MODELING THE CUSTOMER-FIRM RELATIONSHIP WITH
SALES PROFILES: ESTIMATION AND APPLICATION

by

Glen L. Urban*
Magid Ibrahim**

Massachusetts Institute of Technology, Cambridge, MA
Alfred P. Sloan School of Management, W.P. #1682-85

*Alfred P. Sloan School of Management
**Information Resources, Inc.

ABSTRACT

This paper proposes that individual client sales histories or what we term "sales profiles" be analyzed to describe the overall structure of the relationship between an industrial firm and its customers. We propose a clustering methdology to: (1) group customers who have similar profiles of sales over time, (2) estimate the shape of the average profile in each group, and (3) examine the stationarity of profiles. A flow model is formulated to examine the strategic implications of the derived sales profiles. An application of the methodology to a computer software company is presented. The paper closes with an indication of future research needs and the implications of the model in linking strategic and annual plans.

INTRODUCTION

Standard financial reporting systems for industrial products display
sales and profits by month; often these are classified by division, region,
product, and/or salesman. While these figures of merit are useful in
planning and control, they do not assess the quality of the critical
continuing relationship between the firm and its base of customers. Failure
to understand the relationship may lead to problems and missed
opportunities. For example, total sales may be increasing because of the
arrival of many new customers, but if sales to older customers are
decreasing, a possible long run overall sales decline may go unrecognized.
It is important to know if total sales are comprised of a number of segments
of customers who vary by volume per period and sales growth rate. In this
paper we propose a segmentation method based on the similarities of customer
sales patterns over time to answer this question and examine the
implications of such a segmentation scheme on marketing strategy formulation.

Figure 1a displays several patterns of the sales relationship between a
customer and the producer. The sales of a buying firm per period are
plotted against the number of periods the firm has been a client. We call
such curves "sales profiles". They represent the sales to a client over the
history of that client's relationship with the firm. The sales profile
labeled number one appears to be a healty pattern; it is growing and is at a
high level. Profile two represents a similar growth pattern but at a lower
level. Profile three shows a client whose sales grew and then declined.
This could reflect a life cycle of use at the client company level.
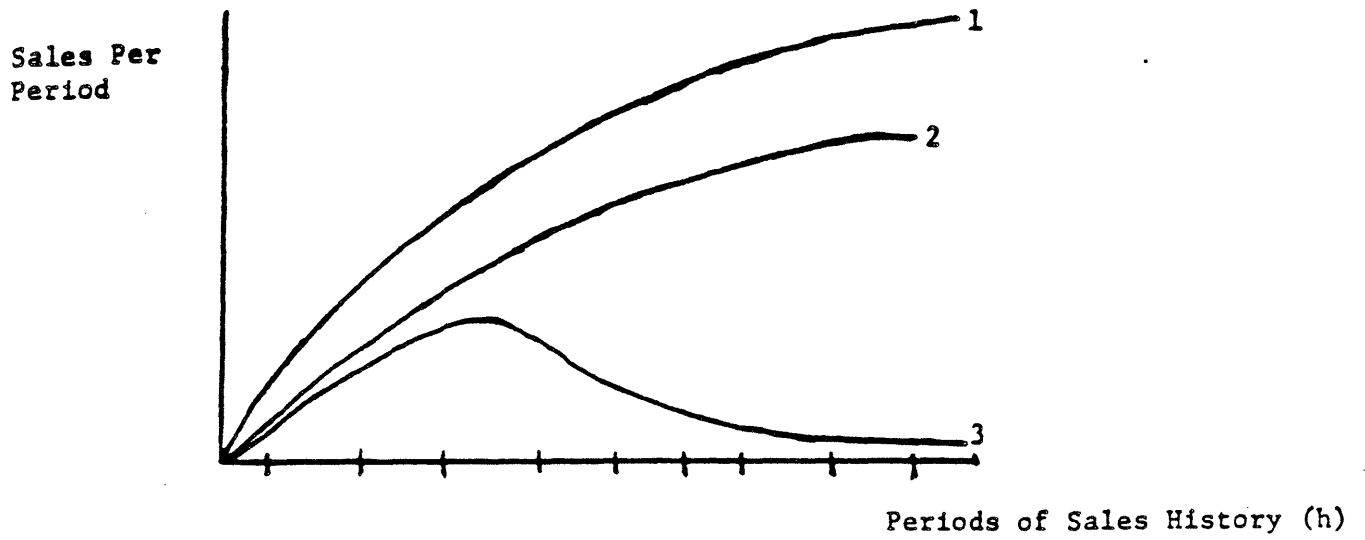
FIGURE 1 - SALES PROFILES

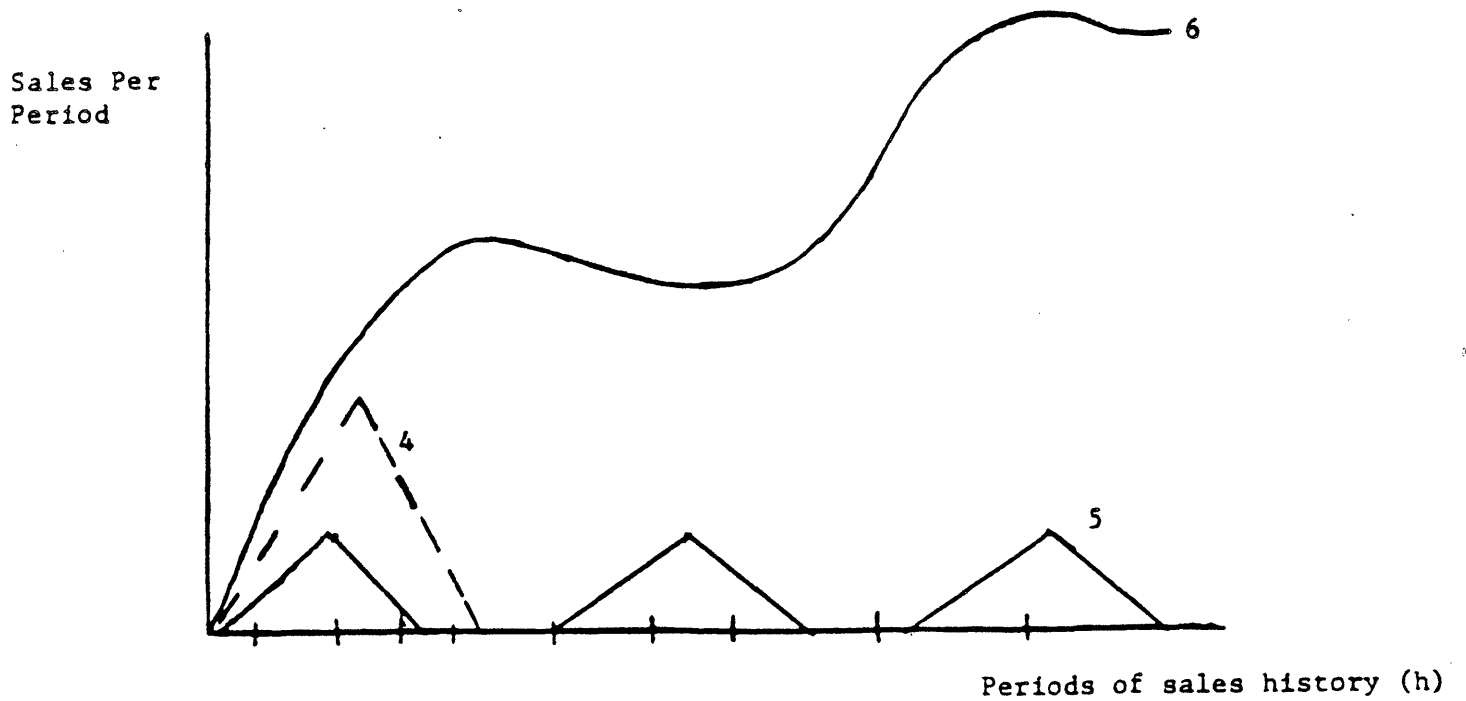Figure 1(a) - Typical Sales Patterns

Figure 1(b) - Other Possible Patterns

Figure 1b shows some alternative sales patterns. Curve four displays a customer who bought only once and pattern five shows a firm who made a series of discrete purchases over a period of time. Sales profile six depicts a cycle around a growth trend.

With these individual customer sales profiles, the customer-firm relationship can be diagnosed. An important question is why the profiles differ. Is it differences in the characteristics of the firms (e.g., size, type of industry), the importance of product attributes (e.g., quality, service, price), competition, or the result of the original firms own sales force allocation strategy? One attribute that is likely to be associated with the profile is the product buying characteristic. Buy/rebuy products are likely to be described by profiles 1, 2, or 3. Modified rebuy products fit profile 5 and one time buys could be shown in profile 4.

In today's market, profiles such as number 6 in figure 1b are becoming common. The customer buys not just one product, but a continuing service. Jet engines are a large purchase item, but spare parts provide a continuing stream of revenue and the source of most of the profit for the manufacturers. Automated production tools are major purchases, but a successful manufacturer attempts to make repeat sales to its clients. Expansion of the production capacity or sales of other automated machines represent a potential basis for a continuing pattern of sales. Another example is computers where the continuing sales relationship is based on service, software, training, expansion and upgrading for new technologies. A final example is in the communication industry where a firm may sell a range of products (e.g. phones, PBX, video conferencing, networking, and data transmission) and strive to build a continuing sales relationship with

each client. Emphasis remains on a specific sales, but major attention is
being directed at building a positive sales relationship with a client.
Understanding and putting in place strategies to nurture this relationship
is an important marketing function. If correctly managed, a healthy, and
growing sales profile should result.

Grouping individual customers with similar sales profiles provides a
market segmentation based on the client-firm relationship. Managerial
insights and decision implications may be gained through such a segmentation
and a management science model can be built based around these segment sales
profiles to simulate decision implications.

In this paper we propose a clustering method to define segments of
customers with similar sales profiles. These profiles are the basic
elements utilized in a proposed model to forecast sales and evaluate pricing
and sales force decisions. We present an application of sales profiles and
the model to a firm in the computer software business. We close with a
consideration of the implications for industrial marketing and future
research needs.

ESTIMATION OF SALES PROFILES

The data necessary to define individual sales profiles are readily
available in the firm's order files. Table 1 depicts some typical client
sales data for each calendar period. Table 2 shifts the time scale so that
first period of sales history is the first entry. Data entries end for a
client when the current calendar time period is reached. For example, in

TABLE 1 — ILLUSTRATIVE SALES PROFILE DATA

Time Periods (t)

	1980				1981				1982				1983				1984			
	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q
Client A	5	10	8	12	20	18	15	24	25	24	28	30	32	34	40	32	34	36	40	38
Client B								3	5	8	10	20	11	10	8	10	12	15	13	18
Client C																	6	15	10	20
Client D										15	12	10	8	6	6	4	2	0	1	0

TABLE 2 — SALES UNITES BY CLIENT VERSUS CALENDAR

Periods of Sales History (h)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Client A	5	10	8	12	20	18	15	24	25	24	28	30	32	34	40	32	34	36	40	35
Client B	3	5	8	10	20	11	10	8	10	12	15	13	18							
Client C	6	15	10	20																
Client D	15	12	10	8	6	6	4	2	0	1	0									

Table 1, client A began purchasing the first quarter of 1980 and continued to the current period (fourth quarter 1984). In table 2, it is shown as a client with 20 periods of data. Client B began purchasing in the fourth quarter of 1981 and is currently still a customer so the client is shown with 13 quarters of data. Client C first purchased in the first quarter of 1984 and is shown in table 2 with 4 quarters of data. Typically the matrix would be very large because many customers exist and many periods of sales history exist.

The individual sales profiles data can be cluster analyzed to determine the number of segments and their average profiles. A number of clustering algorithms are available (Everitt, 1974) and a number of proximity measures can be used (Green and Rao 1972). In the cluster analysis only clients with at least some minimum number of quarters of sales history are included. For example, requiring at least eleven quarters of data would call for clustering over clients A, B, and D but not C in table 2. In most cases, many customers have more than the minimum number of observations. The output at this point is a set of clusters and their average profiles that best describe the disaggregate sales history data (see Table 2).

The cluster analysis can be done over alternate samples based in the calendar date of the first sale to examine the stationarity of the profiles. We may cluster analyze clients with 5 years of sales history that first became a customer in 1975 and compare the results to a cluster analysis of clients with 5 years of sales history that first purchased in 1980. Comparison of the results allows an assessment of the similarity of the sales profiles. If they are alike, we may want to assume in forecasting that they will remain similar. If they are different, accurate forecasting

would require a mechanism for updating the profiles for clients that start in 1985. For example, if the clients who started in 1975 can be grouped into two life cycle shaped profiles (say for simplicity, one for large sales and one for small sales rates) and if those who first purchased in 1980 show two profiles identical to the profiles of those who started in 1975, we might assume this stability continues and apply these profiles to future new customers. If, however, the customers who began in 1980 show, for example, a shorter life cycle in both profiles it may be appropriate to forecast a further shortening of the life cycle for those customers starting in 1985.

Another nonstationarity may occur in the fraction of clients who are characterized by each of the profiles. In the example above, the fraction of clients who are described by the large sales rate life cycle may be increasing. In this case future forecasts require an assumption as to whether this trend towards more customers being classified in the large sales rate life cycle profile will continue. Forecasting the trends in sales profiles and the number of customers in each profile requires managerial judgement (see application below for an example of consideration of nonstationarity and its implications).

If the profiles are stationarity, increasing the minimum value on the number of required periods of sales history can increase the duration of the estimated sales profiles. For example, one analysis may be done on a first sample of 500 clients with 20 quarters of experience and another second analysis may be done on the 350 clients with 30 quarters of experience. These profiles can be combined if the profiles do not vary over the first 20 quarters and it is assumed that the underlying causal phenomena are similar

across the two groups. That is, the duration of the sales profile obtained from the first analysis of 20 quarters can be extended from 20 to 30 quarters by the second analysis. If the profiles are not similar and stationary, extension of the duration of the sales profiles for more recent customers would require managerial judgment.

Few statistics are available to aid in determining the best number of clusters or the significance of differences between clusters of subsamples. Most clustering algorithms display the percentage of variation explained by the clustering and some provide a cophenetic correlation, but formal significance testing is not possible[1]. The percent variation explained by adding another cluster is a good indicator, but one must use judgment to select the best number of clusters.

One informal test can be based on comparing the values of the sales profiles across clusters. If the average values represented by the clusters are "very" different, it would seem that separate clusters are appropriate. An analysis of variance can be done for each period to compare differences between profiles. This is not a formally correct test because clustering inflates the F statistic, but it may help in making judgments. For example, if F's are very large for all periods, it might suggest that differences exist. But if the F's are low, indeed, separate profiles are not justified.

---

[1]See Hartigan (1978) and Milligan and Cooper (1983) for a discussion of the statistical properties in cluster analysis.

After determining the best number of clusters and the average sales
profiles in each cluster, the reasons for differences between profiles can
be examined. Discriminant analysis could be used to test the association of
variables such as client size and with membership in a cluster. This stat-
istical analysis can be supplemented by an examination by managers of the
firms assigned to specific clusters. This may elicit hypotheses on why
specific firms that had similar profiles were assigned to the same cluster
segment. Similarily a review of the fraction of clients in each cluster and
the shapes of the sale profiles may lead to additional managerial diagnosis
and problem finding.

MODEL STRUCTURE

The managerial implications of sales profiles can be explored with a
quantitative model. In this section, we describe a customer flow model
(Urban and Hauser 1980) based on sales profiles and outline how it can be
used for sales forecasting and sales force size and pricing decision support.

OVERALL CUSTOMER FLOW

The basic structure begins with sales profiles to characterize existing
customers and then models the inflow of new customers as the results of
marketing effort (see figure two). Sales to existing customers are
represented by values from the profile for each set of customers' and the
number of periods of sales history for each customer. New customers are
added to the profile segments and assigned to the first period of profile.

Summing up sales over the new and old all customers in the profiles for the relevant period gives a sales forecast.

Formally, let us denote the number of customers in a sales profile segment p at time t with h periods of sales history as $N_{p,t,h}$. The number of new customers in each profile is:
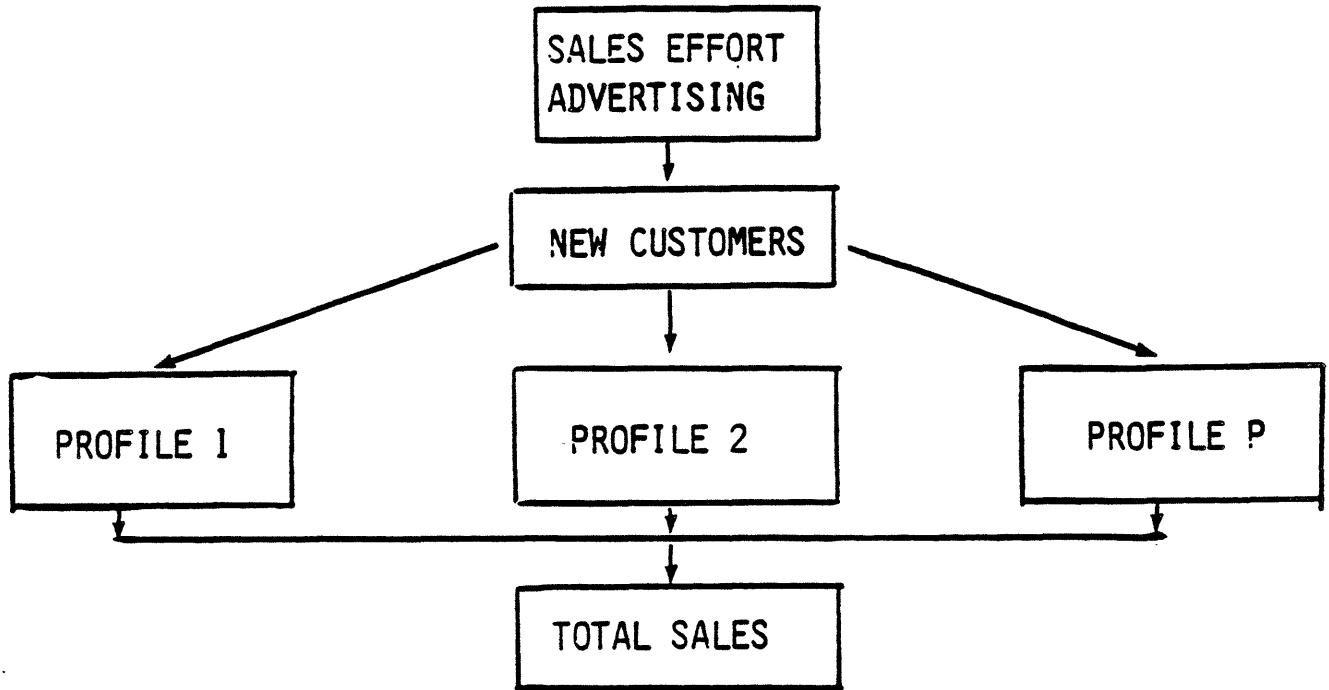
(1) $\quad N_{p,t,1} = F_p W_t$

$\quad N_{p,t,1}$ = number of customers in profile p at time t in their first period of sales history.

$\quad W_t$ = number of new customers in period t (t=1, 2, ... T)

$\quad F_p$ = fraction of new customers described by profile p (p=1, 2, ...P).

Let the average sales for customers in profile p with a history of h periods be the sales profile and be denoted by $S_{p,h}$. The sales revenue in period t for a profile p are:

(2) $\quad R_{p,t} = \sum_{h=1}^{t} N_{p,t,h} S_{p,h}$

$\quad R_{p,t}$ = sales revenue at period t for customers in profile p

$\quad N_{p,t,h}$ = number of customers in profile p at time t with h periods of sales history.

$\quad S_{p,h}$ = sales of customers in profile p with h periods of sales history h (h=1, 2... H).

The summation is from h=1 to t because we establish the convention that t=1 at h=1 for the client segment with the longest sales history. Therefore summing h=1 to t in equation 2 adds together all clients as of time t for all levels of sales history from the youngest (h=1) to the oldest (h=t).

FIGURE 2: CUSTOMER FLOW

```
                        ┌──────────────────┐
                        │  SALES EFFORT    │
                        │  ADVERTISING     │
                        └────────┬─────────┘
                                 │
                                 ▼
                        ┌──────────────────┐
              ┌─────────│  NEW CUSTOMERS   │─────────┐
              │         └────────┬─────────┘         │
              ▼                  ▼                    ▼
     ┌───────────────┐  ┌───────────────┐   ┌───────────────┐
     │   PROFILE 1   │  │   PROFILE 2   │   │   PROFILE P   │
     └───────┬───────┘  └───────┬───────┘   └───────┬───────┘
             │                  │                   │
             └──────────────────┼───────────────────┘
                                ▼
                        ┌───────────────┐
                        │  TOTAL SALES  │
                        └───────────────┘
```

Total revenue in period t is:

$$(3) \quad TR_t = \sum_{p=1}^{P} R_{p,t}$$

Equations 1 to 3 can be used to forecast future sales given estimates of new customer arrivals (W_t), sales profiles ($S_{p,h}$), and the fraction of customers in each sales profile segment (F_p). Before making a forecast, it is wise to use past data on new customer arrivals (W_t), fraction of new customers in each profile (F_p), and sales profiles ($S_{p,h}$) in equation 1 to 3 to calculate the predicted sales and compare them to what was _actually_ observed for total revenue. If the model does not pass this minimum fitting test it should not be used for forecasting.

NEW CUSTOMER GENERATION

The basic flows can be modified to reflect marketing and personal selling resources. The number of customers obtained is usually a function of the number of salesmen, sales calls by each salesman and the average number of calls necessary to obtain a new client. The number of new clients is:

$$(4) \quad W_t = \sum_{i=1}^{I} C_{t,i} / A(a_i)$$

W_t = number of new customers in period t.

$C_{t,i}$ = number of sales calls on new customers by salesman i ($i = 1,2,\ldots I$) in period t

$A(a_i)$ = average number of calls by salesman i ($i = 1, 2, \ldots I$) with a years of experience ($a = 1, 2, \ldots A$) necessary to obtain a new customer.

I = total number of salesmen.

Because of the high variance in sales performance we consider different salesmen (i) and their years of experience (a). The variable "a" reflects the increasing effectiveness of salesmen as they learn the product and gain expertise. The number of calls a salesman must make to gain a new client should decline as he/she gains experience. If the firm to be modeled has a small sales force (less than 50) each salesman is treated individually. Aggregate classes of salesmen are developed if the sales force is large. The aggregation would be based on groups of salesmen with similar performance and experience. Each group would be denoted by a value of the i subscript.

The number of calls a salesman makes on new clients $(C_{t,i})$ depends on the amount of prospecting time he/she has available. Three major activities that fill the salesman's time are: (1) account maintenance and service, (2) administration and training, and (3) calls on new customers. We assume calls on new customers are made after maintaining existing accounts and completing training and administrative duties. The number of calls on new consumers $(C_{t,i})$ is:

(5) $\qquad C_{t,i} = (HT_i - HM_{t,i} - HA_{t,i}) \,/\, AC_i$

$\qquad HT_i \qquad$ = Total hours worked per period for salesman i.

$\qquad HM_{t,i} \qquad$ = Hours on account maintenance in period t by salesman i.

$\qquad HA_{t,i} \qquad$ = Hours on administrative and training in period t by salesman i.

$\qquad AC_i \qquad$ = Average hours necessary for a sales call and travel to customer for salesman i.

If the time spent on service of existing accounts increases $(HM_{t,i})$, the number of hours available for calling on potential new clients decreases

- 13 -

(eq. 5) and therefore the number of new clients declines (eq. 4). The time

spent on maintaining an account and thereby moving it along its sales profile

$(HM_{t,i})$ is:

(6) $\quad HM_{t,i} \quad = \quad \sum_{p,h} \quad N_{p,t,h,i} \quad M_{p,h,i}$

where $N_{p,t,h,i}$ = number of clients in profile p at time t with h periods

of sales history that are assigned to salesman i.

$M_{p,h,i}$ = number of hours per period necessary to sustain sales pro-

file level p for clients with h periods of history for

salesman i.

Administration time $(HA_{t,i})$ is the fixed time necessary per period for

reporting and planning plus the time for training new salesmen. If new

salesmen are hired we increment the administrative time by the allocation of

training responsibilities assigned to specific salesman i. We assume after

initial corporate training, field training responsibilities are assigned to

the "best" salesmen. "Best" is operationalized in this model by the lowest

number of calls necessary to sell a new account $(A(a_i)$ in eq. 4). This

assignment rule produces a feedback as the salesforce is increased. If new

hiring is done, sales training time for the best salesman increases and by

equation 5 calls on new customers decline for this salesman. This decline

of new sales is greater if movement of existing clients down their sales

profiles increases account maintenance demands $(HM_{t,i})$. Existing account

sales increases and new hiring may reduce the number of new customers in the

short run due to a reduction in calls on new customers $(C_{t,i})$ by existing

salesmen but it should increase it in the long run as new salesmen are

trained and become productive.

PROFIT ACCOUNTING

After sales have been determined by the generation of new customers and the aging of existing customers down the sales profiles, profits can be calculated:

$$(7) \quad PR_t = \sum_p (R_{p,t} (1-V) - M_{p,t}) - KM_t - KO_t$$

PR_t = profit in period t

$R_{p,t}$ = revenue from customers in profile p in period t (eq. 2).

V = variable cost as a fraction of revenue.

$(1-V)$ = contribution margin fraction.

$M_{p,t}$ = sales force expenditure allocated for existing customers in profile p at time t

$$= \sum_{i,h} N_{p,t,h,i} \, M_{p,h,i} \, B_i$$

B_i = cost per hour of salesman i's time.

KM_t = fixed marketing costs in period t.

KO_t = other fixed costs in period t.

The fixed marketing cost (KM_t) would reflect advertising, direct mail, promotion, sales time not allocated to account maintenance, and new salesmen in corporate training. After corporate training, salesmen are assigned a subscript i value and an experience level a. They are counted in the number of selling hours and calls on new customers (eq. 5) and influence the resulting number of new customers (eq. 4). As the new salesmen gain experience their productivity increases ($A(a_i)$ eq. 4). In some cases the mix of products or the margins in specific profiles will vary. The revenue is calculated for each profile ($R_{p,h}$) so pricing and mix differences are

captured, but in some cases, the variable cost fraction (V) may have to be subscripted by p and specified for each profile to reflect cost differences.

If an experience curve is present the variable cost (V) can be systematically reduced over time as cumulative volume increases.

$$(8) \quad V_t = V_0 \left(\frac{Y_t}{Y_0}\right)^{-\alpha}$$

V_t = variable unit cost at time t.

Y_t = cumulative past sales as of period t.

Y_0 = initial level of cumulative sales level.

V_0 = initial level of variable cost at initial cumulative sales. level Y_0.

α = experience parameter.

As sales increase, costs decline via equation 8 and simulation can indicate the profit impact.

The relatively simple model proposed here enables a wide range of simulation scenarios. Forecasts can be made by projecting past profiles ahead with existing sales strategies. The sensitivity of sales to shifts in profile ($S_{p,h}$ in eq. 2) or in the mix of new customers (F_p, eq. 1) can be evaluated over a future planning period. These simulations would reflect dynamics of diversion of time for sales training, sales experience, production experience and client growth. Simulations can be run with more sales effort (adding salesmen i=I+1), new productivity (change $A(a_i)$ in eq. 4), new targeting (modify F_p, eq. 1), or new methods of account maintenance (revise $HM_{t,i}$, eq. 5, or $M_{p,h,i}$, eq. 6). Competitive changes

could be reflected in reduction of sales productivity ($A(a_i)$ in eq. 4), revenue ($S_{p,h}$, eq. 2) or margin (1-V in eq. 7), or by an increase in marketing fixed costs (KM_t, eq. 7). Profits over time can be calculated for various sales costs (B_i, KM_t, eq. 7) and contribution margin fractions (1-V in eq. 7). The model is a decision aid that enables managers to use empirically derived sales profiles and judgment to make forecasts and evaluate future decision alternatives.

FLOW MODEL EXTENSIONS

The simple macro flow structure can be customized to individual decision situations. For example, an index to reflect advertising (IA) could be multiplied times the average time necessary to sell a new customer ($A(a_i)$ in eq. 4) to reflect the ability of advertising to shorten the selling cycle. IA would be 1.0 at its past base level, less than one for higher levels of spending, and greater than one for lower levels. The value of this index could be set based on judgment or by the results of field experiments.

Pricing: An evolutionary step in increasing the model's flexibility and usefulness is the inclusion of price as a variable. When prices change, three effects can occur:

 a) More new clients may be attracted if prices are reduced and vice versa.

 b) The revenues generated by clients in a specific profile can change, reflecting the new revenue value for each unit sold and the change

in number of units sold as the result of the price elasticity. For instance, the units sold might go up due to a price cut, but the net impact on revenues depends on whether the price elasticity is high enough to compensate for the revenue loss from reduced prices. The impact of a price change on the shape of a revenue profile is modeled with an index with a value of 1 at the current prices and a value different than 1 when the price increase takes effect.

c) A pricing change could also result in shifting a fraction of clients from one profile to another. For example, some buying firm may lease equipment for a period of time and then purchase it. This would be modeled by shifting a fraction of clients with a sales history of h years from a sales profile that reflects leasing (e.g. curve 2 in figure 1a) to a new profile to represent a large purchase (e.g. curve 4 in figure 1b).

Mathematically this is simply:

(9) $N_{p',t,h} = N_{p',t-1,h} + T_{p,p',h} \; N_{p,t,h}$

$N_{p,t,h}$ = number of clients in profile p at time t with h period of sales history.

$T_{p,p',h}$ = fraction of clients who make a transition from profile p to new profile p' when they have h periods of sales history.

The transition matrix $T_{p,p',h}$ is a function of price, π. It can be operationalized by using an index $I(\pi)_{p,p',h}$ which would have a value of 1.0 at the reference price, π_0, greater than one at lower prices, and lower than one at higher prices.

Investment in new products: Development expenditures can be represented by increases in fixed costs (KO_t in eq. 7) over the specific investment periods. Subsequent sales charges can be modeled by adding a new profile to reflect the sales of the new product to new clients (eq. 2 for P+1), by inserting a new profile to which old customers can switch (eq. 9 with p' replaced by P+1), or the direct modification of sales profiles at the specific time ($S_{p,h,t}$).

The flow structure allows a flexible path for evolution from simple sales forecasting (eq. 1, 2, 3) to consideration of selling, advertising, and pricing levels (eq. 4-9). In all cases the sales profiles and the aging of clients down them is the underlying phenomenon.

APPLICATION

The proposed model and sales profiling methodology has been applied to one division of a highly diversified large company. This division sells computer software for management information systems. The software can be supplied on a time sharing computing basis or sold outright for installation on client "in-house" computers. The division also sells, but does not manufacture computers to support the software for "in-house" installation. Consulting services are available along with the software or as a separate service. The division has three strategic business units (SBU) which serve different segments of the business market. The division has about 225 clients and about 15 million dollars in sales volume per year. It has been growing at 30 percent per year and has been viewed as highly successful by its parent firm.

SALES PROFILES

The sales data for 223 clients was arrayed in a matrix of clients by eight quarters of sales history (see table 2 for an example) and cluster analyzed with the Howard-Harris Clustering Program. Table 3 shows the percent of the overall variation explained by the cluster solutions (one minus the sum of squares for n clusters divided by the sum of squares for one cluster). The three cluster solution was judged as most appropriate. Although the four and five cluster solution did explain additional variation, the marginal explanation decreased and the sample size became small in some clusters.

TABLE 3: CLUSTER ANALYSIS RESULTS

	Percent of Variation Explained
Two Clusters	31.3%
Three Clusters	43.4%
Four Clusters	52.2%
Five Clusters	56.5%

The average sales profile for each cluster is shown in figure 3. The profiles in each segment are very different. Segment S is characterizd by a small sales volume and sales are declining; segment M sales achieve a medium sales level but show some weakness in late quarters; and segment L has large sales volumes and they are growing.

The fraction of customers in each profile was surprising. 63 percent were small, 32 percent medium, and only 5 percent large. The 30 percent overall sales growth the division experienced was due to a small number of

large clients. Hidden below the positive overall sales trend was the large
proportion of clients that tried the company's service and did not use it at
an increasing rate or whose sales fell to very low levels.

The significance of the differences in the profiles can not be
statistically determined, but table 4 gives the results of an analysis of
variance across the three segments in each period. The F values are large
for all periods after the first and substantiate a judgement that
differences between the profiles exist. The range of plus and minus one
standard deviation of the mean estimate is shown in figure 3 (i.e., the
standard deviation in sales in cluster divided by the square root of the
number of clients in the profile cluster).

Figure 3a: Large Sales Volume Client Profile (L)
[5% of clients]
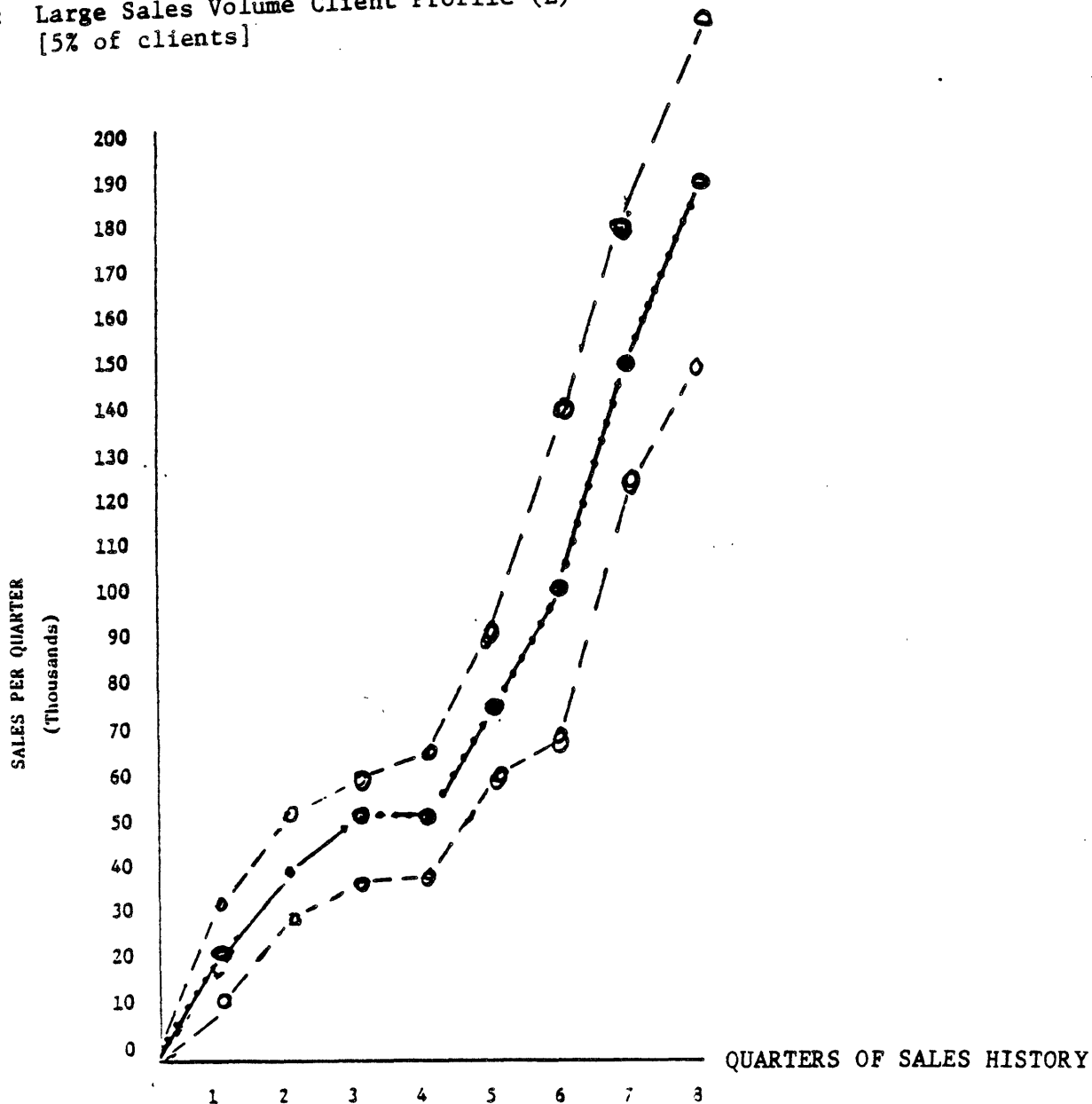
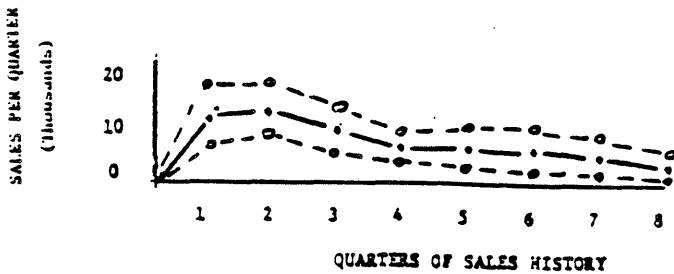Figure 3b: Medium Sales Client Profile (M)
[32% of clients]
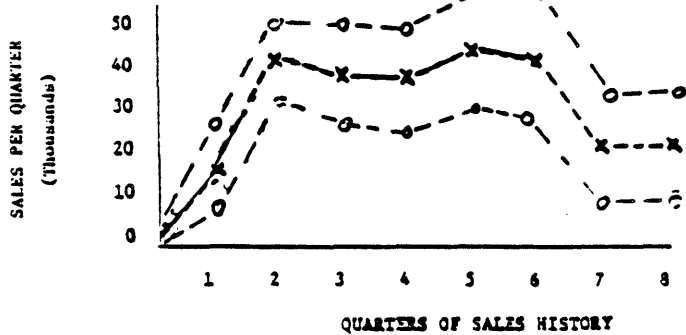
Figure 3c: Small Sales Client Profile (S)
[63% of clients]

Key: _____ = mean $\bar{x}$

-------- = mean $\pm \sigma_{\bar{x}}$

TABLE 4: ANALYSIS OF VARIANCE BETWEEN PROFILES

Sales History Periods	F Value
1	1.1
2	20.5
3	36.1
4	92.4
5	81.0
6	61.1
7	129.5
8	204.5

One obvious hypothesis on why the three segments differ is the size of the client company. A discriminant analysis indicated size of company and other customer demographics were not significant predictors of cluster membership. Rather it appeared that based on discussions with the sales-force and personal interviews with a sample of customers in segment S, customers in the small sales volume segment (S) tried the software and were not satisfied with the benefits they derived relative to the costs within the context of their organizational constraints and culture. Those in segment M were pleased with the system and expanded its use, but unlike segment L the system did not become a corporate standard and thereby enjoy wide spread and growing use.

The stationarity of the profiles was examined by conducting a separate cluster analysis for clients who first purchased from the firm before 1978, in the period 1978 to 1980, and after 1980. In all cases three profiles such as those for segment S, M, and L were obtained. Table 5 shows the F values by cluster type (S,M, or L) for the differences across each group

TABLE 5: F Values for Analysis of Variance Across Cluster Groups with Different Starting Dates within Segment

Sales History Periods	SEGMENT		
	S	M	L
1	2.1	.5	16.7
2	.7	1.8	1.1
3	5.7	.1	12.5
4	7.5	2.4	8.5
5	.4	2.3	5.0
6	.6	6.8	27.6
7	1.4	9.0	113.2
8	1.2	2.6	436.6

TABLE 6: F Values for SBU Level Analysis of Variance (Across Profiles within SBU's)

Sales History Periods	SBU 1	SBU 2	SBU 3
1	2.8	3.4	26.1
2	2.9	24.9	21.5
3	1.0	32.3	11.4
4	34.8	89.1	15.3
5	85.4	33.3	23.0
6	21.4	114.2	26.2
7	28.1	117.3	24.4
8	38.5	125.7	9.0

defined by starting time that resulted from an analysis of variance. The
values of the F are relatively low from segment S and M and do not strongly
contradict the judgment that these sales profiles are stationary. The large
sales segment does have large F values which indicate some dynamics.
Investigations into differences over time indentified a trend away from time
sharing as the basis of computation towards the purchase of the software for
use on in-house computers. This trend was reflected in the model by the
creation of a new profile for those large customers who moved from large
time sharing use to "in-house" implementation. (see eq. 9) On transition to
the new profile, $200,000 in revenue occured and then a quarterly
maintenance fee of $10,000 accrued.

With an overall understanding of the sales profiles and their stability,
the next issue was the level of detail necessary for managerial analysis.
In this firm the SBU's made separate plans, so the clustering was repeated
for each SBU. In two of the SBU's three clusters were selected and in the
remaining SBU, two clusters were judged to be adequate descriptions. Table
six gives the analysis of variance results across the segments within each
SBU. Overall, the values are consistent with a judgment that differences
exist between the sales profiles within each SBU.

MODEL

The model described above (eqs. 1-9) was implemented with three SBU's.
In SBU one and two, the large client segment was modeled as having a
probability of shifting to a "in-house" sales profile. This translation
probability was a function of the number of quarters of sales history in the
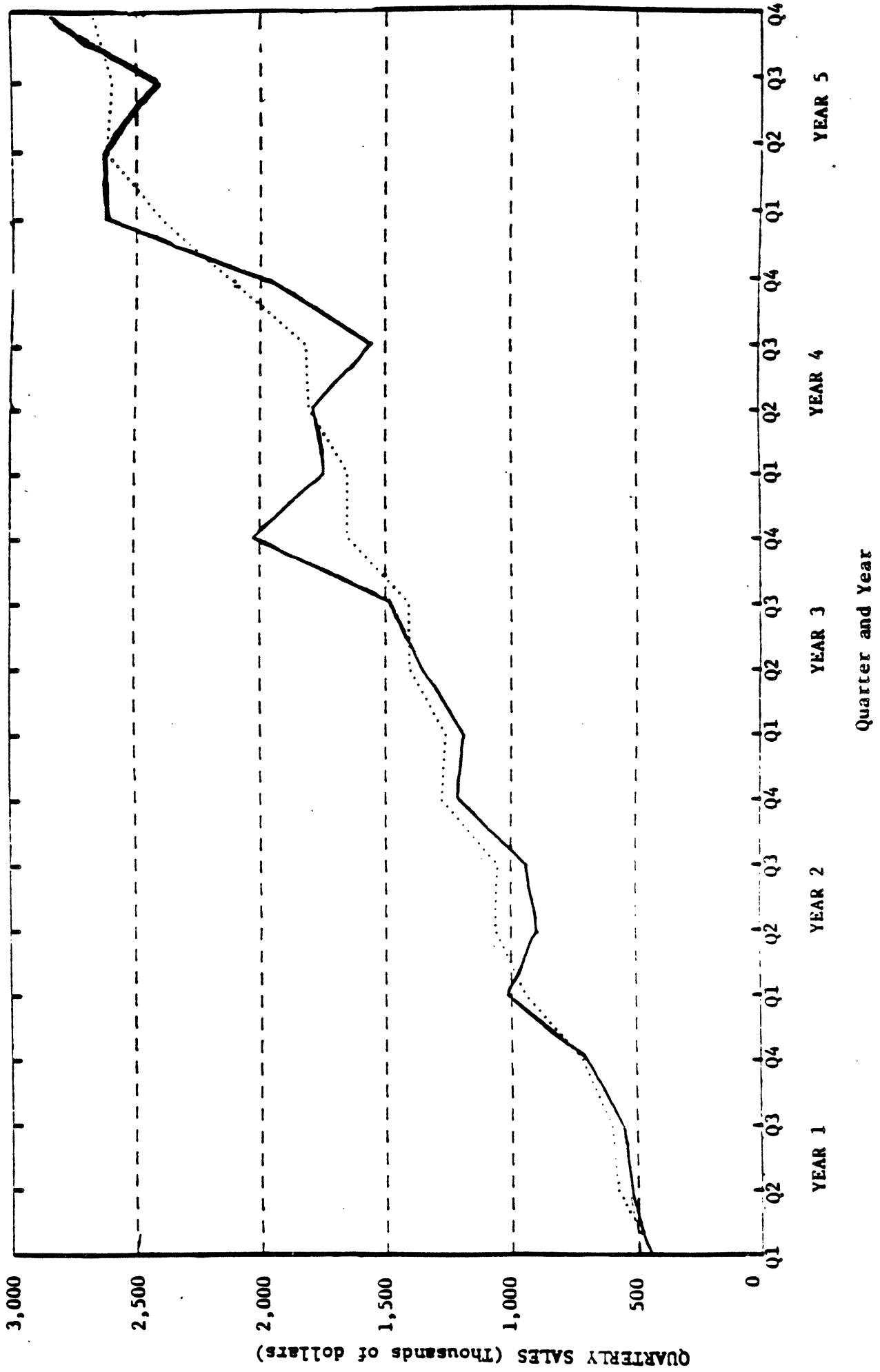large segment. The reasoning was that as the clients' purchases per quarter

rose over time, the purchase of the software for implementation on the

client's computer looked more attractive. Judgmentally determined values

for this function began as .1 for the first quarter and rose to .9 at the

end of the fourth quarter.

In addition to the "in-house" transition another new trend was emerging.

This was for an immediate "in-house" purchase. A profile was created to

represent this process based on first quarter sales of $10,000 to reflect

trial use and then zero for future quarters, but a transition probability of

.3 to the "in-house" conversion profile (see $T_{p,p',h}$ on eq. 9) was defined

to reflect a direct sales. Those who did not make the purchase and its

associated transition were assigned zero future sales because they were

assumed to purchase a competitive product.

The empirical sales profile estimation was done on the basis of eight

quarters, but in the model we needed profiles that extended for longer sales

histories. Empirically, the number of clients with long sales histories was

small, so judgment and these longer sales history sample mean profiles were

used to extend the profiles in each segment to 20 quarters.

Fitting: With the estimated profiles and the observed number of new

clients per period, the total sales were calculated with the model for the

past 20 quarters and compared to the actual results over that period. See

Figure 4. The fits are quite good and suggest the empirical analysis and

judgmental extension of the profiles were reasonable. The standard

deviation is $160,000 or 11 percent of the mean. The corelation of the

actual and fitted values is .99.

FIGURE 4 - FITTING GIVEN NEW CLIENT ARRIVAL


QUARTERLY SALES (Thousands of dollars)

3,000

2,500

2,000

1,500

1,000

500

0

Q1 Q2 Q3 Q4 | Q1 Q2 Q3 Q4 | Q1 Q2 Q3 Q4 | Q1 Q2 Q3 Q4 | Q1 Q2 Q3 Q4
YEAR 1   YEAR 2   YEAR 3   YEAR 4   YEAR 5

Quarter and Year

Fit = ...
Actual = _____

The number of new clients was predicted for years four and five with the new customer generation model described above (eq. 4-6). See Figure 5. Each salesman was considered individually and judgments were modified to fit the observed number of new clients per quarter. The eight man full time sales force was formed in year three. The number of new customers rose in year four as the salesforce gained experience. Late in the fourth year the salesforce size increased from 8 to 14 men. The resulting training burden and the increased demands for servicing of past accounts led to a decrease in the time spent in calling on potential customers and therefore a drop in new client arrivals. This was reflected in the model structure, its predictions, and the actual values. At the end of year four, the new salesforce training was substantially complete so the time available for calling on clients increased $(C_{t,i}$ in eq. 5) and productivity for new salesmen was growing $(A(a_i)$ in eq. 4).

Sales Forecast: With the confidence that the model fit past new customer arrivals and total sales, forecasts were made for years 6, 7 and the first half of year 8. The base forecast was made on the assumption of no new sales hiring and stationarity in the profiles $(S_{p,h}$ eq. 2) and proportions of new clients flowing to each profile $(F_p$, eq. 1). This is shown as the base forecast in Figure 6. Sales grow from about 3 million dollars per quarter to 6.5 million per quarter. The number of new clients averages about 15 per quarter in the period. The forecast is very sensitive to the new customer arrival rate. If the arrival rate drops to 10 per quarter, sales would remain constant at 3 million per quarter. With many of the old clients and 63 percent of the new clients falling in the small (S) segment which is characterized by declining sales (see figure 3c), a substantial inflow of

FIGURE 5 - FITTING NEW CUSTOMER ARRIVAL

New Customers Per Quarter

	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
		YEAR 4					YEAR 5	
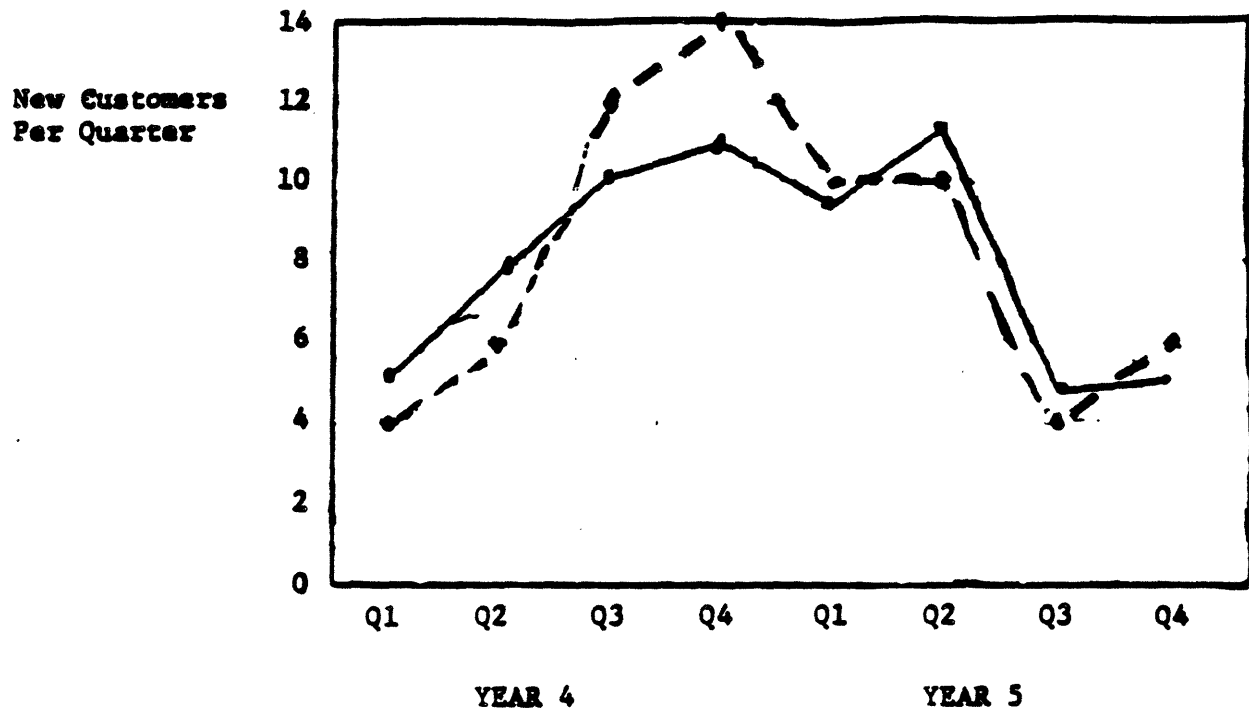
——————— = Predicted

– – – – – – – – = Actual

new customers is necessary just to maintain sales levels. If selling results fall below 10 new customers per quarter, total sales would decline as small (S) and medium (M) segment customers age down their sales profiles and the base revenue erodes.

Simulations that increased the fraction of customers that fall in the medium (M) and large (L) sales volume segments showed substantial improvements in sales. Such changes could result from better pre-screening and targeting of prospects or improvements in the price value relationship of the software product. Although the discriminant analysis was not successful in describing a demographic set of classification variables, the sales manager of the firm felt that a better assessment of the benefits of the software for a client and selection of a meaningful first application in each company would increase the likelihood of a firm following the medium or large sales profiles.

Salesforce Size: The sensitivity of the forecast to the number of new customers suggests a possible gain to increasing the salesforce. A scenario to reflect more selling effort was formulated based on hiring two new salesmen in the first quarter and three in the second quarter of year 6. The resulting forecast showed little change in sales in year 6 and an increase in years 7 and 8 (see figure 6 for a larger sales force). This occured because the new salesmen had to gain experience and their training took some time from existing salesmen and sales managers. The lag of about one year between the hiring and incremental sales for the company resulted in profit reductions in year 6 (see figure 7). The incremental profit impact per salesman per quarter is shown in figure seven. The $140,000
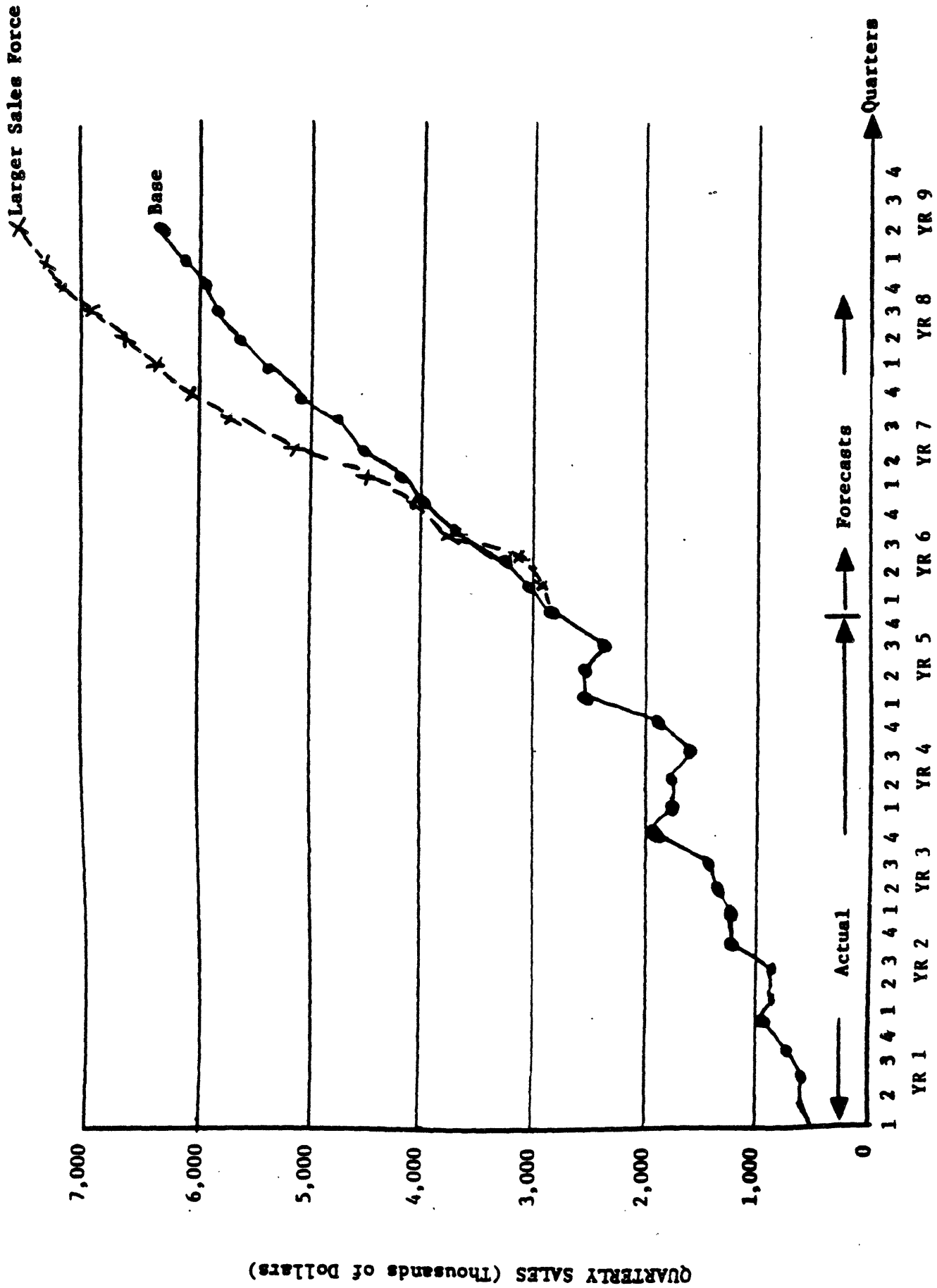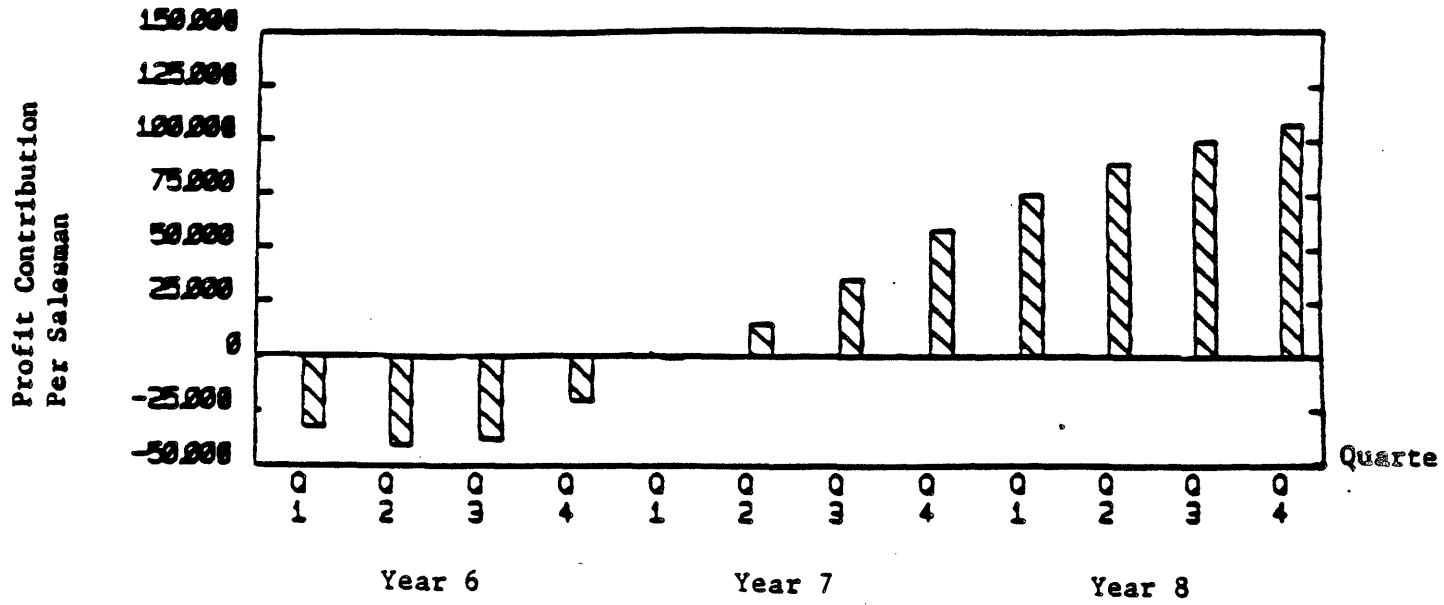
FIGURE 6 – SALES FORECASTS

QUARTERLY SALES (Thousands of Dollars)

- 31 -

FIGURE 7: IMPACT OF ONE MORE SALESMAN

profit loss results in the first year for each salesman reflects the incremental investment necessary to increase the sales capacity. The second and third years generate profit and the investment is paid back by the end of the second year. Each salesman generates $350,000 of incremental profit over the three years. These simulations indicate the need for more than an annual planning horizon for decisions on salesforce size. In this case, short run (12 months) salesforce hiring would not increase sales and would decrease profit. In fact, simulations showed short run profits could be increased at the expense of future growth by firing the division's least experienced salesmen. The division adopted a longer run perspective and enlarged the salesforce realizing this would do little to improve the next year's results.

Pricing: The large number of customers in the small (S) sales segment and the sensitivity to increasing the fraction of new customers in medium (M) and large (L) sales segments suggested the possibility of lowering prices to improve the price-value perception of the product. When price is changed a number of effects result. New client inflow and sales profiles work to increase revenue, but the dollar sales to existing clients fall unless their price elasticity is high enough to more than compensate by increased unit sales for the price decline. In this application another complication was present because prices for time shared computing and "in-house" purchase prices could be set separately. If the purchase price was reduced, the transition to the "in-house" profile (eq. 9) would increase. Although this results in a large single payment, its present value is less than a client remaining as a time share user.

Simulations were run for changes in time sharing prices alone, "in-house" prices alone, and for combinations of both of these. The elasticities for existing sales, new client arrival rates, and transitions were judgmentally set after Delphi procedures were used with salesmen, small client segment rejectors were interviewed, and industry data was studied.

The results indicated that time sharing price reductions over a wide range of values were unprofitable and lowering "in-house" conversion prices by fifty percent would have little effect on total sales revenue and profit. In the case of lowered "in-house" prices, the loss of profit from existing customers was balanced by a gain of profits from an increased input of new clients. The sales revenue in dollars was constant, but the number of "in-house" installations doubled and the firm's market share increased.

The division decided it would cut price to build market share. Simulation of new competitive entry also suggested it would be advisable to cut price in an attempt to preempt competition.

MANAGERIAL RESULTS

The model and sales profiles were useful to the division in understanding its sources of sales. New attention was directed at nurturing and growing small clients (e.g. changing their sales profile) and qualifying leads so as to increase the fraction of medium and large segment clients. The sales profiles and mix of segments led to recognition of the critical nature of a continuing inflow of new customers. The sales force size increase and fifty percent price reduction indicated by the simulation runs

were implemented. Systems software developments were undertaken to make the software more productive for users and improve its performance price ratio.

Subsequent to the model forecasting runs (figure 6), actual year six sales results became available. The predicted sales results in year six were 14.4 million versus the prediction of 13.9 million. This reflects reasonable accuracy. Future forecasts were made to reflect the past sales force hiring and price reductions.

Several new phenomena have been added to the model and its inputs to reflect changes in the firm's business. First, a micro computer version of the software has been introduced. It is envisioned as an additional product to add to the sales profile for clients. Prices of the mainframe product have continued to suffer further erosion due to technological change and new competition. These changes have been reflected in new total revenue profiles. Sales requirements measured by the number of calls necessary to sell a new account ($A(a_i)$) have increased because more information system managers have been added to end users in the decision making unit. Finally, time sharing has been disappearing at an ever more rapid rate so the transition to "in-house" and the fraction of customers going directly "in-house" has been increased. With these new inputs, simulation runs produced three year forecasts that showed little growth in year seven but a return to 30 percent growth in years eight and nine.

CONCLUSION

In this paper we have proposed the concept of sales profiles as an aggregate description of the relationship of the industrial firm and its customers. A straight forward application of cluster analysis of readily available sales data provided estimates of these profiles. A macro flow simulation model based on sales profiles for segments of customers was useful in one application in understanding the underlying source of sales, forecasting, and the analysis of marketing decisions.

Future research will be focused on additional testing and application of the model, sales profile estimation methods, and evolutionary developments of the flow structure. New clustering procedures aimed at the splitting of a segment over time into different profiles will be explored (Wong 1983). The model will be extended to include sales profiles of revenue gained through a distribution channel as well as directly from users. The proposed model only considers the sales a firm receives from a client. It would be desirable to extend the analysis to include competitive effects. Perhaps the share of business a client gives a manufacturer could be the basis of the customer profiles. The propensity of a client to switch from one supplier to another (Gensch, 1984) could be considered for inclusion in a more comprehensive representation of the customer-firm relationship. We envision an evolutionary modeling system to produce a better model for managerial decision support (Urban and Karash 1971).

One of the important lessons learned from the application presented in this paper is the need to look at a multi-year horizon in sales force

staffing. Recall only considering an annual planning period would have not led to investment in the salesforce because profit declined in the first year after hiring. Perhaps this model with its longer horizon can be a link between annual planning and five year strategic plans. The five year plan often suggests investment in some SBU to build share, but does provide an operational plan to achieve the sales increase. The annual plan in many cases reflects relatively small changes from the past and is rather myopic. The model proposed in this paper may be able to translate strategic resource commitments into detailed oprational variables and make the annual plan consistent with the five year plan. In this context we could call the model a "strategic operations model" because it details the specific resource commitments over a multiyear period and formulates annual plans that follow the strategic time path. Future research will assess the use of sales profiles and the flow model in the operationalization of marketing strategies.

ACKNOWLEDGEMENTS

BIBLIOGRAPHY

Everitt, B., _Cluster Analysis_ (London, England, Heinemann Educational Books Ltd., 1974).

Gensch, D.H., (1984), "Targeting the Switchable Industrial Customer," _Marketing Science_, Vol. 3, No. 1 , pp. 41-54.

Green, P.E. and V.R. Rao, _Applied Multidimensional Scaling_, (New York, Holt, Rinehart, and Winston Inc., 1972).

Hartigan, J.A. (1978), "Asymptotic Districtions for Clustering Criteria," _Annals of Statistics_, 6, pp. 117-131.

Milligan, G.W., and M.C. Cooper (1983), "An Examination of Procedures for Determining the Number of Clusters in a Data Set," College of Administrative Science, WP #83-51, The Ohio State University.

Urban, G.L. and J.R. Hauser, _Design and Marketing New Products_, (Englewood Cliffs, N.J., Prentice Hall Inc. 1980), pp. 433-60.

Urban, G.L. and R. Karash, "Evolutionary Model Building," _Journal of Marketing Research_, Vol. VIII (February 1971), pp. 62-66.

Wong, M. Anthony, "A Bootstrap Testing Procedure for Investigating the Number of Subpopulations," _Sloan School of Management_ (Cambridge, MA, MIT, 1984).