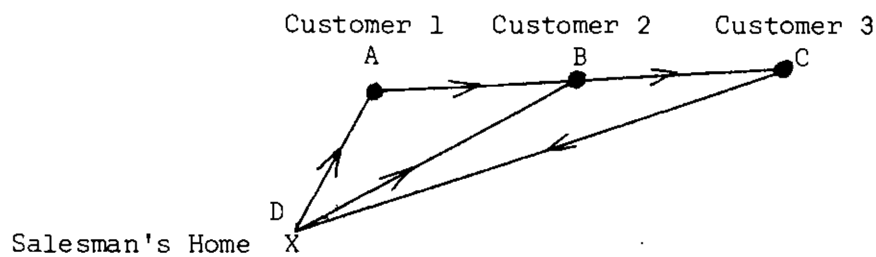


Figure 6. Illustration of Method III Inaccuracy

When three customers lie on one linear route, no mileage is charged under this method to customer 2, the second of the three. If three points are nearly collinear, such as the salesman's home, and customers 1 and 2, the proration is also rather unreasonable. All in all, were the allocation of travel costs of the above tour to be evaluated under Method III, it would be found that customer 3 would be charged with nearly all of the costs. Therefore, Method III will not be considered further.

While any of the above methods would probably seem *fair* from a post-facto accounting standpoint, they each have definite drawbacks for use in forecasting; from that viewpoint, they are not reasonable. Any other travel cost allocation method would very likely be no more reasonable for use in forecasting than the three already presented. To show just how difficult a problem the formulation of the *ideal* cost allocation method is, consider this hypothetical case: A call that is itself rather short in length could so extend a salesman's stay at a given location that he arbitrarily decides to stay over the night rather than drive on home or to another customer. The costs of this overnight stay might well run from $15 to $25; from a practical forecasting, not post-facto, standpoint, this sort of situation is extremely difficult to model. In addition, it must be remembered that all of the above methods allocate on the basis of mileage travelled, not time--this is not in line with conventions already adopted on other items.

Since this is such a troublesome problem, it appears reasonable to discontinue our approach on this basis. Assume instead that average travel costs per mile will not appreciably vary over the long run. Then the total costs of a tour associated with travel might well be estimated to some degree of accuracy, while it may not be possible to realistically formulate travel costs in terms of the individual calls made while on the tour.

In other words, the concept of *marginal* time and/or cost associated with a particular call is not a *sine qua non* in the development of an approximate solution to optimal sales call allocation, since marginal tour costs can instead be estimated fairly accurately in terms of time and/or dollar costs.

More specifically, the multi-customer case lends itself to the same sort of analysis as the one-customer case insofar as determining time-to-call and call-to-sales relationships (Steps One and Two of the example in the one-customer case). Given certain additional constraints relating to tour length and time, then an approach can be made to the sort of results incumbent in a directed search technique that constituted Step Three in the one-customer example.

Computationally, the problem is further simplified, since distances between customers are known (or are easily learned), and these can be converted generally both to the time and dollar cost of the travel portion of a salesman's activities.

Therefore, this *marginal tour* concept will be employed in seeking optimizing algorithms that would effect the same general purpose as *Patternsearch* in Step Three of the example of the one-customer case. First, however, it should be understood specifically what data are given, what initial conditions affect the givens, and what is sought by way of a *solution*. Basically, the givens to this problem reflect those stated in "Mathematical Statement of the General Case" in Chapter I. Specifically, however, the case under study concerns only one salesman's activities, and are assuming that the potential volume of customers and prospects is reflected in the sales-to-call relationships that will be determined; embraced in these relationships also will be a measure of the ability of the salesman to sell the various customers and product lines.

Insofar as the initial conditions of the problem are concerned, all those reflected in "The General Case" except sales performance (3) will be required.

In addition, historically determined average length of a sales call for product will be used as the length of such calls for estimates in future periods. Additional items of information needed are:

(a) maximum permissible tour time, e.g., 40 hours; (b) maximum permissible number of calls for a given product line and customer on a given tour.

Based on this information, it can now be estimated what an *optimal* ordered set of sales calls (by product and customer) would be, and what sales and gross margin dollars (profits) would result from such a strategy. At the same time the judgment and ability of a salesman with respect to a given customer and product can be gauged to some extent; in other words, this analysis can be used as an evaluation aid by sales management.

An Algorithm for the One-Salesman Multi-Customer Case

1. Divide customers for a product into a few groups, so that s_{ij} will be a composite function for these similarly important customers. This will reduce computational aspects of the problem while providing a larger set of observations on which to base s_{ij} functions.

2. Assume two time constraints are known: (a) total time in a single sales period; and (b) maximum allowed tour time.

3. Assume one constraint on calls: There is a known maximum number of calls per tour (by product and customer).

4. Use multiple linear regression and polynomial regression to determine:

(a) $T_t = \psi_{total} + f(n_i)$

(b) $s_{ij} = f(n_{ij})$

This corresponds to Steps One and Two of the example in the one-customer case.

5. Determine the travel time for all $\frac{n^2 - n}{2}$ routes from one of the n customers to another.

6. From s_{ij} functions, determine marginal dollar returns for all possible calls.

7. Compute the average total amount of time per time period presently devoted to travel and travel-related tasks. Subtract this from 2(a). This then is the time spent in directly and indirectly making the sales calls.

8. Adding direct and indirect time per call, divide marginal s_{ij} returns by time per call. These figures represent the expected sales returns per time unit of all calls, excepting travel time.

9. Arrange these in rank order from greatest return per unit time to the least. Multiply these by α_i.

10. It is now possible, through an iterative procedure, to determine *trial* tours. Basically, this procedure involves *using* marginal time-return units figured in step 8 in the *best* fashion (highly dependent on the results of step 9) until a tour's available time span (Figure 2(b)) is taken up with these calls.

11. Establish a minimum acceptable return per unit time figure. Initially, successively add the greatest time-return values from step 9 until the time available in step 7 would have been *used* if only direct and indirect time had been considered.

12. Establish an initial customer.

Choose a customer:

(a) Add round-trip travel time.

(b) Successively chose time-return units, starting with the most *profitable* call, until:

1. The 2(b) constraint would be unheeded if additional calls were made, or

2. Calls of value less than that determined in step 11 would be chosen, or

3. The over-all average return per unit time for calls made to that customer would decline, or

4. The constraint in step 3 would not be heeded (in this case, neglect computing in this call; proceed with the next).

13. Divide the expected return from this sub-tour by the expected time.

14. Do steps 12 and 13 for all customers.

15. Select the customer whose result in step 13 is largest. This is the initial customer of the tour.

16. If condition 12(b) 1 was the limiting factor, a basic tour has been selected; proceed to step 22. Otherwise, some tour time remains to be *used*, and another customer and more calls need to be planned into this tour.

17. Select another customer. Compute the marginal travel time associated with including this customer on the tour. This will be equal to the smallest of:

(a) The time associated with a round-trip to and from that customer and:

(1) the salesman's home, or

(2) the nearest customer previously included in this tour; or

(b) The time difference between the direct route of two previously included customers on the tour and the time of a route beginning with and ending with these two customers but including the new customer, or

(c) The time difference between the direct route between the salesman's home and a previously included customer, and the route beginning and ending with these points, but including the new customer.

18. Repeat Steps 12(b) and 13 for this customer.

19. Repeat steps 17 and 18 for all remaining customers.

20. Select the largest step 13 value; this customer is now added to the tour.

21. Repeat steps 16 through 20 until no more new customers can be added to the tour.

22. The basic tour is now complete. Check sequence of the tour with the Gani[24] algorithm; recompute time remaining, and check steps 16-20 to see if more customers can be added; if so, start this step again.

23. Reduce the figure in step 11 to 0. Repeat steps 18 and 20 to determine if additional calls can be made to a customer previously included on the tour. When this procedure can yield no more additional calls, a final tour itinerary has been calculated.

24. Figure additional tours within the 2(a) timespan.

25. Add up marginal sales and time units; multiply sales by gross margins. These totals represent the optimal call allocations, sales and gross margins. The individual tours give detailed itineraries.

The multi-customer case is noteworthy in that two distinct data sets are of interest:

(1) The optimal expected sales, gross margin and call distribution obtainable from a given set of resources, i.e., with a fixed amount of sales time that can be distributed.

(2) The same figures if manpower is not restricted to that presently available.

The algorithm in the preceding pages effects an approximation of the first set; the solution to the second data set will now be investigated.

Without any understood concept of the *marginal* (per call) travel costs, the solution to the second desired data set can be well approximated, assuming an extension of several existing conditions. If we assume that costs and sales trends could be extended, we can see how many salesmen of the caliber of the existing one could be profitably employed in this sales territory. From a planning standpoint, this might be more valuable than the first data set, since the addition or possibly removal of salesmen from a territory might affect profits more favorably than a shifting of sales efforts by present salesmen.

In addition to the information required for the first data set, the direct costs (all costs other than travel) per hour of the salesman, and the per-mile travel costs must be known; after all, this algorithm employs a marginal return to marginal cost concept. As before, we will assume that information analogous to that obtained in steps 1 and 2 of the one-customer case has been obtained.

An Algorithm for the Multi-Customer Unrestricted Manpower Case

1. Compute an *average* cost per call. This is to act only as a base cost as a starting point in an iterative procedure.

   (a) Using Method I for determining marginal travel costs, figure the historical average cost of travel for calls by product and customer.

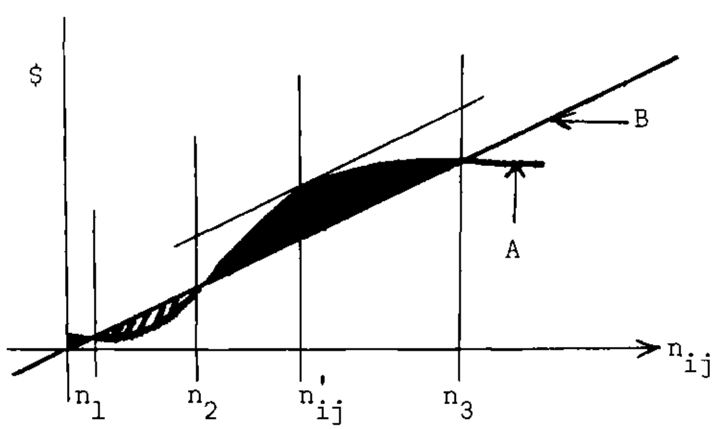
   (b) Add to this the product of the direct hourly sales expense and the sum of the direct and indirect time associated with a call (by product and customer). Call this figure c_{ij}; it represents costs associated with all three aspects of a salesman's activities.

2. Since s_{ij} is an nth order polynomial and c_{ij} is a constant, ordinary differential calculus can be applied to the equation

$$\frac{d}{dn}\int_{o}^{n} (\alpha_i s_{ij} - c_{ij})dn = 0$$

to determine the optimal number of calls for a given product and territory.

Graphically, this is reflected in Figure 7:



NOTE: "A" represents the cumulative expected gross profit ($\alpha_i s_{ij}$) neglecting, however, any sales expenses.

"B" represents the cumulative cost of calls, i.e., sales expenses.

Figure 7. Determination of an Optimal n_{ij} Figure in an Unrestricted Manpower Case.

The shaded portion of the figure represents profits before taxes; the striped portion represents losses. If, for example, n_1 calls were made, the expected profit would be zero. The same is true if n_2 or n_3

calls are made. However, if n'_{ij} calls are made, the gap between profits without sales expenses and sales expenses is greatest. Notice also that the slope of s_{ij} is equal to the slope of the sales expense curve at the point whose x-coordinate is n'_{ij}. This is strictly in accordance with the elementary theorems of differential calculus relating to maxima and minima; all continuously differentiable functions having extreme points between their end points, such as polynomials, obey this behavior. Therefore, the optimal number of calls to make in this case is n'_{ij}. The optimal numbers of calls for other products and customers should also be determined in the same manner at this step.

3. (a) Assume that these calls are to be equally spaced time-wise within the total time period, and that we should not prohibit a given call from being made anytime during the total time period.

For example, if the total time period is T and a certain n'_{ij} equals 5, then equidistant time spacing would dictate the mean times of 0.1T, 0.3T, 0.5T, 0.7T and 0.9T for the five calls. Since we would not restrict a call from being made anytime during the time period, *tolerance* limits of $\pm \frac{1}{2}\, T/n_{ij}$ surround these five calls.

(b) Using a table of random numbers, establish an exact time for each call.

If the first random number selected was 0.0150, then the first call is to be made at (0.0150) $(T)(1)/n_{ij}$, or 0.0300T. Similarly, if the second random number was 0.8500, then the second call on behalf of product i to customer j is to be made at (0.8500) $2T/n_{ij}$, or 0.3700T.

(c) On the basis of the findings of (3)b, and the longest permissible tour time, categorize all calls into a given tour.

If for example, the total time period is six months, this might represent 25 working weeks, and the longest permissible tour might be 40 working hours. All calls to be made from time 0.0000T to 0.0400T must be made in *sales trips* no longer than 40 hours that began at time 0.0000T. In this case it is assumed that all manpower necessary to service these call *needs* is available. The costs of these calls, the amount of manpower necessary, and the sequencing of these calls will now be determined.

4. Determine the sequencing of calls, and the manpower requirements of a tour that meets a subset of these calls *needed*.

If this 40-hour tour time limit is designated T_x, then a tour that meets the requirements of our sales trips could be planned in the following way:

A. Starting with the salesman's home, select the nearest customer with *needed* calls.

1. Make all *needed* calls to that customer.

2. Compute $\phi_{ij} + \theta_{ij} = t_{ij} + \beta_{ij}$

3. Add in the travel time associated with visiting that customer. This is the smallest of:

(1) The round-trip time involved in a tour having the customer and

(a) a previously-visited customer, or

(b) the salesman's home
as an end point,

or

(2) the additional travel time necessitated by including this customer in the tour.

NOTE: On the first pass of this step, alternative (1)(b) is required.

4. Add the time in steps A2 and A3 to that previously *used* on this tour. If this equals or is less than T_x, select another customer and begin this iteration at step A1.

B. As long as the making of all *needed* calls at a customer is possible, continue steps A1 through A4. If this is not possible, the following method will determine if part of the *needed* calls can be made at any other customers:

1. Select a customer having *needed* calls who has not been *visited* (included) previously on this tour.

2. Compute the marginal travel time associated with including this customer in the tour. This will be equal to the smallest of:

(a) Time associated with a round-trip to and from that customer and:

(1) the salesman's home, or

(2) the nearest customer previously included in this tour;

or

(b) The time difference between the direct route of two previously included customers on the tour and the time of a route beginning and ending with these customers but including this new customer;

or

(c) The time difference between the direct route between the salesman's home and a previously-included customer, and the route beginning and ending with these points, but including this new customer.

Add this marginal travel time to that of the last feasible time figured in step 4 A.4. Call this value λ'.

3. Repeat steps 4 B.1 through 4 B.2 until no other customers remain. If for any of these, $\lambda' < T_x$, select the smallest λ'; otherwise, proceed with step 4 B.6.

4. Compute marginal *per call time values* for the direct and indirect time associated with the *needed* calls for that customer.

5. Arrange these in ascending order in terms of their time length; subtract their times successively until no more calls can be made in this manner. While no further calls can likely be made, retrace this procedure from step 4 B.1 on to this point.

6. Now resequence this tour. Notice that this constitutes a traveling salesman problem. It would appear the method of Gani[25] is applicable. Then check and see if this resequencing

results in appreciable additional free time, i.e., the time required for the tour is appreciably different from T_x. If it is, begin the procedure at step A1, after choosing a possible customer to call upon who has either not been solicited on this tour or who has only some *needs* met.

7. After this has been done, a single tour of time $\leq T_x$ has been selected.

5. If other calls *should* be made during this subperiod (week), calculate additional tours until all *needed* calls in the subperiod have been made; in other words, repeat step 4 for remaining customers and calls.

6. Proceed to satisfy call requirements for other subperiods (weeks) until all *needed* calls in the entire time period (T) have been made.

7. Figure the time involved in the travel represented by this group of tours. Using prior cost per mile estimates, compute the estimated total travel costs for the time period T. Divide this by the number of calls to be made in the time period. This is a new *average* travel cost per call; this can easily be determined if the new c_{ij} (which will be designated c'_{ij}) is much changed from that found in (1). If it is, begin a new iteration.

For example, c'_{ij} might be considered significantly different from c_{ij} if it is either 5 per cent above or below the previous figure. Naturally this would cause the n'_{ij} for the various products and customers to be affected; this would affect the number of tours and sequencing. Therefore, if c'_{ij} is different the entire process starting with step 2 must be redone.

Reviewing what has been accomplished in the course of th steps, it is seen that:

1. The addition of all tour times will give an indication of the general manpower requirements for this unrestricted case.

2. The largest number of tours in one subperiod is the maximum number of salesmen needed in this territory at one time, if the results of one salesman's experiences can be projected to other salesmen's work.

3. In the course of computing the various n'_{ij}, the optimal sales and gross margins for each product and customer were calculated; the costs of calls were also established.

4. The sequencing of each tour has been done.

While it might be tempting to either reject or accept such an answer in its totality (if it is agreed that the general assumptions upon which it was constructed seem reasonable), it is probably unwise to do so for any real example. It must be remembered that a number of simplifying assumptions have been made, and that tour time length is in reality not so inflexible as it is figured in the above algorithm. Also mean time, sales and cost estimates were used; these will vary somewhat from a particular set of such values in any real data set. Nevertheless, it does represent a relatively unbiased forecast.

In summary, the second data set has been figured; it *may*, after careful analysis, prove to be a good estimator of manpower requirements for a territory.

CHAPTER IV

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

1. Given past data on sales, sales calls, customer placement, and salesman availability generally available to a firm, approximate solutions towards optimizing the sales call *mix* according to a given objective function does seem at least theoretically possible.

2. The determination of an optimal *mix* by the methods used in this research would require extensive computer analysis of given data in order to overcome severe combinatorial and computational burdens.

3. Forecaster bias could seriously alter the results of computations of this nature; therefore, the accuracy of the methods used in this research is dependent on: (1) amount of past data on sales and sales calls; (2) variability of effectiveness of sales calls; (3) variation in competitive efforts; (4) variations in the length of sales calls, and the indirect time associated with sales calls; (5) economic and other factors beyond the control of the salesman; and (6) forecast assumptions concerning sales trends.

Recommendations

1. Techniques for simplifying the combinatorial aspects incumbent in the multi-customer algorithms should be investigated.

2. The multi-customer algorithms should be formulated into computer programs.

3. The techniques of this research should be applied in the analysis of data of actual salesmen.

4. The sensitivity and reliability of the models presented should be tested with respect to changes in forecaster predictions and considerations beyond the control of the firm.

BIBLIOGRAPHY

LITERATURE CITED

1. Alan Mercer, "Applications of Operations Research in Marketing," *Operational Research Quarterly*, Volume 17, Number 3 (September, 1966), pp. 235-252.

2. R. L. Ackoff and P. E. Green, "Status Report: The Role of Personal Selling in the Promotional Mix." *Report of the Management Science Center*, Wharton School of Finance and Commerce, University of Pennsylvania (1965).

3. William Lazar, "Operations Research in Marketing--Some Critical Comments," *Operational Research Quarterly*, Volume 15, Number 3 (September, 1965), pp. 275-278.

4. William Lazar, *Ibid.*

5. Alan Mercer, *op. cit.*

6. R. L. Ackoff and P. E. Green, *op. cit.*

7. Robert S. Reichard, *Practical Techniques of Sales Forecasting*, McGraw-Hill Book Company, New York (1966).

8. B. V. Wagle, "A Review of Two Statistical Aids in Forecasting," *Statistician*, Volume 15, Number 2 (1965), pp. 191-196.

9. P. J. Harrison and O. L. Davies, "The Use of Cumulative Sum (Cusum) Techniques for the Control of Routine Forecasts of Product Demand," *Operations Research*, Volume 12, Number 2 (March-April, 1964), pp. 325-333.

10. Isidor Heller, "The Traveling-Salesman's Problem, Part I: Basic Facts," The George Washington University Logistics Research Project, June, 1954.

11. H. W. Kuhn, "The Traveling-Salesman Problem," *Proceedings of the Sixth Symposium in Applied Mathematics* (American Mathematical Society), McGraw-Hill Book Company, New York (1956).

12. C. E. Miller, A. W. Tucker, R. A. Zemlin, "Integer Programming Formulation of Traveling Salesman Problems," *Journal of the Association of Computing Machinery*, Volume 7, Number 4.

13. G. Dantzig, R. Fulkerson, and S. Johnson, "Solution of a Large-Scale Traveling-Salesman Problem," *Journal of the Operations Research Society of America,* Volume 2, Number 2 (1954), pp. 393-410.

14. John D. C. Little, Katta G. Murty, Dura W. Sweeney, Caroline Karel, "An Algorithm for the Traveling-Salesman Problem," Operations Research, Volume 11, Number 6 (November-December, 1963), pp. 972-989.

15. Frank M. Bass and Ronald T. Lonsdale, "An Exploration of Linear Programming in Media Selection," *Journal of Marketing, Research,* Volume 3, Number 2 (May, 1966), pp. 179-188.

16. Ralph L. Day, "Linear Programming in Media Selection," *Journal of Advertising Research,* Volume 2, Number 2 (June, 1962), pp. 40-44.

17. Frank J. Charvat and W. T. Whitman, "Marketing Management," Simmons-Boardman Publishing Corporation, New York, 1964.

18. Charles H. Sevin, "Marketing Productivity Analysis," McGraw-Hill Book Company, New York (1965), pp. 133ff.

19. Arthur A. Brown, Frank T. Hulswit, and John D. Kettelle, "A Study of Sales Operation," *Operations Research,* Volume 4, Number 3 (June, 1956), pp. 296-308.

20. Clark Waid, Donald F. Clark, and Russell L. Ackoff, "Allocation of Sales Effort in the Lamp Division of the General Electric Company," *Operations Research,* Volume 4, Number 6 (December, 1956), pp. 629-647.

21. Arthur A. Brown, Frank T. Hulswit, and John D. Ketelle, *op. cit.*

22. Clark Waid, Donald F. Clark, and Russell L. Ackoff, *op. cit.*

23. H. P. Künzi, and Wilhelm Krelle, *Nonlinear Programming,* Blaisdell Publishing Company, Waltham, Massachusetts, 1966.

24. A. Z. Gani, *Evaluation of Alternative Materials Handling Flow Pattern,* an unpublished paper submitted as an I.E. 704 project, Georgia Institute of Technology, April, 1965.

25. A. Z. Gani, *op. cit.*

Other References

William W. Cooley and Paul R. Lohnes, *Multivariate Procedures for the Behavioral Sciences,* John Wiley & Sons, Inc., New York, 1962.

R. J. Duffin, Elmor L. Peterson, and Clarence Zener, *Geometric Programming,* John Wiley & Sons, Inc., New York, 1967.

Walter J. Talley, "How to Design Sales Territories," *Journal of Marketing,* Volume 25, Number 3 (January, 1960), pp. 7-13.

E. J. Williams, *Regression Analysis,* John Wiley & Sons, Inc., New York, 1959.