Dynamic Pricing with a Prior on Market Response

Vivek F. Farias
Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, vivekf@mit.edu

Benjamin Van Roy
Stanford University, Stanford, California 94305, bvr@stanford.edu

We study a problem of dynamic pricing faced by a vendor with limited inventory, uncertain about demand, and aiming to maximize expected discounted revenue over an infinite time horizon. The vendor learns from purchase data, so his strategy must take into account the impact of price on both revenue and future observations. We focus on a model in which customers arrive according to a Poisson process of uncertain rate, each with an independent, identically distributed reservation price. Upon arrival, a customer purchases a unit of inventory if and only if his reservation price equals or exceeds the vendor's prevailing price.

We propose a simple heuristic approach to pricing in this context, which we refer to as decay balancing. Computational results demonstrate that decay balancing offers significant revenue gains over recently studied certainty equivalent and greedy heuristics. We also establish that changes in inventory and uncertainty in the arrival rate bear appropriate directional impacts on decay balancing prices in contrast to these alternatives, and we derive worst-case bounds on performance loss. We extend the three aforementioned heuristics to address a model involving multiple customer segments and stores, and provide experimental results demonstrating similar relative merits in this context.

Subject classifications: dynamic pricing; demand learning; revenue management.
Area of review: Revenue Management.
History: Received March 2007; revisions received November 2007, July 2008, January 2009; accepted March 2009.
    Published online in *Articles in Advance* November 19, 2009.

1. Introduction

Consider a vendor of winter apparel. New items are stocked in the autumn and sold over several months. Because of significant manufacturing lead times and fixed costs, items are not restocked over this period. Evolving fashion trends generate great uncertainty in the number of customers who will consider purchasing these items. To optimize revenue, the vendor should adjust prices over time. But how should these prices be set as time passes and units are sold? This is representative of problems faced by many vendors of seasonal, fashion, and perishable goods.

There is a substantial literature on pricing strategies for such a vendor (see Talluri and van Ryzin 2004 and references therein). Gallego and van Ryzin (1994), in particular, formulated an elegant model in which the vendor starts with a finite number of identical indivisible units of inventory. Customers arrive according to a Poisson process, with independent, identically distributed reservation prices. In the case of exponentially distributed reservation prices, the optimal pricing strategy is easily derived. The analysis of Gallego and van Ryzin (1994) can be used to derive pricing strategies that optimize expected revenue over a finite horizon and is easily extended to the optimization of discounted expected revenue over an infinite horizon. Resulting strategies provide insight into how prices should depend on the arrival rate, expected reservation price, and the length of the horizon or discount rate.

Our focus is on an extension of this model in which the arrival rate is uncertain and the vendor learns from sales data. Incorporating such uncertainty is undoubtedly important in many industries that practice revenue management. For instance, in the winter fashion apparel example, there might be great uncertainty in how the market will respond to the product at the beginning of a sales season; the vendor must take into account how price influences both revenue and future observations from which he can learn.

In this setting, it is important to understand how uncertainty should influence price. However, uncertainty in the arrival rate makes the analysis challenging. Optimal pricing strategies can be characterized by a Hamilton-Jacobi-Bellman (HJB) equation, but this approach is typically not analytically tractable (at least for the models we consider). Furthermore, for arrival rate distributions of interest, grid-based numerical methods require discommoding computational resources and generate strategies that are difficult to interpret. As such, researchers have designed and analyzed heuristic approaches.

Aviv and Pazgal (2005a) studied a *certainty equivalent* heuristic for exponentially distributed reservation prices, which at each point in time computes the conditional expectation of the arrival rate, conditioned on observed sales data, and prices as though the arrival rate is equal to this expectation. Araman and Caldentey (2009) recently proposed a more sophisticated heuristic that takes arrival

rate uncertainty into account when pricing. The idea is to use a strategy that is *greedy* with respect to a particular approximate value function. In this paper, we propose and analyze *decay balancing*, a new heuristic approach which makes use of the same approximate value function as the greedy approach of Araman and Caldentey (2009).

Several idiosyncrasies distinguish the models studied in Aviv and Pazgal (2005a) and Araman and Caldentey (2009). The former models uncertainty in the arrival rate in terms of a Gamma distribution, whereas the latter uses a two-point distribution. The former considers maximization of expected revenue over a finite horizon, whereas the latter considers expected discounted revenue over an infinite horizon. To elucidate relationships among the three heuristic strategies, we study them in the context of a common model. In particular, we take the arrival rate to be distributed according to a finite mixture of Gamma distributions. This is a very general class of priors and can closely approximate any bounded continuous density. We take the objective to be maximization of expected discounted revenue over an infinite horizon. It is worth noting that in the case of exponentially distributed reservation prices, such a model is equivalent to one without discounting but where expected reservation prices diminish exponentially over time. This might make it an appropriate model for certain seasonal, fashion, or perishable products. Our modeling choices were made to provide a simple yet fairly general context for our study. We expect that our results can be extended to other classes of models such as those with finite time horizons, although this is left for future work.

The certainty equivalent heuristic is natural and simple to implement. It does not, however, take uncertainty in the arrival rate into account. While the greedy heuristic does attempt to do this, we demonstrate through computational experiments that the performance of this approach can degrade severely at high levels of uncertainty in arrival rate. While being only slightly more complex to implement than the certainty equivalent heuristic and typically less so than greedy pricing, we demonstrate that decay balancing offers significant performance gains over these heuristics especially in scenarios where the seller begins with a high degree of uncertainty in arrival rate. From a more qualitative perspective, uncertainty in the arrival rate and changes in inventory bear appropriate directional impacts on decay balancing prices: uncertainty in the arrival rate increases price, while a decrease in inventory increases price. In contrast, uncertainty in the arrival rate has no impact on certainty equivalent prices, while greedy prices can increase or decrease with inventory.

In addition to our base model, we consider a generalization that involves a vendor with multiple branches that can offer different prices and tend to attract different classes of customers. The branches share and learn from each others' data, and they price to maximize aggregate expected discounted revenue. We extend the three heuristics to the context of this model and present computational results demonstrating relative merits analogous to our base case of a single branch and single-customer class.

We establish bounds on performance loss incurred by decay balancing relative to an optimal policy. These bounds indicate a certain degree of robustness. For example, when customer reservation prices are exponentially distributed and their arrival rate is Gamma distributed, we establish that decay balancing always garners at least one third of the maximum expected discounted revenue. Allowing for a dependence on arrival rate uncertainty and/or restricting attention to specific classes of reservation price distributions leads to substantially stronger bounds. It is worth noting that no performance loss bounds (uniform or otherwise) have been established for the certainty equivalent and greedy heuristics. Furthermore, computational results demonstrate that our bounds are not satisfied by the greedy heuristic.

Aside from Aviv and Pazgal (2005a) and Araman and Caldentey (2009), there is a significant literature on dynamic pricing while learning about demand. A recent paper in this regard is Aviv and Pazgal (2005b) which considers, in a discrete time setting, a partially observable Markov modulated demand model. As we will discuss further in §6, a special case of the heuristic they develop is closely related to decay balancing. Lin (2006) considers a model identical to Aviv and Pazgal (2005a) and develops heuristics which are motivated by the behavior of a seller who knows the arrival rate and anticipates all arriving customers. Bertsimas and Perakis (2006) develop several algorithms for a discrete, finite time-horizon problem where demand is an unknown linear function of price plus Gaussian noise. This allows for least-squares-based estimation. Lobo and Boyd (2003) study a model similar to Bertsimas and Perakis (2006) and propose a "price-dithering" heuristic that involves the solution of a semidefinite convex program. All the aforementioned work is experimental; no performance guarantees are provided for the heuristics proposed. Cope (2006) studies a Bayesian approach to pricing where inventory levels are unimportant (this is motivated by sales of online services) and there is uncertainty in the distribution of reservation price. His work uses a very general prior distribution—a Dirichlet mixture—on reservation price. Modeling this type of uncertainty within a framework where inventory levels do matter represents an interesting direction for future work. In contrast with the above work, Burnetas and Smith (1998) and Kleinberg and Leighton (2004) consider nonparametric approaches to pricing with uncertainty in demand. However, those models again do not account for inventory levels. Recently, Besbes and Zeevi (2006) presented a nonparametric algorithm for "blind" pricing of multiple products that use multiple resources, similar to the model considered in Gallego and van Ryzin (1997). Their algorithm requires essentially no knowledge of the demand function and is optimal in a certain asymptotic regime. However, the algorithm requires trying each alternative among a multidimensional grid in the space of price vectors and, therefore, could take a long time to adequately learn about demand.

The remainder of this paper is organized as follows. In §2, we formulate our model and cast our pricing problem as one of stochastic optimal control. Section 3 develops the HJB equation for the optimal pricing problem in the contexts of known and unknown arrival rates. Section 4 first introduces existing heuristics for the problem and then introduces a new heuristic—decay balancing—which is the focus of this paper. This section also discusses structural properties of the decay balancing policy. Section 5 presents a computational study that compares decay balancing to certainty equivalent and greedy pricing heuristics as well as a clairvoyant algorithm. Section 6 is devoted to a theoretical performance analysis of the decay balancing heuristic. We provide worst-case performance guarantees that depend on initial uncertainty in market response. When the arrival rate is Gamma distributed and reservation prices are exponentially distributed, we prove a uniform performance guarantee for our heuristic. Section 7 discusses an extension of our heuristic to a multidimensional version of the problem which involves a vendor with multiple stores and several customer segments with different demand characteristics. This section presents computational results that are qualitatively similar to those in §5. Section 8 concludes.

2. Problem Formulation

We consider a problem faced by a vendor who begins with x_0 identical indivisible units of a product and dynamically adjusts price p_t over time $t \in [0, \infty)$. Customers arrive according to a Poisson process with rate λ. As a convention, we will assume that the arrival process is right continuous with left limits. Each customer's reservation price is an independent random variable with cumulative distribution $F(\cdot)$. A customer purchases a unit of the product if it is available at the time of his arrival at a price no greater than his reservation price; otherwise, the customer permanently leaves the system.

For convenience, we introduce the notation $\bar{F}(p) = 1 - F(p)$ for the tail probability. We place the following restrictions on $F(\cdot)$:

ASSUMPTION 1. 1. *$F(\cdot)$ has a differentiable density $f(\cdot)$ with support $\mathbb{R}_+$.*

2. F has a nondecreasing hazard rate. That is, $\rho(p) = f(p)/\bar{F}(p)$ is nondecreasing in p.

Assumption 1 is satisfied by many relevant distributions including the exponential, extreme value, logistic, and Weibull, to name a few. We introduce this assumption to facilitate use of first-order optimality conditions to characterize solutions of various optimization problems that will arise in our discussion. It is possible to extend our results to reservation price distributions that do not satisfy this assumption, although that would require additional technical work.

Let t_k denote the time of the kth purchase and $n_t = |\{t_k: t_k \leqslant t\}|$ denote the number of purchases made by customers arriving at or before time t. The vendor's expected revenue, discounted at a rate of $\alpha > 0$, is given by

$$E\left[\int_{t=0}^{\infty} e^{-\alpha t} p_t \, dn_t \right].$$

Let $\tau_0 = \inf\{t: x_t = 0\}$ be the time at which the final unit of inventory is sold. For $t \leqslant \tau_0$, n_t follows a Poisson process with intensity $\lambda \bar{F}(p_t)$. Consequently (see Theorem III.T9 in Bremaud 1981), one may show that

$$E\left[\int_{t=0}^{\infty} e^{-\alpha t} p_t \, dn_t \right] = E\left[\int_{t=0}^{\tau_0} e^{-\alpha t} p_t \lambda \bar{F}(p_t) \, dt \right].$$

We now describe the vendor's optimization problem. We first consider the case where the vendor knows λ and later allow for arrival rate uncertainty. In the case with known arrival rate, we consider pricing policies π that are measurable real-valued functions of the inventory level. The price is irrelevant when there is no inventory, and as a convention, we will require that $\pi(0) = \infty$. We denote the set of policies by Π_λ. A vendor who employs pricing policy $\pi \in \Pi_\lambda$ sets price according to $p_t = \pi(x_t)$, where $x_t = x_0 - n_t$, and receives expected discounted revenue

$$J_\lambda^\pi(x) = E_{x, \pi}\left[\int_{t=0}^{\tau_0} e^{-\alpha t} p_t \lambda \bar{F}(p_t) \, dt \right],$$

where the subscripts of the expectation indicate that $x_0 = x$ and $p_t = \pi(x_t)$. The optimal discounted revenue is given by $J_\lambda^*(x) = \sup_{\pi \in \Pi_\lambda} J_\lambda^\pi(x)$, and a policy π is said to be optimal if $J_\lambda^* = J_\lambda^\pi$.

Suppose now that the arrival rate λ is not known, but rather, the vendor starts with a prior on λ that is a finite mixture of Gamma distributions. A Kth order mixture of this type is parameterized by vectors $a_0, b_0 \in \mathbb{R}_+^K$ and a vector of K weights $w_0 \in \mathbb{R}_+^K$ that sum to unity. The density, g, for such a prior is given by

$$g(\lambda) = \sum_k w_{0, k} \frac{b_{0, k}^{a_{0, k}} \lambda^{a_{0, k} - 1} e^{-\lambda b_{0, k}}}{\Gamma(a_{0, k})},$$

where Γ denotes the Gamma function: $\Gamma(x) = \int_{s=0}^{\infty} s^{x-1} e^{-s} \, ds$. The expectation and variance are $E[\lambda] = \sum_k w_{0, k} a_{0, k} / b_{0, k} \triangleq \mu_0$ and $\mathrm{Var}[\lambda] = \sum_k w_{0, k} a_{0, k} (a_{0, k} + 1) / b_{0, k}^2 - \mu_0^2$. Any prior on λ with a continuous, bounded density can be approximated to an arbitrary accuracy within such a family (see Dalal and Hall 1983). Moreover, as we describe below, posteriors on λ continue to remain within this family rendering such a model parsimonious as well as relatively tractable.

The vendor revises his beliefs about λ as sales are observed. In particular, at time t, the vendor obtains a posterior that is a kth order mixture of Gamma distributions with parameters

$$a_{t, k} = a_{0, k} + n_t \quad \text{and} \quad b_{t, k} = b_{0, k} + \int_{\tau=0}^{t} \bar{F}(p_\tau) \, d\tau,$$

and weights that evolve according to

$$dw_{t,k} = w_{t,k}\left(\frac{a_{t,k}/b_{t,k} - \sum_k w_{t,k}(a_{t,k}/b_{t,k})}{\sum_k w_{t,k}(a_{t,k}/b_{t,k})} dn_t \right.$$
$$\left. + \left(a_{t,k}/b_{t,k} - \sum_k w_{t,k}(a_{t,k}/b_{t,k}) \right) dt \right).$$

Note that the vendor does not observe all customer arrivals but only those that result in sales. Furthermore, lowering price results in more frequent sales and therefore more accurate estimation of the demand rate.

We consider pricing policies π that are measurable real-valued functions of the inventory level and arrival rate distribution parameters. As a convention, we require that $\pi(0, a, b, w) = \infty$ for all arrival rate distribution parameters a, b, and w. We denote the domain by $\mathscr{S} = \mathbb{N} \times \mathbb{R}_+^k \times \mathbb{R}_+^k \times \mathbb{R}_+^k$ and the set of policies by Π. Let $z_t = (x_t, a_t, b_t, w_t)$. A vendor who employs pricing policy $\pi \in \Pi$ sets price according to $p_t = \pi(z_t)$ and receives expected discounted revenue

$$J^\pi(z) = E_{z,\pi}\left[\int_{t=0}^{\tau_0} e^{-\alpha t} p_t \lambda \bar{F}(p_t)\, dt \right],$$

where the subscripts of the expectation indicate that $z_0 = z$ and $p_t = \pi(z_t)$. Note that, unlike the case with known arrival rate, λ is a random variable in this expectation. The optimal discounted revenue is given by $J^*(z) = \sup_{\pi \in \Pi} J^\pi(z)$, and a policy π is said to be optimal if $J^* = J^\pi$. We will use the notation $J^{*,\alpha}$ for the optimal value function when we wish to emphasize the dependence on α.

3. Optimal Pricing

An optimal pricing policy can be derived from the value function J^*. The value function in turn solves the HJB equation. Unfortunately, direct solution of the HJB equation, either analytically or computationally, does not appear to be a feasible task and one must resort to heuristic policies. With an end to deriving such heuristic policies, we characterize optimal solutions to problems with known and unknown arrival rates and discuss some of their properties.

3.1. The Case of a Known Arrival Rate

We begin with the case of a known arrival rate. For each $\lambda \geq 0$ and $\pi \in \Pi_\lambda$, define an operator $H_\lambda^\pi \colon \mathbb{R}^{x_0+1} \to \mathbb{R}^{x_0+1}$ by

$$(H_\lambda^\pi J)(x) = \lambda \bar{F}(\pi(x))(\pi(x) + J(x-1) - J(x)) - \alpha J(x).$$

Recall that $\pi(0) = \infty$. In this case, we interpret $\bar{F}(\pi(0))\pi(0)$ as a limit, and Assumption 1 (which ensures a finite, unique static revenue maximizing price) implies that $(H_\lambda^\pi J)(0) = -\alpha J(0)$. Furthermore, we define the dynamic programming operator

$$(H_\lambda J)(x) = \sup_{\pi \in \Pi_\lambda} (H_\lambda^\pi J)(x).$$

It is easy to show that J_λ^* is the unique solution to the HJB equation $H_\lambda J = 0$. The first-order optimality condition for prices yields an optimal policy of the form

$$\pi_\lambda^*(x) = 1/\rho(\pi_\lambda^*(x)) + J_\lambda^*(x) - J_\lambda^*(x-1)$$

for $x > 0$. By Assumption 1 and the fact that $J_\lambda^*(x) \geq J_\lambda^*(x-1)$, the above equation always has a solution in $\mathbb{R}_+$.

Given that J^* satisfies the HJB equation, we have

$$\alpha J_\lambda^*(x)$$
$$= \begin{cases} \sup_{p \geq 0} \lambda \bar{F}(p)(p + J_\lambda^*(x-1) - J_\lambda^*(x)) & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Assumption 1 guarantees that $\sup_{p \geq 0} \bar{F}(p)(p - c)$ is a decreasing function of c on $\mathbb{R}_+$. This allows one to compute $J_\lambda^*(x)$ given $J_\lambda^*(x-1)$ via bisection search. This offers an efficient algorithm that computes $J_\lambda^*(0), J_\lambda^*(1), \ldots, J_\lambda^*(x)$ in x iterations. As a specific concrete example, consider the case where reservation prices are exponentially distributed with mean $r > 0$. We have

$$\alpha J_\lambda^*(x) = \begin{cases} \lambda r \exp\left(\frac{1}{r}(J_\lambda^*(x-1) - J_\lambda^*(x)) - 1 \right) & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

It follows that

$$J_\lambda^*(x) = rW((e^{-1}\lambda/\alpha) \exp(J_\lambda^*(x-1)/r)) \quad (2)$$

for $x > 0$, where $W(\cdot)$ is the Lambert W-function (the inverse of xe^x).

We note that a derivation of the optimal policy for the case of a known arrival rate can also be found in Araman and Caldentey (2009), among other sources.

The value function for the case of a known arrival rate will be used in the design of our heuristic for the case with arrival rate uncertainty. We establish here properties of this value function J_λ^* and its associated optimal policy π_λ^*, which we will later use. We will make the following assumption to simplify our analysis.

ASSUMPTION 2. $J_\lambda^*(x)$ *is a differentiable function of λ on* $\mathbb{R}_+$ *for all* $x \in \mathbb{N}$.

Note that this assumption is satisfied for the case of exponential reservation prices. The following comparative statics for π_λ^* are proved in the online appendix.

An electronic companion to this paper is available as part of the online version that can be found at http://or.journal.informs.org/.

LEMMA 1. $\pi_\lambda^*(x)$ *is decreasing in x (on $\mathbb{N}$) and nondecreasing in λ (on $\mathbb{R}_+$).*

For a fixed inventory level, it is natural to expect decreasing returns to increases in the arrival rate λ; the following lemma, proved in the online appendix, formalizes this intuition.

LEMMA 2. *For all $x \in \mathbb{N}$, $J_\lambda^*(x)$ is an increasing, concave function of λ on* $\mathbb{R}_+$.

3.2. The Case of an Unknown Arrival Rate

Let $S_{\tilde{x}, \tilde{a}, \tilde{b}} = \{(x, a, b, w) \in S : a + x = \tilde{a} + \tilde{x}, \tilde{b} \leqslant b, w \geqslant 0, 1'w = 1\}$ denote the set of states that might be visited starting at a state with $x_0 = \tilde{x}$, $a_0 = \tilde{a}$, $b_0 = \tilde{b}$. Let $\mathcal{J}$ denote the set of functions $J: \mathcal{S} \mapsto \mathbb{R}$ such that $\sup_{z \in \mathcal{S}_{\tilde{x}, \tilde{a}, \tilde{b}}} |J(z)| < \infty$ for all $\tilde{x}$ and $\tilde{b} > 0$ and that have bounded derivatives with respect to the third and fourth arguments. We define $\mu(z)$ to be the expectation for the prior on arrival rate in state z, so that $\mu(z) = \sum_k w_k a_k / b_k$.

For each policy $\pi \in \Pi$, we define an operator

$$(H^\pi J)(z) = \bar{F}(\pi(z))(\mu(z)(\pi(z) + J(z') - J(z))$$
$$+ (DJ)(z)) - \alpha J(z),$$

where $z \in \mathcal{S}_{\tilde{x}, \tilde{a}, \tilde{b}}$, $z = (x, a, b, w)$ and $z' = (x - 1, a + 1, b, w')$. Here, w' is defined according to $w'_k = (w_k a_k)/(b_k \mu(z))$, and D is a differential operator given by

$$(DJ)(z) = \sum_k w_k (\mu(z) - a_k/b_k) \frac{d}{dw_k} J(z) + \frac{d}{db_k} J(z).$$

We now define the dynamic programming operator H according to

$$(HJ)(z) = \sup_\pi (H^\pi J)(z). \tag{3}$$

Using standard dynamic programming arguments, one can show that the value function J^* solves the HJB equation

$$(HJ)(z) = 0. \tag{4}$$

Furthermore, a policy π is optimal if and only if $H^\pi J^* = 0$. This equation provides a prescription for efficient computation of an optimal policy given the value function. Unfortunately, there is no known analytical solution to the HJB equation when the arrival rate is unknown, even for special cases such as a Gamma or two-point prior with exponential reservation prices. Furthermore, grid-based numerical solution methods require discommoding computational resources and generate strategies that are difficult to interpret. As such, simple effective heuristics are desirable.

4. Heuristics

This section presents three heuristic pricing policies. The first two have been considered in prior literature and the third is one we propose and analyze in this paper.

4.1. Certainty Equivalent

Aviv and Pazgal (2005a) studied a *certainty equivalent* heuristic that at each point in time computes the conditional expectation of the arrival rate, conditioned on observed sales data, and prices as though the arrival rate is equal to this expectation. In our context, the price function for such a heuristic uniquely solves

$$\pi_{\text{ce}}(z) = \frac{1}{\rho(\pi_{\text{ce}}(z))} + J^*_{\mu(z)}(x) - J^*_{\mu(z)}(x - 1)$$

for $x > 0$. The existence of a unique solution to this equation is guaranteed by Assumption 1. As derived in the preceding section, this is an optimal policy for the case where the arrival rate is known and equal to $\mu(z)$, which is the expectation of the arrival rate given a prior distribution with parameters a, b, and w. The certainty equivalent policy is computationally attractive because J^*_λ is easily computed numerically (and in some cases, even analytically) as discussed in the previous section. As one would expect, prices generated by this heuristic increase as the inventory x decreases. However, arrival rate uncertainty bears no influence on price—the price only depends on the arrival rate distribution through its expectation $\mu(z)$. Hence, this pricing policy is unlikely to appropriately address information acquisition.

4.2. Greedy Pricing

We now present another heuristic which was recently proposed by Araman and Caldentey (2009) and does account for arrival rate uncertainty. To do so, we first introduce the notion of a *greedy policy*. A policy π is said to be *greedy* with respect to a function J if $H^\pi J = HJ$. The first-order necessary condition for optimality and Assumption 1 imply that the greedy price is given by the solution to

$$\pi(z) = \left(\frac{1}{\rho(\pi(z))} + J(z) - J(z') - \frac{1}{\mu(z)} (DJ)(z) \right)^+$$

for $z = (x, a, b, w)$ with $x > 0$ and $z' = (x - 1, a + 1, b, w')$ with $w'_k = (w_k a_k)/(b_k \mu)$.

Perhaps the simplest approximation one might consider to $J^*(z)$ is $J^*_{\mu(z)}(x)$, the value for a problem with known arrival rate $\mu(z)$. One troubling aspect of this approximation is that it ignores the variance (as also higher moments) of the arrival rate. The alternative approximation proposed by Araman and Caldentey (2009) takes variance into account. In particular, their heuristic employs a greedy policy with respect to the approximate value function $\tilde{J}$ which takes the form

$$\tilde{J}(z) = E[J^*_\lambda(x)],$$

where the expectation is taken over the random variable λ, which is drawn from a Gamma mixture with parameters a, b, and w. $\tilde{J}(z)$ can be thought of as the expected optimal value if λ is to be observed at the next time instant. It is interesting to note that this approximation is in the same spirit as the "full information" approximation considered in Aviv and Pazgal (2005b). The greedy price, however, is distinct from the full information price considered there.

Because it can only help to know the value of λ, $J^*_\lambda(x) \geqslant E[J^*(z) \mid \lambda]$. Taking expectations of both sides of this inequality, we see that $\tilde{J}$ is an upper bound on J^*. The approximation $J^*_{\mu(z)}(x)$ is a looser upper bound on $J^*(z)$ (which follows from concavity of J^*_λ in λ). Consequently, we have the following result, whose proof can be found in the online appendix.

LEMMA 3. *For all $z \in \mathcal{S}$, $\alpha > 0$,*

$$J^*(z) \leqslant \tilde{J}(z) \leqslant J^*_{\mu(z)}(x) \leqslant \frac{\bar{F}(p^*)p^*\mu(z)}{\alpha},$$

where p^* is the static revenue maximizing price.

The greedy price in state z is thus the solution to

$$\pi_{\mathrm{gp}}(z) = \left(\frac{1}{\rho(\pi_{\mathrm{gp}}(z))} + \tilde{J}(z) - \tilde{J}(z') - \frac{1}{\mu(z)}(D\tilde{J})(z) \right)^+$$

for $z = (x, a, b, w)$ with $x > 0$ and $z' = (x-1, a+1, b, w')$ with $w'_k = (w_k a_k)/(b_k \mu(z))$.

We have observed through computational experiments (see §6) that when reservation prices are exponentially distributed and the vendor begins with a Gamma prior with scalar parameters a and b, greedy prices can increase or decrease with the inventory level x, keeping a and b fixed. This is clearly not optimal behavior.

4.3. Decay Balancing

We now describe *decay balancing*, a new heuristic that will be the primary subject of the remainder of the paper. To motivate the heuristic, we start by deriving an alternative characterization of the optimal pricing policy. The HJB Equation (4) yields

$$\max_{p \geqslant 0} \bar{F}(p)(\mu(z)(p + J^*(z') - J^*(z)) + (DJ^*)(z)) = \alpha J^*(z)$$

for all $z = (x, a, b, w)$ and $z' = (x-1, a+1, b, w')$ with $x > 0$ and $w'_k = (w_k a_k)/(b_k \mu(z))$. This equation can be viewed as a balance condition. The right-hand side represents the rate at which value decays over time; if the price were set to infinity so that no sales could take place for a time increment dt but an optimal policy is used thereafter, the current value would become $J^*(z) - \alpha J^*(z)dt$. The left-hand side represents the rate at which value is generated from both sales and learning. The equation requires these two rates to balance so that the net value is conserved.

Note that the first-order optimality condition implies that if $J(z') - J(z) + (1/\mu(z))(DJ)(z) < 0$ (which must necessarily hold for $J = J^*$),

$$\frac{\bar{F}(p^*)}{\rho(p^*)}\mu(z) = \max_{p \geqslant 0} \bar{F}(p)(\mu(z)(p + J(z') - J(z)) + (DJ)(z)),$$

if p^* attains the maximum in the right-hand side. Interestingly, the maximum depends on J only through p^*. Hence, the balance equation can alternatively be written in the following simpler form:

$$\frac{\bar{F}(\pi^*(z))}{\rho(\pi^*(z))}\mu(z) = \alpha J^*(z),$$

which implicitly characterizes π^*.

This alternative characterization of π^* makes obvious two properties of optimal prices. Note that $\bar{F}(p)/\rho(p)$ is

decreasing in p. Consequently, holding a, b, and w fixed, as x decreases, $J^*(z)$ decreases, and therefore $\pi^*(z)$ increases. Furthermore, because $J^*(z) \leqslant J^*_{\mu(z)}(x)$, we see that for a fixed inventory level x and expected arrival rate $\mu(z)$, the optimal price in the presence of uncertainty is higher than in the case where the arrival rate is known exactly.

Like greedy pricing, the decay balancing heuristic relies on an approximate value function. We will use the same approximation $\tilde{J}$. But instead of following a greedy policy with respect to $\tilde{J}$, the decay balancing approach chooses a policy π_{db} that satisfies the balance condition

$$\frac{\bar{F}(\pi_{\mathrm{db}}(z))}{\rho(\pi_{\mathrm{db}}(z))}\mu(z) = \alpha \tilde{J}(z), \tag{5}$$

with the decay rate approximated using $\tilde{J}(z)$. The following lemma guarantees that the above balance equation always has a unique solution so that our heuristic is well defined. The result is a straightforward consequence of Assumption 1 and the fact that $(\bar{F}(p^*)/(\alpha\rho(p^*)))\mu(z) \geqslant \tilde{J}(z) \geqslant J^*(z) = (\bar{F}(\pi^*(z))/(\alpha\rho(\pi^*(z))))\mu(z)$, where p^* is the static revenue maximizing price.

LEMMA 4. *For all $z \in \mathcal{S}$, there is a unique $p \geqslant 0$ such that $(\bar{F}(p)/\rho(p))\mu(z) = \alpha\tilde{J}(z)$.*

Unlike certainty equivalent and greedy pricing, uncertainty in the arrival rate and changes in inventory level have the correct directional impact on decay balancing prices. Holding a, b, and w fixed, as x decreases, $\tilde{J}(z)$ decreases, and therefore $\pi_{\mathrm{db}}(z)$ increases. Holding x and the expected arrival rate $\mu(z)$ fixed, $\tilde{J}(z) \leqslant J^*_{\mu(z)}(x)$, so that the decay balance price with uncertainty in arrival rate is higher than when the arrival rate is known with certainty.

5. Computational Study

This section will present computational results that highlight the performance of the decay balancing heuristic. We consider both Gamma and Gamma mixture priors, and we restrict attention to exponentially distributed reservation prices. Furthermore, we consider only problem instances where $\alpha = e^{-1}$ and $r = 1$; as we will discuss in §6, this is not restrictive (see Lemmas 5 and 10).

Performance Relative to a Clairvoyant Algorithm. Consider a "clairvoyant" algorithm that has access to the realization of λ at $t = 0$ and subsequently uses the pricing policy π^*_λ. The expected revenue garnered by such a pricing policy upon starting in state z is simply $E[J^*_\lambda(x)] = \tilde{J}(z)$ which, by Lemma 3, is an upper bound on $J^*(z)$. Our first experiment measures the average revenue earned using decay balancing with that earned using such a clairvoyant algorithm. The results are summarized in Table 1. We consider two cases In the first, λ is drawn from a Gamma distribution with shape parameter $a = 0.04$ and scale parameter $b = 0.001$, which corresponds to a mean of 40 and a coefficient of variation of 5. In the second,

Table 1. Performance vs. a clairvoyant algorithm.

Inventory level	$x_0 = 1$ (%)	$x_0 = 2$ (%)	$x_0 = 5$ (%)	$x_0 = 10$ (%)	$x_0 = 20$ (%)	$x_0 = 40$ (%)
Performance gain (Gamma)	−13	−10	−6	−3.7	−2	−0.5
Performance gain (Gamma mixture)	−14.9	−12.3	−6.7	−4.3	−2.7	−2.4

λ is drawn from a two-point Gamma mixture with parameters $a = [0.01023\ 0.07161]'$, $b = [0.00102\ 0.00102]'$, $w = [0.5\ 0.5]'$, which correspond to a mean of 40 and a coefficient of variation of 5. These parameter values are representative of a high level of uncertainty in λ. As is seen in Table 1, the performance of the decay balancing heuristic is surprisingly close to that of the clairvoyant algorithm for both prior distributions.

Performance Relative to Certainty Equivalent and Greedy Pricing Heuristics. We finally turn to studying the performance gains offered by decay balancing relative to the certainty equivalent and greedy policies.

1. *Dependence on initial inventory*: The gains offered by decay balancing relative to the certainty equivalent and greedy heuristic are pronounced at lower initial inventory levels, i.e., in regimes where judiciously managing inventory is crucial. We consider relative performance gains for inventory levels between 1 and 10. We consider two initial priors for all heuristics: a Gamma prior with $a = 0.04$ and $b = 0.001$ (corresponding to a mean of 40 and a coefficient of variation of 5) and a Gamma mixture prior with parameters $a = [0.01033\ 0.09297]'$, $b = [0.00103\ 0.00103]'$, $w = [0.5\ 0.5]'$ (corresponding to a mean of 50 and a coefficient of variation of 4.47). Figures 1 and 2 indicate that we offer a substantial gain in performance over the certainty equivalent and greedy pricing heuristics, especially at lower initial inventory levels. The greedy policy performs particularly poorly. In addition, as discussed earlier, that policy exhibits qualitative behavior that is clearly suboptimal: For a problem with a Gamma prior, mean reservation price 1, and discount factor e^{-1}, we compute $\pi_{gp}(1, 0.1, 0.1)\ (=1.26) < \pi_{gp}(4, 0.1, 0.1)\ (=1.61) > \pi_{gp}(10, 0.1, 0.1)\ (=1.25)$ so that, all other factors remaining the same, prices might increase or decrease with an increase in inventory level. Our gain in performance falls at higher initial inventory levels. This is not surprising; intuitively, the control problem at hand is simpler there because we are essentially allowed to sacrifice a few units of inventory early on so as to learn quickly without incurring much of a penalty.

Figure 1. Performance gain over the certainty equivalent and greedy pricing heuristics at various inventory levels for Gamma prior.
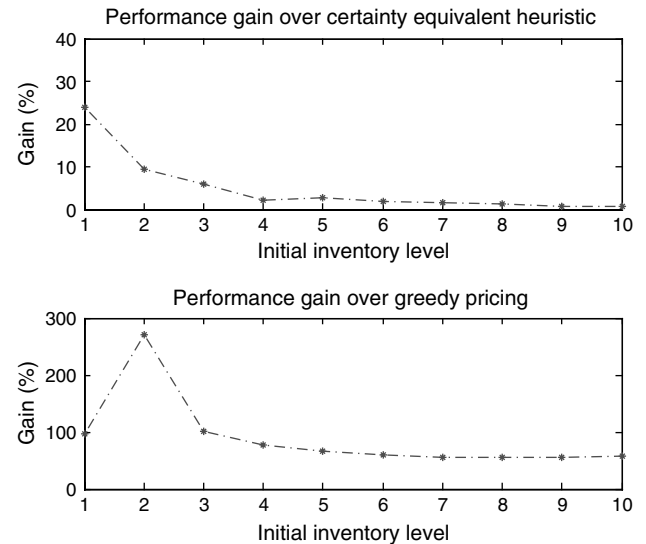

2. *Dependence on initial uncertainty*: The performance gains offered by the decay balancing heuristic are higher at higher initial levels of uncertainty. We present a lower bound on the maximal performance gain over the greedy and certainty equivalent heuristics for various coefficients of variation of an initial Gamma prior on λ. See Figure 3, wherein the data point for each coefficient of variation c corresponds to an experiment with $a = 1/c^2$, $b = 0.001$ (which corresponds to a mean of $1{,}000/c^2$), and an inventory level of 1 and 2 for the certainty equivalent and greedy pricing heuristics, respectively. These experiments indicate that the potential gain from using decay balancing increases with increasing uncertainty in λ, and that in fact the gain over certainty equivalence can be as much as a factor of 1.3 and over the greedy policy can be as much as a factor of 4.

3. *Value of "active" learning*: We examine the "learning gains" offered by decay balancing relative to the certainty equivalent and greedy heuristics. In particular, we do this by comparing the gains offered by all three pricing heuristics relative to a naive no-learning heuristic that, at every point in time, prices assuming that the arrival rate λ is equal to the initial prior mean. We observe that these gains are higher at lower initial levels of uncertainty. As before, we consider a Gamma prior with mean 40 and coefficient of variation 5 and measure performance for a range of starting inventory levels for each of the three heuristics as also the no-learning heuristic. We define the learning gain relative to the certainty equivalent heuristic

Table 2. Relative learning gains offered by decay balancing.

Inventory	$x_0 = 1$	$x_0 = 2$	$x_0 = 3$	$x_0 = 4$	$x_0 = 5$	$x_0 = 6$	$x_0 = 7$	$x_0 = 8$	$x_0 = 9$	$x_0 = 10$
π_{ce}	∞	∞	109%	26.2%	27.7%	16.8%	13.1%	10.0%	5.7%	4.4%

Figure 2. Performance gain over the certainty equivalent and greedy pricing heuristics at various inventory levels for Gamma mixture prior.
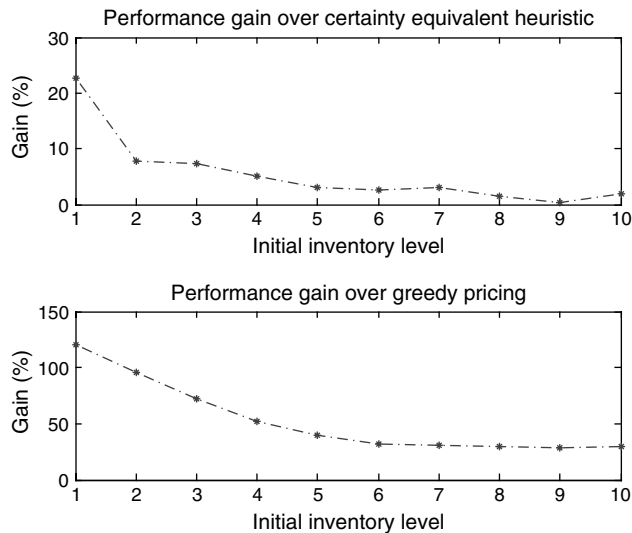
as $100 \times ((J^{\pi_{db}} - J^{\pi_{nl}})/(J^{\pi_{ce}} - J^{\pi_{nl}})^+) - 100\%$ and that relative to the greedy pricing heuristic is defined similarly. Table 2 summarizes these relative learning gains. Because the greedy pricing policy consistently underperforms the no-learning policy, we only report learning gains relative to the certainty equivalent heuristic. In addition to noting that the decay balancing heuristic offers consistent gains relative to the no-learning heuristic, the data reported suggest that the decay balancing heuristic likely captures a substantial portion of the gains to be had from using a learning scheme relative to a naive heuristic that does not learn.

To summarize our computational experience with Gamma and Gamma mixture priors, we observe that even at high levels of uncertainty in market response, the decay balancing heuristic offers a level of performance not far from that of a clairvoyant algorithm. Moreover, in all our computational experiments, the decay balancing heuristic dominates both the certainty equivalent and the greedy pricing heuristics; this is especially so at high levels of uncertainty in the initial prior and at high "load factors," i.e., scenarios where judicious inventory management is important. One final issue is robustness—in its quest to intelligently leverage uncertainty in market response, the greedy heuristic experiences a drastic loss in performance. With this issue in mind, the following section provides a performance analysis that rules out the drastic performance decay experienced with greedy pricing and sheds some light on the critical determinants of performance for our pricing problem.

6. Bounds on Performance Loss

For the decay balancing price to be a good approximation to the optimal price at a particular state, one requires only a good approximation to the value function at that state (and *not* its derivatives). This section characterizes the quality of our approximation to J^* and uses such a characterization to ultimately bound the performance loss incurred by decay balancing relative to optimal pricing. Our analysis will focus primarily on the case of a Gamma prior and exponential reservation prices (although we will also provide performance guarantees for other types of reservation price distributions). A rigorous proof of the existence and uniqueness of a solution to the HJB equation for this case may be found in the online appendix. We will show that in this case, decay balancing captures at least 33.3% of the expected revenue earned by the optimal algorithm for all

Figure 3. Lower bound on maximal performance gain over the certainty equivalent heuristic (left) and greedy heuristic (right) for various coefficients of variation.
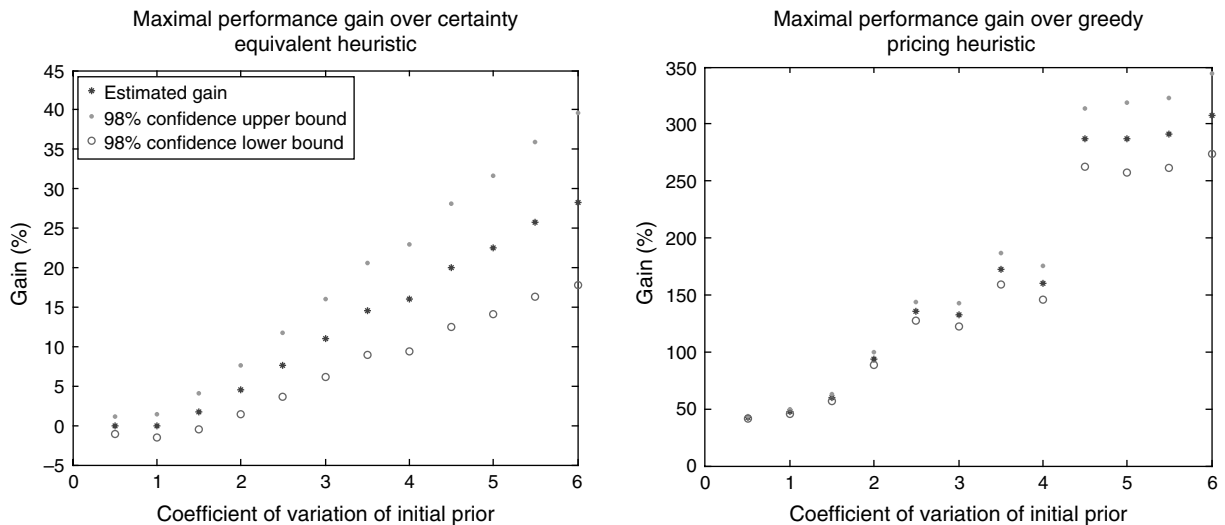
choices of $x_0 > 1$, $a_0 > 0$, $b_0 > 0$, $\alpha > 0$, and $r > 0$ when reservation prices are exponentially distributed with mean $r > 0$. Such a bound is an indicator of robustness across all parameter regimes. Decay balancing is the first heuristic for problems of this type for which a uniform performance guarantee is available.

Before we launch into the proof of our performance bound, we present an overview of the analysis. Because our analysis will focus on a Gamma prior, we will suppress the state variable w in our notation, and a and b will be understood to be scalars. Without loss of generality, we will restrict attention to problems with $\alpha = e^{-1}$; in particular, the value function exhibits the following invariance where the notation $J^{\pi, \alpha}$ makes the dependence of the value function on α explicit.

LEMMA 5. *Let $\pi\colon \mathcal{S} \to \mathbb{R}_+$ be an arbitrary policy and let $\pi'\colon \mathcal{S} \to \mathbb{R}_+$ be defined according to $\pi'(x, a, b) = \pi(x, a, b/\alpha)$. Then, for all $z \in \mathcal{S}$, $\alpha > 0$, $J^{\pi, \alpha}(z) = J^{\pi', 1}(x, a, \alpha b)$, and, in particular, $J^{*, \alpha}(z) = J^{*, 1}(x, a, \alpha b)$.*

Via the above lemma, we see that any performance bound established for a heuristic assuming a discount factor e^{-1} applies to other discount factors $\alpha > 0$ as well. In particular, given a heuristic policy $\widetilde{\pi}$ designed for discount factor e^{-1} and satisfying $J^{\widetilde{\pi}, e^{-1}}(z) \leqslant \beta J^{*, e^{-1}}(z)$ for some $z \in \mathcal{S}$, the above lemma tells us that the policy $\widetilde{\pi}'$ defined according to $\widetilde{\pi}'(x, a, b) = \widetilde{\pi}(x, a, \alpha b/e^{-1})$ satisfies $J^{\widetilde{\pi}', \alpha}(x, a, e^{-1}b/\alpha) \leqslant \beta J^{*, \alpha}(x, a, e^{-1}b/\alpha)$.

As a natural first step, we attempt to find upper and lower bounds on $\pi_{\mathrm{db}}(z)/\pi^*(z)$, the ratio of the decay balancing price in a particular state to the optimal price in that state. We are able to show that $1 \geqslant J^*(z)/\tilde{J}(z) \geqslant 1/\kappa(a)$, where $\kappa(\cdot)$ is a certain decreasing function. By then specializing attention to specific reservation price distributions, this suffices to establish that $1/f(\kappa(a)) \leqslant \pi_{\mathrm{db}}(z)/\pi^*(z) \leqslant 1$, where f is some increasing, nonnegative function dependent upon the reservation price distribution under consideration.

For example, in the case of exponential reservation price distributions, $f(x) = 1 + \log x$.

By considering a certain system under which revenue is higher than the optimal revenue, we then use the bound above and a dynamic programming argument to show that $1/f(\kappa(a)) \leqslant J^{\pi_{\mathrm{db}}}(z)/J^*(z) \leqslant 1$, where $J^{\pi_{\mathrm{db}}}(z)$ denotes the expected revenue earned by the decay balancing heuristic starting in state z. If z is a state reached after i sales, then $a = a_0 + i > i$, so that the above bound guarantees that the decay balancing heuristic will demonstrate performance that is within a factor of $f(\kappa(i))$ of optimal moving forward after i sales.

Our general performance bound can be strengthened to a uniform bound in the special case of exponential reservation prices. In particular, a coupling argument that uses a refinement of the general bound above along with an analysis of the maximal loss in revenue up to the first sale

for exponential reservation prices, establishes the uniform bound $1/3 \leqslant J^{\pi_{\mathrm{db}}}(z)/J^*(z) \leqslant 1$.

We begin our proof with a simple dynamic programming result that we will have several opportunities to use. The proof is a consequence of Dynkin's formula and can be found in the online appendix.

LEMMA 6. *Let $J \in \mathcal{J}$ satisfy $J(0, a, b) = 0$, let $\tau = \inf\{t\colon J(z_t) = 0\}$, and let $z_0 \in \mathcal{S}_{\tilde{x}, \tilde{a}, \tilde{b}}$. Then,*

$$E\left[\int_0^\tau e^{-\alpha t} H^\pi J(z_t)\, dt \right] = J^\pi(z_0) - J(z_0).$$

Let $J\colon \mathbb{N} \to \mathbb{R}$ be bounded and satisfy $J(0) = 0$, let $\tau = \inf\{t\colon J(x_t) = 0\}$, and let $x_0 \in \mathbb{N}$. Then,

$$E\left[\int_0^\tau e^{-\alpha t} H_\lambda^\pi J(x_t)\, dt \right] = J_\lambda^\pi(x_0) - J(x_0).$$

6.1. Decay Balancing vs. Optimal Prices

As discussed in the preceding outline, we will establish a lower bound on $J^*(z)/\tilde{J}(z)$ to establish a lower bound on $\pi_{\mathrm{db}}(z)/\pi^*(z)$. Let $J^{nl}(z)$ be the expected revenue garnered by a pricing scheme that does not learn, upon starting in state z. Delaying a precise description of this scheme for just a moment, we will have $J^{nl}(z) \leqslant J^*(z) \leqslant \tilde{J}(z) \leqslant J_{a/b}^*(x)$. It follows that $J^{nl}(z)/J_{a/b}^*(x) \leqslant J^*(z)/\tilde{J}(z)$, so that a lower bound on $J^{nl}(z)/J_{a/b}^*(x)$ is also a lower bound on $J^*(z)/\tilde{J}(z)$. We will focus on developing a lower bound on $J^{nl}(z)/J_{a/b}^*(x)$.

Upon starting in state z, the "no-learning" scheme assumes that $\lambda = a/b = \mu$ and does not update this estimate over time. Assuming we begin with a prior of mean μ, such a scheme would use a pricing policy given implicitly by

$$\pi^{nl}(z) = \pi_\mu^*(x) = 1/\rho(\pi_\mu^*(x)) + J_\mu^*(x) - J_\mu^*(x - 1). \quad (6)$$

Using the definition of $H^{\pi^{nl}}$ and the fact that $H_\mu J_\mu^* = 0$, some simplification yields

$$H_\lambda^{\pi^{nl}} J_\mu^*(x) = (\lambda/\mu - 1)\alpha J_\mu^*(x).$$

The following two results are then essentially immediate consequences of Lemma 6; proofs can be found in the online appendix.

LEMMA 7. *If $\lambda < \mu$, $J_\lambda^{\pi^{nl}}(x) \geqslant (\lambda/\mu) J_\mu^*(x)$ for all $x \in \mathbb{N}$.*

LEMMA 8. *If $\lambda \geqslant \mu$, $J_\lambda^{\pi^{nl}}(x) \geqslant J_\mu^*(x)$ for all $x \in \mathbb{N}$.*

Armed with these two results, we can establish a lower bound on $J^{nl}(z)/J_{a/b}^*(x)$.

THEOREM 1. *For all $z \in \mathcal{S}$,*

$$\frac{J^{nl}(z)}{J_{a/b}^*(x)} \geqslant \frac{\Gamma(a+1) - \Gamma(a+1, a) + a\Gamma(a, a)}{a\Gamma(a)} \equiv 1/\kappa(a).$$

PROOF. Setting $\mu = a/b$, we have

$$
\begin{aligned}
J^{nl}(z) &= E_\lambda[J_\lambda^{\pi^{nl}}(x)] \\
&\geqslant E_\lambda[1_{\lambda < \mu}\lambda/\mu J_\mu^*(x) + 1_{\lambda \geqslant \mu}J_\mu^*(x)] \\
&= \frac{\Gamma(a+1) - \Gamma(a+1, a) + a\Gamma(a, a)}{a\Gamma(a)}J_\mu^*(x),
\end{aligned}
$$

where the inequality follows from the two preceding lemmas and the equality by direct integration of the Gamma(a, b) density. $\Gamma(\cdot, \cdot)$ is the incomplete Gamma function and is given by $\Gamma(x, y) = \int_y^\infty s^{x-1}e^{-s}\,ds$. $\square$

The decay balance equation allows one to use the above bound on the quality of our approximation $\tilde{J}$ to compute bounds on the decay balance price relative to the optimal price at a given state. In particular, Corollary 1 establishes such bounds for exponential and logit reservation price distributions; see the online appendix for bounds one might establish for general classes of reservation price distributions.

COROLLARY 1. *For all $z \in \mathcal{S}$, and exponential reservation price distributions with parameter r,*

$$
\frac{1}{1 + \log \kappa(a)} \leqslant \frac{\pi_{db}(z)}{\pi^*(z)} \leqslant 1.
$$

For all $z \in \mathcal{S}$, and logit reservation price distributions with parameter r,

$$
\frac{1.27}{1.27 + \log \kappa(a)} \leqslant \frac{\pi_{db}(z)}{\pi^*(z)} \leqslant 1.
$$

6.2. An Upper Bound on Performance Loss

We next establish a lower bound on $J^{\pi_{db}}(z)/J^*(z)$ that will depend on the coefficient of variation of the prior on λ, $1/\sqrt{a}$.

Let

$$
R^{db}(z) = \sum_{k: t_k \leqslant \tau} e^{-e^{-1}t_k}\pi_{db}(z_{t_k})
$$

be the revenue under the decay balancing policy for a particular sample path of the sales process, starting in state z, and define

$$
R^{ub}(z) = \sum_{k: t_k \leqslant \tau} e^{-e^{-1}t_k}\pi^*(z_{t_k^-}).
$$

This describes a system whose state evolution is identical to that under the decay balancing policy but whose revenues on a sale correspond to those that would be earned if the price set prior to the sale was that of the optimal pricing algorithm.

Of course, $J^{\pi_{db}}(z) = E_z[R^{db}(z)]$. Define $J^{ub}(z) = E_z[R^{ub}(z)]$, where the expectation is over $\{t_k\}$ and assumes that an arriving consumer at time t_k makes a purchase with probability $\bar{F}(\pi_{db}(z_{t_k}))$. That is, the expectation is understood to be according to the dynamics of the system con-

trolled by π_{db}. The following result should be intuitive given our construction of the upper-bounding system and the fact that because $\pi_{db}(z) \leqslant \pi^*(z)$, the probability that a customer arriving in state z chooses to purchase is higher in the system controlled by the decay balancing policy. The proof uses a dynamic programming argument and may be found in the online appendix.

LEMMA 9. *For all $z \in \mathcal{S}$, and reservation price distributions satisfying Assumptions 1 and 2,*

$$
J^{ub}(z) \geqslant J^*(z).
$$

Now observe that because $\kappa(a)$ is decreasing in a, we have from Corollary 1 that

$$
\frac{1}{f(\kappa(a))} \leqslant \frac{R^{db}(z)}{R^{ub}(z)} \leqslant 1,
$$

where $f(x) = 1 + \log(x)$ for exponential reservation price distributions and $f(x) = 1 + \log(x)/1.27$ for Logit distributions. Taking expectations, and employing Lemma 9, we then immediately have the following theorem.

THEOREM 2. *For all $z \in \mathcal{S}$, and exponential reservation price distributions with parameter r,*

$$
\frac{1}{1 + \log \kappa(a)} \leqslant \frac{J^{\pi_{db}}(z)}{J^*(z)} \leqslant 1,
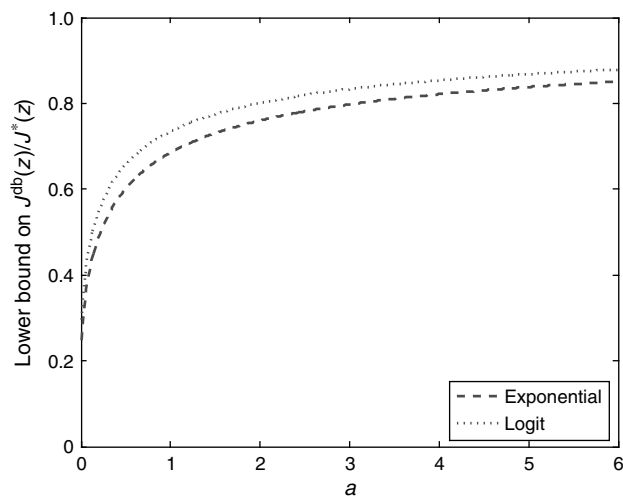$$

while for logit reservation price distributions with parameter r,

$$
\frac{1.27}{1.27 + \log \kappa(a)} \leqslant \frac{J^{\pi_{db}}(z)}{J^*(z)} \leqslant 1.
$$

See Figure 4 for an illustration of these bounds.

It is worth pausing to reflect on the bound we have just established. Our performance bound does *not* depend on x,

Figure 4. Lower bound on decay balancing performance.

b, α, or the parameters of the reservation price distribution. Because the coefficient of variation of a Gamma prior with parameters a and b is given by $1/\sqrt{a}$, the bound illustrates how decay balancing performance approaches optimal performance as the coefficient of variation of the initial prior is decreased; below coefficients of variation of 0.5, the bound guarantees performance levels that are at least within 80% of optimal. Nonetheless, the bound can be arbitrarily poor at high coefficients of variation. Next, we further specialize our attention to exponential reservation price distributions and present a uniform performance guarantee for that case.

6.3. A Uniform Performance Bound for Exponential Reservation Prices

We now consider the case of exponentially distributed prices. In particular, we have $\bar{F}(p) = \exp(-p/r)$, where $r > 0$. In light of the following lemma (where the notation $J^{\pi, \alpha, r}$ makes the dependence of the value function on α and r explicit), we can assume without loss of generality that the mean reservation price, r, is 1.

LEMMA 10. *Let $\pi: \mathscr{S} \to \mathbb{R}_+$ be an arbitrary policy and let $\pi': \mathscr{S} \to \mathbb{R}_+$ be defined according to $\pi'(z) = (1/r)\pi(z)$. Then, for all $z \in \mathscr{S}$, $\alpha > 0$, $r > 0$, $J^{\pi, \alpha, r}(z) = r J^{\pi', \alpha, 1}(z)$, and, in particular, $J^{*, \alpha, r}(z) = r J^{*, \alpha, 1}(z)$.*

Our proof of a uniform performance bound will use Theorem 2 along with a coupling argument to bound performance loss up to the time of the first sale. Begin by considering the following coupling (a superscript "db" on a variable indicates that the variable is relevant to a system controlled by the π_{db} policy). For an arbitrary policy $\pi(\cdot) \in \Pi$, the sales processes n_t^{db} and n_t^{π} are coupled in the following sense: Denote by $\{t_k\}$ the points of the Poisson process corresponding to customer *arrivals* (not sales) to both systems. Assume that $\pi_{\mathrm{db}\,t_k^-} \leqslant \pi_{t_k^-}$. Then, a jump in n^{π} at time t_k can occur if and only if a jump occurs in n^{db} at time t_k. Furthermore, conditioned on a jump in n^{db} at t_k, the jump in n^{π} occurs with probability $\exp(-(\pi_{t_k^-} - \pi_{\mathrm{db}\,t_k^-}))$. The situation is reversed if $\pi_{\mathrm{db}\,t_k^-} > \pi_{t_k^-}$. Let τ denote the time of the first sale for the π_{db} system, i.e., $\tau = \inf\{t: n_t^{\mathrm{db}} = 1\}$. In the context of this coupling, consider the optimal (i.e., π^*) and π_{db} controlled systems. Denoting

$$J^*(z \mid \tau) = E\left[\int_{t=0}^{\infty} e^{-\alpha t} \pi^*(z_t) \lambda \bar{F}(\pi^*(z_t))\, dt \mid \tau, z_0 = z\right],$$

we then have the following lemma.

LEMMA 11. *For all $z \in \mathscr{S}$,*

$$J^*(z \mid \tau) \leqslant e^{-e^{-1}\tau}\big(e^{-(\pi^* - \pi_{\mathrm{db}})}[\pi^* + J^*(x-1, a+1, b_\tau^{\mathrm{db}})] + (1 - e^{-(\pi^* - \pi_{\mathrm{db}})})J^*(x, a+1, b_\tau^{\mathrm{db}})\big),$$

where $\pi^* = \pi^*(x, a, b_\tau^*)$ and $\pi_{\mathrm{db}} = \pi_{\mathrm{db}}(x, a, b_\tau^{\mathrm{db}})$.

The result above is essentially a consequence of the fact that it is never the case that the π^* controlled system sells its first item before the π_{db} system, and moreover, that conditioning on τ, and the information available in both systems up to τ^-, yields a posterior with shape parameter $a + 1$ and scale parameter b_τ^{db}. We will also find the following technical lemma useful.

LEMMA 12. *For $x > 1$, $a > 1$, $b > 0$, $J^*(x, a, b) \leqslant 2.05 J^*(x-1, a, b)$.*

The result above is intuitive. It would follow, for example, from decreasing returns to an additional unit of inventory. Unfortunately, we are not able to show such a "decreasing returns" property directly and a certain coupling argument is used to prove the lemma (see the online appendix). We are now poised to prove a uniform (over $x > 1$) performance bound for our pricing scheme.

THEOREM 3. *For all $z \in \mathscr{S}$ with $x > 1$,*

$$\frac{J^{\pi_{\mathrm{db}}}(z)}{J^*(z)} \geqslant 1/3.$$

PROOF. In Lemma 11, we showed

$$J^*(z) \leqslant E[e^{-e^{-1}\tau}(e^{-(\pi^* - \pi_{\mathrm{db}})}[\pi^* + J^*(x-1, a+1, b_\tau^{\mathrm{db}})]$$
$$+ (1 - e^{-(\pi^* - \pi_{\mathrm{db}})})J^*(x, a+1, b_\tau^{\mathrm{db}}))].$$

Now,

$$e^{-e^{-1}\tau}(e^{-(\pi^* - \pi_{\mathrm{db}})}[\pi^* + J^*(x-1, a+1, b_\tau^{\mathrm{db}})]$$
$$+ (1 - e^{-(\pi^* - \pi_{\mathrm{db}})})J^*(x, a+1, b_\tau^{\mathrm{db}}))$$
$$\leqslant e^{-e^{-1}\tau}(e^{-(\pi^* - \pi_{\mathrm{db}})}\pi^* + J^*(x, a+1, b_\tau^{\mathrm{db}}))$$
$$\leqslant e^{-e^{-1}\tau}(e^{-(\pi^* - \pi_{\mathrm{db}})}\pi^* + 2.05 J^*(x-1, a+1, b_\tau^{\mathrm{db}}))$$
$$\leqslant e^{-e^{-1}\tau}(\pi_{\mathrm{db}} + 2.05(1 + \log\kappa(a+1))J^{\pi_{\mathrm{db}}}(x-1, a+1, b_\tau^{\mathrm{db}}))$$
$$\leqslant e^{-e^{-1}\tau}2.05(1 + \log\kappa(a+1))(\pi_{\mathrm{db}} + J^{\pi_{\mathrm{db}}}(x-1, a+1, b_\tau^{\mathrm{db}})),$$

where the first inequality is because J^* is nondecreasing in x. The second inequality follows from Lemma 12. The third inequality follows from the fact that $\pi^* \geqslant \pi_{\mathrm{db}} \geqslant 1$ so that $\pi^* e^{-\pi^*} \leqslant \pi_{\mathrm{db}} e^{-\pi_{\mathrm{db}}}$ and from Theorem 2. Finally, taking expectations of both sides, we get

$$E[e^{-e^{-1}\tau}(e^{-(\pi^* - \pi_{\mathrm{db}})}[\pi^* + J^*(x-1, a+1, b_\tau^{\mathrm{db}})]$$
$$+ (1 - e^{-(\pi^* - \pi_{\mathrm{db}})})J^*(x, a+1, b_\tau^{\mathrm{db}}))]$$
$$\leqslant 2.05(1 + \log\kappa(a+1))$$
$$\cdot E[e^{-e^{-1}\tau}(\pi_{\mathrm{db}} + J^{\pi_{\mathrm{db}}}(x-1, a+1, b_\tau^{\mathrm{db}}))]$$
$$\leqslant 2.05(1 + \log\kappa(1))J^{\pi_{\mathrm{db}}}(z).$$

Thus,

$$\frac{J^{\pi_{\mathrm{db}}}(z)}{J^*(z)} \geqslant \frac{1}{2.05(1 + \log\kappa(1))} \geqslant 1/3. \quad \square$$

Perhaps the most crucial point borne out in the performance analysis we have just presented is that the decay balancing heuristic is *robust*; our analysis precludes the drastic performance breakdown observed for the greedy pricing heuristic in our computational experiments.

7. Multiple Stores and Consumer Segments

We now explore extensions of decay balancing to a model with multiple stores and consumer segments. We do not attempt to extend our performance analysis to this more general model but instead present numerical experiments, the goal being to show that decay balancing demonstrates the same qualitative behavior as in the one-store—one-customer segment case we have studied to this point.

More formally, we consider a model with N stores and M consumer segments. Each store is endowed with an initial inventory $x_{0,i}$ for $i \in \{1, \ldots, N\}$. Customers from class j, for $j \in \{1, \ldots, M\}$, arrive according to a Poisson process of rate λ_j, where λ_j is a Gamma distributed random variable with shape parameter $a_{0,j}$ and scale parameter $b_{0,j}$. An arriving segment j customer considers visiting a single store and will consider store i with probability α_{ij}. Consequently, each store i sees a Poisson stream of customers having rate $\sum_j \alpha_{ij} \lambda_j$. We assume without loss of generality that $\sum_i \alpha_{ij} = 1$. We assume that customers in each segment have exponential reservation price distributions with mean r, and moreover, that upon a purchase the store has a mechanism in place to identify what segment the purchasing customer belongs to.

Let $p_t \in \mathbb{R}^N$, $t \in [0, \infty)$ represent the process of prices charged at the stores over time. Let $n_{t,i}^j$ represent the total number of type j customers served at store i up to time t, and let $n_t^j = \sum_i n_{t,i}^j$. The parameter vectors a and b are then updated according to

$$a_{t,j} = a_{0,j} + n_{t,j} \quad \text{and} \quad b_{t,j} = b_{0,j} + \int_{\tau=0}^{t} \sum_i e^{-p_{\tau,i}/r} \, d\tau.$$

Our state at time t is now $z_t = (x_t, a_t, b_t)$. As before, we will consider prices generated by policies π that are measurable, nonnegative vector-valued functions of state, so that $p_t = \pi(z_t) \geqslant 0$. Letting Π denote the set of all such policies, our objective will be to identify a policy $\pi^* \in \Pi$ that maximizes

$$\hat{J}^\pi(z) = E_{z,\pi} \left[\sum_i \int_0^{\tau^i} \hat{\rho}_{t,i} e^{-p_{t,i}/r} \, dt \right],$$

where $\tau^i = \inf\{t: \sum_j n_{t,i}^j = x_{0,i}\}$ and $\hat{\rho}_i = \sum_j \alpha_{i,j}(a_j/b_j)$. We define the operator

$$(\tilde{H}^\pi J)(z) = \sum_i \left[\hat{\rho}_i e^{-\pi(z)_i/r} \left(\pi(z)_i + \sum_j \frac{\alpha_{i,j}(a_j/b_j)}{\hat{\rho}_i} \right.\right.$$
$$\left.\left. \cdot J(x - e_i, a + e_j, b) - J(z) \right) \right.$$
$$\left. + \sum_j e^{-\pi(z)_i/r} \frac{d}{db_j} J(z) \right] - \alpha J(z),$$

where e_k is the vector that is 1 in the kth coordinate and 0 in other coordinates. One might show that $\hat{J}^* = \hat{J}^{\pi^*}$ is the unique solution to

$$\sup_{\pi \in \Pi} (\tilde{H}^\pi J)(z) = 0 \quad \forall z,$$

satisfying $\hat{J}^*(0, a, b) = 0$, and that the corresponding optimal policy for $x_i > 0$ is given by

$$(\pi^*(z))_i = r + \hat{J}^*(z) - \sum_j \frac{\alpha_{i,j}(a_j/b_j)}{\hat{\rho}_i} \hat{J}^*(x - e_i, a + e_j, b)$$
$$- \frac{1}{\hat{\rho}_i} \sum_j \frac{d}{db_j} \hat{J}^*(z). \tag{7}$$

Now, assuming that the λ_js are known perfectly a priori, it is easy to see that the control problem decomposes across stores. In particular, the optimal strategy simply involves store i using as its pricing policy $p_{t,i} = \pi^*_{\rho_i}(x_{t,i})$, where $\rho_i = \sum_j \alpha_{i,j} \lambda_j$. Consequently, a certainty equivalent policy would use the pricing policy $(\pi_{CE}(z))_i = \pi^*_{\hat{\rho}_i}(x_i)$.

We can also consider as an approximation to $\hat{J}^*$ the following upper bound (which is in the spirit of the upper bound we derived in §4):

$$\bar{J}(z) = E \left[\sum_i J^*_{\rho_i}(x_i) \right].$$

The analogous greedy pricing policy π_{gp} is then given by (7) upon substituting $\bar{J}(\cdot)$ for $\hat{J}^*(\cdot)$ in that expression.

Motivated by the decay balancing policy derived for the single-store case, we consider using the following pricing policy at each store:

$$(\pi_{\text{db}}(z))_i = r \log \left(\frac{r \hat{\rho}_i}{\alpha E[J^*_{\rho_i}(x_i)]} \right).$$

The above pricing equation assumes that, moving forward, each store will operate as a separate entity. Nonetheless, the heuristic incorporates joint learning across stores and continues to account for the level of uncertainty in market size in the pricing process. Furthermore, the structural properties discussed in the single-store case are retained.

Figure 5. Performance relative to a clairvoyant algorithm.

Performance gain over clairvoyant algorithm
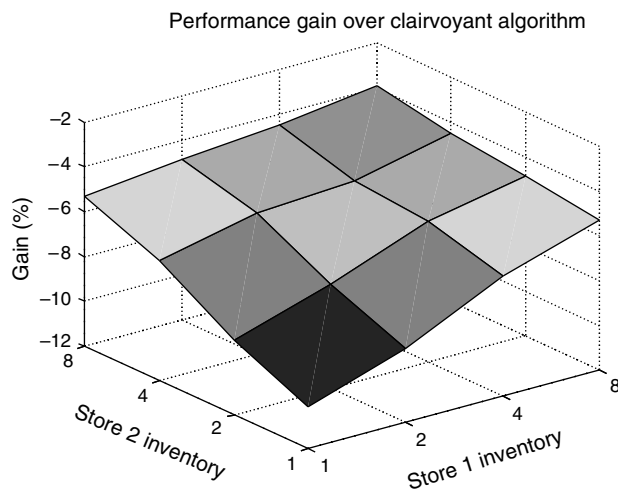
Figure 6.   Performance relative to the certainty equivalent heuristic (left) and greedy pricing heuristic (right).
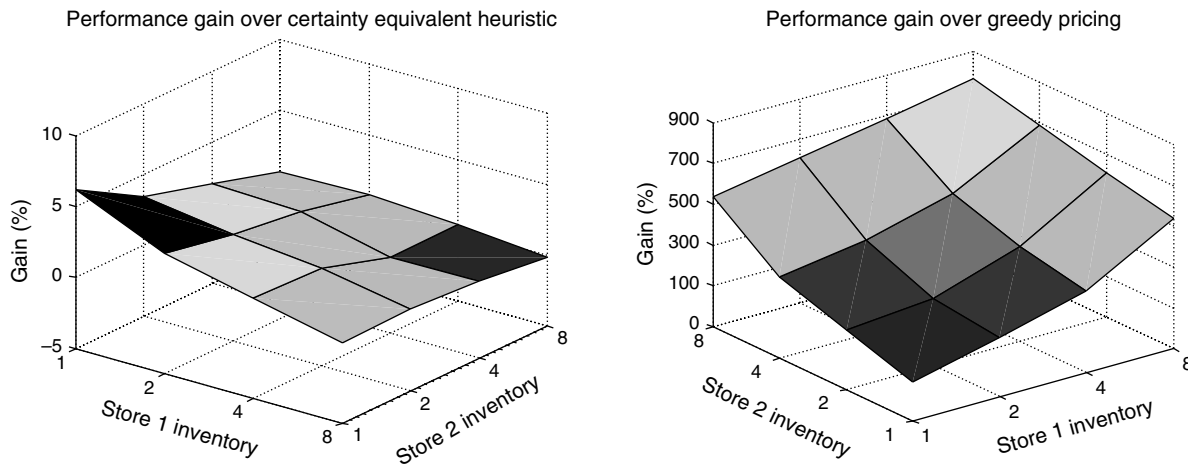
The joint learning under this heuristic does not, however, account for inventory levels across stores.

We now present computational results for the three heuristics. Our experiments will use the following model parameters. We take $N = 2$, $M = 2$ and assume that $\alpha_{i,j} = 1/2$ for all i, j, and further that we begin with prior parameters $a_1 = a_2 = 0.04$ and $b_1 = b_2 = 0.001$ (which corresponds to a mean of 40 and a coefficient of variation of 5). As usual, $\alpha = e^{-1}$, $r = 1$. Our first set of results (Figure 5) compares the decay balancing heuristic's performance against that of a clairvoyant algorithm which as in §5 has perfect a priori knowledge of λ. As in the $N = 1$, $M = 1$ case, our performance is quite close to that of the clairvoyant algorithm. Figure 6 compares decay balancing performance to the certainty equivalent heuristic and the greedy heuristic. Figure 6 is indicative of performance that is qualitatively similar to that observed for the $N = 1$, $M = 1$ case; there is a significant gain over certainty equivalence at lower inventory levels, but this gain shrinks as inventory level increases. The performance of the greedy heuristic is particularly dismal; one explanation for which is that $\sum_j (d/db_j) \bar{J}(z)$ is a potentially poor approximation to $\sum_j (d/db_j) \tilde{J}^*(z)$.

8. Discussion and Conclusions

The dynamic pricing model proposed by Gallego and van Ryzin (1994) is central to a large body of the revenue management literature. This work considered an important extension to that model. In particular, we considered incorporating uncertainty in the customer arrival rate or "market response," which is without doubt important in many industries that practice revenue management.

We proposed and analyzed a simple new heuristic for this problem: decay balancing. Decay balancing is computationally efficient and leverages the solution to problems with no uncertainty in market response. Our computational experiments (which focused on Gamma priors and exponentially distributed reservation prices) suggest that decay balancing

achieves near-optimal performance even on problems with high levels of uncertainty in market response. Pricing policies generated by decay balancing have the appealing structural property that, all other factors remaining the same, the price in the presence of uncertainty in market response is higher than the corresponding price with no uncertainty. This is reasonable from an operational perspective and is in fact a property possessed by the optimal policy. Our analysis also demonstrated a uniform performance guarantee for decay balancing when reservation prices are exponentially distributed, which is an indicator of robustness.

Two heuristics proposed for problems of this nature prior to our work were the certainty equivalent heuristic (by Aviv and Pazgal 2005a) and the greedy pricing heuristic (by Araman and Caldentey 2009). Our computational results suggest that decay balancing offers significant performance advantages over these heuristics. These advantages are especially clear at high levels of uncertainty in market response, which is arguably the regime of greatest interest. Decay balancing relies only on a good approximation to the value of an optimal policy at a given state. This is in contrast with greedy pricing that requires a good approximation not only to value but further to derivatives of value with respect to the scale parameter. At the same time, uncertainty in arrival rate and changes in inventory levels bear the appropriate directional impact on decay balancing prices: uncertainty in the arrival rate calls for higher prices than in corresponding situations with no uncertainty, while a decrease in inventory calls for an increase in prices. In contrast, uncertainty in the arrival rate has no impact on certainty equivalent prices, while greedy prices can increase or decrease with decreasing inventory.

Our computational study and performance analysis were focused on exponentially distributed reservation prices and gamma priors, but we expect favorable performance for other distributions as well. In particular, the analysis of Theorem 1 can be extended to Gamma mixture priors yielding encouraging estimates on the quality of approximation

provided by $\tilde{J}$. Because the decay balancing price at state z is likely to be a good approximation to the optimal price at z if $\tilde{J}(z)$ is a good approximation to $J^*(z)$, this suggests that decay balancing is likely to do a good job of approximating the optimal price for general reservation price distributions and priors on arrival rate, which in turn should lead to superior performance.

There is ample room for further work in the general area of pricing with uncertainty in market response and other factors that impact demand. One direction is considering more complex models. We considered applying the decay balancing heuristic to a problem with multiple stores and consumer segments; our computational results there are promising. There are other models one might hope to consider; for example, the multiproduct model proposed by Gallego and van Ryzin (1997). Another potential direction is exploring new approximations to the value function beyond the approximation considered here and applying such approximations with either the greedy pricing or decay balancing heuristics. Finally, it would be interesting to extend the approaches in this paper to problems with uncertainty in other factors that impact demand, such as price elasticity.

9. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://or.journal.informs.org/.

Acknowledgments

References

Araman, V. F., R. Caldentey. 2009. Dynamic pricing for nonperishable products with demand learning. *Oper. Res.* **57**(5) 1169–1188.

Aviv, Y., A. Pazgal. 2005a. Pricing of short life-cycle products through active learning. Working paper, Washington University, St. Louis.

Aviv, Y., A. Pazgal. 2005b. A partially observed Markov decision process for dynamic pricing. *Management Sci.* **51**(9) 1400–1416.

Bertsimas, D., G. Perakis. 2006. Dynamic pricing: A learning approach. D. Hearn, S. Lawphongpanich, eds. *Mathematical and Computational Models for Congestion Charging*. Springer, Berlin, 45–79.

Besbes, O., A. Zeevi. 2006. Blind nonparametric revenue management: Asymptotic optimality of a joint learning and pricing method. Submitted.

Bremaud, P. 1981. *Point Processes and Queues: Martingale Dynamics*, 1st ed. Springer-Verlag, Berlin.

Burnetas, A. N., C. E. Smith. 1998. Adaptive ordering and pricing for perishable products. *Oper. Res.* **48**(3) 436–443.

Cope, E. 2006. Bayesian strategies for dynamic pricing in e-commerce. *Naval Res. Logist.* **54**(3) 265–281.

Dalal, S. R., W. J. Hall. 1983. Approximating priors by mixtures of natural conjugate priors. *J. Roy. Statist. Soc. B* **45**(2) 278–286.

Gallego, G., G. van Ryzin. 1994. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Sci.* **40**(8) 999–1020.

Gallego, G., G. van Ryzin. 1997. A multiproduct dynamic pricing problem and its applications to network yield management. *Oper. Res.* **45**(1) 24–41.

Kleinberg, R., F. T. Leighton. 2004. The value of knowing a demand curve: Bounds on regret for on-line posted-price auctions. *Proc. 44th IEEE Sympos. Foundations Comput. Sci.* (*FOCS 2003*), 594–605.

Lin, K. Y. 2006. Dynamic pricing with real-time demand learning. *Eur. J. Oper. Res.* **174**(1) 522–538.

Lobo, M. S., S. Boyd. 2003. Pricing and learning with uncertain demand. Submitted.

Talluri, K. T., G. J. van Ryzin. 2004. *The Theory and Practice of Revenue Management*. Springer Science+Business Media, Berlin.