# Bayesian Dynamic Pricing with Two-Sided Censored Customer Willingness-to-Pay Data

Li Chen

*The Fuqua School of Business, Duke University, Durham, NC 27708*
*li.chen@duke.edu*

---

We consider a Bayesian dynamic pricing problem with an unknown distribution of customer willingness to pay (WTP). The posted price serves as either a left- or right-censoring point of the customer's true WTP. This two-sided censoring effect makes the Bayesian dynamic pricing problem difficult to analyze. Little is known in the literature about how to solve this problem in its general form. In this paper, we develop computational strategies to tackle this challenge. We first quantify the loss of informational value due to censoring and derive the first-order derivative for the problem. Leveraging on these results, we derive a sequence of lower and upper bounds for the value function of the problem. We then propose three easy-to-compute heuristics based on the bound results. Our numerical studies show that one of the proposed heuristics is near optimal in all cases. We further discuss several interesting insights obtained from our analytical and numerical analysis.

---

Key words: Dynamic pricing, two-sided censored observations, partially-observable Markov decision processes, Bayesian dynamic programming.

History: Revised on January 15, 2014.

## 1. Introduction

Imagine that you have a few spare tickets for a sold-out basketball game. Each ticket has a face value of \$120. The game is starting in less than two hours. There are quite a few interested fans wanting to buy the tickets from you. How would you price your tickets? Would you price them at an exorbitant rate of \$600 or at a more moderate rate of \$160? Airline revenue managers and online retailers face similar pricing problems every day (see Talluri and van Ryzin 2004). More recently, many professional sports teams have begun the practice of dynamically adjusting ticket prices based on demand; Forbes magazine has ranked "dynamic pricing" as the No. 3 sports business story to watch in 2012 (Rishe 2011). Empirically, Sweeting (2012) found that sellers increasingly use dynamic pricing strategies in secondary markets for Major League Baseball tickets and that dynamic pricing raises the average seller's expected payoff by around 16%.

Clearly, the seller's pricing decision would depend on how much time and how many tickets are left. The closer it gets to the game start time, the more the price would drop; and the smaller the number of tickets available, the higher the price is likely to rise (see Gallego and van Ryzin 1994 for a formal analysis). Moreover, things would become a little easier if the distribution of each customer's willingness to pay (WTP) is known. Unfortunately, however, this is usually not the case. Worse, no customer is willing to reveal her actual WTP for the tickets. All one can do is to learn from the observed purchase decisions: if a customer buys the ticket, one learns that her WTP is greater than or equal to the asking price; if she does not buy the ticket, her WTP must be below the asking price. In other words, the asking price serves as either a left- or right-censoring point of the customer's true WTP.

In this paper, we consider a Bayesian dynamic pricing problem with an unknown customer WTP distribution, in which we need to take into account the two-sided censoring effect as described above. To isolate the customer WTP distribution uncertainty from the arrival rate uncertainty, we assume the customer arrival rate is known in our problem (see a discussion in §1.1 about the difference of these two types of uncertainty). We formulate this dynamic pricing problem as a finite-horizon partially-observable Markov decision process (POMDP). Because no simple conjugate prior exists under two-sided censoring (see Braden and Freimer 1991), the optimal solution to the problem is difficult to compute due to the curse of dimensionality. Lazear (1986) provided a formal analysis of the problem for a special case in which each customer has the same (unknown) fixed WTP. When we relax this assumption and allow customer WTP to be a random draw from a general probability distribution, the problem becomes much more complex. Kandel (1990, Chapter 1) formulated the problem for selling one unit of inventory, but was only able to provide a solution approach for a simple two-period case. More recently, Cope (2007) considered the case in which the available inventory quantity is unlimited and relied on approximate posteriors to solve the problem. Beyond these special cases, little is known in the literature about how to solve this problem in its general form. The main goal of this paper is thus to develop computational strategies to tackle this challenge.

We make five contributions to the literature. First, we quantify the loss of informational value due to censoring and derive the first-order derivative for the problem in its general form. Leveraging on these results, we derive a sequence of lower and upper bounds for the value function of the original problem. The upper bound systems assume that the seller obtains either exact or right-censored observations of customer WTP after observing certain number of left-censored observations; as a result, such systems achieve higher expected revenue than the original system with either left- or right-censored observations throughout the decision horizon (value of better information). Symmetrically, the lower bound systems assume that the seller employs a suboptimal pricing policy after a certain future period; as a result, such systems achieve lower expected revenue than the original system with two-sided censoring (value of better decision).

Second, we propose three easy-to-compute heuristics based on the value function bounds. The first heuristic, termed the "exact-observation heuristic," is the optimal solution to a special upper bound system where the seller obtains only exact customer WTP observations in every period. This heuristic is easy to compute under conjugate prior distributions because it assumes away the two-sided censoring issue. The second heuristic, termed the "one-step myopic heuristic," is the optimal solution to a lower bound system where the seller accounts for two-sided censoring in the first period and then employs a myopic pricing policy (based on the updated prior) afterwards. This heuristic is easy to compute because it reduces the original Bayesian dynamic program to a two-stage optimization problem; the second stage (the myopic policy stage) problem only involves a simple forward evaluation. The third heuristic, termed the "one-step dynamic heuristic," has a similar structure to the one-step myopic heuristic. In particular, we replace the myopic pricing policy in the second stage with a dynamic pricing policy that accounts for the marginal value of inventory capacity (which is a better approximation to the optimal policy than the myopic policy). This heuristic is also an optimal solution to a lower bound system for the original problem. It requires a little more computational effort than the one-step myopic heuristic because the second stage involves a dynamic pricing policy, but is still fairly easy to compute.

Third, we obtain some interesting qualitative results by comparing the exact-observation heuristic with the optimal solution. When there is one unit inventory available for sale, the original problem only involves left-censored observations. In this case, we show that the optimal solution is less than the exact-observation heuristic due to the motive for reducing the left-censoring effect. This insight is consistent with the "price low to obtain information" insight obtained by Braden and Oren (1994). When the available inventory is abundant, the "price low" result may, however, reverse to "price high"—the optimal solution becomes greater than the exact-observation heuristic. In particular, we show the "price high" result holds in a two-period problem with exponential distribution, suggesting that reducing the right-censoring effect is more valuable than reducing left-censoring in this case.

Fourth, for illustrative purposes, we provide detailed computational formulas under the exponential distribution. The exponential distribution case is a common case studied in the literature (see for example, Gallego and van Ryzin 1994 and Aviv and Pazgal 2005a). When the initial prior follows a gamma distribution, we show that both the optimal solution and the heuristics possess a scaling property that enables dimensionality reduction in computations. This scaling property is reminiscent of those discovered in Bayesian inventory control problems that feature either no-censoring or right-censoring (Scarf 1960; Azoury 1985; Lariviere and Porteus 1999; Bisi et al. 2011). Here we extend the scaling result to the Bayesian dynamic pricing problem with two-sided censoring. In §5, we further generalize the result to the Weibull distribution family.

Finally, we conduct extensive numerical studies to evaluate and compare the performance of our proposed heuristics. Our numerical results show that the exact-observation and the one-step dynamic heuristics perform better than the heuristic suggested in Aviv and Pazgal (2005a) (their heuristic assumes that the seller acquires full knowledge of customer WTP distribution after the pricing decision). Among the three heuristics proposed in this paper, the one-step dynamic heuristic is near optimal in all cases. As a robustness check, we conduct an additional numerical study based on the normal distributions with a two-point prior. The results again show that the one-step dynamic heuristic is near optimal in all cases.

Our numerical results also suggest that when inventory is in short supply and customer WTP is unknown (such as in the sporting event tickets selling example), one should use more sophisticated heuristics, such as the one-step dynamic heuristic, for pricing decisions to maximize total revenue, because the stake in pricing correctly is high in this case. When, however, inventory is abundant or customer demand is predictable, a simple myopic pricing decision or a dynamic pricing decision without Bayesian learning may just be sufficient, because the marginal value of inventory capacity or the value of learning is low in this case. Another interesting and also very comforting insight from our numerical results is that active learning under two-sided censoring helps keep price gouging behavior in check. When available inventory runs low, a seller equipped with better information, such as in the full-information and exact-observation scenarios, may engage in price gouging to extract the maximum value from such information. In contrast, under two-sided censoring, if the price is set too high and no purchase is observed, the seller learns little about the customer WTP. To avoid such a situation, a learning seller would instead set the price at a more moderate level.

### 1.1 Literature Review

There is an extensive literature on dynamic pricing and revenue management (see Bitran and Caldentey 2003 and Elmaghraby and Keskinocak 2003 for comprehensive reviews). However, active learning under two-sided censoring has not received sufficient attention in this literature. To date, the literature has primarily focused on problems with an unknown arrival rate and a known distribution of customer WTP (e.g., Aviv and Pazgal 2005ab, Araman and Caldentey 2009, and Farias and Van Roy 2010). In such a case, the seller observes the aggregated customer purchase quantity during a period and the two-sided censoring effect is masked by the aggregated data. When we instead assume that the customer arrival rate is known and the distribution of customer WTP is unknown, such as in this paper, the seller observes individual customers' binary purchase decisions and the two-sided censoring effect emerges.

Our paper is closely related to Aviv and Pazgal (2005a). As mentioned earlier, we compare the performance between our heuristics and the heuristic suggested in their paper. However, our model differs from theirs in three aspects. First, we allow for a continuous set for the (unknown) distribution parameter, whereas they considered a finite discrete set. Second, the decision period in our model is finer than in their model. We only allow for one customer arrival per period, whereas they allowed for multiple customer arrivals in a period. The finer decision period in our model gives rise to the two-sided censoring effect. Third, for ease of exposition, we assume that the customer WTP is independently and identically-distributed (i.i.d.), whereas they studied a richer Markov-modulated demand model. In §5 we note that our results can also be generalized to the Markov-modulated process.

In the dynamic pricing literature, both continuous-time models (e.g., Kincaid and Darling 1963; Gallego and van Ryzin 1994; Zhao and Zheng 2000) and discrete-time models (e.g., Bitran and Mondschein 1997; Aviv and Pazgal 2005a) have been studied. Our model belongs to the latter category. Our learning framework is in the category of parametric Bayesian models (e.g., Balvers and Cosimano 1990; Aviv and Pazgal 2005ab; Araman and Caldentey 2009; Farias and Van Roy

2010; Harrison et al. 2012). For other learning models, such as maximum likelihood estimation, least squares, and nonparametric models, we refer the readers to Carvalho and Puterman (2005), Bertsimas and Perakis (2006), Lim and Shanthikumar (2007), Besbes and Zeevi (2009), Keskin and Zeevi (2013), and Wang et al. (2013).

Our POMDP formulation extends the classic formulation by allowing for additional observable state variables, such as the available inventory capacity, in the decision process (see Monahan 1982; Lovejoy 1993). The two-sided censoring process studied in this paper also generalizes the one-sided censoring problems studied in the Bayesian inventory control literature (e.g., Harpaz et al. 1982, Lariviere and Porteus 1999, Ding et al. 2002, Lu et al. 2005, 2008, Bensoussan et al. 2005, 2007, 2008, 2009, Bisi and Dada 2007, Chen and Plambeck 2008, Chen 2010, and Bisi et al. 2011).

The remainder of the paper is organized as follows. We first introduce the problem formulation in §2. We then present general analysis of the problem and value function bounds and heuristics in §3. In §4, we provide detailed computation formulas and results from our numerical studies. §5 contains extensions and our concluding remarks. All proofs are presented in the Appendix.

## 2. Problem Formulation

Consider a finite-horizon dynamic pricing problem for a single product. Inventory replenishment is not possible during the selling horizon, and the terminal value at the end of the horizon is zero. At the beginning of a period, given the available inventory quantity $q$, the seller determines the unit price $p$ for the product. The goal is to maximize expected total revenue over the finite horizon.

Specifically, we divide the finite selling horizon into $T$ periods such that there is only one customer arrival possible during a period (e.g., Talluri and van Ryzin 2004, p. 203). We index the time periods in reverse order, with the first period being period $T$ and the last selling period being period 1. Let $m_t$ denote the (time-varying) probability of a customer arrival in period $t$. For example, $m_t$ can be interpreted as the arrival probability in a discretized nonhomogeneous Poisson arrival process (e.g., Bitran and Mondschein 1997). We assume that $m_t$ is known to the seller.

Each arriving customer has her own willingness to pay, which is a random draw from an i.i.d. distribution with a continuous density $f(x|\theta)$, where $x$ ($x \geq 0$) is the actual WTP and $\theta$ ($\theta \in \Theta$) is the unknown parameter of the distribution, such as the mean of the distribution. The seller has a prior belief concerning the value of $\theta$ at the beginning of each period $t$, which is denoted by $\pi_t(\theta)$. In what follows, we assume that $\Theta$ is a continuous set and $\pi_t(\theta)$ is a density over this set. When $\Theta$ is a discrete set, all our analysis will carry through by treating $\pi_t(\theta)$ as a probability mass function.

Given the price decision $p$ in period $t$, the customer purchase decision $Z_t$ is binary ($Z_t = 0$ means no purchase; and $Z_t = 1$ means purchase):

$$Z_t = \begin{cases} 0 & \text{if } 0 \leq X_t < p, \\ \\ 1 & \text{if } p \leq X_t < \infty, \end{cases}$$

where $X_t$ is the (random) customer WTP.

Because the customer WTP is not directly observed, the seller has to infer it from the observed purchase decision. Following Braden and Freimer (1991), we call the observation of a purchase decision $Z_t = 0$ the left-censored observation, because in this case we know only that the actual WTP is less than the posted price $p$. Symmetrically, we call the observation of $Z_t = 1$ the right-censored observation, because we know only that the actual WTP is greater than or equal to $p$. Clearly, in this problem the posted price $p$ serves as either a left- or a right-censoring point, depending on the observed purchase decision. Based on the above definition, the likelihood of an observation $Z_t = z$, with $z \in \{0, 1\}$, can be written as

$$k(Z_t = z|\theta) = \begin{cases} F(p|\theta) & \text{if } z = 0, \\ \\ \bar{F}(p|\theta) & \text{if } z = 1, \end{cases} \tag{1}$$

where $F(p|\theta) = \int_0^p f(x|\theta)dx$ and $\bar{F}(p|\theta) = 1 - F(p|\theta)$. Given the observation $Z_t = z$, the updated prior of $\theta$ at the beginning of period $t-1$, according to Bayes' rule, is given by

$$\pi_{t-1}(\theta) = \frac{k(Z_t = z|\theta) \cdot \pi_t(\theta)}{\int_\Theta k(Z_t = z|\theta) \cdot \pi_t(\theta)d\theta}, \tag{2}$$

where $\pi_t(\theta)$ is the prior at the beginning of period $t$. In what follows, we shall use $\pi_t$ and $\pi_t(\theta)$ interchangeably and suppress the subscript $t$ whenever appropriate within the context.

Given a price $p$ and a prior $\pi$, the expected single-period revenue can be written as

$$R(p, \pi) = p \int_p^\infty \int_\Theta f(x|\theta)\pi(\theta)d\theta dx.$$

We note that if the predictive distribution $\int_\Theta f(x|\theta)\pi(\theta)d\theta$ has an increasing generalized failure rate (IGFR), then the single-period revenue function $R(p, \pi)$ is unimodal, and the revenue-maximizing price is unique (see Ziya et al. 2004, Lariviere 2006). The IGFR condition is met if, for instance, $f(x|\theta)$ is an exponential distribution with $\theta$ either known or subject to a gamma prior distribution.

The dynamic pricing problem specified above belongs to the class of Markov decision processes with incomplete information. According to Dynkin and Yushkevich (1979, pp. 214-217), this problem can be reduced to an equivalent model with complete information by replacing the unobservable state $\theta$ with its prior distribution $\pi(\theta)$. Furthermore, based on our problem definition, it can be verified that the model satisfies the conditions of the general Borel model (Dynkin and Yushkevich 1979, pp. 46-47). Thus, the problem can be formulated by the following dynamic programming optimality equations: For $t = T, ..., 1$,

$$\begin{aligned} V_t(q,\pi) &= \max_{p\geq 0}\Bigg\{ R(p,\pi)\cdot m_t + V_{t-1}\left(q, \frac{F(p|\cdot)\pi}{\int_\Theta F(p|\theta)\pi(\theta)d\theta}\right)\cdot \int_\Theta F(p|\theta)\pi(\theta)d\theta \cdot m_t \\ &\quad + V_{t-1}\left(q-1, \frac{\bar{F}(p|\cdot)\pi}{\int_\Theta \bar{F}(p|\theta)\pi(\theta)d\theta}\right)\cdot \int_\Theta \bar{F}(p|\theta)\pi(\theta)d\theta \cdot m_t + V_{t-1}(q,\pi)\cdot(1-m_t)\Bigg\}, \end{aligned}$$

with $V_t(0,\cdot) = 0$ and $V_0(\cdot,\cdot) = 0$. In the above optimality equation, the first term is the current-period reward, the second and third terms are the expected total future rewards when the current-period observation is left- and right-censored, respectively, and the last term is the expected total future rewards when there is no arrival in the current period (and thus the posterior is not updated). Note that the last term does not depend on the pricing decision. With some rearrangement of terms, we have

$$\begin{aligned} V_t(q,\pi) &= m_t\cdot\max_{p\geq 0}\Bigg\{ R(p,\pi)\cdot + V_{t-1}\left(q, \frac{F(p|\cdot)\pi}{\int_\Theta F(p|\theta)\pi(\theta)d\theta}\right)\cdot \int_\Theta F(p|\theta)\pi(\theta)d\theta \\ &\quad + V_{t-1}\left(q-1, \frac{\bar{F}(p|\cdot)\pi}{\int_\Theta \bar{F}(p|\theta)\pi(\theta)d\theta}\right)\cdot \int_\Theta \bar{F}(p|\theta)\pi(\theta)d\theta\Bigg\} \\ &\quad + V_{t-1}(q,\pi)\cdot(1-m_t). \end{aligned}$$

From the above expression, it is clear that the optimal pricing decision in period $t$ does not depend on the (known) arrival probability $m_t$. Thus, without loss of generality, we shall focus on the the normalized case with $m_t \equiv 1$ henceforth (for the general $m_t$ case, we can simply adjust the value function obtained in the normalized case accordingly). The normalized case is given as follows: for $t = T, ..., 1$,

$$\begin{aligned} V_t(q,\pi) &= \max_{p\geq 0}\{G_t(p,q,\pi)\} \\ &= \max_{p\geq 0}\Bigg\{ R(p,\pi) + V_{t-1}\left(q, \frac{F(p|\cdot)\pi}{\int_\Theta F(p|\theta)\pi(\theta)d\theta}\right)\cdot \int_\Theta F(p|\theta)\pi(\theta)d\theta \\ &\quad + V_{t-1}\left(q-1, \frac{\bar{F}(p|\cdot)\pi}{\int_\Theta \bar{F}(p|\theta)\pi(\theta)d\theta}\right)\cdot \int_\Theta \bar{F}(p|\theta)\pi(\theta)d\theta\Bigg\}, \end{aligned} \tag{3}$$

with $V_t(0,\cdot) = 0$ and $V_0(\cdot,\cdot) = 0$. Let $p_t^*$ denote the optimal solution to the above problem.

Lazear (1986) considered a special case of the above problem in which each customer has the same fixed (unknown) WTP $\theta$. In other words, the WTP distribution is a singleton, i.e.,

$f(x|\theta) = \delta(x - \theta)$, where $\delta(x - \theta)$ is the Dirac delta function. Under this assumption, based on the customer purchase decision, the seller learns that the underlying (fixed) WTP is either above or below the posted price. Thus, the range of possible values of $\theta$ is narrowed after each observation and the dynamic program can be solved fairly easily. When we relax the fixed WTP assumption and allow the WTP to be a random draw from a general probability distribution, the problem, as defined above in (3), becomes much more complex. Kandel (1990, Chapter 1) was able to provide a solution approach for a simple two-period problem with one unit inventory for sale. But beyond that special case, little is known in the literature about how to solve this problem in its general form. The main goal of this paper is thus to develop computational strategies to tackle this challenge.

Let us first introduce a transform technique to simplify the dynamic programming formulation defined in (3). From the Bayesian updating formula (2), let us define an unnormalized prior $\tilde{\pi}_{t-1}$ as

$$\tilde{\pi}_{t-1}(\theta) = k(Z_t = z|\theta) \cdot \tilde{\pi}_t(\theta)$$

for $t = T, ..., 2$, with $\tilde{\pi}_T(\theta) = \pi_T(\theta)$. It is straightforward to show that $\tilde{\pi}_t = \pi_t \cdot \int_\Theta \tilde{\pi}_t(\theta)d\theta$. Thus, $\tilde{\pi}_t(\theta)$ is a derived measure on $\Theta$. Based on this measure, we redefine the single-period revenue as

$$\tilde{R}(p, \tilde{\pi}) = p \int_p^\infty \int_\Theta f(x|\theta)\tilde{\pi}(\theta)d\theta dx,$$

and the optimality equations as: for $t = T, ..., 1$,

$$\begin{aligned} \tilde{V}_t(q, \tilde{\pi}) &= \max_{p \geq 0} \left\{ \tilde{G}_t(p, q, \tilde{\pi}) \right\} \\ &= \max_{p \geq 0} \left\{ \tilde{R}(p, \tilde{\pi}) + \tilde{V}_{t-1}(q, F(p|\cdot)\tilde{\pi}) + \tilde{V}_{t-1}(q-1, \bar{F}(p|\cdot)\tilde{\pi}) \right\}, \end{aligned} \tag{4}$$

with $\tilde{V}_t(0, \cdot) = 0$ and $\tilde{V}_0(\cdot, \cdot) = 0$. The following lemma establishes the equivalence between (4) and the original formulation (3):

**Lemma 1.** *For $t = T, ..., 1$, $\tilde{G}_t(p, q, \tilde{\pi}) = G_t(p, q, \pi) \cdot \int_\Theta \tilde{\pi}(\theta)d\theta$ and $\tilde{V}_t(q, \tilde{\pi}) = V_t(q, \pi) \cdot \int_\Theta \tilde{\pi}(\theta)d\theta$.*

From the above lemma, it is clear that the dynamic program defined in (4) is a scaled version of the original program (3). Thus, they are equivalent in the sense that both programs share the same optimal solution. This unnormalized-prior transform technique was first used by Bensoussan et al. (2005, 2007, 2008, 2009) to study various Bayesian inventory control problems that involve right-censored observations. Here we apply the technique to a Bayesian dynamic pricing problem with two-sided censoring. In what follows, we shall focus only on the problem defined by (4) for our analysis, and we shall drop the tilde above all functions including $\tilde{\pi}$ in the formulation to reduce notation.

## 3. Dynamic Program Analysis

In this section, we first derive some general properties for the dynamic program (4). Based on these properties, we develop a series of upper and lower bounds for the value function. These bound results are then used to devise effective heuristics for the problem.

To begin with our analysis, we first define the following function:

$$\begin{aligned}\Delta_t(q,\pi',\pi) &= V_t(q,\pi) - G_t(\hat{p},q,\pi) + \Delta_{t-1}(q, F(\hat{p}|\cdot)\pi', F(\hat{p}|\cdot)\pi) \\ &\quad + \Delta_{t-1}(q-1, \bar{F}(\hat{p}|\cdot)\pi', \bar{F}(\hat{p}|\cdot)\pi), \qquad (5)\end{aligned}$$

with $\Delta_t(0,\cdot,\cdot) = 0$, $\Delta_0(\cdot,\cdot,\cdot) = 0$, and $\hat{p} = \arg\max_{p\geq 0}\{G_t(p,q,\pi')\}$ (when $\hat{p}$ is not unique, we pick the solution with the largest value). The function $\Delta_t(q,\pi',\pi)$ can be interpreted as the difference between the value function $V_t(q,\pi)$ and the expected revenue when a suboptimal pricing policy, computed along each sample path assuming a prior $\pi'$, is evaluated based on the true system prior $\pi$. Thus, by a simple backward induction, it is straightforward to verify that $\Delta_t(q,\pi',\pi) \geq 0$ and the equality holds when $\pi' = \pi$.

With the generalized envelope theorem of Milgrom and Segal (2002) and a backward induction, we can establish the following result:

**Theorem 1.** *The value function $V_t(q, F(p|\cdot)\pi)$ is absolutely continuous in $p$ (implying it is differentiable in $p$ almost everywhere) and can be written as*

$$V_t\left(q, F(p|\cdot)\pi\right) = \int_0^p V_t(q, f(x|\cdot)\pi)dx - \int_0^p \Delta_t\left(q, F(x|\cdot)\pi, f(x|\cdot)\pi\right)dx,$$

*where $\Delta_t$ is defined in (5). Symmetrically, $V_t(q,\bar{F}(p|\cdot)\pi)$ is also absolutely continuous in $p$, with*

$$V_t\left(q, \bar{F}(p|\cdot)\pi\right) = \int_p^\infty V_t(q, f(x|\cdot)\pi)dx - \int_p^\infty \Delta_t\left(q, \bar{F}(x|\cdot)\pi, f(x|\cdot)\pi\right)dx.$$

The term $\int_0^p \Delta_t\left(q, F(x|\cdot)\pi, f(x|\cdot)\pi\right)dx$ in the above result can be viewed as the loss of informational value when $p$ is a left-censored observation—it quantifies the informational value of having an exact observation as compared to having a left-censored observation. Similarly, the term $\int_p^\infty \Delta_t\left(q, \bar{F}(x|\cdot)\pi, f(x|\cdot)\pi\right)dx$ quantifies the loss of informational value when $p$ is a right-censored observation. It is worth commenting here that in our dynamic pricing model the left- and right-censoring points happen to coincide at the price point $p$. When these points do not coincide, the result of Theorem 1 still holds under some proper modification of the definition of $\Delta_t$. For problems involving only the right-censored observations such as those studied in the Bayesian inventory control literature, Theorem 1 can be used to prove the well-known "stock more" results (e.g., Harpaz et al. 1982; Lariviere and Porteus 1999; Ding et al. 2002; Lu et al. 2005, 2008; Bensoussan et al. 2007, 2009; and Chen and Plambeck 2008). The following corollary provides a sufficient condition to guarantee the differentiability of $G_t(p,q,\pi)$:

**Corollary 1.** *If* $\arg\max_{p\geq 0}\{G_i(p,q,\pi)\}$ *is unique for any given* $q$*,* $\pi$*, and* $i \leq t-1$*, then* $G_t(p,\cdot,\cdot)$ *is differentiable in* $p$*, with the first-order derivative given by*

$$\begin{aligned} G_t'(p,q,\pi) &= R'(p,\pi) + V_{t-1}(q, f(p|\cdot)\pi) - V_{t-1}(q-1, f(p|\cdot)\pi) \\ &\quad -\Delta_{t-1}\left(q, F(p|\cdot)\pi, f(p|\cdot)\pi\right) + \Delta_{t-1}\left(q-1, \bar{F}(p|\cdot)\pi, f(p|\cdot)\pi\right), \end{aligned}$$

*where* $\Delta_{t-1}$ *is defined in (5).*

In the above derivative expression, the first term is the current-period marginal revenue, the second and third terms are the expected marginal value of capacity, and the fourth and fifth terms are the expected loss of informational value due to the two-sided censoring. From this result, we can get a sense of the computational challenges in the problem. First, as discussed earlier, when the predictive distribution is an IGFR distribution, the single-period revenue function $R(p,\pi)$ has a unique maximizer. Thus, the condition of Corollary 1 is met for $i = 1$. For cases with $i > 1$, it is generally difficult to establish specific conditions to ensure the uniqueness of the optimal solution. Thus, the differentiability of the objective function may not always be guaranteed. Second, even if the differentiability is guaranteed, computing the $\Delta_{t-1}$ terms in the derivative expression is still cumbersome: one needs to determine the optimal solution $\hat{p}$ along each sample path according to the expression (5). Third, and most importantly, we cannot leverage on the conjugate prior structure to reduce the dimensionality of the dynamic program, because no simple conjugate prior exists under two-sided censoring (see Braden and Freimer 1991). In §4.1.1, we will demonstrate that the dimensionality of the problem increases linearly in the length of the decision horizon when customer WTP follows an exponential distribution with a gamma prior. For these reasons, we need to resort to approximation techniques to reduce the computational complexity of the problem.

## 3.1 Value Function Upper Bounds

From Theorem 1, we observe that there exists informational value of having exact observations. Intuitively, the value function of the two-sided censoring problem should be bounded above by the value function of a system in which customer WTP is observed exactly. Let us first define such an exact-observation system: for $t = T, ..., 1$,

$$V_t^e(q,\pi) = \max_{p\geq 0}\{G_t^e(p,q,\pi)\}$$

$$= \max_{p\geq 0}\left\{R(p,\pi)+\int_0^p V^e_{t-1}(q,f(x|\cdot)\pi)dx+\int_p^\infty V^e_{t-1}(q-1,f(x|\cdot)\pi)dx\right\}, \quad (6)$$

with $V^e_t(0,\cdot)=0$ and $V^e_0(\cdot,\cdot)=0$. Let $p^e_t$ be the solution to the above dynamic program (the superscript "$e$" stands for exact observation in this case). It is easy to verify that $G^e_t(p,q,\pi)$ is differentiable in $p$, and its derivative is given by

$$G^{e\prime}_t(p,q,\pi)=R'(p,\pi)+V^e_{t+1}(q,f(p|\cdot)\pi)-V^e_{t+1}(q-1,f(p|\cdot)\pi).$$

Thus, the optimal solution $p^e_t$ can be computed easily based on the first-order condition. Moreover, because there is no data-censoring in the formulation, we can leverage on the conjugate prior structure to further reduce the problem dimensionality (see §4.1.2 for a detailed illustration).

We can also define a system with either exact or right-censored customer WTP observations as follows: for $t=T,...,1$,

$$\begin{aligned} V^r_t(q,\pi) &= \max_{p\geq 0}\{G^r_t(p,q,\pi)\} \\ &= \max_{p\geq 0}\left\{R(p,\pi)+\int_0^p V^r_{t-1}(q,f(x|\cdot)\pi)dx+V^r_{t-1}(q-1,\bar{F}(p|\cdot)\pi)\right\}, \end{aligned} \quad (7)$$

with $V^r_t(0,\cdot)=0$ and $V^r_0(\cdot,\cdot)=0$ (the superscript "$r$" stands for right-censoring only in this case). Symmetrically, we can define a system with either exact or left-censored observations. According to Braden and Freimer (1991), the conjugate prior distribution family for exact and right-censored observations is much broader than that for exact and left-censored observations. Thus, for computational purposes, we shall focus only on the above right-censoring only system. By Theorem 1, we can establish the following result:

**Proposition 1.** *For any given $q$ and $\pi$, the following holds: $V_T(q,\pi)\leq V^r_T(q,\pi)\leq V^e_T(q,\pi)$.*

The above proposition confirms the intuition that a system with better information achieves higher total revenue: the exact-observation system achieves higher revenue than the right-censoring only system, and the latter achieves higher revenue than the two-sided censoring system. Thus, $V^r_T(q,\pi)$ serves as a tighter upper bound for $V_T(q,\pi)$ than does $V^e_T(q,\pi)$. To further improve the value function upper bound, we can define the following dynamic program based on the right-censoring only system (7): for $t=T,...,1$, let $n_t$ be the number of left-censored observations obtained by the beginning of period $t$,

$$\begin{aligned} V^{u_i}_t(q,\pi) &= \max_{p\geq 0}\{G^{u_i}_t(p,q,\pi)\} \\ &= \begin{cases} \max_{p\geq 0}\left\{R(p,\pi)+V^{u_i}_{t-1}(q,F(p|\cdot)\pi)+V^{u_i}_{t-1}(q-1,\bar{F}(p|\cdot)\pi)\right\} & \text{if} \quad n_t<i, \\ \max_{p\geq 0}\{G^r_t(p,q,\pi)\}=V^r_t(q,\pi) & \text{if} \quad n_t\geq i, \end{cases} \end{aligned} \quad (8)$$

with $V_t^{u_i}(0,\cdot) = 0$ and $V_0^{u_i}(\cdot,\cdot) = 0$. According to the above definition, $V_t^{u_i}(q,\pi)$ is the value function of a system in which either exact or right-censored customer WTP observations become available after the seller observed $i$ left-censored observations. In the case of $i = T$, we have $V_t^{u_T}(q,\pi) = V_t(q,\pi)$, so the above dynamic program reduces to the original one. The following proposition establishes that $V_T^{u_i}(q,\pi)$ forms a sequence of upper bounds for the original value function $V_T(q,\pi)$:

**Proposition 2.** *For any given $q$ and $\pi$, the following holds:*

*(a)* $G_T(p,q,\pi) = G_T^{u_T}(p,q,\pi) \leq ... \leq G_T^{u_1}(p,q,\pi) \leq G_T^{u_0}(p,q,\pi) = G_T^r(p,q,\pi)$*;*

*(b)* $V_T(q,\pi) = V_T^{u_T}(q,\pi) \leq ... \leq V_T^{u_1}(q,\pi) \leq V_T^{u_0}(q,\pi) = V_T^r(q,\pi)$*.*

The sequence of value function upper bounds is obtained by assuming the seller acquires either exact or right-censored customer WTP observations after observing certain number of left-censored observations. From the proposition, we observe that as $i$ increases from 0 to $T$, the value function upper bound $V_T^{u_i}$ becomes tighter, but the computational complexity also increases. The least tight but also the easiest to compute upper bound is $V_T^{u_0}(q,\pi)$, or equivalently $V_T^r(q,\pi)$. Based on our numerical experience in §4, the one-step upper bound $V_T^{u_1}(q,\pi)$ is much tighter than $V_T^r(q,\pi)$, with only a moderate increase of computational complexity.

### 3.2 Value Function Lower Bounds

Now let us derive the lower bound for the value function $V_t(q,\pi)$. Intuitively, we can compute the expected system revenue under a suboptimal pricing policy and use it as a lower bound. Let $p_t^s(q,\pi)$ be a predetermined pricing policy based on the system state $(q,\pi)$. The expected system revenue under such a policy is given by: for $t = T, ..., 1$,

$$\hat{V}_t(q,\pi) = R(p_t^s(q,\pi),\pi) + \hat{V}_{t-1}(q, F(p_t^s|\cdot)\pi) + \hat{V}_{t-1}(q-1, \bar{F}(p_t^s|\cdot)\pi), \tag{9}$$

where $\hat{V}_t(0,\cdot) = 0$ and $\hat{V}_0(\cdot,\cdot) = 0$. In the above expression, we note that the second and third terms are evaluated at the decision points based on the updated state variables, i.e., $p_{t-1}^s(q, F(p_t^s|\cdot)\pi)$ and $p_{t-1}^s(q-1, \bar{F}(p_t^s|\cdot)\pi)$, respectively. Given an easy-to-compute pricing policy $p_t^s$, the above revenue function can be evaluated easily by a forward iteration from period $t$ to 1.

Now based on the revenue function defined in (9), we can define the following dynamic program: for $t = T, ..., 1$,

$$V_t^{l_i}(q,\pi) = \max_{p\geq 0} \left\{ G_t^{l_i}(p,q,\pi) \right\}$$

$$= \begin{cases} \max\limits_{p\geq 0}\left\{R(p,\pi)+V_{t-1}^{l_i}(q,F(p|\cdot)\pi)+V_{t-1}^{l_i}(q-1,\bar{F}(p|\cdot)\pi)\right\} & \text{if} \quad t>T-i, \\ \max\limits_{p\geq 0}\left\{\hat{V}_t(q,\pi)\right\}=\hat{V}_t(q,\pi) & \text{if} \quad t\leq T-i, \end{cases} \tag{10}$$

with $V_t^{l_i}(0,\cdot)=0$ and $V_0^{l_i}(\cdot,\cdot)=0$. According to the above definition, $V_t^{l_i}(q,\pi)$ is the value function of a system in which a suboptimal pricing policy $p_t^s$ is employed from period $T-i$ onwards. In the case of $i=T$, we have $V_t^{l_T}(q,\pi)=V_t(q,\pi)$, so the above dynamic program reduces to the original one. The following proposition establishes that $V_T^{l_i}(q,\pi)$ forms a series of lower bounds for the original value function $V_T(q,\pi)$:

**Proposition 3.** *For any given $q$ and $\pi$, the following holds:*

*(a)* $G_T(p,q,\pi)=G_T^{l_T}(p,q,\pi)\geq ...\geq G_T^{l_1}(p,q,\pi)\geq G_T^{l_0}(p,q,\pi)=\hat{V}_T(q,\pi)$;

*(b)* $V_T(q,\pi)=V_T^{l_T}(q,\pi)\geq ...\geq V_T^{l_1}(q,\pi)\geq V_T^{l_0}(q,\pi)=\hat{V}_T(q,\pi)$.

From the proposition, we observe that as $i$ increases from 0 to $T$, the value function lower bound $V_T^{l_i}$ becomes tighter, but the computational complexity also increases. The least tight but also the easiest to compute lower bound is $V_T^{l_0}(q,\pi)$, or equivalently $\hat{V}_T(q,\pi)$ (i.e., when a predetermined suboptimal policy is employed from period $T$ to 1). Based on our numerical experience in §4, the one-step lower bound $V_T^{l_1}(q,\pi)$ is much tighter than $V_T^{l_0}(q,\pi)$, with only a moderate increase in computational complexity.

## 3.3 Heuristic Solutions

In this section, we propose and discuss three easy-to-compute heuristics based on the value function bounds, with the first one derived from the upper bound result and the second and third ones from the lower bound result.

### 3.3.1 Exact-Observation Heuristic

From our discussion in §3.1, an easy-to-compute heuristic for the two-sided censoring problem would be the optimal pricing decision $p_t^e$ defined in (6), i.e., when the customer WTPs are assumed to be observed exactly. Similar exact-observation heuristics have been proposed by Chen (2010) for an Bayesian inventory control problem with unobserved lost sales. In that problem, only right-censored observations are involved and it can be shown that the exact-observation heuristic is always less than the optimal solution. Our Bayesian dynamic pricing problem involves more complex two-sided censored observations. Below we provide two instances in which the exact-observation heuristic can be either greater or less than the optimal solution.

**Proposition 4.** *When $q = 1$, for any $T \geq 1$ and any given initial prior $\pi$, $p_T^e \geq p_T^*$. When $q \geq 2$, if the customer willingness to pay follows an exponential distribution with a gamma prior, then $p_T^e < p_T^*$ for $T = 2$.*

It is easy to verify that when $q = 1$, our problem reduces to a left-censoring only problem. Thus, one would price lower than the exact-observation heuristic to reduce the chance of a left-censored observation, and hence we have $p_T^e \geq p_T^*$. This result is consistent with the "price low to obtain information" insight from Braden and Oren (1994). However, their result was derived from a different learning model in which the seller learns from a customer's total consumption quantity rather than from the WTP observations. When $q \geq 2$, the problem involves two-sided censoring and the opposite insight may be true. In particular, in a two-period problem with exponential distribution, one would actually price higher than the exact-observation heuristic, which suggests that the informational gain from reducing right-censoring outweighs that from reducing left-censoring in this case.

### 3.3.2 One-Step Myopic Heuristic

To further improve on the exact-observation heuristic, we propose a heuristic based on the value function lower bound result. Specifically, we focus on the one-step lower bound. The motivation for the heuristic is as follows. Because the original problem is difficult to compute, to evaluate the performance of a heuristic solution $p^h$, we can use the relative error bound determined by the one-step lower bound, i.e.,

$$\frac{V_T(q,\pi) - G_T(p^h,q,\pi)}{V_T(q,\pi)} = 1 - \frac{G_T(p^h,q,\pi)}{V_T(q,\pi)} \leq 1 - \frac{G_T^{l_1}(p^h,q,\pi)}{V_T(q,\pi)},$$

where the inequality follows from Proposition 3. Thus, a heuristic that achieves the smallest error upper bound is given by $\arg\max_{p\geq 0} G_T^{l_1}(p,q,\pi)$. Based on the definition of $G_T^{l_1}(p,q,\pi)$ given in (10), such a heuristic is determined by solving the following optimization problem:

$$\max_{p\geq 0} \left\{ R(p,\pi) + \hat{V}_{T-1}(q, F(p|\cdot)\pi) + \hat{V}_{T-1}(q-1, \bar{F}(p|\cdot)\pi) \right\}, \tag{11}$$

where $\hat{V}_{T-1}(\cdot,\cdot)$ is defined by (9). To determine $\hat{V}_{T-1}(\cdot,\cdot)$, we need to define a (suboptimal) pricing policy $p_t^s$ along each sample path. The simplest policy for this purpose is the myopic pricing policy that maximizes the current-period revenue based on the updated prior along each sample path, i.e., $p_t^s(q,\pi) = \arg\max_{p>0} R(p,\pi)$. Specifically, $p_t^s(q,\pi)$ can be determined by the first-order condition:

$$R'(p,\pi) = \int_\Theta \bar{F}(p|\theta)\pi(\theta)d\theta - p\int_\Theta f(p|\theta)\pi(\theta)d\theta = 0.$$

This heuristic is essentially the optimal solution to a system in which the seller accounts for two-sided censoring in the first period and then employs a myopic pricing policy afterwards (based on the updated prior). Thus, we term this heuristic the "one-step myopic heuristic." According to Proposition 3, this heuristic attains a lower bound for the value function of the original problem.

#### 3.3.3 One-Step Dynamic Heuristic

Besides the myopic pricing policy, we can use an alternative pricing policy in determining $\hat{V}_{T-1}(\cdot,\cdot)$ in (11). Specifically, we propose to use a dynamic pricing policy $p_t^s(q,\pi)$ that is obtained by solving the following simple dynamic program: for $t,...,1$,

$$V_t^n(q,\pi) \quad = \quad \max_{p\geq 0}\left\{R(p,\pi)+V_{t-1}^n(q,\pi)\cdot\int_\Theta F(p|\theta)\pi(\theta)d\theta+V_{t-1}^n(q-1,\pi)\cdot\int_\Theta \bar{F}(p|\theta)\pi(\theta)d\theta\right\}, \tag{12}$$

with $V_t^n(0,\cdot)=0$ and $V_0^n(\cdot,\cdot)=0$. We note that the above dynamic program does not involve Bayesian updating (thus, the superscript "$n$" stands for no Bayesian updating in this case). It is easy to verify that $p_t^s(q,\pi)$ can be determined by the following first-order condition:

$$R'(p,\pi)+\left[V_{t-1}^n(q,\pi)-V_{t-1}^n(q-1,\pi)\right]\cdot\int_\Theta f(p|\theta)\pi(\theta)d\theta=0,$$

where $V_{t-1}^n(q,\pi)$ can be computed based on (12). Comparing the above equation to the first-order condition used in the one-step myopic heuristic, it is clear that $p_t^s(q,\pi)$ in this case takes into account the marginal value of inventory capacity. Thus, this pricing policy is a better approximation to the optimal policy than the myopic policy. We note that computing $p_t^s(q,\pi)$ requires a little more computational effort than the myopic policy, but is still quite easy to compute (because there is no Bayesian updating involved).

With this heuristic, we are effectively determining the optimal solution to a system in which the seller accounts for two-sided censoring in the first period and then employs a dynamic pricing policy afterwards (based on the updated prior but with no Bayesian updating in future periods). Thus, we term this heuristic the "one-step dynamic heuristic." Like the one-step myopic heuristic, according to Proposition 3, this heuristic also attains a lower bound for the value function of the original problem.

## 4. Numerical Studies

In this section, we conduct numerical studies based on the case of exponential distribution. Formally, we assume customer WTP follows an exponential density $f(x|\theta)=\theta e^{-\theta x}$, where $\theta$ is an (unknown) parameter. We choose this case for two reasons. First, this is a common case considered in the literature (see for example, Gallego and van Ryzin 1994 and Aviv and Pazgal 2005a). Second, as will be demonstrated below, we can compute the optimal solution in this case when the decision horizon is not too long. This enables us to compare the heuristics directly against the optimal solution. For illustrative purposes, we first present the detailed computational formulas for the optimal solution and the heuristic solutions in §4.1. To keep things concise, we omit much of the calculus derivation steps for the formulas; they are all easy to verify nevertheless. In §4.2, we present the numerical results based on these formulas.

## 4.1 Computational Formulas

Consider first the full information case in which the parameter $\theta$ is known. In this case, the dynamic program is given by the following: for $t = T, ..., 1$,

$$V_t^{FI}(q|\theta) = \max_{p \geq 0} \left\{ pe^{-\theta p} + (1 - e^{-\theta p}) \cdot V_{t-1}^{FI}(q|\theta) + e^{-\theta p} \cdot V_{t-1}^{FI}(q-1|\theta) \right\}, \tag{13}$$

with $V_t^{FI}(0|\cdot) = 0$ and $V_0^{FI}(\cdot|\cdot) = 0$. Bitran and Mondschein (1997) showed that the optimal solution to this case is unique and can be determined by

$$p_t^{FI}(q|\theta) = \theta^{-1} + V_{t-1}^{FI}(q|\theta) - V_{t-1}^{FI}(q-1|\theta),$$

where $V_{t-1}^{FI}(q|\theta)$ can be computed easily based on the dynamic program (13). The following scaling result can further simplify computation:

**Proposition 5.** *For any given $q > 0$ and $1 \leq t \leq T$, $p_t^{FI}(q|\theta) = \theta^{-1} p_t^{FI}(q|1)$ and $V_t^{FI}(q|\theta) = \theta^{-1} V_t^{FI}(q|1)$.*

Based on the above result, to obtain the optimal solution for any $\theta$ value, all we need to do is to solve the case with $\theta = 1$ and then scale the solution by $1/\theta$.

### 4.1.1 Optimal Solution under Two-Sided Censoring

Now consider the case where the parameter $\theta$ is unknown and the set of all possible $\theta$ is $\Theta = [0, \infty)$. Let us assume that the initial prior $\pi_T(\theta)$ follows an (unnormalized) gamma distribution given by $\pi_T(\theta) = \theta^{a-1} e^{-S\theta}$, with $a > 1$ and $S > 0$ (where $a$ is the shape parameter and $S$ the scale parameter).

If there is an exact observation $x$, the updated (unnormalized) prior becomes $f(x|\theta)\pi(\theta) = \theta^a e^{-(S+x)\theta}$, which is a gamma distribution with parameters $(a+1, S+x)$. If there is a right-censored observation $x$, the updated prior becomes $\bar{F}(x|\theta)\pi(\theta) = \theta^{a-1} e^{-(S+x)\theta}$, which is also a gamma distribution with parameters $(a, S+x)$. However, when there is a left-censored observation $x$, the updated prior becomes $F(x|\theta)\pi(\theta) = (1-e^{-\theta x})\theta^{a-1}e^{-S\theta} = \theta^{a-1}e^{-S\theta} - \theta^{a-1}e^{-(S+x)\theta}$, which is a mixture of two gamma distributions with parameters $(a, S)$ and $(a, S+x)$.

Suppose that at the beginning of period $t$, there have been $n_t$ left-censored observations $y_1$, ..., $y_{n_t}$, and $T-t-n_t$ right-censored observations $x_1$, ..., $x_{T-t-n_t}$. Let $a_t = a$, $S_t = S + \sum_{i=1}^{T-t-n_t} x_i$ and $\mathbf{y}_t = [y_1, ..., y_{n_t}]$. Then the updated prior for period $t$ is given by

$$\pi_t(\theta|a_t, S_t, \mathbf{y}_t) = \theta^{a-1}e^{-S_t\theta}\prod_{i=1}^{n_t}\left(1-e^{-y_i\theta}\right). \tag{14}$$

The above expression is a mixture of gamma distributions with the same shape parameter $a$ but different scale parameters involving $S_t$ and all possible combinations of $y_1$, ..., $y_{n_t}$. Thus, we can use $(a, S_t, \mathbf{y}_t)$ to represent the state of the prior distribution at the beginning of period $t$, with the understanding that a new left-censored observation will be added at the end of the vector $\mathbf{y}_t$ and will thus increase the vector dimensionality by one.

To facilitate our subsequent analysis, we define the following function:

$$\begin{aligned}
\varphi(p|a, S_t, \mathbf{y}_t) &= \int_p^\infty \int_\Theta f(x|\theta)\pi_t(\theta|a_t, S_t, \mathbf{y}_t)d\theta dx \\
&= \int_p^\infty \int_0^\infty \theta^a e^{-(x+S_t)\theta}\prod_{i=1}^{n_t}\left(1-e^{-y_i\theta}\right)d\theta dx \\
&= \frac{\Gamma(a)}{(p+S_t)^a} - \sum_{i=1}^{n_t}\frac{\Gamma(a)}{(p+S_t+y_i)^a} + \sum_{i=1}^{n_t}\sum_{j>i}\frac{\Gamma(a)}{(p+S_t+y_i+y_j)^a} \\
&\quad -... + \frac{(-1)^{n_t}\Gamma(a)}{(p+S_t+\sum_{i=1}^{n_t} y_i)^a},
\end{aligned} \tag{15}$$

where $\Gamma(\cdot)$ is the gamma function and the last equality follows from standard calculus. It is easy to verify that the derivative of $\varphi(p|a, S_t, \mathbf{y}_t)$ with respect to $p$ is given by

$$\varphi'(p|a, S_t, \mathbf{y}_t) = -\varphi(p|a+1, S_t, \mathbf{y}_t).$$

Thus, the single-period revenue function can be written as $R(p|a, S_t, \mathbf{y}_t) = p\varphi(p|a, S_t, \mathbf{y}_t)$. The derivative of the single-period revenue function is given by

$$R'(p|a, S_t, \mathbf{y}_t) = \varphi(p|a, S_t, \mathbf{y}_t) - p\varphi(p|a+1, S_t, \mathbf{y}_t).$$

The two-sided censoring problem under the exponential distribution can be formulated as the following dynamic program: for $t = T, ..., 1$,

$$\begin{aligned}
V_t(q|a, S_t, \mathbf{y}_t) &= \max_{p\geq 0}\{G_t(p, q|a, S_t, \mathbf{y}_t)\} \\
&= \max_{p\geq 0}\{R(p|a, S_t, \mathbf{y}_t) + V_{t-1}(q|a, S_t, [\mathbf{y}_t, p]) + V_{t-1}(q-1|a, S_t+p, \mathbf{y}_t)\}, \quad (16)
\end{aligned}$$

with $V_t(0|\cdot,\cdot,\cdot) = 0$ and $V_0(\cdot|\cdot,\cdot,\cdot) = 0$. Let $p_t^*(q|a, S_t, \mathbf{y}_t)$ denote the optimal solution for period $t$ to the above dynamic program. We can establish the following scaling result to simplify computation:

**Proposition 6.** *For any $q \geq 0$ and $1 \leq t \leq T$, $p_t^*(q|a, S, \mathbf{y}) = S \cdot p_t^*(q|a, 1, \mathbf{y}/S)$, $G_t(p, q|a, S, \mathbf{y}) = S^{-a+1} \cdot G_t(p/S, q|a, 1, \mathbf{y}/S)$, and $V_t(q|a, S, \mathbf{y}) = S^{-a+1} \cdot V_t(q|a, 1, \mathbf{y}/S)$.*

The above scaling result is reminiscent of those discovered in Bayesian inventory control problems that feature either no-censoring or right-censoring (Scarf 1960; Azoury 1985; Lariviere and Porteus 1999; Bisi et al. 2011). Here we extend the result to a Bayesian dynamic pricing problem with two-sided censoring. Also, our result is slightly different from those in the literature in that the scaling factor for the objective and value functions is $S^{-a+1}$ rather than $S$. This is due to the fact that our dynamic program formulation is based on the unnormalized prior. With this result, for any value of $S$, we can simply solve the problem with $S = 1$ and then scale the solution by $S$. Also, we note that $a$ does not change across periods; therefore, we can suppress $a$ in the problem formulation. Specifically, applying the result of Proposition 6 to (16), with some calculus, we can define the following dynamic program problem: for $t = T, ..., 1$,

$$\begin{aligned} v_t(q|\mathbf{y}) &= \max_{p \geq 0}\{g_t(p, q|\mathbf{y})\} \\ &= \max_{p \geq 0}\left\{p\varphi(p|a, 1, \mathbf{y}) + v_{t-1}\left(q|[\mathbf{y}, p]\right) + (p+1)^{-a+1} \cdot v_{t-1}\left(q-1|\mathbf{y}/(p+1)\right)\right\}, \end{aligned} \tag{17}$$

with $v_t(0|\cdot) = 0$, $v_0(\cdot|\cdot) = 0$. It is straightforward to verify that $V_t(q|a, 1, \mathbf{y}) = v_t(q|\mathbf{y})$. Because $\mathbf{y}$ is an empty vector $\emptyset$ at the beginning of period $T$, from Proposition 6, it follows that the optimal solution to the original problem is given by $p_T^*(q|a, S, \emptyset) = S \cdot \arg\max_{p \geq 0} g_T(p, q|\mathbf{y} = \emptyset)$. We note that the dimensionality of the dynamic program (17) increases linearly in the length of the decision horizon. Thus, if the decision horizon is relatively long (e.g., $T \geq 5$), the optimal solution becomes very difficult to compute due to the curse of dimensionality.

### 4.1.2 Heuristic Formulas under Two-Sided Censoring

In this section, we provide detailed formulas for five heuristics. The first one is a baseline heuristic, which is the solution without Bayesian updating. The second one is the full-information heuristic proposed by Aviv and Pazgal (2005a). The remaining three heuristics are those proposed in §3.3.

**Baseline (No-Learning) Heuristic Formula**

A baseline heuristic for our problem is the solution without Bayesian updating. Given the initial gamma prior with parameters $(a, S)$, let us define the dynamic program without Bayesian updating as follows: $t = T, ..., 1$,

$$V_t^n(q) = \max_{p \geq 0} \left\{ \frac{\Gamma(a)p}{(p+S)^a} + V_{t-1}^n(q) \cdot \left(1 - \frac{\Gamma(a)}{(p+S)^a}\right) + V_{t-1}^n(q-1) \cdot \frac{\Gamma(a)}{(p+S)^a} \right\},$$

with $V_t^n(0) = 0$ and $V_0^n(\cdot) = 0$. By the first-order condition, we have

$$\frac{\Gamma(a)[S-(a-1)p]}{(p+S)^{a+1}} + [V_{t-1}^n(q) - V_{t-1}^n(q-1)] \cdot \frac{\Gamma(a+1)}{(p+S)^{a+1}} = 0.$$

With some rearrangement, we obtain that the no-learning heuristic, denoted by $p_T^n(q|a,S)$, is given by:

$$p_T^n(q|a,S) = \frac{S + a[V_{T-1}^n(q) - V_{T-1}^n(q-1)]}{a-1},$$

where $V_{t-1}^n(q)$ can easily be determined from the above one-dimensional dynamic program. From the formula, it is clear that $p_T^n(q|a,S)$ is increasing in the marginal value of inventory capacity $V_{T-1}^n(q) - V_{T-1}^n(q-1)$ in this case.

**Full-Information Heuristic Formula**

Aviv and Pazgal (2005a) proposed two heuristics: the full-information heuristic and an enhanced version termed the "information-structure modification (IMS) heuristic." These two heuristics assume that the parameter $\theta$ becomes known after the pricing decision is made. In our model, because the parameter set includes the point of $\theta = 0$, it is easy to verify that the IMS heuristic reduces to the full-information heuristic (see p. 1407 of Aviv and Pazgal 2005a). Therefore, we only need to consider the full-information heuristic here. Given the initial gamma prior with parameters $(a, S)$, it can be shown that the full-information heuristic for period $T$ can be determined by the following first-order condition:

$$\frac{\Gamma(a)[S-(a-1)p]}{(p+S)^{a+1}} + \int_0^\infty [V_{T-1}^{FI}(q|\theta) - V_{T-1}^{FI}(q-1|\theta)] \cdot \theta^a e^{-(S+p)\theta} d\theta = 0.$$

Now apply the scaling result of Proposition 5 to the above equation. With some calculus, we obtain that the full-information heuristic, denoted by $p_T^f(q|a,S)$, is given by:

$$p_T^f(q|a,S) = \frac{S\left[1 + V_{T-1}^{FI}(q|1) - V_{T-1}^{FI}(q-1|1)\right]}{a - 1 - V_{T-1}^{FI}(q|1) + V_{T-1}^{FI}(q-1|1)},$$

where $V_{T-1}^{FI}(q|1)$ can be determined from the one-dimensional dynamic program (13). From the formula, it is easy to verify that $p_T^f(q|a,S)$ is increasing in the marginal value of inventory capacity $V_{T-1}^{FI}(q|1) - V_{T-1}^{FI}(q-1|1)$ in this case.

**Exact-Observation Heuristic Formula**

Given a mixed gamma prior with parameters $(a, S, \mathbf{y})$ defined in (14), when the customer WTP $x$ is observed exactly, it is easy to verify that the updated prior remains a mixed gamma distribution with updated parameters $(a+1, S+x, \mathbf{y})$. Thus, the dynamic program for the exact observation case can be formulated as follows: for $t = T, ..., 1$,

$$\begin{aligned} V_t^e(q|a,S,\mathbf{y}) &= \max_{p\geq 0}\{G_t^e(p,q|a,S,\mathbf{y})\} \\ &= \max_{p\geq 0}\left\{R(p|a,S,\mathbf{y}) + \int_0^p V_{t-1}^e\left(q|a+1,S+x,\mathbf{y}\right)dx\right. \\ &\qquad \left. + \int_p^\infty V_{t-1}^e\left(q-1|a+1,S+x,\mathbf{y}\right)dx\right\}, \end{aligned} \tag{18}$$

with $V_t^e(0|\cdot,\cdot,\cdot) = 0$ and $V_0^e(\cdot|\cdot,\cdot,\cdot) = 0$. Let $p_t^e(q|a,S,\mathbf{y})$ be the optimal solution to the above dynamic program. We can establish the following scaling result similar to that of Proposition 6:

**Proposition 7.** *For any $q \geq 0$ and $1 \leq t \leq T$, $p_t^e(q|a,S,\mathbf{y}) = S \cdot p_t^e(q|a,1,\mathbf{y}/S)$, $G_t^e(p,q|a,S,\mathbf{y}) = S^{-a+1} \cdot G_t^e(p/S,q|a,1,\mathbf{y}/S)$, and $V_t^e(q|a,S,\mathbf{y}) = S^{-a+1} \cdot V_t^e(q|a,1,\mathbf{y}/S)$.*

With the above result, for any value of $S$, we can simply solve the problem with $S = 1$ and then scale the solution by $S$. When the initial prior at the beginning of period $T$ is a gamma distribution with parameters $(a, S)$, we have $\mathbf{y} = \emptyset$. Thus, we can suppress $\mathbf{y}$ from the above notation. Also, we note that $a$ increases by one after each period; therefore, we can further suppress $a$ in the problem formulation. Now apply the scaling result of Proposition 7 to (18). With some calculus, we can define the following one-dimensional dynamic program: for $t = T, ..., 1$,

$$v_t^e(q) = \max_{p\geq 0}\left\{\frac{\Gamma(a+T-t)p}{(p+1)^{a+T-t}} + \frac{[(p+1)^{a+T-t-1}-1]v_{t-1}^e(q) + v_{t-1}^e(q-1)}{(a+T-t-1)(p+1)^{a+T-t-1}}\right\}, \tag{19}$$

with $v_t^e(0) = 0$ and $v_0^e(\cdot) = 0$.

It is straightforward to verify that $V_{T-1}^e(q|a+1, S+p, \emptyset) = (S+p)^{-a} \cdot v_{T-1}^e(q)$. Thus, the exact-observation heuristic for period $T$ can be determined by the following first-order condition:

$$\frac{\Gamma(a)[S-(a-1)p]}{(p+S)^{a+1}} + \frac{v_{T-1}^e(q) - v_{T-1}^e(q-1)}{(p+S)^a} = 0. \tag{20}$$

With some rearrangement, we obtain the exact-observation heuristic for period $T$ is given by the following formula:

$$p_T^e(q|a,S) = \frac{S\left[\Gamma(a) + v_{T-1}^e(q) - v_{T-1}^e(q-1)\right]}{\Gamma(a+1) - \Gamma(a) - v_{T-1}^e(q) + v_{T-1}^e(q-1)},$$

where $v_{T-1}^e(q)$ can easily be determined from the one-dimensional dynamic program (19). From the formula, it is easy to verify that $p_T^e(q|a,S)$ is increasing in the marginal value of inventory capacity $v_{T-1}^e(q) - v_{T-1}^e(q-1)$ in this case.

**One-Step Myopic Heuristic Formula**

Given the initial gamma prior with parameters $(a, S)$ in period $T$, based on our discussion in §3.3.2, the one-step myopic heuristic, denoted by $p_T^{sm}(q|a, S)$, can be computed by solving the following problem:

$$p_T^{sm}(q|a, S) = \arg\max_{p\geq 0} \left\{ R(p|a, S) + \hat{V}_{T-1}\left(q|a, S, p\right) + \hat{V}_{T-1}\left(q-1|a, S+p, \emptyset\right) \right\},$$

where $\hat{V}_{T-1}$ is determined by the following equations: for $t = T-1, ..., 1$,

$$\hat{V}_t(q|a, S_t, \mathbf{y}_t) = R(p_t^s|a, S_t, \mathbf{y}_t) + \hat{V}_{t-1}(q|a, S_t, [\mathbf{y}_t, p_t^s]) + \hat{V}_{t-1}(q-1|a, S_t + p_t^s, \mathbf{y}_t), \qquad (21)$$

with $\hat{V}_t(0|\cdot, \cdot, \cdot) = 0$ and $\hat{V}_0(\cdot|\cdot, \cdot, \cdot) = 0$. In the above expression, $p_t^s$ is determined based on a myopic policy, i.e., $p_t^s = \arg\max_{p\geq 0} R(p|a, S_t, \mathbf{y}_t)$. Specifically, $p_t^s$ can easily be determined by the following first-order condition:

$$\varphi(p|a, S_t, \mathbf{y}_t) - p\varphi(p|a+1, S_t, \mathbf{y}_t) = 0,$$

where the function $\varphi(p|a, S_t, \mathbf{y}_t)$ is defined in (15).

**One-Step Dynamic Heuristic Formula**

Given the initial gamma prior with parameters $(a, S)$ in period $T$, based on our discussion in §3.3.3, the one-step dynamic heuristic, denoted by $p_T^{sd}(q|a, S)$, can be computed by solving the problem:

$$p_T^{sd}(q|a, S) = \arg\max_{p\geq 0} \left\{ R(p|a, S) + \hat{V}_{T-1}\left(q|a, S, p\right) + \hat{V}_{T-1}\left(q-1|a, S+p, \emptyset\right) \right\},$$

where $\hat{V}_{T-1}$ is determined by (21), except that the (suboptimal) policy $p_t^s$ in the $\hat{V}_{T-1}$ expression is now determined by the following dynamic program: for $t, t-1, ..., 1$,

$$V_t^n(q) = \max_{p\geq 0} \left\{ p\varphi(p|a, S_t, \mathbf{y}_t) + V_{t-1}^n(q) \cdot (1 - \varphi(p|a, S_t, \mathbf{y}_t)) + V_{t-1}^n(q-1) \cdot \varphi(p|a, S_t, \mathbf{y}_t) \right\}, \quad (22)$$

with $V_t^n(0) = 0$ and $V_0^n(\cdot) = 0$. In the above expression, the function $\varphi(p|a, S_t, \mathbf{y}_t)$ is defined in (15). By the first-order condition, it follows that $p_t^s$ in this case is determined by the following equation:

$$\varphi(p|a, S_t, \mathbf{y}_t) - p\varphi(p|a+1, S_t, \mathbf{y}_t) + [V_{t-1}^n(q) - V_{t-1}^n(q-1)] \cdot \varphi(p|a+1, S_t, \mathbf{y}_t) = 0,$$

where it factors in the marginal value of inventory capacity $V_{t-1}^n(q) - V_{t-1}^n(q-1)$ and this term can easily be determined from the one-dimensional dynamic program (22).

### 4.1.3 Value Function Upper Bound Formula

Below we provide the computation formula for the one-step value function upper bound, which will be needed in evaluating the relative error bound defined in (24). Given the initial gamma prior with parameters $(a, S)$, by the definition given in §3.1, the one-step upper bound $V_T^{u_1}(q|a, S)$ can be computed by the following dynamic program: for $t = T, ..., 1$,

$$V_t^{u_1}(q|a,S) = \max_{p\geq 0} \left\{ R(p|a,S) + V_{t-1}^r(q|a,S,p) + V_{t-1}^{u_1}(q-1|a,S+p) \right\},$$

with $V_t^{u_1}(0|\cdot,\cdot) = 0$ and $V_0^{u_1}(\cdot|\cdot,\cdot) = 0$. In the above expression, $V_{t-1}^r$ can be computed as the following: For $t = T-1, ..., 1$,

$$V_t^r(q|a,S,y) = \max_{p\geq 0} \left\{ R(p|a,S,y) + \int_0^p V_{t-1}^r(q|a+1,S+x,y)dx + V_{t-1}^r(q-1|a,S+p,y) \right\},$$

with $V_t^r(0|\cdot,\cdot,\cdot) = 0$ and $V_0^r(\cdot|\cdot,\cdot,\cdot) = 0$. To further simplify the computation, with some calculus, we can define the following two normalized dynamic programs: for $t = T, ..., 1$,

$$v_t^{u_1}(q) = \max_{p\geq 0} \left\{ \frac{\Gamma(a)p}{(p+1)^a} + v_{t-1}^r(q|a,p) + \frac{v_{t-1}^{u_1}(q-1)}{(p+1)^{a-1}} \right\},$$

with $v_t^{u_1}(0) = 0$ and $v_0^{u_1}(\cdot) = 0$, and for $t = T-1, ..., 1$,

$$v_t^r(q|a,y) = \max_{p\geq 0} \left\{ p\varphi(p|a,1,y) + \int_0^p \frac{v_{t-1}^r(q|a+1, \frac{y}{x+1})}{(x+1)^a} dx + \frac{v_{t-1}^r(q-1|a, \frac{y}{p+1})}{(p+1)^{a-1}} \right\},$$

with $v_t^r(0|\cdot,\cdot) = 0$ and $v_0^r(\cdot|\cdot,\cdot) = 0$. By applying the scaling property, it is easy to verify that the one-step value function upper bound $V_T^{u_1}(q|a,S) = S^{-a+1} \cdot v_T^{u_1}(q)$, where $v_T^{u_1}(q)$ can be determined by the above two normalized dynamic programs.

## 4.2 Numerical Results

As discussed in §4.1.1, we can compute the optimal solution for the original two-sided censoring problem when the decision horizon $T < 5$. Thus, in our first numerical study, we keep the decision horizon at $T = 4$, so that we can compute the optimal solution and use it to benchmark the heuristics. Also, given the gamma prior parameters $(a, S)$, it is easy to verify that the coefficient of variation of the gamma prior (which is a measure of the prior uncertainty) is given by $\sqrt{1/a}$. In our numerical studies, we set $a = 2, 3, 4, 5$. As a result, the level of the prior uncertainty decreases from 0.707 to 0.447. Moreover, the myopic pricing decision that maximizes the single-period reward is equal to the predictive distribution mean $S/(a-1)$ in our model. Therefore, to have a consistent comparison, we choose the gamma parameters $(a, S)$ to be $(2, 10)$, $(3, 20)$, $(4, 30)$, and $(5, 40)$, so that the predictive distribution mean is kept at a constant value of 10 across all scenarios in our numerical studies. The numerical results for the optimal and heuristic solutions are presented in Table 1. In the table, we also place the relative error percentage of revenue loss compared to the optimal solution in parentheses next to the corresponding heuristics. For a heuristic solution $p^h$, the relative error percentage is defined as follows:

$$\text{Relative error percentage} = \frac{V_T(q,\pi) - G_T(p^h, q, \pi)}{V_T(q,\pi)} \times 100\%. \tag{23}$$

Table 1: Comparison of optimal and heuristic solutions ($T = 4$)

| $(a, S)$ | $q$ | $p_T^*$ | $p_T^n$ (%) | $p_T^f$ (%) | $p_T^e$ (%) | $p_T^{sm}$ (%) | $p_T^{sd}$ (%) |
|---|---|---|---|---|---|---|---|
| $(2, 10)$ | 1 | 30.7 | 22.4 (0.8%) | 101.1 (5.8%) | 46.4 (1.1%) | 27.5(0.1%) | 31.0 (0.0%) |
| | 2 | 17.1 | 12.3 (0.7%) | 16.5 (0.0%) | 16.5 (0.0%) | 18.7 (0.1%) | 16.7 (0.0%) |
| | 3 | 12.5 | 10.2 (0.3%) | 10.8 (0.1%) | 11.0 (0.1%) | 12.9 (0.0%) | 12.5 (0.0%) |
| | 4 | 11.5 | 10.0 (0.1%) | 10.0 (0.1%) | 10.0 (0.1%) | 11.5 (0.0%) | 11.4 (0.0%) |
| $(3, 20)$ | 1 | 23.6 | 20.6 (0.2%) | 30.8 (0.8%) | 25.7 (0.1%) | 22.0 (0.1%) | 23.2 (0.0%) |
| | 2 | 14.4 | 12.4 (0.2%) | 14.2 (0.0%) | 14.2 (0.0%) | 15.6 (0.1%) | 14.2 (0.0%) |
| | 3 | 11.2 | 10.3 (0.0%) | 10.6 (0.0%) | 10.7 (0.0%) | 11.4 (0.0%) | 11.2 (0.0%) |
| | 4 | 10.5 | 10.0 (0.0%) | 10.0 (0.0%) | 10.0 (0.0%) | 10.4 (0.0%) | 10.4 (0.0%) |
| $(4, 30)$ | 1 | 21.6 | 19.9 (0.1%) | 25.0 (0.3%) | 22.5 (0.0%) | 20.7 (0.0%) | 21.3 (0.0%) |
| | 2 | 13.6 | 12.4 (0.1%) | 13.5 (0.0%) | 13.5 (0.0%) | 14.7 (0.1%) | 13.5 (0.0%) |
| | 3 | 10.8 | 10.3 (0.0%) | 10.5 (0.0%) | 10.6 (0.0%) | 11.1 (0.0%) | 10.8 (0.0%) |
| | 4 | 10.2 | 10.0 (0.0%) | 10.0 (0.0%) | 10.0 (0.0%) | 10.2 (0.0%) | 10.2 (0.0%) |
| $(5, 40)$ | 1 | 20.7 | 19.5 (0.1%) | 22.9 (0.2%) | 21.2 (0.0%) | 18.8 (0.2%) | 20.8 (0.0%) |
| | 2 | 13.3 | 12.4 (0.1%) | 13.3 (0.0%) | 13.2 (0.0%) | 14.4 (0.1%) | 13.2 (0.0%) |
| | 3 | 10.7 | 10.3 (0.0%) | 10.5 (0.0%) | 10.6 (0.0%) | 10.8 (0.0%) | 10.8 (0.0%) |
| | 4 | 10.1 | 10.0 (0.0%) | 10.0 (0.0%) | 10.0 (0.0%) | 10.0 (0.0%) | 10.0 (0.0%) |

*Note: $p_T^*$ is the optimal solution, $p_T^n$ the baseline (no-learning) heuristic, $p_T^f$ the full-information heuristic, $p_T^e$ the exact-observation heuristic, $p_T^{sm}$ the one-step myopic heuristic, and $p_T^{sd}$ the one-step dynamic heuristic. The percentage in the parentheses is the relative error percentage.*

From Table 1, several observations are in order. First, the one-step dynamic heuristic is near optimal in all cases. It is almost identical to the optimal solution, with differences only in decimal points. As a result, the relative error percentage of revenue loss from this heuristic is negligible. Second, both the exact-observation heuristic and the one-step dynamic heuristic perform better than the full-information heuristic suggested by Aviv and Pazgal (2005a) in all cases. Third, when $q \geq 2$, the no-learning heuristic performs the worst among all heuristics. However, when $q = 1$ and $a = 2$, both the full-information and exact-observation heuristics perform significantly worse than the no-learning heuristic. Thus, these numerical results suggest that when inventory is in short supply and customer WTP is unknown (such as when a new product is just being introduced to the market), one should use more sophisticated heuristics, such as the one-step dynamic heuristic, for pricing decisions. When, however, inventory is abundant and customer demand is predictable (such as in the case of seasonal products that sell year after year), a simple myopic pricing decision or a dynamic pricing decision without Bayesian learning may just be sufficient, because the marginal value of inventory capacity or the value of learning is low in this case. It is also interesting to note that the no-learning heuristic tends to underprice compared to the optimal solution, suggesting an underestimation of the marginal value of inventory capacity in this heuristic.

In addition, we observe that the "price low" versus "price high" insight from Proposition 4 appears to be robust for problems with more than two periods. When $q$ is small, we observe that the optimal price is less than the exact-observation heuristic. As $q$ increases, the comparison result reverses. Finally, the optimal price under two-sided censoring decreases with the available inventory quantity, and increases with both the time horizon and the prior uncertainty level. This observation is consistent with predictions from dynamic pricing models without the data censoring effect (e.g., Gallego and van Ryzin 1994 and Aviv and Pazgal 2005a).

For problems with a relatively long decision horizon, e.g., $T \geq 5$, computing the optimal solution becomes intractable due to the curse of dimensionality. In this case, we cannot use the relative error percentage defined in (23) as a performance measure. However, leveraging the one-step lower and upper bounds from Propositions 2 and 3, for a heuristic solution $p^h$, we have

$$\frac{V_T(q,\pi) - G_T(p^h,q,\pi)}{V_T(q,\pi)} = 1 - \frac{G_T(p^h,q,\pi)}{V_T(q,\pi)} \leq 1 - \frac{G_T^{l_1}(p^h,q,\pi)}{V_T^{u_1}(q,\pi)},$$

where the righthand side can easily be computed. Thus, we can define an upper bound measure of the relative error percentage for a heuristic solution $p^h$ as follows:

$$\text{Relative error percentage bound} = \left(1 - \frac{G_T^{l_1}(p^h,q,\pi)}{V_T^{u_1}(q,\pi)}\right) \times 100\%. \tag{24}$$

In our second numerical study, we choose the decision horizon to be $T = 10$ and keep the gamma prior parameters the same as in the first study. For brevity, we report below the results with gamma prior parameters $(a, S) = (3, 20)$ and $(4, 30)$ in Table 2. In the table, we place the relative error percentage bound computed according to (24) in parentheses next to the corresponding heuristics.

From Table 2, we observe again that the one-step dynamic heuristic performs best among all heuristic solutions, with a maximum error bound at 4.2%. When $q = 1$ and $a = 3$, the full-information and exact-observation heuristics set prices significantly higher than does the one-step dynamic heuristic, suggesting a significant overestimation of the marginal value of inventory capacity in these two heuristics. This numerical observation can also be interpreted as follows:

Table 2: Comparison of heuristic solutions ($T = 10$)

| $(a, S)$ | $q$ | $p_T^n$ (%) | $p_T^f$ (%) | $p_T^e$ (%) | $p_T^{sm}$ (%) | $p_T^{sd}$ (%) |
|---|---|---|---|---|---|---|
| $(3, 20)$ | 1 | 32.2 (4.5%) | 107.4 (6.7%) | 63.3 (5.1%) | 25.8 (5.7%) | 38.8 (4.2%) |
| | 2 | 19.9 (2.8%) | 32.6 (2.6%) | 30.1 (2.5%) | 25.8 (2.4%) | 25.4 (2.4%) |
| | 3 | 14.8 (1.9%) | 17.8 (1.6%) | 19.8 (1.5%) | 22.6 (1.6%) | 20.2 (1.5%) |
| | 4 | 12.1 (1.3%) | 14.5 (1.1%) | 14.9 (1.0%) | 19.2 (1.2%) | 16.0 (1.0%) |
| | 5 | 10.8 (0.8%) | 11.9 (0.7%) | 12.2 (0.7%) | 15.0 (0.7%) | 13.2 (0.6%) |
| | 6 | 10.2 (0.5%) | 10.6 (0.5%) | 10.8 (0.5%) | 12.4 (0.4%) | 12.2 (0.4%) |
| | 7 | 10.0 (0.3%) | 10.2 (0.3%) | 10.2 (0.3%) | 11.2 (0.3%) | 11.0 (0.3%) |
| | 8 | 10.0 (0.3%) | 10.0 (0.3%) | 10.0 (0.3%) | 11.0 (0.3%) | 11.0 (0.3%) |
| | 9 | 10.0 (0.3%) | 10.0 (0.3%) | 10.0 (0.3%) | 11.0 (0.3%) | 11.0 (0.3%) |
| | 10 | 10.0 (0.3%) | 10.0 (0.3%) | 10.0 (0.3%) | 11.0 (0.3%) | 11.0 (0.3%) |
| $(4, 30)$ | 1 | 30.0 (2.8%) | 51.6 (3.4%) | 41.3 (2.8%) | 24.0 (3.8%) | 34.2 (2.5%) |
| | 2 | 19.6 (1.6%) | 26.1 (1.5%) | 24.8 (1.5%) | 24.3 (1.5%) | 24.9 (1.5%) |
| | 3 | 14.9 (1.1%) | 17.8 (0.9%) | 17.9 (0.9%) | 21.0 (1.1%) | 17.7 (0.9%) |
| | 4 | 12.3 (0.7%) | 13.8 (0.6%) | 14.1 (0.6%) | 17.1 (0.7%) | 14.4 (0.6%) |
| | 5 | 10.9 (0.4%) | 11.6 (0.4%) | 11.9 (0.4%) | 13.8 (0.4%) | 12.3 (0.4%) |
| | 6 | 10.3 (0.3%) | 10.6 (0.2%) | 10.7 (0.2%) | 11.7 (0.2%) | 11.1 (0.2%) |
| | 7 | 10.1 (0.2%) | 10.1 (0.2%) | 10.2 (0.2%) | 10.8 (0.2%) | 10.5 (0.2%) |
| | 8 | 10.0 (0.1%) | 10.0 (0.1%) | 10.0 (0.1%) | 10.5 (0.1%) | 10.5 (0.1%) |
| | 9 | 10.0 (0.1%) | 10.0 (0.1%) | 10.0 (0.1%) | 10.5 (0.1%) | 10.5 (0.1%) |
| | 10 | 10.0 (0.1%) | 10.0 (0.1%) | 10.0 (0.1%) | 10.5 (0.1%) | 10.5 (0.1%) |

*Note: $p_T^n$ the baseline (no-learning) heuristic, $p_T^f$ the full-information heuristic, $p_T^e$ the exact-observation heuristic, $p_T^{sm}$ the one-step myopic heuristic, and $p_T^{sd}$ the one-step dynamic heuristic. The percentage in the parentheses is the relative error percentage bound.*

When the product is in short supply, a seller equipped with better information, such as in the case of the full-information and exact-observation heuristics, is induced to engage in price gouging to extract the maximum value from such information. In contrast, under two-sided censoring (e.g., in the one-step dynamic heuristic), if the price is set too high and no purchase is observed, the seller learns little about the customer WTP. To avoid such a situation, a learning seller would instead set the price at a more moderate level. In this case, active learning under two-sided censoring helps curb price gouging behavior.

### 4.3 Robustness Check with a Two-Point Prior

We note that in Table 2, we used the relative error bound as an approximation performance measure. To further evaluate the performance of the one-step dynamic heuristic, we conduct an additional numerical study with a simple two-point prior. Under a two-point prior, the dimensionality of the dynamic program does not increase with the decision horizon length. Thus, we can compute the optimal solution and use it to benchmark the one-step dynamic heuristic as well as the relative error bound measure. Specifically, we assume customer WTP follows one of two candidate exponential distributions: one with $\theta_1 = 1/5$ and the other with $\theta_2 = 1/15$. Thus, the expected WTP under these two distributions are 5 and 15, respectively. We consider three initial prior cases: $(0.2, 0.8)$, $(0.5, 0.5)$, and $(0.8, 0.2)$, where the two entries in the vector represent the prior probabilities of the unknown parameter being $\theta_1$ and $\theta_2$, respectively. The numerical results are presented in Table 3 below. In the table, we include two percentage measures in the parentheses: the first one is the relative error percentage defined in (23) and the second one is the relative error percentage bound defined in (24).

Table 3: Solution comparison under exponential distribution with a two-point prior ($T = 10$)

| | Prior $= (0.2, 0.8)$ | | | Prior $= (0.5, 0.5)$ | | | Prior $= (0.8, 0.2)$ | | |
|---|---|---|---|---|---|---|---|---|---|
| $q$ | $p_T^*$ | $p_T^{sd}$ | Error % | $p_T^*$ | $p_T^{sd}$ | Error % | $p_T^*$ | $p_T^{sd}$ | Error % |
| 1 | 37.7 | 37.6 | (0.0%, 2.3%) | 35.7 | 35.2 | (0.0%, 6.8%) | 24.5 | 24.4 | (0.0%, 3.3%) |
| 2 | 27.2 | 27.2 | (0.0%, 1.1%) | 24.3 | 24.1 | (0.0%, 2.6%) | 16.0 | 16.1 | (0.0%, 2.0%) |
| 3 | 21.4 | 21.4 | (0.0%, 0.6%) | 18.4 | 18.3 | (0.0%, 1.3%) | 12.1 | 12.0 | (0.0%, 1.2%) |
| 4 | 17.8 | 17.8 | (0.0%, 0.4%) | 14.7 | 14.8 | (0.0%, 0.7%) | 9.7 | 9.7 | (0.0%, 0.7%) |
| 5 | 15.6 | 15.6 | (0.0%, 0.3%) | 12.6 | 12.6 | (0.0%, 0.5%) | 8.3 | 8.4 | (0.0%, 0.4%) |
| 6 | 14.4 | 14.4 | (0.0%, 0.3%) | 11.5 | 11.5 | (0.0%, 0.4%) | 7.6 | 7.6 | (0.0%, 0.3%) |
| 7 | 13.9 | 13.9 | (0.0%, 0.3%) | 11.0 | 11.0 | (0.0%, 0.3%) | 7.2 | 7.3 | (0.0%, 0.1%) |
| 8 | 13.8 | 13.8 | (0.0%, 0.3%) | 10.9 | 10.9 | (0.0%, 0.3%) | 7.2 | 7.2 | (0.0%, 0.1%) |
| 9 | 13.8 | 13.8 | (0.0%, 0.3%) | 10.8 | 10.9 | (0.0%, 0.3%) | 7.2 | 7.1 | (0.0%, 0.1%) |
| 10 | 13.8 | 13.8 | (0.0%, 0.3%) | 10.8 | 10.9 | (0.0%, 0.3%) | 7.2 | 7.1 | (0.0%, 0.1%) |

*Note: $p_T^*$ is the optimal solution, and $p_T^{sd}$ the one-step dynamic heuristic. The first percentage in the parentheses is the relative error percentage, and the second one is the relative error percentage bound.*

From Table 3, we observe that the one-step dynamic heuristic is almost identical to the optimal solution in all cases. Moreover, in the cases of $q = 1$, the relative error percentage bound appears to be exaggerating the actual relative error: it varies from 2.3% to 6.8%, but the actual relative error is 0.0% nonetheless. This numerical observation further corroborates the performance of the one-step dynamic heuristic in the gamma prior case, where the actual relative error percentages could also be much lower than the relative error percentage bounds reported in Table 2. We also conduct additional numerical assessment with different $\theta_1$ and $\theta_2$ values. The results are all similar to those shown in Table 3; we omit them for brevity.

Finally, to demonstrate that the one-step dynamic heuristic works for other distributions, we conduct an additional numerical study under the normal distribution with a two-point prior. Specifically, we assume customer WTP follows one of two candidate normal distributions: one with mean 5 and the other with mean 15. These two normal distributions share the same standard deviation of 5. As in the exponential case, we consider three initial prior cases: $(0.2, 0.8)$, $(0.5, 0.5)$, and $(0.8, 0.2)$. The numerical results are presented in Table 4 below. We also include two percentage measures in the parentheses: the first one is the relative error percentage defined in (23) and the second one is the relative error percentage bound defined in (24).

Table 4: Solution comparison under normal distribution with a two-point prior ($T = 10$)

| | Prior = $(0.2, 0.8)$ | | | Prior = $(0.5, 0.5)$ | | | Prior = $(0.8, 0.2)$ | | |
|---|---|---|---|---|---|---|---|---|---|
| $q$ | $p_T^*$ | $p_T^{sd}$ | Error % | $p_T^*$ | $p_T^{sd}$ | Error % | $p_T^*$ | $p_T^{sd}$ | Error % |
| 1 | 20.5 | 20.4 | (0.0%, 5.1%) | 19.2 | 19.1 | (0.0%, 7.9%) | 16.0 | 16.0 | (0.0%, 3.2%) |
| 2 | 18.3 | 18.4 | (0.0%, 2.9%) | 16.9 | 16.8 | (0.0%, 4.2%) | 13.6 | 13.5 | (0.0%, 1.8%) |
| 3 | 16.8 | 16.8 | (0.0%, 1.8%) | 15.2 | 15.2 | (0.0%, 2.5%) | 11.7 | 11.7 | (0.0%, 1.4%) |
| 4 | 15.5 | 15.5 | (0.0%, 1.2%) | 13.9 | 13.7 | (0.0%, 1.5%) | 10.3 | 10.3 | (0.0%, 1.1%) |
| 5 | 14.4 | 14.4 | (0.0%, 0.8%) | 12.8 | 12.8 | (0.0%, 1.0%) | 9.3 | 9.4 | (0.0%, 0.7%) |
| 6 | 13.5 | 13.4 | (0.0%, 0.5%) | 11.9 | 12.0 | (0.0%, 0.7%) | 8.7 | 8.7 | (0.0%, 0.5%) |
| 7 | 12.6 | 12.6 | (0.0%, 0.4%) | 11.1 | 11.2 | (0.0%, 0.5%) | 8.2 | 8.3 | (0.0%, 0.4%) |
| 8 | 11.8 | 11.8 | (0.0%, 0.3%) | 10.5 | 10.5 | (0.0%, 0.3%) | 7.9 | 8.0 | (0.0%, 0.2%) |
| 9 | 11.3 | 11.4 | (0.0%, 0.3%) | 10.1 | 10.1 | (0.0%, 0.3%) | 7.7 | 7.7 | (0.0%, 0.2%) |
| 10 | 11.2 | 11.2 | (0.0%, 0.3%) | 9.9 | 10.0 | (0.0%, 0.3%) | 7.7 | 7.6 | (0.0%, 0.2%) |

*Note: $p_T^*$ is the optimal solution, and $p_T^{sd}$ the one-step dynamic heuristic. The first percentage in the parentheses is the relative error percentage, and the second one is the relative error percentage bound.*

From Table 4, we observe again that the one-step dynamic heuristic is almost identical to the optimal solution in all cases. In the cases of $q = 1$, the relative error percentage bound also appears to be exaggerating the actual relative error: it varies from 3.2% to 7.9%, but the actual relative error is 0.0% nonetheless. These numerical results are very similar to those under the exponential distribution shown in Table 3, suggesting that the one-step dynamic heuristic is robust with respect to different distribution assumptions. We also conduct additional numerical assessment with different mean and standard deviation values for the normal distribution. The results are all similar to those shown in Table 4; we omit them for brevity.

## 5. Extensions and Concluding Remarks

The Bayesian dynamic pricing problem considered in this paper can be further extended in several directions. First, the scaling results of Propositions 6 and 7 are derived under the exponential distribution. We can further extend the result to the Weibull distribution family. Specifically, let us assume customer WTP follows a Weibull distribution given by $f(x|\theta) = k\theta x^{k-1}e^{-\theta x^k}$ with $k \geq 1$ (when $k = 1$, the Weibull distribution reduces to the exponential distribution) and the initial prior is an (unnormalized) gamma distribution with parameters $(a, S)$.

In this case, if there is an exact observation $x$, it can be shown that the updated (unnormalized) prior is a gamma distribution with parameters $(a + 1, S + x^k)$. If there is a right-censored observation $x$, the updated prior becomes a gamma distribution with parameters $(a, S + x^k)$. However, when there is a left-censored observation $x$, the updated prior becomes a mixture of two gamma distributions with parameters $(a, S)$ and $(a, S + x^k)$.

In our two-sided censoring problem, let us suppose that at the beginning of period $t$, there have been $n_t$ left-censored observations $y_1$, ..., $y_{n_t}$, and $T - t - n_t$ right-censored observations $x_1$, ..., $x_{T-t-n_t}$. Let $a_t = a$, $S_t = S + \sum_{i=1}^{T-t-n_t} x_i^k$ and $\mathbf{y}_t = [y_1^k, ..., y_{n_t}^k]$. Then the updated prior for period $t$ is given by

$$\pi_t(\theta|a_t, S_t, \mathbf{y}_t) = \theta^{a-1}e^{-S_t\theta}\prod_{i=1}^{n_t}\left(1 - e^{-y_i^k\theta}\right) \tag{25}$$

which is a mixture of gamma distributions with the same shape parameter $a$ but different scale parameters involving $S_t$ and all possible combinations of $y_1^k$, ..., $y_{n_t}^k$. Thus, like in the exponential distribution case, we can use $(a, S_t, \mathbf{y}_t)$ to represent the state of the prior distribution at the beginning of period $t$ in this case. With some proper modification of the dynamic program formulations, we can generalize Propositions 6 and 7 to the Weibull distribution case as follows:

**Proposition 8.** *Suppose that customer willingness to pay follows a Weibull distribution with an initial gamma prior. For any prior distribution state* $(a, S, \mathbf{y})$, $q \geq 0$, *and* $1 \leq t \leq T$, *the following holds:*

*(a)* $p_t^*(q|a, S, \mathbf{y}) = S^{1/k} \cdot p_t^*(q|a, 1, \mathbf{y}/S^{1/k})$, $G_t(p, q|a, S, \mathbf{y}) = S^{-a+1/k} \cdot G_t(p/S^{1/k}, q|a, 1, \mathbf{y}/S^{1/k})$, *and* $V_t(q|a, S, \mathbf{y}) = S^{-a+1/k} \cdot V_t(q|a, 1, \mathbf{y}/S^{1/k})$;

*(b)* $p_t^e(q|a, S, \mathbf{y}) = S^{1/k} \cdot p_t^e(q|a, 1, \mathbf{y}/S^{1/k})$, $G_t^e(p, q|a, S, \mathbf{y}) = S^{-a+1/k} \cdot G_t^e(p/S^{1/k}, q|a, 1, \mathbf{y}/S^{1/k})$, *and* $V_t^e(q|a, S, \mathbf{y}) = S^{-a+1/k} \cdot V_t^e(q|a, 1, \mathbf{y}/S^{1/k})$.

With the above scaling result, all the heuristic formulas derived in §4.1 can be extended to the Weibull distribution family.

Another extension to our model is to generalize customer WTP from the i.i.d. process to the Markov-modulated process (e.g., Iglehardt and Karlin 1962, Song and Zipkin 1993). If we assume that the unknown parameter $\theta$ evolves between periods according to a *known* stochastic transition function $\tau(\theta|\theta')$ (given the value of $\theta'$, $\tau(\cdot|\theta')$ is a probability distribution) and if we also assume that $\pi_t(\theta)$ is first transitioned to a new distribution by $\int_\Theta \tau(\theta|\theta')\pi_t(\theta')d\theta'$ before observing $Z_t = z$, then the posterior $\pi_{t-1}$ is given by

$$\pi_{t-1}(\theta) = \frac{k(Z_t = z|\theta) \cdot \int_\Theta \tau(\theta|\theta')\pi_t(\theta')d\theta'}{\int_\Theta k(Z_t = z|\theta) \cdot \int_\Theta \tau(\theta|\theta')\pi_t(\theta')d\theta' d\theta},$$

where the likelihood function $k(Z_t = z|\theta)$ is defined in (1). In the i.i.d. special case, $\tau(\theta|\theta')$ is the Dirac delta function $\delta(\theta - \theta')$; hence, $\int_\Theta \tau(\theta|\theta')\pi_t(\theta')d\theta' = \pi_t(\theta)$ and the above expression reduces to (2). Following an analysis analogous to the i.i.d. case, we can extend Theorem 1 to the Markov-modulated process, and all the subsequent results follow.

Finally, given the finite-horizon formulation of our dynamic pricing model, it is not clear whether the Bayesian updating process will lead to a close approximation to the true distribution under two-sided censoring. To study this Bayesian consistency problem, we need to consider an infinite-horizon problem formulation (see, for example, Scarf 1959). The required analysis is beyond the scope of the present work, but the problem merits investigation in future research.


### Acknowledgements

The author would like to thank Chengyu Wu for carrying out most of the numerical studies. Thanks also go to Paul Zipkin, Jeannette Song, Marty Lariviere, Erica Plambeck, Kaijie Zhu, and seminar participants at Duke University, University of North Carolina, University of Michigan, Shanghai Jiao Tong University, and the M&SOM Conference.


## References


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

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

Aviv, Y., A. Pazgal. 2005b. Dynamic pricing of short life-cycle products through active learning. Working Paper. Washington University, St Louis.

Azoury, K.S. 1985. Bayes solution to dynamic inventory models under unknown demand distribution. *Management Sci.* **31** 1150–1160.

Balvers, R.J., T.F. Cosimano. 1990. Actively learning about demand and the dynamics of price adjustment. *Econom. J.* **100** 882–898.

Bensoussan, A., M. Cakanyildirim, S.P. Sethi. 2005. On the optimal control of partially observed inventory systems. *Comptes Rendus Mathematique.* **341** 419–426.

Bensoussan, A., M. Cakanyildirim, S.P. Sethi. 2007. A multi-period newsvendor problem with partially observed demand. *Math. of Oper. Res.* **32**(2) 322–344.

Bensoussan, A., M. Cakanyildirim, S.P. Sethi. 2008. Inventory problems with partially observed demands and lost sales. *J. Optim. Theory and Applications.* **136**(2) 321–340.

Bensoussan, A., M. Cakanyildirim, S.P. Sethi. 2009. A note on "The censored newsvendor and the optimal acquisition of information". *Oper. Res.* **57**(3) 791–794.

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

Besbes, O., A. Zeevi. 2009. Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Oper. Res.* **57**(6) 1407–1420.

Bisi, A., M. Dada. 2007. Dynamic learning, pricing, and ordering by a censored newsvendor. *Naval Res. Logist. Quart.* **54** 448–461.

Bisi, A., M. Dada, S. Tokdar. 2011. A censored-data multiperiod inventory problem with newsvendor demand distributions. *Manufacturing & Service Oper. Management.* **13**(4) 525–533.

Bitran, G. R., S.V. Mondschein. 1997. Periodic pricing of seasonal products in retailing. *Management Sci.* **43**(1) 64–79.

Bitran, G. R., R. Caldentey. 2003. An overview of pricing models for revenue management. *Manufacturing & Service Oper. Management.* **5**(3) 203–229.

Braden, D.J., M. Freimer. 1991. Informational dynamics of censored observations. *Management Sci.* textbf37(11) 1390–1404.

Braden, D.J., S.S. Oren. 1994. Nonlinear pricing to produce information. *Marketing Sci.* **13**(3) 310–326.

Chen, L. 2010. Bounds and heuristics for optimal Bayesian inventory control with unobserved lost sales. *Oper. Res.* Forthcoming.

Chen, L., E.L. Plambeck. 2008. Dynamic inventory management with learning about the demand distribution and substitution probability. *Manufacturing & Service Oper. Management.* **10**(2) 236–256.

Carvalho, A.X., M.L. Puterman. 2005. Learning and pricing in an internet environment with binomial demands. *J. Revenue and Pricing Management.* **3**(4) 320–336.

Cope, E. 2007. Bayesian strategies for dynamic pricing in e-commerce. *Naval Res. Logist.* **54** 265–281.

Ding, X., M.L. Puterman, A. Bisi. 2002. The censored newsvendor and the optimal acquisition of information. *Oper. Res.* **50**(3) 517–527.

Dynkin, E., A. Yushkevich. 1979. *Controlled Markov Processes.* Springer, New York.

Elmaghraby, W., P. Keskinocak. 2003. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management Sci.* **49**(10) 1287–1309.

Farias, V., B. Van Roy. 2010. Dynamic pricing with a prior on market response. *Oper. Res.* **58**(1) 16–29.

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

Harpaz, G., W.Y. Lee, R.L. Winkler. 1982. Learning, experimentation, and the optimal output decisions of a competitive firm. *Management Sci.* **28**(6) 589–603.

Harrison, J.M., B. Keskin, A. Zeevi. 2012. Bayesian dynamic pricing policies: Learning and earning under a binary prior distribution. *Management Sci.* **58**93) 570–586.

Iglehardt, D., S. Karlin. 1962. Optimal policy for dynamic inventory process with nonstationary stochastic demands. *Studies in Applied Probability and Management Science.* K. Arrow, S. Karlin, H. Scarf (eds.). Chapter 8, Stanford University Press, Stanford, California.

Kandel, E. 1990. Essays on pricing under uncertainty. Unpublished Ph.D. Thesis, Graduate School of Business, University of Chicago.

Keskin, B., A. Zeevi. 2013. Dynamic pricing with an unknown demand model: Asymptotically optimal semi-myopic policies. Working paper. University of Chicago.

Kincaid, W.M., D.A. Darling. 1963. An inventory pricing problem. *J, Math. Annal. Appl.* **7** 183–208.

Lariviere, M.A., E.L. Porteus. 1999. Stalking information: Bayesian inventory management with unobserved lost sales. *Management Sci.* **45**(3) 346–363.

Lazear, E.P. 1986. Retail pricing and clearance sales. *Amer. Econom. Rev.* **76** 14–32.

Lim, A., J. Shanthikumar. 2007. Relative entropy, exponential utility, and robust dynamic pricing. *Oper. Res.* **55**(2) 198–214.

Lovejoy, W.S. 1993. Suboptimal policies, with bounds, for parameter adaptive decision processes. *Oper. Res.* **41**(3) 583–599.

Lu, X., J.-S. Song and K. Zhu. 2005. On "The censored newsvendor and the optimal acquisition of information". *Oper. Res.* **53**(6) 1024-1026.

Lu, X., J.S. Song, K. Zhu. 2008. Analysis of perishable inventory systems with censored data. *Oper. Res.* **56**(4) 1034–1038.

Milgrom, P., I. Segal. 2002. Envelope theorems for arbitrary choice sets. *Econometrica.* **70**(2) 583–601.

Monahan, G.E. 1982. A survey of partially observable Markov decision processes: model, theory and algorithms. *Management Sci.* **28**(1) 1–16.

Rishe, R. 2011. http://www.forbes.com/sites/prishe/2011/12/22/ 10-sports-business-stories-to-watch-in-2012/

Scarf, H.E. 1959. Bayes solution of the statistical inventory problem. *Ann. Math. Statist.* **30** 490–508.

Scarf, H.E. 1960. Some remarks on Bayes solutions to the inventory problem. *Naval Res. Logist. Quart.* **7** 591–596.

Song, J.-S., P. Zipkin. 1993. Inventory control in a fluctuating demand environment. *Oper. Res.* **41**(2) 351–370.

Sweeting, A. 2012. Dynamic pricing behavior in perishable goods markets: Evidence from secondary markets for major league baseball tickets. *J. Political Econom.* **120**(6) 1133–1172.

Talluri, K.T., G.J. van Ryzin. 2004. *The Theory and Practice of Revenue Management.* Springer Science, New York.

Wang, Z., S. Deng, Y. Ye. 2013. Close the gaps: A learning-while-doing algorithm for a class of single-product revenue management Problems. Working Paper. University of Minnesota.

Zhao, W., Y-S Zheng. 2000. Optimal dynamic pricing for perishable assets with nonhomogeneous demand. *Management Sci.* **46**(3) 375–388.

Ziya, S., H. Ayhan, R.D. Foley. 2004. Relationships among three assumptions in revenue management. *Oper. Res.* **52**(5) 804–809.

## Appendix for Online Companion

**Proof (Lemma 1)** It suffices to show that $\tilde{G}_t(p,q,\tilde{\pi}) = G_t(p,q,\pi)\cdot\int_\Theta \tilde{\pi}(\theta)d\theta$. We prove the result by backward induction. For period $t=1$, $\tilde{G}_1(p,q,\tilde{\pi}) = \tilde{R}(p,\tilde{\pi})$. By the definition of $\tilde{R}(p,\tilde{\pi})$, it is straightforward to show that $\tilde{R}(p,\tilde{\pi}) = R(p,\pi)\cdot\int_\Theta \tilde{\pi}(\theta)d\theta$. Now, assume the result holds for period $t-1$. For period $t$, we have

$$\begin{aligned}
\tilde{G}_t(p,q,\tilde{\pi}) &= \tilde{R}(p,\tilde{\pi}) + \tilde{V}_{t-1}(q, F(p|\cdot)\tilde{\pi}) + \tilde{V}_{t-1}(q-1, \bar{F}(p|\cdot)\tilde{\pi}) \\
&= R(p,\pi)\cdot\int_\Theta \tilde{\pi}(\theta)d\theta + V_{t-1}\left(q, \frac{F(p|\cdot)\tilde{\pi}}{\int_\Theta F(p|\theta)\tilde{\pi}(\theta)d\theta}\right)\cdot\int_\Theta F(p|\theta)\tilde{\pi}(\theta)d\theta \\
&\quad + V_{t-1}\left(q-1, \frac{\bar{F}(p|\cdot)\tilde{\pi}}{\int_\Theta \bar{F}(p|\theta)\tilde{\pi}(\theta)d\theta}\right)\cdot\int_\Theta \bar{F}(p|\theta)\tilde{\pi}(\theta)d\theta \\
&= \left\{R(p,\pi) + V_{t-1}\left(q, \frac{F(p|\cdot)\pi}{\int_\Theta F(p|\theta)\pi(\theta)d\theta}\right)\cdot\int_\Theta F(p|\theta)\pi(\theta)d\theta\right. \\
&\quad \left. + V_{t-1}\left(q-1, \frac{\bar{F}(p|\cdot)\pi}{\int_\Theta \bar{F}(p|\theta)\pi(\theta)d\theta}\right)\cdot\int_\Theta \bar{F}(p|\theta)\pi(\theta)d\theta\right\}\cdot\int_\Theta \tilde{\pi}(\theta)d\theta \\
&= G_t(p,q,\pi)\cdot\int_\Theta \tilde{\pi}(\theta)d\theta,
\end{aligned}$$

where the second equality follows from the induction assumption, the third equality follows from the fact that $\tilde{\pi} = \pi\cdot\int_\Theta \tilde{\pi}(\theta)d\theta$, and the last equality follows from definition (3). $\square$

**Proof (Theorem 1)** Let us prove the result by backward induction. For period $t=1$,

$$\begin{aligned}
V_1(q, F(p|\cdot)\pi) &= \max_{p'\geq 0}\left\{G_1(p', q, F(p|\cdot)\pi)\right\} \\
&= \max_{p'\geq 0}\left\{R(p', F(p|\cdot)\pi)\right\} = \max_{p'\geq 0}\left\{p'\int_{p'}^{\infty}\int_\Theta f(x|\theta)F(p|\theta)\pi(\theta)d\theta dx\right\}.
\end{aligned}$$

Note that $G_1(p', q, F(p|\cdot)\pi)$ is differentiable and absolutely continuous in $p$ because the integrand is continuous in $p$, with $dG_1(p', q, F(p|\cdot)\pi)/dp = G_1(p', q, f(p|\cdot)\pi)$. Also, note that $V_1(q, F(0|\cdot)\pi) = 0$. Then, by Theorem 2 of Milgrom and Segal (2002, p. 586), we immediately have $V_1(q, F(p|\cdot)\pi)$ is absolutely continuous in $p$. Based on the definition of $\Delta_1$ given in (5), we conclude that

$$V_1(q, F(p|\cdot)\pi) = \int_0^p V_1(q, f(x|\cdot)\pi)dx - \int_0^p \Delta_1(q, F(x|\cdot)\pi, f(x|\cdot)\pi)dx.$$

Now assume this holds for period $t-1$. For period $t$, we have

$$G_t(p', q, F(p|\cdot)\pi) = R(p', F(p|\cdot)\pi) + V_{t-1}(q, F(p'|\cdot)F(p|\cdot)\pi) + V_{t-1}(q-1, \bar{F}(p'|\cdot)F(p|\cdot)\pi).$$

By the induction assumption, we know that $G_t(p', q, F(p|\cdot)\pi)$ is absolutely continuous in $p$, which implies differentiability almost everywhere. Thus, at any given differentiable point $p$, we have

$$\begin{aligned}
& \frac{d}{dp} G_t(p', q, F(p|\cdot)\pi) \\
= \;& \frac{d}{dp} R(p', F(p|\cdot)\pi) + \frac{d}{dp} V_{t-1}(q, F(p'|\cdot)F(p|\cdot)\pi) + \frac{d}{dp} V_{t-1}(q-1, \bar{F}(p'|\cdot)F(p|\cdot)\pi) \\
= \;& R(p', f(p|\cdot)\pi) + V_{t-1}(q, F(p'|\cdot)f(p|\cdot)\pi) - \Delta_{t-1}(q, F(p'|\cdot)F(p|\cdot)\pi, F(p'|\cdot)f(p|\cdot)\pi) \\
& + V_{t-1}(q-1, \bar{F}(p'|\cdot)f(p|\cdot)\pi) - \Delta_{t-1}(q-1, \bar{F}(p'|\cdot)F(p|\cdot)\pi, \bar{F}(p'|\cdot)f(p|\cdot)\pi) \\
= \;& G_t(p', q, f(p|\cdot)\pi) - \Delta_{t-1}(q, F(p'|\cdot)F(p|\cdot)\pi, F(p'|\cdot)f(p|\cdot)\pi) \\
& - \Delta_{t-1}(q-1, \bar{F}(p'|\cdot)F(p|\cdot)\pi, \bar{F}(p'|\cdot)f(p|\cdot)\pi),
\end{aligned}$$

where the second equality follows from the induction assumption and the last equality follows from the definition of $G_t(\cdot,\cdot,\cdot)$. Let $\hat{p} = \arg\max_{p' \geq 0}\{G_t(p', q, F(p|\cdot)\pi)\}$. Based on the definition of $\Delta_t$ given in (5), we have

$$\frac{d}{dp} G_t(\hat{p}, q, F(p|\cdot)\pi) = V_t(q, f(p|\cdot)\pi) - \Delta_t(q, F(p|\cdot)\pi, f(p|\cdot)\pi).$$

Thus, by applying Theorem 2 of Milgrom and Segal (2002, p. 586) again, we have $V_t(q, F(p|\cdot)\pi)$ is absolutely continuous in $p$, and is given by

$$V_t(q, F(p|\cdot)\pi) = \int_0^p \frac{d}{dx} G_t(\hat{p}, q, F(x|\cdot)\pi) dx = \int_0^p V_t(q, f(x|\cdot)\pi) dx - \int_0^p \Delta_t\left(q, F(x|\cdot)\pi, f(x|\cdot)\pi\right) dx.$$

This completes the induction proof. Symmetrically, we can show that the corresponding result holds for $V_t(q, \bar{F}(p|\cdot)\pi)$. □

**Proof (Corollary 1)** Note that $G_1(p', q, F(p|\cdot)\pi) = R(p', F(p|\cdot)\pi)$ is differentiable in $p$ and $dG_1(p', q, F(p|\cdot)\pi)/dp$ is continuous in both $p'$ and $p$. Also, $G_1(p', q, F(p|\cdot)\pi)$ has a unique maximizer for any given $q$ and $F(p|\cdot)\pi$. Hence, by Corollary 4 of Milgrom and Segal (2002, p. 592), it immediately follows that $V_1(q, F(p|\cdot)\pi)$ is differentiable in $p$. Symmetrically, we can also show that $V_1(q, \bar{F}(p|\cdot)\pi)$ is differentiable in $p$. Thus, for any given $q$ and $\pi$, $G_2(p, q, \pi) = R(p, \pi) + V_1(q, F(p|\cdot)\pi) + V_1(q-1, \bar{F}(p|\cdot)\pi)$ is differentiable in $p$. Repeatedly applying Corollary 4 of Milgrom and Segal (2002, p. 592) from period 2 to period $t$, we arrive at that $G_t(p, q, \pi)$ is differentiable in $p$, with the derivative expression as given in the proposition. □

**Proof (Proposition 1)** By definition, in the last period we have $V_1(q, \pi) = V_1^r(q, \pi) = V_1^e(q, \pi) = \max_{p \geq 0} R(p, \pi)$. By Theorem 1, we have

$$G_2(p, q, \pi) \;\; = \;\; R(p, \pi) + V_1(q, F(p|\cdot)\pi) + V_1(q-1, \bar{F}(p|\cdot)\pi)$$

$$
\begin{aligned}
&\leq R(p,\pi) + \int_0^p V_1(q, f(x|\cdot)\pi)dx + V_1(q-1, \bar{F}(p|\cdot)\pi) \\
&= R(p,\pi) + \int_0^p V_1^r(q, f(x|\cdot)\pi)dx + V_1^r(q-1, \bar{F}(p|\cdot)\pi) \\
&= G_2^r(p,q,\pi).
\end{aligned}
$$

By maximizing over $p$ on both sides of the inequality, we have $V_2(q,\pi) \leq V_2^r(q,\pi)$. Similarly, using Theorem 1 again, we have

$$
\begin{aligned}
G_2^r(p,q,\pi) &= R(p,\pi) + \int_0^p V_1^r(q, f(x|\cdot)\pi)dx + V_1^r(q-1, \bar{F}(p|\cdot)\pi) \\
&\leq R(p,\pi) + \int_0^p V_1^r(q, f(x|\cdot)\pi)dx + \int_p^\infty V_1^r(q-1, f(x|\cdot)\pi)dx \\
&= R(p,\pi) + \int_0^p V_1^e(q, f(x|\cdot)\pi)dx + \int_p^\infty V_1^e(q-1, f(x|\cdot)\pi)dx \\
&= G_2^e(p,q,\pi).
\end{aligned}
$$

By maximizing over $p$ on both sides of the inequality, we have $V_2^r(q,\pi) \leq V_2^e(q,\pi)$. Repeat the above process from period 2 to $T$, we arrive at the desired result: $V_T(q,\pi) \leq V_T^r(q,\pi) \leq V_T^e(q,\pi)$. $\square$

**Proof (Proposition 2)** We prove parts (a) and (b) together. It is easy to verify by definition that $G_T(p,q,\pi) = G_T^{u_T}(p,q,\pi)$, $G_T^{u_0}(p,q,\pi) = G_T^r(p,q,\pi)$, $V_T(q,\pi) = V_T^{u_T}(q,\pi)$ and $V_T^{u_0}(q,\pi) = V_T^r(q,\pi)$. Let us first show that $V_t^{u_i}(q,\pi) \leq V_t^r(q,\pi)$ for any $i$. By definition, we have $V_1^{u_i}(q,\pi) = V_1^r(q,\pi)$. Now suppose that $V_{t-1}^{u_i}(q,\pi) \leq V_{t-1}^r(q,\pi)$. For period $t$, consider two cases. Case (1): if $n_t \geq i$, we have $V_t^{u_i}(q,\pi) = V_t^r(q,\pi)$. Case (2): if $n_t < i$, we have

$$
\begin{aligned}
G_t^{u_i}(p,q,\pi) &= R(p,\pi) + V_{t-1}^{u_i}(q, F(p|\cdot)\pi) + V_{t-1}^{u_i}(q-1, \bar{F}(p|\cdot)\pi) \\
&\leq R(p,\pi) + \int_0^p V_{t-1}^{u_i}(q, f(x|\cdot)\pi)dx + V_{t-1}^{u_i}(q-1, \bar{F}(p|\cdot)\pi) \\
&\leq R(p,\pi) + \int_0^p V_{t-1}^{r}(q, f(x|\cdot)\pi)dx + V_{t-1}^{r}(q-1, \bar{F}(p|\cdot)\pi) \\
&= G_t^r(p,q,\pi),
\end{aligned}
$$

where the first inequality follows from Theorem 1 and the second inequality follows from the induction assumption. Thus, by maximizing over $p$ on both sides of the inequality, it follows that $V_t^{u_i}(q,\pi) \leq V_t^r(q,\pi)$ in this case. Combining these two cases completes the induction proof. Now let us show that $V_t^{u_{i+1}}(q,\pi) \leq V_t^{u_i}(q,\pi)$ for any $t$. By definition, we have $V_1^{u_{i+1}}(q,\pi) = V_1^{u_i}(q,\pi)$. Now suppose that $V_{t-1}^{u_{i+1}}(q,\pi) \leq V_{t-1}^{u_i}(q,\pi)$. For period $t$, consider three cases. Case (1): if $n_t \geq i+1$, we have $V_t^{u_{i+1}}(q,\pi) = V_t^{u_i}(q,\pi) = V_t^r(q,\pi)$. Case (2): if $n_t = i$, by definition, we have $V_t^{u_i}(q,\pi) =$

$V_t^r(q,\pi)$. Using the result we have just shown, we conclude $V_t^{u_{i+1}}(q,\pi) \le V_t^r(q,\pi) = V_t^{u_i}(q,\pi)$. Case (3): if $n_t < i$, we have

$$
\begin{aligned}
G_t^{u_{i+1}}(p,q,\pi) &= R(p,\pi) + V_{t-1}^{u_{i+1}}(q, F(p|\cdot)\pi) + V_{t-1}^{u_{i+1}}(q-1, \bar{F}(p|\cdot)\pi) \\
&\le R(p,\pi) + V_{t-1}^{u_i}(q, F(p|\cdot)\pi) + V_{t-1}^{u_i}(q-1, \bar{F}(p|\cdot)\pi) \\
&= G_t^{u_i}(p,q,\pi),
\end{aligned}
$$

where the inequality follows from the induction assumption. Thus, by maximizing over $p$ on both sides of the inequality, it follows that $V_t^{u_{i+1}}(q,\pi) \le V_t^{u_i}(q,\pi)$ in this case. Combining the three cases, we arrive at the desired result. $\square$

**Proof (Proposition 3)** We prove parts (a) and (b) together. It is easy to verify by definition that $G_T(p,q,\pi) = G_T^{l_T}(p,q,\pi)$, $G_T^{l_0}(p,q,\pi) = \hat{V}_T(q,\pi)$, $V_T(q,\pi) = V_T^{l_T}(q,\pi)$ and $V_T^{l_0}(q,\pi) = \hat{V}_T(q,\pi)$. By definition, when $t = T - i$, we have

$$
\begin{aligned}
V_t^{l_{i+1}}(q,\pi) &= \max_{p\ge 0} \left\{ R(p,\pi) + \hat{V}_{t-1}(q, F(p|\cdot)\pi) + \hat{V}_{t-1}(q-1, \bar{F}(p|\cdot)\pi) \right\} \\
&\ge R(p_t^s(q,\pi),\pi) + \hat{V}_{t-1}(q, F(p_t^s|\cdot)\pi) + \hat{V}_{t-1}(q-1, \bar{F}(p_t^s|\cdot)\pi) \\
&= \hat{V}_t(q,\pi) = V_t^{l_i}(q,\pi),
\end{aligned}
$$

where the inequality follows from the fact that $p_t^s(q,\pi)$ is a special solution to the maximization problem. Substitute this result into the expression of $G_{t+1}^{l_{i+1}}(p,q,\pi)$. It follows that $G_{t+1}^{l_{i+1}}(p,q,\pi) \ge G_{t+1}^{l_i}(p,q,\pi)$, which implies $V_{t+1}^{l_{i+1}}(q,\pi) \ge V_{t+1}^{l_i}(q,\pi)$. Repeat this substitution process from period $t+1$ to $T$. The desired result follows. $\square$

**Proof (Proposition 4)** When $q = 1$, it is easy to verify that $G_t(p,1,\pi)$ is differentiable almost everywhere. At the differentiable point, by Corollary 1, we have

$$
\begin{aligned}
G_t'(p,1,\pi) &= R'(p,\pi) + V_{t-1}(1, f(p|\cdot)\pi) - \Delta_{t-1}\left(1, F(p|\cdot)\pi, f(p|\cdot)\pi\right) \\
&\le R'(p,\pi) + V_{t-1}^e(1, f(p|\cdot)\pi) = {G^e}'_t(p,1,\pi),
\end{aligned}
$$

where the inequality follows from Proposition 1 and the fact that $\Delta_{t-1}\left(1, F(p|\cdot)\pi, f(p|\cdot)\pi\right) \ge 0$. Thus, we have $G_t'(p,1,\pi) \le {G^e}'_t(p,1,\pi)$. Now suppose that for a given $\pi$, $p_t^* > p_t^e$. Then we have

$$
\begin{aligned}
G_t(p_t^*,1,\pi) &= G_t(p_t^e,1,\pi) + \int_{p_t^e}^{p_t^*} G_t'(p,1,\pi)dp \\
&\le G_t(p_t^e,1,\pi) + \int_{p_t^e}^{p_t^*} {G^e}'_t(p,1,\pi)dp
\end{aligned}
$$

$$
\begin{aligned}
&= G_t(p_t^e,1,\pi) + [G_t^e(p_t^*,1,\pi) - G_t^e(p_t^e,1,\pi)] \\
&< G_t(p_t^e,1,\pi),
\end{aligned}
$$

which contradicts the definition of $p_t^*$. Thus, we conclude that for any given $\pi$, $p_t^* \le p_t^e$.

When $q \ge 2$, in a two-period problem under exponential distribution and an initial gamma prior with parameters $(a, S)$, according to the first-order condition (20), it is easy to verify that when $q \ge 2$, $p_2^e = S/(a-1)$, where $S/(a-1)$ is the myopic revenue-maximizing price. By Corollary 1, it is also easy to verify that the first-order condition for the original two-sided censoring problem can be rewritten as

$$
\frac{\Gamma(a)\left[S-(a-1)p\right]}{(p+S)^{a+1}} + \frac{\Gamma(a+1)\hat{p}^l(p)}{(\hat{p}^l(p)+p+S)^{a+1}} - \frac{\Gamma(a+1)\hat{p}^r(p)}{(\hat{p}^r(p)+p+S)^{a+1}} = 0,
$$

where

$$
\begin{aligned}
\hat{p}^l(p) &= \arg\max_{p'>0}\left\{R(p'|a,S,p)\right\} = \arg\max_{p'>0}\left\{\frac{\Gamma(a)p'}{(p'+S)^a} - \frac{\Gamma(a)p'}{(p'+p+S)^a}\right\}, \\
\hat{p}^r(p) &= \arg\max_{p'>0}\left\{R(p'|a,S+p)\right\} = \arg\max_{p'>0}\left\{\frac{\Gamma(a)p'}{(p'+p+S)^a}\right\}.
\end{aligned}
$$

By the first-order condition, it is easy to show that $\hat{p}^r(p) = (S+p)/(a-1)$. For $\hat{p}^l(p)$, we can show the following. First, we take the first-order derivative of $R(p'|a,S,p)$ with respect to $p'$ and obtain

$$
R'(p'|a,S,p) = \frac{\Gamma(a)[S-(a-1)p']}{(p'+S)^{a+1}} - \frac{\Gamma(a)[S+p-(a-1)p']}{(p'+p+S)^{a+1}}.
$$

Now let $g(x) = (x-(a-1)p')/(p'+x)^{a+1}$. Then, it is straightforward to show that $g(x)$ is increasing over $[0, ap')$ and decreasing over $(ap', \infty)$. Therefore, $R'(p'|a,S,p) < 0$ for $S+p < ap'$ and $R'(p'|a,S,p) > 0$ for $S > ap'$. In other words, when $p' > (S+p)/a$, we have $R'(p'|a,S,p) < 0$; and when $p' < S/a$, we have $R'(p'|a,S,p) > 0$. Thus, we conclude that it must follow that $\hat{p}^l(p) \in [S/a, (S+p)/a]$. Because the function $\Gamma(a+1)p'/(p'+p+S)^{a+1}$ is unimodal with the peak at $p' = (S+p)/a$, it immediately follows that

$$
\begin{aligned}
\frac{\Gamma(a+1)\hat{p}^l(p)}{(\hat{p}^l(p)+p+S)^{a+1}} - \frac{\Gamma(a+1)\hat{p}^r(p)}{(\hat{p}^r(p)+p+S)^{a+1}} &\ge \frac{\Gamma(a+1)S/a}{(S/a+p+S)^{a+1}} - \frac{\Gamma(a+1)(S+p)/(a-1)}{((S+p)/(a-1)+p+S)^{a+1}} \\
&= \frac{\Gamma(a)\left(\frac{a}{a+1}\right)^{a+1}S}{\left(S+\frac{a}{a+1}p\right)^{a+1}} - \frac{\Gamma(a)\left(\frac{a-1}{a}\right)^a}{(S+p)^a}.
\end{aligned}
$$

Let $h(p)$ denote the function of the righthand side of the above equality. It is easy to verify that $h'(p) < -ah(p)/(S+p)$. Also note that

$$
h(S/(a-1)) = \frac{\Gamma(a)}{S^a}\cdot\left[\left(\frac{a(a-1)}{a^2+a-1}\right)^{a+1} - \left(\frac{a-1}{a}\right)^{2a}\right].
$$

Thus, to show that $h(S/(a-1)) > 0$ for $a > 1$, it suffices to show the ratio

$$r(a) = \frac{\left(\frac{a(a-1)}{a^2+a-1}\right)^{a+1}}{\left(\frac{a-1}{a}\right)^{2a}} = \frac{a^{3a+1}}{(a^2+a-1)^{a+1}(a-1)^{a-1}} > 1.$$

By L'Hopital's rule, it is easy to verify that $\lim_{a\to 1} r(a) = 1$ and $\lim_{a\to\infty} r(a) = 1$. Also note that

$$r'(a) = \frac{a^{3a+1}}{(a^2+a-1)^{a+1}(a-1)^{a-1}} \cdot \left[\ln\left(\frac{a^3}{(a^2+a-1)(a-1)}\right) - \frac{2a+1}{a(a^2+a-1)}\right].$$

From the expression we have $\lim_{a\to 1} r'(a) = \infty$ and $\lim_{a\to\infty} r'(a) = 0$. Let $w(a)$ denote the term in the bracket of the expression of $r'(a)$. Taking the derivative of $w(a)$, we can verify that there exists a point $a_0 = (3+\sqrt{5})/4$, such that $w(a)$ is decreasing for $1 < a < a_0$ and increasing for $a > a_0$, with $w(a_0) < 0$. Because $r'(1) = \infty$ and $r'(\infty) = 0$, it follows that $r'(a)$ must crosses zero once and only once, i.e., there exists a point $a_1$, such that $r'(a) > 0$ for $1 < a < a_1$ and $r'(a) < 0$ for $a > a_1$. Now recall that $r(1) = r(\infty) = 1$. Hence, we conclude that $r(a) > 1$ for all $a > 1$, or equivalently we have $h(S/(a-1)) > 0$ for all $a > 1$. Now suppose that there exists a point $p_0 \in [0, S/(a-1)]$ such that $h(p_0) < h(S/(a-1))$. Because $h(p)$ is a continuous and differentiable function in $p$, there must exist a point $p_1 \in [p_0, S/(a-1)]$, such that $h(p_1) > 0$ and $h'(p_1) > 0$. But this point contradicts the relationship $h'(p) < -ah(p)/(S+p)$. Therefore, it must be the case that $h(p) \geq h(S/(a-1)) > 0$ for $p \in [0, S/(a-1)]$. Substituting this result back to the first-order condition, we arrive at $G_2'(p, q|a, S) > 0$ for $p \in [0, S/(a-1)]$. Therefore, it follows that $p_2^* > S/(a-1) = p_2^e$. $\square$

**Proof (Proposition 5)** Prove by backward induction. By the result that $p_t^{FI}(q|\theta) = \theta^{-1} + V_{t-1}^{FI}(q|\theta) - V_{t-1}^{FI}(q-1|\theta)$, it is straightforward to verify that $p_1^{FI}(q|\theta) = \theta^{-1} p_1^{FI}(q|1)$ and $V_1^{FI}(q|\theta) = \theta^{-1} V_1^{FI}(q|1)$. Now assume this is true for the case of $t-1$. Then, for the case of $t$, we have $p_t^{FI}(q|\theta) = \theta^{-1} + V_{t-1}^{FI}(q|\theta) - V_{t-1}^{FI}(q-1|\theta) = \theta^{-1}[1 + V_{t-1}^{FI}(q|1) - V_{t-1}^{FI}(q-1|1)] = \theta^{-1} p_t^{FI}(q|1)$, where the second equality follows from induction assumption and the last equality follows from the definition of $p_t^{FI}(q|1)$. From (13), it is also easy to verify that $V_t^{FI}(q|\theta) = \theta^{-1} V_t^{FI}(q|1)$. Thus, this completes the induction proof. $\square$

**Proof (Proposition 6)** Prove by backward induction. It is straightforward to verify that $R(p|a, S, \mathbf{y}) = S^{-a+1} \cdot R(p/S|a, 1, \mathbf{y}/S)$. Therefore, we have $p_1^*(q|a, S, \mathbf{y}) = S \cdot p_1^*(q|a, 1, \mathbf{y}/S)$, $G_1(p, q|a, S, \mathbf{y}) = S^{-a+1} \cdot G_1(p/S, q|a, 1, \mathbf{y}/S)$, and $V_1(q|a, S, \mathbf{y}) = S^{-a+1} \cdot V_1(q|a, 1, \mathbf{y}/S)$. Now assume this is true for the case of $t-1$. Then, for the case of $t$, according to (16), we have

$$G_t(p, q|a, S, \mathbf{y})$$

$$
\begin{aligned}
&= S^{-a+1} \cdot [R(p/S|a,1,\mathbf{y}/S) + V_{t-1}(q|a,1,[\mathbf{y}/S,p/S]) + V_{t-1}(q-1|a,p/S+1,\mathbf{y}/S)] \\
&= S^{-a+1} \cdot G_t(p/S,q|a,1,\mathbf{y}/S),
\end{aligned}
$$

where the first equality follows from the induction assumption and the second equality follows from the definition of $G_t(p/S,q|a,1,\mathbf{y}/S)$. Given the above result, it immediately follows that $p_t^*(q|a,S,\mathbf{y}) = S \cdot p_t^*(q|a,1,\mathbf{y}/S)$ and $V_t(q|a,S,\mathbf{y}) = S^{-a+1} \cdot V_t(q|a,1,\mathbf{y}/S)$. This completes the induction proof. □

**Proof (Proposition 7)** Prove by backward induction. It is straightforward to verify that $p_1^e(q|a,S,\mathbf{y}) = S \cdot p_1^e(q|a,1,\mathbf{y}/S)$, $G_1^e(p,q|a,S,\mathbf{y}) = S^{-a+1} \cdot G_1^e(p/S,q|a,1,\mathbf{y}/S)$, and $V_1^e(q|a,S,\mathbf{y}) = S^{-a+1} \cdot V_1^e(q|a,1,\mathbf{y}/S)$. Now assume this is true for the case of $t-1$. Then, for the case of $t$, according to (18), we have

$$
\begin{aligned}
& G_t^e(p,q|a,S,\mathbf{y}) \\
=\; & R(p|a,S,\mathbf{y}) + \int_0^p V_{t-1}^e(q|a+1,S+x,\mathbf{y})\,dx + \int_p^\infty V_{t-1}^e(q-1|a+1,S+x,\mathbf{y})\,dx \\
=\; & S^{-a+1} \cdot \Bigg[R(p/S|a,1,\mathbf{y}/S) + \int_0^{p/S} V_{t-1}^e(q|a+1,x+1,\mathbf{y}/S)\,dx \\
& \qquad\qquad + \int_{p/S}^\infty V_{t-1}^e(q-1|a+1,x+1,\mathbf{y}/S)\,dx\Bigg] \\
=\; & S^{-a+1} \cdot G_t^e(p/S,q|a,1,\mathbf{y}/S),
\end{aligned}
$$

where the second equality follows from the induction assumption and some integration calculus, and the last equality follows from the definition of $G_t^e(p/S,q|a,1,\mathbf{y}/S)$. Given the above result, it immediately follows that $p_t^e(q|a,S,\mathbf{y}) = S \cdot p_t^e(q|a,1,\mathbf{y}/S)$ and $V_t^e(q|a,S,\mathbf{y}) = S^{-a+1} \cdot V_t^e(q|a,1,\mathbf{y}/S)$. This completes the induction proof. □

**Proof (Proposition 8)** The proof is essentially the same as that of Propositions 6 and 7. We omit it here for brevity. □