OPERATIONS RESEARCH
Vol. 58, No. 2, March–April 2010, pp. 396–413
ISSN 0030-364X | EISSN 1526-5463 | 10 | 5802 | 0396
DOI 10.1287/opre.1090.0726

Bounds and Heuristics for Optimal Bayesian Inventory Control with Unobserved Lost Sales

Li Chen
Fuqua School of Business, Duke University, Durham, North Carolina 27708, li.chen@duke.edu

In most retail environments, when inventory runs out, the unmet demand is lost and not observed. The sales data are effectively censored by the inventory level. Factoring this censored data effect into demand estimation and inventory control decision makes the problem difficult to solve. In this paper, we focus on developing bounds and heuristics for this problem. Specifically, we consider a finite-horizon inventory control problem for a nonperishable product with unobserved lost sales and a demand distribution having an unknown parameter. The parameter is estimated sequentially by the Bayesian updating method. We first derive a set of solution upper bounds that work for all prior and demand distributions. For a fairly general monotone likelihood-ratio distribution family, we derive relaxed but easily computable lower and upper bounds along an arbitrary sample path. We then propose two heuristics. The first heuristic is derived from the solution bound results. Computing this heuristic solution only requires the evaluation of the objective function in the observed lost-sales case. The second heuristic is based on the approximation of the first-order condition. We combine the first-order derivatives of the simpler observed lost-sales and perishable-inventory models to obtain the approximation. For the latter case, we obtain a recursive formula that simplifies the computation. Finally, we conduct an extensive numerical study to evaluate and compare the bounds and heuristics. The numerical results indicate that both heuristics perform very well. They outperform the myopic policies by a wide margin.

Subject classifications: inventory/production: heuristics; unknown demand distribution; Bayesian updating; unobserved lost sales; dynamic programming/optimal control: bounds.
Area of review: Manufacturing, Service, and Supply Chain Operations.
History: Received January 2007; revisions received October 2007, November 2008, January 2009; accepted April 2009. Published online in *Articles in Advance* October 7, 2009.

1. Introduction

A key assumption in the classic inventory control model is that the demand distribution is known a priori. However, in reality, this information is usually not available. To estimate the unknown demand distribution parameters, one needs to rely on the historical sales data. Such estimation works well when the sales data reflect the real demand, but in most retail environments, when inventory runs out, the unmet demand is lost and not observed. In other words, the sales data are censored by the inventory level. If the censored data are not factored into the estimation procedure, then the demand estimate will be biased low (see Nahmias 1994). Worse, if the low demand estimate is subsequently used to determine an inventory-stocking decision, the resulting inventory level will also be biased low, and thus lead to more lost sales and an even lower future demand estimate. To avoid this potential vicious cycle, it is thus important to take into account the censored data effect in the demand estimation and inventory control decisions.

In this paper, we consider a finite-horizon inventory control problem for a nonperishable product with unobserved lost sales and a demand distribution having an unknown parameter. This problem can be formulated as a Bayesian dynamic program. Under the formulation, the demand parameter is sequentially updated according to the Bayes rule. When lost sales are observed, a common approach in the literature is to assume that the prior distribution belongs to a conjugate family. The prior distribution can be characterized by one or two sufficient statistics, and thus the dimensionality of the problem can be reduced. For certain conjugate families, the dimensionality can be further reduced by a state-space reduction technique developed by Scarf (1960) and Azoury (1985). Under suitable conditions, Lovejoy (1990) has shown that the Bayesian dynamic program can be simplified to a single-period optimization problem. However, when lost sales are not observed, these techniques are generally not applicable because the censoring destroys the conjugate prior distribution structure (see Braden and Freimer 1991). One of the few known exceptions is the gamma-Weibull family, for which Lariviere and Porteus (1999) have shown that the dimensionality of the problem can be reduced, and thus the exact solution is computable. For all other cases, it remains a challenge to compute the exact optimal solution. In this paper, we focus on developing bounds and heuristics for this problem.

There are a few existing bound results in the literature. Chen and Plambeck (2008) and Chen (2009) have shown under a general demand distribution that the optimal solution is bounded below by the optimal solution in the observed lost-sales case. Lu et al. (2005b) have derived an upper bound for the optimal solution based on the first-order condition. Their upper bound result works for certain prior and demand distributions but not for all distributions (see §4.1 for a discussion). Can we construct a more general upper bound? If so, can we use it to design a heuristic policy? And how much improvement can we get from these heuristic policies over the myopic policies? These are the questions we seek to address in this paper.

Our contributions to the literature are threefold. First, we prove a sequence of lower and upper bounds for the value function of the Bayesian dynamic program. These results reinforce the intuition that (1) a system with observed lost sales achieves better performance than a system with unobserved lost sales (value of information), and (2) a system with Bayesian updating achieves better performance than a system without Bayesian updating (value of learning). Based on the value function bounds, we derive a sequence of solution upper bounds. These upper bounds work for all prior and demand distributions. For a fairly general monotone likelihood-ratio distribution family, we further develop relaxed but easily computable lower and upper bounds for the optimal solution along an arbitrary sample path.

Second, we propose two heuristics. The first heuristic is derived from the solution bound results. Computing this heuristic solution only requires the evaluation of the objective function in the observed lost sales; but one needs to choose a weighting parameter based on judgment or experience. The second heuristic is based on the approximation of the first-order condition. We combine the derivatives of two simplified cases to obtain the approximation. These two simplified cases are: the observed lost-sales case (which captures the inventory carryover effect but ignores the censored data effect) and the perishable-inventory case (which ignores the inventory carryover effect but captures the censored data effect). Whereas it is fairly straightforward to compute the derivative in the observed lost-sales case (Scarf 1959), computing the derivative in the perishable inventory case remains difficult (see Harpaz et al. 1982; Lariviere and Porteus 1999; Ding et al. 2002; Lu et al. 2005a, 2008). To overcome this difficulty, we obtain a recursive formula for the derivative in the perishable inventory case. This new formula simplifies the computation, and therefore we can use it to compute the heuristic solution from the approximate first-order condition. This heuristic method is more robust than the first one because there is no need for choosing an ad hoc weighting parameter.

Third, we conduct an extensive numerical study to evaluate and compare the bounds and heuristics developed in this paper. Our numerical results indicate that the solution lower bound obtained by Chen and Plambeck (2008) and Chen (2009) is consistently tighter than the solution upper bounds. We also compare our solution upper bounds with that of Lu et al. (2005b). We find that their upper bound, when it exists, falls in-between the most and the least tight upper bounds of ours in all cases. The heuristic suggested by Lu et al. (2005b) is a weighted average of the solution upper bound and lower bound. With a properly chosen weighting parameter, their heuristic is essentially the same as the first heuristic proposed in this paper. Our numerical results indicate that under a carefully chosen weighting parameter, the weighted-average heuristic is near optimal in all cases (within 0.04%), but the second heuristic obtained from the approximate first-order condition is even better (within 0.01%). Finally, we compare the heuristic policies with three myopic policies. The three myopic policies are defined as follows: Myopic-1 employs the Bayesian updating method and accounts for the censored data effect, Myopic-2 employs the Bayesian updating method but ignores the censored data effect, and Myopic-3 is a static policy with no Bayesian updating. The numerical results show that the heuristics outperform all three myopic policies, and the magnitude of improvement increases in the level of prior uncertainty. In addition, the results suggest that employing Bayesian updating as well as factoring the censored data effect into the estimation procedure can greatly improve system performance.

The rest of this paper is organized as follows. Section 2 introduces the problem formulation. Section 3 contains the value function bound results. The solution upper bounds are derived in §4. In §5, we develop two heuristics. Numerical results of the bounds and heuristics are presented in §6. Section 7 concludes the paper. All proofs are presented in the appendix.

2. Problem Formulation

Consider a periodic-review inventory control problem for a single product. The product is stocked and sold for T periods. At the beginning of each period t ($t = 1, \ldots, T$), an inventory level y is chosen to minimize the total inventory-holding and stockout penalty costs. The production lead time is assumed to be negligible, so the inventory level is achieved immediately after the decision. The product is nonperishable, so leftover inventory can be used to satisfy demand in subsequent periods.

At the end of each period, a unit holding cost h or a unit penalty cost p is charged for any leftover or shorted inventory, respectively. The purchase cost of the product is omitted in our formulation because it can be normalized to zero with the standard technique of Heyman and Sobel (1984). The terminal value at the end of the planning horizon is assumed to be zero. For any nonzero convex terminal value function, our results remain valid if the last-period cost function is properly adjusted.

The demands in each period, denoted by X_t, are independently and identically distributed with a common, general probability density function. This function, denoted by $f(x \mid \theta)$, has an unknown parameter θ, with $\theta \in \Theta$. Also, let $F(x \mid \theta)$ denote the cumulative distribution function (CDF) and $\bar{F}(x \mid \theta) = 1 - F(x \mid \theta)$ the complement CDF.

At the beginning of each period t, the unknown parameter θ is subject to a prior distribution $\pi_t(\theta)$. We will use π_t and $\pi_t(\theta)$ interchangeably when the meaning is clear from the context. The predictive demand density in period t given the prior distribution π_t is defined as

$$m(x \mid \pi_t) = \int_\Theta f(x \mid \theta)\pi_t(\theta)\, d\theta. \tag{1}$$

Given the inventory level y and demand x, the cost function in a period is defined as

$$\begin{aligned} c(y-x) &= h(y-x)^+ + p(y-x)^- \\ &= h \cdot \max(y-x, 0) + p \cdot \max(x-y, 0). \end{aligned}$$

The expected cost in a period can then be written as

$$C(y \mid \pi_t) = E_{X_t \mid \pi_t}\{c(y - X_t)\} = \int_0^\infty c(y-x)m(x \mid \pi_t)\, dx.$$

From the prior distribution π_t, the posterior distribution π_{t+1} can be updated based on the observation of the random demand X_t. If an exact demand x is observed, i.e., $X_t = x$, then based on Bayes rule, we have

$$\pi_{t+1}(\theta \mid \pi_t, X_t = x) = \frac{f(x \mid \theta)\pi_t(\theta)}{\int_\Theta f(x \mid \theta')\pi_t(\theta')\, d\theta'}. \tag{2}$$

For ease of notation, we will use $\pi_{t+1}(\theta \mid \pi_t, x)$ as shorthand for $\pi_{t+1}(\theta \mid \pi_t, X_t = x)$ when the meaning is clear from the context.

If a censored demand x is observed, which means the actual demand X_t can be greater than or equal to x, i.e., $X_t \geqslant x$, then based on Bayes rule, we have

$$\pi_{t+1}(\theta \mid \pi_t, X_t \geqslant x) = \frac{\bar{F}(x \mid \theta)\pi_t(\theta)}{\int_\Theta \bar{F}(x \mid \theta')\pi_t(\theta')\, d\theta'}. \tag{3}$$

Below we consider two scenarios depending on whether or not lost sales are observed.

2.1. Bayesian Updating with Observed Lost Sales

When lost sales are observed, the demand observation is always exact. To distinguish this case from the unobserved lost-sales case, let us add a superscript "o" (stands for "observed") to all corresponding dynamic programming value functions and optimal solutions. Specifically, for $t = 1, \ldots, T$, the optimality equations with observed lost sales are given by

$$\begin{aligned} V_t^o(z \mid \pi_t) &= \min_{y \geqslant z}\{G_t^o(y \mid \pi_t)\} \\ &= \min_{y \geqslant z}\{C(y \mid \pi_t) + E_{X_t \mid \pi_t}\{V_{t+1}^o((y - X_t)^+ \\ &\qquad \mid \pi_{t+1}(\cdot \mid \pi_t, X_t))\}\}, \end{aligned} \tag{4}$$

with the terminal value $V_{T+1}^o(\cdot \mid \cdot) = 0$ and $\pi_{t+1}(\cdot \mid \pi_t, X_t)$ as defined in (2). The optimal solution is denoted by $y_t^o(z \mid \pi_t)$. It is known that $G_t^o(y \mid \pi_t)$ is convex in y (Scarf 1959).

2.2. Bayesian Updating with Unobserved Lost Sales

When lost sales are not observed, demand information X_t is censored by the inventory level y. Let us use the notation $X_t \wedge y$ to indicate that X_t is censored by inventory level y. Given an observation $X_t \wedge y = x$, if $x < y$, then it is equivalent to the event $X_t = x$; otherwise, if $x = y$, then it is equivalent to the event $X_t \geqslant y$. Thus, the posterior distribution given an observation $X_t \wedge y = x$ is

$$\pi_{t+1}(\theta \mid \pi_t, X_t \wedge y = x) = \begin{cases} \pi_{t+1}(\theta \mid \pi_t, X_t = x) & \text{if } x < y, \\ \pi_{t+1}(\theta \mid \pi_t, X_t \geqslant y) & \text{if } x = y, \end{cases}$$

where $\pi_{t+1}(\theta \mid \pi_t, X_t = x)$ and $\pi_{t+1}(\theta \mid \pi_t, X_t \geqslant y)$ are defined in (2) and (3), respectively. In this unobserved lost-sales case, the optimality equations are given by, for $t = 1, \ldots, T$,

$$\begin{aligned} V_t(z \mid \pi_t) &= \min_{y \geqslant z}\{G_t(y \mid \pi_t)\} \\ &= \min_{y \geqslant z}\{C(y \mid \pi_t) + E_{X_t \mid \pi_t}\{V_{t+1}((y - X_t)^+ \\ &\qquad \mid \pi_{t+1}(\cdot \mid \pi_t, X_t \wedge y))\}\}, \end{aligned} \tag{5}$$

with the terminal value $V_{T+1}(\cdot \mid \cdot) = 0$. The optimal solution in this case is denoted by $y_t^*(z \mid \pi_t)$. From (5), we can see that the inventory-level decision y influences not only the on-hand inventory of the subsequent period, but also the posterior distribution of the future periods. Due to this added complexity, the convexity of the Bayesian dynamic program is difficult to establish.

3. Bounds for the Value Function

In this section, we develop bounds for the value function of for the Bayesian dynamic program (5). Let us first introduce two key observations in the following lemma:

Lemma 1. *For* $t = 1, \ldots, T$, $y \geqslant 0$, $z \geqslant 0$, *the following holds*:

(a) $E_{X_t \mid \pi_t}\{C(z \mid \pi_{t+1}(\cdot \mid \pi_t, X_t \wedge y))\} = C(z \mid \pi_t)$;

(b) $V_{t+1}^o(z \mid \pi_{t+1}(\cdot \mid \pi_t, X_t \geqslant y)) \geqslant E_{X_t \mid \pi_t}\{V_{t+1}^o(z \mid \pi_{t+1}(\cdot \mid \pi_t, X_t)) \mid X_t \geqslant y\}$.

Part (a) of Lemma 1 is essentially a special case of the law of total expectation under Bayesian updating with unobserved lost sales. Part (b) shows that censoring increases expected cost under Bayesian updating. In other words, censoring makes the observation less informative, and thus reduces the value of information. This result is essentially in the same vein as the concavity theorem of Bayes risk (DeGroot 1970, p. 125).

It is useful to consider two variant cases of demand updating here. First, let us consider the case when lost sales are initially not observed, but become observable from period i ($1 \leqslant i \leqslant T+1$) onwards. Let $V_t^{I_i}$ denote the value function in this case. The corresponding dynamic programming equations are given by

$$V_t^{l_i}(z \mid \pi_t) = \min_{y \geqslant z}\{G_t^{l_i}(y \mid \pi_t)\}$$
$$= \begin{cases} \min_{y\geqslant z}\{C(y \mid \pi_t) + E_{X_t\mid\pi_t}\{V_{t+1}^{l_i}((y-X_t)^+ \\ \qquad \mid \pi_{t+1}(\cdot \mid \pi_t, X_t \wedge y))\}\} & \text{if } t < i, \\ \min_{y\geqslant z}\{G_t^o(y \mid \pi_t)\} = V_t^o(z \mid \pi_t) & \text{if } t \geqslant i, \end{cases}$$

where $G_t^o(y \mid \pi_t)$ is given in (4). From the above definition, for $t \geqslant i$, $G_t^{l_i}(y \mid \pi_t) \equiv G_t^o(y \mid \pi_t)$. Furthermore, when $i = T$, which means lost sales are not observable until period T, we recover the original problem (5), i.e., $G_t^{l_T}(y \mid \pi_t) \equiv G_t(y \mid \pi_t)$.

Second, let us consider the case in which Bayesian updating stops from period i onwards and a stationary order-up-to policy is used in the ensuing periods. Let $V_t^{u_i}$ denote the value function in this case. The corresponding dynamic programming equations are

$$V_t^{u_i}(z \mid \pi_t) = \min_{y \geqslant z}\{G_t^{u_i}(y \mid \pi_t)\}$$
$$= \begin{cases} \min_{y\geqslant z}\{C(y \mid \pi_t) + E_{X_t\mid\pi_t}\{V_{t+1}^{u_i}((y-X_t)^+ \\ \qquad \mid \pi_{t+1}(\cdot \mid \pi_t, X_t \wedge y))\}\} & \text{if } t < i, \\ (T-t+1)\cdot \min_{y\geqslant z}\{C(y \mid \pi_t)\} & \text{if } t \geqslant i. \end{cases} \quad (6)$$

From the above definition, we observe that for $t \geqslant i$, $G_t^{u_i}(y \mid \pi_t) \equiv (T-t+1)\cdot C(y \mid \pi_t)$. Furthermore, when $i = T$, which means Bayesian updating is carried on until period T, we recover the original problem (5), i.e., $G_t^{u_T}(y \mid \pi_t) \equiv G_t(y \mid \pi_t)$.

Proposition 1. *For $t = 1, \ldots, T$, $y \geqslant 0$, $z \geqslant 0$, the following holds*:

(a) $G_t(y \mid \pi_t) = G_t^{l_T}(y \mid \pi_t) \geqslant \cdots \geqslant G_t^{l_{t+1}}(y \mid \pi_t) \geqslant G_t^{l_t}(y \mid \pi_t) = G_t^o(y \mid \pi_t)$;

(b) $V_t(z \mid \pi_t) = V_t^{l_T}(z \mid \pi_t) \geqslant \cdots \geqslant V_t^{l_{t+1}}(z \mid \pi_t) \geqslant V_t^{l_t}(z \mid \pi_t) = V_t^o(z \mid \pi_t)$;

(c) $G_t(y \mid \pi_t) = G_t^{u_T}(y \mid \pi_t) \leqslant \cdots \leqslant G_t^{u_{t+1}}(y \mid \pi_t) \leqslant G_t^{u_t}(y \mid \pi_t) = (T-t+1)\cdot C(y \mid \pi_t)$;

(d) $V_t(z \mid \pi_t) = V_t^{u_T}(z \mid \pi_t) \leqslant \cdots \leqslant V_t^{u_{t+1}}(z \mid \pi_t) \leqslant V_t^{u_t}(z \mid \pi_t) = (T-t+1)\cdot \min_{y\geqslant z}\{C(y \mid \pi_t)\}$.

Proposition 1 establishes a sequence of bounds for the objective (value) function of the Bayesian dynamic program (5): The lower bounds are the objective (value) functions when lost sales are assumed to be observable after a certain future period; and the upper bounds are the objective (value) functions when Bayesian updating stops after a certain future period. When $z \leqslant \arg\min_{y\geqslant 0}\{C(y \mid \pi_t)\}$, the rightmost of the inequalities in Proposition 1(d) is equivalent to the value function without Bayesian updating. These results reinforce the intuition that: (1) a system with observed lost sales achieves better performance than a system with unobserved lost sales (value of information); and (2) a system with Bayesian updating, regardless of whether lost sales are observed or not, achieves better performance than a system without Bayesian updating (value of learning).

4. Bounds for the Optimal Solution

Under a general discrete-demand distribution, Chen and Plambeck (2008) have shown that the optimal solution to (5) is bounded below by the optimal solution in the observed lost-sales case. A proof for the general continuous-demand distribution is given by Chen (2009). The following proposition restates their result:

Proposition 2 (Chen and Plambeck 2008, Chen 2009). *Given the same starting inventory and the same prior distribution, the optimal solution to the unobserved lost-sales problem* (5) *is bounded below by the optimal solution to the observed lost-sales problem* (4), *i.e.,* $y_t^*(z \mid \pi_t) \geqslant y_t^o(z \mid \pi_t)$.

The lower bound $y_t^o(z \mid \pi_t)$ is easy to compute, benefiting from two facts: first, the corresponding Bayesian dynamic program is convex; and second, for a fairly broad class of conjugate prior distribution families, the state-space reduction technique developed by Scarf (1960) and Azoury (1985) is applicable. In the following, we will focus on deriving upper bounds for the optimal solution.

4.1. Upper Bounds for the Optimal Solution

Intuitively, if we can find an inventory level such that any excess over this limit will result in an expected cost higher than the optimal cost, then we know this limit must be an upper bound to the optimal solution. Following this idea, with the aid of the value function bounds derived in Proposition 1, we can construct a sequence of upper bounds for $y_t^*(z \mid \pi_t)$ as given below:

Proposition 3. *The optimal solution to the unobserved lost-sales problem* (5) *is bounded above by* $y_t^{u_i}(z \mid \pi_t)$ $(t \leqslant i \leqslant T)$, *with* $y_t^{u_i}(z \mid \pi_t)$ *determined by solving the equation*

$$G_t^o(y \mid \pi_t) = V_t^{u_i}(z \mid \pi_t), \quad \text{s.t.} \quad y \geqslant y_t^o(z \mid \pi_t),$$

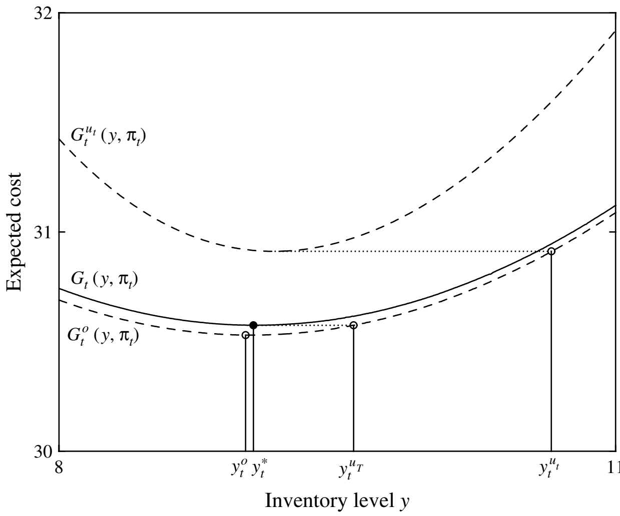
where $V_t^{u_i}(z \mid \pi_t)$ *is defined in* (6). *Furthermore, the upper bounds satisfy* $y_t^{u_{i+1}}(z \mid \pi_t) \leqslant y_t^{u_i}(z \mid \pi_t)$ *for* $t \leqslant i < T$.

By solving the equation given in Proposition 3, we have for any $y \geqslant y_t^{u_i}(z \mid \pi_t)$,

$$G_t(y \mid \pi_t) \geqslant G_t^o(y \mid \pi_t) \geqslant G_t^o(y_t^{u_i} \mid \pi_t)$$
$$= V_t^{u_i}(z \mid \pi_t) \geqslant V_t(z \mid \pi_t),$$

where the first and last inequalities follow from the bound result of Proposition 1 and the second inequality follows from the convexity property of $G_t^o(y \mid \pi_t)$. Therefore, we know that $y_t^{u_i}(z \mid \pi_t)$ must be an upper bound for the optimal solution $y_t^*(z \mid \pi_t)$.

Figure 1. Illustration of the lower and upper bounds for optimal inventory level.



A visual illustration of the upper bounds of Proposition 3 is given in Figure 1, where we show two upper bounds $y_t^{u_t}(0 \mid \pi_t)$ and $y_t^{u_T}(0 \mid \pi_t)$ as well as the lower bound $y_t^o(0 \mid \pi_t)$. From Proposition 1, we know that as i increases from t to T, the value function upper bound $V_t^{u_i}(z \mid \pi_t)$ becomes tight. Therefore, the resulting solution upper bound from solving the equation of Proposition 3 will also be tight, but the computational complexity will increase as well. The easiest case to compute is when $i = t$, i.e., $V_t^{u_t}(z \mid \pi_t) = (T-t+1) \cdot \min_{y \geqslant z}\{C(y \mid \pi_t)\}$, where we only need to solve a single-period optimization problem.

Lu et al. (2005b) have suggested an alternative solution upper bound based on the lower bound of the first-order derivative $dG_t(y \mid \pi_t)/dy$. Specifically, they show that

$$\frac{dG_t(y \mid \pi_t)}{dy} \geqslant \frac{dC(y \mid \pi_t)}{dy} + [V_{t+1}(0 \mid \pi) - \tilde{G}_{t+1}(\mathbf{y}^*_{t+1}(\pi') \mid \pi)] m(y \mid \pi_t),$$

where $\pi = \pi_{t+1}(\theta \mid \pi_t, X_t = y)$, $\pi' = \pi_{t+1}(\theta \mid \pi_t, X_t \geqslant y)$, and $\tilde{G}_{t+1}(\mathbf{y}^*_{t+1}(\pi') \mid \pi)$ is the expected cost when an "optimal" inventory policy $\mathbf{y}^*_{t+1}(\pi')$ determined by the starting prior π' is employed along the sample path, but the demand process evolves according to another starting prior π. Note that we use the bold font $\mathbf{y}^*_{t+1}$ to denote the inventory policy (as opposed to the single-period inventory decision) employed along the sample path. As is expected, it is generally difficult, if not impossible, to compute the term $\tilde{G}_{t+1}(\mathbf{y}^*_{t+1}(\pi') \mid \pi)$. Therefore, Lu et al. (2005b) propose to use the following relaxed lower bound instead:

$$\frac{dG_t(y \mid \pi_t)}{dy} \geqslant \frac{dC(y \mid \pi_t)}{dy} + V_{t+1}(0 \mid \pi) \cdot m(y \mid \pi_t) - V_{t+1}(0 \mid \pi') \cdot \int_\Theta \bar{F}(x \mid \theta)\pi_t(\theta)\, d\theta \cdot \max_{\theta \in \Theta}\left\{\frac{f(y \mid \theta)}{\bar{F}(y \mid \theta)}\right\}.$$

The problem with this relaxation is that the term $\max_{\theta \in \Theta}\{f(y \mid \theta)/\bar{F}(y \mid \theta)\}$ is not guaranteed to be finite when Θ is a continuous set like $\Theta = [0, \infty)$. For example, under the gamma-exponential conjugate prior distribution family, $\max_{\theta \in \Theta}\{f(y \mid \theta)/\bar{F}(y \mid \theta)\} = \max_{\theta \geqslant 0}\{\theta\} = \infty$, which makes their method inapplicable. In contrast, our upper-bound result given in Proposition 3 works for all prior and demand distributions.

4.2. Solution Bounds on a Sample Path

According to Propositions 2 and 3, we can compute the lower and upper bounds by solving a corresponding Bayesian dynamic program with observed lost sales. As discussed earlier, for certain conjugate families, we can employ the state-space reduction technique to simplify the computation. However, in many cases, this technique becomes inapplicable after the first censored observation because the censoring destroys the conjugate prior structure. In this section, we develop relaxed bounds that can be computed by the state-space reduction technique after a censored observation. Let us first introduce a few definitions:

Likelihood-Ratio Ordering. Let $f(x)$ and $g(x)$ denote two probability density distributions. We say that f is greater than g in likelihood ratio, denoted by $f \geqslant_{LR} g$, if, for all $x \geqslant x'$, $f(x)/g(x) \geqslant f(x')/g(x')$.

First-Order Stochastic Dominance (*FSD*). We say that f dominates g under first-order stochastic dominance, denoted by $f \geqslant_{FS} g$, if, for all y, $\int_y^\infty f(x)\,dx \geqslant \int_y^\infty g(x)\,dx$. It is known that likelihood-ratio ordering implies first-order stochastic dominance (Ross 1983).

Monotone Likelihood Ratio (*MLR*). A distribution family $f(\cdot \mid \theta)$ parameterized by a parameter θ is said to have monotone likelihood ratio if $f(\cdot \mid \theta) \geqslant_{LR} f(\cdot \mid \theta')$ for all $\theta \geqslant \theta'$. Distributions possessing this property include the normal with known variance, the binomial, the Poisson, the gamma, the Weibull with known shape parameter, and other less-common distributions (see Karlin and Rubin 1956). Note that θ is an arbitrary parameter; it does not have to be the mean of the distribution—for example, it could be the scale parameter in the gamma distribution.

For the monotone likelihood-ratio demand distributions, the following lemma establishes the stochastic ordering of the posterior distributions defined in (2) and (3) and the predictive demand distribution defined in (1):

LEMMA 2. *If $f(x \mid \theta)$ satisfies the monotone likelihood-ratio property, the following holds:*

(a) $\pi_{t+1}(\theta \mid \pi_t, X_t = x_1) \geqslant_{LR} \pi_{t+1}(\theta \mid \pi_t, X_t = x_2)$ *for all* $x_1 \geqslant x_2 \geqslant 0$;

(b) $\pi_{t+1}(\theta \mid \pi_t, X_t \geqslant x) \geqslant_{LR} \pi_{t+1}(\theta \mid \pi_t, X_t = x)$ *for all* $x \geqslant 0$;

(c) *if* $\pi_t(\theta) \geqslant_{LR} \pi'_t(\theta)$, *then* $\pi_{t+1}(\theta \mid \pi_t, X_t = x) \geqslant_{LR} \pi_{t+1}(\theta \mid \pi'_t, X_t = x)$ *for all* $x \geqslant 0$;

(d) *if* $\pi_t(\theta) \geqslant_{LR} \pi'_t(\theta)$, *then* $m(x \mid \pi_t) \geqslant_{FS} m(x \mid \pi'_t)$.

Part (d) of the above lemma is a generalized version of the result of Ding et al. (2002, Proposition 1). For monotone likelihood-ratio demand distributions, these results indicate that the posterior distribution after Bayesian updating is likelihood-ratio increasing in the value of the demand observation; and the posterior distribution with a censored observation is larger in likelihood ratio than that with an exact observation. From the lemma, we can establish the following monotonicity result for the optimal solution in the observed lost-sales case:

Proposition 4. *Assume that $f(x \mid \theta)$ satisfies the monotone likelihood-ratio property. Given two prior distributions $\pi_t(\theta)$ and $\pi'_t(\theta)$ subject to $\pi_t(\theta) \geqslant_{LR} \pi'_t(\theta)$, the optimal solutions to the observed lost-sales problem* (4) *satisfy $y_t^o(z \mid \pi_t) \geqslant y_t^o(z \mid \pi'_t)$.*

The above proposition generalizes the result of Eppen and Iyer (1997, Theorem 4) to demand distributions that possess the monotone likelihood-ratio property. It shows that for monotone likelihood-ratio demand distributions, when lost sales are observed, the optimal Bayesian inventory level is increasing in the likelihood-ratio ordering of the prior distribution.

A relaxed lower bound for the optimal solution with unobserved lost sales can be derived based on the above proposition. The idea is the following: Whenever we have a censored observation x, if we treat it as an exact observation, then $\pi_{t+1}(\theta \mid \pi_t, X_t \geqslant x) \geqslant_{LR} \pi_{t+1}(\theta \mid \pi_t, X_t = x)$ (Lemma 2b). By Proposition 4, it follows that $y_{t+1}^o(z \mid \pi_{t+1}(\cdot \mid \pi_t, X_t \geqslant x)) \geqslant y_{t+1}^o(z \mid \pi_{t+1}(\cdot \mid \pi_t, X_t = x))$. Then, according to Proposition 2, we have

$$y_{t+1}^*(z \mid \pi_{t+1}(\cdot \mid \pi_t, X_t \geqslant x)) \geqslant y_{t+1}^o(z \mid \pi_{t+1}(\cdot \mid \pi_t, X_t \geqslant x)) \geqslant y_{t+1}^o(z \mid \pi_{t+1}(\cdot \mid \pi_t, X_t = x)).$$

Thus, $y_{t+1}^o(z \mid \pi_{t+1}(\cdot \mid \pi_t, X_t = x))$ is a relaxed lower bound for $y_{t+1}^*(z \mid \pi_{t+1}(\cdot \mid \pi_t, X_t \geqslant x))$.

In general, given t observations $X_i \wedge y_i = x_i$ $(1 \leqslant i \leqslant t)$, according to Bayes' rule, the posterior distribution is given by

$$\pi_{t+1}(\theta \mid \pi_1, X_i \wedge y_i = x_i, 1 \leqslant i \leqslant t) = \frac{\prod_{i \in \{x_i < y_i\}} f(x_i \mid \theta) \cdot \prod_{i \in \{x_i = y_i\}} \bar{F}(x_i \mid \theta) \cdot \pi_1(\theta)}{\int_\Theta \prod_{i \in \{x_i < y_i\}} f(x_i \mid \theta') \cdot \prod_{i \in \{x_i = y_i\}} \bar{F}(x_i \mid \theta') \cdot \pi_1(\theta')\, d\theta'}.$$

Because the expression has a product form, we note that π_{t+1} does not depend on the sequence order of the past observations $X_i \wedge y_i = x_i$.

Proposition 5. *Assume that $f(x \mid \theta)$ satisfies the monotone likelihood-ratio property. Given t observations $X_i \wedge y_i = x_i$ $(1 \leqslant i \leqslant t)$, the optimal solution to the unobserved lost-sales problem* (5) *is bounded below by $y_{t+1}^o(z \mid \pi)$, where $\pi = \pi_{t+1}(\cdot \mid \pi_1, X_i = x_i, 1 \leqslant i \leqslant t)$.*

The above proposition shows that for monotone likelihood-ratio demand distributions, the lower bound of Proposition 2 can be relaxed by treating all observations along the sample path as exact observations.

Now, let us turn to the computation of an upper bound on a sample path. Given the posterior $\pi_{t+1} = \pi_{t+1}(\theta \mid \pi_1, X_i \wedge y_i = x_i, 1 \leqslant i \leqslant t)$, let us define a relaxed objective function in the observed lost-sales case as follows:

$$G_{t+1}^{o'}(y \mid \pi_{t+1}) = C(y \mid \pi_{t+1}) + E_{X_1, \ldots, X_{t+1} \mid \pi_1}\{V_{t+2}^o(0 \mid \pi_{t+2}(\cdot \mid \pi_1, X_1, \ldots, X_{t+1})) \mid X_i \wedge y_i = x_i, 1 \leqslant i \leqslant t\}. \quad (7)$$

In this function, the posterior π_{t+2} in the value function $V_{t+2}^o(0 \mid \pi_{t+2})$ treats all past observations as exact observations. Therefore, the state-space reduction technique can continue to apply to the computation of $V_{t+2}^o(0 \mid \pi_{t+2})$. Based on this relaxed objective function, the upper bound result of Proposition 3 can be extended:

Proposition 6. *Suppose that $\pi_{t+1} = \pi_{t+1}(\theta \mid \pi_1, X_i \wedge y_i = x_i, 1 \leqslant i \leqslant t)$. Then, the following holds: $G_{t+1}^{o'}(y \mid \pi_{t+1}) \leqslant G_{t+1}^o(y \mid \pi_{t+1})$, where $G_{t+1}^o(y \mid \pi_{t+1})$ is the objective function in the observed lost-sales case* (4). *Also, the optimal solution to the unobserved lost-sales problem* (5) *is bounded above by $y_{t+1}^{u'_i}(z \mid \pi_{t+1})$ $(t < i \leqslant T)$, with $y_{t+1}^{u'_i}(z \mid \pi_{t+1})$ determined by solving the equation*

$$G_{t+1}^{o'}(y \mid \pi_{t+1}) = V_{t+1}^{u_i}(z \mid \pi_{t+1}), \quad \text{s.t.} \quad y \geqslant y_{t+1}^{o'}(z \mid \pi_{t+1}).$$

The upper bounds satisfy $y_{t+1}^{u'_{i+1}}(z \mid \pi_{t+1}) \leqslant y_{t+1}^{u'_i}(z \mid \pi_{t+1})$ for $t < i < T$.

Below we choose the gamma-gamma distribution (Scarf 1960) as an example to illustrate how to calculate the relaxed lower and upper bounds using the state-space reduction technique. Extending the calculation to other suitable distributions identified by Azoury (1985) is straightforward.

4.3. The Gamma-Gamma Example

Suppose that demand is gamma distributed with known shape parameter $k > 0$ and unknown scale parameter θ (which satisfies the monotone likelihood-ratio property). The probability density function is given by

$$f(x \mid \theta) = \frac{\theta^k x^{k-1} e^{-\theta x}}{\Gamma(k)}.$$

Furthermore, suppose that the initial prior on θ has a gamma distribution with shape parameter a and scale parameter S, which is given by

$$\pi_1(\theta) = \frac{S^a \theta^{a-1} e^{-\theta S}}{\Gamma(a)}.$$

Given exact demand observations $x_1, \ldots, x_t$, it is easy to show that the posterior distribution $\pi_{t+1}(\theta)$ is a gamma distribution with shape parameter $a_{t+1} = a + kt$ and scale parameter $S_{t+1} = S + \sum_{j=1}^{t} x_j$. Define the normalized predictive demand density in period t as

$$\phi_t(x) = \frac{\Gamma(k+a_t)}{\Gamma(k)\Gamma(a_t)} \cdot \frac{x^{k-1}}{(1+x)^{k+a_t}}. \tag{8}$$

When lost sales are observed, Scarf (1960) has shown that the problem can be reduced to the following one-dimensional dynamic program: for $t = 1, \ldots, T$,

$$v_t^o(z) = \min_{y \geqslant z} \left\{ L_t(y) + \int_0^\infty (1+x) v_{t+1}^o\left(\frac{(y-x)^+}{1+x}\right) \phi_t(x)\, dx \right\}, \tag{9}$$

with $v_{T+1}^o(\cdot) = 0$, and $L_t(y) = E_{\phi_t}\{c(y - X_t)\}$. Let $\bar{y}_t^o(z)$ be the optimal solution to the above dynamic program. The optimal solution to the original problem is simply $y_t^o(z \mid \pi_t) = S_t \cdot \bar{y}_t^o(z/S_t)$, and the value function is $V_t^o(z \mid \pi_t) = S_t \cdot v_t^o(z/S_t)$.

Now, consider the case when lost sales are not observed. Given t observations $X_i \wedge y_i = x_i$ $(1 \leqslant i \leqslant t)$. By Proposition 5, the lower bound for $y_{t+1}^*(z \mid \pi_{t+1})$ is simply $S_{t+1} \cdot \bar{y}_{t+1}^o(z/S_{t+1})$, with $S_{t+1} = S + \sum_{j=1}^{t} x_j$, and $\bar{y}_{t+1}^o(z/S_{t+1})$ is the solution to the one-dimensional dynamic program (9). Based on definition (7), we have

$$G_{t+1}^{o'}(y \mid \pi_{t+1}) = C(y \mid \pi_{t+1}) + v_{t+2}^o(0) \cdot E_{X_1, \ldots, X_{t+1} \mid \pi_1}\left\{ \sum_{j=1}^{t+1} X_i \mid \pi_1, X_i \wedge y_i = x_i, 1 \leqslant i \leqslant t \right\}.$$

Plugging the above expression into the equation given in Proposition 6, we obtain the upper bound for the optimal solution $y_{t+1}^*(z \mid \pi_{t+1})$.

5. Heuristic Solutions

In this section, we propose two heuristic solutions to the unobserved lost-sales problem (5) based on the results derived in the previous sections.

5.1. Heuristic Solution I

Combining Propositions 2 and 3, we have $y_t^o(z \mid \pi_t) \leqslant y_t^*(z \mid \pi_t) \leqslant y_t^{u_T}(z \mid \pi_t)$, where both $y_t^o(z \mid \pi_t)$ and $y_t^{u_T}(z \mid \pi_t)$ are determined by evaluating the objective function $G_t^o(y \mid \pi_t)$. Motivated by this observation, we can construct an approximate solution, defined as $y_t^a(z \mid \pi_t)$, by solving the following equation:

$$G_t^o(y \mid \pi_t) = V_t^o(z \mid \pi_t) \cdot (1+\rho), \quad \text{s.t.} \quad y \geqslant y_t^o(z \mid \pi_t), \quad \rho \geqslant 0.$$

In the above equation, when $\rho = 0$, we have $y_t^a(z \mid \pi_t) = y_t^o(z \mid \pi_t)$; and when $\rho = (V_t(z \mid \pi_t) - V_t^o(z \mid \pi_t))/V_t^o(z \mid \pi_t)$, we have $y_t^a(z \mid \pi_t) = y_t^{u_T}(z \mid \pi_t)$. Therefore, to obtain an approximate optimal solution, we can choose ρ between zero and $(V_t(z \mid \pi_t) - V_t^o(z \mid \pi_t))/V_t^o(z \mid \pi_t)$ (the relative percentage of cost increase over the observed lost-sales case). The choice of ρ value can be determined by judgement or experience. Based on our numerical experience, a small ρ value tends to give a good approximation because the objective function $G_t^o(y \mid \pi_t)$ is flat around the minimum as illustrated in Figure 1. We will discuss the performance of this heuristic further in the numerical study section.

5.2. Heuristic Solution II

The above heuristic solution uses just the objective function in the observed lost-sales case. Below we propose an alternative heuristic solution based on the approximation of the first-order condition.

Let us first consider the special case when inventory is perishable. To distinguish this case from the nonperishable inventory case, let us add a superscript "p" (stands for "perishable") to all corresponding dynamic programming value functions and optimal solutions. Because there is no inventory carryover between periods, the on-hand inventory at the beginning of a period is always zero; the dynamic program optimality equations can be written as: for $t = 1, \ldots, T$,

$$V_t^p(0 \mid \pi_t) = \min_{y \geqslant 0}\{G_t^p(y \mid \pi_t)\} = \min_{y \geqslant 0}\{C(y \mid \pi_t) + E_{X_t \mid \pi_t}\{V_{t+1}^p(0 \mid \pi_{t+1}(\cdot \mid \pi_t, X_t \wedge y))\}\}, \tag{10}$$

with the terminal value $V_{T+1}^p(0 \mid \cdot) = 0$. The optimal solution to the above dynamic program is denoted by $y_t^p(\pi_t)$. Lu et al. (2008) have shown that

$$\frac{dG_t^p(y \mid \pi_t)}{dy} = \frac{dC(y \mid \pi_t)}{dy} + [V_{t+1}^p(0 \mid \pi) - \tilde{G}_{t+1}^p(\mathbf{y}_{t+1}^p(\pi') \mid \pi)] m(y \mid \pi_t),$$

where $\pi = \pi_{t+1}(\theta \mid \pi_t, X_t = y)$, $\pi' = \pi_{t+1}(\theta \mid \pi_t, X_t \geqslant y)$, and $\tilde{G}_{t+1}^p(\mathbf{y}_{t+1}^p(\pi') \mid \pi)$ is the expected cost when an "optimal" inventory policy $\mathbf{y}_{t+1}^p(\pi')$ determined by a starting prior π' is employed along the sample path, but the demand process evolves according to another starting prior π. As in the nonperishable inventory case, it is generally difficult, if not impossible, to compute the term $\tilde{G}_{t+1}^p(\mathbf{y}_{t+1}^p(\pi') \mid \pi)$. Below we show that there exists a recursive formula for $dG_t^p(y \mid \pi_t)/dy$, which simplifies the computation.

Proposition 7. *In a perishable inventory system with unobserved lost sales, the following holds*:

$$\frac{dG_t^p(y \mid \pi_t)}{dy} = [V_{t+1}^p(0 \mid \pi) - G_{t+1}^p(\hat{y} \mid \pi)] m(y \mid \pi_t) + E_{X_t \mid \pi_t}\left\{ \frac{d}{dy} G_{t+1}^p(y \mid \pi_{t+1}(\cdot \mid \pi_t, X_t \wedge \hat{y})) \right\},$$

where $\pi = \pi_{t+1}(\theta \mid \pi_t, X_t = y)$, $\pi' = \pi_{t+1}(\theta \mid \pi_t, X_t \geqslant y)$, $\hat{y} = y^p_{t+1}(\pi')$, *and* $G^p_t(\cdot \mid \cdot)$ *is the objective function defined in* (10).

It is worth commenting here that this nice recursive formula is obtained by relying on the fact that the optimal base-stock level is always achievable in the perishable inventory case. For the more general nonperishable inventory case, this premise is no longer true, and therefore we are unable to establish a similar result.

According to the observed lost-sales case (4), the first-order derivative of the objective function is given by

$$\frac{dG^o_t(y\mid\pi_t)}{dy} = \frac{dC(y\mid\pi_t)}{dy} + \int_0^y \frac{d}{dy} V^o_{t+1}(y-x\mid\pi_{t+1}(\cdot\mid\pi_t,x))m(x\mid\pi_t)\,dx.$$

In the unobserved lost-sales case (5), as shown by Lu et al. (2005b) and Chen (2009), the first-order derivative of the objective function is given by

$$\begin{aligned}\frac{dG_t(y\mid\pi_t)}{dy} &= \frac{dC(y\mid\pi_t)}{dy} \\ &\quad + \int_0^y \frac{d}{dy} V_{t+1}(y-x\mid\pi_{t+1}(\cdot\mid\pi_t,x))m(x\mid\pi_t)\,dx \\ &\quad + [V_{t+1}(0\mid\pi) - \tilde{G}_{t+1}(\mathbf{y}^*_{t+1}(\pi')\mid\pi)]m(y\mid\pi_t).\end{aligned}$$

We propose the following approximation for the above expression:

$$\begin{aligned}\frac{dG_t(y\mid\pi_t)}{dy} &\approx \frac{dC(y\mid\pi_t)}{dy} \\ &\quad + \int_0^y \frac{d}{dy} V^o_{t+1}(y-x\mid\pi_{t+1}(\cdot\mid\pi_t,x))m(x\mid\pi_t)\,dx \\ &\quad + [V^p_{t+1}(0\mid\pi) - \tilde{G}^p_{t+1}(\mathbf{y}^p_{t+1}(\pi')\mid\pi)]m(y\mid\pi_t) \\ &= \frac{dG^o_t(y\mid\pi_t)}{dy} + \frac{dG^p_t(y\mid\pi_t)}{dy} - \frac{dC(y\mid\pi_t)}{dy},\end{aligned}$$

where we use $dV^o_{t+1}(y\mid\cdot)/dy$ to approximate $dV_{t+1}(y\mid\cdot)/dy$ and $V^p_{t+1}(0\mid\pi) - \tilde{G}^p_{t+1}(\mathbf{y}^p_{t+1}(\pi')\mid\pi)$ to approximate $V_{t+1}(0\mid\pi) - \tilde{G}_{t+1}(\mathbf{y}^*_{t+1}(\pi')\mid\pi)$. As a result, an approximate optimal solution to the original problem (5) can be obtained by solving the approximate first-order condition

$$\frac{dG^o_t(y\mid\pi_t)}{dy} + \frac{dG^p_t(y\mid\pi_t)}{dy} - \frac{dC(y\mid\pi_t)}{dy} = 0, \qquad (11)$$

where $dG^p_t(y\mid\pi_t)/dy$ can be computed from the recursive formula given in Proposition 7.

It is known that $dG^o_t(y\mid\pi_t)/dy \geqslant dC(y\mid\pi_t)/dy$ (Scarf 1959) and $dG^p_t(y\mid\pi_t)/dy \leqslant dC(y\mid\pi_t)/dy$ (Lu et al. 2008). Therefore, it immediately follows that the solution obtained from Equation (11) lies between $y^o_t(0\mid\pi_t)$ and $y^p_t(\pi_t)$. When lost sales are observed, the optimal solution in the perishable inventory case is an upper bound for the optimal solution in the nonperishable inventory case (see Scarf 1959). This is due to the fact that the potential overstock in the subsequent period induces one to stock less. When lost sales are not observed, we are unable to establish this upper bound result analytically because in this case, a lower inventory level would also mean less learning about the demand distribution. Therefore, there is a trade-off between the overstock risk and the information acquisition. However, based on our numerical results shown in the next section, it appears that this upper-bound result continues to hold when lost sales are not observed.

6. Numerical Results

In this section, we present numerical study results on the bounds and heuristics developed in this paper. We first report results based on the gamma-exponential distribution. As we discussed in §4.1, the upper bound suggested by Lu et al. (2005b) is not applicable in this case, so we conduct a second numerical study based on the normal demand distribution with a two-point prior, where we can compare their result.

6.1. Gamma-Exponential Distribution

The one-dimensional dynamic program (9) for the gamma-gamma example can be further simplified. Note that $\phi_t(x)$ in (8) is the standard Pearson Type VI density (Johnson and Kotz 1970, p. 51). By the transformation $\tau = x/(1+x)$, $\phi_t(x)$ can be transformed to a standard beta density function with parameters k and $a_t = a + k(t-1)$:

$$\begin{aligned}\beta_{k,a_t}(\tau) &= \frac{\Gamma(k+a_t)}{\Gamma(k)\Gamma(a_t)} \cdot \left(\frac{\tau}{1-\tau}\right)^{k-1} \left(\frac{1}{1-\tau}\right)^{-(k+a_t)} \frac{1}{(1-\tau)^2} \\ &= \frac{\Gamma(k+a_t)}{\Gamma(k)\Gamma(a_t)} \cdot \tau^{k-1}(1-\tau)^{a_t-1}.\end{aligned}$$

Hence, by the property of the standard beta density, we can thus rewrite (9) as

$$\begin{aligned}v^o_t(z) &= \min_{y\geqslant z}\left\{L_t(y) + \int_0^1 \frac{1}{1-\tau} v^o_{t+1}((y(1-\tau)-\tau)^+) \cdot \beta_{k,a_t}(\tau)\,d\tau\right\} \\ &= \min_{y\geqslant z}\left\{L_t(y) + \frac{k+a_t-1}{a_t-1} \cdot \int_0^1 v^o_{t+1}((y(1-\tau)-\tau)^+)\beta_{k,a_t-1}(\tau)\,d\tau\right\},\end{aligned}$$

with $v^o_{T+1}(\cdot) = 0$, and

$$L_t(y) = p\left(\frac{k}{a_t-1} - y\right) + (h+p) \cdot \int_0^{y/(y+1)} \left(y - \frac{\tau}{1-\tau}\right)\beta_{k,a_t}(\tau)\,d\tau$$

$$= p\left(\frac{k}{a_t-1}-y\right)+(h+p)y\cdot I_{y/(y+1)}(k,a_t) - (h+p)\frac{k}{a_t-1}\cdot I_{y/(y+1)}(k+1,a_t-1), \qquad (12)$$

where $I_x(\cdot,\cdot)$ is the standard incomplete beta function. The minimizer of $L_t(y)$ is determined by solving the first-order condition $I_{y/(y+1)}(k,a_t)=p/(h+p)$.

To precisely evaluate the performance of the bounds and heuristics, we need to compute the exact optimal solution. As we have commented in the introduction section, the exact solution is computable in the case of the gamma-Weibull conjugate prior distribution family. Fortunately, the gamma-Weibull family shares a common special case with the gamma-gamma family—the gamma-exponential case ($k=1$ in the gamma-gamma family). Therefore, we can actually compute the exact solution in this special case. According to Lariviere and Porteus (1999), after some transformation, the gamma-exponential case with unobserved lost sales can be solved by the following two-dimensional dynamic program: for $t=1,\ldots,T$,

$$v_t(z\,|\,a_t)=\min_{y\geqslant z}\left\{L_t(y\,|\,a_t)+\int_0^{y/(y+1)}\frac{1}{1-\tau}\cdot v_{t+1}(y(1-\tau)-\tau\,|\,a_t+1)\cdot\beta_{1,a_t}(\tau)\,d\tau + (1+y)\cdot v_{t+1}(0\,|\,a_t)\cdot\int_{y/(y+1)}^1\beta_{1,a_t}(\tau)\,d\tau\right\}$$

$$=\min_{y\geqslant z}\left\{L_t(y\,|\,a_t)+\frac{a_t}{a_t-1}\cdot\int_0^{y/(y+1)}v_{t+1}(y(1-\tau)-\tau\,|\,a_t+1)\cdot\beta_{1,a_t-1}(\tau)\,d\tau + (1+y)\cdot v_{t+1}(0\,|\,a_t)\cdot(1-I_{y/(y+1)}(1,a_t))\right\},$$

with $v_{T+1}(\cdot\,|\,\cdot)=0$ and $L_t(y\,|\,a_t)$ specified as in (12) with $k=1$. Let $\bar{y}_t^*(z\,|\,a_t)$ be the optimal solution to the above dynamic program; then the optimal solution to the original problem is simply $y_t^*(z\,|\,\pi_t)=S_t\cdot\bar{y}_t^*(z/S_t\,|\,a_t)$, and the value function is $V_t(z\,|\,\pi_t)=S_t\cdot v_t(z/S_t\,|\,a_t)$. In the following, we conduct our numerical study based on this special case.

The numerical study is designed in the following manner. First, we hold the inventory-holding cost at $h=1$ and vary the stockout penalty cost by setting $p=5$ and 10. The implied critical ratios from these cost parameters are 83% and 91%, respectively. Second, we vary the gamma prior parameters (a,S) to study the effect of prior uncertainty on the bounds and heuristics. Note that the mean of the gamma prior is a/S and the coefficient of variation is $\sqrt{1/a}$. We choose (a,S) to be (3, 10) and (6, 10), so that the prior uncertainty (coefficient of variation) decreases ($\sqrt{1/a}=0.577$ and 0.408, respectively), whereas the prior mean remains the same ($a/S=0.3$). Finally, we vary the planning horizon by setting $T=3$, 5, and 10, so that we can study the effect of the planning horizon on the bounds and heuristics.

The numerical results for the bounds performance are presented in Table 1. Specifically, we present the optimal solution y_1^*, the lower bound y_1^o, the optimal solution in the perishable inventory case y_1^p, and three upper bounds: $y_1^{u_T}$, $y_1^{u_2}$, and y_1^{Lu} (the first two upper bounds are from Proposition 3, and the third one is from Lu et al. 2005b). In the table, we put the relative error percentage in parentheses next to the corresponding bounds. The relative error percentage is defined as

$$\%\ \text{error}(y)=\frac{G_1(y\,|\,\pi_1)-V_1(0\,|\,\pi_1)}{V_1(0\,|\,\pi_1)}. \qquad (13)$$

We choose this metric to measure the performance degradation due to the deviation from the optimal decision in the first period (note that the system is assumed to resume optimal control after the first period). Therefore, we can compare system performance under different initial inventory levels. This relative error percentage metric serves as the lower bound for the relative error percentage at the

Table 1. Bounds performance under the gamma-exponential distribution.

p	(a,S)	T	y_1^*	y_1^o	y_1^p	$y_1^{u_T}$	$y_1^{u_1}$	y_1^{Lu}
5	(3, 10)	3	7.81	7.58 (0.02%)	8.49 (0.18%)	9.87 (1.58%)	11.94 (5.95%)	N/A
		5	7.79	7.43 (0.04%)	8.66 (0.21%)	10.50 (1.96%)	14.04 (9.81%)	N/A
		10	7.75	7.38 (0.04%)	8.84 (0.14%)	11.42 (1.92%)	18.27 (15.9%)	N/A
	(6, 20)	3	6.85	6.81 (0.00%)	7.02 (0.01%)	7.66 (0.38%)	8.65 (1.87%)	N/A
		5	6.85	6.78 (0.00%)	7.06 (0.01%)	8.09 (0.56%)	9.83 (3.28%)	N/A
		10	6.89	6.78 (0.01%)	7.12 (0.02%)	8.82 (0.70%)	12.31 (5.67%)	N/A
10	(3, 10)	3	11.38	11.09 (0.02%)	12.66 (0.39%)	13.87 (1.42%)	17.67 (8.12%)	N/A
		5	11.10	10.76 (0.03%)	12.85 (0.52%)	14.29 (1.73%)	20.45 (13.6%)	N/A
		10	11.06	10.58 (0.03%)	13.03 (0.41%)	14.98 (1.65%)	26.00 (22.4%)	N/A
	(6, 20)	3	9.59	9.54 (0.00%)	9.90 (0.04%)	10.46 (0.31%)	12.11 (2.53%)	N/A
		5	9.56	9.48 (0.00%)	9.94 (0.04%)	10.85 (0.45%)	13.66 (4.45%)	N/A
		10	9.59	9.46 (0.00%)	10.00 (0.03%)	11.54 (0.54%)	16.97 (7.77%)	N/A

Note. y_1^* is the optimal solution; y_1^o is the lower bound; y_1^p is the optimal solution to the perishable inventory problem; $y_1^{u_T}$ and $y_1^{u_1}$ are two upper bounds based on Proposition 3; y_1^{Lu} is the upper bound based on Lu et al. (2005b).

policy level, which involves different inventory levels along the sample path.

A few observations can be made from Table 1. First, the lower bound is consistently tighter than the upper bounds in all cases. Second, the optimal solution in the perishable inventory case is a fairly tight upper bound in all cases. Third, as a increases, which means the prior uncertainty (coefficient of variation) decreases, the upper bounds all become tighter. However, when the horizon T increases, the opposite is true. As we have discussed in §4, the tightness of the solution upper bound is determined by the tightness of the value function upper bound. These numerical results reinforce this insight: When the prior uncertainty is reduced, the potential informational gain from Bayesian updating becomes smaller, and therefore the value function upper bound becomes tighter. On the other hand, as T increases, the potential informational gain from Bayesian updating increases, and therefore the value function upper bound becomes less tight.

In Table 2, we present the optimal solution, the Heuristic (I) given in §5.1 (with $\rho = 0.01\%$), the Heuristic (II) given in §5.2, and the myopic solution. In parentheses next to the corresponding heuristic solution is the relative error percentage as defined in (13). From the table, we can see that both Heuristic (I) and Heuristic (II) are consistently near-optimal in all cases, with Heuristic (II) giving a better performance. As a increases, which means the prior uncertainty decreases, the performance of the myopic solution improves. Finally, as the stockout penalty cost p increases, the performance of the myopic solution degrades, but the two heuristic solutions continue to be near-optimal.

6.2. Normal Distribution with a Two-Point Prior

In this section, we present a second set of numerical results based on the normal demand distribution with a two-point prior. In this numerical study, there are two candidate normal demand distributions with mean and standard deviation denoted by $N(\mu_1, \sigma_1)$ and $N(\mu_2, \sigma_2)$. The prior distribution is a two-point distribution given by $\pi_1 = (p_1, p_2)$, with $p_1 + p_2 = 1$.

Table 2. Comparison of heuristic solutions under the gamma-exponential distribution.

p	(a, S)	T	y_1^*	y_1^{h1}	y_1^{h2}	y_1^m
5	(3, 10)	3	7.81	7.74 (0.00%)	7.79 (0.00%)	8.17 (0.05%)
		5	7.79	7.63 (0.01%)	7.73 (0.00%)	8.17 (0.04%)
		10	7.75	7.61 (0.01%)	7.77 (0.00%)	8.17 (0.01%)
	(6, 20)	3	6.85	6.95 (0.00%)	6.85 (0.00%)	6.96 (0.01%)
		5	6.85	6.96 (0.00%)	6.86 (0.00%)	6.96 (0.00%)
		10	6.89	6.94 (0.00%)	6.90 (0.00%)	6.96 (0.00%)
10	(3, 10)	3	11.38	11.30 (0.00%)	11.35 (0.00%)	12.24 (0.18%)
		5	11.10	10.99 (0.01%)	11.09 (0.00%)	12.24 (0.21%)
		10	11.06	10.87 (0.01%)	10.99 (0.00%)	12.24 (0.14%)
	(6, 20)	3	9.59	9.70 (0.01%)	9.59 (0.00%)	9.83 (0.02%)
		5	9.56	9.67 (0.00%)	9.55 (0.00%)	9.83 (0.02%)
		10	9.59	9.59 (0.00%)	9.58 (0.00%)	9.83 (0.01%)

Note. y_1^* is the optimal solution; y_1^{h1} is the Heuristic (I) with $\rho = 0.01\%$; y_1^{h2} is the Heuristic (II); and y_1^m is the myopic solution.

We choose this example for two reasons. First, the upper bound suggested by Lu et al. (2005b) is computable, so we can make a direct comparison. Second, in this simple two-point prior scenario, the state-space of the prior distribution is one-dimensional, which makes the exact optimal solution computable; so, we can use it to precisely evaluate and compare the bounds and heuristics. We have also tried cases with more than two candidate demand distributions and found that the results of our bounds and heuristics are consistent, but the exact optimal solution becomes much more difficult to compute. Therefore, in the following, we will conduct our numerical study based on this simple two-point prior case.

The numerical study is designed in the following manner. First, we consider two parameter scenarios: (1) $\mu_1 = 100$, $\mu_2 = 200$, and $\sigma_1 = \sigma_2 = 100$; (2) $\mu_1 = 100$, $\mu_2 = 400$, and $\sigma_1 = \sigma_2 = 100$. Second, as in the gamma-exponential case, we hold the inventory-holding cost at $h = 1$ and vary the stockout penalty cost by setting $p = 5$ and 10. The implied critical ratios from these cost parameters are 83% and 91%. Third, we vary the prior parameters by setting $(p_1, p_2) = (0.2, 0.8)$, $(0.5, 0.5)$, and $(0.8, 0.2)$, so that we can study the effect of prior uncertainty on the bounds and heuristics. Finally, we vary the planning horizon by setting $T = 3$, 5, and 10. In the numerical computation, we truncate the normal distribution to ensure nonnegative demand.

The numerical results for the bounds performance are presented in Table 3. We observe that the lower bound is tighter than the upper bounds in all cases, and the optimal solution in the perishable inventory case is a fairly tight upper bound in all cases. These observations are consistent with what we have observed in the gamma-exponential case. We also observe that in the case of $\mu_2 = 200$, y_1^*, y_1^o, and y_1^p are almost identical. This indicates that when the two candidate distributions are close, the need for active learning is reduced. In addition, it appears that the upper bound becomes tighter as the stockout penalty cost p increases. In this two-point prior case, increasing the stockout penalty cost results in higher inventory level, which reduces the incremental informational gain from Bayesian updating; as a result, the upper bound for the value function becomes tight, and therefore the solution upper bound also becomes tight.

From Table 3, we note that the upper bound suggested by Lu et al. (2005b) falls in-between the most tight ($y_1^{u_T}$) and least tight ($y_1^{u_1}$) upper bounds of Proposition 3 in all cases, with the relative error percentage ranging from 0.15% to 18.2%. Lu et al. (2005b) have proposed to use a weighted average of the solution upper bound and lower bound as a heuristic solution. Their solution lower bound is a close approximate to that of Chen and Plambeck (2008), i.e, y_1^o. So, if the weight is chosen properly, their heuristic solution is essentially the same as our Heuristic (I) (see §5.1).

Table 3. Bounds performance under the normal distribution with a two-point prior.

μ_2	p	(p_1, p_2)	T	y_1^*	y_1^o	y_1^p	$y_1^{u_T}$	$y_1^{u_1}$	y_1^{Lu}
200	5	(0.2, 0.8)	3	286	286 (0.00%)	286 (0.00%)	288 (0.01%)	303 (0.47%)	309 (0.84%)
			5	286	286 (0.00%)	286 (0.00%)	288 (0.00%)	317 (0.90%)	324 (1.32%)
			10	286	286 (0.00%)	286 (0.00%)	293 (0.03%)	353 (1.89%)	349 (1.69%)
		(0.5, 0.5)	3	262	262 (0.00%)	263 (0.00%)	265 (0.02%)	292 (1.39%)	283 (0.70%)
			5	262	262 (0.00%)	263 (0.00%)	269 (0.05%)	316 (2.56%)	299 (1.26%)
			10	262	262 (0.00%)	263 (0.00%)	275 (0.08%)	369 (4.50%)	326 (1.78%)
		(0.8, 0.2)	3	231	231 (0.00%)	232 (0.00%)	238 (0.07%)	258 (1.16%)	245 (0.31%)
			5	232	231 (0.00%)	232 (0.00%)	242 (0.11%)	278 (2.04%)	258 (0.69%)
			10	232	231 (0.00%)	232 (0.00%)	249 (0.15%)	322 (3.55%)	284 (1.29%)
	10	(0.2, 0.8)	3	324	324 (0.00%)	324 (0.00%)	325 (0.00%)	339 (0.36%)	338 (0.32%)
			5	324	324 (0.00%)	324 (0.00%)	325 (0.00%)	353 (0.77%)	348 (0.54%)
			10	324	324 (0.00%)	324 (0.00%)	328 (0.01%)	390 (1.71%)	367 (0.80%)
		(0.5, 0.5)	3	302	301 (0.00%)	303 (0.00%)	304 (0.01%)	332 (1.36%)	316 (0.32%)
			5	302	301 (0.00%)	303 (0.00%)	306 (0.02%)	356 (2.43%)	326 (0.54%)
			10	302	301 (0.00%)	303 (0.00%)	310 (0.04%)	413 (4.36%)	345 (0.82%)
		(0.8, 0.2)	3	270	270 (0.00%)	271 (0.00%)	275 (0.04%)	299 (1.27%)	280 (0.15%)
			5	270	270 (0.00%)	271 (0.00%)	278 (0.06%)	321 (2.22%)	288 (0.30%)
			10	271	270 (0.00%)	271 (0.00%)	282 (0.07%)	371 (3.87%)	307 (0.61%)
400	5	(0.2, 0.8)	3	469	469 (0.00%)	481 (0.21%)	470 (0.00%)	553 (8.51%)	536 (5.69%)
			5	467	467 (0.00%)	481 (0.18%)	476 (0.07%)	622 (15.9%)	560 (6.77%)
			10	468	467 (0.00%)	481 (0.09%)	477 (0.05%)	769 (23.2%)	591 (5.91%)
		(0.5, 0.5)	3	412	410 (0.00%)	443 (1.17%)	422 (0.13%)	562 (22.8%)	508 (10.3%)
			5	410	409 (0.00%)	443 (0.89%)	428 (0.26%)	657 (38.3%)	535 (11.9%)
			10	410	409 (0.00%)	443 (0.50%)	429 (0.17%)	848 (54.3%)	569 (10.2%)
		(0.8, 0.2)	3	302	298 (0.02%)	327 (0.49%)	325 (0.49%)	502 (34.7%)	435 (15.8%)
			5	304	298 (0.02%)	327 (0.35%)	334 (0.59%)	595 (50.5%)	475 (18.2%)
			10	304	298 (0.01%)	327 (0.19%)	335 (0.34%)	770 (66.5%)	520 (15.9%)
	10	(0.2, 0.8)	3	511	511 (0.00%)	521 (0.15%)	511 (0.00%)	588 (6.98%)	556 (2.70%)
			5	510	509 (0.00%)	521 (0.13%)	515 (0.03%)	657 (13.4%)	575 (3.52%)
			10	509	509 (0.00%)	521 (0.07%)	516 (0.02%)	808 (19.8%)	602 (3.38%)
		(0.5, 0.5)	3	465	465 (0.00%)	491 (0.91%)	468 (0.01%)	602 (18.8%)	531 (5.31%)
			5	463	463 (0.00%)	491 (0.71%)	473 (0.09%)	699 (32.7%)	552 (6.45%)
			10	464	463 (0.00%)	491 (0.40%)	474 (0.06%)	897 (47.3%)	681 (5.91%)
		(0.8, 0.2)	3	377	376 (0.00%)	412 (1.24%)	391 (0.21%)	567 (31.8%)	470 (8.47%)
			5	376	376 (0.00%)	413 (0.93%)	396 (0.27%)	667 (48.4%)	498 (9.91%)
			10	376	376 (0.00%)	413 (0.52%)	397 (0.16%)	860 (65.2%)	535 (9.15%)

Note. y_1^* is the optimal solution; y_1^o is the lower bound; y_1^p is the optimal solution to the perishable inventory problem; $y_1^{u_T}$ and $y_1^{u_1}$ are two upper bounds based on Proposition 3; y_1^{Lu} is the upper bound based on Lu et al. (2005b).

In Table 4, we present the performance of Heuristic (I) and Heuristic (II), together with the myopic solution. Consistent with the gamma-exponential example, both Heuristic (I) and Heuristic(II) are near-optimal, with Heuristic (II) giving a better performance. They outperform the myopic solution by a wide margin in the case when the two candidate distributions are far apart. When the two candidate distributions are close (i.e., the case of $\mu_2 = 200$), the myopic solution performs as well as Heuristic (II). This again confirms that the need for active learning is reduced when the two candidate distributions are close. Given the fact that the solution upper bound is not always tight, choosing a proper weighting parameter in Heuristic (I) and that of Lu et al. (2005b) may not always be easy. In this sense, Heuristic (II) is not only a superior method, but also a more robust method because there is no need for choosing an ad hoc weighting parameter.

6.3. Comparison with Myopic Policies

In the previous sections, we have compared the bounds and heuristics with the optimal solution in the first period. In this section, we will compare the cost performance of employing various heuristic policies along the entire sample path. In particular, we will compare the two heuristic policies proposed in this paper with three myopic policies. The three myopic policies are defined as follows. Myopic-1 is the best-possible myopic policy, in which the posterior updating takes into account the censored data effect; this policy is optimal when inventory is perishable and future lost sales are observed. Myopic-2 is a "naive" heuristic policy, in which censored observations are treated as exact observations in the posterior updating process. Myopic-3 is a static policy with no Bayesian updating: The inventory level is determined in the first period and then remains fixed along the entire sample path. We define the relative

Table 4. Comparison of heuristic solutions under the normal distribution with a two-point prior.

μ_2	p	(p_1, p_2)	T	y_1^*	y_1^{h1}	y_1^{h2}	y_1^m
200	5	(0.2, 0.8)	3	286	289 (0.02%)	286 (0.00%)	286 (0.00%)
			5	286	289 (0.01%)	286 (0.00%)	286 (0.00%)
			10	286	291 (0.01%)	286 (0.00%)	286 (0.00%)
		(0.5, 0.5)	3	262	265 (0.02%)	262 (0.00%)	263 (0.00%)
			5	262	265 (0.01%)	262 (0.00%)	263 (0.00%)
			10	262	267 (0.01%)	262 (0.00%)	263 (0.00%)
		(0.8, 0.2)	3	231	234 (0.01%)	232 (0.00%)	231 (0.00%)
			5	232	235 (0.01%)	232 (0.00%)	231 (0.00%)
			10	232	236 (0.01%)	232 (0.00%)	231 (0.00%)
	10	(0.2, 0.8)	3	324	327 (0.02%)	324 (0.00%)	324 (0.00%)
			5	324	327 (0.01%)	324 (0.00%)	324 (0.00%)
			10	324	329 (0.01%)	324 (0.00%)	324 (0.00%)
		(0.5, 0.5)	3	302	305 (0.02%)	302 (0.00%)	302 (0.00%)
			5	302	305 (0.01%)	302 (0.00%)	302 (0.00%)
			10	302	306 (0.01%)	302 (0.00%)	302 (0.00%)
		(0.8, 0.2)	3	270	273 (0.01%)	271 (0.00%)	271 (0.00%)
			5	270	274 (0.01%)	271 (0.00%)	271 (0.00%)
			10	271	275 (0.01%)	271 (0.00%)	271 (0.00%)
400	5	(0.2, 0.8)	3	469	472 (0.01%)	469 (0.00%)	481 (0.21%)
			5	467	471 (0.01%)	468 (0.00%)	481 (0.18%)
			10	468	472 (0.01%)	468 (0.00%)	481 (0.09%)
		(0.5, 0.5)	3	412	414 (0.01%)	411 (0.00%)	443 (1.17%)
			5	410	413 (0.01%)	409 (0.00%)	443 (0.89%)
			10	410	414 (0.01%)	409 (0.00%)	443 (0.50%)
		(0.8, 0.2)	3	302	302 (0.00%)	302 (0.00%)	325 (0.49%)
			5	304	303 (0.00%)	302 (0.00%)	325 (0.29%)
			10	304	304 (0.00%)	302 (0.00%)	325 (0.16%)
	10	(0.2, 0.8)	3	511	514 (0.01%)	511 (0.00%)	521 (0.15%)
			5	510	513 (0.01%)	510 (0.00%)	521 (0.13%)
			10	509	514 (0.01%)	510 (0.00%)	521 (0.07%)
		(0.5, 0.5)	3	465	468 (0.01%)	465 (0.00%)	491 (0.91%)
			5	463	467 (0.01%)	464 (0.00%)	491 (0.71%)
			10	464	468 (0.01%)	464 (0.00%)	491 (0.40%)
		(0.8, 0.2)	3	377	380 (0.01%)	377 (0.00%)	412 (1.24%)
			5	376	380 (0.01%)	376 (0.00%)	412 (0.88%)
			10	376	381 (0.01%)	376 (0.00%)	412 (0.49%)

Note. y_1^* is the optimal solution; y_1^{h1} is the Heuristic (I) with $\rho = 0.01\%$; y_1^{h2} is the Heuristic (II); and y_1^m is the myopic solution.

percentage error in expected cost for each policy as

$$\%\ \text{error}(H) = \frac{V_1^H(0 \mid \pi_1) - V_1(0 \mid \pi_1)}{V_1(0 \mid \pi_1)},$$

where $V_1^H(0 \mid \pi_1)$ is the expected cost from applying the corresponding policy along the sample path.

Table 5 shows the policy cost performance under the gamma-exponential demand distribution with stockout penalty cost $p = 10$. The results from other scenarios are similar in order and thus provide no additional insights, so we omit them for brevity. From the table, it is clear that both Heuristic (I) and Heuristic (II) outperform the three myopic policies. We can also see the increased benefits from employing more sophisticated demand estimation and inventory optimization procedures. Specifically, the most significant improvement comes from upgrading Myopic-3 to Myopic-2, which shows that employing the Bayesian updating method when the demand distribution is not known is of first-order importance. The second-most significant improvement comes from upgrading Myopic-2 to Myopic-1, which demonstrates that factoring the censored data effect into the estimation procedure is important. The improvement from Myopic-1 to either Heuristic (I) or (II) is most prominent when the prior uncertainty is high. In this case, Myopic-1 runs the risk of overstock in the subsequent periods because it fails to account for the inventory carryover effect. Both Heuristic (I) and Heuristic (II) produce near-optimal performance in all cases, with Heuristic (II) giving a better performance. It is worth summarizing here the trade-offs between Heuristic (I) and Heuristic

Table 5. Cost performance of the heuristic policies under the gamma-exponential distribution.

p	(a, S)	T	Optimal cost V_1	Heuristic (I) ($\rho = 0.01\%$) (%)	Heuristic (II) (%)	Myopic-1 (best) (%)	Myopic-2 (naive) (%)	Myopic-3 (static) (%)
10	(3, 10)	3	51.46	0.01	0.00	0.21	1.58	7.03
		5	81.69	0.01	0.00	0.31	2.47	12.37
		10	151.25	0.04	0.01	0.24	3.56	21.38
	(6, 20)	3	34.59	0.01	0.00	0.03	0.35	2.26
		5	56.64	0.01	0.00	0.03	0.59	4.09
		10	109.79	0.03	0.01	0.03	1.01	7.39

(II): On the one hand, Heuristic (I) requires less computational effort because it only involves the evaluation of the objective function in the observed lost-sales case, but one needs to choose an ad hoc parameter ρ based on judgement or experience; on the other hand, Heuristic (II) requires additional computational effort because one needs to compute the derivatives, but it does provide a more precise and robust approximation to the optimal solution.

7. Concluding Remarks

In this paper, we have focused on developing bounds and heuristics for the Bayesian inventory control problem with unobserved lost sales. We have proved a sequence of lower and upper bounds for the value function of the problem. Based on the value function bounds, we have derived a sequence of solution upper bounds. These solution upper bounds work for all prior and demand distributions. For a fairly general monotone likelihood-ratio distribution family, we have developed relaxed but easily computable lower and upper bounds for the optimal solution along an arbitrary sample path.

We have also proposed two heuristics. The first heuristic is derived from the solution bound results. Computing this heuristic solution only requires the evaluation of the objective function in the observed lost sales. The second heuristic is based on the approximation of the first-order condition. We combine the derivatives of the observed lost-sales case and the perishable inventory case to obtain the approximation. For the latter case, we have established a recursive formula that simplifies the computation.

To evaluate and compare the bounds and heuristics, we have conducted an extensive numerical study. We compare our upper bounds with that of Lu et al. (2005b). We find that their upper bound, when it exists, falls in-between the most and the least tight bounds of ours in all cases. Our numerical results suggest that the optimal solution in the perishable inventory case is a fairly tight solution upper bound. It would be interesting to study whether this result can be established analytically in future research. The numerical results also indicate that both heuristics proposed in this paper are near-optimal (within 0.04%), with the second heuristic giving a better performance. They outperform the myopic policies by a wide margin. Based on our numerical experiences, we believe our second heuristic is a robust approximate for the optimal solution in more general cases where the exact optimal solution is difficult to compute.

Appendix

For ease of notation, let us define the predictive complement CDF given the prior π_t as

$$M(x \mid \pi_t) = \int_\Theta \bar{F}(x \mid \theta)\pi_t(\theta)\,d\theta. \tag{14}$$

Now let us first introduce a few identities to facilitate the proofs given in this appendix:

Lemma A3. *The following identities hold true*:
(a) $\pi_{t+1}(\theta \mid \pi_t, X_t = x)m(x \mid \pi_t) = f(x \mid \theta)\pi_t(\theta)$;
(b) $\pi_{t+1}(\theta \mid \pi_t, X_t \geqslant x)M(x \mid \pi_t) = \bar{F}(x \mid \theta)\pi_t(\theta)$;
(c) $m(x' \mid \pi_t(\cdot \mid \pi_{t-1}, x))m(x \mid \pi_{t-1}) = m(x \mid \pi_t(\cdot \mid \pi_{t-1}, x'))m(x' \mid \pi_{t-1})$;
(d) $m(x \mid \pi_t(\cdot \mid \pi_{t-1}, X_{t-1} \geqslant y))M(y \mid \pi_{t-1}) = M(y \mid \pi_t(\cdot \mid \pi_{t-1}, x))m(x \mid \pi_{t-1})$;

Proof. Identity (a) follows from definitions of (1) and (2). Identity (b) follows from definitions of (14) and (3). For part (c), we have

$$\begin{aligned} &m(x' \mid \pi_t(\cdot \mid \pi_{t-1}, x))m(x \mid \pi_{t-1})\\ &\quad= \int_\Theta f(x' \mid \theta)\pi_t(\cdot \mid \pi_{t-1}, x)\,d\theta \cdot m(x \mid \pi_{t-1})\\ &\quad= \int_\Theta f(x' \mid \theta)f(x \mid \theta)\pi_{t-1}(\theta)\,d\theta\\ &\quad= m(x \mid \pi_t(\cdot \mid \pi_{t-1}, x'))m(x' \mid \pi_{t-1}), \end{aligned}$$

where the second equality follows from part (a) and the last equality follows from the symmetry of the product form of $\int_\Theta f(x' \mid \theta)f(x \mid \theta)\pi_{t-1}(\theta)\,d\theta$. Part (d) can be shown by the same logic. □

Proof (Lemma 1). Part (a) can be shown as follows:

$$\begin{aligned} &E_{X_t \mid \pi_t}\{C(z \mid \pi_{t+1}(\cdot \mid \pi_t, X_t \wedge y))\}\\ &\quad= \int_0^y C(z \mid \pi_{t+1}(\cdot \mid \pi_t, x)) \cdot m(x \mid \pi_t)\,dx\\ &\qquad\quad + C(z \mid \pi_{t+1}(\cdot \mid \pi_t, X_t \geqslant y)) \cdot M(y \mid \pi_t)\\ &\quad= \int_0^y \left\{ \int_0^\infty c(z - x') \int_\Theta f(x' \mid \theta) \right.\\ &\qquad\qquad \left. \cdot\, \pi_{t+1}(\cdot \mid \pi_t, X_t = x) \cdot m(x \mid \pi_t)\,d\theta\,dx' \right\} dx \end{aligned}$$

$$+\int_0^\infty c(z-x')\int_\Theta f(x'\mid\theta)\cdot\pi_{t+1}(\cdot\mid\pi_t,X_t\geqslant y)\cdot M(y\mid\pi_t)\,d\theta\,dx'$$
$$=\int_0^y\left\{\int_0^\infty c(z-x')\int_\Theta f(x'\mid\theta)f(x\mid\theta)\pi_t(\theta)\,d\theta\,dx'\right\}dx$$
$$+\int_0^\infty c(z-x')\int_\Theta f(x'\mid\theta)\bar F(y\mid\theta)\pi_t(\theta)\,d\theta\,dx'$$
$$=\int_0^\infty c(z-x')\int_\Theta f(x'\mid\theta)\{F(y\mid\theta)+\bar F(y\mid\theta)\}\pi_t(\theta)\,d\theta\,dx'$$
$$=\int_0^\infty c(z-x')\int_\Theta f(x'\mid\theta)\pi_t(\theta)\,d\theta\,dx'=C(z\mid\pi_t),$$

where the third equality follows from Lemma A3(a) and (b), and the change of integration order in the fourth equality is permissible because the integrand is a nonnegative function. To show part (b), it suffices to show that

$$\int_y^\infty V^o_{t+1}(z\mid\pi_{t+1}(\cdot\mid\pi_t,x))\cdot m(x\mid\pi_t)\,dx\leqslant V^o_{t+1}(z\mid\pi_{t+1}(\cdot\mid\pi_t,X_t\geqslant y))\cdot M(y\mid\pi_t).$$

We prove the result by backward induction. For period $t=T$, it is trivial. Now, assume that the result holds for period $t+1$. For period t, we have

$$\int_y^\infty V^o_t(z\mid\pi_t(\cdot\mid\pi_{t-1},x))\cdot m(x\mid\pi_{t-1})\,dx$$
$$=\int_y^\infty\min_{y'\geqslant z}\{C(y'\mid\pi_t(\cdot\mid\pi_{t-1},x))+E_{X_t\mid\pi_t}\{V^o_{t+1}((y'-X_t)^+\mid\pi_{t+1}(\cdot\mid\pi_{t-1},X_{t-1}=x,X_t))\}\}m(x\mid\pi_{t-1})\,dx$$
$$\leqslant\min_{y'\geqslant z}\left\{\int_y^\infty C(y'\mid\pi_t(\cdot\mid\pi_{t-1},x))\cdot m(x\mid\pi_{t-1})\,dx+\int_y^\infty E_{X_t\mid\pi_t}\{V^o_{t+1}((y'-X_t)^+\mid\pi_{t+1}(\cdot\mid\pi_{t-1},X_{t-1}=x,X_t))\}\cdot m(x\mid\pi_{t-1})\,dx\right\}$$
$$=\min_{y'\geqslant z}\left\{\int_y^\infty\int_0^\infty c(y'-x')\cdot\int_\Theta f(x'\mid\theta)f(x\mid\theta)\pi_{t-1}(\theta)\,d\theta\,dx'\,dx+\int_0^\infty\int_y^\infty V^o_{t+1}((y'-x')^+\mid\pi_{t+1}(\cdot\mid\pi_{t-1},X_{t-1}=x,X_t=x'))\cdot m(x'\mid\pi_t(\cdot\mid\pi_{t-1},x))\cdot m(x\mid\pi_{t-1})\,dx\,dx'\right\}$$
$$=\min_{y'\geqslant z}\left\{C(y'\mid\pi_t(\cdot\mid\pi_{t-1},X_{t-1}\geqslant y))\cdot M(y\mid\pi_{t-1})+\int_0^\infty\int_y^\infty V^o_{t+1}((y'-x')^+\mid\pi_{t+1}(\cdot\mid\pi_{t-1},X_{t-1}=x',X_t=x))\cdot m(x\mid\pi_t(\cdot\mid\pi_{t-1},x'))\,dx\cdot m(x'\mid\pi_{t-1})\,dx'\right\}$$
$$\leqslant\min_{y'>z}\left\{C(y'\mid\pi_t(\cdot\mid\pi_{t-1},X_{t-1}\geqslant y))\cdot M(y\mid\pi_{t-1})+\int_0^\infty V^o_{t+1}((y'-x')^+\mid\pi_{t+1}(\cdot\mid\pi_{t-1},X_{t-1}=x',X_t\geqslant y))\cdot M(y\mid\pi_t(\cdot\mid\pi_{t-1},x'))\cdot m(x'\mid\pi_{t-1})\,dx'\right\}$$
$$=\min_{y'>z}\left\{C(y'\mid\pi_t(\cdot\mid\pi_{t-1},X_t\geqslant y))\cdot M(y\mid\pi_{t-1})+\int_0^\infty V^o_{t+1}((y'-x')^+\mid\pi_{t+1}(\cdot\mid\pi_{t-1},X_{t-1}\geqslant y,X_t=x'))\cdot m(x'\mid\pi_t(\cdot\mid\pi_{t-1},X_{t-1}\geqslant y))\,dx'\cdot M(y\mid\pi_{t-1})\right\}$$
$$=V^o_t(z\mid\pi_t(\cdot\mid\pi_{t-1},X_{t-1}\geqslant y))\cdot M(y\mid\pi_{t-1}),$$

where the second inequality follows from the induction assumption; the equality before the second inequality follows from Lemma A3(c) and the fact that $\pi_{t+1}(\cdot\mid\pi_{t-1},X_{t-1}=x,X_t=x')=\pi_{t+1}(\cdot\mid\pi_{t-1},X_{t-1}=x',X_t=x)$; and the equality after the second inequality follows from Lemma A3(d) and the fact that $\pi_{t+1}(\cdot\mid\pi_{t-1},X_{t-1}=x',X_t\geqslant y)=\pi_{t+1}(\cdot\mid\pi_{t-1},X_{t-1}\geqslant y,X_t=x')$. This completes the induction proof. □

Proof (Proposition 1). We show parts (a) and (b) together. It suffices to show that for any i, $G^{l_{i+1}}_t(y\mid\pi_t)\geqslant G^{l_i}_t(y\mid\pi_t)$ for all t. By taking the minimum over the inequality, we obtain $V^{l_{i+1}}_t(z\mid\pi_t)\geqslant V^{l_i}_t(z\mid\pi_t)$ for all $z\geqslant 0$. By definition, we have $G^{l_{i+1}}_t(y\mid\pi_t)=G^{l_i}_t(y\mid\pi_t)$ for $t\geqslant i+1$. For $t=i$, by definition, we have

$$G^{l_{i+1}}_i(y\mid\pi_i)$$
$$=C(y\mid\pi_i)+\int_0^y V^{l_{i+1}}_{i+1}(y-x\mid\pi_{i+1}(\cdot\mid\pi_i,x))m(x\mid\pi_i)\,dx+V^{l_{i+1}}_{i+1}(0\mid\pi_{i+1}(\cdot\mid\pi_i,X_i\geqslant y))M(y\mid\pi_i)$$
$$=C(y\mid\pi_i)+\int_0^y V^o_{i+1}(y-x\mid\pi_{i+1}(\cdot\mid\pi_i,x))m(x\mid\pi_i)\,dx+V^o_{i+1}(0\mid\pi_{i+1}(\cdot\mid\pi_i,X_i\geqslant y))M(y\mid\pi_i)$$
$$\geqslant C(y\mid\pi_i)+\int_0^y V^o_{i+1}(y-x\mid\pi_{i+1}(\cdot\mid\pi_i,x))m(x\mid\pi_i)\,dx+\int_y^\infty V^o_{i+1}(0\mid\pi_{i+1}(\cdot\mid\pi_i,x))m(x\mid\pi_i)\,dx$$
$$=G^o_i(y\mid\pi_i)=G^{l_i}_i(y\mid\pi_i),$$

where the inequality follows from Lemma 1(b). Hence, $V^{l_{i+1}}_i(z\mid\pi_i)\geqslant V^{l_i}_i(z\mid\pi_i)$ for all $z\geqslant 0$. Now assume that the result holds for $t=j$ $(j\leqslant i)$. For $t=j-1$, we have

$$G^{l_{i+1}}_{j-1}(y\mid\pi_{j-1})=C(y\mid\pi_{j-1})+E_{X_{j-1}\mid\pi_{j-1}}\{V^{l_{i+1}}_j((y-X_{j-1})^+\mid\pi_j(\cdot\mid\pi_{j-1},X_{j-1}\wedge y))\}$$

$$\geqslant C(y\mid\pi_{j-1})+E_{X_{j-1}\mid\pi_{j-1}}\{V_j^{l_i}((y-X_{j-1})^+ \mid \pi_j(\cdot\mid\pi_{j-1},X_{j-1}\wedge y))\}$$
$$=G_{j-1}^{l_i}(y\mid\pi_{j-1}),$$

where the inequality follows from the induction assumption that $V_j^{l_{i+1}}(z\mid\cdot)\geqslant V_j^{l_i}(z\mid\cdot)$. This completes the induction proof for Parts (a) and (b).

Analogously, to prove parts (c) and (d), it suffices to show that for any i, $G_t^{u_{i+1}}(y\mid\pi_t)\leqslant G_t^{u_i}(y\mid\pi_t)$ for all t. By taking the minimum over the inequality, we obtain $V_t^{u_{i+1}}(z\mid\pi_t)\leqslant V_t^{u_i}(z\mid\pi_t)$ for all $z\geqslant 0$. By definition, we have $G_t^{u_{i+1}}(y\mid\pi_t)=G_t^{u_i}(y\mid\pi_t)$ for $t\geqslant i+1$. For $t=i$, by definition, we have

$$\begin{aligned}
&G_i^{u_{i+1}}(y\mid\pi_i)\\
&=C(y\mid\pi_i)+\int_0^y V_{i+1}^{u_{i+1}}(y-x\mid\pi_{i+1}(\cdot\mid\pi_i,x))m(x\mid\pi_i)\,dx\\
&\quad+V_{i+1}^{u_{i+1}}(0\mid\pi_{i+1}(\cdot\mid\pi_i,X_i\geqslant y))M(y\mid\pi_i)\\
&=C(y\mid\pi_i)+\int_0^y (T-i)\cdot\min_{y'\geqslant y-x}\{C(y'\mid\pi_{i+1}(\cdot\mid\pi_i,x))\}m(x\mid\pi_i)\,dx\\
&\quad+(T-i)\cdot\min_{y'\geqslant 0}\{C(y'\mid\pi_{i+1}(\cdot\mid\pi_i,X_i\geqslant y))\}M(y\mid\pi_i)\\
&\leqslant C(y\mid\pi_i)+\int_0^y (T-i)\cdot C(y\mid\pi_{i+1}(\cdot\mid\pi_i,x))m(x\mid\pi_i)\,dx\\
&\quad+(T-i)\cdot C(y\mid\pi_{i+1}(\cdot\mid\pi_i,X_i\geqslant y))M(y\mid\pi_i)\\
&=C(y\mid\pi_i)+(T-i)\cdot C(y\mid\pi_i)\\
&=G_i^{u_i}(y\mid\pi_i),
\end{aligned}$$

where the second-to-last equality follows from Lemma 1(a). Hence, $V_i^{u_{i+1}}(z\mid\pi_i)\leqslant V_i^{u_i}(z\mid\pi_i)$ for all $z\geqslant 0$. Now assume that the result holds for $t=j$ $(j\leqslant i)$. For $t=j-1$, we have

$$\begin{aligned}
G_{j-1}^{u_{i+1}}(y\mid\pi_{j-1})&=C(y\mid\pi_{j-1})+E_{X_{j-1}\mid\pi_{j-1}}\{V_j^{u_{i+1}}((y-X_{j-1})^+ \mid\pi_j(\cdot\mid\pi_{j-1},X_{j-1}\wedge y))\}\\
&\leqslant C(y\mid\pi_{j-1})+E_{X_{j-1}\mid\pi_{j-1}}\{V_j^{u_i}((y-X_{j-1})^+ \mid\pi_j(\cdot\mid\pi_{j-1},X_{j-1}\wedge y))\}\\
&=G_{j-1}^{u_i}(y\mid\pi_{j-1}),
\end{aligned}$$

where the inequality follows from the induction assumption that $V_j^{u_{i+1}}(z\mid\cdot)\leqslant V_j^{u_i}(z\mid\cdot)$. This completes the induction proof for parts (c) and (d). □

Proof (Proposition 3). Let $y\geqslant y_t^o(z\mid\pi_t)$. By Proposition 1, we have

$$\begin{aligned}
G_t(y\mid\pi_t)-G_t(y^*\mid\pi_t)&=G_t(y\mid\pi_t)-V_t(z\mid\pi_t)\\
&\geqslant G_t^o(y\mid\pi_t)-V_t^{u_i}(z\mid\pi_t).
\end{aligned}$$

If we determine a $y_t^{u_i}$ such that $G_t^o(y_t^{u_i}\mid\pi_t)-V_t^{u_i}(z\mid\pi_t)=0$ and $y_t^{u_i}(z\mid\pi_t)\geqslant y_t^o(z\mid\pi_t)$, then we have $G_t(y\mid\pi_t)>G_t(y^*\mid\pi_t)$ for all $y>y_t^{u_i}(z\mid\pi_t)\geqslant y_t^o(z\mid\pi_t)\geqslant z$. This is because $G_t^o(y\mid\pi_t)$ is strictly increasing in y for $y\geqslant y_t^o(z\mid\pi_t)$ (convexity property). Hence, we conclude that the optimal Bayesian inventory level $y_t^*(z\mid\pi_t)$ must be no greater than $y_t^{u_i}(z\mid\pi_t)$. Furthermore, because $V_t^{u_i}(z\mid\pi_t)$ is decreasing in i, we have $y_t^{u_i}(z\mid\pi_t)$ also decreasing in i. □

Proof (Lemma 2). It is straightforward to verify parts (a) and (b) by the definition of likelihood-ratio ordering and the monotone likelihood-ratio property of $f(\cdot\mid\theta)$. We omit the proof here. To show part (c), we note, for $\theta\geqslant\theta'$,

$$\begin{aligned}
&\frac{\pi_{t+1}(\theta\mid\pi_t,X_t=x)}{\pi_{t+1}(\theta\mid\pi_t',X_t=x)}\\
&=\frac{f(x\mid\theta)\pi_t(\theta)}{\int_\Theta f(x\mid\theta)\pi_t(\theta)\,d\theta}\cdot\frac{\int_\Theta f(x\mid\theta)\pi_t'(\theta)\,d\theta}{f(x\mid\theta)\pi_t'(\theta)}\\
&=\frac{f(x\mid\theta)}{f(x\mid\theta)}\cdot\frac{\pi_t(\theta)}{\pi_t'(\theta)}\cdot\frac{\int_\Theta f(x\mid\theta)\pi_t'(\theta)\,d\theta}{\int_\Theta f(x\mid\theta)\pi_t(\theta)\,d\theta}\\
&\geqslant\frac{f(x\mid\theta')}{f(x\mid\theta')}\cdot\frac{\pi_t(\theta')}{\pi_t'(\theta')}\cdot\frac{\int_\Theta f(x\mid\theta)\pi_t'(\theta)\,d\theta}{\int_\Theta f(x\mid\theta)\pi_t(\theta)\,d\theta}\\
&=\frac{\pi_{t+1}(\theta'\mid\pi_t,X_t=x)}{\pi_{t+1}(\theta'\mid\pi_t',X_t=x)},
\end{aligned}$$

where the inequality follows from the definition of $\pi_t(\theta)\geqslant_{LR}\pi_t'(\theta)$. Hence, we conclude that $\pi_{t+1}(\theta\mid\pi_t,X_t=x)\geqslant_{LR}\pi_{t+1}(\theta\mid\pi_t',X_t=x)$. To show part (d), we note that likelihood-ratio ordering implies first-order stochastic dominance (Ross 1983). Because $f(x\mid\theta)$ satisfies the monotone likelihood-ratio property, it follows that $f(x\mid\theta)\geqslant_{FS}f(x\mid\theta')$ for all $\theta\geqslant\theta'$, or equivalently, $\bar F(x\mid\theta)$ is increasing in θ. Also, $\pi_t(\theta)\geqslant_{LR}\pi_t'(\theta)$ implies $\pi_t(\theta)\geqslant_{FS}\pi_t'(\theta)$. By the property of first-order stochastic dominance (Porteus 2002), we immediately have $\int_\Theta\bar F(x\mid\theta)\pi_t(\theta)\,d\theta\geqslant\int_\Theta\bar F(x\mid\theta)\pi_t'(\theta)\,d\theta$ or $\int_\Theta\int_x^\infty f(\xi\mid\theta)\,d\xi\,\pi_t(\theta)\,d\theta\geqslant\int_\Theta\int_x^\infty f(\xi\mid\theta)\,d\xi\,\pi_t'(\theta)\,d\theta$. By changing the order of integration, we have $\int_x^\infty\int_\Theta f(\xi\mid\theta)\pi_t(\theta)\,d\theta\,d\xi\geqslant\int_x^\infty\int_\Theta f(\xi\mid\theta)\pi_t'(\theta)\,d\theta\,d\xi$ or $\int_x^\infty m(\xi\mid\pi_t)d\xi\geqslant\int_x^\infty m(\xi\mid\pi_t')\,d\xi$. By the definition of FSD, we have $m(x\mid\pi_t)\geqslant_{FS}m(x\mid\pi_t')$. □

Proof (Proposition 4). It suffices to show that $dG_t^o(y\mid\pi_t)/dy\leqslant dG_t^o(y\mid\pi_t')/dy$. We prove the result by backward induction. For period $t=T$, because $\pi_T(\theta)\geqslant_{LR}\pi_T'(\theta)$, by Lemma 2(d), $m(x\mid\pi_T)\geqslant_{FS}m(x\mid\pi_T')$. Because the last-period problem is a standard newsvendor problem, it is straightforward to show that

$$\frac{dG_T^o(y\mid\pi_T)}{dy}=\frac{dC(y\mid\pi_T)}{dy}=-p+(h+p)F_T(y),$$

where $F_T(y)=\int_0^y m(x\mid\pi_T)\,dx$. Note that $m(x\mid\pi_T)\geqslant_{FS}m(x\mid\pi_T')$ implies $\int_0^y m(x\mid\pi_T)\,dx\leqslant\int_0^y m(x\mid\pi_T')\,dx$. Hence, $dG_T^o(y\mid\pi_T)/dy\leqslant dG_T^o(y\mid\pi_T')/dy$.

Now, assume that the result holds for period $t+1$. For period t, we have

$$\frac{dG_t^o(y\mid\pi_t)}{dy}=\frac{dC(y\mid\pi_t)}{dy}+\int_0^y\frac{dV_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t,x))}{dy}m(x\mid\pi_t)\,dx.$$

Because $\pi_t(\theta)\geqslant_{LR}\pi_t'(\theta)$, from the case of period T, we have $dC(y\mid\pi_t)/dy\leqslant dC(y\mid\pi_t')/dy$. Therefore, it remains to show that

$$\int_0^y\frac{dV_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t,x))}{dy}m(x\mid\pi_t)\,dx\leqslant\int_0^y\frac{dV_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t',x))}{dy}m(x\mid\pi_t')\,dx.$$

From Lemma 2(c), we have $\pi_{t+1}(\cdot\mid\pi_t,x)\geqslant_{LR}\pi_{t+1}(\cdot\mid\pi_t',x)$. Hence,

$$\begin{aligned}\frac{d}{dy}V_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t,x))&=\max\left\{0,\frac{d}{dy}G_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t,x))\right\}\\&\leqslant\max\left\{0,\frac{d}{dy}G_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t',x))\right\}\\&=\frac{d}{dy}V_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t',x)),\end{aligned}\tag{15}$$

where the inequality follows from the induction assumption.

Now let us show that $dV_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t',x))/dy$ is decreasing in x for $y\geqslant x\geqslant 0$. For any $y\geqslant x_1\geqslant x_2\geqslant 0$, from Lemma 2(a), we have $\pi_{t+1}(\cdot\mid\pi_t',x_1)\geqslant_{LR}\pi_{t+1}(\cdot\mid\pi_t',x_2)$. Hence,

$$\begin{aligned}\frac{d}{dy}V_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t',x_1))&=\max\left\{0,\frac{d}{dy}G_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t',x_1))\right\}\\&\leqslant\max\left\{0,\frac{d}{dy}G_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t',x_2))\right\}\\&=\frac{d}{dy}V_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t',x_2)),\end{aligned}$$

where the inequality follows from the induction assumption.

Because $\pi_t(\theta)\geqslant_{LR}\pi_t'(\theta)$, from Lemma 2(d), we have $m(x\mid\pi_t)\geqslant_{FS}m(x\mid\pi_t')$. By the property of first-order stochastic dominance, $dV_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t',x))/dy$ decreasing in x implies

$$\int_0^y\frac{dV_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t',x))}{dy}m(x\mid\pi_t)\,dx\leqslant\int_0^y\frac{dV_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t',x))}{dy}m(x\mid\pi_t')\,dx.$$

Combining this with inequality (15), we obtain

$$\begin{aligned}&\int_0^y\frac{dV_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t,x))}{dy}m(x\mid\pi_t)\,dx\\&\quad\leqslant\int_0^y\frac{dV_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t',x))}{dy}m(x\mid\pi_t)\,dx\\&\quad\leqslant\int_0^y\frac{dV_{t+1}^o(y-x\mid\pi_{t+1}(\cdot\mid\pi_t',x))}{dy}m(x\mid\pi_t')\,dx.\end{aligned}$$

This completes the induction proof. □

Proof (Proposition 5). By repeatedly applying Lemma 2(b) and (c) for cases when $x_i=y_i$, i.e., $X_i\geqslant x_i$, we have

$$\pi_{t+1}(\theta\mid\pi_1,X_i\wedge y_i=x_i,1\leqslant i\leqslant t)\geqslant_{LR}\pi_{t+1}(\theta\mid\pi_1,X_i=x_i,1\leqslant i\leqslant t).$$

Then, by Proposition 4, the result follows. □

Proof (Proposition 6). It suffices to show that $G_{t+1}^o(y\mid\pi_{t+1})\geqslant G_{t+1}^{o'}(y\mid\pi_{t+1})$. Because the posterior is $\pi_{t+1}=\pi_{t+1}(\theta\mid\pi_1,X_i\wedge y_i=x_i,1\leqslant i\leqslant t)$, we have

$$\begin{aligned}&G_{t+1}^o(y\mid\pi_{t+1})\\&\quad=C(y\mid\pi_{t+1})+E_{X_{t+1}\mid\pi_{t+1}}\{V_{t+2}^o((y-X_{t+1})^+\mid\pi_{t+2}(\cdot\mid\pi_{t+1},X_{t+1}))\}\\&\quad\geqslant C(y\mid\pi_{t+1})+E_{X_{t+1}\mid\pi_{t+1}}\{V_{t+2}^o(0\mid\pi_{t+2}(\cdot\mid\pi_{t+1},X_{t+1}))\}\\&\quad=C(y\mid\pi_{t+1})+E_{X_1,\ldots,X_{t+1}\mid\pi_1}\{V_{t+2}^o(0\mid\pi_{t+2}(\cdot\mid\pi_1,X_i\wedge y_i=x_i,X_{t+1},1\leqslant i\leqslant t)\}\\&\quad\geqslant C(y\mid\pi_{t+1})+E_{X_1,\ldots,X_{t+1}\mid\pi_1}\{V_{t+2}^o(0\mid\pi_{t+2}(\cdot\mid\pi_1,X_1,\ldots,X_{t+1}))\mid X_i\wedge y_i=x_i,1\leqslant i\leqslant t\},\end{aligned}$$

where the last inequality follows from applying Lemma 1(b) repeatedly for cases when $x_i=y_i$, i.e., $X_i\geqslant y_i$. Now, by the same reasoning used in the proof of Proposition 3, the result follows. □

Proof (Proposition 7). First, let $\pi=\pi_{t+1}(\theta\mid\pi_t,X_t=y)$ and $\pi'=\pi_{t+1}(\theta\mid\pi_t,X_t\geqslant y)$. From Lu et al. (2008), we know that

$$\begin{aligned}&\frac{dG_t^p(y\mid\pi_t)}{dy}-\frac{dC(y\mid\pi_t)}{dy}\\&\quad=[V_{t+1}^p(0\mid\pi)-\tilde G_{t+1}^p(\mathbf{y}_{t+1}^p(\pi')\mid\pi)]m(y\mid\pi_t)\\&\quad=[V_{t+1}^p(0\mid\pi)-G_{t+1}^p(y_{t+1}^p(\pi')\mid\pi)]m(y\mid\pi_t)\\&\qquad+[G_{t+1}^p(y_{t+1}^p(\pi')\mid\pi)-\tilde G_{t+1}^p(\mathbf{y}_{t+1}^p(\pi')\mid\pi)]\\&\qquad\cdot m(y\mid\pi_t).\end{aligned}\tag{16}$$

Let us focus on the second term. Let $\hat{y} = y^p_{t+1}(\pi')$. We have

$$
\begin{aligned}
&[G^p_{t+1}(y^p_{t+1}(\pi') \mid \pi) - \tilde{G}^p_{t+1}(\mathbf{y}^p_{t+1}(\pi') \mid \pi)]m(y \mid \pi_t)\\
&\quad = \int_0^{\hat{y}} [V^p_{t+2}(0 \mid \pi_{t+2}(\cdot \mid \pi, x)) - \tilde{G}^p_{t+2}(\mathbf{y}^p_{t+2}(\pi_{t+2}(\cdot \mid \pi', x))\\
&\qquad\qquad \mid \pi_{t+2}(\cdot \mid \pi, x))]m(x \mid \pi)m(y \mid \pi_t)\, dx\\
&\qquad + [V^p_{t+2}(0 \mid \pi_{t+2}(\cdot \mid \pi, X_{t+1} \geqslant \hat{y}))\\
&\qquad\qquad - \tilde{G}^p_{t+2}(\mathbf{y}^p_{t+2}(\pi_{t+2}(\cdot \mid \pi', X_{t+1} \geqslant \hat{y}))\\
&\qquad\qquad\qquad \mid \pi_{t+2}(\cdot \mid \pi, X_{t+1} \geqslant \hat{y}))]\\
&\qquad \cdot M(\hat{y} \mid \pi)m(y \mid \pi_t).
\end{aligned} \tag{17}
$$

Now let $\pi''(x) = \pi_{t+1}(\theta \mid \pi_t, X_t = x)$. By the product form of posterior distribution, it is easy to verify that $\pi_{t+2}(\cdot \mid \pi, x) = \pi_{t+2}(\cdot \mid \pi''(x), y)$ and $\pi_{t+2}(\cdot \mid \pi', x) = \pi_{t+2}(\cdot \mid \pi''(x), X_{t+1} \geqslant y)$. From Lemma A3(c), we have $m(x \mid \pi)m(y \mid \pi_t) = m(y \mid \pi''(x))m(x \mid \pi_t)$.

Similarly, let $\pi'''(\hat{y}) = \pi_{t+1}(\theta \mid \pi_t, X_t \geqslant \hat{y})$. By the product form of posterior, it is easy to verify that $\pi_{t+2}(\cdot \mid \pi, X_{t+1} \geqslant \hat{y}) = \pi_{t+2}(\cdot \mid \pi'''(\hat{y}), y)$ and $\pi_{t+2}(\cdot \mid \pi', X_{t+1} \geqslant \hat{y}) = \pi_{t+2}(\cdot \mid \pi'''(\hat{y}), X_{t+1} \geqslant y)$. From Lemma A3(d), we have $M(\hat{y} \mid \pi)m(y \mid \pi_t) = m(y \mid \pi'''(\hat{y}))M(\hat{y} \mid \pi_t)$. Substitute these identities into (17). We arrive at

$$
\begin{aligned}
&[G^p_{t+1}(y^p_{t+1}(\pi') \mid \pi) - \tilde{G}^p_{t+1}(\mathbf{y}^p_{t+1}(\pi') \mid \pi)]m(y \mid \pi_t)\\
&\quad = \int_0^{\hat{y}} [V^p_{t+2}(0 \mid \pi_{t+2}(\cdot \mid \pi''(x), y))\\
&\qquad\qquad - \tilde{G}^p_{t+2}(\mathbf{y}^p_{t+2}(\pi_{t+2}(\cdot \mid \pi''(x), X_{t+1} \geqslant y))\\
&\qquad\qquad\qquad \mid \pi_{t+2}(\cdot \mid \pi''(x), y))]\\
&\qquad \cdot m(y \mid \pi''(x))m(x \mid \pi_t)\, dx\\
&\qquad + [V^p_{t+2}(0 \mid \pi_{t+2}(\cdot \mid \pi'''(\hat{y}), y))\\
&\qquad\qquad - \tilde{G}^p_{t+2}(\mathbf{y}^p_{t+2}(\pi_{t+2}(\cdot \mid \pi'''(\hat{y}), X_{t+1} \geqslant y))\\
&\qquad\qquad\qquad \mid \pi_{t+2}(\cdot \mid \pi'''(\hat{y}), y))]\\
&\qquad \cdot m(y \mid \pi'''(\hat{y}))M(\hat{y} \mid \pi_t)\\
&\quad = \int_0^{\hat{y}} \left[\frac{dG^p_{t+1}(y \mid \pi''(x))}{dy} - \frac{dC(y \mid \pi''(x))}{dy}\right] m(x \mid \pi_t)\, dx\\
&\qquad + \left[\frac{dG^p_{t+1}(y \mid \pi'''(\hat{y}))}{dy} - \frac{dC(y \mid \pi'''(\hat{y}))}{dy}\right] M(\hat{y} \mid \pi_t)\\
&\quad = E_{X_t \mid \pi_t}\left\{\frac{d}{dy} G^p_{t+1}(y \mid \pi_{t+1}(\cdot \mid \pi_t, X_t \wedge \hat{y}))\right\}\\
&\qquad - \frac{dC(y \mid \pi_t)}{dy},
\end{aligned} \tag{18}
$$

where the second equality follows from the original formula for $dG^p_{t+1}(y \mid \cdot)/dy$ and the last equality follows from the fact that

$$
\begin{aligned}
&\int_0^{\hat{y}} \frac{dC(y \mid \pi''(x))}{dy} m(x \mid \pi_t)\, dx + \frac{dC(y \mid \pi'''(\hat{y}))}{dy} M(\hat{y} \mid \pi_t)\\
&\quad = \int_0^{\hat{y}} \left[(h+p)\int_0^y m(\xi \mid \pi''(x))\, d\xi - p\right] m(x \mid \pi_t)\, dx\\
&\qquad + \left[(h+p)\int_0^y m(\xi \mid \pi'''(\hat{y}))\, d\xi - p\right] M(\hat{y} \mid \pi_t)\\
&\quad = (h+p)\cdot\left[\int_0^{\hat{y}}\int_0^y f(\xi \mid \theta) f(x \mid \theta)\pi_t(\theta)\, d\theta\, d\xi\, dx\right.\\
&\qquad\qquad \left. + \int_0^y f(\xi \mid \theta)\bar{F}(\hat{y} \mid \theta)\pi_t(\theta)\, d\theta\, d\xi\right] - p\\
&\quad = (h+p)\int_0^y f(\xi \mid \theta)\pi_t(\theta)\, d\theta\, d\xi - p = \frac{dC(y \mid \pi_t)}{dy},
\end{aligned}
$$

where the second equality follows from Lemma A3(c) and (d). Now, substitute (18) back into (16). We obtain the recursive formula. □

Acknowledgments

The author thanks Erica Plambeck, Paul Zipkin, the associate editor, and two anonymous referees for their helpful comments and suggestions.

References

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

Braden, D. J., M. Freimer. 1991. Informational dynamics of censored observations. *Management Sci.* **37**(11) 1390–1404.

Chen, L. 2009. An envelope theorem for Bayesian dynamic program and its application to an inventory problem. Working paper, Duke University, Durham, NC.

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.

DeGroot, M. H. 1970. *Optimal Statistical Decisions*. McGraw-Hill Book Company, New York.

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

Eppen, G. D., A. V. Iyer. 1997. Improved fashion buying with Bayesian updates. *Oper. Res.* **45**(6) 805–819.

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.

Heyman, D. P., M. J. Sobel. 1984. *Stochastic Models in Operations Research, Volume II: Stochastic Optimization*. McGraw-Hill Book Company, New York.

Johnson, N. L., S. Kotz. 1970. *Continuous Univariate Distributions—2*. John Wiley & Sons, New York.

Karlin, S., H. Rubin. 1956. Distributions possessing a monotone likelihood ratio. *J. Amer. Statist. Assoc.* **51**(276) 637–643.

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

Lovejoy, W. S. 1990. Myopic policies for some inventory models with uncertain demand distributions. *Management Sci.* **36**(6) 724–738.

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

Lu, X., J. S. Song, K. Zhu. 2005b. Inventory control with unobservable lost sales and Bayesian updates. Working paper, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong.

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

Nahmias, S. 1994. Demand estimation in lost sales inventory systems. *Naval Res. Logist.* **41** 739–757.

Porteus, E. 2002. *Foundations of Stochastic Inventory Theory*. Stanford University Press, Stanford, CA.

Ross, S. 1983. *Stochastic Processes*. John Wiley & Sons, New York.

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.