IIE Transactions (2004) **36**, 941–952

ISSN: 0740-817X print / 1545-8830 online
DOI: 10.1080/07408170490487777

Salesforce design with experience-based learning

SANJOG MISRA, EDIEAL J. PINKER* and ROBERT A. SHUMSKY

William E. Simon School of Business Administration, University of Rochester, Rochester, NY 14627, USA
E-mail: misra or pinker or shumsky@simon.rochester.edu

Received September 2002 and accepted January 2004

This paper proposes and analyzes an integrated model of salesforce learning, product portfolio pricing and salesforce design. We consider a firm selling two products, with a pool of sales representatives that is split into separate salesforces, one for each product. The salesforce assigned to each product is faced with an independent stream of sales leads. The salesforce may also handle leads that overflow from other product salesforces. In addition, salespeople "learn by doing" over their tenure on the job. In particular, the more time they spend selling a particular product, the more productive the sales effort. The objective of the firm is to maximize profits by optimizing the size of all salesforces as well as the prices of all products. Using data collected from the salesforce of a large manufacturer, we provide evidence for the link between experience and sales, and we demonstrate how parameters of the model may be estimated from real data. Numerical experiments using parameters derived from the data analysis indicate that the optimal salesforce size increases with both sales productivity and the learning rate, and decreases with salesforce costs (e.g., wage per representative), product production costs and consumer price sensitivity. We also find that worker learning can significantly dampen the effect of rising costs (or decreasing margins) on staffing levels. Finally, we examine the impact of learning on both the optimal salesforce structure (specialists versus generalists) as well as the optimal routing of sales leads to sales representatives.

1. Introduction

By broadening the range of tasks assigned to individual workers, firms hope to create workforces that are responsive to variability in workloads. In the case of a salesforce, such job flexibility is often implemented by pooling salespeople across product lines. In a sales organization with pooling, each salesperson may have a primary responsibility for some product line but is also able to sell some subset of the other product lines. Two important factors that influence the performance of such a salesforce are pricing and experience. Product price will influence the demand for a product, while the salesperson's experience influences the likelihood of making a sale. Therefore, when sizing and structuring a salesforce, a firm must consider pricing and the impact of staffing decisions on the experience of the salesforce. Understaffing the sales team for a product may lead to lost sales and possibly to lost future market share. Overstaffing the sales team can be expensive since good sales people are typically well compensated. Assigning staff to product lines must also be done carefully. Some products are complex and sales success depends upon experience, while simpler products require little experience on the part of the salesforce.

In this paper, we examine the interactions among staffing, learning, and pricing in the management of a salesforce. We develop a model of a salesforce that receives sales leads with the sales volume generated from each lead depending upon the experience of the salesperson and the price of the product. The firm's goal is to maximize profit by adjusting product prices as well as how many salespeople are allocated to each product. Our model applies to firms with the following three attributes: (i) sufficient market power to have control over pricing; (ii) complex products that require experience to sell effectively; and (iii) large marketing efforts, such as advertising in the media and at tradeshows, so that most of the leads are generated by activities outside of the salesforce. Specifically, a firm selling high-technology industrial products in a mature market would satisfy all of these criteria (such a firm satisfies attribute (iii), for in a mature market there are few untapped sales leads, and its salesforce focuses its energy on following up requests by existing customers rather than unearthing new customers). After developing the model, we estimate its parameters using data collected from the salesforce of one such firm.

This paper makes use of the framework developed in Pinker and Shumsky (2000) for modeling learning in service systems. We combine a model of job tenure with a model of experience-based learning and apply it to the problem of salesforce design. The current paper, however, has four significant differences and extensions that lead to interesting results not previously seen in the literature. Firstly, the service level, defined as the throughput of sales leads, is determined endogenously while it is exogenously determined

*Corresponding author

in Pinker and Shumsky (2000). Therefore, we are able to see how optimal staffing responds to changes to cost parameters. Given this additional degree of freedom, we also find that worker learning can significantly dampen the effect of rising costs (or decreasing margins) on staffing levels. For example, if learning is a significant factor, an increase or decrease in the cost of salesforce compensation has a relatively small impact on the optimal salesforce size. Secondly, we model a more complex routing of customers (sales leads) to workers (salespeople) that more accurately reflects a common practice in sales organizations. Given this routing, we arrive at the surprising result that, with learning, pooling of workers may lead to optimal staffing levels that are higher than when workers specialize. This contradicts the conventional wisdom that the economy of scale provided by pooling reduces staffing requirements. Thirdly, the model incorporates pricing decisions and their effect on demand. As a result we are able to study the relationship among staffing levels, job flexibility, and price levels. In particular, we find that the optimal salesforce size declines as price sensitivity increases. We also find that when learning is a significant factor in determining sales volume, a specialized (or "exclusive") salesforce leads to higher optimal prices than the optimal prices for a pooled salesforce. Fourthly, motivated by data collected from the salesforce of a large manufacturer, we propose a learning-curve model different from the model in Pinker and Shumsky (2000), and we demonstrate how parameters of the model may be estimated from the data.

In the next section, we present an overview of the relevant literature. In Section 3, we formulate our model by integrating a service process model with an employee tenure model, a model of experience-based learning, and a model of consumer demand. Section 4 contains analytical results that describe the impact of various parameters on the optimal price. Section 5 describes the analysis of industry sales data that provide baseline parameters for the numerical experiments of Section 6. These numerical experiments provide insights into how learning effects staffing and pricing. Section 7 summarizes our results and discusses possible extensions of the model.

2. Literature review

As noted above, this paper is most closely related to Pinker and Shumsky (2000). Other researchers have also considered parts of the problem addressed in this paper but we believe that ours is the first to integrate all of them into one model. Some researchers have studied the control problem of how to hire, fire and promote workers to maintain appropriate staff levels when career paths are stochastic, and these are listed in Pinker and Shumsky (2000). None of these studies consider the effect of pricing on staffing and therefore do not connect staffing and learning to sales, limiting their applicability to salesforce design.

The literature on salesforce management has focused on either salesforce incentives, see for example Basu *et al.* (1985), Lal and Srinivasan (1993), Joseph and Thevaranjan (1998) and Bhardwaj (2001), or salesforce sizing and mix issues. In this paper, we do not focus our attention on compensation issues. In particular, we assume that effort is perfectly observable and hence a fixed wage, forcing contract is optimal.

Montgomery and Urban (1969) and Lucas *et al.* (1975) use profit maximization models to solve for the optimal salesforce size. A limitation of their approach is that they ignore the presence of multiple products and/or territories. Lodish *et al.* (1988) use a more sophisticated approach to modeling the issue in the case of one particular firm. They showed, for this one firm, that adding salespeople and redeploying them would result in increased profits. Zoltners (1976), Lodish (1976, 1980), Rangaswamy *et al.* (1990), and Mantrala *et al.* (1992) consider the problem of finding the optimal allocation of salespeople to territories, products or customers. These studies use static frameworks that do not incorporate learning effects within the salesforce. Another issue that has not been addressed adequately in the literature is specialization and the effect it has on structuring the salesforce. Given that salespeople often specialize in particular products and that such specialists are scarce, there are instances when a non-specialist serves a customer, which may have an impact on sales.

Dewan and Mendelson (1990), Stidham (1992) and So and Song (1998) are examples of works in which both capacity and pricing are endogenous to the firm's decision problem. In all of these, the firm is modeled as a single-server queue and capacity is determined by the service rate. In this paper, we are explicitly modeling capacity as the staffing level in a multi-server queue. Furthermore, we consider the interaction of parallel queues serving different customer types. Finally, in our model price determines the sales quantity rather than the customer arrival process. As we mentioned in the Introduction, our formulation is appropriate for environments in which sales leads are "handed off" to the salesforce.

To summarize, we are studying a set of problems that have been looked at in isolation from various perspectives in the marketing and operations literature. We aim to integrate these diverse perspectives into a unified model that will help us to understand the dynamics of learning and its impact on pricing, salesforce size and salesforce design.

3. Model formulation

Consider a firm that sells two products, A and B, and has two types of salespeople, A and B. We assume that sales leads representing customers interested in each of these products arrive according to a Poisson process with arrival rates of λ_A and λ_B respectively. In the following, we refer to customers and sales leads interchangeably. We can state the profit function of the firm in very general terms as a function of the staffing $\mathbf{S} = (S_A, S_B)$ and the price of each product $\mathbf{p} = (\mathbf{p}_A, \mathbf{p}_B)$ as follows:

$$\Pi(\mathbf{S}, \mathbf{p}) = \sum_{i=A,B} \sum_{j=A,B} r_{ij} q_{ij} (\mathbf{p}_i - c_i) - d_i S_i, \qquad (1)$$

where r_{ij} is the throughput of type-i leads through type-j salespeople, q_{ij} is the expected quantity of a product i sold by a type j salesperson pursuing a type-i lead, c_i is the production cost of each unit of product i and d_i is the cost per unit time of each salesperson. Both c_i and d_i are exogenous parameters. The following sections describe how we find the throughput of each customer type (r_{ij}) and the quantity sold (q_{ij}).

3.1. *Throughput statistics*

In practice, sales leads must be allocated to individual salespeople. Each salesperson may work on many leads simultaneously with a particular lead being active for days, weeks or months depending upon the nature and characteristics of the product class. As we discussed in the Introduction, it is common for salespeople to be assigned primary responsibility for one set of products and secondary responsibility for others. For example, the firm would prefer that type-i salespeople sell type-i products, but if all i salespeople are occupied the firm may want a type-j salesperson to follow-up on the lead rather than lose the sales opportunity altogether. In practice, salesforce compensation is often designed as a matrix that assigns a commission to salesperson-type and product-type pairs. The purpose of such a matrix is to encourage salespeople to focus on their primary product lines while keeping the option open for cross-selling. To simplify our analysis, we assume that the assignment of customers to salespeople occurs as follows. When a type-i sales lead arrives it is directed to a type-i salesperson. However, if all type-i salespeople are busy pursuing other leads the lead is routed to a type-j salesperson. If all salespeople (i.e., of both types) are busy, the lead/customer is lost. Within each group of salespeople, arrivals are routed so that work is shared equitably among the salespeople. Figure 1 shows the routing of leads through the salesforce in our model. We refer to this as a *hierarchical* salesforce design.

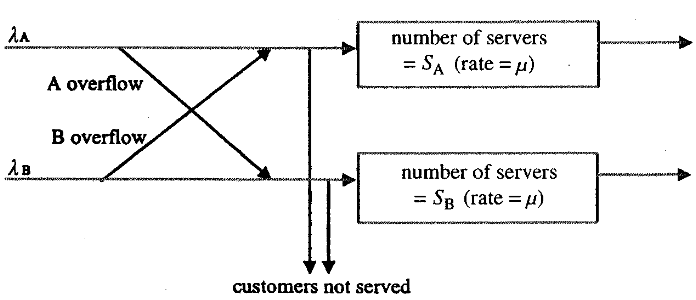

Fig. 1. Routing of leads through a hierarchical salesforce.

We are modeling the salesforce as a pair of multi-server service systems with exponential service times that operate in parallel and receive their own independent Poisson arrival streams with rates λ_A and λ_B but also allow leads to overflow into each other. Later we will also examine systems in which a group of salespeople are dedicated to a single product. Because the "single-product" model is relatively simple (e.g., the sales leads flow through an $M/M/S/S$ queueing system), we will focus on the more general two-product model here. We will assume that the service rate of each salesperson, μ, is the same (although it is not difficult to relax this assumption). Note that we are modeling each lead as being processed sequentially by a salesperson while in practice a salesperson would be pursuing multiple leads simultaneously. We make this abstraction to simplify the calculation of queueing statistics, and we believe that explicitly modeling the simultaneous processing of leads would be an interesting topic for further research. In addition, we assume that sales quantity does not affect the time it takes to pursue a sales lead.

The sales quantity model described in the next section uses two sets of statistics from this model: throughput and utilization. Let r_{AB} represent the throughput of type A leads through type B salespeople; r_{BA}, r_{AA}, and r_{BB} have similar interpretations. Note that r_{ij} is a function of the staffing vector $\mathbf{S} = (S_A, S_B)$. Utilization is represented as ρ_{AB}, ρ_{BB}, ρ_{AA}, and ρ_{BA}, and each of these is calculated easily from the appropriate values of r_{ij}.

Calculating the throughput statistics of the salespeople is more difficult than finding the throughput of a standard loss system because the arrival process to each group of salespeople is not purely Poisson but is instead a combination of a Poisson process and bursts of arrivals that are sent when the other sales group is fully occupied. For this system we calculate throughput statistics numerically. Specifically, we define a two-dimensional state space (N_A, N_B) where N_i represents the number of busy salespeople of type i. The balance equations for this state space are relatively simple to enumerate, and we use these equations to solve iteratively for the steady-state probabilities of (N_A, N_B) (Gross and Harris, 1985, p. 437). Given the steady-state probabilities, we calculate the expected throughput and utilization.

3.2. *Quantity sold*

The quantity of the product sold as a result of pursuing a lead depends upon the price of the product and the experience of the salesperson with that product. Here we develop a model of demand that is based upon customer sensitivity to both price and the experience of the salesforce as well as a model of the career path and experience accrual of an individual salesperson.

Given that a type-j salesperson pursues a lead for product i, we assume that the sales generated by that lead is a random variable $D_{ij} = \alpha_{ij} - \beta_i p_i$, where α_{ij} is a random variable that depends upon the (random) experience level of the salesperson encountered by a customer and β_i is the sensitivity of demand to price for product i. Therefore, in the profit function of Equation (1):

$$q_{ij} = E[D_{ij}].$$

The important difference between our specification and the traditional, linear demand model is that we allow the intercept to vary across salespeople. In the marketing literature sales volume is typically described as a function of the skill of a salesperson. Rao (1990) for example, proposes the functional form:

$$s = s_0(1 - e^{-nb}), \tag{2}$$

where s is the sales, s_0 is the maximum achievable sales level, b is the skill of the salesperson and n is a parameter determining the rate at which s_0 is approached. From the learning-curve literature (Yelle, 1979; Badiru, 1992), we can see that often heterogeneity in skill is a result of heterogeneity in experience levels. In this paper we combine these two perspectives by expressing sales volume as a function of experience. We use the same functional form as in Equation (2), in particular we assume that $\alpha_{ij} = K_{ij}(1 - \exp^{(-nw_{ij})})$ where w_{ij} is the accrued experience of a type-j salesperson selling product i (measured in units of time), K_{ij} is a constant representing the upper limit of sales ability, i.e., the sales volume of a salesperson with infinite experience, and n is a learning parameter. Much of the traditional literature on learning curves uses units of work, for example widgets built, as a measure of experience. In manufacturing settings where unit labor costs decrease with learning because workers become faster, and it is easy to measure costs, this approach is appropriate. In our application the work unit is a sales lead and it is difficult to obtain data on the number of leads handled. Furthermore, it is not clear that the time spent per lead will decrease with experience. Rather, as we model it, the likelihood of a sale will increase with experience. Therefore, additional experience selling a product has the effect of increasing the intercept of the demand for that product. This functional form is appealing for a number of reasons. First it is consistent with the marketing literature, second, it is consistent with the learning-curve literature in which most learning curves are asymptotic, and finally we have found that it provides a reasonable fit to actual sales-force performance data. The learning-curve literature typically models production costs as a decreasing function of experience that asymptotically approaches zero. In our case we are modeling the effect of experience on sales (or revenue generation) and therefore use an increasing function of experience that asymptotically approaches some upper limit. Note that the proposed function differs from the unbounded "power function" used in Pinker and Shumsky (2000) to describe the relationship between experience and service quality. We found that the bounded function proposed here provided a better fit to the sales data that will be described in Section 5.

Given our discussion above we can now define quantity sold as:

$$D_{ij} = K_{ij}(1 - e^{-nw_{ij}}) - \beta_i p_i. \tag{3}$$

Let y be the tenure of a salesperson. Then the expected sales quantity is:

$$q_{ij} = E[D_{ij}] = K_{ij} - K_{ij}E_y[E[e^{-nw_{ij}} \mid y]] - \beta_i p_i. \tag{4}$$

The outer expected value is over the tenure of the type-j salesperson serving a given customer and the inner expected value is over the experience of the salesperson with the type-i product, given tenure y.

It can be shown that when the average time spent on a sales lead is small relative to the tenure of a salesperson then $E[\exp^{(-nw_{ij})} \mid y]$ can be closely approximated by $\exp^{(-n\rho_{ij}y)}$. This approximation is similar to one that appears in Pinker and Shumsky (2000) and its accuracy here has been verified using simulation. Therefore:

$$q_{ij} \approx \int_0^{\infty} K_{ij}(1 - e^{-n\rho_{ij}t})g_y(t)\mathrm{d}t - \beta_i p_i. \tag{5}$$

The probability density function for y, $g_y(t)$, is derived from a model of a salesperson's tenure process in which a career is divided into stages, so that the stages of the career can be modeled as states of a continuous-time Markov chain. The tendency to end employment (by being fired or quitting) varies from stage to stage, and the time a salesperson stays in a stage before leaving is exponentially distributed. The parameter λ_1 is the rate at which salespeople move from the first stage to the second stage, λ_2 is the rate at which they end their employment in the first stage, λ_3 is the rate at which workers in the second stage end their employment, and $\lambda_2 > \lambda_3$.

Using this model of the tenure process it can be shown that:

$$q_{ij} \approx K_{ij}\frac{n\rho_{ij}}{\lambda_1 + \lambda_2 + n\rho_{ij}}\left(1 + \frac{\lambda_1^2 + \lambda_1\lambda_2}{(\lambda_1 + \lambda_3)(\lambda_3 + n\rho_{ij})}\right) - \beta_i \mathbf{p}_i. \tag{6}$$

Equation (6) accounts for price, learning and the tenure process to determine the quantity sold. Since ρ_{ij} is a by-product of the staffing levels, $\mathbf{S} = (S_A, S_B)$, Equation (6) also links staffing to sales.

3.3. *The complete objective function*

We can now restate the optimization problem faced by the firm as:

$$\operatorname*{Max}_{\mathbf{S},\mathbf{p}} \Pi(\mathbf{S}, \mathbf{p})$$

where

$$\Pi(\mathbf{S}, \mathbf{p}) = \sum_{i=A,B}\sum_{j=A,B}\left\{ r_{ij}(\mathbf{S})\left[K_{ij}\left(\frac{r_{ij}(\mathbf{S})}{\mu S_j}\right) \times \left(\frac{n}{\lambda_1 + \lambda_2 + n(r_{ij}(\mathbf{S})/\mu S_j)}\right)\right.\right.$$

$$\times\left(1+\frac{\lambda_1^2+\lambda_1\lambda_2}{(\lambda_1+\lambda_3)(\lambda_3+n(r_{ij}(\mathbf{S})|\mu S_j))}\right)$$
$$-\beta_i\mathbf{p}_i\Bigg](\mathbf{p}_i-c_i)-d_iS_i\Bigg\}. \qquad (7)$$

There are a number of trade offs explicitly represented in this objective function. First we know that throughput (r_{ij}) is increasing in staffing and therefore there is a trade-off between the additional revenue brought by increased staffing and the marginal cost of an additional salesperson d_i. However, while increasing staffing increases the number of sales leads that can be pursued, increasing staffing also reduces the number of units sold per lead because it reduces the utilization and therefore the experience of the salesforce. This complex effect of utilization on experience and profits can be seen by the appearance of the variable S in the denominator of some of the terms in Equation (7). This trade-off is clear when the salesforce sells a single product. In Section 6 we will see that this *utilization effect* also has a significant influence on salesforce design decisions when there are multiple products, e.g., whether to deploy a specialized or pooled salesforce.

4. Optimizing prices, given salesforce size

The complex interactions among staffing, sales and experience make it difficult to derive an analytical characterization of the optimal decision. However, given the expected profit function, Equations (1) and (7), it is relatively straightforward to solve for optimal prices given a specific staffing of each type of salesperson.

To obtain the optimal prices given staffing we differentiate expected profit, Equation (1), with respect to price and obtain:

$$\sum_{j=\mathrm{A,B}}\left[r_{ij}\frac{\partial q_{ij}}{\partial p_i}(p_i-c_i)+r_{ij}q_{ij}\right]=0 \quad \text{for } i=\mathrm{A,B}. \qquad (8)$$

Solving for price and using the notation of Equation (7) we find:

$$p_i=\frac{c}{2}+\sum_j r_{il}K_{ij}\left(\frac{r_{ij}(\mathbf{S})}{\mu S_j}\right)\left(\frac{n}{\lambda_1+\lambda_2+n(r_{ij}(\mathbf{S})/\mu S_j)}\right)$$
$$\times\left(1+\frac{\lambda_1^2+\lambda_1\lambda_2}{(\lambda_1+\lambda_3)(\lambda_3+n(r_{ij}(\mathbf{S})/\mu S_j))}\right)\Bigg/2\beta_i\sum_j r_{ij},$$
$$\text{for } i=\mathrm{A,B}. \qquad (9)$$

This expression is similar to the standard monopolistic price for a linear demand curve, except that the intercept is a weighted average of the intercepts of the two sources of demand. If there were only one product, no learning effects (so that the demand intercept is a constant, α), and throughput were equal to one, then the price equation is:

$$p=\frac{c}{2}+\frac{\alpha}{2\beta},$$

which is the standard monopoly price.

We now describe a few properties that follow directly from the price equation. The price of product i,

1. increases with the cost, c_i, of product i;
2. increases with the maximum productivity, K_{ij}, of a type-j salesperson with product i;
3. decreases with the price sensitivity β_i, of product i;
4. increases with the learning rate, n (while this is not obvious from Equation (9), it can be shown that $\partial p_i/\partial n>0$).

One property we do not specify here is the relationship between staffing and pricing. While it might seem appropriate to conjecture that prices and staffing (for a given product) move in the same direction, this is not clear from our model. While throughput is increasing in staffing, utilization is not, and this may create a non-monotone relationship between the two. We revisit this issue in the numerical experiments of Section 6.

5. Industry data analysis

The model described in Section 3 assumes that sales productivity grows with the experience of a particular salesperson. While the impact of experience on manufacturing productivity has been well documented by empirical research (see the summary by Yelle (1979)), to our knowledge there have been no published studies linking sales and experience in a salesforce. The data analysis in this section helps us to identify reasonable learning-curve parameters that will be used in the numerical experiments of the next section.

For our analysis we have obtained sales data from one particular company, "Firm A," a market leader in office products with an annual sales revenue of over \$10 billion and over 40 000 employees. Although the firm operates in various product and service markets we restrict our focus to the division that is the flagship of the company and accounts for a substantial proportion of its revenues. The business environment of this division conforms to the assumptions of our model: the firm is a market leader and has some pricing power, the product is complex, and the market is mature so that salespeople primarily respond to requests from existing customers, rather than finding new leads. The division has two primary salesforces, "Representatives" (or "Reps") and "Specialists". Specialists sell technologically advanced, high-priced equipment to large corporations while Reps focus on less-complex and less expensive products for small and medium-sized firms. Our data set is cross-sectional: it records the number of years a salesperson has been with the firm (tenure) and the most recent annual sales figure for each employee. Table 1 contains a summary of the data. Figures 2 and 3 display the relationship between tenure and sales in each sale force. Each '•' in Fig. 2 represents the average sales of 50 salespeople, while each data point in Fig. 3 represents a group of 40 salespeople. For example, the first point on the lower left of Fig. 2 shows the average sales of the 50 most inexperienced Reps: their average tenure was 5 months, and

Table 1. Summary of salesforce data

		Employee tenure (years)				*Annual sales per employee ($, ×$10^6$)*			
Salesforce	*Number of employees*	*Mean*	*Std. dev.*	*Min.*	*Max.*	*Mean*	*Std. dev.*	*Min.*	*Max.*
Reps	1239	9.1	9.3	0.1	35.6	1.3	1.1	0.005	9.5
Specialists	409	10.0	7.8	0.2	34.8	3.1	2.2	0.005	14.9

the average sales in that group was $330 000/year. In the figures we see a relationship between tenure and sales that could be attributed to learning, and that in each salesforce there is a large number of relatively inexperienced salespeople on the "steep" part of the learning curve. For example, over half of the Reps have less than 4 years of experience.

To quantify this relationship between tenure and sales we estimated the parameters of the following model for each salesforce:

$$sales_i = H(1 - e^{-NT_i}) + \varepsilon_i, \tag{10}$$

where $sales_i$ represents the dollar value of sales made by a salesperson i, T_i is the length of tenure, and ε_i is a stochastic error term assumed to be distributed identically and independently normal with mean zero. We used the maximum likelihood method to estimate H and N from each data set. These estimates are presented in Table 2 and the associated functions are plotted as dotted lines in Figs. 2 and 3. As one might expect, the learning curve is more gradual and the asymptote H is higher for the specialists, who handle more complex and expensive products.

The proposed model seems to provide a good fit with the data, although there are clearly other factors besides experience that influence sales ($R^2 = 0.10$ and 0.11). There are also some limitations to this data set that restrict our ability to precisely estimate the learning-curve parameters n and K_{ij} (or, K when there is just one product). Because of the aggregate nature of the data, the statistical model presented in this section cannot control for a variety of complicating factors, including selection bias, variations in utilization among salespeople, and differences in the quality of leads assigned to each worker. Therefore, the parameter values should be interpreted as approximations, and we will only use them to guide our choice of parameter ranges in the following numerical experiments.

In particular, we believe that the two estimates of N (0.046 and 0.066) represent relatively low values of n: both the Specialists and Reps deal with complex markets and products, and these values of n would indicate a slow rate of growth in sales as experience increases (e.g., with $n = 0.046$, a salesperson with 1 year of experience expects to achieve 42% of maximum sales). As a lower bound for this learning-curve parameter, we will use $n = 0.02$ (in this case, a salesperson reaches just 20% of the maximum after 1 year and requires almost 17 years to reach 98%). On the other hand, some products and markets are relatively simple, so that salespeople have a rapid ascent up the learning curve, relative to their tenure. After this rapid climb, sales do not increase significantly with experience. To represent such environments we use upper bound of $n = 4$ (a salesperson reaches 98% of the maximum within 1 month).

While the estimates of N derived from the industry data lead directly to our estimates of n in the general model, the connection between H and the parameter K is more complex. There are two complications when trying to derive K from H: (i) K represents a *quantity* of product sold per sales lead, while H is an upper bound on the annual sales per salesperson; and (ii) K is the number of items sold per lead, given infinite sales experience and a price of zero (see

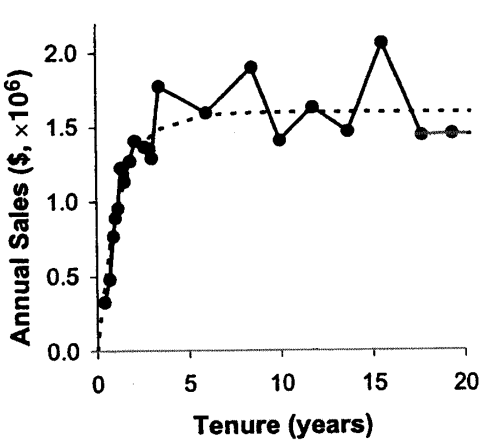

Fig. 2. Reps data and model.

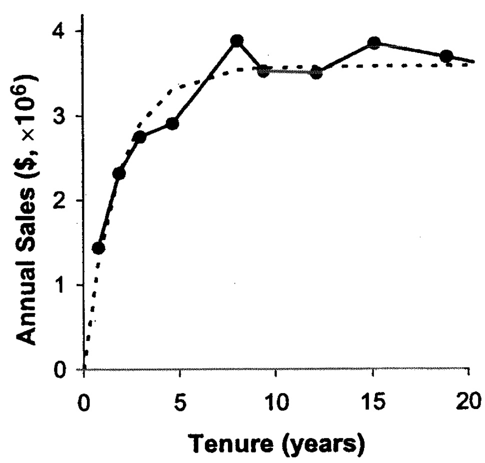

Fig. 3. Specialists data and model.

Table 2. Maximum likelihood estimation results

Salesforce	*Parameter*	*Estimate*	*Std. err*	*T-stat*	*P-value*
Reps	H	1596065	35655	44.8	<0.0001
($R^2 = 0.10$)	N	0.066	0.0083	8.0	<0.0001
Specialists	H	3579766	124986	28.6	<0.0001
($R^2 = 0.11$)	N	0.046	0.0011	4.44	<0.0001

Equation (3)). To overcome complications (i) and (ii) we make a few additional assumptions and then find a value of K that is consistent with the observed sales data. Details of this procedure are included in the next section. In practice, however, a firm's prices, cost per unit, and quantity sold are often observable, and these data can be used to more directly estimate the cost and demand parameters of the model.

6. Numerical experiments

Here we explore the interactions among the exogenous parameters of the model (e.g., learning rate, price sensitivity and cost parameters) and the endogenous decisions (staffing levels, prices, and routing decisions). In Section 6.1, we describe baseline parameters for the model that are derived from the analysis of Section 5 and are used in subsequent comparisons. In Section 6.2, we find the optimal staffing levels for a single-product system, given the baseline parameters, and we see how the optimal staffing level changes as the cost, learning curve, and productivity parameters change. In Section 6.3 we consider a firm with two products and two salesforces and investigate the relative benefits of a pooled salesforce versus a system with two completely specialized salesforces. We also examine the impact of the routing decision within a pooled salesforce by comparing two sales-lead assignment procedures: (i) random assignment; and (ii) the assignment of primary and secondary products to each salesperson, the hierarchical system described in Section 3.1.

6.1. *Baseline scenario*

The parameters for the baseline scenario are based upon the sales "Rep" data from the previous section.

- From the analysis above, $n = 0.066$ in the baseline model. However, we will vary n from 0.02 to 4.
- The tenure parameters (λ_1, λ_2, and λ_3) have been set so that the distribution of tenure found by a random arrival to the system is similar to the distribution of tenure in the Rep data set. In particular, the model is configured so that the average tenure of a sales Rep seen by a customer is just over 9 years, with a large percentage of relatively inexperienced salespeople: 52% below 4 years.
- $d = \$350$/day. This is the average rate of compensation in the industry.
- $\beta = 2$. Below we experiment with a range of β and describe the impact of changes in β.
- $\mu = 1$/day for all products and salespeople. According to the industry data, the average total time spent on a single lead is approximately 1 day.
- $\lambda = 40$/day. The size of an entire salesforce can often be measured in the thousands, but an offered load (λ/μ) of sales leads equivalent to 40 salespeople corresponds to a medium-sized regional salesforce for a single product. In Section 6.2, we will consider the salesforce for a single product taken in isolation, with $\lambda = 40$/day, while in Section 6.3 we apply our model to two products sold by two salesforces. In the two-product case we assume that all parameters for each product and salesforce are equal to the baseline parameters described here, except that $\lambda_A = \lambda_B = 20$/day (for a *total* load of 40 on the system). To simplify the exposition we are reporting results of experiments with a completely symmetric system in which the parameters for each salesforce and product are the same. We have conducted numerous experiments with asymmetric systems without revealing any major additional insights.

Unfortunately, the data set described in Section 5 does not contain sufficient information to find c or K. The model developed from the industry data does indicate that the average Rep, given essentially infinite experience, can earn \$1600000 in annual revenue. Under our assumption that the average lead requires 1 day of work, and assuming 250 workdays/year, these most experienced Reps average \$6400 in revenue per lead. To use this information to find the maximum possible *quantity* of product sold per lead (K) and the cost per unit (c), we must make two additional assumptions. Assume that firm A: (i) uses the optimal price, as described in Section 4; and (ii) earns a 25% margin on its sales (including the cost of the sale force itself). Then, $K = 7.2$ and $c = \$1120$ are the only parameter values that are *consistent* with these assumptions, the parameters above, and the observed maximum revenue of \$6400/lead. These values were found by "reverse-engineering" the model described in Sections 3 and 4. While useful as a baseline, we will also experiment with a range of both K and c.

6.2. *A single product*

First we consider a firm with a single product to sell and a single salesforce. For the single-product case, the objective function of Equation (7) has a single term in the summation, and the subscripts i and j are removed (e.g., K_{ij} replaced by K). Throughput and utilization statistics are calculated from the Erlang-B formula. Given the baseline parameters, we find the optimal price (Equation (9)) and total profit (Equation (7)) as a function of salesforce size, S. The results are shown in Fig. 4. The optimal price is

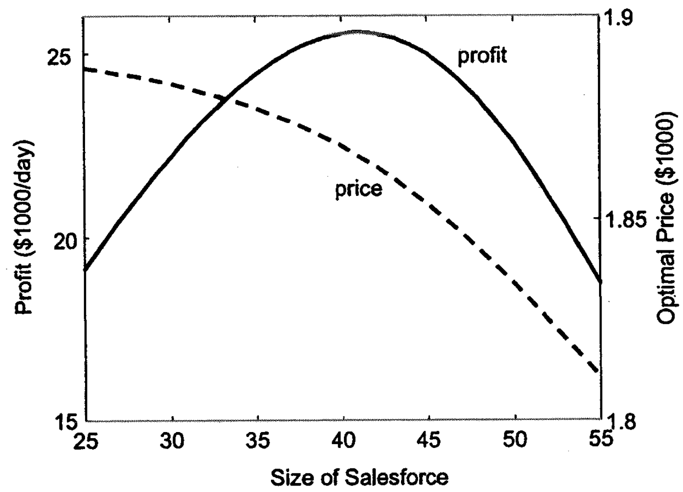

Fig. 4. Profit and price as staffing varies in the baseline model.

$1870/unit, the profit-maximizing staffing level is 41 sales representatives, and the optimal profit is $25 000/day. We examined the objective function for hundreds of cases and in each case the objective function was unimodal. This was true for both the one-product and two-product scenarios. However, to be thorough, all results presented here were found by searching for the optimum over the entire range of reasonable staffing configurations.

In Fig. 4, large unit profits explain the rapid rise in profitability on the left-hand side of the graph: as we add salespeople, throughput, r, rises and profits increase. This increase is partially balanced by the cost of each additional sales representative. However, in this baseline case $d = \$350$, and the decrease in profits on the right-hand side of the graph is much more rapid than the rate implied by d. In this case, the primary cost of additional servers is the decrease in the utilization of each server and the concurrent decrease in experience. As utilization decreases, both the demand-curve intercept and the optimal price decrease (see Equation (9)). Here the utilization effect first described at the end of Section 3 has a strong impact on both the profit and the size of the optimal salesforce. We also see, in Fig. 4, that the optimal price decreases with the size of the salesforce. We found that this was the case in all our numerical experiments. This observation is consistent with intuition: as we increase staffing we want to increase volume and therefore must decrease prices.

To determine the impact of cost, productivity and learning parameters on the optimal levels of staffing, we varied each model parameter over a wide range around the baseline value described above. For example, β took on values from one-half to three. With $\beta = 0.5$, the customer is barely price sensitive and the firm's profit margin is high (80%), given the other baseline parameters. With $\beta = 3$, the consumer is extremely price sensitive, and given the other baseline parameters the firm cannot be profitable (in this case, the optimal size of the salesforce is zero).

Table 3. Effect of exogenous variables on the optimal salesforce size

Parameter	K	c	β	d	n
Direction of change of the *optimal salesforce size* as the parameter increases	+	−	−	−	+

For each parameter combination, we solved for the optimal salesforce size. In all, we conducted over 100 000 of these numerical experiments (contact the authors for a detailed description of the parameters used). For five of the parameters the impact on optimal staffing levels was monotonic. These results are presented in Table 3. In the table, a + (−) indicates that the optimal salesforce size is non-decreasing (non-increasing) as the parameter increases.

These results are intuitive. The impact of K, the productivity parameter, is positive for salesforce size. Because K increases the demand intercept term, an increase in K leads to an increase in salesforce productivity thereby making the addition of salespeople profitable. On the other hand, an increase in the cost of the product, c, has a negative impact on salesforce size. Again, this is because an increase in product cost decreases the marginal revenue gained by adding an additional salesperson. A similar pattern is seen for the price sensitivity parameter, β: as price sensitivity increases, price goes down, and this diminishes the marginal revenue of each salesperson. This leads to a decrease in the salesforce size. An increase in d, the cost of a salesperson, reduces the optimal salesforce size. Finally, an increase in n both increases salesforce productivity and reduces the impact of the utilization effect described above. Both of these effects lead to an increase in salesforce size. However, we will see in the next section that the optimal salesforce size may not be monotone in n when the salesforce handles two products, rather than one product.

In addition, the existence of a learning curve for the salesforce can affect the impact of changes in other parameters; if altering a parameter changes staffing levels, then learning can dampen this effect. For example, in Fig. 5 we see the optimal salesforce size under a range of compensation rates for our model with learning (the solid line) and for a model without learning (the dashed line). In the model without learning, salesforce productivity is fixed so that the two models have the same optimal staffing level, given the baseline parameters. The figure shows how an increase in the cost per salesperson leads to a decline in the optimal salesforce size (as suggested in Table 3), and the figure also shows that the rate of decline is much more gradual, given employee learning. This is because any reduction in staffing also increases utilization. Therefore, in an environment with a learning curve, the marginal contribution of each salesperson is larger and the optimal number of salespeople remains high as d grows. We found a similar effect as we varied

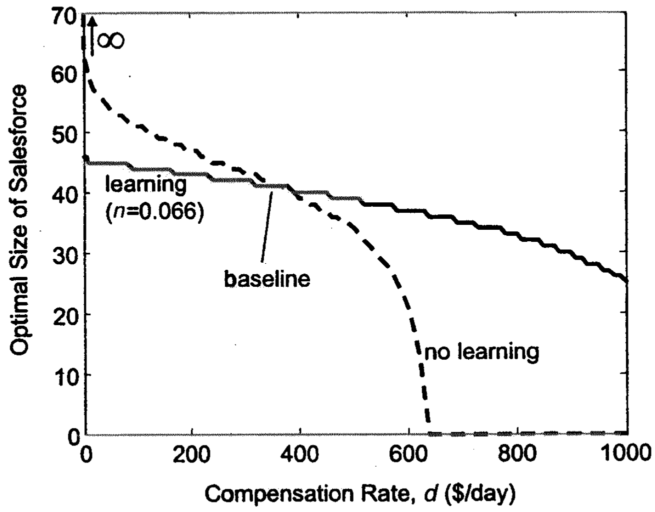

Fig. 5. Optimal size of the salesforce as the cost of a salesperson, d, varies.

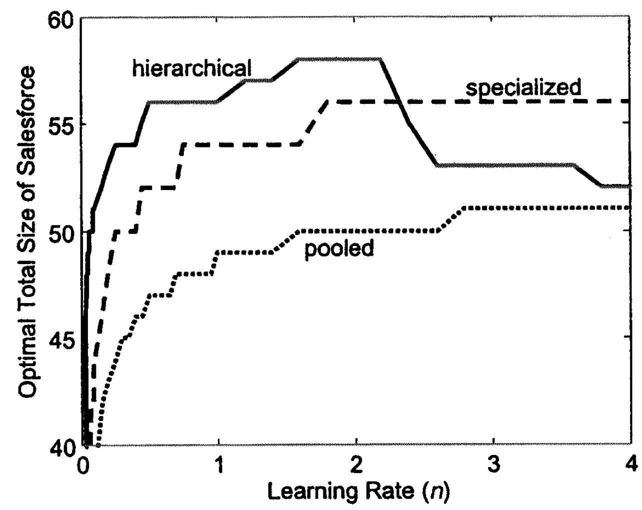

Fig. 6. Comparing salesforce sizes as the learning rate varies.

parameters β and c. In each case, the presence of learning moderates the impact of changes in the other parameters.

6.3. *Two products*

In this section, we consider a firm with two distinct products and salesforces. We consider three options for managing these salesforces: (i) specialized; (ii) hierarchical; and (iii) pooled. Under the *specialized* structure, each salesforce has exclusive selling rights for one of the products and does not receive leads for the other. Therefore, the two products and their salesforces can be managed independently, as if there were only one product and one salesforce. Under the *hierarchical* structure we assume that each product has a primary salesforce, but if all the primary salespeople are busy then an arriving lead is passed to the other, secondary salesforce. Finally, under the *pooled* structure there is a single salesforce responsible for both products, and arriving leads are assigned randomly to an available salesperson. Note that both the hierarchical and pooled systems take advantage of queueing economies of scale; in fact, these two systems have identical throughputs, and the overall average utilization of a salesperson is the same in both systems. However, the two systems differ in their routing, how they make the initial assignment of leads to salespeople, and this leads to significant differences in optimal staffing levels and profitability.

In Fig. 6 we plot the optimal staffing level as a function of the learning rate parameter n for the three salesforce structures. As in a traditional staffing problem, economies of scale in the pooled system lead to a smaller workforce than the specialized system. In addition, the optimal salesforce sizes of both the pooled and specialized systems increase with n, as suggested in the previous section's experiments with one product. However, staffing for the hierarchical system rises, and then falls, as n rises. The hierarchical system has a larger salesforce than the specialized system when n is low, and the optimal staffing level of the hierarchical system can be dramatically higher than that of the pooled system, even though both of these systems benefit from economies of scale. Why?

The staffing pattern for the hierarchical system is due to the presence of experience-based learning in the model. When n is low, a salesperson with little experience is extremely unproductive, so experience gained by a salesperson must be focused on one product to maximize learning. In the specialized system, the learning is focused by design, but in the hierarchical system each salesperson receives overflow leads for their secondary products. To prevent this overflow, it is optimal when n is low to staff each salesforce at higher levels in the hierarchical system than in the specialized system. For high n, however, most salespeople reach the plateau of the learning curve for both primary and secondary products, and the staffing level for the hierarchical system is close to the level for the simple pooled system.

The dynamics described above also have an effect on pricing. In Fig. 7, we plot the optimal price as a function of the learning rates for all three systems. As the learning rate increases, the optimal price increases, because each salesperson becomes much more effective; increased learning leads to a rise in the demand curve. We also see that the specialized system has higher prices than the pooled system because the specialized system has higher-skilled salespeople, producing a higher demand curve.

Finally, we examine which system is preferable, for a given value of n. Figures 8 and 9 compare the profitability of the three systems. For the lowest values of n, none of the systems are profitable. In this case, for a wide range of n the specialized system is more profitable because of that system's ability to focus its salespeople on a single product. Only with

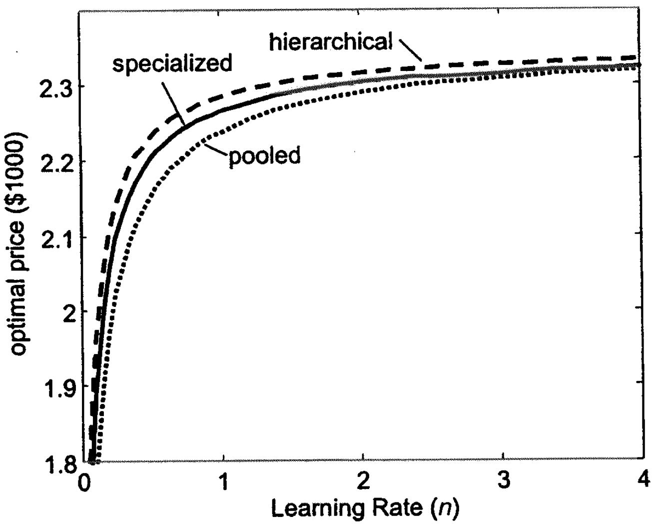

Fig. 7. Optimal prices as n varies.

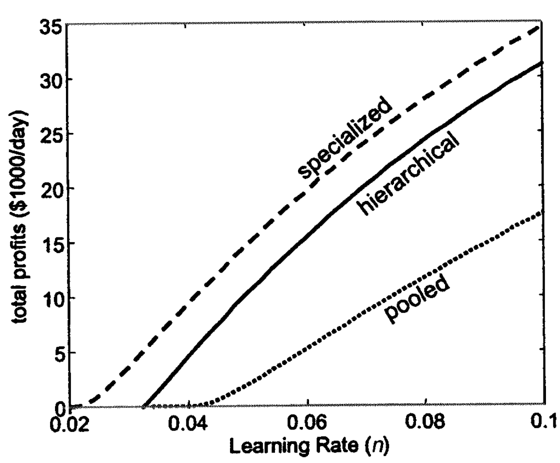

Fig. 8. Profits for each system, low n .

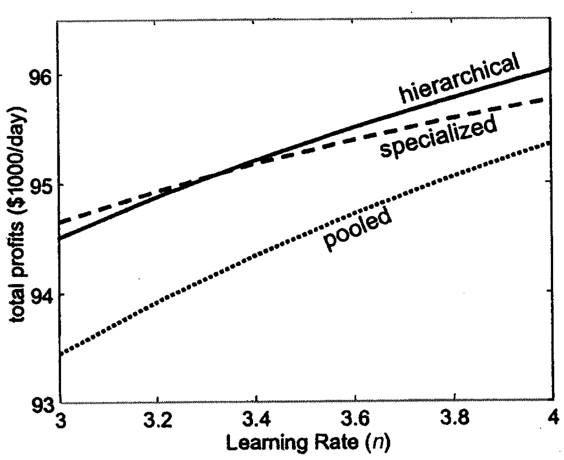

Fig. 9. Profits for each system, high n.

the highest values of n do economies of scale dominate, and the hierarchical system is preferred. The "crossing point," is shown in Fig. 9, and occurs at $n = 3.3$, a learning rate that enables a salesperson to reach 96% of the maximum sales effectiveness within 1 month. Such a rapid rise in the learning curve could be due to a variety of factors, including the simplicity of the product and market, an extremely effective salesforce training program, or the ability of the firm to consistently hire salespeople who are already experienced in selling the firm's product.

For no value of n is the pooled system best, and its performance is often significantly worse than the alternatives. For low n, pooling prevents the salespeople from learning enough about either product. For the highest values of n, the hierarchical system takes advantage of economies of scale while allowing each salesforce to focus on its primary product.

Of course, the relative profitability of the three types of systems depends upon all the parameters. For example, reducing the system size and/or increasing the expected tenure length increases the benefits of the hierarchical system, while increasing the wage d amplifies the differences between the specialized system and the other two systems. In addition, increasing d lowers the crossing point where the profits of the specialized and hierarchical systems are equal. If d is doubled from $350/day to $700/day, the crossing point moves down to $n = 2.0$. With this parameter, a salesperson achieves 86% of maximum sales effectiveness within 1 month and reaches 98% after 2 months.

7. Discussion and conclusions

In this section, we summarize our major findings, discuss the model's contributions and limitations, and point out directions for future work. We have applied a previous model (Pinker and Shumsky, 2000) combining learning, experience, and service systems to an important new application. For many products, the salesforce's effectiveness is linked to their knowledge of the product and the customers, and much of this knowledge can only be obtained through experience. This model is the first to show how this learning curve affects salesforce management. To apply the model in an appropriate manner we have extended it to include pricing, endogenous service levels, and a routing scheme that more accurately reflects salesforce management practice.

By making the service level endogenous we are able to see how optimal staffing responds to changes to cost parameters. For example, our results in Fig. 5 show that when salesforce skill is viewed as fixed and independent of staffing levels, changing labor costs lead to much stronger effects on staffing than when experience-based learning is taken into account. By modeling pricing we are able to study the interaction among learning, staffing levels and prices. The results indicate that higher prices work best with smaller, specialized and thus more experienced salesforces while lower prices correspond to larger, less specialized, and thus less experienced salesforces.

By modeling a more complex routing of sales leads to salespeople we obtain the result that, with learning, pooling of workers may lead to optimal staffing levels that are higher than when workers specialize. This contradicts the conventional wisdom that the economy of scale provided by pooling reduces staffing requirements. We also show that, with learning, pooling using a hierarchical routing scheme is always preferred to pure pooling that randomly assigns leads to salespeople. Finally, in this paper, we use data collected from the salesforce of a large manufacturer, and fit the learning-curve and tenure-process parameters of the model to this data.

While this paper has contributed to the literature it is not without limitations. One extension of the current model would be to add a cost for lost leads to the profit function. Within the objective function of the model Equation (1), including such a cost is equivalent to including some additional revenue for each sale, revenue that does not depend upon the price. Therefore, a significant cost for each lost lead would increase the value of throughput, increase the significance of queueing economies of scale, and thus increase the value of the hierarchical model over the specialist approach.

There are additional limitations of the model that are opportunities for future research. One area of improvement would be to model a N-product salesforce with more complex spill-over effects. Another improvement to the realism of the model would be to explicitly model how salespeople work multiple leads simultaneously rather than sequentially. In doing this one could draw upon the literature on the performance analysis of shared processors and polling systems. It would also be interesting to allow the salesperson to be a more active participant in the system, making both effort and pricing decisions, as in agency theoretic models. We have modeled the arrival process as exogenous, but in some environments, such as new or volatile markets in which the customer base is growing or changing quickly, the same salespeople are generating the leads and turning them into sales. It would be interesting to model how salespeople decide to allocate effort between generating leads and following leads up, within the context of a model of staffing with learning effects. We recognize that this is not an easy task because, among other complications, the service rates of the salespeople would be endogenous.

References

Badiru, A. B. (1992) Computational survey of univariate and multivariate learning curve models. *IEEE Transactions on Engineering Management*, **39**, 176–188.

Basu, A., Lal, R., Srinivasan, V. and Staelin, R. (1985) Salesforce compensation plans: an agency theoretic perspective. *Marketing Science*, **4**, 267–291.

Bhardwaj, P. (2001) Delegating pricing decisions. *Marketing Science*, **20**(2), 143–169.

Dewan, S. and Mendelson, H. (1990) User delay costs and internal pricing for a service facility. *Management Science*, **36**, 1502–1517.

Gross, D. and Harris C. (1985) Fundamentals of Queueing Theory. John Wiley and Sons,

Joseph, K. and Thevaranjan, A. (1998) Monitoring and incentives in sales organizations: an agency-theoretic perspective. *Marketing Science*, **17**(2), 107–124.

Lal, R. and Srinivasan, V. (1993) Compensation plans for single and multi-product salesforces: an application of the Holmstrom-Milgrom model. *Management Science*, **39**, 777–793.

Lodish, L. (1976) Assigning salesmen to accounts to maximize profits. *Journal of Marketing Research*, **13**, 440–444.

Lodish, L. (1980) A user oriented model for salesforce size, product and market allocation decisions. *Journal of Marketing*, **44**, 70–78.

Lodish, L., Curtis, E., Ness, E., and Simpson, M.K. (1988) Sales force sizing and deployment using a decision calculus model at Syntex Laboratories. *Interfaces*, **18**, 5–20.

Lucas, H., Weinberg, C. and Clowes, K. (1975) Sales response as a function of territorial potential and sales representative workload. *Journal of Marketing Research*, **12**, 298–305.

Mantrala, M., Sinha, P. and Zoltners, A. (1992) Impact of resource allocation rules on marketing investment-level decisions and profitability. *Journal of Marketing Research*, **29**(2), 162–176.

Montgomery, D. and Urban, G. (1969) *Management Science in Marketing*, Prentice Hall, Englewood Cliffs, NJ.

Pinker, E.J. and Shumsky, R.A. (2000) The efficiency-quality trade-off of cross-trained workers. *Manufacturing & Service Operations Management*, **2**(1), 32–48.

Rangaswamy, A., Sinha, P. and Zoltners, A. (1990) An integrated model-based approach for sales force structuring. *Marketing Science*, **9**(4), 279–299.

Rao, R. (1990) Compensating heterogeneous salesforces—some explicit solutions. *Marketing Science*, **9**(4), 319–341.

So, K. and Song, J. (1998) Price, delivery time guarantees and capacity selection. *European Journal of Operational Research*, **111**, 28–49.

Stidham, S. (1992) Pricing and capacity decisions for a service facility: stability and multiple local optima. *Management Science*, **38**(8), 1121–1139.

Yelle, L. (1979) The learning curve: historical review and comprehensive study. *Decision Sciences*, **10**, 302–328.

Zoltners, A. (1976) Integer programming models for sales territory alignment to maximize profits. *Journal of Marketing Research*, **13**, 426–430.

Biographies

Sanjog Misra is an Assistant Professor of Marketing in the William E. Simon School of Business Administration at the University of Rochester. He has a M.S. in Statistics and a Ph.D. in Marketing from the State University of New York at Buffalo. His research interests include the application of analytical and econometric methods to understand sales and marketing related problems. His recent research papers focus on the design of efficient salesforce structures, salesforce compensation, salesforce training and the use of the salesforce as a strategic competitive tool. He is also interested in the development of econometric methods relating to limited dependent variables and discrete choice models. His recent consulting projects include Adventis, IMS Health, Lucent Technologies and Xerox.

Edieal J. Pinker is an Associate Professor of Computers and Information Systems at the Simon School of Business of the University of Rochester, in Rochester NY. He earned his M.S. and Ph.D. in Operations Research from the Massachusetts Institute of Technology. His research interests are on issues of business process design, electronic commerce and workforce management. He has published research on the use of contingent workforces, cross-training and experience-based learning in service sector environments as it applies to work and workflow design. He is currently studying the optimal design of medical practices, how ERP systems enable information sharing between firms engaged in B-to-B electronic commerce, and online auctions. His research has been published in *Management Science*, *M&SOM*, *EJOR*, *IIE Transactions* and *CACM*. He has consulted for the United States Postal Service, the financial services industry, and the auto industry. He serves on the editorial review boards of *M&SOM* and *POMS*.

Robert A. Shumsky is an Associate Professor of Operations Management at the Simon School of Business, University of Rochester. He has research and teaching interests in the modeling and control of service systems. His current research focuses on the dynamic use of flexible capacity, the use of incentives for operational control of service systems, and the application of revenue management under competition. He has conducted research on the US air traffic management system and studied transportation operations for the Massachusetts Port Authority and the Federal Aviation Administration. His research has been published in *Management Science*, *Operations Research*, *M&SOM*, and *Air Traffic Control Quarterly*. He is an associate editor for *Management Science* and *Operations Research* and serves on the editorial review boards of *M&SOM* and the *Journal of Revenue and Pricing Management*. He earned his M.S. and Ph.D. in Operations Research from the Massachusetts Institute of Technology.