2005

Dynamic pricing with real-time demand learning

Lin, Kyle Y.

þÿEuropean Journal of Operational Research, Volume 174, (2006), pp. 522

http://hdl.handle.net/10945/43128

Available online at www.sciencedirect.com

SCIENCE DIRECT®

European Journal of Operational Research 174 (2006) 522–538

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

www.elsevier.com/locate/ejor

Production, Manufacturing and Logistics

Dynamic pricing with real-time demand learning

Kyle Y. Lin *

Operations Research Department, Naval Postgraduate School, Monterey, CA 93943, USA

Received 5 January 2004; accepted 21 January 2005
Available online 25 March 2005

Abstract

In many service industries, the firm adjusts the product price dynamically by taking into account the current product inventory and the future demand distribution. Because the firm can easily monitor the product inventory, the success of dynamic pricing relies on an accurate demand forecast. In this paper, we consider a situation where the firm does not have an accurate demand forecast, but can only roughly estimate the customer arrival rate before the sale begins. As the sale moves forward, the firm uses real-time sales data to fine-tune this arrival rate estimation. We show how the firm can first use this modified arrival rate estimation to forecast the future demand distribution with better precision, and then use the new information to dynamically adjust the product price in order to maximize the expected total revenue. Numerical study shows that this strategy not only is nearly optimal, but also is robust when the true customer arrival rate is much different from the original forecast. Finally, we extend the results to four situations commonly encountered in practice: unobservable lost customers, time dependent arrival rate, batch demand, and discrete set of allowable prices.
Published by Elsevier B.V.

Keywords: Pricing; Dynamic pricing; Forecasting; Learning; Revenue management

1. Introduction

Dynamic pricing is a business strategy that adjusts the product price in a timely fashion in order to allocate the right service to the right customer at the right time. The rationale of dynamic pricing can be understood with an example of an airline company. When an airline sells seats in the same class, it offers different fares depending on time to departure and current seat inventory. The airline has the incentive to promote sale when the departure time is approaching with a lot of vacancies on hand, because each empty seat is

* Tel.: +1 831 656 2648; fax: +1 831 656 2595.
E-mail address: kylin@nps.edu

0377-2217/$ - see front matter Published by Elsevier B.V.
doi:10.1016/j.ejor.2005.01.041

worth nothing after the airplane takes off. On the other hand, the airline still wants to reserve a certain number of seats for possible last-minute travelers who are willing to pay substantially more in price. As a consequence, airfare often fluctuates in its selling horizon.

Products such as airline seats are called perishable products, which have three major characteristics: (1) the quantity is fixed and reordering is not possible; (2) there is a deadline for sale; and (3) the marginal cost of selling one more item is little, so most part of revenue goes directly to profit. Because of these characteristics, perishable products are particularly suitable for dynamic pricing. Besides being used extensively in the airline industry, dynamic pricing can also be found in other travel industries—such as hotel rooms [4], rental cars, and cruise lines [11]—to incorporate seasonal fluctuation in demand. Interested readers are referred to survey papers, such as [20] and [17], for an overview of dynamic pricing and its role in revenue management.

In general, there are two major sources of randomness in demand: customer arrival rate and customer reservation price distribution. Most existing literature concerning dynamic pricing assumes that both customer arrival rate and customer reservation price distribution are well known before the sale begins. In many service industries, however, whereas the seller can use historical data to estimate the customer reservation price to a good extent, it is rather difficult to accurately forecast the customer arrival rate before the sale begins. For example in the travel industry, the demand rates for air travel services and for hotel rooms on different days may be different if an event—such as a commencement, a trade show, or a conference—takes place at the destination city. For another example in the entertainment industry, when a pop singer goes on an international tour, it is relatively easy to know how much a loyal fan is willing to pay for a ticket, but it is rather difficult to know how many fans there are in each city and how many of those fans will be aware of the event. In these cases, if the seller roughly estimates the customer arrival rate and dynamically sets the product price based on this rough estimation, he faces a significant risk. If the true customer arrival rate is much lower than the estimated rate, the seller will end up with many unsold items at the end. On the other hand, if the true arrival rate is much higher, the seller will be out of stock quickly and loses the opportunity to take advantage of the excess demand. The dynamic pricing literature does not adequately address this risk.

In this paper, we present a dynamic pricing model where customers arrive in accordance with a conditional Poisson process, whose rate is not known to the seller in advance. Instead, through preliminary presale market research, the seller obtains a prior distribution of the customer arrival rate. As the sale moves forward, the seller uses real-time sales data from the realized demand to fine-tune the arrival rate estimation, and then uses the fine-tuned arrival rate estimation to better understand the demand curve in the future. Consequently, the seller updates the future demand distribution in real time, and then dynamically sets the product price to maximize the expected total revenue.

In recent years, the problem of dynamic pricing has drawn much attention. Most research on dynamic pricing assumes that the customers arrive according to a stochastic process that has independent increments; that is, the numbers of customers in disjoint time intervals are independent random variables. With this assumption, knowing the number of customers that have shown up so far provides no additional information about how many more customers will show up later on, so learning is not possible. For example, in a continuous-time setting, a common assumption is that customers arrive according to a Poisson process with a given intensity function [6,8,9]. In a discrete time setting, time is divided into small intervals such that in each time interval there is a small probability a customer will arrive, independent of everything else [13,19]. With the assumption that the demand process has independent increments, the problem is often formulated as a Markov decision process. In most cases, it can be shown that the optimal product price increases in the remaining time and decreases in the current inventory level. However, because the optimal policy is difficult to derive, most research focuses on developing heuristic policies.

Learning models have been studied in the operations management literature to better forecast the future demand curve. Most work assumes that price is exogenous, while the firm decides how much inventory to replenish in each time period [2,12,16]. Learning models that incorporate both price and replenishment decisions include [3] and [18]. In both papers, the demand curve in each time period is a deterministic and identical function of the price, while the parameters of the function are unknown to the decision maker. Based on the realized demand in early periods, the decision maker learns about the demand curve in order to set a proper price later on. Burnetas and Smith [5] considered a similar problem except that the demands in different periods are independent and identically distributed random variables. They developed a policy with which the realized profit converges with probability one to the optimal value under complete information. These learning models are different from our model because the seller learns from repetition of identical experiments (same flight number through different days), and in our model the seller learns throughout the sales horizon of a single event.

The rest of this paper is organized as follows. In Section 2, we introduce a dynamic pricing model where customers arrive according to a conditional Poisson process. We show how the seller can improve the estimation on the customer arrival rate from the real-time sales data as the sale moves forward. Motivated by these preliminary results, we consider a surrogate dynamic pricing model and derive its optimal policy in Section 3. Then in Section 4, we use the results from this surrogate model to develop the *variable-rate* policy for the original problem described in Section 2. The numerical experiments show that this variable-rate policy is not only nearly optimal, but also robust even when the pre-sale estimation on the customer arrival rate is relatively poor. In Section 5, we extend the results to four settings that are often encountered in practice: (1) lost customers are not observable; (2) the customer arrival process is non-stationary; (3) each customer can request more than one item; and (4) the allowable price set is discrete. Finally we conclude the paper and discuss future research directions in Section 6.

2. The model and preliminaries

Consider a dynamic pricing model where a seller sells a given stock of identical items over a finite time horizon $[0, T]$. Customers arrive according to a conditional Poisson process with an unknown rate Λ. Upon arrival, a customer will purchase one item if the posted product price is lower than her reservation price, or leaves empty-handed otherwise. We assume the reservation prices of all customers are independent and identically distributed with a continuous cumulative distribution function F. In other words, when the seller sets the product price p, the probability an arriving customer will purchase one item is $\overline{F}(p) \equiv 1 - F(p)$, $p \geqslant 0$. We assume that $\overline{F}(p)$ is strictly decreasing in p and let $\overline{F}^{-1}(v)$, $0 \leqslant v \leqslant 1$, denote its inverse function. With this assumption, we can interpret the seller's decision as choosing the probability of successfully selling one item instead of choosing the price p. In the rest of the paper, we will refer the probability of successfully selling one item as the *policy* of the seller.

Let $g(v) \equiv v\overline{F}^{-1}(v)$ denote the expected revenue collected from the arriving customer when the seller uses the policy v. We assume that $g(v)$ is continuous, twice differentiable, and bounded for $v \in [0,1]$. In particular, because $g(0) = 0$ when the seller decides not to sell to a particular customer, we let $\lim_{v \to 0^+} g(v) = 0$ so that the expected revenue tends to zero as the price approaches infinity. The sale ends either when the seller sells out the stock, or at time T, whichever occurs first. We assume that any unsold items at time T have no salvage value. This assumption is without loss of generality if the salvage value of the stock is linear in the number of unsold items.

The customer arrival rate Λ, although not known in advance, can be estimated by pre-sale market research. As soon as the sale begins, the seller starts to observe the arrival stream of customers. By counting the number of arriving customers as the sale moves forward, the seller then fine-tunes the estimator of Λ so as to better predict the demand distribution in the future. The seller's objective is to use the pre-sale market research and the real-time sales data to dynamically adjust the product price in order to maximize the expected total revenue when the sale ends.

2.1. Forecast the customer arrival rate

Before the sale begins, the seller conducts market research such as market surveys or historical data analysis, to forecast the customer arrival rate Λ. In particular, we assume that this pre-sale market research reveals the first two moments of the random quantity Λ: the estimator μ and the standard error σ. The seller can endure more effort in market research to get a better estimation on Λ, that is, a smaller value in σ. The extreme case that $\sigma = 0$ reduces the model to the one where the seller can accurately predict the customer arrival rate—a common assumption used in most existing literature.

With estimations for the first two moments of the arrival rate Λ, a natural choice is to assume Λ follows a normal distribution with the estimated mean and variance. A normal distribution prior, however, has at least two drawbacks: (1) the normal distribution has positive support for the whole real line, while the customer arrival rate can never be negative; and (2) as the sale moves forward, the resulting posterior distribution of Λ does not have a closed form and is mathematically intractable. For these reasons, choosing a normal distribution prior would add unnecessary complexity to the problem.

In this paper, we assume the customer arrival rate Λ follows a gamma distribution with parameters (k, a), with the following density function,

$$f_\Lambda(\lambda) = \frac{a\mathrm{e}^{-a\lambda}(a\lambda)^{k-1}}{\Gamma(k)},$$

where $\Gamma(k) = \int_0^\infty \mathrm{e}^{-x}x^{k-1}\mathrm{d}x$ is the gamma function. To fit the gamma distribution with the estimated mean and variance, we choose the parameters k and a such that

$$\mu = E[\Lambda] = \frac{k}{a} \quad \text{and} \quad \sigma^2 = \mathrm{Var}(\Lambda) = \frac{k}{a^2},$$

or equivalently, $k = \mu^2/\sigma^2$ and $a = \mu/\sigma^2$. The gamma distribution does not have the drawbacks of the normal distribution aforementioned, while its bell shape is close to normal distribution to a good extent. This is particularly true when k is large (a gamma distribution converges to a normal distribution as k tends to infinity), which is the case if the pre-sale market research is reasonably effective ($\mu \gg \sigma$). The most significant advantage of choosing a gamma distribution is that the demand distribution becomes negative binomial, which we shall discuss next.

2.2. Compute the demand distribution

To calculate the probability distribution of the total number of customers N, we compute $P(N = n)$ by conditioning on the value of Λ:

$$\begin{aligned} P(N = n) &= \int_0^\infty P(N = n|\Lambda = \lambda)f_\Lambda(\lambda)\,\mathrm{d}\lambda = \int_0^\infty \mathrm{e}^{-\lambda T}\frac{(\lambda T)^n}{n!}\frac{a\mathrm{e}^{-a\lambda}(a\lambda)^{k-1}}{\Gamma(k)}\,\mathrm{d}\lambda \\ &= \frac{\Gamma(n+k)}{n!\,\Gamma(k)}\left(\frac{a}{a+T}\right)^k\left(\frac{T}{a+T}\right)^n. \end{aligned}$$

If k is an integer, the preceding becomes

$$P(N = n) = \binom{n+k-1}{n}\left(\frac{a}{a+T}\right)^k\left(\frac{T}{a+T}\right)^n,$$

which is a negative binomial distribution with parameters $(k, a/(a + T))$. We will restrict the choice of k to be an integer because it improves mathematical tractability significantly. This restriction, however, does not hinder the rich shape the gamma distribution possesses.

In most revenue management literature, the arrival process of customers is assumed to be a Poisson process. For both homogeneous and nonhomogeneous Poisson processes, the resulting distribution of total demand follows a Poisson distribution whose coefficient of variation is the reciprocal of the square root of the mean. This coefficient of variation is much lower than what is encountered in practice [17]. In our model, the total demand follows a negative binomial distribution, whose coefficient of variation is

$$\frac{\sqrt{\operatorname{Var}(N)}}{E[N]} = \frac{\sqrt{\frac{T(a+T)}{a^2}k}}{\frac{T}{a}k} = \frac{1}{\sqrt{E[N]}}\sqrt{\frac{a+T}{a}} > \frac{1}{\sqrt{E[N]}}.$$

In other words, the coefficient of variation of the negative binomial demand distribution is larger than that of the Poisson distribution with the same expected value. For that reason, the negative binomial demand distribution of our model helps explain what is observed in practice.

There are extensive studies on the negative binomial demand distribution in marketing science literature [14]. In particular, Agrawal and Smith [1] found evidence in a major retailing chain that the negative binomial demand fits significantly better than the Poisson or the normal distribution.

2.3. Update the demand distribution in real time

In this section, we first show how to fine-tune the customer arrival rate estimation in real time, and then use the modified arrival rate estimation to forecast the future demand distribution.

Let $\{N(t), 0 \leqslant t \leqslant T\}$ denote the counting process of customer arrivals, where $N(t)$ represents the number of customers that have shown up through time t. After observing the value of $N(t)$, the seller would have a better idea about the true customer arrival rate. That is, if $N(t)$ is relatively large, it is more likely the arrival rate Λ is also large. Specifically, suppose i customers have shown up and at time t, the $(i+1)$st customer arrives. Because the density of this event is equal to the density that the arrival time of the $(i+1)$st customer—which follows a gamma distribution with parameters $(i+1, \lambda)$—is exactly equal to t, we can compute the posterior probability density function of Λ as follows:

$$f_{\Lambda|N(t^-)=i,\ \text{arrival at time}\ t}(\lambda) = \frac{f_\Lambda(\lambda)\ \lambda \mathrm{e}^{-\lambda t}\frac{(\lambda t)^i}{i!}}{\int_0^\infty f_\Lambda(\lambda)\ \lambda \mathrm{e}^{-\lambda t}\frac{(\lambda t)^i}{i!}\ \mathrm{d}\lambda} = \frac{(a+t)\mathrm{e}^{-(a+t)\lambda}((a+t)\lambda)^{k+i}}{(k+i)!},$$

which follows a gamma distribution with parameters $(k+i+1, a+t)$. The seller can then use this density function to calculate the distribution of the number of future customers that will show up in the interval $(t, T]$. That is,

$$\begin{aligned} &P(N(T)-N(t)=n \mid N(t^-)=i,\ \text{arrival at time}\ t) \\ &= \int_0^\infty \frac{\mathrm{e}^{-(T-t)\lambda}((T-t)\lambda)^n}{n!}\frac{(a+t)\mathrm{e}^{-(a+t)\lambda}((a+t)\lambda)^{k+i}}{(k+i)!}\ \mathrm{d}\lambda = \binom{n+k+i}{n}\left(\frac{a+t}{a+T}\right)^{k+i+1}\left(\frac{T-t}{a+T}\right)^n. \end{aligned} \tag{1}$$

In other words, when the seller needs to set the product price for the $(i+1)$st customer at time t, he knows the number of future customers follows a negative binomial distribution with parameters $(k+i+1, (a+t)/(a+T))$. This key result motivates a surrogate model that we will introduce in Section 3.

One special case is that the seller sets $\sigma = 0$ by assuming the pre-sale estimation on the customer arrival rate is perfect. By substituting $a = k/\mu$ into Eq. (1) and letting $k \to \infty$ (this is equivalent to letting $\sigma \to 0$), we can see that $N(T) - N(t)$ converges to a Poisson distribution with mean equal to $\mu(T-t)$. This special case reduces the model to the one in which the seller assumes the customer arrival process (Poisson process) has independent increments. In other words, the seller assumes the customer arrival rate is μ and does not change this assumption no matter what he observes during the sales horizon.

The assumption of a perfect forecast on the customer arrival rate is commonly used in the literature. When the customer's reservation price is exponentially distributed, the optimal policy can be derived in closed forms [10]. We refer this closed form optimal policy as the *fixed-rate* (FR) policy. In particular, we call it the FR(μ) policy if the seller uses the policy that would be optimal if the customer arrival rate were equal to μ. Although the FR(μ) policy can be easily carried out, its performance can be poor if the true customer arrival rate is different from μ. We will further discuss this issue in Section 4.

3. A surrogate model

It is difficult, if not impossible, to derive the optimal policy for the continuous-time dynamic pricing problem in the previous section, because the distribution of the customer arrival rate depends on both time elapsed and the number of customers that have shown up. In Section 3.1, we consider a surrogate dynamic pricing problem, which is motivated by the observation that whenever the seller needs to set the product price for an arriving customer, the number of future customers follows a negative binomial distribution. We then solve the surrogate dynamic pricing problem in Section 3.2.

3.1. The seller's new problem

Suppose when a new customer arrives, the seller needs to set the product price while knowing the number of future customers follows a negative binomial distribution with parameters (c, α). Instead of assuming some functional forms of the customers' arrival times in a continuous-time setting, we assume that the future customers will show up one immediately after another. The sale ends either when the entire stock is sold out, or when there are no more customers. The seller's objective is to set a potentially different price upon arrival of each new customer in order to maximize the expected total revenue when the sale ends.

Let X denote a negative binomial random variable with parameters (c, α) and that

$$P(X = i) = \binom{i + c - 1}{i} \alpha^c (1 - \alpha)^i, \qquad i = 0, 1, 2, \ldots$$

The random variable X can be interpreted as the number of times the tail appeared before the cth time the head appears, when a biased coin with head probability α is continually flipped. With this interpretation, X can be written as

$$X = \sum_{j=1}^{c} Y_j,$$

where Y_j, $j = 1, \ldots, c$, are independent geometric random variables with parameter α, and that

$$P(Y_j = i) = (1 - \alpha)^i \alpha, \qquad i = 0, 1, 2, \ldots$$

Note that we define the geometric distribution by counting the number of *tails* until the first head appears, whereas a more conventional definition is to count the number of *flips*.

For each realization of X, we can divide the customers into c distinct groups, with group j having Y_j customers, $j = 1, \ldots, c$. One can imagine that, after Y_1 customers show up, a *dummy* customer steps in before the second group of Y_2 customers, which is then followed by another dummy customer, and so on. In other words, these dummy customers between successive customer groups signal the beginning of a new group. Of course these dummy customers are invisible to the seller, and the seller cannot tell which group each *real* customer belongs to.

3.2. A supernatural sales agent

Consider a supernatural sales agent who has the magical ability to see dummy customers. In other words, this supernatural sales agent is informed which group each customer belongs to. Because this supernatural sales agent has more information than the normal seller, it follows that the optimal expected total revenue generated by the supernatural sales agent is higher than the normal seller.

Let $J(s,c,\alpha)$ denote the optimal expected total revenue the supernatural sales agent can generate with s items, if the total number of customers follows a negative binomial distribution with parameters (c,α). Because the supernatural sales agent can see both real and dummy customers, she can formulate the optimality equation by conditioning on whether the next customer is real. Specifically, with probability α, the next customer is dummy and the resulting expected total revenue becomes $J(s,c-1,\alpha)$, and with probability $1-\alpha$, the next customer is real and the supernatural sales agent needs to choose v to maximize the expected total revenue knowing that the additional number of customers is still a negative binomial random variable with parameters (c,α) because of the memoryless property of the geometric distribution. Consequently,

$$\begin{aligned} J(s,c,\alpha) &= \alpha J(s,c-1,\alpha) + (1-\alpha)\max_{v}\left\{v\left(\overline{F}^{-1}(v) + J(s-1,c,\alpha)\right) + (1-v)J(s,c,\alpha)\right\} \\ &= \alpha J(s,c-1,\alpha) + (1-\alpha)J(s,c,\alpha) + (1-\alpha)\max_{v}\{g(v) - v(J(s,c,\alpha) - J(s-1,c,\alpha))\}. \end{aligned} \tag{2}$$

The optimization problem in the preceding equation is to find v that maximizes

$$g(v) - v(J(s,c,\alpha) - J(s-1,c,\alpha)). \tag{3}$$

Let $v^*(s,c,\alpha)$ denote the least maximizer of the preceding, that is,

$$v^*(s,c,\alpha) \equiv \min\arg\max_{v}\ g(v) - v(J(s,c,\alpha) - J(s-1,c,\alpha)). \tag{4}$$

Because the supernatural sales agent needs to set the product price only if the next customer is real, $v^*(s,c,\alpha)$ is the optimal policy for the arriving (real) customer if the agent still has s items and the additional number of customers (besides the arriving one) follows a negative binomial distribution with parameters (c,α).

To maximize Eq. (3), first note that it is twice differentiable because $g(v)$ is twice differentiable. Therefore, the optimal policy $v^*(s,c,\alpha)$ satisfies the first-order condition of optimality:

$$g'(v^*(s,c,\alpha)) = J(s,c,\alpha) - J(s-1,c,\alpha). \tag{5}$$

Substituting $J(s,c,\alpha) = g'(v^*(s,c,\alpha)) + J(s-1,c,\alpha)$ back into Eq. (2) shows that $v^*(s,c,\alpha)$ satisfies v in

$$\alpha(g'(v) + J(s-1,c,\alpha)) = \alpha J(s,c-1,\alpha) + (1-\alpha)(g(v) - vg'(v)). \tag{6}$$

For a given value of α, we can use the preceding equation to compute $v^*(s,c,\alpha)$ if we know $J(s-1,c,\alpha)$ and $J(s,c-1,\alpha)$. We can then use Eq. (5) to find the value of $J(s,c,\alpha)$. In other words, we can compute $v^*(s,c,\alpha)$ and $J(s,c,\alpha)$ recursively with the boundary conditions

$$J(s,0,\alpha) = J(0,c,\alpha) = 0, \qquad c \geqslant 0,\ s \geqslant 0.$$

To solve for $v^*(s,c,\alpha)$ from Eq. (6), we need to find all roots in the region where $g''(v) < 0$. A simple search algorithm is efficient because when $g''(v) < 0$, the left-hand side of Eq. (6) decreases in v and the right-hand side increases in v.

The monotonic properties for the optimal revenue function J can be understood as follows. Because with more items or (stochastically) more customers, the agent can generate at least as much expected revenue by using the same policy, it follows that $J(s,c,\alpha)$ increases in s, in c, and decreases in α. We next present three second-order properties for $J(s,c,\alpha)$ in Lemma 3.1, which leads to the monotonic properties for the optimal policy $v^*(s,c,\alpha)$ in Theorem 3.1. The proofs can be found in Appendix A.

Lemma 3.1. *The optimal revenue function* $J(s,c,\alpha)$ *is concave in* s, *supermodular in* (s,c), *and submodular in* (s,α).

Theorem 3.1. *The optimal policy* $v^*(s,c,\alpha)$ *increases in* s, *decreases in* c, *and increases in* α.

Because the optimal price $\overline{F}^{-1}(v^*(s,c,\alpha))$ bears an inverse relationship with $v^*(s,c,\alpha)$, the supernatural sales agent should increase the price when c increases, and should decrease the price when s increases or when α increases. Generally speaking, these properties state that the price should be set higher with a larger demand or with a smaller supply, which are consistent with the results in other dynamic pricing models; for example, see Gallego and van Ryzin [9] and Lin [15].

Now consider the following scenario. A seller with s items hires a supernatural sales agent to dynamically set the product price. The supernatural sales agent, however, has not yet reported to work when the first customer arrives. The seller knows that the additional number of customer follows a negative binomial distribution with parameters (c,α), and that the supernatural sales agent will arrive momentarily so he can delegate the agent to set the product price for all future customers. However, because the first customer does not want to wait, the seller has to set the price for this first customer. What should the seller do? The answer is that the optimal policy of the seller is to use the policy $v^*(s,c,\alpha)$ for this first customer. To see why this is the case, first note that the seller cannot do better than the supernatural sales agent because the latter has more information. By using the policy $v^*(s,c,\alpha)$, however, the seller can do equally well to the supernatural sales agent if the latter were to set the price for the first customer. Therefore, the policy $v^*(s,c,\alpha)$—which is a feasible policy for the seller—is also the seller's optimal policy. The policy $v^*(s,c,\alpha)$, however, would not be optimal if the seller had to set the price for all future customers.

This observation motivates a heuristic policy for the real-time demand learning model we introduced in Section 2. We shall discuss this heuristic policy in the next section.

4. Dynamic pricing with real-time demand learning

In this section, we return to the model in Section 2, where the seller updates the demand distribution in real time. In Section 4.1, we present a dynamic pricing policy such that the seller sets the product price based on the updated demand distribution. We then present numerical examples to demonstrate this policy's efficiency in Section 4.2 and its robustness in Section 4.3.

4.1. The variable-rate policy

Consider the model in Section 2, where the customers arrive according to a conditional Poisson process with rate Λ that follows a gamma distribution with parameters (k,a). Recall that in Eq. (1), if the $(i+1)$st customer arrives at time t—which implies that i customers have shown up in the interval $[0,t)$—the number of future customers follows a negative binomial distribution with parameters $(k+i+1,(a+t)/(a+T))$. If the current inventory level is s, then we propose to use the policy

$$v^*\left(s,\ k+i+1,\ \frac{a+t}{a+T}\right),$$

which would be optimal if future customers were to arrive sequentially. In other words, we use the policy that would be optimal when considering only the distribution of the number of future customers but not the functional forms of their arrival times. We call this dynamic pricing policy the *variable-rate* (VR) policy because the seller fine-tunes the estimation of the customer arrival rate in real time so that he can set the product price properly. In particular, we call it the VR(μ,σ) policy if the prior distribution of the customer arrival rate has mean μ and standard deviation σ.

The rationale of the VR policy can be understood as follows. Imagine each arriving customer triggers a new puzzle that requires the seller to dynamically set the product price when facing a negative binomial demand distribution. The seller, however, has to solve only the first step of the puzzle by setting the product price for the arriving customer, but can delegate an imaginary supernatural sales agent (as described in Section 3.2) to solve the rest of the puzzle. In other words, the seller's job is simply to solve the first step of each new puzzle upon arrival of each new customer. Because the seller can compute the function $v^*(\cdot,\cdot,\cdot)$, which is the optimal policy the supernatural sales agent would use if she were to solve that first step, the function $v^*(\cdot,\cdot,\cdot)$ is also what the seller should use.

To implement the VR policy, the seller needs to post a price that changes in real time in anticipation of a new customer. Specifically, if the current inventory level is s and i customers have shown up, the price at time t is

$$\overline{F}^{-1}\left(v^*\left(s,\ k+i+1,\ \frac{a+t}{a+T}\right)\right).$$

Because k, a, T are constants, during the sale the price is a function of s (current inventory level), i (number of customers arrived so far), and t (amount of time elapsed). Intuitively, the optimal price should decrease in s because with a lot of items the seller should lower the price to promote sales to avoid unwanted inventory at the end; the optimal price should increase in i because seeing more customers implies a higher customer arrival rate and therefore a higher demand in the future. The optimal price should also decrease in t for two reasons: (1) seeing the same number of customers in a longer time period implies the customer arrival rate is likely to be lower; and (2) with less time to go the demand tends to be lower. The results in Theorem 3.1 are consistent with these conjectures.

Note that computing $v^*(s, k+i+1, (a+t)/(a+T))$ requires solving a set of recursive equations, and the computational effort increases as s increases, or as $k+i+1$ increases. For that reason, computing $v^*(s,\ k+i+1, (a+t)/(a+T))$ in real time may be impractical if the seller initially has a large inventory, or if a lot of customers have shown up. A remedy for this computational problem is to construct a table of $v^*(s, c, \alpha)$ with three dimensions: $s = 1, 2, \ldots$; $c = 1, 2, \ldots$; and different values of α between 0 and 1. The seller can then quote the policy from this table and use linear interpolation on α when necessary.

4.2. Numerical experiments

To evaluate the performance of the VR policy, we consider an example where the initial inventory level $s = 10$, and the total time for sale $T = 5$. Recall in Section 2.3 that the FR policy has a closed form solution when the customer's reservation price is exponentially distributed [10]. For that reason, we assume an exponential reservation price distribution with mean equal to 1 in the numerical example, so that we can compare the FR policy with the VR policy. Nevertheless, it is worth noting that the VR policy also works for general reservation price distributions.

In the numerical study, we assume the arrival rate Λ follows a gamma distribution with mean μ and standard deviation σ, and simulate the expected total revenue generated by the FR policies and the VR policies with different parameters. In each simulation run, we first generate the customer arrival rate from the gamma distribution, and then generate a sample path of customer arrivals and their respective reservation prices. We collect the data of total revenues for each policy to obtain the estimator of the expected total revenue, and also the standard error of the estimator.

In order to assess the performance of each policy, we derive an upper bound for the optimal expected total revenue. Consider the following scenario that the seller has perfect foresight to accurately predict the value of Λ before the sale begins. Once the value of the true customer arrival rate is revealed, say λ, the seller then uses the optimal policy, that is, the FR(λ) policy. The expected total revenue produced in this scenario is then an upper bound for the optimal expected total revenue of the original problem. It is

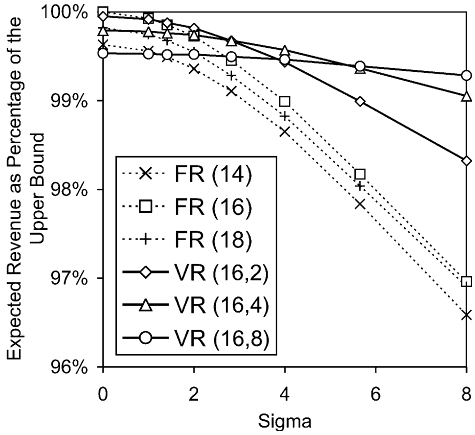

Fig. 1. Efficiency of the VR policies.

quite straightforward to simulate this upper bound because once we generate the arrival rate λ, the optimal expected total revenue has a closed form solution.

The simulation results are plotted in Fig. 1. We choose $\mu = 16$ and plot the performance of each policy for different values of σ. The performance of each policy is reported as the ratio of its expected total revenue to the upper bound aforementioned. We simulate 10 million runs for each scenario so that the 99.75% confidence interval covers at most $(100 \pm 0.03)\%$ of the estimator in all scenarios.

As shown in Fig. 1, the performances of the FR policy drops quickly if the uncertainty in the customer arrival rate estimation, namely σ, increases. In addition, the performance of the FR policy also drops if the customer arrival rate assumed by the FR policy differs from $E[\Lambda]$. On the other hand, the VR policies are very effective. The performance of the VR(16,2) policy drops slowly as σ increases, while that of the VR(16,8) policy almost stays flat across different values of σ. The self-learning mechanism of the VR policy allows the seller to fine-tune the customer arrival rate estimation as the sale moves forward. In doing so, the seller is able to adapt the prices to an unexpected demand curve.

Although the difference in revenues between FR policies and VR policies is in the order of 1%, the increase in profit can be significant in many service industries because the profit margin is usually very thin. For example, Feldman [7] reports that the average profit margin for the airline industry is 1.6% during 1978 to 1988. For this industry, an increase of 1% in revenue is equivalent to 60% increase in profit.

4.3. Robustness and sensitivity analysis

This section discusses the robustness of the VR policy against different customer arrival rates. Specifically, we simulate the expected total revenue generated by various FR policies and VR policies for a given customer arrival rate, say λ. After 10 million simulation runs, we calculate the performance of these policies as the ratio of their respective expected total revenue to that generated by the FR(λ) policy. Finally we change the value of λ and plot the results in Fig. 2.

As shown in Fig. 2, although the FR(μ) policy does well if λ is close to μ, its performance drops quickly as $|\lambda - \mu|$ increases. On the other hand, the VR(μ, σ) policy gives up a little revenue in the case when λ is close to μ, in exchange for robustness when λ is away from μ. The larger the value of σ, the more significant this tradeoff becomes. For example, the VR(16,8) policy yields almost the same expected revenue for a wide range of customer arrival rates. Finally, it is very encouraging that by employing a learning mechanism, the

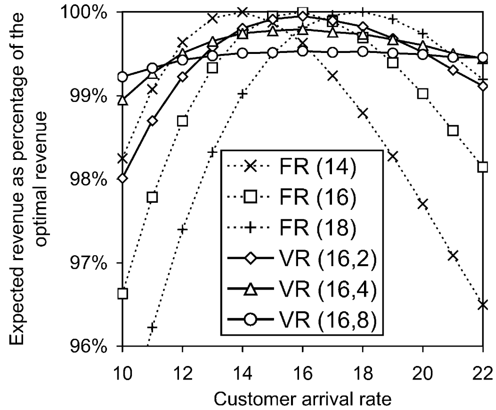

Fig. 2. Robustness of the VR policies.

VR(16,2) policy outperforms the FR(16) policy with a noticeable margin in most cases while giving up very little only when the customer arrival rate $\lambda \approx 16$.

5. Extensions

We next consider several extensions to the basic model. The first extension deals with the situation when the seller can track only the number of items sold but not the number of customers. The second extension considers the case when the customer arrival rate is time dependent. For these two extensions, we modify the VR policy so that the seller can quote the price from the same three-dimensional table discussed in Section 4.1. The third extension allows each customer to buy multiple items, and the last extension restricts the seller to choose the price from a given discrete set. For these two extensions, the seller needs to first construct a new price table before he can quote the price properly.

5.1. Unobservable lost customers

In previous sections, we assume that the seller can track the number of customers who inquire the product even though the inquiry does not eventually lead to a purchase. This is a reasonable assumption for online retailers, but it may not be the case for traditional retail stores. In this section, we consider the situation when the seller can only observe the number of items sold, but not the number of customers who purchased nothing.

Let $\{N_S(t), t \geqslant 0\}$ denote the counting process where $N_S(t)$ is the number of items sold up to time t. In addition, let $\{p(s), 0 \leqslant s < t\}$ denote the price history up to time t^-, where $p(s)$ is the product price at time s. At time t, the seller's objective is to use $\{N_S(s), 0 \leqslant s < t\}$ and $\{p(s), 0 \leqslant s < t\}$ to properly set the product price in order to maximize the expected total revenue from time t and onward.

We first consider the relation between $N_S(t)$ and $\{p(s), 0 \leqslant s < t\}$. Because a customer arriving at time s will independently buy one item with probability $\overline{F}(p(s))$, the counting process $\{N_S(t), t \geqslant 0\}$ is a sampling of the original customer arrival process. Conditional on $\Lambda = \lambda$, $\{N_S(t), t \geqslant 0\}$ is a nonhomogeneous Poisson process with intensity function $\lambda\overline{F}(p(s))$. Given $\Lambda = \lambda$ and the price history $\{p(s), 0 \leqslant s < t\}$, the number of items sold up to time t follows a Poisson distribution with mean equal to

$$m(t) \equiv \lambda \int_0^t \overline{F}(p(s))\mathrm{d}s = \lambda\Theta(t),$$

where

$$\Theta(t) \equiv \int_0^t \overline{F}(p(s))\mathrm{d}s \tag{7}$$

can be viewed as the effective time length during which customer arrivals are observable. If j items have been sold and a customer arrives at time t, the posterior distribution of Λ becomes

$$f_{\Lambda|\{p(s),0\leqslant s<t\},N_S(t^-)=j,\text{ arrival at time }t}(\lambda) = \frac{f_\Lambda(\lambda)\ \lambda \mathrm{e}^{-m(t)} \frac{(m(t))^j}{j!}}{\int_0^\infty f_\Lambda(\lambda)\ \lambda \mathrm{e}^{-m(t)} \frac{(m(t))^j}{j!}\ \mathrm{d}\lambda} = \frac{(a+\Theta(t))\ \mathrm{e}^{-(a+\Theta(t))\lambda}\ ((a+\Theta(t))\lambda)^{k+j}}{(k+j)!}.$$

We can then calculate the probability mass function of the future number of customers in $(t, T]$:

$$\begin{aligned}
&P(N(T)-N(t)=n|\{p(s),0\leqslant s<t\},N(t^-)=j,\text{ arrival at time }t)\\
&=\int_0^\infty \frac{\mathrm{e}^{-(T-t)\lambda}((T-t)\lambda)^n}{n!}\frac{(a+\Theta(t))\mathrm{e}^{-(a+\Theta(t))\lambda}((a+\Theta(t))\lambda)^{k+j}}{(k+j)!}\ \mathrm{d}\lambda\\
&=\binom{n+k+j}{n}\left(\frac{a+\Theta(t)}{a+\Theta(t)+T-t}\right)^{k+j+1}\left(\frac{T-t}{a+\Theta(t)+T-t}\right)^n.
\end{aligned}$$

In other words, if a new customer arrives at time t and j items have been sold, the number of future customers in $(t, T]$ follows a negative binomial distribution with parameters $(k+j+1, (a+\Theta(t))/(a+\Theta(t)+T-t))$. Therefore, if the current inventory level is s, the seller should use the policy

$$v^*\left(s,\ k+j+1,\ \frac{a+\Theta(t)}{a+\Theta(t)+T-t}\right).$$

It is worth noting that $\Theta(t)$ can be viewed as the effective time length during which the arriving customers are observable in the time period $[0, t)$. It is seen from Eq. (7) that $\Theta(t) \leqslant t$, where the equality holds only if all customers are observable—the case in Section 2. In general, one can imagine that the price history $\{p(s), 0 \leqslant s < t\}$ screens the customer arrival process so that the seller can only see those customers with higher reservation prices. The higher the price history $\{p(s), 0 \leqslant s < t\}$ has been, the smaller the value $\Theta(t)$ becomes, and therefore less information is collected because more customers have turned away from the high price.

5.2. Time dependent arrival rate

In many industries, it is often desirable to model the customer arrival process as a nonhomogeneous Poisson process to capture different demand curves over time. For example, fashion retail stores may see more customers on weekends than weekdays. Similarly, a sports event may see a surge of online customers in the first few hours when the tickets go on sale. Although it is difficult to predict the strength of the market response before the sale begins, it is often easier to understand the relative shape of the demand curve over time from the historical data of similar events. Motivated by this observation, we suppose the arrival rate function $\Lambda(t)$ can be represented by

$$\Lambda(t) = R\,\lambda(t), \qquad 0 \leqslant t \leqslant T,$$

where $\lambda(t)$ is a constant baseline arrival rate that exhibits the relative shape of the demand curve over time, while R is a gamma random variable with parameters (k, a) that measures the strength of market response.

Although we allow the customer arrival rate to be time dependent, this model is mathematically equivalent to the one with a constant arrival rate because we can transform a nonhomogeneous Poisson process into a homogeneous Poisson process by properly scaling the time horizon. Similar to the derivation in Section 2, it is quite straightforward to show that if the $(i+1)$st customer arrives at time t, the number of future customers follows a negative binomial distribution with parameters

$$\left(k+i+1, \frac{a+\int_0^t \lambda(s)\mathrm{d}s}{a+\int_0^T \lambda(s)\mathrm{d}s}\right).$$

Consequently, we can directly generalize the VR policy by using the same three dimensional table of $v^*(\cdot,\cdot,\cdot)$ defined in Section 4.1.

5.3. Batch demand

In many service industries, a customer often would like to purchase two or more units of the same product. Examples include airline seats, hotel rooms, and concert/sport event tickets. When inquiring the price for four airline seats in planning a family vacation, a customer typically will either buy all four tickets, or buy nothing. In other words, the seller loses the customer if he cannot or is not willing to fulfill the entire order. In this section, we consider the same model as in Section 2 except that each customer will independently request j items with probability q_j, $j = 0,1,2,\ldots$, where $\sum_{j=0}^{\infty} q_j = 1$. When requesting j items, the customer either will buy all j items if the unit price is less than her reservation price, or will leave empty-handed otherwise. Although the formulation does not explicitly assume a maximum number of units a customer is allowed to buy, it can accommodate such a restriction by setting $q_j = 0$ for $j > m$ if each customer can buy up to m units, and letting q_0 capture the probability a customer would like to buy more than m units.

To extend the model to allow batch demand, first note that the number of future customers upon a new arrival still follows a negative binomial distribution as in the basic model. Therefore, we consider the same surrogate model as in Section 3, except that each customer now requests j items with probability q_j. We replace the optimality equation in Eq. (2) by

$$\begin{aligned}
J_{\mathrm{B}}(s,c,\alpha) &= \alpha J_{\mathrm{B}}(s,c-1,\alpha) + (1-\alpha)\max_v \left\{ \sum_{j=1}^{s} q_j \Big(v(j\overline{F}^{-1}(v) + J_{\mathrm{B}}(s-j,c,s)) \right. \\
&\qquad \left. + (1-v) J_{\mathrm{B}}(s,c,\alpha)\Big) + \left(1-\sum_{j=1}^{s} q_j\right) J_{\mathrm{B}}(s,c,\alpha) \right\} \\
&= \alpha J_{\mathrm{B}}(s,c-1,\alpha) + (1-\alpha) J_{\mathrm{B}}(s,c,\alpha) \\
&\qquad + (1-\alpha)\max_v \left\{ \sum_{j=1}^{s} q_j (jg(v) - v(J_{\mathrm{B}}(s,c,\alpha) - J_{\mathrm{B}}(s-j,c,\alpha))) \right\},
\end{aligned}$$

where the subscript B denotes that $J_{\mathrm{B}}(\cdot,\cdot,\cdot)$ is the optimal value function in this batch demand extension. Similarly, let $v^*_{\mathrm{B}}(s,c,\alpha)$ denote the least maximizer of the preceding, that is,

$$v^*_{\mathrm{B}}(s,c,\alpha) = \min\arg\max_v \sum_{j=1}^{s} q_j (jg(v) - v(J_{\mathrm{B}}(s,c,\alpha) - J_{\mathrm{B}}(s-j,c,\alpha))).$$

We can use the boundary conditions $J_{\mathrm{B}}(s, 0, \alpha) = J_{\mathrm{B}}(0, c, \alpha) = 0$ to solve for $v^*_{\mathrm{B}}(s,c,\alpha)$ and $J_{\mathrm{B}}(s,c,\alpha)$ recursively. The algorithm is similar to that in Section 3.2, and the computation is quite straightforward. Consequently, we can generalize the VR policy and let the seller set the product price equal to

$$\overline{F}^{-1}\left(v^*_{\mathrm{B}}\left(s,\ k+i+1,\ \frac{a+t}{a+T}\right)\right)$$

at time t if i customers have shown up and the current inventory level is s.

5.4. Discrete set of allowable prices

The assumption that the seller can choose any price from a continuous set possesses some mathematical elegance, but may not be practical in some cases. In many retail industries, the seller has incentive to restrict the price to be chosen from a discrete price set—such as \$99, \$129, \$149, etc.—for various marketing reasons. In this section, we extend the VR policy to incorporate this situation.

Let $p_1, p_2, \ldots, p_m$ denote the m distinct prices from which the seller can choose, and v_i the probability a customer's reservation price is higher than p_i, $i = 1, \ldots, m$. We first consider the same surrogate model as in Section 3 except that the seller can choose the price only from the discrete set $\{p_1, \ldots, p_m\}$. We replace the optimality equation in Eq. (2) by

$$J_D(s,c,\alpha) = \alpha J_D(s,c-1,\alpha) + (1-\alpha)J_D(s,c,\alpha) + (1-\alpha)\max_{i\in\{1,\ldots,m\}}\{v_i p_i - v_i(J_D(s,c,\alpha) - J_D(s-1,c,\alpha))\}, \tag{8}$$

where the subscript D denotes that $J_D(\cdot,\cdot,\cdot)$ is the optimal value function in this discrete price set extension.

To solve for $J_D(s,c,\alpha)$ in Eq. (8), we first assume the values of $J_D(s-1,c,\alpha)$ and $J_D(s,c-1,\alpha)$ are given. If the seller chooses the policy v_i in Eq. (8), the resulting expected total revenue, denoted by r_i, is equal to

$$r_i = \frac{\alpha J_D(s,c-1,\alpha) + (1-\alpha)v_i p_i + (1-\alpha)v_i J_D(s-1,c,\alpha)}{\alpha + (1-\alpha)v_i},$$

where the preceding follows by substituting $J_D(s,c,\alpha) = r_i$ into Eq. (8). Hence, the maximized expected total revenue

$$J_D(s,c,\alpha) = \max\{r_1, \ldots, r_m\},$$

while the optimal policy $v_D^*(s,c,\alpha)$ is the policy v_i such that $r_i = \max\{r_1,\ldots,r_m\}$. With the boundary conditions $J_D(s,0,\alpha) = J_D(0,c,\alpha) = 0$, the preceding allows us to recursively compute $J_D(s,c,\alpha)$ and $v_D^*(s,c,\alpha)$ for all c and s. Consequently, the idea of the VR policy can be directly generalized.

Remark 5.1. Although we consider each of these four extensions individually, it is worth noting that we can generalize our results to any combination from these four extensions. For example, if the seller can only observe the number of customers who made a purchase (extension 1), and each customer can buy multiple items (extension 3), then the seller can use the policy

$$v_B^*\left(s, k+j+1, \frac{a+\Theta(t)}{a+\Theta(t)+T-t}\right)$$

at time t if j customers have made a purchase and the current inventory level is s, where $\Theta(t)$ is defined in Eq. (7).

6. Conclusions

In this paper we present a dynamic pricing model where the seller needs to sell a given stock of identical items by a deadline. Unlike traditional dynamic pricing models where the seller knows the customer arrival rate, a key assumption in our model is that the seller can only estimate the arrival rate. As the sale moves forward, the seller collects the sales data in real time to fine-tune the customer arrival rate estimation. He then uses this fine-tuned arrival rate estimation to better forecast the future demand curve. We develop a variable-rate policy such that the seller incorporates this learning mechanism in setting the product price. Numerical study demonstrates the efficiency of the variable-rate policy and its robustness when the customer arrival rate changes.

The proposed model arises in many retail service industries. For instance, when a performing art event goes on a nationwide tour, there is little historical data to support the estimation of the demand rate at different cities. In the sport industry, the attendance of a team's home game varies depending on the time of the game, the opposing team, etc. Even in the airline industry where tons of historical data are available, the demand for air travel services on a particular day may possess a distinct pattern if a one-time event, such as the Olympic Games, takes place at the destination city. In these situations, it is more reasonable to allow some margin of errors in the initial demand estimation, rather than to assume one can accurately predict the demand curve.

As mentioned in the introduction, the customer arrival rate and the customer reservation price distribution are two major sources of demand uncertainty. Whereas this paper focuses on learning the customer arrival rate, there are some cases where learning the customer reservation price distribution is important. For example, when a fashion store introduces a newly designed seasonal garment, it may be easy to forecast the number of customers that will visit the store, but it is often difficult to predict how much money those customers are willing to pay for the new product. In addition, could lowering the price in the beginning be an effective method to learn more about the demand curve in order to make more profit later on? These observations suggest important future research directions.

Acknowledgements

The author thanks anonymous referees for careful reviews and helpful comments. This material is based upon work supported by the National Science Foundation under Grant No. 0223314. Most of the work was done when the author was in the Grado Department of Industrial and Systems Engineering at Virginia Tech.

Appendix A

Proof of Lemma 3.1. We first prove that $J(s, c, \alpha)$ is supermodular in (s, c); that is,

$$J(s,c,\alpha) + J(s-1,c-1,\alpha) \geqslant J(s,c-1,\alpha) + J(s-1,c,\alpha). \tag{9}$$

Consider an otherwise-identical model where each customer reveals his reservation price after the supernatural agent posts the price, and let $\widetilde{J}(\cdot,\cdot,\cdot)$ denote the optimal revenue function in this otherwise-identical model. Because the reservation prices are independent, knowing the reservation prices of previous customers does not help the agent to make future decisions; therefore $\widetilde{J} = J$. We will first prove that $\widetilde{J}$ is supermodular in (s, c), which then validate the same result for J.

Consider four agents: agents 1, 2, $\bar{1}$, $\bar{2}$. Agents 1 and $\bar{1}$ each have s items, and agents 2 and $\bar{2}$ each have $s-1$ items. The numbers of customers for agents 1 and $\bar{2}$ follow a negative binomial distribution with parameters $(c-1, \alpha)$, and those for agents 2 and $\bar{1}$ follow a negative binomial distribution with parameters (c, α). To use a sample path argument, we couple the customer arrival processes for four agents such that the number of customers in each of the first $c-1$ groups are identical for four agents, and that each corresponding customer has the same reservation price. We will present feasible policies for agents $\bar{1}$ and $\bar{2}$ such that, for each sample path, the total revenue is no less than that generated by agents 1 and 2 with the optimal policy. We can then conclude that

$$\widetilde{J}(s,c,\alpha) + \widetilde{J}(s-1,c-1,\alpha) \geqslant \widetilde{J}(s,c-1,\alpha) + \widetilde{J}(s-1,c,\alpha), \tag{10}$$

and complete the proof.

To do so, let agents 1 and 2 use the optimal policy, and let agents $\bar{1}$ and $\bar{2}$ respectively follow agents 1's and 2's policies, and switch to the optimal policy when one of the following three events eventually occurs: (a) agent 1 sells one more item than agent 2 does; (b) agents 1 and $\bar{2}$ exhaust their customers; and (c) agents

2 and $\bar{2}$ sell out their items. Note that the preceding is indeed an feasible policy for agents $\bar{1}$ and $\bar{2}$ because in this otherwise-identical model each customer reveals his reservation price after the agent posts the price. Therefore, agents $\bar{1}$ and $\bar{2}$ each can keep track of the number of items sold by both agents 1 and 2.

We next show that in each scenario, agents $\bar{1}$ and $\bar{2}$ together can generate more revenue than agents 1 and 2 together.

(a) Agent 1 sells one more item than agent 2 does: At that moment all four agents have exactly the same number of items in inventory. Therefore, the additional revenue for agents $\bar{1}$ and $\bar{2}$ together is equal to that for agents 1 and 2 together.
(b) Agents 1 and $\bar{2}$ exhaust their customers: At that moment agents $\bar{1}$ and 2 each have one group of customers, and that agent $\bar{1}$ has at least one more item than agent 2 does (otherwise (a) must have occurred). Therefore, agent $\bar{1}$ will generate at least as much revenue as agent 2.
(c) Agents 2 and $\bar{2}$ sell out their items: At that moment agents 1 and $\bar{1}$ have the same number of items, and agent $\bar{1}$ has one more group of customers than agent 1. Therefore, agent $\bar{1}$ will generate at least as much revenue as agent 1.

Because in each sample path, agents $\bar{1}$ and $\bar{2}$ together generate at least as much revenue as agents 1 and 2, taking expected value over all sample paths yields Eq. (10). Eq. (9) then follows because $\widetilde{J} = J$.

Using a similar argument we can show $J(s, c, \alpha)$ is submodular in (s, α).

To show that $J(s, c, \alpha)$ is concave in s, we first suppose the statement is false, and then draw a contradiction. Specifically, if $J(s+1, c, \alpha) - J(s, c, \alpha) > J(s, c, \alpha) - J(s-1, c, \alpha)$, then for all v,

$$g(v) - v(J(s+1, c, \alpha) - J(s, c, \alpha)) < g(v) - v(J(s, c, \alpha) - J(s-1, c, \alpha)),$$

and therefore,

$$\frac{1-\alpha}{\alpha}\max_v\{g(v) - v(J(s+1, c, \alpha) - J(s, c, \alpha))\} < \frac{1-\alpha}{\alpha}\max_v\{g(v) - v(J(s, c, \alpha) - J(s-1, c, \alpha))\}.$$

Note that Eq. (2) implies that

$$J(s, c, \alpha) - J(s, c-1, \alpha) = \frac{1-\alpha}{\alpha}\max_v\{g(v) - v(J(s, c, \alpha) - J(s-1, c, \alpha))\},$$

which leads to the following inequality

$$J(s+1, c, \alpha) - J(s+1, c-1, \alpha) < J(s, c, \alpha) - J(s, c-1, \alpha).$$

Because this last equation contradicts Eq. (9), we can conclude that $J(s, c, \alpha)$ is concave in s.

Proof of Theorem 3.1. We first prove that $v^*(s, c, \alpha) \leqslant v^*(s+1, c, \alpha)$. Consider any policy $y > v^*(s+1, c, \alpha)$ in the following:

$$\begin{aligned} g(y) - y(J(s, c, \alpha) - J(s-1, c, \alpha)) &= g(y) - y(J(s+1, c, \alpha) - J(s, c, \alpha)) \\ &\quad - \underbrace{y}_{>v^*(s+1,c,\alpha)} \underbrace{(J(s, c, \alpha) - J(s-1, c, \alpha) - J(s+1, c, \alpha) + J(s, c, \alpha))}_{\geqslant 0 \text{ by Lemma3.1}} \\ &\leqslant g(v^*(s+1, c, \alpha)) - v^*(s+1, c, \alpha)(J(s+1, c, \alpha) - J(s, c, \alpha)) \\ &\quad - v^*(s+1, c, \alpha)(J(s, c, \alpha) - J(s-1, c, \alpha) - J(s+1, c, \alpha) + J(s, c, \alpha)) \\ &= g(v^*(s+1, c, \alpha)) - v^*(s+1, c, \alpha)(J(s, c, \alpha) - J(s-1, c, \alpha)), \end{aligned}$$

where the inequality follows from the definition of $v^*(s+1,c,\alpha)$ in Eq. (4), the assumption $y > v^*(s+1,c,\alpha)$, and that $J(s,c,\alpha)$ is concave in s from Lemma 3.1. According to Eq. (2), the preceding equation states that using a policy $y > v^*(s+1,c,\alpha)$ in state (s,c,α) can do at most as well as using $v^*(s+1,c,\alpha)$. For that reason, any policy greater than $v^*(s+1,c,\alpha)$ cannot be optimal in state (s,c,α), implying $v^*(s,c,\alpha) \leqslant v^*(s+1,c,\alpha)$.

By applying the other two parts in Lemma 3.1 with a similar argument, the fact that $J(s,c,\alpha)$ is supermodular in (s,c) leads to the conclusion that $v^*(s,c,\alpha)$ decreases in c, and the fact that $J(s,c,\alpha)$ is submodular in (s,α) leads to the conclusion that $v^*(s,c,\alpha)$ increases in α.

References

[1] N. Agrawal, S.A. Smith, Estimating negative binomial demand for retail inventory management with unobservable lost sales, Naval Research Logistics 43 (6) (1996) 839–861.
[2] K.S. Azoury, Bayes solutions to dynamic inventory models under unknown demand distribution, Management Science 31 (9) (1985) 1150–1160.
[3] R.J. Balvers, T.F. Cosimano, Actively learning about demand and the dynamics of price adjustment, The Economic Journal 100 (402) (1990) 882–898.
[4] G.R. Bitran, S.V. Mondschein, An application of yield management to the hotel industry considering multiple day stays, Operations Research 43 (3) (1995) 427–443.
[5] A.N. Burnetas, C.E. Smith, Adaptive ordering and pricing for perishable products, Operational Research 48 (3) (2000) 436–443.
[6] R.E. Chatwin, Optimal dynamic pricing of perishable products with stochastic demand and a finite set of prices, European Journal of Operational Research 125 (2000) 149–174.
[7] J.M. Feldman, Fares: To raise or not to raise. Air Transport World 27 (6) (1990) 58–59.
[8] Y. Feng, B. Xiao, A continuous-time yield management model with multiple prices and reversible price changes, Management Science 46 (5) (2000) 644–657.
[9] G. Gallego, G. van Ryzin, Optimal dynamic pricing of inventories with stochastic demand over finite horizons, Management Science 40 (8) (1994) 999–1020.
[10] W.M. Kincaid, D. Darling, An inventory pricing problem, Journal of Mathematical Analysis and Applications 7 (1963) 183–208.
[11] S.P. Ladany, A. Arbel, Optimal cruise-liner passenger cabin pricing policy, European Journal of Operational Research 55 (1991) 136–147.
[12] M.A. Lariviere, E.L. Porteus, Stalking information: Bayesian inventory management with unobserved lost sales, Management Science 45 (3) (1999) 346–363.
[13] C.J. Lautenbacher, S. Stidham, The underlying markov decision process in the single-leg airline yield management problem, Transportation Science 33 (2) (1999) 136–146.
[14] G.L. Lilien, P. Kotler, K. Sridhar Moorthy, Marketing Models, Prentice Hall, Upper Saddle River, NJ, 1992, 07458.
[15] K.Y. Lin, A sequential dynamic pricing model and its applications, Naval Research Logistics 51 (4) (2004) 501–521.
[16] W.S. Lovejoy, Myopic policies for some inventory models with uncertain demand distributions, Management Science 36 (6) (1990) 724–738.
[17] J.I. McGill, G.J. van Ryzin, Revenue management: Research overview and prospects, Transportation Science 33 (2) (1999) 233–256.
[18] N.C. Petruzzi, M. Dada, Dynamic pricing and inventory control with learning, Naval Research Logistics 49 (3) (2002) 303–325.
[19] J. Subramanian, S. Stidham, C.J. Lautenbacher, Airline yield management with overbooking, cancellations, and no-shows, Transportation Science 33 (2) (1999) 147–167.
[20] L.R. Weatherford, S.E. Bodily, A taxonomy and research overview of perishable asset revenue management: Yield management, overbooking, and pricing, Operations Research 40 (5) (1992) 831–844.