REVENUE MANAGEMENT THROUGH
DYNAMIC CROSS-SELLING IN CALL CENTERS

E. Lerzan Örmeci† and O. Zeynep Akşin††

† Department of Industrial Engineering
Koç University
34450, Sarıyer - İstanbul, Turkey, Tel:+90 212 338 1534, Fax: +90 212 338 1548

†† Corresponding author. College of Administrative Sciences and Economics
Koç University
34450, Sarıyer - İstanbul, Turkey, Tel:+90 212 338 1545, Fax:+90 212 338 1653
lormeci@ku.edu.tr, zaksin@ku.edu.tr

last revised, April 2009
Revenue Management through Dynamic Cross-Selling in Call Centers

E. Lerzan Örmeci † O. Zeynep Akşin †

April 14, 2009

This paper models the cross-selling problem of a call center as a dynamic service rate control problem. The question of when and to whom to cross-sell is explored using this model. The analysis shows that under the optimal policies cross-selling targets may be a function of the operational system state. Sufficient conditions are established for the existence of preferred calls; i.e. calls that will always generate a cross-sell attempt. These provide guidelines in segment formation for marketing managers, and lead to a static heuristic policy. Numerical analysis establishes the value of different types of information, and different types of automation available for cross-selling. Increased staffing for the same call volume is shown to have a positive and increasing return on revenue generation via cross-selling, suggesting the need to staff for lower loads in call centers that aim to be revenue generators. The proposed heuristic leads to near optimal performance in a wide range of settings.

Keywords: Dynamic programming, call center, customer relationship management, cross-selling heuristic.

1 Introduction

Many firms in mature industries, like the financial services industry, resort to growth by deepening customer relationships and making them more profitable. A significant part of this profitability comes from revenues generated by the sale of additional products and services to existing customers, who are typically better sales prospects compared to new customers. Given the growing dislike among consumers for telemarketing, this type of selling is increasingly being performed via cross-selling and up-selling initiatives (Krebsbach 2002). According to Kamakura et al. (1991) cross-selling is emerging as one of the important customer relationship management (CRM) tools used to strengthen relationships.

---

*Koç University, Department of Industrial Engineering
†Koç University, College of Administrative Sciences and Economics
Particularly in banking, inbound call centers are an important point of contact with the customer, where this type of selling takes place. Given the increasing percentage of these centers that are organized as profit centers, focus is shifting to cross-selling. According to a Wells Fargo executive (The Economist 2004) 80% of the bank’s growth is coming from selling additional products to existing customers. As the leader in cross-selling, this bank’s customers hold an average of a little over four products per household. Given that an average American household has sixteen financial products, the opportunity for cross-selling growth in this industry is apparent.

A major concern for managers is identifying the right person and the right time to attempt a sale. While it is believed that cross-selling ensures that customers acquire multiple products of a firm, improves customer retention (Marple and Zimmerman 1999) and reduces customer churn, excessive selling can motivate a customer to switch (Kamakura et al. 2003). Database marketing techniques that address this issue are being developed (Paas and Kuijlen 2001, Kamakura et al. 1991, 2003), and software that helps insurance agents or bankers cross-sell more effectively is becoming more common (Insurance Advocate 2003, American Banker 2003) as companies embrace this tactic.

Cross-selling in a call center requires a customer service agent to transform an inbound service call into a sales call. According to an article in the Call Center Magazine (2003), call centers can use integrated predictive analysis and service automation software to make real-time recommendations to banking customers. However, in a review of existing products Chambers (2002) states that real-time automation is relatively immature and many products offer only the option of setting preset business rules that make promotion recommendations based on previously captured and stored data. Common practice is to segment the customer base into groups based on their sales potential, and to target sales to high potential segments. In the absence of real-time automation, the customer service representative will use segment based estimates to determine whether it is appropriate to attempt a cross-sell to a particular call.

Irrespective of the type of automation in place, a cross-sell attempt in a call center implies additional talk time from the agent. Thus cross-selling will influence the load of a call center, as documented in Akşin and Harker (1999). The biggest challenge of a call center manager is to manage the tension between costs and customer service. While for the long-term this corresponds to determining the right number of service representatives to hire, in the shorter term it is resolved through capacity allocation.
This description of cross-selling in a call center identifies a key challenge for managers: When should a cross-selling attempt be made such that revenue generation is maximized while congestion costs are kept as low as possible? Current practice identifies off-peak times during the day for cross-selling. However it is clear that a dynamic policy will utilize valuable capacity more effectively. It is this question of dynamic capacity allocation that motivates the research herein.

In a more general setting, Güneş and Akşin (2004) consider this tradeoff between revenue generation and service costs, and analyze the interaction with a market segmentation decision and server incentives. That paper does not consider the queue state information in its optimization of the problem. Other studies have considered dynamic cross-selling in call centers (Byers and So, 2007; Armony and Gurvich, 2006; Gurvich et al., 2007).

Byers and So (2007) model a call center as a single server Markovian queue, and show that a threshold type control policy is optimal. Their analysis compares the performance of cross-selling policies that consider queue state information as well as customer profile information for two customer segments. It is illustrated numerically that using both types of information is most desirable in highly congested systems with high sales revenues. We build on these results by analyzing a more realistic setting. Focusing on a multi-server loss model, and modeling sales revenues by a continuous revenue distribution (thus modeling each customer as a separate segment), we perform an extensive numerical analysis demonstrating when different types of information (queue state information versus revenue distribution information) are preferred and when multiple state-dependent thresholds may be replaced by simple threshold-based controls obtained via a heuristic. The proposed heuristic control rule is the first in the literature combining simplicity and ease of implementation with close to optimal performance in a variety of settings.

Two recent papers Armony and Gurvich (2006) and Gurvich et al. (2007) explore the cross-selling control problem in conjunction with staffing. The joint optimization of cross-selling and staffing is the main contribution of these papers. We do not explicitly model the staffing problem, however explore the interaction with staffing through a set of numerical examples. An important distinction between the models in these papers comes from the way service is modeled. In the models by Armony and Gurvich (2006) and Gurvich et al. (2007), service is modeled as consisting of two phases, thus allowing for the possibility of delaying cross-sell decisions to the second phase of service. This provides the possibility for finer control of the service versus sales tradeoff. Nevertheless, in many large call centers
that rely on automation and human resource policies that do not support judgement by service representatives, the single phase service as modeled herein is relevant. The value of servers having the capability to incorporate information from the first phase of service into their cross-sell decision in the second phase remains to be tested in future research. While features of the models herein and those in Armony and Gurvich (2006) and Gurvich et al. (2007) differ in several respects, both approaches lead to the recommendation of threshold type controls for cross-selling. A numerical analysis allows us to characterize operating environments that justify the use of complex dynamic controls as well as the value that can be generated through the use of different types of information.

We formulate the cross-selling problem as a dynamic service rate control problem in a multi-server loss system, where a customer’s revenue potential is taken as a random variable. Weber and Stidham (1987) consider dynamic optimal control of service rates in queueing systems. Our work also relates to dynamic pricing and admission control problems, where random rewards have been considered (Ghoneim and Stidham, 1985, Örmeci et al. 2002, Gans and Savin 2007). All of these papers show the existence of optimal threshold policies. Örmeci et al. (2002) and Gans and Savin (2007) further characterize conditions for the existence of preferred jobs, where preferred jobs are defined as those calls which are always admitted to the system whenever there is at least one available server. The fact that all calls have to be admitted for service and the decision is to choose a service rate, as opposed to admission control with pre-determined service rates for each class, constitutes the key difference of the model studied herein.

In this paper, preferred calls are defined as calls which always receive a cross-sell attempt, and whenever all calls of a segment are preferred, that segment is labeled as preferred. Sufficient conditions for preferred calls are used as a building block for the heuristic control rule, that is proposed as an alternative to dynamic controls. The technique introduced by Örmeci et al. (2002) is used to derive sufficient conditions for the system to have preferred calls and segments. The analysis brings insights on both operations and marketing: To guide the operational policies on cross-selling, an easy-to-implement heuristic based on these sufficient conditions is developed. The threshold identified by the heuristic provides marketing managers a definition of segments that incorporates both customer value and operational considerations.

In summary, a loss model with random revenues is presented in this paper, to formulate the dynamic cross-selling problem of a call center. Structural properties of this model lead
to the first major contribution of the paper in the form of a heuristic control rule that is threshold based, requires less system information than the optimal control, and furthermore is shown to perform well in a wide array of settings. An extensive numerical analysis, which constitutes the second major contribution of the paper, demonstrates how the performance of dynamic cross-selling is tied to staffing at the call center and to revenue information availability. The analysis further identifies environments where simpler technology and/or simpler controls can be used without sacrificing revenue performance.

The paper is organized as follows: The basic model is given in Section 2. Section 3 presents sufficient conditions for the existence of preferred calls, as well as a cross-selling heuristic. In Section 4 we generalize our discussion of the model to features found in practice. In Section 5, a set of numerical examples motivated by a real call center, allow us to demonstrate the value of dynamic controls, the availability of different types of information, and identify settings where these would be useful. The robustness of the heuristic under different operating conditions is illustrated. An analysis of the interaction between cross-selling and staffing in call centers illustrates that slack capacity leads to increasing returns in terms of revenue generation. Provided that the additional staffing costs remain below the revenue that is generated, this suggests that cross-selling call centers should be designed to operate in a lower load regime. We end with concluding remarks.

2 A Dynamic Cross-Selling Model

In order to study the dynamic cross-selling problem, we model an inbound call center as a loss system with \( c \) identical parallel servers. Customers arrive to the system according to a Poisson process with rate \( \lambda \). The center is primarily concerned with service provision, so treats all call requests that are not blocked due to capacity limitations. Each time a call arrives to a system with at least one available server, there will be a decision to attempt a cross-sell or not. If the decision is not to cross-sell, then the call is treated as a service call with a fixed revenue \( r \), which requires an exponentially distributed service time with rate \( \mu \). If the decision is to attempt a cross-sell, the call will generate a random revenue \( r + \rho \) and the service time will be distributed exponentially with rate \( \mu_1 = \mu - k \), where \( k \) is a positive constant that reflects the impact of the selling activity on the duration of the call. The objective of the call center is to maximize the total expected discounted revenues over an infinite time horizon and/or maximize long-run average revenue of the center.
The random revenue, $\rho$, corresponds to the so-called score of the incoming customer. In marketing terminology, the score of a customer reflects customer profitability or revenue potential estimate having two basic components: (1) the probability of purchase (2) likely purchase volume and profit margin or revenue information. These estimates are generated for the whole customer database using data on individual customer characteristics, such as demographic, past purchase, and psychographic information. In our model, we assume that customers’ random revenues are independent of each other and have a common probability distribution $F$. We do not have any specific assumptions on the distribution $F$, except that the random revenues are bounded, i.e., $0 \leq \underline{\rho} \leq \rho \leq \bar{\rho} < \infty$.

Now we present a discrete time Markov decision process for the system. The formulation is given for the objective of maximizing total expected discounted returns over a finite time horizon with a discount rate $\beta$. Define $x_1$ and $x_2$ as the number of cross-sell calls and the number of service calls in the system, respectively. The system changes its state at service completions and at arrivals. Because the decision to attempt a cross-sell depends on the customer revenue potential, the state at arrival instants is defined as $(x; \rho) = (x_1, x_2; \rho)$. At all other times the state information is described by $x = (x_1, x_2)$.

Discounting is interpreted as exponential failures with rate $\beta$ (for the equivalence of discounting and an exponential deadline, see e.g., Walrand 1988). We use uniformization and normalization to build a discrete time equivalent of the original system by assuming $\lambda + c\mu + \beta = 1$, so that the system will be observed at exponentially distributed intervals with mean 1. There will be an arrival with probability $\lambda$. A real service completion due to a standard service call occurs with probability $x_2\mu$ and due to a cross-sell call with probability $x_1\mu_1$, while a fictitious service completion occurs with probability $c\mu - x_2\mu - x_1\mu_1$.

We next develop the optimality equations for the transformed system over a finite horizon. Let $u_n(x)$ and $v_n(x; \rho)$ be the maximal expected reward, starting in state $x$ and $(x; \rho)$, respectively, until $n$ transitions occur. Upon a call arrival, if all servers are busy, then the call has to be rejected. Otherwise, computing $v_n(x; \rho)$ requires comparison of two actions: cross-selling the incoming call to move to state $x + e_1$ with a reward of $\rho + r$, and giving the standard service to enter state $x + e_2$ with a reward of $r$, where $e_j$ is the two-dimensional unit vector with the $j^{th}$ component equal to 1. The optimality equations of this model are as follows. For $x_1 + x_2 < c$:

$$v_n(x; \rho) = \max\{\rho + r + u_n(x + e_1), r + u_n(x + e_2)\}$$

(1)
\[ u_{n+1}(x) = \lambda E[v_n(x; \rho)] + x_2 \mu u_n(x - e_2) + x_1 \mu_1 u_n(x - e_1) \]

\[ + (c\mu - x_2 \mu - x_1 \mu_1) u_n(x), \]

where we set \( u_n(-1, x_2) = u_n(0, x_2), u_n(x_1, -1) = u_n(x_1, 0), \) and \( E[.] \) denotes the expectation with respect to random reward \( \rho. \) For \( x_1 + x_2 = c, \) no calls can be accepted so that \( v_n(x; \rho) = u_n(x). \) We assume that ties in equation (1) are broken by selecting pure service.

The model is analyzed with the objective of maximizing total expected \( \beta \)-discounted reward for a finite number of transitions, \( n. \) The state space of the model is compact, while the action space in each state is finite. Since \( \rho \) is bounded, the expectation \( E(v_n(x; \rho)) \) is finite for all possible policies, ensuring that the optimal value functions are bounded. Then, \( u(x) = \lim_{n \to \infty} u_n(x) \) exists, where \( u(x) \) denotes the total expected \( \beta \)-discounted reward of a system which starts in state \( x \) over an infinite horizon (Hernandez-Lerma and Lasserre, 1999). The existence of \( u(x) \) implies that the optimal solution can be found using value iteration. The same analysis applies under the long-run average gain criterion so that all stated results are also valid in that case. We refer to the technical report Örmeci and Akşin (2004) for the details.

### 3 Analysis of Being Preferred

The main purpose of this section is to characterize preferred calls, with the intent of using this characterization in constructing a heuristic control rule for the cross-selling problem. Preferred calls are those that always generate a cross-sell attempt, irrespective of system state. Existence of such calls is of interest since their identification paves the way for the derivation of simpler controls that do not need to consider system occupancy dynamics like in the optimal policies. We aim to derive sufficient conditions for having preferred calls. Our starting point to derive sufficient conditions is equation (1). We observe that it is optimal to cross-sell in a state \( (x; \rho) \) if and only if \( u_n(x + e_2) - u_n(x + e_1) < \rho, \) since we choose “pure service” when both actions are optimal. A call with a reward of \( \rho \) is said to be preferred if \( u_n(x + e_2) - u_n(x + e_1) < \rho \) for all \( x. \) In words, whenever a call generates a random revenue that is high enough, it is always optimal to cross-sell to this call, regardless of the number of cross-sell and service calls in the system.
3.1 Identifying Preferred Calls

Making use of the idea of preferred calls, we next derive a threshold to identify such calls. The main duty of the call center we consider is to answer regular service calls to obtain a fixed reward of \( r \). The call center cannot fulfill this duty if all servers are busy. On the other hand, cross-sell decisions increase the total load of the call center, which, in turn, increases the probability of having all servers busy. Hence, in cross-sell decisions comparison of the fixed reward \( r \) with the random reward \( \rho \) plays an important role. However, this is not the only aspect. A cross-sell decision for the current call affects the potential cross-sell decision for a future call, which requires a comparison of the random reward offered now with the random reward to be offered in the future. These two comparisons, i.e., current random reward versus future fixed reward \( r \) and current versus future random rewards, together will determine the sufficient conditions for the existence of preferred calls. We specify these conditions by deriving upper bounds on the relative benefits \( u_n(x + e_2) - u_n(x + e_1) \). The derivation compares the current reward with the maximum possible revenue that can be obtained from an incoming call, namely with \( \bar{\rho} + r \), regardless of how probable it is to obtain such a revenue. This allows us to obtain upper bounds that are valid for all sample paths. These bounds are then used to derive an explicit threshold on the cross-sell revenues. The following proposition identifies certain preferred calls using this threshold. Its proof can be found in the Appendix.

**Proposition 1** All calls bringing at least a reward of \( \zeta \) are preferred, where

\[
\zeta = \frac{\lambda}{\lambda + \mu_1 + \beta} \frac{\mu - \mu_1}{\mu + \beta} (\bar{\rho} + r).
\]

Essentially, Proposition 1 compares the cross-sell revenue of a current call with the potential maximum revenue that can be obtained from a cross-sell call, \( \bar{\rho} + r \), by explicitly specifying a threshold on the cross-sell revenues, \( \zeta \). For all \( \rho > \zeta \), it is always optimal to cross-sell, regardless of the system state. In other words, these calls are preferred. It is also possible to have preferred calls with \( \rho < \zeta \) but we cannot detect them.

The threshold \( \zeta \) decreases when the total arrival rate \( (\lambda) \), the highest possible cross-sell revenue \( (\bar{\rho}) \), the base service revenue \( (r) \) and the service rate difference due to cross-selling \( (k = \mu - \mu_1) \) decrease, in fact even when only the regular service rate \( (\mu) \) decreases or only the cross-sell service rate \( \mu_1 \) increases. Since the arrival rates in call centers are usually high, the first term of the threshold \( \zeta \) will be very close to 1. Therefore, being preferred mostly
depends on the service time and revenue characteristics of the calls. Finally we note that by construction, the threshold is independent of the revenue distribution and only depends on the upper bound of revenues.

### 3.2 A Heuristic for Cross-Selling

Proposition 1 presents two pieces of practical information: (i) Structure of a good static policy to be specified by a threshold on the random revenues, so that it is always optimal to cross-sell a call which brings a random reward higher than this threshold, and (ii) a starting point to specify a good threshold for such a policy. In principle, an optimal threshold can be computed for this type of static policy. However, this computation is very complex, since it requires carrying out a line search over the range of random rewards to find a threshold: For each threshold value considered in the search, a Markov reward model needs to be solved. This requires computation of the expected return using the random revenue distribution, for a given threshold value. Instead of computing an optimal threshold, we use the information in (ii) to find a good threshold value for this type of policy. More explicitly, we use the expression for the threshold $\zeta$, to construct a heuristic which identifies static policies that exploit the structure of the problem.

The purpose of introducing this heuristic is two-fold: one is to suggest easy-to-implement and effective cross-selling policies to call centers that wish to avoid the complexities associated with state dependent controls that emerge from a dynamic optimization. The other is to provide the threshold identified by the heuristic as feedback from the operations side to the marketing department enabling segment formation that combines customer value as well as operational considerations.

We concentrate on the second term of the threshold $\zeta$, described as a function:

$$\alpha(t) = \frac{\mu - \mu_1}{\mu + \beta} (t + r).$$

For the long-run average criterion, $\beta$ is set to 0 in this expression. Recall that $t = \bar{\rho}$ in $\zeta$, and Proposition 1 compares a current cross-sell revenue with the maximum possible revenue, $\bar{\rho} + r$. The shape of the probability distribution, or even the probability of observing the revenue $\bar{\rho} + r$, does not appear in this expression since the derivation follows from sufficient conditions. Obviously, the threshold $\zeta$ is overestimating the optimal threshold of this type of static policy. In order to make up for this discrepancy, we replace the benchmark, $\bar{\rho}$, with a more attainable revenue. For this purpose, consider the following cross-selling policy: The
current customer brings a cross-sell reward of $R$, and we take it as a service call. Moreover, we cross-sell the future customers only if their cross-sell reward is strictly greater than $R$. Now define $h(R)$ as the expected reward that one can gain from cross-selling a future customer who offers a revenue of more than $R$. Then, $h(R)$ is given by the following conditional expectation: $h(R) = E[\rho|\rho > R]$. In the resulting cross-sell policy, the current revenue, $R$, is compared with $\alpha(h(R))$, the “expected future gain”. The heuristic decides to cross-sell the current customer only if $R > \alpha(h(R))$.

This rule can be specified with only one parameter in most systems, since the equation $R = \alpha(h(R))$ has a solution for a wide range of probability distributions of $\rho$. We set $R^*$ as the solution of $R = \alpha(h(R))$. $R^*$ facilitates the implementation of the heuristic. While requiring some pre-processing to compute the threshold, it necessitates no further real-time computation. The computation of $R^*$ involves only the service rates ($\mu, \mu_1$) and the discounting rate ($\beta$). In particular, the heuristic provides a control that is independent of the current state ($x_1, x_2$) of the system, as well as the arrival rate $\lambda$ and the number of servers $c$. The practical implications of these properties on call center operations are discussed further in Section 4.

In typical marketing practice, where the decision of whom to cross-sell is based on an individual’s score $\rho$, the cross-selling policy that emerges is to identify a threshold such that whenever an individual’s $\rho$ is above this threshold the decision is to sell. The parameter of the heuristic, $R^*$, is a good candidate for such a threshold. Unlike the thresholds determined in practice, $R^*$ includes additional information from the operations of the call center. We evaluate the performance of the heuristic, or in marketing terms the segment formed by $R^*$, in the numerical analysis section.

4 Considerations from Practice and Model Implications

Our model analyzes the dynamic control of a call center with the primary goal of providing service and the secondary goal of cross-selling new products. Like most call center management problems, it involves issues of relevance to both marketing and operations. In the following subsections, we compare and contrast our model and its analysis to reality and current cross-selling practices, organized around marketing and operations assumptions.
4.1 Marketing Issues

On the marketing side, we consider two main issues from practice, namely the desire to segment customers and the role of marketing automation. We present different systems, depending on the choices available to marketing managers in making cross-selling decisions along these two issues. We show that all of these systems can be represented by our model, and discuss the implications of our results in all of them.

To enable comparison of our optimal cross-selling policies to current practice, we now describe how customers are typically classified by managers for cross-selling purposes. When answering the question of whom to cross-sell, typical approaches do not take the operational capacity tradeoff into account. Instead, marketing managers select cross-selling targets using data on individual customer characteristics, which are converted to scores as we have explained in Section 2. Cross-selling policies are either individual based, where targeting is done at the individual level, or segment-based, where targeting is done at the group level. Segments are formed by aggregating customers in homogeneous groups according to their cross-sell potentials, based on demographic, past purchase, and psychographic information.

The model as presented up to this point, can be directly related to a setting with individual level targeting based on the score \( \rho \). Sometimes marketing managers prefer establishing cross-sell decisions for entire groups of customers. Segment-based cross-selling policies will emerge where customers are grouped into \textit{good-to-sell} \( (s = G) \) and \textit{bad-to-sell} \( (s = B) \) categories up-front, where \( s \) is an index denoting the segment. Whenever a customer is identified as type-\( G \) a cross-sell will be attempted. In typical practice, these policies are formed from a value perspective, disregarding the ties to the operational aspect. To establish the link between the optimal dynamic policies and segment-based policies, we need to represent this setting in our model.

We assume that \( s \)-segment customers call according to a Poisson process with rate \( \lambda_s \) for \( s = G, B \), so that the total arrival rate \( \lambda = \lambda_G + \lambda_B \). Moreover, customers of segment \( s \) generate a random revenue \( \rho_s \), which follows a probability distribution \( F_s \). We do not have any specific assumptions on the distributions \( F_s \), except for \( 0 \leq \underline{\rho}_s \leq \rho_s \leq \bar{\rho}_s < \infty \). The random revenue distribution \( F \) in our model is given by \( F(y) = (\lambda_G F_G(y) + \lambda_B F_B(y)) / \lambda \). Hence, our model can be used to analyze segment-based cross-selling policies. In particular, the sufficient condition for being preferred is valid for the segmented case.

We consider two possible scenarios for the segmentation, a discrete and an overlapping
Scenario 1: $\bar{\rho}_B \leq \hat{\rho}_B \leq \bar{\rho}_G \leq \hat{\rho}_G$

Scenario 2: $\bar{\rho}_B \leq \bar{\rho}_G < \hat{\rho}_B \leq \hat{\rho}_G$

Segments are formed using statistical procedures that take data on certain customer characteristics as input and group customers based on these. Scenario 1 represents a case where the procedure clearly distinguished between two groups of customers in terms of their cross-selling potential, placing them in non-overlapping segments. Scenario 2 represents the more realistic case. The overlap can be thought of as the type one and type two errors of this statistical procedure. Our numerical analysis will build on these two scenarios in order to observe the effect of such errors, as well as of the segmentation.

Availability of different types of marketing automation results in different estimation capabilities for call centers. One where the revenue potential is estimated individually for each customer in real-time and another where only historical averages for the customer base or for a particular segment are estimated. The system we have described so far will be referred to as the model with revenue realizations due to the underlying assumption about the possibility to estimate the revenue potential of a customer at the time of the decision: In this case, it is assumed that a server can observe the realization of the random revenue $\rho$ before taking the decision to cross-sell or not. The model with revenue realizations represents the case of a call center where marketing and technology support is such that as soon as a customer call arrives, the system is capable of identifying the customer and displaying its revenue realization. This represents a setting with software that has real-time automation capability as described in Section 1. However, it is also possible that the server takes a
decision based on historical segment analysis and then the revenue realization is observed at service completion. We label this model as the *model with expected revenues*. In such a setting, managers look at historical data, segment customers and then just use the average for the segment each time a member of this segment is served. How much revenue is eventually realized from a particular customer will be determined at call completion.

When the revenue potentials cannot be estimated for individual customers, all customers of a segment are assumed to bring the same expected reward. Since the aim of our model is to maximize the total expected reward, the optimal decision will be a result of comparing the expected reward of a segment with the additional load that the system will observe due to cross-selling. Then, the model with expected revenues becomes a special case of the model with random revenues: when all customers constitute one population, \( \rho = \bar{\rho} = E[\rho] \), while in the segmented case, \( \rho_s = \bar{\rho}_s = E[\rho_s] \) for \( s = G, B \). Now, we can specify the sufficient condition for being preferred in the model with expected revenues and segmentation, as a corollary to Proposition 1. Before stating this result we note that a segment is called preferred if all calls of that segment always generate a cross-sell attempt.

**Corollary 1** (For the model with expected revenues)

i. When customers form one population, all customers are preferred if \( \lambda (\mu - \mu_1) r < (\lambda + \mu + \beta)(\mu_1 + \beta)E[\rho] \).

ii. When customers form two populations: (a) segment-G is preferred if \( (\lambda_G + \lambda_B)(\mu - \mu_1) r < (\lambda_G + \lambda_B + \mu + \beta)(\mu_1 + \beta)E[\rho_G] \).

(b) all customers are preferred if: \( (\lambda_G + \lambda_B)(\mu - \mu_1)(E[\rho_G] + r) < (\lambda_G + \lambda_B + \mu_1 + \beta)(\mu + \beta)E[\rho_B] \).

### 4.2 Operational Issues

We model call centers as loss systems with \( c \) identical parallel servers. The arrival rate \( \lambda \) and \( c \) are assumed to be constant. Moreover, we assume that we always fully observe the state of the system. However call centers may not possess these characteristics. They are queueing, rather than loss systems, the arrival rates, and so the number of servers vary during the day, and it is not always possible to observe the number of service and cross-sell calls in progress. This section discusses these issues in some detail, while the next section numerically evaluates the performance of our model in various settings that do not agree with the basic assumptions of our model. Finally, we note that the service-sales tradeoff addressed
by our model will fundamentally change under different staffing levels. Optimal dynamic
cross-selling policies vary with the staffing levels. We can think of the extreme cases: a very
low staffing level will induce a policy of never cross-selling, whereas a very high level will
urge the system to always cross-sell. How the optimal dynamic policies change in between
these two extremes is not clear, so it is not possible to predict the effect of staffing levels on
the maximal expected revenues before computing the exact policies and the corresponding
values. As a result, we solve the dynamic cross-selling model for fixed staffing levels, and
then compare the solutions, i.e., we numerically analyze the relationship between staffing
levels and the value generated from cross selling.

First consider the choice of modeling a call center as a loss system instead of a queueing
system. The main motivation for this assumption is one of tractability, since a Markov
decision process based analysis of a model with waiting space is significantly more complex.
Call centers have been modeled as loss systems before, particularly to simplify analysis in
staffing or routing problems with multi-skill servers (Chevalier et al. 2004, Chevalier and van
illustrate that a loss system captures the basic performance characteristics of corresponding
queueing systems quite well. In the cross-selling context, the main issue arises from the fact
that the waiting time of the customers will affect their tendency to buy, which cannot be
explicitly considered in a loss model. At times of heavy load, the loss system will block many
customers, thus reducing congestion in the system. This benefit is not present in a model
with an infinite waiting room, however abandonment from the queue will dampen this effect
to some extent. These aspects of the loss assumption are explored numerically in Section 5.
We use the heuristic to numerically verify that treating the call center as a loss system does
not distort our results.

The arrival pattern of calls is a determining factor in the operations of call centers. Arrival
rates are known to differ by day-of-the-week and time-of-day. In addition a number of recent
studies show that arrival rates are random (see e.g. Avramidis et al., 2004, Brown et al.,
2005, Whitt, 2006). Since call centers have to keep up with fluctuating and random arrival
rates, the number of operators also changes during the day. As a result, implementing
optimal dynamic cross-selling policies may require solving our model for every 15 or 30-
minute time interval. Another requirement for implementation is to observe the number of
service and cross-sell calls in progress, information that cannot be obtained in settings where
the technology in place does not enable real time monitoring of the system state. This is
frequently the case in multi-site structures. Our heuristic is affected by none of these, since it provides a static policy which depends neither on the parameters, \( \lambda \) and \( c \), nor on the state of the system \((x_1, x_2)\). The simple threshold-based control rule facilitates operational implementation further, due to the ease with which it can be explained and motivated to managers and call center agents.

The heuristic, as well as dynamic optimization require estimation of the service rates \( \mu \) and \( \mu_1 \). In an environment where cross-selling is pursued, the observed service times will clearly depend on the cross-selling policies being used, and care needs to be taken in estimating these rates. The numerical analysis, by considering the effect of different service rates on the value from cross-selling, provides qualitative understanding of the role that these rates play.

5 Numerical Analysis

In this section we have several objectives. We first develop a set of numerical examples that reflect real call center operations as closely as possible, and that cover a wide range of possible customer types and operating regimes. We explore interactions with staffing in Section 5.1. This analysis demonstrates that cross-selling benefits are closely tied to staffing. In Section 5.2, we perform an analysis to guide managerial choices along technological issues. In particular, the analysis aims to characterize environments where real-time marketing automation can be foregone for simpler technology, or where dynamic controls can be replaced by simpler controls as in the proposed heuristic. The comparisons in this section, as illustrated in Figure 3, allow us to identify the value of different types of information, the value of real-time marketing automation, and the performance of the heuristic, all as a function of staffing. We end the section with an analysis of the heuristic’s robustness to various modeling assumptions made in the paper.

All of the test problems are solved to maximize the long-run average revenue, i.e., \( \beta = 0 \), using the value iteration algorithm. In order to evaluate the effects of different marketing policies discussed in Section 4, the test suite is based on two segments of customers with known revenue distributions. We assume that segment-\( s \) revenues are uniformly distributed with parameters \((\rho_s, \bar{\rho}_s)\) for \( s = G, B \). For the model with expected revenues, we find the corresponding expected value of the revenues for the given parameters.

In all these test problems, we vary base service call lengths \( (\mu) \), cross-sell durations
(\(k\) or \(\mu_1\)), basic service revenues \((r)\), the upper and lower bounds on random revenues \((\rho_B, \bar{\rho}_B, \rho_G, \bar{\rho}_G)\), the size of the call center in terms of number of servers \((c)\), and the load \((\lambda/c\mu)\) of the call center. Of these, service and cross-sell durations are customer-related features, but are also expected to depend on the organizational structure of the call center as documented in Aksin and Harker (1999). Revenues are customer-related features, while load and the number of servers are supply-related features driven by operational decisions.

Case 1 (C1) is motivated by a real retail banking call center. Using data estimated for this center, the average length of a service call is taken as 2.7 minutes. Based on a scenario analysis performed at this bank, increase in talk times can be 27% or 220% depending on organizational assumptions being made. While our analysis, as summarized in Section 6 considers both of these settings, we only report numerical values obtained with a 220% increase herein. In the test suite, a 27% increase in talk time leads to small differences in value (around one to two percent) between dynamic optimization and simpler approaches to the cross-selling problem. We qualitatively summarize our results for this case in our concluding remarks.

Average revenue from a call with cross-selling is estimated as 75 units. We take this as the lowest value of the upper limit of \(G\)-segment revenues and consider three values \(\bar{\rho}_G \in \{75, 125, 175\}\). Revenues from basic service calls are taken as \(r = 1\) to reflect the situation that service calls generate very low revenue compared to sales in this call center. For all our examples \(\rho_B = 0, \bar{\rho}_B \in \{0.3\bar{\rho}_G, 0.6\bar{\rho}_G, 0.9\bar{\rho}_G\}\) and \(\rho_G \in \{0.3\bar{\rho}_G, 0.6\bar{\rho}_G, 0.9\bar{\rho}_G\}\) to allow us to cover a wide range of possible revenue distributions, wide or narrow, with or without overlap.

We consider call centers with \(c \in \{100, 150, 200\}\) servers. Call volumes are obtained to ensure four different loads \((\lambda/c\mu)\), characterizing quality driven (0.75), quality-efficiency driven (0.9), and efficiency driven (1.05, 1.2) centers (for precise definitions of these terms see Gans et al. 2003). The total call volume is split in three different ways, such that \(\lambda_G/\lambda_B \in \{0.1, 0.25, 0.4\}\), representing settings with increasing proportions of calls offering a cross-selling potential. To simplify the presentation, we only report a subset of our results for selected values of these parameters.

Case 2 (C2) is constructed taking C1 as a basis. It is assumed to represent the setting of an insurance call center or an investment bank, where the basic service length is longer. We assume an average service call length of 5.5 minutes, taking the instance of a major insurance call center in the U.S. The only other difference from C1 is the assumption that basic service
calls generate more revenue, so that $r = 20$. The two cases result in 3888 problem instances.

The impact of dynamic optimization is explored from a value perspective. We consider different managerial levers that can be used to affect the value from cross-selling. The first of these is an operational lever. We provide a preliminary analysis on the effect of available capacity on optimal policies which shows us that managers can enhance the value of dynamic optimization through an appropriate choice of staffing. Other managerial levers relate to information availability and thus depend on the technology in place. These are grouped under technology related factors below.

5.1 Exploring the Role of Operational Factors

In order to explore the interaction between cross-selling and staffing, we consider a set of examples where we keep the same loads for given arrival rates and vary the number of servers. We explore the interaction with staffing for the service rates of C1 and C2, for loads of 75 %, 90 %, 105 %, and 120 %, and for two sets of revenue distributions from the test suite. One where additional revenues can be quite high with $\rho_B = 0$, $\bar{\rho}_B = 157.5$, $\rho_G = 157.5$, $\bar{\rho}_G = 175$ and another with lower additional revenues with $\rho_B = 0$, $\bar{\rho}_B = 22.5$, $\rho_G = 22.5$, $\bar{\rho}_G = 75$. Recall that for C1 the revenue from service calls is 1 whereas this value is 20 for C2. We concentrate on the settings where the proportion of good segment calls is 1 whereas this value is 0.1.

For the C1 setting, taking total arrival rate $\lambda = 40$, we compute the number of servers as 144, 120, 102, 90 for the loads 75 %, 90 %, 105 %, and 120 % respectively. Similarly for the C2 setting, taking total arrival rate $\lambda = 21$, we obtain the number of servers as 154, 128, 110, 96 for the loads 75 %, 90 %, 105 %, and 120 % respectively. Figure 2 compares total revenues

Figure 2: Revenue as a function of the four loads 75 % (1), 90 % (2), 105 % (3), 120 % (4) for C1 in (a) and C2 in (b)
obtained under no cross-selling and optimal cross-selling for all of these cases. We observe that compared to no cross-selling, optimal cross-selling increases revenues for a given staffing level. This benefit is higher for higher cross-sell revenue values, and also for the cases where the service revenue is lower as in C1. As we increase the staffing level (or decrease the load) the additional revenue benefit obtained from cross-selling vis-a-vis no cross-selling increases. In other words, revenues show an increasing marginal returns feature for percent reduction in load. As we go from a load of 120 % to 75 %, the revenue difference between the no cross-selling and the optimal cross-selling policies increases. The revenue increase for each additional server in C1 can be calculated as 2.08 per server per minute and 11.58 per server per minute for the low and high cross-sell revenue settings respectively, while in C2 these values are 1.32 per server per minute and 5.84 per server per minute respectively. For an estimated cost of an employee ranging between $0.5-$1 per minute (summing to an annual salary of around $62,400-$125000), one observes that an additional server costs well below the benefit that s/he will bring to the call center. In additional results not reported here for brevity, we perform a similar analysis for systems with finite buffers. Even though less revenue is generated in these systems compared to an equivalent loss system, we find revenue increases per additional server that qualitatively parallel these results. This illustrates that provided staffing costs are not too high, there are operational synergies between service levels and sales effort, suggesting that managers should staff according to a quality regime in order to maximize the benefits obtained from cross-selling. The demonstrated interaction between staffing and cross-selling control motivates our presentation of the subsequent numerical analysis, detailed as a function of the number of servers and the load.

5.2 Exploring the Role of Technology Driven Factors

The value comparisons we make in this section, along with which model (with revenue realizations or with expected revenues) and what type of control we use (dynamic or heuristic based static) for each comparison are illustrated in Figure 3. The value iteration algorithm is used to compute the expected gains of both the optimal dynamic control policies and the heuristic. We optimize the actions (i.e., to cross-sell or not) in each state \((x_1, x_2; \rho)\) for optimal dynamic control, whereas in the heuristic we only need to compute the expected gain since the decisions are determined by the thresholds, \(R^*\), which are computed off-line.
5.2.1 The Value of Real-Time Marketing Automation

We compare the performance of the model with expected revenues and the one with revenue realizations, the former using only average revenue information while the latter makes use of information on the entire revenue distribution. A comparison of their optimal revenues under dynamic optimization will quantify the value of having real-time automation, as is assumed for the revenue realizations case. Table 1 displays the averages for each call volume split and overall range of the ratio of the optimal gain from the model with expected revenues to the optimal gain in the model with revenue realization.

The effect of operational characteristics: The value of real-time automation does not show high sensitivity to the operational factors size and load that we consider (though it is relatively more sensitive to the latter). In most of the examples, the model with expected revenues performs slightly better in heavier loaded, larger centers.

The effect of call durations and segment characteristics: We observe that the difference in revenues between the two models can be as high as nineteen percent. In Scenario 2, $E[\rho_C]$ and $E[\rho_B]$ take closer values, thus making it harder for the model with expected revenues to be selective. Similarly, the case represented by C1 makes selective selling more difficult due to the less distinct cost of cross-selling implied by short call durations and low service revenues in this setting. Indeed, the model with expected revenues performs the best in the C2-Scenario 1 case. Finally, we note that the volume split between $G$-segment and $B$-segment customers has an impact, as demonstrated by the two averages in each case. The model with expected revenues performs better for higher proportions of $G$-segment calls. As
### 5.2.2 Performance of the Heuristic

We next implement the heuristic described in Section 3.2. Table 2 reports averages and the overall range of the ratio of the gain obtained with this heuristic to the optimal gain. This comparison captures the value of replacing the sophisticated dynamic control with a simple static rule, while making use of information on the entire revenue distribution in both cases. The heuristic approaches optimal performance on the average consistently across all operating environments considered.

**The effect of operational characteristics:** In these examples, we observe that the heuristic performance is better in larger and heavier loaded centers. To further explore the generalizability of this observation to small systems, an additional set of tests are per-

---

### Table 1: Ratio of the optimal gain with expected revenues to that with revenue realizations

<table>
<thead>
<tr>
<th>Case, load</th>
<th>c: volume split</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>avg. %</td>
<td>100: 0.1-0.4</td>
<td>200: 0.1-0.4</td>
</tr>
<tr>
<td>C1, 75 %</td>
<td>avg. %</td>
<td>84.97</td>
<td>84.97</td>
</tr>
<tr>
<td></td>
<td>range %</td>
<td>82-100</td>
<td>81-100</td>
</tr>
<tr>
<td>C1, 90 %</td>
<td>avg. %</td>
<td>85.97</td>
<td>85.97</td>
</tr>
<tr>
<td></td>
<td>range %</td>
<td>81-100</td>
<td>81-100</td>
</tr>
<tr>
<td>C1, 105 %</td>
<td>avg. %</td>
<td>85.97</td>
<td>85.97</td>
</tr>
<tr>
<td></td>
<td>range %</td>
<td>81-100</td>
<td>81-100</td>
</tr>
<tr>
<td>C2, 75 %</td>
<td>avg. %</td>
<td>93.99</td>
<td>93.99</td>
</tr>
<tr>
<td></td>
<td>range %</td>
<td>82-100</td>
<td>81-100</td>
</tr>
<tr>
<td>C2, 90 %</td>
<td>avg. %</td>
<td>95.99</td>
<td>95.99</td>
</tr>
<tr>
<td></td>
<td>range %</td>
<td>81-100</td>
<td>81-100</td>
</tr>
<tr>
<td>C2, 105 %</td>
<td>avg. %</td>
<td>95.99</td>
<td>95.99</td>
</tr>
<tr>
<td></td>
<td>range %</td>
<td>81-100</td>
<td>81-100</td>
</tr>
</tbody>
</table>

The two segment’s volumes approach each other there is less need for selective selling, which once again helps the model with expected revenues. All settings in these examples, except for the C2-Scenario 1 case, represent situations where real-time automation would add value, particularly for the 0.1-0.9 volume split case. Clearly segment characteristics interact with service durations in identifying the cases with highest value from real-time automation.
formed: In these, the number of servers takes values in \{10, 20, \ldots, 50\}, the volume split is \(\lambda_G/(\lambda_G + \lambda_B) = 0.1\), and all other parameters remain as before, resulting in 1080 instances.

For these systems with less than 50 servers, the performance of the heuristic improves for smaller systems: The ratio of the gain with the heuristic to the optimal gain changes in the range 87.74 - 100 \%, with averages of 92 \%, 99 \% for C1 and C2 of Scenario 1, respectively, and 94 \% and 99 \% for C1 and C2 of Scenario 2, respectively. A closer look at the test suite for small systems also reveals that the heuristic performs the worst under efficiency regimes. These seemingly contradictory observations, however, have an intuitive explanation: The heuristic cross-sells to all customers above a certain revenue. This implies that the proportion of customers to whom cross-selling attempts are made depends only on the revenue distribution (and not on arrival rates or the number of servers) and as a result remains a fixed proportion of the customers in all these systems. Optimal policies, on the other hand, change this proportion according to the level of resources as well as of the load. This proportion can be labeled as small in heavily loaded small systems, as medium for small systems under a quality regime and for large systems under a heavy load, and finally as large in large systems under a quality regime. As a result, the heuristic estimates this proportion well for larger systems when they operate under high loads and for smaller systems under a quality regime, while it under-utilizes the capacity when large systems operate in a quality regime, and over-utilizes it when small systems operate with a heavier load.

In summary, based on these results it is possible to state that the dynamic control is more valuable in large centers operating under a quality regime, and in small centers operating under an efficiency regime.

The effect of call durations and segment characteristics: The heuristic’s worst case behavior is in C1 under Scenario 1 when \(\lambda_G/(\lambda_G + \lambda_B) = 0.1\). All systems with performances less than 70\% correspond to problems with a wide range of \((\tilde{\rho}_G, \tilde{\rho}_B)\) with \(\tilde{\rho}_B = \tilde{\rho}_G\), and \(\lambda_G/(\lambda_G + \lambda_B) = 0.1\). In words, these systems receive a very heterogeneous \(G\)-segment from a revenue perspective, generating one tenth of all calls, and barely differentiated from the \(B\)-segment calls.

Comparing the heuristic to an alternative static policy: We compare the performance of the heuristic with the marketing policy which cross-sells only to the \(G\)-segment, referred to as the \(G\)-only policy. Table 3 shows that the \(G\)-only policy performs poorly, particularly when \(\lambda_G/(\lambda_G + \lambda_B) = 0.1\). Even in Scenario 1, where both the heuristic and the \(G\)-only policy are the same kind of threshold policy, the performance of \(G\)-only is inferior to that
of the heuristic. The difference between the performances of these two policies reflects the
device to using the information from the operations side in marketing decisions as highlighted
earlier in Section 3.2.

<table>
<thead>
<tr>
<th>Case, load</th>
<th>c: volume split</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>100: 0.1,0.4</td>
<td>200: 0.1,0.4</td>
</tr>
<tr>
<td>C1, 75 %</td>
<td>avg. % 82.98, 83.98</td>
<td>97.98, 98.98</td>
<td></td>
</tr>
<tr>
<td></td>
<td>range % 55-100, 55-100</td>
<td>96-100, 96-100</td>
<td></td>
</tr>
<tr>
<td>C1, 90 %</td>
<td>avg. % 85.98, 86.98</td>
<td>98.98, 98.98</td>
<td></td>
</tr>
<tr>
<td></td>
<td>range % 60-100, 61-100</td>
<td>96-100, 96-100</td>
<td></td>
</tr>
<tr>
<td>C1, 105 %</td>
<td>avg. % 86.98, 87.98</td>
<td>98.98, 98.98</td>
<td></td>
</tr>
<tr>
<td></td>
<td>range % 62-100, 63-100</td>
<td>96-100, 96-100</td>
<td></td>
</tr>
<tr>
<td>C2, 75 %</td>
<td>avg. % 91.98, 91.98</td>
<td>93.96, 93.97</td>
<td></td>
</tr>
<tr>
<td></td>
<td>range % 76-100, 77-100</td>
<td>75-99, 75-99</td>
<td></td>
</tr>
<tr>
<td>C2, 90 %</td>
<td>avg. % 95.98, 95.99</td>
<td>95.97, 96.97</td>
<td></td>
</tr>
<tr>
<td></td>
<td>range % 82-100, 83-100</td>
<td>84-99, 85-99</td>
<td></td>
</tr>
<tr>
<td>C2, 105 %</td>
<td>avg. % 96.99, 96.99</td>
<td>96.97, 97.98</td>
<td></td>
</tr>
<tr>
<td></td>
<td>range % 84-100, 85-100</td>
<td>87-99, 88-99</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Ratio of the gain with the heuristic to the optimal gain

5.2.3 Value of state versus revenue distribution information

The heuristic explicitly uses the probability distributions of the random revenues along with
the service rates, while completely ignoring the current state of the system. On the other
hand, the model with expected revenues uses the information about the current state of the
system, while it has no reference to the random revenue distributions. Hence, comparing their
performances relative to the optimal may show the value of these two kinds of information.
We make use of the results reported in Tables 1 and 2.

The average performance of the heuristic is almost always superior, where the most signif-
icant improvements are observed under Scenario 2. These represent settings with overlapping
segments, where clearly the value of revenue distribution information is important.

In terms of the worst case behavior, the expected revenue model generally performs
better, particularly in the C1-Scenario 1 setting. Except for the C1-Scenario 2 setting, its
Table 3: Ratio of the gain from $G$-only policy to the optimal gain

<table>
<thead>
<tr>
<th>Case, load</th>
<th>c: volume split</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>avg. %</td>
<td>range %</td>
<td>avg. %</td>
</tr>
<tr>
<td>C1, 75 %</td>
<td>61.96</td>
<td>40-100</td>
<td>62.97</td>
</tr>
<tr>
<td>C1, 90 %</td>
<td>66.97</td>
<td>43-100</td>
<td>67.97</td>
</tr>
<tr>
<td>C1, 105 %</td>
<td>67.97</td>
<td>44-100</td>
<td>68.97</td>
</tr>
<tr>
<td>C2, 75 %</td>
<td>91.99</td>
<td>66-100</td>
<td>91.99</td>
</tr>
<tr>
<td>C2, 90 %</td>
<td>94.99</td>
<td>71-100</td>
<td>94.99</td>
</tr>
<tr>
<td>C2, 105 %</td>
<td>95.99</td>
<td>72-100</td>
<td>95.99</td>
</tr>
</tbody>
</table>

worst case performance is also superior to that of the heuristic when loads of 75 % are coupled with a volume split of 0.1-0.9 in our examples.

We conclude that as long as one of the valuable information kinds, the current state of the system or the probability distributions of the revenues, is used effectively, the system will perform very closely to an optimal one. If a single type of information were available, then using the state information is more beneficial in the C1 case of Scenario 1, and more generally for centers operating in a quality regime with the good segment of customers generating a small volume of the calls. All other cases are better off relying on the revenue distribution information. From an implementation standpoint, the heuristic has an advantage over the model with expected revenues, by not requiring information on the current arrival rate and the current number of servers.

5.3 Heuristic Robustness to Modeling Assumptions

The first question we explore in this section is how well the heuristic, which was designed taking the structural characteristics of a loss system into account, performs when customers are allowed to wait in queue and to abandon if the wait exceeds their patience threshold.
The second question is related to the effect of this waiting on a customer’s propensity to buy. The setting with a finite buffer allows us to explore the difference between customers whose buying behavior is not affected by how long they wait and those who will be less likely to buy as their wait increases. To capture this difference, we consider two extremes in our cross-selling strategies: one which cross-sells to all customers if it is profitable to do so (which we call can-wait-cross-sell), and the other which attempts to cross-sell only to customers who do not wait (which we call no-wait-cross-sell). The first strategy implicitly assumes customers are indifferent to waiting times in their reactions to cross-selling, while the second strategy implies an assumption that customers will not buy if cross-sold after waiting for service.

We consider a call center with 50 servers and a buffer capacity of 15 with $G$-segment call volume proportions of 0.1 and 0.4. All other parameters are the same as before. In order to observe the effect of abandonment, we solve all problems with and without abandonment. In systems with abandonment, we take the abandonment rate as $3\mu$. Allowing for the two cross-sell strategies, this gives us a set of 3456 examples. In all the problems, the performance of the heuristic when used to control an $M/M/c/K+M$ system is compared to the optimal policy for the same system, where both policies are once again determined via value iteration.

Under the can-wait-cross-sell strategy, ignoring the queue has a negligibly small effect on the performance of the heuristic vis-a-vis the optimal gain. For brevity we do not report these results, but note that the averages as well as ranges observed for the different settings being considered are almost identical to those in Table 2, with differences of less than one percent. The overall average of the heuristic to the optimal gain is 97.5% in systems with no abandonment, while it is 97.3% in those with abandonment. Under no-wait-cross-sell, the performance of the heuristic improves, yielding an average of 98.8% in both with and without abandonment. This is due to the limited benefit that can be obtained from optimal cross-selling when the number of customers to be attempted is restricted. The suggested heuristic is very robust since it is performing well with finite buffers in a number of different settings. We can conclude that the structure of an optimal policy in a finite-waiting-room system is similar to that in a loss system, since the heuristic is based on the structural properties of optimal policies in loss systems. Computationally, derivation of the optimal dynamic policies in an $M/M/c/K+M$ queue becomes increasingly harder as system size and buffer size is increased, since the state space is three dimensional in this case. The heuristic is also useful in overcoming this curse of dimensionality.
6 Concluding Remarks

An extensive numerical analysis allows us to draw recommendations for managers in terms of choice of technology and cross-selling controls as a function of investment in capacity. Understanding these choices is important in practice: we show that in many settings, important savings in technology investment and/or managerial costs can be obtained without making significant sacrifices in revenue performance. Throughout, we have assumed that whenever the performance of cross-selling using simpler technology and/or controls results in performance that is within 95% of the optimal, these simpler approaches may be preferred. We consider medium to large call centers with more than 50 servers for the recommendations made in this section.

For environments with a low talk time increase due to cross-selling (represented by a 27% increase in talk times in the test suite), it is possible to obtain almost all the benefits of dynamic optimization by making use of the proposed heuristic.

When the G-segment calls constitute an important part of the call volume (as represented by the volume split 0.4-0.6 in the test suite) making use of real-time automation coupled with the heuristic as a control leads to near optimal performance in all settings. If furthermore segmentation capability is such that discrete segments can be formed, managers may forego real-time automation, but use dynamic controls to obtain close to optimal performance. For this volume split (0.4-0.6) we thus find that making use of only one type of information (revenue distribution or system state) ensures performance within 95% of the optimal in all of the settings considered.

For a select G-segment (0.1-0.9 volume split) and higher increase in call content due to cross-selling, the recommendations may vary as a function of customer characteristics and operational factors, as illustrated in Figure 4. The only environment where real-time automation is not critical for high revenue performance is for a call center operating under the quality efficiency or efficiency regime, with long call durations (C2), and having the capability of forming discrete segments. In all other environments, real-time automation is essential in obtaining within 95% of the optimal performance. While revenue distribution information seems quite valuable in this setting with a select group of G-type calls, dynamic controls are not always necessary to ensure within 95% of optimal performance. Except for the C2 case operating in a quality regime, all settings with overlapping segments may simplify cross-selling initiatives by replacing dynamic controls with simpler controls based
on the heuristic. Whenever discrete segments occur together with call characteristics as in C1, or whenever a call center operating under a quality regime has call characteristics as in C2, coupling real-time automation with dynamic optimization will generate significant value. We conclude with the observation that most settings for a call center designed to operate in a quality regime will benefit from the most sophisticated cross-selling controls, combining revenue distribution information with dynamic control policies. On the other hand for the most part, if investment in excess capacity has not been made, not much value is added from expensive cross-selling controls. When this is the case, the heuristic provides a viable control alternative.

It is possible to show the optimality of threshold-type policies for this cross-selling problem, if concavity of the value functions $u_n$ in $x_1$ are assumed (see Örmeci and Akşin, 2004). Concavity of value functions even in simpler systems cannot be shown (Örmeci et al., 2002), although no example of such systems is reported that violated concavity. In the large test suite herein, we have similarly not observed a single value function which is not concave.

**Acknowledgement:** This research was partially supported by The Scientific & Technological Research Council of Turkey, TÜBİTAK. Zeynep Akşin thanks the MEDS Department, Kellogg School of Management, Northwestern University, where part of this work was done.
Appendix: Proof of Proposition 1

In this section, we restrict ourselves to systems with two revenue distributions, and prove Proposition 1 only for these systems. The proof for single revenue distribution is straightforward, and is omitted.

The following lemma identifies simple bounds on the effects of an additional customer in the system and of changing a regular-service customer to a cross-sell customer. The system collects the regular service fee as well as the random cross-sell revenue in the beginning of the service. Hence, the customers who are already in the system bring only a burden by preventing the admission of new customers and/or the decision of a cross-sell. This is represented in part (a) of Lemma 1. Similarly, changing a regular-service customer to a cross-sell customer induces a positive burden due to the difference in their service times, which corresponds to the first inequality of part (b). To understand the second inequality of part (b), we consider two systems, one in state \((x + e_2)\) and the other in \((x + e_1)\). Under the conditions specified in the proof of the lemma, these systems will preserve the difference between them except for the two events: (1) service completion of the extra regular-service customer and (2) system close-down. The difference between the two systems vanishes with the system close-down (with probability \(\beta\)) and when both additional customers finish their services (with probability \(\mu_1\)), whereas the difference converts to the effect of an additional cross-sell customer when the regular-service customer departs from the system leaving the cross-sell customer behind (with probability \(\mu - \mu_1\)). The second inequality of part (b) quantifies this observation.

**Lemma 1** For all \(x\) with \(x_1 + x_2 < c\) and for all \(n \geq 0\): (a) \(0 \leq u_n(x) - u_n(x + e_i)\) for \(i = 1, 2\); (b) \(0 \leq u_n(x + e_2) - u_n(x + e_1) \leq \frac{\mu - \mu_1}{\mu + \beta}(u_n(x) - u_n(x + e_1))\).

**Proof.** We prove only the second inequality of part (b) in detail. The other statements can be proved similarly. Setting \(u_0(x) = 0\) for all \(x\) satisfies the inequality, so now assume that it is true for \(n\), and consider \(n + 1\).

We use a coupling argument: Assume that system A starts in state \(x + e_2\) and system B starts in \(x + e_1\). We couple the two systems via the service and interarrival times, so that all the departure and arrival times are the same in both systems except for the additional calls. Moreover, the additional regular call, say call \(d_2\), in system A is coupled with the additional cross-sell call, say call \(d_1\), in system B, so that both \(d_1\) and \(d_2\) leaves the system
with probability $\mu_1$, and $d_1$ remains in the system while $d_2$ leaves with probability $\mu - \mu_1$. Then, we can let system A follow the optimal policy and system B imitate all the decisions of system A. Defining $u^B_n(x + e_2)$ as the expected discounted return of system B, we have:

$$u_{n+1}(x + e_2) - u_{n+1}(x + e_1) \leq u_{n+1}(x + e_2) - u^B_{n+1}(x + e_1)$$

$$\leq (\lambda + (c-1)\mu)(u_n(x + e_2) - u_n(x + e_1)) + (\mu - \mu_1)(u_n(x) - u_n(x + e_1))$$

$$\leq (1 - \mu - \beta)\frac{\mu - \mu_1}{\mu + \beta}(u_n(x) - u_n(x + e_1)) + (\mu - \mu_1)(u_n(x) - u_n(x + e_1))$$

$$= \frac{\mu - \mu_1}{\mu + \beta}(u_n(x) - u_n(x + e_1)),$$

where the first inequality follows from the description of the policies for systems A and B, the second inequality is due to the coupling, and the third follows by uniformization and the induction hypothesis.

**Proof of Proposition 1:** We first note that if $\zeta > \bar{\rho}$, there are no preferred calls. Hence, we assume, without loss of generality, $\zeta \leq \bar{\rho}$. The statement is satisfied for $u_0(x) = 0$ for all $x \in \tilde{S}$. Now we assume that it is true for $n$, and show it holds for $n+1$.

For $\zeta$ to be a threshold on the cross-sell revenues, it has to satisfy $u_n(x + e_2) - u_n(x + e_1) \leq \zeta$. By Lemma 1, it is enough to show $u_n(x) - u_n(x + e_1) \leq \eta$, where we let $\eta = \frac{\mu + \beta}{\mu - \mu_1}\zeta$. For this purpose, we use a coupling argument, similar to the one for Lemma 1, so that we assume system A is in state $x$ and system B in $x + e_1$ in period $n+1$. System A follows an optimal policy, whereas system B imitates system A whenever possible. Note that it cannot imitate system A only if all of its servers are busy so that it has to reject all incoming calls. Then, letting $u^B_{n+1}(x + e_1)$ be the total expected discounted reward of system B in period $n+1$:

$$u_{n+1}(x) - u_{n+1}(x + e_1) \leq u_{n+1}(x) - u^B_{n+1}(x + e_1)$$

$$\leq \lambda \max\{u_n(x) - u_n(x + e_1), u_n(x + e_2) - u_n(x + e_1) + r, \rho + r\} + (c\mu - \mu_1)(u_n(x) - u_n(x + e_1))$$

$$\leq \lambda \max\{\eta, \zeta + r, \bar{\rho} + r\} + (c\mu - \mu_1)\eta \leq \lambda \max\{\eta, \bar{\rho} + r\} + (c\mu - \mu_1)\eta$$

where the first inequality is due to the optimality of $u_n$’s, the second due to coupling, the third due to the definition of $\bar{\rho}$, the induction hypothesis and Lemma 1, and the fourth
inequality follows since \( \zeta \leq \bar{\rho} \). If \( \eta \geq \bar{\rho} + r \), the statement is proven; otherwise:

\[
\begin{align*}
  u_{n+1}(x) - u_{n+1}(x + e_1) & \leq \lambda(\bar{\rho} + r) + (1 - \lambda - \mu_1 - \beta)\eta \\
  &= (\lambda + \mu_1 + \beta)\eta + (1 - \lambda - \mu_1 - \beta)\eta = \eta
\end{align*}
\]

where the inequality is due to uniformization and normalization, and the equality follows by the definition of \( \eta \). Thus, the statement is true for all \( x \) with \( x_1 + x_2 < c \) and for all \( n \geq 0 \).

References

Akşin O. Z. and Harker P. T. To Sell or Not to Sell: Determining the Tradeoffs Between Service and Sales in Retail Banking Phone Center. Journal of Service Research 1999; 2(1); 19-33.


Chambers, B. Marketing Automation Connects with Contact Centers, Information Week 2002; 873; 72.


Franx, G. J., G. Koole, and A. Pot. Approximating Multi-Skill Blocking Systems by Hyper-Exponential Decomposition, Performance Evaluation 2006; 63(8); 799-824.
Insurance Advocate 2003. AIG Small Business Center has New Web-Enabled Risk Profiling Tool for Agents’/Brokers’ Use; 114(21); 8.
Krebsbach, K. Banks’ New Mantra: Cross Sell Like Crazy, Bank Investment Consultant 2002; 10(8); 22-28.

Whitt, W. Staffing a Call Center with Uncertain Arrival Rate and Absenteeism. Production and Operations Management 2006; 15:1; 88-102.