Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Managing Hotel Reservations with Uncertain Arrivals

Gabriel R. Bitran, Stephen M. Gilbert,

MANAGING HOTEL RESERVATIONS WITH UNCERTAIN ARRIVALS

GABRIEL R. BITRAN

Massachusetts Institute of Technology, Cambridge, Massachusetts

STEPHEN M. GILBERT

Case Western Reserve University, Cleveland, Ohio

(Received July 1992; revisions received August 1993, February 1994; accepted August 1994)

Based on our interactions with managers at two large hotels, we present a realistic model of the hotel reservation problem. Unlike traditional models, ours does not assume that all customers arrive simultaneously on the targeted booking date. We explain why this assumption may not be appropriate for the hotel industry and develop a model of reservation booking which explicitly includes the room allocation decisions which are made on the targeted booking date. Based on observations of how the problem is solved in practice as well as the insights gained from this analysis, we develop simple heuristic procedures for accepting reservations. Computational results demonstrate that these heuristics perform well relative to an upper bound that is based on perfect information about reservations requests and customer arrivals.

The problem that hotel managers face in managing the demand for reservations is quite difficult. They must respond to requests for a variety of types of reservations to balance the expected loss of revenue from unsold rooms against both the tangible and intangible costs of "walking" customers, i.e., failing to honor reservations. Since customers who have booked rooms may either cancel or fail to show up with some probability, it is not unusual for hotels to "overbook," i.e., to accept more reservations than they have rooms. Doing this successfully requires a thorough understanding of market dynamics and consumer behavior in several different segments of the market.

There are several hierarchical levels in the reservation planning process. A summary of the planning process that we observed at one hotel, which is representative of what we have observed at many hotels, is given in Table 1 and Figure 1. The highest level, aggregate planning, involves both corporate and local hotel management. This aggregate planning team convenes about once per month to agree upon selective sales targets, which represent the anticipated fraction of room sales to groups and transients. The planning horizon for this high level planning extends as far into the future as necessary to accommodate requests for group reservations. At some hotels, which cater to large business conventions, this may be several years.

At the aggregate planning level, two selective sales targets are specified for each month: one for weekends (Friday–Saturday), and another for weekdays (Sunday–Thursday). These targets serve two purposes: They represent booking limits on discounted rooms, and they facilitate the development of a budget. The targets are based on the aggregate planning team's anticipation, i.e., forecast, of demand. It is the generation of demand forecasts that consumes most of the team's attention. They use historical data as a base case, and make adjustments according to the city's calendar of special events, conferences, etc. as well as any changes that may have occurred in the competitive environment.

Room merchandizing is an intermediate level of planning which involves only the local hotel management. It is reviewed about once per week with a planning horizon of 60–90 days. The purpose of this level of planning is to refine the monthly selective sales targets to booking limits for individual days. The local management has authority to adjust the booking limits on discounted reservations on specific days as long as the aggregate level of discounted sales for each month remains consistent with the selective sales targets from the aggregate planning level. The reservations manager at a particular hotel adjusts the booking limits on the basis of three types of information:

1. the profile of reservations already booked, including discounts, rack rate, guaranteed and nonguaranteed;
2. the amount of time remaining until the booking date;
3. the event calendar for the city.

If the demand for rack rate rooms relative to the time remaining is either higher or lower than was expected, the reservations manager does two things. First he/she contacts other hotels to see if this is a city-wide phenomena. Second, he/she looks at the city's calendar of events to ascertain the source(s) of the demand. This information is important because some types of travelers tend to book sooner than others. For example, conference attendees tend to book sooner than leisure travelers. Combining this information with experience and intuition, the reservations manager then decides whether to increase or decrease the booking limit on discounted reservations.

Subject classifications: Industries, hotel/motel: reservations planning. Inventory/production, perishable/aging items: booking limits for hotel reservations.
Area of review: MANUFACTURING, OPERATIONS AND SCHEDULING (SPECIAL ISSUE ON NEW DIRECTIONS IN OPERATIONS MANAGEMENT).

0030-364X/96/4401-0035 $01.25

Table I
Reservation Planning Levels

Planning Level	Frequency of Review	Length of Horizon	Management Involvement
Aggregate Planning	Monthly	$\geq$1 Year	Corporate and Local
Room Merchandizing	Weekly	2–3 Months	Local
Inventory Management	Daily	1 Day	Local

Daily inventory management is the lowest level of reservation planning. It involves the allocation of rooms to customers as they arrive at the hotel. Early in the day, the reservations manager reviews the profile of reservations. This profile includes the numbers of guaranteed and nonguaranteed reservations, as well as a list of early departures of guests with multiple day reservations and other late cancellations. Based on this information and the historical rate of no-shows for various types of reservations, he/she decides whether to begin admitting walk-in customers immediately. If walk-ins are to be admitted, priority is given to current guests who request to extend their stay.

On days when the hotel is likely to be filled, the reservations manager monitors the inventory position, i.e., the number of unclaimed rooms net of outstanding reservations, several times during the day. He/she attempts to avoid either winding up with unrented rooms or having to deny rooms to customers who arrive with reservations. By monitoring the rate at which customers are registering for rooms, conferring with competing hotels regarding their pace of registration, and considering the source(s) of demand, the reservations manager develops a "feel" for the likelihood of filling the hotel. If he/she feels that corrective action is required, he/she will either admit walk-ins or refer arriving customers to competing hotels. Note that customers are often referred to other hotels prior to the time when every room has been registered to a guest. The reason for this is that, as the day goes on, not only does it become increasingly difficult for the reservations manager to arrange with competing hotels to accommodate a referral, the customer who is referred experiences greater inconvenience.

If there are reservations that have not been guaranteed with a credit card, a significant decision point occurs at the time of their expiration. At this point, the only remaining outstanding reservations are the guaranteed ones, which tend to have a high probability of showing up. Since the expiration of the nonguaranteed

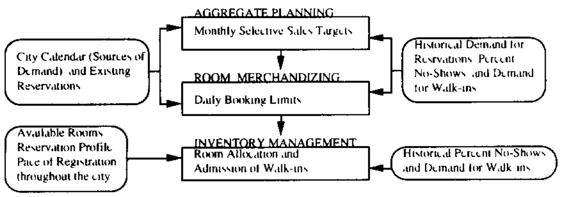

Figure 1. Decision hierarchy.

reservations eliminates much of the uncertainty associated with the remaining arrivals of reservation customers, it is often at this point that the reservations manager begins admitting walk-ins, if he/she has not done so already. Our paper formalizes the relationship between the room merchandizing and inventory management levels of planning.

A number of authors have studied versions of reservations yield management problems in airlines, hotels, rental vehicles, and a variety of other service industries. All of these problems share a common thread: Uncertain demand for products which cannot be inventoried. Yet each specific application of yield management has unique features.

The airline yield management problem has been particularly well studied. Rothstein (1985) and Belobaba (1987) provide detailed reviews of mathematical approaches which have been taken to the problem of maximizing the expected revenue associated with a given flight in which capacity is fixed. Much of this work studies the allocation of sales of a single product to customers paying different fares, and is relevant to the reservations acceptance phase of the hotel problem. Belobaba (1989) has taken a marginal pricing approach in what he calls the Expected Marginal Seat Revenue (EMSR) model. In this model, he assumes that lower fare classes purchase before high ones. He shows that for each fare class i, seats should be "protected" or withheld in such a way that the marginal revenue (with respect to the number of seats withheld) from sales to higher paying classes is exactly equal to the fare for class i. Although he extends this work to account for no shows, the emphasis is placed upon the allocation of a fixed number of tickets to different fare categories.

The EMSR approach has been recognized as being nonoptimal. However, Brumelle and McGill (1993) demonstrate that although it produces seat allocations that differ significantly from optimal ones, the associated loss in expected revenue is quite small.

Rothstein (1974) did some early work in bridging the gap between the hotel and the airline reservations management problems. He contrasts the two, and proposes a Markovian sequential decision model. His focus is upon how to adjust overbooking limits at various decision points that lead up to a target date. Requests for reservations, cancellations, and show rates are all sources of uncertainty.

Ladany (1976) proposes a dynamic decision model for a hotel with both single and double rooms. In each stage of his dynamic program, a random number of reservation requests and cancellations are received. The controls are the limits on the number of each type of reservation to accept. Ladany provides a concise dynamic programming formulation of the problem.

Alstrup et al. (1986) consider a similar model in an application to airlines with two types of seats. However, because of the computational effort required to solve the dynamic program for an airplane with 110 seats, they suggest that an approximation problem be solved instead. In the approximation model, passengers are treated as groups rather than individually. That is, the limits on ticket sales can only be set at multiples of the group size. They show that it is possible to obtain accurate results using a group size of five or fewer.

Liberman and Yechiali (1978) propose another dynamic decision model. Although they consider only one type of room, they assume that in addition to limiting the number of reservations to accept, management can either cancel previously confirmed reservations, or acquire additional bookings at some specified cost. They show that the optimal strategy is a 3-region policy. In each period, upper and lower limits on reservation inventory create three regions. The optimal action depends only upon where the current reservation inventory lies within these regions.

Williams (1977) considers a slightly different perspective than these other authors. He models a particular date that represents a peak in demand. He assumes that demand for rooms on this date comes from three sources, listed in decreasing order of priority: stay-overs—guests occupying rooms on the day preceding the critical date; reservations—guests arriving on the critical date with reservations; and walk-ins—guests arriving without reservations. He further assumes that the occupancy of the hotel on the day before the critical one is known with certainty, and calculates the expected costs of forgone revenues and overbooking that are associated with various numbers of reservations. Although Williams' suggests methods of designing decision aids, his model is concerned more with estimating the costs of specific policies than with optimization.

All of these models implicitly assume that, on the date of the rental, no rooms are allocated until after all of the walk-ins and customers with reservations have arrived. This is equivalent to assuming that all customers arrive at the same time. In practice, customers arrive throughout the day. If a room is not available when a given customer arrives, he will not wait for several hours to find out whether he will be given a room.

In contrast to the models of hotel reservations problems that were just described, we do not assume that all customers arrive simultaneously on the targeted booking date. Our approach integrates features of existing models. We begin with a model of the room allocation problem which resembles the airline seat/booking allocation models that have been studied by Belobaba, Brumelle and McGill, among others. However, the cost structure of our problem differs from those of the airline booking allocation models where the only penalty associated with denying a booking to a potential customer is the opportunity cost if it is not sold for a higher fare. In our problem, there is a substantial penalty to the hotel for denying a room to a customer with a reservation, regardless of whether the room is eventually filled.

This room allocation problem is embedded within the broader problem of reservations acceptance to reflect the fact that a hotel controls the distribution of arrivals on the targeted date through its booking policies. Although, in practice, an airline controls the distribution of demand for bookings through its pricing policies, we are unaware of a model which captures the relationship between pricing and booking allocation. Moreover, even if such a model were developed for an airline, its structure would be fundamentally different from ours because of the fact that, for hotels, the reservation acceptance and room allocation problems are solved sequentially rather than simultaneously.

For the sake of simplicity, we consider only one type of room. Although, in practice, hotels may have several types of room: singles, doubles, luxury suites, etc., it is not unusual for them to have most of their capacity concentrated in rooms which are indistinguishable to the customer. We also consider only independent, single night, single room reservations. There are two potential sources of dependency among reservations: groups that request multiple rooms, and individuals who request multiple nights. Because of the risk that would be involved with having a multiroom reservation fail to show up, hotels generally require that group reservations be made far in advance, and that they be prepaid. Thus, the presence of group reservations affects only the capacity of the hotel which is available to individual customers. Although multiple day reservations can pose problems in practice, Ladany suggests that for many hotels they represent only a small portion of total reservations. Rothstein (1974) assumes that the individual dates of a multiday reservation cancel or show-up independently. Bitran, Gilbert and Leong (1992) present empirical evidence which validates this assumption. Thus, a multiday reservation can be approximated by a series of independent, single night reservations.

In general, the probability distributions for reservation requests and walk-ins can best be modeled by Poisson distributions. Those for the numbers of reservations to "survive" from one period to the next are best modeled as binomials. However, these discrete distributions are difficult to work with analytically. Since hotels often have hundreds of rooms, it is reasonable to approximate the discrete distributions for reservation requests and cancellations with continuous distributions.

The remainder of the paper is organized as follows: We first present a formal model of the room allocation problem. In Section 2, we analyze the properties of the formulation. In Section 3 we use the insights from this analysis and our observations of what is done in practice to develop heuristic methods for solving the reservations acceptance problem. In Section 4, we describe an upper bound on the value of an optimal solution to a dynamic programming formulation of the problem. Finally, in Sections 5 and 6, we discuss the performance of our heuristics and areas for future research. Before presenting the formal model, let us introduce the following notation.

1. MODEL OF THE PROBLEM

Parameters

π_G, π_H: The per room revenue net of variable cost for renting a room to a guaranteed or 6 p.m. hold customer;

p_G, p_H: the per room cost of failing to honor a guaranteed or 6 p.m. hold reservation;

C: the capacity of the hotel in number of rooms;

G_1, H_1: the number of guaranteed and 6 p.m. hold reservations that are outstanding at the beginning of the day in which customers will be arriving.

Random Variables

r_W: The number of room requests from "walk-in" customers on the target date;

S_{G0}: the number of guests with guaranteed reservations to show up on the target date;

S_{H0}: the number of guests with 6 p.m. hold reservations to show up on the target date;

$\mathbf{S}_0$: a 2-dimensional vector whose components are S_{G0} and S_{H0};

C_{WG}: the number of rooms that remain available for walk-ins and guaranteed reservations after the arrival of the 6 p.m. hold guests;

C_G: the number of rooms that remain available for guaranteed reservations after the arrival of walk-ins and the 6 p.m. hold guests.

Decision Variables

N_{H0}: The limit on the total number of rooms that will be assigned to customers arriving with 6 p.m. hold reservations on the target date;

N_W: the limit on the total number of walk-in requests that will be accepted on the target date.

Other Notation

$f_{\mathbf{X}}(\mathbf{x})$: The probability density function for random vector $\mathbf{X}$ evaluated at $\mathbf{x}$;

$F_X(x)$: the cumulative distribution function for random variable X evaluated at x;

$F_X^c(x)$: the complementary cumulative distribution function for the random variable X evaluated at x;

$E_{\mathbf{X}}[g(\mathbf{X})]$: the expectation of a function $g(\mathbf{X})$ with respect to the vector $\mathbf{X}$;

$(x)^+$: Max $(x, 0)$.

Consider the following model of the room allocation problem.

$$MER_0(G_1, H_1, C) = \operatorname*{Max}_{\substack{N_{H0} \le C \\ N_{H0} \le H_1}} \{E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})]\} = \operatorname*{Max}_{\substack{N_{H0} \le C \\ N_{H0} \le H_1}} \left\{ \int_{S_{H0}=0}^{H_1} R_H(G_1, S_{H0}, C, N_{H0}) \cdot f_{S_{H0}}(S_{H0})\, dS_{H0} \right\}, \tag{1}$$

where

$$R_H(G_1, S_{H0}, C, N_{H0}) = \pi_H \cdot \text{Min}\,(N_{H0}, S_{H0}) - p_H(S_{H0} - N_{H0})^+ + MER_W(G_1, C - \text{Min}\,(N_{H0}, S_{H0})). \tag{2}$$

$$MER_W(G_1, C_{WG}) = \operatorname*{Max}_{N_W \le C_{WG}} \{E_{r_W}[R_W(G_1, r_W, C_{WG}, N_W)]\} = \operatorname*{Max}_{N_W \le C_{WG}} \left\{ \int_{r_W=0}^{\infty} \left(\pi_W W + \int_{S_{G0}=0}^{G_1} R_G(S_{G0}, C_{WG} - W) f_{S_{G0}}(S_{G0})\, dS_{G0} \right) \cdot f_{r_W}(r_W)\, dr_W \right\}, \tag{3}$$

where

$$W = \text{Min}\,(N_W, r_W), \quad \text{and} \tag{4}$$

$$R_G(S_{G0}, C_G) = \text{Max: } \pi_G \cdot \text{Min}\,(S_{G0}, C_G) - p_G(S_{G0} - C_G)^+. \tag{5}$$

As shown in Figure 2, this model has the following interpretation: At the beginning of a targeted day ($t = 0$), the hotel manager has given quantities for guaranteed and 6 p.m. reservations, G_1 and H_1, respectively, and some knowledge of the distribution of demand from walk-ins, $f_{r_W}(r_W)$. Note that the model does not explicitly include requests for same-day reservations. Since reservations made for the same day are virtually certain to show up, they are equivalent to walk-ins. Problem $MER_0(G_1, H_1, C)$ represents the maximum expected profit given that the hotel has a capacity of C rooms, and that G_1 and H_1

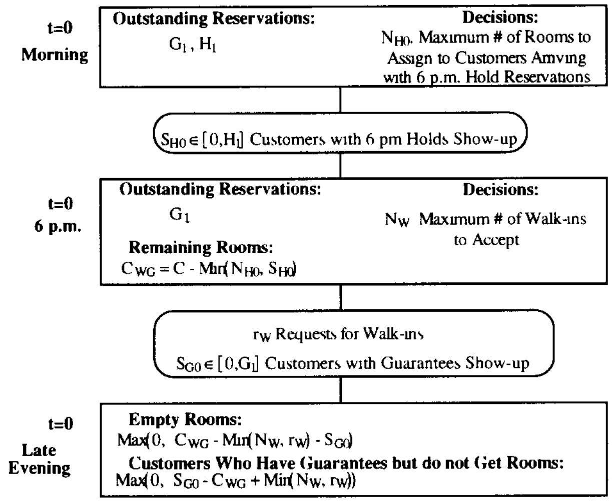

Figure 2. Room allocation.

reservations are held in the morning of the target date. The actual profits are determined by the show rates of these reservations, the number of walk-in customers, and the hotel manager's decisions about room assignments.

We have assumed that the three different types of guests arrive at nonoverlapping times of the day. Specifically, we assume that customers with 6 p.m. hold reservations arrive first, followed by walk-ins, and finally those with guaranteed reservations. In practice, the latest arriving customers tend to have guaranteed reservations. Because of the high cost of "walking" customers, when capacity is tight management tends to delay accepting walk-ins until late in the day. On other days, when the hotel is not likely to be filled, the solution to the problem is trivial: All arriving customers are assigned rooms. Thus, for the difficult cases, when capacity of the hotel is constrained there is little overlap between the arrival of the 6 p.m. hold customers and walk-ins. Although some of the customers with guaranteed reservations generally arrive along with the other guests, our model represents the worst case. Note that as the number of outstanding guaranteed reservations at 6 p.m. decreases, so does the risk of having to "walk" late arrivals. In the extreme case, if guaranteed reservation customers arrive prior to the walk-ins, we could decide how many walk-ins to admit with full information about the realization of the yield from reservations. Thus, the problem becomes easier rather than harder when we relax the assumption that guaranteed reservation customers arrive after the other two customer types.

The first decision to be made is N_{H0}, the limit on the number of rooms to assign to customers with 6 p.m. hold reservations. The actual number that is assigned depends on this limit as well as on the number (S_{H0}) of customers who arrive with valid 6 p.m. hold reservations. In many cases, it is less expensive to "walk" a customer at 6 p.m. than to do so later in the evening. Not only is it easier to find an alternative hotel at 6 p.m., the customer's perception of the inconvenience of being walked is likely to increase with the time of evening. Because customers with guaranteed reservations often arrive very late in the evening, it may be prudent to "walk" some customers who have 6 p.m. hold reservations to serve these late arrivals.

The next decision occurs after the 6 p.m. hold customers have either arrived or their reservations have expired. During the period between 6 p.m. and the end of the evening, the hotel manager may receive requests for rooms from walk-in customers. Because these customers do not have reservations, there is no penalty, other than the foregone profit margin for turning them away. $MER_W(G_1, C_{WG})$ represents the decision regarding N_W, the limit on the number of rooms to give to walk-in customers. It maximizes the expected profits from the arrival of walk-ins and guaranteed reservations given that:

1. C_{WG} rooms remain available after the arrival of the 6 p.m. hold guests;
2. an uncertain number, r_W, of walk-in customers will request rooms;
3. the minimum of r_W and N_W rooms will actually be given rooms;
4. $C_G = C_{WG} - \text{Min}\,(r_W, N_W)$ rooms will be available after the arrival of 6 p.m. hold guests and walk-ins;
5. G_1 guaranteed reservations are booked, and some uncertain number S_{G0} will actually show up.

Notice that N_W is constrained to be no larger than the number of available rooms. The expectation is taken with respect to r_W and S_{G0}, the actual number of walk-in requests and arrivals of guests with guaranteed reservations.

2. MODEL ANALYSIS

The primary objective of our analysis of problem $MER_0(G_1, H_1, C)$ is to develop insights which will help hotel managers to make better decisions on how to allocate rooms to customers as they arrive at the hotel. A secondary objective is to use our understanding of this room allocation problem to guide decisions about accepting guaranteed and 6 p.m. hold reservations.

Before discussing optimality conditions we need to address the concavity of the problem. We first consider the revenue associated with the arrival of customers with guaranteed reservations $R_G(S_{G0}, C_G)$. It is easy to show the following lemmas.

Lemma 1. *The function $R_G(S_{G0}, C_G)$, given in* (5), *is nondecreasing and concave in C_G.*

Lemma 2. *The gradient of $\mathbf{E}_{S_{G0}}[R_G(S_{G0}, C_G)]$ with respect to C_G can be expressed as:*

$$\frac{\partial}{\partial C_G} \mathbf{E}_{S_{G0}}[R_G(S_{G0}, C_G)] = (\pi_G + p_G) F^c_{S_{G0}}(C_G), \quad (6)$$

where $F^c_{S_{G0}}(C_G)$ is the probability that $S_{G0} \geqslant C_G$. These results are consistent with the intuition that the expected profit associated with the arrival of customers with guaranteed reservations increases as a function of C_G, the number of rooms that are available for them. However, as C_G increases, the marginal value of an additional room diminishes until it reaches zero when $C_G = G_1$. In other words, there is nothing to be gained from having more rooms than there are customers to claim them. We are now prepared to begin discussing the optimality conditions for each of the various stages of the room allocation problem.

Claim 1. *Given that after the 6 p.m. hold reservations expire, there remain C_{WG} unassigned rooms and G_1 outstanding guaranteed reservations, the following is an optimal limit on the number of walk-ins to accept:*

$$N^*_W(C_{WG}) = \begin{cases} 0 & \text{if } C_{WG} \leqslant F^{c\ -1}_{S_{G0}}\left(\dfrac{\pi_W}{\pi_G + p_G}\right) \\ C_{WG} - F^{c\ -1}_{S_{G0}}\left(\dfrac{\pi_W}{\pi_G + p_G}\right) & \textit{otherwise}, \end{cases} \tag{7}$$

where $F^{c\ -1}_{S_{G0}}(\alpha) = G$: Prob $(S_{G0} \geqslant G) = \alpha$.

The proof of this is given in Appendix A. The optimal solution to the walk-in allocation problem MER_W, which is similar to the result obtained by Belobaba (1989) for the booking of airline tickets, is intuitively appealing. Note that (7) sets N^*_W to the maximum of zero or the solution to:

$$N^*_W\colon \frac{\pi_W}{\pi_G + p_G} = F^c_{S_{G0}}(C_{WG} - N^*_W)$$
$$= (1 - F_{S_{G0}}(C_{WG} - N^*_W)).$$

The marginal benefit from giving an additional room to walk-ins is the revenue π_W. The expected marginal cost of not having that room available for later arriving guaranteed reservations is equal to the probability = Prob ($S_{G0} > C_{WG} - N_W$) that we will be unable to honor at least one reservation multiplied by the cost ($\pi_G + p_G$) of doing so. Thus, it is in our interest to accept walk-ins as long as the probability of failing to honor a guaranteed reservation is less than $\pi_W/(\pi_G + p_G)$. Note the similarity between the optimal solution to this problem and the classic newsboy problem. In that problem, it is optimal to order enough newspapers that the probability of a stockout equals the ratio of backorder costs to the sum of backorder and holding costs. Here, rooms should be allocated to walk-in requests until the probability of failing to provide a room to a guaranteed reservation exceeds the ratio: $\pi_w/(\pi_G + p_G)$. Recall that $C_G = C_{WG} - N_W$ is the number of rooms that are available for guaranteed reservations if N_W rooms are given to walk-ins.

In Figure 3, we can see that, for a given number of guaranteed reservations G_1, there is an optimal number of rooms, $N^*_W(C_{WG})$, which should be saved for these guests. Walk-ins should be accepted until the number of remaining rooms equals this critical value. If the number of rooms remaining after the arrival of the 6 p.m. hold guests is less than this critical value, then no walk-ins should be accepted. Note that the point at which $N^*_W(C_{WG})$ equals zero depends upon the cumulative distribution function for arrivals of guaranteed reservations, which is a function of G_1, the number of outstanding guaranteed reservations. Thus, for a larger (smaller) value of G_1, the line representing $N^*_W(C_{WG})$ in Figure 3 will shift to the right (left).

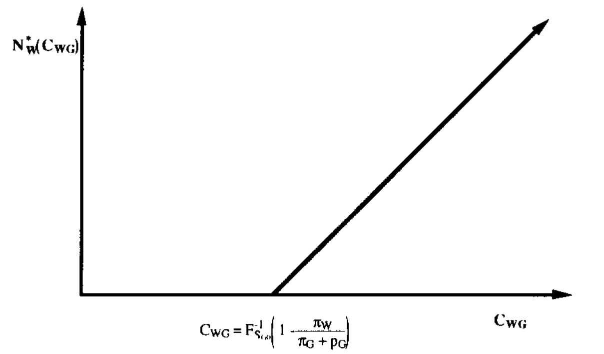

Figure 3. Allocation to walk-ins after 6 p.m. holds expire.

We can now investigate the optimality conditions for the problem of deciding the number of rooms to allocate to customers who arrive with 6 p.m. hold reservations, i.e., problem $MER_0(G_1, H_1, C)$. Claim 2 characterizes the optimal solution under the assumption that $\pi_W < \pi_H + p_H$, which is commonly accepted in the industry.

Claim 2. *If $\pi_W < \pi_H + p_H$, then $E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})]$, the expected revenue associated with the arrival of customers, with and without reservations, is maximized with respect to N_{H0} at the point:*

$$N^*_{H0} = \begin{cases} 0 & \text{if } \pi_H + p_H - (\pi_G + p_G)F^c_{S_{G0}}(C) < 0 \\ C - F^{c\ -1}_{S_{G0}}\left(\dfrac{\pi_H + p_H}{\pi_G + p_G}\right) & \textit{otherwise}. \end{cases}$$

The proof of this result is given in Appendix A. The intuition is the following: Suppose that the capacity of the hotel is large enough to accommodate all of the customers with guaranteed reservations. As we begin to allocate rooms to 6 p.m. hold customers, we tradeoff the increased revenue and savings in overbooking penalties associated with them against the potential lost revenue from turning away walk-ins. However, as we continue to increase N_{H0}, we decrease the amount of capacity that is available to walk-ins and guarantees to the point at which no walk-ins will be accepted. Note that since walk-ins pay less than the sum of the lost revenue and penalty associated with denying a 6 p.m. hold customer a room, it is never advantageous to turn away 6 p.m. hold customers to accommodate walk-ins. As we continue to increase N_{H0} beyond the point at which no walk-ins will be accepted, the increased revenue and savings in "overbooking" penalties from allocating an additional room to a 6 p.m. hold is balanced against the potential lost revenue and more severe penalties associated with being unable to honor a guaranteed reservation later in the evening. This result is consistent with our observation that, in situations when his/her hotel is overbooked, managers often walk arriving reservation customers long before their supply of rooms is exhausted.

It is interesting to compare this room allocation problem and the seat/booking allocation problem that airlines face. In both cases, a manager must decide when to deny a room (booking) to a current customer to avoid a future denial which would be more expensive. However, the only control that an airline has over the distribution of demand for bookings is through the fares that it offers. In contrast, a hotel directly controls the distribution of the number of customers who will arrive with reservations through its reservations acceptance policies. In fact, the parameters of the room allocation problem are determined by the reservations acceptance problem. In the following section, we propose a model of the reservations acceptance problem which makes this relationship explicit.

3. THE RESERVATIONS ACCEPTANCE PROBLEM

As discussed before, the parameters G_1 and H_1 of the room allocation problem depend upon the hotel's reservations acceptance policies. To model this dependency, let t represent an index of periods prior to some target date ($t = 0$) for which reservations are being taken. Let G_t and H_t represent the number of guaranteed and 6 p.m. hold reservations booked at the end of period t, and let r_{Gt} and r_{Ht} represent the numbers of requests for the two reservation types during period t. At the beginning of period t, hotel management observes the current uncancelled bookings (G_{t+1} and H_{t+1}) and imposes limits (N_{Gt} and N_{Ht}) on the number of reservation requests to accept in period t. Let S_{Gt} and S_{Ht} represent the numbers of reservations that "survive" period t, i.e., the number that are held at the beginning of period t and do not cancel by the end of t.

Using this notation, the reservations acceptance problem can be formulated as follows.

$$\begin{aligned} &MER_t(G_{t+1}, H_{t+1}, C) \\ &= \operatorname*{Max}_{N_{Gt}, N_{Ht}} \{E_{\mathbf{r}_t \mathbf{S}_t}[R_t(G_{t+1}, H_{t+1}, C, N_{Gt}, N_{Ht})]\} \\ &= \operatorname*{Max}_{N_{Gt}, N_{Ht}} \Bigg\{ \int_{r_{Gt}}^{\infty} \int_{r_{Ht}}^{\infty} \int_{S_{Gt}=0}^{G_{t+1}} \int_{S_{Ht}=0}^{H_{t+1}} MER_{t-1} \\ &\qquad \cdot (G_t, H_t, C) f_{\mathbf{r}_t \mathbf{S}_t}(\mathbf{r}_t, \mathbf{S}_t) \Bigg\}, \qquad (8) \\ &\qquad\qquad \text{for } t = T, \dots, 1, \end{aligned}$$

where,

$$H_t = \operatorname{Min}\{r_{H,t}, N_{H,t}\} + S_{H,t} \quad \text{for } t = 1, \dots, T; \quad (9)$$

$$G_t = \operatorname{Min}\{r_{G,t}, N_{G,t}\} + S_{G,t} \quad \text{for } t = 1, \dots, T; \quad (10)$$

and $MER_0(G_1, H_1, C)$ is as defined in (1)–(5).

Analysis of this problem is not easy. If we use discrete distributions to represent the reservations that survive from one period to the next, then for small problems we can identify optimal solutions through enumeration. Unfortunately, this approach is not practical for realistically sized problems. Even if we approximate the discrete distributions with continuous ones, we cannot compute gradients because of the recursive structure of the optimization problem.

Although these difficulties make it hard to identify an optimal solution to the reservations acceptance problem, we have developed heuristic methods which perform well *vis-à-vis* an upper bound on the value of an optimal solution. In Section 4, we present an upper bound on the value of an optimal solution to the dynamic program against which the heuristics can be compared.

In practice, managers tend to rely on "rules of thumb" and "gut feel" to determine when to begin refusing reservations requests. The heuristics which we describe were motivated by the insight which we obtained from analyzing the room allocation problem as well as from discussions that we held with reservations managers from The OMNI Parker House and a Marriott Hotel in Boston. During these discussions, it was our objective to understand the general approach and the rules of thumb that are used to solve the reservations acceptance problem in practice.

In each of the three heuristics, a target number of walk-ins is selected. This number represents the number of rooms that are targeted for walk-ins. It is based upon the size of the premium that a walk-in pays as well as the distribution function for walk-in requests. The target is:

$$TargW = F_{r_W}^{c^{-1}}\left(\frac{\pi_R}{\pi_W}\right), \qquad (11)$$

where π_R is the average price paid by a customer with a reservation, guaranteed, or otherwise. This target is based on the fact that when $TargW$ rooms are allocated to walk-ins, the marginal benefit of having an additional room to allocate is equal to the walk-in price multiplied by the probability that demand for walk-ins would be sufficient to fill all of the $TargW$ rooms. As we make more rooms available to walk-ins (increasing $TargW$), the probability of filling them all ($F_{r_W}^c(TargW)$) decreases. Since making an additional room available to walk-ins means foregoing the opportunity to fill it with a reservation customer, the opportunity cost of doing so is equal to π_R, the average revenue associated with reservation customers. The marginal cost and benefit are balanced when the probability of having to turn away at least one walk-in is equal to the ratio of the two revenues.

The first heuristic for setting limits on the numbers of guaranteed and 6 p.m. hold reservations to accept in period t is simple. Let us define $ETrG_t$ to be the expected total number of requests for guaranteed reservations in periods $t, \dots, 1$. Also, let $q_{Gt}(q_{Ht})$ be the probability that a customer with a guaranteed (6 p.m. hold) reservation does not cancel during period t. Recall from the model of the problem that is defined in (8)–(10) that there are G_{t+1} and H_{t+1} reservations booked at the beginning of period t, and S_{Gt} and S_{Ht} of those reservations survive, i.e., do not cancel, during period t. The decisions N_{Gt} and N_{Ht} are the maximum numbers of guaranteed and hold reservations to accept during period t.

$\mathbf{HEUR1}_t(G_{t+1}, H_{t+1})$

STEP 1. Calculate $TargW$ using (11).

STEP 2. Set $TargR = C - TargW$.

STEP 3. Calculate $ETrG_t = E[r_{Gt} + \dots + r_{G1}]$.

STEP 4a. If $G_{t+1}q_{G0} + H_{t+1}q_{H0} \geqslant TargR$, Then $N_{Gt} = N_{Ht} = 0$.

STEP 4b. Else:

$$N_{Gt} = \frac{TargR - G_{t+1} \cdot q_{G0} - H_{t+1} \cdot q_{H0}}{q_{G0}},$$

$$N_{Ht} = \mathrm{Max}\left\{0, \frac{TargR - (G_{t+1} + ETrG_t) \cdot q_{G0} - H_{t+1} \cdot q_{H0}}{q_{H0}}\right\}$$

Steps 1 and 2 must be calculated only once. Steps 3 and 4 must be calculated in every period. The condition in Step 4a is that, ignoring cancellations prior to the target date, the expected number of rooms that will be filled with reservation guests is at least as high as the target. Thus, no reservations will be accepted in the next period. If the condition in Step 4a is not met, then guaranteed reservations are accepted until the point where the expected number of rooms that will be filled with reservation guests is equal to the target. Then 6 p.m. hold reservations are accepted only to the extent that the expected future requests ($ETrG_t$) for guaranteed reservations is insufficient to meet the target. Notice that, with the exception of the targeted number of rooms for walk-ins, all calculations are made using expected values of the random variables. Also note that although **HEUR1** considers no-show rates, it ignores cancellations that occur prior to the target date. Although this is consistent with the rules of thumb that we observed in practice, we propose a modified heuristic. **HEUR2**, which explicitly considers the effects of these early cancellations, attrition, upon reservation limits. Let us define $EYrG_t$ to be the expected total "yield" from requests for guaranteed reservations in periods $t, \dots, 1$, where yield refers to the number of requests net of cancellations and no-shows. The rule is as follows.

$\mathbf{HEUR2}_t(G_{t+1}, H_{t+1})$

STEPS 1 and 2. These steps are as in heuristic $\mathbf{HEUR1}_t(G_{t+1}, H_{t+1})$.

STEP 3. Calculate

$$EYrG_t = E\left[\sum_{\tau=1}^{t} \left(r_{G\tau} \prod_{i=0}^{\tau-1} q_{Gi}\right)\right].$$

STEP 4a. If $G_{t+1}\Pi_{\tau=0}^{t} q_{G\tau} + H_{t+1}\Pi_{\tau=0}^{t} q_{H_\tau} \leqslant TargR$, then $N_{Gt} = 0$, and $N_{Ht} = 0$.

STEP 4b. Else:

$$N_{Gt} = \frac{TargR - G_{t+1} \cdot \prod_{\tau=0}^{t} q_{G\tau} - H_{t+1} \cdot \prod_{\tau=0}^{t} q_{H\tau}}{\prod_{\tau=0}^{t-1} q_{G_\tau}},$$

and

$$N_{Ht} = \mathrm{Max}\left\{0, \frac{\mathrm{TargR} - G_{t+1} \cdot \prod_{\tau=0}^{t} q_{G\tau} - EYrG_t - H_{t+1} \cdot \prod_{\tau=0}^{t} q_{H\tau}}{\prod_{\tau=0}^{t-1} q_{H\tau}}\right\}.$$

HEUR2 is identical to **HEUR1** except for the fact that it explicitly inflates the reservations limits to allow for cancellations prior to the target date. The third heuristic which we consider differs from **HEUR2** only in the way in which $TargR$, the number of rooms that are targeted to be filled with reservation customers, is calculated.

$\mathbf{HEUR3}_t(G_{t+1}, H_{t+1})$

STEPS 1 and 2. These steps are as in $\mathbf{HEUR1}_t(G_{t+1}, H_{T+1})$.

STEP 2a. Set ϵ such that: Prob $[S_{G0} \leqslant C - (TargW - \epsilon)] = p_G/\pi_G + p_G$, where S_{G0} is the sum of $TargR$ independent Bernoulli trials, each with parameter q_{G0}. Set $TargR = TargR - \epsilon$.

STEPS 3 and 4 as in **HEUR1**.

The intuition behind Step 2 of **HEUR3** is similar to that of the standard newsboy problem. The idea is that when $TargR$ reservations are booked, the probability that no more than $C - TargW$ of these bookings show up should equal $p_G/(\pi_G + p_G)$. However, because $TargR$ is a parameter in the binomial distribution for S_{G0}, the number that show up, adjusting $TargR$ directly would be a lengthy iterative procedure. In Step 2a of **HEUR3**, ϵ represents the number of additional rooms that would be required on the target date to provide a service level of $p_G/(\pi_G + p_G)$ if $(C - TargW)/q_{G0}$ guaranteed reservations were booked and $C - TargW$ rooms were available. Thus, after reducing by ϵ the number of rooms targeted for reservations, the probability of being unable to honor a reservation reflects the tradeoff between the cost of doing so and that of an empty room.

4. AN UPPER BOUND FOR THE DYNAMIC PROGRAM

To evaluate the performance of the heuristics, it is necessary to have an upper bound on the value of an optimal solution to the dynamic program $MER_T(G_{T+1}, H_{T+1}, C)$. Let us denote the following vectors with boldface notation, e.g., $\mathbf{r}_t = (r_{Gt}, r_{Ht})$ for $t = 1, \ldots, T$, and $\mathbf{S}_t = (S_{Gt}, S_{Ht})$ for $t = 0, \ldots, T$.

Claim 3. *Define $S_{Ht}(H)$ as a binomial random variable with parameters q_{Ht} and H, and define $S_{Gt}(G)$ as a binomial random variable with parameters q_{Gt} and G. Let $\hat{H}_t = r_{Ht} + S_{Ht}(\hat{H}_{t+1})$ and $\hat{G}_t = r_{Gt} + S_{Gt}(\hat{G}_{t+1})$ for $t = 1, \ldots, T$, where $\hat{H}_{T+1} = H_{T+1}$ and $\hat{G}_{T+1} = G_{T+1}$. An optimal value for $MER_T(G_{T+1}, H_{T+1}, C)$ is less than or equal to $UMER_T(G_{T+1}, H_{T+1}, C)$, the value of the following problem.*

$$E_{\mathbf{r}_T,\ldots,\mathbf{r}_1,r_W,\mathbf{S}_T,\ldots,\mathbf{S}_0}[\text{Max } \pi_H N_{H0} + \pi_W N_W + \pi_G N_{G0} \tag{12a}$$

subject to

$$N_{H0} \le S_{H0}(\hat{H}_1) \tag{12b}$$

$$N_W \le r_W \tag{12c}$$

$$N_{G0} \le S_{G0}(\hat{G}_1) \tag{12d}$$

$$N_{G0} + N_{H0} + N_W \le C \tag{12e}$$

$$N_{H0}, N_{G0}, N_W \le 0]. \tag{12f}$$

Note that the expectation is with respect to r_W and the vectors $\mathbf{r}_t$ for $t = T, \ldots, 1$ and $\mathbf{S}_t$ for $t = T, \ldots, 0$.

Rather than providing the mathematical details of the proof, we will outline it and discuss the underlying intuition. Suppose that a hotel manager could observe the entire process of requests for reservations, cancellations, walk-ins, and no-shows before making any decisions. On the target date, he would simply allocate the capacity of his hotel to walk-ins, uncancelled guaranteed and 6 p.m. hold reservations in order of their profitability. Clearly, he could do at least as well as he can in real life where he must commit himself prior to observing these random events.

The derivation of the bound which has just been described is obtained by successively interchanging the maximization and expectation functions in the dynamic program, and applying Jensen's inequality. That is, for a given function $f(\ )$, random variable x and real valued N:

$$\underset{N\ge 0}{\text{Max}}\ \{E_x[f(N, x)]\} \le E_x\Big[\underset{N\ge 0}{\text{Max}}\ \{f(N, x)\}\Big].$$

The right-hand sides of (12b) and (12d) are upper bounds on the numbers of uncancelled 6 p.m. hold and guaranteed reservations, respectively. That is, regardless of the values of the decision variables representing reservation limits, for any given realization of the random numbers of requests for walk-ins and reservations, cancellations, and no-shows, we have that $S_{G0}(\hat{G}_1) \ge S_{G0}(G_1)$. This follows from the fact that the terms on both sides of this inequality represent the sum of a set of i.i.d. Bernoulli variates, and the variates in the summation on the right are a subset of those on the left. Similarly, we have that $S_{H0}(\hat{H}_1) \ge S_{H0}(\hat{H}_1)$. Thus, the constraints on the allocation of rooms to hold and guaranteed reservations are less constrained than in the original dynamic program (equations 1–5). Constraint 12c represents the fact that the number of rooms that can be assigned to walk-ins is limited to the number of requests. Constraint 12e reflects the capacity of the hotel.

The bound in 12 provides a benchmark against which we can evaluate the performance of heuristic methods of solving the real problem where decisions must be made before the random events (reservations requests, cancellations, and no-shows) have been observed. In the following section we describe the parameters of a Monte Carlo simulation of the heuristics, and discuss their performance.

5. COMPUTATIONAL RESULTS

To evaluate the performance of the reservation acceptance heuristics, we used Monte Carlo techniques to simulate the performance of each one in an environment in which reservation demand, cancellations, no-shows, and the number of walk-ins were random. We assumed that, on the target date, the optimal room allocation policies were followed. By measuring the hotel profits over a series of repeated simulations, we were able to obtain statistical estimates of the expected costs of using each heuristic.

We also used Monte Carlo simulation to determine a statistical estimate of the upper bound. Recall that the upper bound is the expected optimal value of a linear program in which the coefficients of the right-hand side are random. We "estimated" this expected optimal value by repeatedly generating realizations of the random coefficients and solving the resulting linear programs. This "estimate" of the upper bound on the expected revenue of an optimal solution to the original dynamic program provides a benchmark against which we can evaluate the performance of the heuristics.

In the Monte Carlo simulations, we assumed that in each decision period prior to the target date, the number of requests for 6 p.m. holds and guaranteed reservations is drawn from Poisson distributions. On the target date, the number of requests for walk-ins was also assumed to be a Poisson distributed random variable. We assumed that both no-shows and cancellations of reservations are drawn from binomial distributions. Because the customer base was composed primarily of individuals, rather than groups, these distributions are intuitively appealing. Their use is also common in the literature. For example, Alstrup et al. claim that the sales of airline tickets follow Poisson distributions while cancellations and no-shows follow binomials. Rothstein (1974) defends the validity of these distributions for modeling hotel reservations processes, and uses them to test his results.

We developed our tests on the basis of discussions with the management of a Marriott Hotel near downtown

Boston. The parameters that we used approximate those at this hotel. The target dates which present the management of this hotel with the most difficulty are the ones for which there is considerable uncertainty regarding demand. To protect themselves against having empty rooms, a large number of discounts are made available, and the average price paid by a customer with either a 6 p.m. hold or a guaranteed reservation is about \$100 versus \$150 for a walk-in. The cost of failing to honor a reservation before 6 p.m. is \$100 versus \$250 later in the evening. Although a large component of these costs represents intangibles, the substantial difference between them can be explained by the fact that prior to 6 p.m., it is often possible to relocate a customer to another downtown Marriott. By attempting to relocate customers prior to the point when there are no more rooms available, the hotel affords itself the luxury of offering several customers the option of relocating. It is not uncommon to be able to find someone who is quite receptive to spending a free night at another hotel. Alternatively, when the relocation occurs later in the evening, and customers are given no option, it is usually perceived as a far greater inconvenience.

For the purposes of the simulations, we assumed that there are four decision points prior to the target date, and that the length of the intervals between decisions is chosen such that the expected number of reservation requests is the same in each of the four periods. The parameters of the probability distributions for our base case scenario are given in Table II. The first column in the table is the number of periods that remain before the target date. The second, third, and fourth columns contain the expected number of requests for guaranteed reservations, 6 p.m. holds, and walk-ins in each period. Note that no requests for reservations are made on the target date, and that requests for walk-ins occur only then. Columns five and six contain the probability, $q_{Gt}(q_{Ht})$ that a given guaranteed (6 p.m. hold) reservation that is held at the beginning of period t will not cancel before the end of period t. Recall that on the target date, period 0, a cancellation is equivalent to a no-show.

The other scenarios that we tested were variations on this base case. The parameters of these scenarios are given in Appendix B. In scenario B, the expected number of requests for guaranteed reservations in each period leading up to the target date was 100 instead of 70.

Table II
Scenario A

t	$E[r_G]$	$E[r_H]$	$E[r_W]$	q_G	q_H
4	70	40	—	0.9	0.9
3	70	40	—	0.9	0.9
2	70	40	—	0.9	0.9
1	70	40	—	0.9	0.9
0	—	—	30	0.9	0.5

In scenario C, this parameter was 40. In scenario D, we returned the expected demand for guaranteed reservations to the base case level, but increased the expected number of requests for walk-ins to 60.

For each of these scenarios, we ran 500 iterations of the heuristics. By taking sample averages over the 500 iterations, we obtained statistical estimates of the expected costs of using the heuristics in each of the four different conditions. These estimates are presented in Table III as percentages of the upper bound. The standard deviations associated with these estimates is roughly 0.1%.

Recall that the calculation of the upper bound is itself a statistical estimate of the expected value of a linear program in which the right-hand-side coefficients are random variables. This estimate is also based upon 500 generations of the random variables, and has a standard deviation of roughly 0.1%. It can be seen in Table III that **HEUR2** and **HEUR3** perform much better than **HEUR1** for scenarios A, B, and D. Since **HEUR1** does not consider the effects of reservation attrition, i.e., cancellations prior to the target date, it tends to accept fewer reservations in the planning periods furthest from the target date than do the other heuristics. As should be expected, all three heuristics perform very well for scenario C. In this scenario, the expected demand for rooms, net of cancellations, and no-shows, is very low, and nearly everyone who wants a room receives one. Thus, hindsight is of little benefit, and the performance of the heuristics is very close to that of the upper bound. Using expected revenue as a measure of performance, **HEUR3** offers only marginally better performance than **HEUR2**. In an attempt to further differentiate these heuristics, we considered another measure of performance: The frequency with which at least five reservations cannot be honored. Although the hotel managers with whom we spoke indicated that their primary concern is maximizing revenue, they are also concerned about being unable to honor large numbers of reservations because of

Table III
Revenue of Heuristics as a Percent of the Upper Bound

Scenario	**HEUR1** (%)	**HEUR2** (%)	**HEUR3** (%)	Upper Bound (\$)
A	94.5	97.3	97.3	31,262
B	93.3	97.1	97.3	31,496
C	99.9	99.9	99.9	23,646
D	94.7	96.8	96.8	32,978

Table IV
Occasions (out of 500) in Which at Least Five Reservations Could not be Honored

Scenario	HEUR1	HEUR2	HEUR3
A	8	44	42
B	1	61	44
C	0	0	0
D	13	49	43

overbooking. If we consider the likelihood that a large number of reservations cannot be honored, then **HEUR3** looks more attractive. Table IV contains the number of target dates out of 500 in which at least five customer reservations could not be honored. Although by this measure, **HEUR1** dominates the other heuristics, the expected revenue from using it indicates that it might be too conservative. However, **HEUR3** has expected profits that are comparable to those of **HEUR2**, and simultaneously results in fewer occasions in which large number of reservations cannot be honored.

To test the robustness of the heuristics with respect to the probability that a customer with a guaranteed reservation will fail to show-up on the target date ($1 - q_{G0}$), we performed a series of tests in which this probability varied between 0.05 and 0.3. Because we wanted to focus our attention on the effect of increased variability rather than decreased total demand for rooms, we also adjusted the parameters of the requests for reservations. In particular, we adjusted the parameters for reservations requests so that the product of the expected number of requests and the probability that in each test:

$$\sum_{t=1}^{T} \left(E[r_{Gt}] \cdot \prod_{t=0}^{t-1} q_{Gt}\right) = K_G,$$

where K_G is a constant ($\approx$216.657). For example, for $q_{G0} = 0.75$ instead of 0.90, the expected number of requests for guaranteed reservations in each period prior to the target date was 84 instead of 70. The results of these simulations are presented in Tables V and VI. It is interesting to observe in Table V that the heuristics' performance relative to the upper bound does not seem to be adversely affected when the probability that a guaranteed reservation will show-up decreases. However, in Table VI, it can be seen that, for **HEUR2** and **HEUR3** the frequency with which at least five customers must be relocated increases with the no-show probability ($1 - q_{G0}$) for a guaranteed customer. However, this frequency does not increase as quickly for **HEUR3** as for **HEUR2**. For a guaranteed reservation no-show probability of $1 - q_{G0} = 0.05$, **HEUR3** results in 8% fewer occasions with at least five customer relocations than **HEUR2**. For $1 - q_{G0} = 0.3$, **HEUR3** results in 20% fewer occasions. Note that, regardless of the no-show probability for guaranteed reservations, **HEUR1** results in very few occasions where large numbers of customer reservations cannot be honored. However, it achieves this at the expense of expected revenues by turning away large numbers of reservations.

Table VI
Occasions (out of 500) in Which at Least Five Reservations Could not be Honored

q_{G0}	HEUR1	HEUR2	HEUR3
0.95	2	26	24
0.9 (BaseCase)	8	44	42
0.85	2	60	56
0.8	6	67	56
0.75	5	88	73
0.7	3	79	63

Although **HEUR1**, **HEUR2**, and **HEUR3** all perform well with respect to the upper bound under a variety of conditions, **HEUR2** and **HEUR3** perform consistently better than **HEUR1**. This advantage results from the fact that **HEUR2** and **HEUR3** consider the effects of cancellations prior to the target date, while **HEUR1** considers only no-shows in determining the number of reservations to accept. **HEUR3** is slightly more sophisticated than **HEUR2** in that it adds a safety factor into the number of rooms targeted to be filled with reservations customers. This safety factor is based on the tradeoff between the lost revenue of an empty room and the penalty for overbooking. Although the performance of **HEUR3** was only slightly better than that of **HEUR2** in terms of expected revenue, it resulted in substantially fewer occasions in which a large number of customers had to be relocated because of overbooking. By this latter measure, in which the hotel managers expressed a particular interest, the advantage of **HEUR3** over **HEUR2** increased as the show-up probability for the guaranteed reservations decreased.

Table V
Revenue of Heuristics as a Percent of the Upper Bound

$1 - q_{G0}$	HEUR1 (%)	HEUR2 (%)	HEUR3 (%)	Upper Bound ($)
0.05	94.5	97.8	97.8	31,250
0.1 (BaseCase)	94.5	97.3	97.3	31,262
0.15	94.5	97.0	97.1	31,289
0.2	94.7	96.8	96.9	31,237
0.25	94.9	96.5	96.7	31,201
0.3	94.7	96.5	96.7	30,246

6. DISCUSSION

The problems that arise in the hotel industry as a result of the uncertainties associated with room reservations are interesting and complex. Based on interaction with the managements of two hotels near downtown Boston, the OMNI Parker House and the Marriott, we have identified and studied a dimension of the problem that has not, to our knowledge, been studied previously. This problem arises from the fact that rooms must be allocated to customers as they arrive throughout the target date. That is, management must begin allocating rooms before observing the number of no-shows. This contrasts the situation in the airline industry where all of the no-shows are observed prior to the boarding of the flight.

We formulate the room allocation problem as a stochastic dynamic program and derive optimal solution policies which are consistent with the intuitive approaches that are currently used. Based upon both the insights gained from this model and our observations of current practices that are used at two large urban hotels, we propose simple heuristic methods for making reservations acceptance decisions. Finally, we evaluate the performance of the heuristics against an upper bound using Monte Carlo simulation. Although all three of our heuristics for reservations acceptances perform well relative to the upper bound, the superior performance of **HEUR2** and **HEUR3** demonstrate the importance of considering not only no-shows, but also cancellations that occur prior to the target date. In addition to performing well, our heuristics are relatively simple, increasing the likelihood that they will be implemented.

The most obvious direction for future research is to extend these results to cases in which either there are multiple room types, or the arrivals of different types of customers can not be separated temporally. Another direction is to integrate pricing decisions with the reservations acceptance problem. Solutions to either of these more complex models depend ultimately upon the room allocation problem that arises on the target date. As such, we believe that our results represent an encouraging first step toward improved hotel reservations systems.

APPENDIX A

Proof of Claim 1

It is easy to show that $E_{rW}[R_W(G_1, r_W, C_{WG}, N_W)]$ is concave. The proof follows from the fact that for every realization of the random variable r_W, $R_W(G_1, r_W, C_{WG}, N_W)$ is concave, and a convex combination of concave functions is also concave. Thus, the following is a sufficient condition for N_W^* to be a global maximum (Mangasarian, p. 145):

$$(N_W^* + \Delta)\cdot\frac{\partial}{\partial N_W}(E_{r_W}[R_W(G_1, r_W, C_{WG}, N_W^*)]) \leqslant 0, \tag{A.1}$$

for any feasible direction Δ. The gradient of $E_{rW}[R_W(G_1, r_W, C_{WG}, N_W)]$ can be expressed as:

$$\begin{aligned}\frac{\partial}{\partial N_W}&(E_{r_W}[R_W(G_1, r_W, C_{WG}, N_W)])\\ &=\int_{r_W=N_W}^{\infty}(\pi_W-(\pi_G+p_G)\\ &\qquad\cdot(F^c_{S_{G0}}(C_{WG}-N_W)))f_r(r_W)\,dr\\ &=[\pi_W-(\pi_G+p_G)(F^c_{S_{G0}}(C_{WG}\\ &\qquad -N_W))]\cdot F^c_{r_W}(N_W).\end{aligned}\tag{A.2}$$

We can now substitute expressions 7 and A.2 into the optimality condition in (A.1). There are two cases.

Case 1. $C_{WG} < F^{c\ -1}_{S_{G0}}(\pi_W/\pi_G + p_G)$: From (7), $N_W^*(C_{WG}) = 0$, and we have:

$$\begin{aligned}&(N_W^*(C_{WG})+\Delta)\\ &\quad\cdot\frac{\partial}{\partial N_W}\left[\int_{r_W=0}^{\infty}R_W(G_1, r_W, C_{WG}, N_W)f_r(r_W)\,dr\right]\\ &=[0+\Delta][\pi_W-(\pi_G+p_G)(F^c_{S_{G0}}(C_{WG}))][F^c_{r_W}(0)]\\ &<0\quad\text{for all }\Delta>0,\end{aligned}\tag{A.3}$$

where, because the complementary cummulative distribution is a decreasing function:

$$\frac{\pi_W}{\pi_G+p_G}<F^c_{S_{G0}}(C_{WG})\leftrightarrow F^{c\ -1}_{S_{G0}}\left(\frac{\pi_W}{\pi_G+p_G}\right)>C_{WG}.$$

Note that when $N_W^*(C_{WG}) = 0$, a feasible direction Δ must be > 0.

Case 2. $C_{WG} \geqslant F^{c\ -1}_{S_{G0}}(\pi_W/\pi_G + p_G)$: From (7), $N_W^*(C_{WG}) = C_{WG} - F^{c\ -1}_{S_{G0}}(\pi_W/\pi_G + p_G)$, and we have:

$$\begin{aligned}&(N_W^*(C_{WG})+\Delta)\\ &\quad\cdot\frac{\partial}{\partial N_W}\left[\int_{r_W=0}^{\infty}R_W(G_1, r_W, C_{WG}, N_W)f_r(r_W)\,dr\right]\\ &=[N_W^*(C_{WG})+\Delta]\left[\frac{\pi_W}{\pi_G+p_G}\right.\\ &\qquad\left.-(F^c_{S_{G0}}(C_{WG}-N_W^*(C_{WG})))\right]\\ &\qquad\cdot[F^c_{r_W}(N_W^*(C_{WG}))]=0\quad\text{for all }\Delta.\end{aligned}\tag{A.4}$$

Proof of Claim 2

It can be shown that $E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})]$ is a pseudoconcave function of N_{H0}. (See Lemma 1.) Thus, the following is a sufficient condition for $E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})]$ to attain a global maximum at $N_{H0} = N_{H0}^*$:

$$\begin{aligned}(N_{H0}^*+\Delta)\cdot\frac{\partial}{\partial N_{H0}}&(E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0}^*)])\\ &\leqslant 0\quad\text{for any feasible direction }\Delta.\end{aligned}\tag{A.5}$$

We consider two cases.

Case 1

$$C \leqslant F^{c}_{S_{G0}}{}^{-1}\left(\frac{\pi_W}{\pi_G + p_G}\right), \quad \text{or equivalently,}$$

$$F^{c}_{S_{G0}}(C) \geqslant \frac{\pi_W}{\pi_G + p_G}.$$

This condition implies that, regardless of how many rooms are assigned to 6 p.m. hold customers, the optimal walk-in policy (Claim 1) is to accept no walk-ins. It also implies that for all $N_{H0} \geqslant 0$, the gradient of $E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})]$ with respect to N_{H0} can be expressed as the following (the derivation of the gradient is shown in Lemma 1):

$$\frac{\partial}{\partial N_{H0}}(E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})]) = (\pi_H + p_H - (\pi_G + p_G)F^{c}_{S_{G0}}(C - N_{H0}))F^{c}_{S_{G0}}(N_{H0}). \tag{A.6}$$

Since $F^{c}_{S_{G0}}(C - N_{H0})$ is nondecreasing in N_{H0}, and $\pi_H + p_H < \pi_G + p_G$, it follows that condition A.5 is satisfied at the point N^*_{H0} defined in (8) of Claim 2.

Case 2

$$C \geqslant F^{c}_{S_{G0}}{}^{-1}\left(\frac{\pi_W}{\pi_G + p_G}\right), \quad \text{or equivalently,}$$

$$F^{c}_{S_{G0}}(C) \leqslant \frac{\pi_W}{\pi_G + p_G}:$$

This condition implies that, if no rooms were given to 6 p.m. hold customers, some number of rooms would be allocated to walk-ins. Thus, the walk-in allocation depends upon the number of rooms that are given to 6 p.m. holds. It can be shown that at the point $N_{H0} = 0$,

$$\frac{\partial}{\partial N_{H0}}(E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})]) \geqslant \pi_H + p_H - \pi_W \geqslant 0.$$

Since this gradient is nonincreasing in N_{H0}, and is equal to zero at the point $N_{H0} = C - F^{c}_{S_{G0}}{}^{-1}((\pi_H + p_H)/(\pi_G + p_G))$, the optimality condition A.5 is satisfied at the point defined in (8) of Claim 2.

Lemma 1. *If $\pi_W < \pi_H + p_H$, then $E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})]$ is a pseudoconcave function of N_{H0}.*

Proof. To prove that $E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})]$ is pseudoconcave, we must show that for any $N_1, N_2 \in (0, \text{Min}\,(C, N_{H1}),$

$$\nabla_N(E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_1)]) \cdot (N_2 - N_1) \leqslant 0 \text{ implies that:}$$

$$E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_1)] \geqslant E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_2)].$$

This is equivalent to showing that the set

$$\Gamma = (N\colon N \geqslant 0 \quad \text{and} \quad \nabla_N(E_{S_{H0}}[R_H(G_1, S_{H0}, C, N)]) > 0)$$

is convex. Consider the partial derivative of $E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})]$ with respect to N_{H0}:

$$\nabla_{N_{H0}}(E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})]) = \frac{\partial}{\partial N_{H0}}\left[\int_{S_{H0}=0}^{H_1} R_H(G_1, S_{H0}, C, N_{H0}) f_{S_{H0}}(S_{H0})\, dS_{H0}\right]$$

$$= \frac{\partial}{\partial N_{H0}}\left[\int_{S_{H0}=0}^{N_{H0}} (\pi_H S_{H0} + MER_W(G_1, C - S_{H0})) f_{S_{H0}}(S_{H0})\, dS_{H0}\right]$$

$$+ \frac{\partial}{\partial N_{H0}}\left[\int_{S_{H0}=N_{H0}}^{H_1} ((\pi_H + p_H)N_{H0} - p_H S_{H0} + MER_W(G_1, C - N_{H0})) f_{S_{H0}}(S_{H0})\, dS_{H0}\right] \tag{A.7}$$

$$= \left[\int_{S_{H0}=N_{H0}}^{H_1} \left(\pi_H + p_H + \frac{\partial(C - N_{H0})}{\partial N_{H0}} \cdot \frac{\partial(MER_W(G_1, C - N_{H0}))}{\partial(C - N_{H0})}\right) \cdot f_{S_{H0}}(S_{H0})\, dS_{H0}\right] \tag{A.8}$$

$$= \left[\int_{S_{H0}=N_{H0}}^{H_1} \left(\pi_H + p_H - \frac{\partial(MER_W(G_1, C - N_{H0}))}{\partial(C - N_{H0})}\right) \cdot f_{S_{H0}}(S_{H0})\, dS_{H0}\right]. \tag{A.9}$$

Further analysis of expression A.9 necessitates that we evaluate the partial derivative of $MER_W(G_1, C)$ with respect to C:

$$\frac{\partial}{\partial C} MER_W(G_1, C) = \frac{\partial}{\partial C}\left[\int_{r_W=0}^{N^*_W(C)} (\pi_W r_W + E_{S_{G0}}[R_G(S_{G0}, C - r_W)]) \cdot f_r(r_W)\, dr\right]$$

$$+ \frac{\partial}{\partial C}\left[\int_{r_W=N^*_W(C)}^{\infty} (\pi_W N^*_W(C) + E_{S_{G0}}[R_G(S_{G0}, C - N^*_W(C))]) f_r(r_W)\, dr\right] \tag{A.10}$$

$$= \int_{r_W=0}^{N_W^*(C)} \frac{\partial}{\partial C}(\pi_W r_W + E_{S_{G0}}[R_G(S_{G0}, C - r_W)]) f_r(r_W)\, dr$$
$$+ \int_{r_W=N_W^*(C)}^{\infty} \left(\pi_W \frac{\partial N_W^*(C)}{\partial C} + \frac{\partial(E_{S_{G0}}[R_G(S_{G0}, C - N_W^*(C))])}{\partial(C - N_W^*(C))} \cdot \frac{\partial(C - N_W^*(C))}{\partial C}\right) f_r(r_W)\, dr$$
$$= (\pi_G + p_G) \int_{r_W=0}^{N_W^*(C)} F_{S_{G0}}^c(C - r_W) f_r(r_W)\, dr \quad \text{(A.11)}$$
$$+ \left(\pi_W \frac{\partial N_W^*(C)}{\partial C} + (\pi_G + p_G) F_{S_{G0}}^c(C - N_W^*(C)) \frac{\partial(C - N_W^*(C))}{\partial C}\right) F_{r_W}^c(N_W^*(C)). \quad \text{(A.12)}$$

Expression A.12 is obtained from (A.11) by substituting expression A.7 for the partial derivative of ER_G and integrating with respect to r_W. From the definition of $N_W^*(C)$ in (8) of Claim 1 it is easy to see that if:

$$C < F_{S_{G0}}^{c\ -1}\left(\frac{\pi_W}{\pi_G + p_G}\right): N_W^*(C) = 0,$$
$$\frac{\partial N_W^*(C)}{\partial C} = 0, \quad \text{and} \quad \frac{\partial(C - N_W^*(C))}{\partial C} = 1.$$

Otherwise

$$N_W^*(C) = C - F_{S_{G0}}^{c\ -1}\left(\frac{\pi_W}{\pi_G + p_G}\right),$$
$$\frac{\partial N_W^*(C)}{\partial C} = 1, \quad \text{and} \quad \frac{\partial(C - N_W^*(C))}{\partial C} = 0.$$

Thus, the gradient of $MER_W(G_1, C)$ can be expressed in one of two ways.

Case 1

$$C < F_{S_{G0}}^{c\ -1}\left(\frac{\pi_W}{\pi_G + p_G}\right) \Rightarrow \frac{\partial}{\partial C} MER_W(G_1, C) = (\pi_G + p_G) \cdot F_{S_{G0}}^c(C). \quad \text{(A.13a)}$$

Case 2

$$C \geq F_{S_{G0}}^{c\ -1}\left(\frac{\pi_W}{\pi_G + p_G}\right) \Rightarrow \frac{\partial}{\partial C} MER_W(G_1, C) = (\pi_G + p_G) \int_{r_W=0}^{N_W^*(C)} F_{S_{G0}}^c(C - r_W) f_r(r_W)\, dr + \pi_W \cdot F_{r_W}^c\left(C - F_{S_{G0}}^{c\ -1}\left(\frac{\pi_W}{\pi_G + p_G}\right)\right). \quad \text{(A.13a)}$$

Using the definition of $N_W^*(C)$ in Claim 1, it is easy to show that the value of expression A.13b is less than or equal to π_W. We can now substitute expression A.13 and the definition of $N_W^*(C)$ into (A.9) to obtain two alternative expressions for the gradient of $E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})]$.

Case 1

$$N_{H0} < C - F_{S_{G0}}^{c\ -1}\left(\frac{\pi_W}{\pi_G + p_G}\right):$$
$$\nabla_{N_{H0}}(E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})])$$
$$= \left(\pi_H + p_H - \pi_W F_{r_W}^c\left(C - N_{H0} - F_{S_{G0}}^{c\ -1}\left(\frac{\pi_W}{\pi_G + p_G}\right)\right)\right) F_{S_{H0}}^c(N_{H0}) - (\pi_G + p_G) \cdot \int_{r_W=0}^{C - N_{H0} - F_{S_{G0}}^{c\ -1}(\pi_W/\pi_G + p_G)} \cdot F_{S_{G0}}^c(C - N_{H0} - r_W) f_r(r_W)\, dr \geq \pi_H \cdot p_H - \pi_W. \quad \text{(A.14a)}$$

Case 2

$$N_{H0} \geq C - F_{S_{G0}}^{c\ -1}\left(\frac{\pi_W}{\pi_G + p_G}\right):$$
$$\nabla_{N_{H0}}(E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})])$$
$$= (\pi_H + p_H - (\pi_G + p_G) F_{S_{G0}}^c(C - N_{H0})) F_{S_{H0}}^c(N_{H0}), \quad \text{(A.14b)}$$

where the notation $F_X^c(x) = 1 - F_X(x)$ is used to represent the complementary cumulative distribution of random variable X at the point x.

To prove the claim, it remains to be shown that the set $\Gamma = (N_{H0}\colon N_{H0} \geq 0$ and $\nabla_{N_{H0}}(E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})]) > 0)$ is convex. From (A.14a), we can use the fact that $\pi_W < \pi_H + p_H$ to see that:

$$\nabla_{N_{H0}}(E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})]) > 0, \quad \text{for} \quad N_{H0} \in \left(0,\ C - F_{S_{G0}}^{c\ -1}\left(\frac{\pi_W}{\pi_G + p_G}\right)\right).$$

It is easy to show that (A.14a, b) are equal to one another at the point:

$$N_{H0} = C - F_{S_{G0}}^{c\ -1}\left(\frac{\pi_W}{\pi_G + p_G}\right).$$

Finally, we can show that expression A.14b is nonincreasing by using the fact that complementary cumulative distributions are nonincreasing. It follows that the set $\Gamma = (N_{H0}\colon N_{H0} \geq 0$ and $\nabla_{N_{H0}}(E_{S_{H0}}[R_H(G_1, S_{H0}, C, N_{H0})]) > 0)$ is convex.

APPENDIX B

Table VII
Scenario B

t	$E[r_G]$	$E[r_H]$	$E[r_W]$	q_G	q_H
4	100	40	—	0.9	0.9
3	100	40	—	0.9	0.9
2	100	40	—	0.9	0.9
1	100	40	—	0.9	0.9
0	—	—	30	0.9	0.5

Table VIII
Scenario C

t	$E[r_G]$	$E[r_H]$	$E[r_W]$	q_G	q_H
4	40	40	—	0.9	0.9
3	40	40	—	0.9	0.9
2	40	40	—	0.9	0.9
1	40	40	—	0.9	0.9
0	—	—	30	0.9	0.5

Table IX
Scenario D

t	$E[r_G]$	$E[r_H]$	$E[r_W]$	q_G	q_H
4	70	40	—	0.9	0.9
3	70	40	—	0.9	0.9
2	70	40	—	0.9	0.9
1	70	40	—	0.9	0.9
0	—	—	60	0.9	0.5

ACKNOWLEDGMENT

The authors are grateful to Professor Susana Mondschein and an anonymous referee for their comments on earlier versions of this paper.

REFERENCES

ALSTRUP, J., S. BOAS, O. B. G. MADSEN AND R. V. VALQUI. 1986. Booking Policy for Flights With Two Types of Passengers. *Eur. J. Opnl. Res.* **27,** 274–288.

BELOBABA, P. P. 1987. Airline Yield Management—An Overview of Seat Inventory Control. *Trans. Sci.* **21,** 63–73.

BELOBABA, P. P. 1989. Application of a Probabilistic Decision Model to Airline Seat Inventory Control. *Opns. Res.* **37,** 183–197.

BITRAN, G. R., S. M. GILBERT AND T. Y. LEONG. 1992. Hotel Sales and Reservations Planning. In *Proceedings to the 1992 Wharton Conference on Service Management, Technology and Economics: The Service Productivity and Quality Challenge* (to appear).

BRUMELLE, S. L., AND J. I. MCGILL. 1993. Airline Seat Allocation With Multiple Nested Fare Classes. *Opns. Res.* **41,** 127–137.

DRAKE, A. W. 1967. *Fundamentals of Applied Probability Theory*. McGraw-Hill, New York.

LADANY, S. 1976. Dynamic Operating Rules for Motel Reservations. *Dec. Sci.* **7,** 829–840.

LIBERMAN, V., AND U. YECHIALI. 1978. On the Hotel Overbooking Problem—An Inventory System with Stochastic Cancellations. *Mgmt. Sci.* **24,** 1117–1126.

MANGASARIAN, O. L. 1969. *Nonlinear Programming*. McGraw-Hill, New York.

ROTHSTEIN, M. 1974. Hotel Overbooking as a Markovian Sequential Decision Process. *Dec. Sci.* **5,** 389–404.

ROTHSTEIN, M. 1985. OR and the Airline Overbooking Problem. *Opns. Res.* **33,** 237–248.

WILLIAMS, F. E. 1977. Decision Theory and the Innkeeper: An Approach for Setting Hotel Reservation Policy. *Interfaces* 7(4), 18–31.