Manufacturing & Service Operations Management

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

Decision Biases in Revenue Management: Some Behavioral Evidence

J. Neil Bearden, Ryan O. Murphy, Amnon Rapoport,

MSOM MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 10, No. 4, Fall 2008, pp. 625–636
ISSN 1523-4614 | EISSN 1526-5498 | 08 | 1004 | 0625

informs®
DOI 10.1287/msom.1080.0229

Decision Biases in Revenue Management: Some Behavioral Evidence

J. Neil Bearden
Decision Sciences Area, INSEAD, Singapore, 138676, neil.bearden@insead.edu

Ryan O. Murphy
Center for the Decision Sciences, Columbia University, New York, New York 10027, rom2102@columbia.edu

Amnon Rapoport
Department of Management and Organizations, University of Arizona, Tucson, Arizona 85721, amnon@u.arizona.edu

We study a problem of selling a fixed number of goods over a finite and known horizon. After presenting a procedure for computing optimal decision policies and some numerical results of a simple heuristic policy for the problem, we describe results from three experiments involving financially motivated subjects. The experiments reveal that decision makers employ decision policies of the same form of the optimal policy. However, they show systematic biases to demand too much when they have many units to sell and too little when they have few to sell, resulting in significant revenue losses.

Key words: behavioral operations; revenue management; dynamic pricing; decision bias; heuristics
History: Received: June 2, 2006; accepted: July 23, 2007. Published online in *Articles in Advance* August 5, 2008.

1. Introduction

Firms often face the problem of deciding how to best price and control the inventory of perishable products for which demand is stochastic and price sensitive. Airlines must do so for seats on particular flights; hotels must do so for rooms on particular nights; and fashion retailers must do so for seasonal goods. Keeping prices too high can result in unsold items, but keeping them too low can have significant opportunity costs, because costumers would have been willing to pay more. There has been considerable theoretical work in the operations management literature on methods for optimally solving pricing and revenue management (RM) problems, but, as far as we know, there has been no direct experimental work on how well actual decision makers do so. Because managers—who are not necessarily perfectly rational decision makers nor extensively trained in optimization methods—are generally responsible for RM decisions in most firms, investigating how their decisions may be biased should be valuable. In this paper, we use laboratory experiments to investigate decision behavior in a stylized RM problem that captures many of the important features of the problems faced by practicing managers.

Suppose a firm has a finite number of periods—a *season*—in which to sell a fixed number of units of a product. Bids to buy a unit of the product at a particular price arrive sequentially and stochastically in time. Each time it receives a bid, the firm must choose to either accept or reject it on the spot. When it accepts one, it sells a unit of the product to the bidder at the bid price. Otherwise, it irrevocably rejects the bid and must wait for another one, which may or may not come before the end of the season.

There are a number of ways to interpret this general problem. One is to think of it as the one faced by airlines, hotels, and travel agencies that sell their goods on Priceline.com. Visitors to the site make offers to purchase goods (e.g., a single one-way ticket from Tucson to New York on July 5, 2006) at particular prices, and their offers are either accepted or rejected. The visitor's credit card is automatically charged the bid price if the bid is at least as high as the current reservation price for the good, which is determined by Priceline.com, and is, of course, unknown to the visitor; otherwise, she pays nothing and receives nothing. Another interpretation is that goods in different product classes (e.g., fare classes) are priced by the seller and posted. Then, when a buyer attempts to buy a good in a particular class at the posted price, the seller decides whether to make a unit of the good available. Whatever interpretation one assigns, this general problem possesses the fundamental features of problems faced by firms in many industries—namely, a fixed stock of items, a finite selling horizon, and uncertain demand—and is the archetypal problem in RM. It is also the type of problem we study here.

A number of excellent reviews may be consulted for introductions to and overviews of research in pricing and RM (e.g., Bitran and Caldentey 2003, McGill and van Ryzin 1999, Talluri and van Ryzin 2004, Weatherford and Bodily 1992). Therefore, we have chosen instead to focus on some of the prior experimental studies of decision behavior that are most relevant to the RM problem we consider here. These studies all share a common framework: They investigate behavior in sequential decision problems with known optimal decision policies. Our paper employs this same approach.

Optimal stopping is central to many operations management decision problems, such as when to hire a job applicant and when to adopt a new technology. It is also at the core of many RM problems (Brumelle and McGill 1993). The general theory of optimal stopping has received considerable attention (e.g., Chow et al. 1971, Gilbert and Mosteller 1966), and there has been some work in the experimental literature on the stopping behavior of actual decision makers. Rapoport and Tversky (1970), for example, examined decision behavior in the classical full-information optimal stopping problem in which the decision maker (DM) sequentially observes up to N random draws from a distribution with a known density $f(x)$, must irrevocably accept or reject each draw when it is observed, and receives a payoff equal to the value x of the single selected observation. Rapoport and Tversky found that DMs tended to stop sooner than was dictated by the optimal—expected payoff-maximizing—policy. Similar findings regarding behavior in full-information optimal stopping problems were reported, among others, by Cox and Oaxaca (1989) and Schotter and Braunstein (1981).

One conclusion from the experimental studies of optimal stopping problems is that people have a propensity to search inadequately but employ rather complex decision policies that have the same structural form as the optimal policies (see, e.g., Bearden et al. 2005, 2006a, 2006b; Seale and Rapoport 1997, 2000). Assuming that the early stopping bias generalizes to situations beyond the laboratory, one might predict that people do not search enough before making purchase decisions; for example, they may not visit enough sites before purchasing books or airline tickets online. Indeed, Johnson et al. (2004) show that online search is quite limited. For many search problems, such as online shopping, it would be very difficult, if not impossible, to determine what a truly optimal search would require. This is one reason laboratory studies are so powerful. We can confront actual DMs with decision problems for which we know the optimal policies. By comparing actual to optimal decision behavior, we can gain insights into where decision making is done well and where it breaks down.

In all previous experimental studies of optimal stopping, the DM had a *single* unit to sell (e.g., a single position to fill). It is unclear how previous behavioral results on optimal stopping inform the more general RM problem in which the DM has *multiple* units to sell over a fixed period of time. Do the findings suggest that DMs are likely to set their selling prices too low, for example? The problems are sufficiently different that the answer is not known.

Overall, given the dearth work on decision behavior in dynamic decision problems, it is difficult to derive specific predictions for how DMs will perform in RM problems such as the Priceline.com problem described above. We can, however, predict the following: Decision behavior will depart from optimality. Given the relative complexity of making optimal RM decisions, this prediction is obvious and not all that interesting by itself. However, by finding the *ways* decision behavior *systematically departs from* optimality, we can establish a basis for prescription. At the least, experimental work on these problems can be used to warn DMs about the broad ways in which their RM decisions are likely to err. This by itself, we believe, is valuable.

In addition to increasing our understanding of decision behavior in RM settings, the current paper also makes a modest methodological contribution. So far, no clear convention has been established for dealing with data from experiments on multistage decision problems in which each subject performs a large number of trials. Because future experiments on operations management decision problems are likely to produce such data, it is important that we apply procedures that allow us to draw statistically justified inferences from them. Here, we employ a variant of the random regression approach to analyzing repeated measures from multiple subjects in our analyses (cf. Lorch and Myers 1990). Our methods can be extended and applied easily to other domains.

The rest of the paper is organized as follows. In §2, we formally describe a simplified RM problem and present a numerical procedure for computing its optimal decision policies. We then demonstrate in §3 that a relatively simple decision heuristic can perform quite well in the problem. Next, in §4, we present results from three behavioral experiments involving financially motivated subjects in which we examine actual decision behavior in the problem. Finally, §5 contains a discussion of the experimental findings and some suggestions for future experimental research on RM.

2. Problem and Solution

In this section, we describe a stylized RM problem and present a method for computing its optimal policy. The problem has appeared under various guises in the operations literature. Lee and Hersh (1993), for example, present it as a model of airline seat inventory control. Papastavrou et al. (1996) describe it quite generally as a dynamic and stochastic knapsack problem, and relate it to transportation scheduling problems, taking reservations in restaurants, and airline booking. To provide ourselves with a useful shorthand, we will refer to the problem as the *revenue management problem* (RMP).

2.1. RMP

A DM can sell up to S units over a *season* of T discrete time periods. Periods are indexed by t ($t = T-1, \ldots, 0$), which represents the number of *periods remaining* until the end of the season. Units cannot be sold after period $t=1$, and the salvage value for a unit is set at 0. The number of available-to-be-sold units is indexed by s ($s = S, S-1, \ldots, 0$). The *state* of the system is given by (t, s). In each period, an offer to buy a single unit arrives with probability p, and no offers arrive with probability $1-p$. Each offer has an associated bid (revenue) r, which is a random variable taken from a distribution with density $f(r)$. Whenever the DM has $s \geq 1$ units and receives an offer with bid r, she can sell a unit, thereby increasing her total revenue by r and leaving her with $s-1$ units, or she can reject the offer. The decision to either accept or a reject an offer cannot be delayed—it must be made instantaneously. The DM's objective is to maximize her expected (total) revenue for the selling season.

2.2. Dynamic Programming Solution for RMP

Based on Lee and Hersh (1993) and Papastavrou et al. (1996), we know that at stage t with s remaining units, the optimal policy is a threshold rule:

$$\psi^\star(t, s, r) = \begin{cases} \text{accept offer} & \text{if } r \geq R_t^s, \\ \text{reject offer} & \text{if } r < R_t^s. \end{cases}$$

The *thresholds* R_t^s dictate what revenue levels r the DM finds acceptable given t and s. These thresholds are computed from

$$R_t^s = \begin{cases} V_{t-1}^s - V_{t-1}^{s-1} & \text{if } s \geq 1, \\ \infty & \text{if } s < 1, \end{cases} \tag{1}$$

where

$$V_t^s = p\left[\int_0^{R_t^s} f(r) V_{t-1}^s \, dr + \int_{R_t^s}^{\infty} f(r)(r + V_{t-1}^{s-1}) \, dr\right] + (1-p) V_{t-1}^s, \tag{2}$$

with boundary conditions

$$V_0^s = 0, \quad \forall s, \quad \text{and} \quad V_t^0 = 0, \quad \forall t.$$

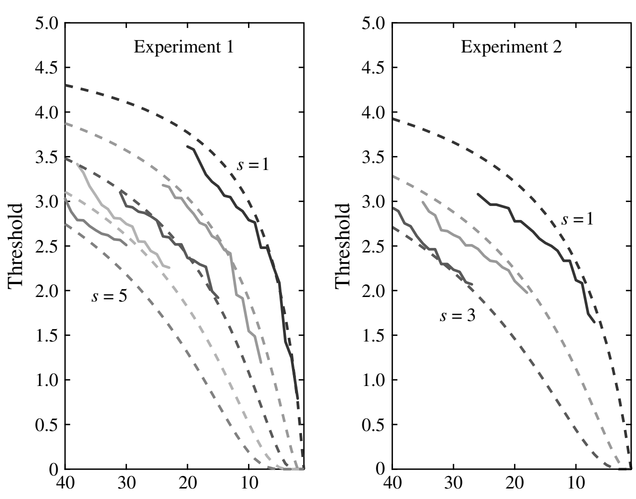
The value function V_t^s gives the DM's expected future revenue for following $\psi^\star(t, s, r)$ from period t to period 0, given that she has s units left. The optimal policy depends only on t, s, and r, and not on the history prior to t. Thus, Bellman's (1957) optimality principle of dynamic programming is satisfied; and the optimal policy for the full problem from period T to period 0 can be obtained by solving Equations (1) and (2) recursively from $t = 0$ to $t = T$. Under the optimal policy, when she receives an offer, the DM simply decides whether the expected marginal value for holding a unit for one more period exceeds the marginal revenue for selling it at the current bid value. If she is (expected to be) better off keeping the unit, she does so; otherwise, she sells it.

From Papastavrou et al. (1996), we know that the optimal policy for the RMP has the following properties:

(i) R_t^s is nonincreasing in s for all t.
(ii) R_t^s is nondecreasing in t for all s.

The optimal DM is less choosy when she has more units to sell and when she has less time to sell them. The intuitions for this result are clear. Because the DM is faced with a deadline beyond which she can no longer sell her units, she must be less demanding when she has a large number of units to sell; otherwise, because future demand is uncertain, she may end up with unsold units, which are worthless. The DM should become less demanding as her deadline approaches, because getting something for a unit is better than getting nothing. These properties can be discerned from the thresholds shown in (the dashed lines in) Figure 1. (We will discuss the estimated thresholds (solid lines) in the figure below when we present our experimental results.) The pattern of thresholds displayed in the figure also holds for other bid distributions such as the normal, exponential, and triangular—it is not peculiar to the uniform distribution.

Figure 1 **Optimal (Dotted Lines) and Mean Estimated (Solid Lines) Thresholds for Experiments 1 (Left) and 2 (Right)**



Note. For each value of s, estimated thresholds are only shown for values of t at which at least 2% of the offers for that s were encountered.

Next, we will present some numerical results on the performance of a simple, nondynamic decision heuristic for the RMP. Then, we will describe three behavioral experiments in which we examine the decision behavior of actual financially motivated DMs in the RMP.

3. Simple Heuristic for RMP

Much has been made in the psychology literature of the usefulness of simple decision heuristics (e.g., Gigerenzer et al. 1999). The now relatively conventional argument—which actually dates back to Simon (1955)—is that human DMs have limited computational capacity but are generally able to make good decisions using simple heuristics. Often, work along these lines proceeds by simply demonstrating that simple decision heuristics *can* perform well, often using Monte Carlo simulation to do so. Less frequently, researchers actually test whether people employ simple heuristics when making decisions (for some exceptions, see Bröder 2000, Newell and Shanks 2003, Johnson et al. 2008).

We showed above that expected revenue maximization in the RMP requires the use of relatively sophisticated dynamic decision policies. Under these, the DM determines her acceptable revenue levels (i.e., her thresholds) in each period after taking into consideration both how many periods she has left to sell units and how many units she has to sell. But how might a DM fare if she employed a simple, static decision policy? The simplest heuristic a DM might employ is the *fixed-threshold policy*: Accept any offer for which $r > R$, where R is fixed for all t and s. If a DM wants to optimize the performance of this policy, how should she set her threshold? Further, how effective would such a fixed-threshold policy be?

The value (expected revenue) of a fixed-threshold policy V_H can be obtained by substituting the single threshold R for each R_t^s in Equation (2), performing the recursion from $t = 0$ to $t = T$, and setting $V_H = V_T^S$. The optimal heuristic threshold R_H is found by solving

$$R_H = \arg\max_R V_H, \qquad (3)$$

Table 1 Optimal Fixed Thresholds R_H, Expected Earnings Under Optimal Fixed-Threshold Policy V_H, Expected Earnings Under (Full-Blown) Optimal Policy V_T^S, and Efficiency of Fixed-Threshold Policy V_H/V_T^S for Some RMPs

Problem	T	S	p	R_H	V_H	V_T^S	V_H/V_T^S
1	40	5	0.30	2.42	17.05	17.66	0.97
2	40	3	0.18	2.29	9.65	10.01	0.97

which can be done using line-search methods. Some numerical results on the performance of the fixed-threshold policy are displayed in Table 1 for some special cases of the RMP. It turns out that very little is lost by using a simple fixed-threshold heuristic in these problems. For each of them, a DM can expect to earn more than 97% of the optimal expected earnings using an optimal fixed threshold. We evaluated V_H/V_T^S for a large number of other combinations of T, S, and p and always found that $V_H/V_T^S > 0.94$. Further, our numerical experiments show that V_H/V_T^S tends to 1 as T grows.

Gallego and van Ryzin (1994) reported a similar result in a dynamic pricing problem. They showed that the expected revenue under a fixed price heuristic was always close to the optimal revenue, and, in fact, converged to the optimal revenue as the selling horizon T grew. Based on their results, they concluded that when demand functions are well known and prices can be set freely, there is likely to be little benefit from dynamic pricing policies. Next, we describe three behavioral experiments in which we examine actual decision behavior in the RMP.

4. Behavioral Studies of RMP

4.1. Overview of Experimental Method

We examined decision behavior in the RMP in three experiments. All had the same general setup and used incentive-compatible payoffs to encourage careful decision making. The experiments differed in the offer arrival probabilities p, the number of to-be-sold units S, and whether subjects could make accept or reject decisions for each offer (Experiments 1 and 2) or were constrained to use a single threshold over the entire course of each season (Experiment 3). We varied the problem parameters in Experiments 1 and 2 to ensure that any observed behavioral patterns in the RMP would not be the consequence of one particular set of parameters. And to increase sensitivity, we selected parameters for the problems that resulted in quite dynamic optimal policies. (For example, we wanted to avoid problem parameterizations whose optimal policies were extremely flat with respect to t.) Of course, one could manipulate any number of parameters of the RMP, including the duration of the selling season T and bid distribution $f(r)$. Our main objective has been to look for broad, replicable patterns of decision behavior in the RMP, and we felt that our manipulations would allow us to cover a large region of the feasible problem space.

We fixed $T = 40$ and $r \sim \text{Uni}[0.01, 5.00]$ for all three experiments. Experiments 1 and 3 had $S = 5$ and $p = 0.30$. Both the number of available units ($S = 3$) and the arrival probability ($p = 0.18$) were lowered in Experiment 2. To maintain some basis for comparison, we held the ratio of expected number of offers pT to available units constant across all three experiments ($pT/S = 2.40$). In Experiments 1 and 2, subjects could make accept or reject decisions for each offer they received. In Experiment 3, subjects were forced to set a single threshold at the beginning of each season, which was then used to automatically make accept or reject decisions for each encountered offer.

Thirty-four, thirty-three, and thirty-six subjects participated in Experiments 1, 2, and 3, respectively. All were recruited through flyers posted around the Columbia University School of Business to take part in a decision-making experiment. Experiments 1 and 2 were conducted simultaneously, with subjects being randomly assigned to one or the other on arrival to the lab. Experiment 3 was run after the first two were completed.

The subjects were paid based on their performance in the experimental task and did not receive any course credit. Specifically, they were paid from one randomly selected trial in Experiments 1 and 3, and from two randomly selected trials in Experiment 2. (This procedure kept the expected earnings across experiments roughly the same, because there were fewer units to sell in Experiment 2.) They earned an average of around $17 for the 1-hour session.

Each subject was provided with extensive written instructions describing the task and the interface of the computer program that administered the experiment. The cover story for the task involved selling "contracts" to "bidders." The instructions described the RMP in nontechnical language, and the values of the parameters of the problem (T, S, p, and $f(r)$) were all presented explicitly. To be clear, the subjects faced the RMP with perfect information about the problem parameters—there was no ambiguity. Once the subjects were confident in their understanding of the task, they performed 50 (independent) trials of the RMP. On each trial, the program automatically advanced through periods in which no offers arrived, pausing for 500 milliseconds in each period. To emphasize that there was a selling deadline, the number of remaining periods was displayed textually (e.g., "Periods Remaining: 20") and also by a graphic progress bar that shrank in each period.

In Experiments 1 and 2, whenever an offer arrived, the subject was shown the bid value and asked to choose to either accept or reject it. In Experiment 3, the subject set his or her threshold for the whole season *prior to* the beginning of each season. Then each time a contract arrived, the subjects observed the value of the contract and whether they accepted or rejected it. No time restrictions were imposed in the accept-reject decisions. The computer program also continuously updated and displayed the number of available contracts, the revenue from each sold contract, and the total revenue to date (for the current trial). A trial terminated either when the deadline was reached or all contracts had been sold, whichever came first. The arrivals and offer values were generated randomly and independently for each subject by the experimental program according to the appropriate experimental parameters.

Given the similarity of the experiments, and to conserve space, we will report all the results together.

4.2. Results

4.2.1. Revenues. Table 2 presents the average revenues for each of the three experiments. The first column contains the averages over all 50 trials in each experiment. The second and third columns present the average earnings in the first and last 20 trials. There are several important findings. First, in each experiment, the average revenues were significantly lower than those expected under the application of the optimal policy. This holds in all three partitionings: overall, first 20 trials, and last 20 trials. Second, in Experiments 1 and 2, where subjects could dynamically make accept-reject decisions, the average revenues in the first and last 20 trials of the experiment are not significantly different from one another. In Experiment 3, on the other hand, we do find evidence of learning: average earnings in the last 20 trials are significantly greater than those in the first 20 trials.

Table 2 **Average Revenues in Experiments 1–3**

Experiment	All	First 20	Last 20	Optimal
1	17.09* (0.50)	17.01* (0.82)	17.20* (0.62)	17.66
2	9.69* (0.37)	9.71* (0.70)	9.58* (0.54)	10.01
3	16.59* (0.72)	16.36* (1.12)	16.83*,** (0.90)	17.05

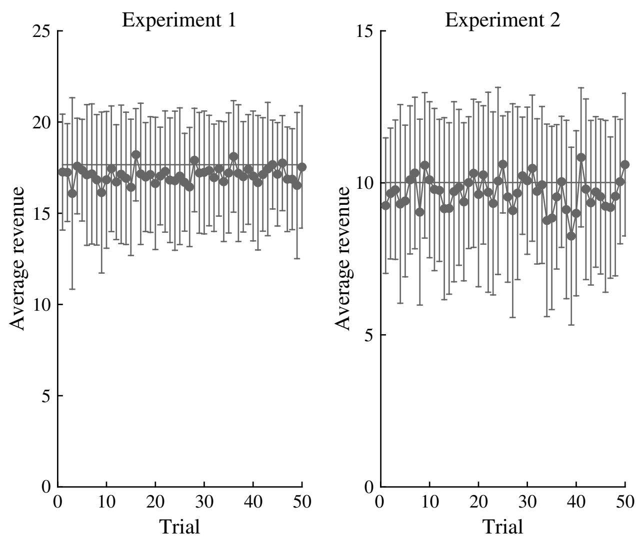
Notes. Values in parentheses represent the standard deviation of average revenues taken across subjects. Each average was compared to the optimal expected revenues using a t-test; those marked with an $*$ were significantly different from the optimal expected revenues at the $\alpha = 0.05$ level. In addition, the averages for the first and last 20 trials were compared using a repeated-measures t-test. Only the differences in mean revenues in Experiment 3 were significant at the $\alpha = 0.05$ level (marked with a $**$).

Importantly, there is no significant difference in the average earnings for the last 20 trials of Experiments 1 and 3. That is, after the learning phase (i.e., the first 30 trials), those subjects who were free to dynamically make accept-reject decisions did not earn significantly higher revenues than those who were forced to use a fixed threshold.

Figures 2 and 3 (left-hand panels) exhibit the average revenues across the 50 experimental trials. To further examine learning, we regressed the average earnings onto trial number. The slope coefficients were not significant for Experiment 1 ($b = 0.005$, $p = 0.23$) or Experiment 2 ($b = -0.002$, $p = 0.65$); however, the slope for Experiment 3 was positive and significant ($b = 0.02$, $p < 0.01$). Thus, there is no evidence that subjects in the first two experiments were able to modify their policies with experience to increase their revenue, whereas they were able to do so in Experiment 3.

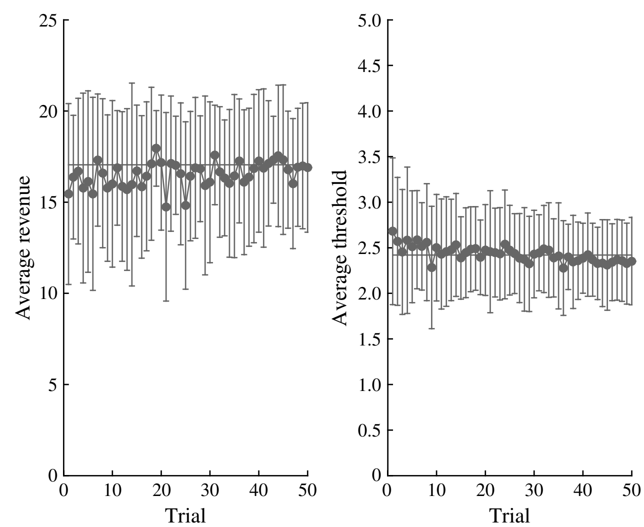
4.2.2. Opportunity Cost Analysis. Obviously, because the subjects did not earn as much as expected under the optimal policy, we may conclude that they used some other, nonoptimal policies. By definition,

Figure 2 **Average Revenues Across Trials in Experiments 1 (Left Panel) and 2 (Right Panel)**



Notes. The error bars represent one standard deviation. The horizontal lines represent the optimal expected earnings.

Figure 3 **Average Revenues (Left Panel) and Average Threshold (Right Panel) from Experiment 3**



Notes. The error bars represent one standard deviation. The horizontal line in the left panel represents the optimal expected earnings, and the one in the right panel represents the optimal single threshold value.

any accept or reject decision that departs from the dictates of the optimal policy will decrease expected revenue. Here, we consider the *implied revenue loss* (or *opportunity costs*) resulting from departures from the application of the optimal policies. There are two types of nonexpected revenue-maximizing errors: rejecting an offer that is good enough or accepting an offer that is not good enough. We will refer to these as *accept* and *reject* errors, respectively. We would like to determine which errors are most common, which tend to be most costly, and under what conditions these errors are most likely to occur.

The *implied revenue loss* for an accept error is

$$L_{\text{acc}}(r_t^s) = r_t^s - R_t^s.$$

Similarly, the implied revenue loss for a reject error is

$$L_{\text{rej}}(r_t^s) = R_t^s - r_t^s.$$

To test whether one of the two error types had a greater impact on revenue losses, we computed the mean implied revenue loss for each type for each subject and used paired-sample t-tests to compare the two. In both Experiments 1 and 2, we find that the average implied revenue losses are greater for accept errors than for reject errors. In contrast, reject errors were more numerous than accept errors in Experiments 1 and 2, though only significantly so in the latter. There was no significant difference in average implied revenue losses from accept and reject errors in Experiment 3; nor was one error type more common than the other. The summary results and test statistics from these analyses are reported in Table 3.

Tables 4 and 5 contain the aggregate proportion of accept errors (conditional on an error) broken down by t and s for Experiments 1 and 2, respectively. The

Table 3 $P(L_{\text{acc}})$ $(P(L_{\text{rej}}))$ **Denotes Average Proportion of Observed Periods on Which an Accept (Reject) Error Was Made**

Experiment	$P(L_{\text{acc}})$	$P(L_{\text{rej}})$	t-value	L_{acc}	L_{rej}	t-value
1	0.05 (0.04)	0.06 (0.05)	$t_{33} = 0.93$	−0.42 (0.15)	−0.31 (0.14)	$t_{33} = 3.19^*$
2	0.04 (0.03)	0.09 (0.05)	$t_{32} = 3.71^*$	−0.48 (0.22)	−0.37 (0.18)	$t_{30} = 2.70^*$
3	0.04 (0.05)	0.04 (0.04)	$t_{35} = 0.31$	−0.27 (0.19)	−0.29 (0.16)	$t_{21} = 0.75$

Notes. Average revenue losses for suboptimal acceptances are denoted by L_{acc} and suboptimal rejections by L_{rej}. The standard deviations of the averages across subjects are shown in parentheses. *T*-tests marked with an asterisk were significant at the $\alpha = 0.05$ level.

Table 4 Aggregate Proportion of Accept Errors (Conditional on Making an Error) Broken Down by s and t for Experiment 1

	t				
	31–40	21–30	11–20	1–10	Average
$s=1$	1.00	0.94	0.78	0.66	0.76
$s=2$	0.80	0.77	0.58	0.71	0.67
$s=3$	0.79	0.55	0.38	0.57	0.51
$s=4$	0.48	0.38	0.41	1.00	0.43
$s=5$	0.25	0.22	0.25	—	0.24
Average	0.39	0.54	0.57	0.67	

Note. The margins provide the weighted (by number of occurrences of errors) average for each t and s.

pattern of errors is quite consistent across the two experiments. We should note that in those few cases where there is a relatively large difference in the averages across the two experiments (e.g., when $s=3$ and $31 \leq t \leq 40$), the average for at least one experiment is based on a relatively small number of observations. For example, having $s=3$ when $31 \leq t \leq 40$ was reasonably rare in Experiment 1 but quite common in Experiment 2.

To draw statistically justified inferences regarding the effects of t and s on tendencies to make accept and reject errors, we must employ a more nuanced line of reasoning. Because we have a large number of observations from each experimental subject, we must take the dependencies between a subject's responses into account in our analyses. To do so, we extend the random regression approach put forward by Lorch and Myers (1990). Specifically, we fit a logistic regression model to each individual subject's data and use the resulting (statistically independent) coefficients in subsequent (nonparametric) analyses.

To determine whether the errors are more strongly associated with time or inventory levels, we fit logistic regression models to each individual subject's error data. Let $q_{t,s}$ denote the proportion of times that a reject error was made in state (t, s) conditional on an error being made. (Therefore, $1-q_{t,s}$ is the proportion of times that an accept error was made in state (t, s) conditional on an error being made.) We fit the following logistic regression model to *each individual* subject's (error) data:

$$\text{logit}(q_{t,s}) = \beta_\varnothing + \beta_t t + \beta_s s + \beta_{t,s}(t \times s). \tag{4}$$

The model allows us to estimate the effects of t and s (and their interaction) on the probability of each error type. What we would like to know is whether there is a *systematic* relationship between t and s and the error types. The reasoning we employ goes as follows: If the signs of a coefficient are systematically positive or negative across subjects, then this would suggest that the errors are systematically related to the variable to which that coefficient corresponds. For instance, if β_s were consistently negative across subjects, this would suggest that subjects tend to make more accept errors as their inventory decreases; but if β_s were just as likely to be positive as negative, then this would suggest no systematic relationship (across subjects). To test the coefficients formally, for Experiments 1 and 2, we conducted (two-tailed) binomial tests on each coefficient in which "successes" were defined as negative coefficients and the probability of success under the null hypothesis was assumed to be 0.50. The proportions of negative coefficients are displayed in Table 6. For the β coefficients, we only find a significant relationship between inventory level (s) and error type. Specifically, in both experiments we find that accept errors tend to increase as inventory levels decrease. Two sample z-tests comparing the proportion of negative coefficients revealed no significant difference between conditions for any of the four

Table 5 Aggregate Proportion of Accept Errors (Conditional on Making an Error) Broken Down by s and t for Experiment 2

	t				
	31–40	21–30	11–20	1–10	Average
$s=1$	0.95	0.87	0.76	0.66	0.81
$s=2$	0.87	0.65	0.59	0.72	0.67
$s=3$	0.52	0.45	0.61	0.00	0.50
Average	0.67	0.66	0.68	0.69	

Note. The margins provide the weighted (by number of occurrences of errors) average for each t and s.

Table 6 Proportion of Negative Logistic Regression Coefficients from Experiments 1 and 2

Experiment	$P(\beta_\varnothing < 0)$	$P(\beta_t < 0)$	$P(\beta_s < 0)$	$P(\beta_{ts} < 0)$
1	0.18*	0.47	0.71*	0.56
2	0.27*	0.39	0.85*	0.58

Note. Entries marked with an asterisk are significant at the $\alpha = 0.05$ level using a two-tailed binomial test assuming the true (population) proportion of negative coefficients is 0.50.

model coefficients. Based on these analyses, there is no systematic relationship between time-remaining (t) and accept or reject errors. We conclude that the departures from optimality tended to be primarily due to inappropriate sensitivity to inventory: subjects were more likely to make reject decisions when inventory levels were higher and to make accept decisions when inventory levels (s) were lower. The errors tended to be independent of time (t).

4.2.3. Estimating Decision Policies. We have seen that the subjects in each experiment earned less than predicted by the application of the appropriate optimal policy. It is easy to determine the policies employed by subjects in Experiment 3, where the subjects were forced to specify a single threshold for each season. In contrast, in Experiments 1 and 2 we only observed whether a subject accepted or rejected each offer; we must, therefore, estimate (or infer) individual decision policies from the observable data.

Some questions we wish to address are: What kind of decision policies do subjects employ in the unconstrained RMP? Do they use sophisticated policies or simple heuristics? It is reasonable to assume that the subjects employ some kind of threshold when making decisions. In what follows, we will assume that each subject employs a threshold decision policy of the form:

$$\begin{cases} \text{accept offer} & \text{if } r \geq \hat{R}_t^s, \\ \text{reject offer} & \text{if } r < \hat{R}_t^s, \end{cases}$$

where $\hat{R}_t^s$ is the subject's (empirical) threshold for state (t, s). Given this assumption, we would like to find the thresholds that best predict the subjects' decision data. There are four obvious threshold-setting policies. The first, which we term a *sophisticated threshold policy*, permits the DM to adjust her threshold as a function of both t and s. The optimal policy is a special case (parameterization) of this policy. Another possible policy that a DM might employ is to set her threshold only on the basis of how many units she has left to sell; a third policy would be to judge acceptable offer values solely on the basis of time remaining to sell units. Finally, the simplest reasonable policy that a DM could employ is to decide on a target marginal revenue and to then only accept offers whose associated revenues exceed her target, regardless of inventory and how much time remains in the selling season. As we showed above in §3, this policy is not as näive as it might appear: a DM who employs it can do quite well.

The last three decision policies are all special cases of the sophisticated threshold policy. Therefore, evidential support for any of them will provide support for the sophisticated policy. On the other hand, in principle, the simpler policies can each be rejected based on the empirical data. Our approach to evaluating the relative success of these policies in accounting for the data is based on elimination. For each policy, we are looking for reasons to reject the hypothesis that subjects used it. Put differently, we cannot show inductively that a particular policy is the correct one, but we can show that a particular policy is an incorrect one. This problem in evaluating models of decision making in dynamic decision problems was discussed in Bearden and Rapoport (2005).

We can estimate *decision policies* by estimating *thresholds* from the decision data. To do this, we find for each subject the set of thresholds that maximize the *proportion of correctly predicted decisions*. The average estimated thresholds for each experiment are shown in Figure 1. These thresholds (solid lines) are based on aggregating all 50 experimental trials. Because we observed no shift in average earnings over the course of Experiments 1 and 2, we make the assumption that each subject employed the same policy over the course of the 50 trials.

Note that the curves do not span the entire range of t. This is because some (t, s) states were either never encountered (e.g., holding 5 units when $t = 10$) or encountered very infrequently (e.g., holding 1 unit when $t = 30$ in Experiment 1); so estimating thresholds for these states was either impossible or likely to be overly sensitive to error (i.e., to response variability). The subjects' data in Experiments 1 and 2 are fit very well by the threshold rule; on average, the policy predicts more than 96% of the subjects' decisions.

Based on the curves in Figure 1, it seems that the subjects did not employ any of the three nonsophisticated policies because the curves are increasing in t and decreasing in s. Only the sophisticated threshold policy simultaneously permits both of these properties. It is important to emphasize that we did not constrain $R_t^s \leq R_t^{s-1}$. Thus, the analyses were not biased in favor of the sophisticated policy. In sum, we

Table 7 Average Thresholds for Experiment 3

	All	First 20	Last 20	Optimal
Average threshold	2.42 (0.35)	2.49 (0.40)	2.37 (0.36)	2.42

Note. Standard deviations of the averages across subjects are shown in parentheses.

conclude that the sophisticated threshold policy best accounts for the decision data from the RMP and that the subjects' policies are in line with the qualitative (structural) predictions that follow from the optimal policy.

No estimation is required for the Experiment 3 policies, because subjects were required to specify their (single) threshold at the beginning of each season. Table 7 shows the mean thresholds. We compared the mean thresholds across 50 trials and those from the first and last 20 trials to the optimal threshold (2.42) using t-tests. None of the tests yielded significant differences at the $\alpha = 0.05$ level. However, the mean thresholds in the last 20 trials are significantly lower than those in the first 20 trials, $t_{35} = 2.52$, $p = 0.02$. We have already seen in Table 2 that the average revenues were significantly greater in the last 20 trials of the experiment. Thus, in sum, subjects learned to increase their average revenues with experience by lowering their thresholds and being less demanding. And, again, after sufficient experience (30 trials), we observed no difference in the mean revenues of those who were free to use dynamic thresholds (Experiment 1) and those who were forced to use a single static threshold (Experiment 3).

5. Discussion

Our results show that the subjects in Experiments 1 and 2 used sophisticated policies that were suboptimally parameterized and that they did no better than those who were forced to use simple heuristic policies in Experiment 3, once the latter had sufficient experience with the problem. It is important to note that the subjects who were free to dynamically make accept and reject decisions showed no evidence of learning, whereas those forced to use the simple heuristic did learn to increase their revenues. It seems that learning from experience is facilitated when the policy used by the subject, even when it is imposed on him or her, is simpler.

Revenue losses resulted from a clear pattern of being too demanding when holding higher levels of inventory and insufficiently demanding when holding lower levels. We term this *inventory mis-sensitivity*. The pattern of subjects' accept-reject decisions is consistent with using a threshold $\tilde{R}_t^s$ that is a convex combination of the optimal threshold R_t^s and a reference threshold $R_t^{g(S)}$:

$$\tilde{R}_t^s = \lambda R_t^s + (1-\lambda) R_t^{g(S)}, \tag{5}$$

where $0 < g(S) < S$. For instance, the results from Experiment 1 are qualitatively consistent with using $g(S) = \lfloor S/2 \rfloor$. In other words, the subjects' thresholds tended to be *regressive*: they were biased toward values that would be optimal for more intermediate inventory levels. Bearden et al. (2007) showed that subjective assessments of quantities that are bounded (e.g., probabilities) tend to be regressive; that is, small values tend to be overestimated and large values underestimated. In the RMP experiments, the subjects could directly observe their inventory levels—there was no ambiguity about how many contracts they had left to sell—but they had to decide on their thresholds, and these tended to display the regressive property consistent with Equation (5).

Generally, it would be difficult to assess the quality of RM decisions in natural environments. To determine optimal policies in these environments, one must make some strong assumptions, and whether these assumptions are (precisely) met would be difficult to ascertain. For instance, pricing models often require that the DM know the demand density function for all feasible prices. The optimal policies are not based on the DM having a "rough sense" of these functions or "good intuitions" about them; rather, these models assume that the DM *knows* the densities with precision. Clearly, conditions such as this are unlikely to be met in most scenarios actual managers face. This fact illustrates one of the reasons experimental studies are so useful. We can place financially motivated DMs in contexts in which they do have all the information that is assumed by the optimal models, which, in turn, allows us to legitimately compare decision behavior to the predictions of the appropriate optimal policies. By examining the ways laboratory behavior departs from optimality, we can establish some basis for making predictions about how DMs are likely to err in real scenarios.

Heching et al. (2002) compared the actual pricing policies of a women's apparel retailer to several model-based pricing schemes. Each of the models they examined required certain assumptions (e.g., knowledge of the demand function), which are unlikely to be perfectly met in reality, in order to derive pricing policies. Nonetheless, accepting these limitations, based on analyses of the company's historical pricing and sales data, Heching et al. concluded that the company's markdown prices were generally lower that those suggested by the models. They also concluded that the company would have increased its revenue significantly by employing smaller price markdowns earlier in the sales season rather than their actual practice of implementing steep markdowns late in the season. Our experimental results on behavior in the RMP are consistent with these empirical findings. In particular, we find that the major driver of revenue losses in the RMP was subjects' tendencies to be insufficiently demanding when they held only a small number of units, which was correlated with nearing the end of the selling season. It is as if the subjects in the experiment employed steep markdown policies and lost revenue for doing so.

Although the biases we documented are compatible with those reported in Heching et al. (2002), this alone does not establish that our results generalize to RM decision making in actual managers. As we stated earlier, one way to gain confidence in the generality of biases observed in laboratory studies is to demonstrate that those biases occur in a range of problems.

So far, we have only discussed problems for which the arrival rate for offers is determined exogenously. Quite often, the DM can affect arrival rates by adjusting selling prices. Generally, demand for a good increases when prices decrease. Dynamic pricing problems, where the DM gets to set prices and thereby influence demand, are another potentially fruitful area for experimental research. Gallego and van Ryzin (1994) have shown that the optimal pricing policy for their continuous-time dynamic pricing problem, where prices can be chosen from an interval, have two important structural properties. First, the optimal price decreases in the number of units left in inventory. Second, for any given inventory level, the optimal price decreases as the end of the selling season approaches. Bitran and Mondschein (1997) proposed a special case of the pricing RMP, in which the price at each period is constrained to be nondecreasing in time, reflecting some retailers' (e.g., clothing retailers) reluctance to increase prices for a good during a selling season. Zhao and Zheng (2000) present results on a related (continuous-time) pricing problem in which demand is time inhomogeneous. Some important questions present themselves: How well do actual DMs solve dynamic pricing problems? Do they tend to set prices too high or too low? How well are their pricing policies adapted to time-inhomogeneous demand? A number of other pricing problems that may be suitable for laboratory investigation can be found in Talluri and van Ryzin (2004).

To the extent that our experimental results have clear practical implications, two major managerial contributions stand out. First, in determining which bids to accept for perishable assets, sellers seem to use reasonably sophisticated policies that are suboptimally parameterized. Specifically, they seem to be too demanding when holding higher levels of inventory and insufficiently demanding when holding relatively lower levels. The second major managerial implication is that sellers who are constrained by one reason or another to use a single threshold (or reservation value) over the course of an entire selling season learn to do as well as those who are free to dynamically adjust their thresholds within a season, though optimal theory reveals that the latter *can* earn more. In contrast, the sellers who are free to use an unconstrained policy do not appear to learn from experience. One plausible account of this pattern is that the unconstrained problem does not provide the kind of feedback that is conducive to learning, whereas the constrained problem does. On this view, sellers in the unconstrained problem simply cannot decide easily how local adjustments to their policies (e.g., being more or less demanding in the first five periods, say) affect their total revenues. Before too much is made of these conclusions, the generality of the behavioral results must be established with a wider and richer range of parameter values (e.g., season duration, size of initial inventory, salvage value, shortage costs, etc.).

Acknowledgments
The authors gratefully acknowledge financial support by Contract F49620-03-1-0377 from the AFOSR-MURI to the Department of Systems and Industrial Engineering and the

Department of Management and Organizations at the University of Arizona.

References

Bearden, J. N., A. Rapoport. 2005. Operations research in experimental psychology. J. C. Smith, ed. *Tutorials in Operations Research: Emerging Theory, Methods, and Applications*. INFORMS, Hanover, MD, 213–236.

Bearden, J. N., R. O. Murphy, A. Rapoport. 2005. A multi-attribute extension of the secretary problem: Theory and experiments. *J. Math. Psych.* **49** 410–425.

Bearden, J. N., A. Rapoport, R. O. Murphy. 2006a. Sequential observation and selection with rank-dependent payoffs: An experimental test. *Management Sci.* **52** 1437–1449.

Bearden, J. N., A. Rapoport, R. O. Murphy. 2006b. Experimental studies of sequential selection and assignment with relative ranks. *J. Behavioral Decision Making* **19** 229–250.

Bearden, J. N., T. S. Wallsten, C. R. Fox. 2007. Contrasting stochastic and support theory accounts of subadditivity. *J. Math. Psych.* **51** 229–241.

Bellman, R. E. 1957. *Dynamic Programming*. Princeton University Press, Princeton, NJ.

Bitran, G. R., R. Caldentey. 2003. An overview of pricing models for revenue management. *Manufacturing Service Oper. Management* **5** 203–229.

Bitran, G. R., S. V. Mondschein. 1997. Periodic pricing of seasonal products in retailing. *Management Sci.* **43** 64–78.

Bröder, A. 2000. Assessing the empirical validity of the "take-the-best" heuristic as a model of human probabilistic inference. *J. Experiment. Psych.: Learn., Memory, Cognition* **26** 1332–1346.

Brumelle, S., J. McGill. 1993. Airline seat allocation with multiple nested fare classes. *Oper. Res.* **41** 127–137.

Chow, Y., H. Robbins, D. Siegmund. 1971. *Great Expectations: The Theory of Optimal Stopping*. Houghton Mifflin Co., Boston.

Cox, J. C., R. L. Oaxaca. 1989. Laboratory experiments with a finite-horizon job-search model. *J. Risk Uncertainty* **2** 301–330.

Gallego, G., G. van Ryzin. 1994. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Sci.* **40** 999–1020.

Gigerenzer, G., P. Todd, The ABC Research Group. 1999. *Simple Heuristics That Make Us Smart*. Oxford University Press, New York.

Gilbert, J., F. Mosteller. 1966. Recognizing the maximum of a sequence. *J. Amer. Statist. Assoc.* **61** 35–73.

Heching, A., G. Gallego, G. van Ryzin. 2002. Mark-down pricing: An empirical analysis of policies and revenue potential at one apparel retailer. *J. Revenue Pricing Management* **1** 139–160.

Johnson, E. J., M. Schulte-Mecklenbeck, M. Willemsen. 2008. Process models deserve process data. A comment on Brandstätter et al. (2006). *Psych. Rev.* **115** 281–290.

Johnson, E., W. Moe, P. Fader, S. Bellman, G. Lohse. 2004. On the depth and dynamics of online search behavior. *Management Sci.* **50** 299–308.

Lee, T. C., M. Hersh. 1993. A model for airline seat inventory control with multiple seat bookings. *Transportation Sci.* **27** 1252–1265.

Lorch, R. F., J. L. Myers. 1990. Regression analyses of repeated measures data in cognitive research. *J. Experiment. Psych.: Learn., Memory, Cognition* **16** 149–157.

McGill, J., G. van Ryzin. 1999. Revenue management: Research overview and prospects. *Transportation Sci.* **33** 233–256.

Newell, B. R., D. R. Shanks. 2003. Take the best or look at the rest? Factors influencing "one-reason" decision making. *J. Experiment. Psych.: Learn., Memory, Cognition* **29** 53–65.

Papastavrou, J. D., S. Rajagopalan, A. J. Kleywegt. 1996. The dynamic and stochastic knapsack problem with deadlines. *Management Sci.* **42** 155–172.

Rapoport, A., A. Tversky. 1970. Choice behaviour in an optimal stopping task. *Organ. Behav. Human Performance* **5** 105–120.

Schotter, A., Y. M. Braunstein. 1981. Economic search: An experimental study. *Econom. Inquiry* **19** 1–25.

Seale, D., A. Rapoport. 1997. Sequential decision making with relative ranks: An experimental investigation of the secretary problem. *Organ. Behav. Human Decision Processes* **69** 221–236.

Seale, D., A. Rapoport. 2000. Optimal stopping behavior with relative ranks: The secretary problem with unknown population size. *J. Behavioral Decision Making* **13** 391–411.

Simon, H. 1955. A behavioral model of rational choice. *Quart. J. Econom.* **69** 99–118.

Talluri, K. T., G. J. van Ryzin. 2004. *The Theory and Practice of Revenue Management*. Kluwer Academic Publishers, Norwell, MA.

Weatherford, L. R., S. E. Bodily. 1992. A taxonomy and research overview of perishable-asset revenue management: Yield management, overbooking, and pricing. *Oper. Res.* **40** 831–844.

Zhao, W., Y. Zheng. 2000. Optimal dynamic pricing for perishable assets with nonhomogenous demand. *Management Sci.* **46** 375–388.