Submitted to
manuscript (Please, provide the mansucript number!)

Bayesian Inventory Management with Potential Change-Points in Demand

Zhe (Frank) Wang, Adam J. Mersereau
Kenan-Flagler Business School, University of North Carolina, zhe_wang@unc.edu, ajm@unc.edu

We consider the inventory management problem of a firm experiencing potential demand shifts whose timings are known but whose impacts are not known. Examples include global news events (e.g., the 9/11 terrorist attacks), local events (e.g., the opening of a nearby attraction), or internal events (e.g., a production redesign). In the periods following such a potential change-point in demand, a manager is torn between relying on a possibly obsolete demand parameters based on a long history and using fresh but less precisely estimated parameters based on recent history. We formulate a Bayesian inventory problem in which a potential change-point has recently occurred. After establishing some structural results about optimal policies, we construct novel cost lower bounds based on particular information relaxations and propose heuristic policies derived from those bounds. Our bounds and heuristics achieve small optimality gaps in numerical studies.

Key words: inventory theory and control; retailing; dynamic programming
History: original version: August 14, 2013; this version: February 17, 2014

1. Introduction

In most real-world inventory control problems, demand changes over time and the true underlying demand distribution is never fully known to the inventory manager. The manager makes dual use of historical demand data to populate the current demand distribution but also to detect fundamental changes in the demand-generating process.

We provide two data examples in Figure 1 to illustrate the complexity of the manager's task. Figure 1(a) shows seasonally adjusted monthly sales by motor vehicle dealers in the United States before and after September 2001. Imagine the situation faced by an automobile dealer in the autumn of 2001. While a reasonable dealer would expect the September 2001 attacks to impact consumer demand for automobiles, the direction and magnitude of the impact would have been difficult to predict from data available at the time. In October 2001 sales spiked substantially, but was this just a temporary surge or an indicator of a new regime in automobile sales? Was pre-October historical data still useful for understanding demand in October and beyond? History shows that demand eventually fell back close to its pre-September levels, but this might have been unclear at the time.

Figure 1(b) shows seasonally adjusted monthly sales for U.S. women's clothing stores in 2008 and 2009. Uncertainty in the financial markets reached a crescendo in September 2008 with the backruptcy of investment bank Lehman Brothers. Even if a women's clothing retailer at the time anticipated a negative impact on garment sales, the magnitude and persistence of the impact would have been harder to anticipate. It turns out that adjusted women's clothing demand bottomed out in December 2008 and stayed close to its December 2008 levels for over a year afterwards. In hindsight, we see that the Lehman Brothers bankruptcy marked a distinct change in women's clothing demand that rendered the previous demand history unsuitable for understanding new demand levels.

These two examples illustrate what we believe is a common challenge faced by retail and other managers, namely how to respond to external events that have the potential to change the demand environment. While September 2001 and the Lehman Brothers bankruptcy are well-known events that impacted many firms across many industries, demand-changing events can also be local. For example, the start of a new marketing campaign, the entrance of a new competitor, and the opening of a nearby attraction can all potentially usher in new demand regimes for a firm. The introduction of a redesigned or a new product could also be interpreted as a potential demand-changing event when historical demand or sales data from a similar product are available for generating a reference forecast. All of these events have in common that their timing is known but their impact is not. In the periods soon after such events, the manager can rely on historical demand to carefully estimate possibly obsolete demand parameters, discard the historical demand data and instead re-estimate demand parameters based on a limited history, or do something in between. The tradeoff inherent to this problem is between the precision brought by a long (but possibly out-of-date) history and the responsiveness that comes from relying on a recent (but limited) history.

We refer to such events as *potential change-points* in demand, and we present and analyze an inventory model that explicitly allows for potential change-points. We focus on the case in which there is a single potential change-point in the recent past, which is relevant to the examples of Figure 1 and to other examples in which change-points occur relatively infrequently. We seek answers to a few research questions in this context: What does an optimal policy look like? Can it be efficiently computed or approximated? Are there general guidelines for dynamically adjusting inventory levels in response to observed demands?

We model the evolution of the manager's belief on the demand process using a Bayesian framework, extending the model pioneered by Scarf (1959). We first leverage the structure of our demand model to characterize the effects of observed demand and the manager's belief on the demand process and the optimal (state-dependent) base-stock levels.

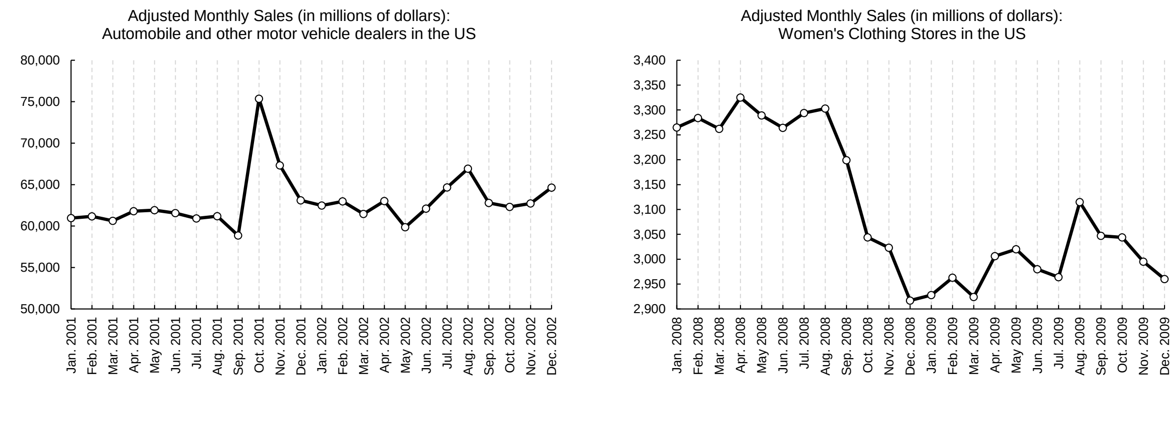
The optimal policy remains computationally challenging, so we pursue heuristic policies coupled with cost lower bounds. Our most sophisticated bounding approach is novel in its formulation of an "independentized" problem that relaxes the dependence between physical demand and demand signals. A particular information relaxation of the demand signal information yields efficient subproblems that are solutions to stochastic multiperiod inventory problems with known demand distributions. An extensive numerical analysis reveals that this bound and an associated policy achieve small optimality gaps relative to other candidate policies, but also that a myopic policy that accounts for potential change-points (but that ignores future inventory dynamics) is near-optimal except in extreme instances.

We also consider the sensitivity of our inventory policies to misspecification of the parameters of the manager's Bayesian prior. Taking a maximin profit perspective, we show that a conservative manager worried about profit downside will follow a policy that assumes the smallest prior (in a sense we will make precise) among a set of candidates.

The remainder of this paper is organized as follows. Section 2 reviews related literature. In Section 3, we formulate our Bayesian demand model and associated inventory control problem, and we present some structural properties about the optimal inventory policy. Section 4 develops lower bounds for the optimal expected cost and several heuristic policies. We numerically study these bounds and policies and measure their performance in Section 5. Section 6 includes discussion on estimating model parameters and sensitivity to parameter misspecification. We conclude in Section 7.

2. Literature Review

This paper is related to the inventory control literature dealing with nonstationary and/or partially observed demand processes. For the situations where the demand is nonstationary but the demand distributions are known, Karlin (1960) analyzes a dynamic inventory system in which demands are stochastic and may vary from period to period and proves the optimality of state-dependent base-stock policies. Song and Zipkin (1993, 1996) propose a continuous-time Markov-modulated Poisson demand framework to model inventory management problems under a fluctuating demand environment. It is assumed that the demand distribution changes regime according to a known Markov chain and the demand distribution in each regime is also fully known. Optimality of state-dependent (s, S) policies is established. Sethi and Cheng (1997) give similar results in a generalized



(a) (b)

Figure 1 Examples of potential changepoints in demand. (Source: U.S. Census Bureau)

discrete-time inventory model with Markov-modulated demands. Graves (1999) characterizes the behavior of an adaptive base-stock policy under an ARIMA demand process. Iida and Zipkin (2006) and Lu et al. (2006) study approximate solutions for inventory planning problems with demand forecasting based on the martingale model of forecast evolution (MMFE).

Using a Bayesian framework, Scarf (1959) pioneers the study of optimal inventory policies under a stationary demand process with an unknown demand distribution parameter. Our work extends this framework to general demand distributions with a more flexible belief structure. Scarf (1960) and Azoury (1985) provide conditions under which the dimensionality of the problem can be reduced and the optimal base-stock levels can be obtained by solving a one-dimensional dynamic program. Our heuristics make possible the computation of approximate solutions to problems with more general prior and demand distributions. Azoury and Miyaoka (2009) study a Bayesian inventory problem where demand in each period depends on side information through a linear regression model. All of these works assume, as we do, that demand is fully observable and backlogged. There is another stream of research looking at inventory management problems when lost sales are unobserved and demand is therefore censored, assuming stationary demand. See, for example, Lariviere and Porteus (1999), Ding et al. (2002), Chen and Plambeck (2008), Bensoussan et al. (2007, 2008), Chen (2010), and Huh et al. (2011). Chen and Mersereau (2013) include a survey of this literature.

The demand process we consider is also related to that of Treharne and Sox (2002), who assume a Markov-modulated demand process in which state transitions are unobserved but the manager knows the transition probability matrix and maintains a belief of the underlying Markov state. They evaluate several heuristics, including limited lookahead policies, numerically. Brown et al.

(2010) apply information relaxation bounds to an extended version of Treharne and Sox (2002)'s model with non-stationary cost parameters. Our model differs from these in two important respects. First, we assume a single potential regime change assumed to be in the past. This simplification yields structure that we take advantage of in deriving new results and bounds. Second, we model component demand distributions that are learned over time, whereas in Treharne and Sox (2002) the demand distribution within each Markov state is assumed known and fixed. With fixed components, the ability to fit a "true" demand distribution outside the set of component distributions is limited, and adding large numbers of components presents computational and estimation challenges. We believe that our model brings distinct advantages in flexibility and parsimony.

Our work is also related to a large stream of the statistics literature on change-point detection — detecting departures of a stochastic process from a known model by monitoring observations drawn from the process over time. We refer readers to Basseville and Nikiforov (1993) and Lai (1995) for reviews. Our paper differs in two major ways. First, in the change-point detection problem the pre-change distribution of the process is often known, whereas in our problem the underlying demand distribution before the potential change-point is also unknown. Second, and more importantly, this literature does not consider operational controls and costs. There is relatively little operations management literature incorporating change-point detection. A recent example is Besbes and Zeevi (2011), in which a decision-maker seeks to detect and exploit a potential change in customer's willingness-to-pay distribution through dynamic pricing.

3. Model and Analysis

In this section we model an inventory management problem over a finite horizon following a potential change in the demand process. We present several structural properties, including certain structure inherited from well-studied inventory problems.

3.1. Inventory Management Following a Single Potential Change-point

Consider a single-item, T-period inventory system. At the beginning of period t, the decision maker (DM) observes the inventory position, x_t, and can place an order to bring the stock level up to $y_t \geq x_t$ at a linear purchasing cost $c \geq 0$. We assume zero lead time such that the order is instantaneously delivered. Demand, denoted by a random variable D_t with realized value d_t, is then realized and satisfied by the inventory on hand. If at the end of the period the DM still has leftover inventory, i.e., $y_t - d_t > 0$, a linear holding cost h is charged; otherwise (i.e., $y_t - d_t \leq 0$), the excess demand is fully backlogged and incurs a linear shortage cost p. The discount factor is $\alpha \in (0,1]$ each period. We assume $p > c(1-\alpha)$ to avoid trivial solutions. The salvage value for leftover inventory at the end of period T is assumed to be zero. We shall omit the subscript t whenever it is clear from the context.

As does Scarf (1959), we assume that past demands are fully observable to the DM and are not censored. This (and our assumption of inventory backlogging) are driven in part by analytical tractability, but we believe it is justified in contexts like the ones described in Section 1 in which changes in demand are likely to impact a whole department, firm, or industry at the same time. In these cases, where aggregate data across stock-keeping units can be utilized for demand estimation, demand censoring may be less relevant.

We extend the Bayesian framework of Scarf (1959). The distinctive feature of our model is how we model the demand process. We assume that the demands D_t's are independently drawn according to a density function $f(\cdot|\theta)$, where $\theta \in \Theta$ is an unknown parameter. The timing of the potential change-point is *known* to the DM and is in the past. Without loss of generality, label the period of the potential change-point as period 1. We assume that the DM has a *historical prior* π^h (h is for "history") on θ which reflects his prior knowledge on the demand parameter based on historical information. In addition, the DM believes *a priori* that with probability γ, θ will be drawn from a prior distribution π^c (c is for "change"). We call γ the *change probability* and π^c the *change prior*. The change probability γ captures the DM's prior belief on the probability that there has been a change in the demand process, and the change prior π^c summarizes the DM's belief on the demand parameter conditional on a change occurring. Let π_t denote the DM's prior belief on the unknown parameter θ at the beginning of period t, then $\pi_1(\theta) = (1-\gamma)\pi^h(\theta) + \gamma\pi^c(\theta)$ by definition, and π_{t+1} is the posterior distribution obtained by updating π_t based on d_t, the demand realization in period t, using Bayes rule. That is,

$$\pi_{t+1}(\theta|\pi_t, D_t = d_t) = \frac{f(d_t|\theta)\pi_t(\theta)}{\int_\Theta f(d_t|\omega)\pi_t(\omega)\mathrm{d}\omega}. \tag{1}$$

The predictive demand density in period t given belief π_t is defined by $\phi(\xi|\pi_t) = \int_\Theta f(\xi|\theta)\pi_t(\theta)\mathrm{d}\theta$. A natural generalization of our model allows for multiple change priors. Most of our results directly extend to this case. (The main exception is Proposition 4 in Section 3.2, which requires further clarification on how priors are ordered and how to handle multi-dimensional change probabilities.)

The DM's objective is to minimize the Bayesian expected discounted total cost over a finite horizon based on his prior belief on the demand process by choosing an order quantity in each period. We use $(a)^+$ to denote $\max\{a, 0\}$ for a real number a. Given an inventory level y after ordering and a demand realization d, the holding and shortage cost in a single period is

$$l(y,d) = h(y-d)^+ + p(d-y)^+,$$

and the expected cost in period t with initial inventory level x and belief π_t is given by

$$\mathbb{E}_{D_t|\pi_t}\left[c(y-x)+l(y,D_t)\right]=c(y-x)+L(y|\pi_t),$$

where $L(y|\pi_t):=\mathbb{E}_{D_t|\pi_t}\left[l(y,D_t)\right]=\int_0^\infty l(y,\xi)\phi(\xi|\pi_t)\mathrm{d}\xi$.

Let $C_t(x|\pi_t)$ be the optimal expected cost for periods $t,t+1,\dots,T$. We can formulate the problem as a Bayesian dynamic program with the following optimality equations for $t=1,\dots,T$:

$$C_t(x|\pi_t)=\min_{y\geq x}\{c(y-x)+L(y|\pi_t)+\alpha\mathbb{E}_{D_t|\pi_t}\left[C_{t+1}(y-D_t|\pi_t\circ D_t)\right]\}, \tag{2}$$

where $\pi_t\circ D_t:=\pi_{t+1}(\cdot|\pi_t,D_t)$ as defined by (1). The terminal cost is given by $C_{T+1}(\cdot|\cdot)=0$.

Although the demand process is complicated by the potential change-points, it is still independent of the ordering decisions. Because of this, the cost functions are convex and a state-dependent base-stock policy is optimal. We state the following result for completeness, but we omit the proof because the result can be obtained by a straightforward modification of proofs in Scarf (1959) and Treharne and Sox (2002).

Proposition 1. *(a) $C_t(x|\pi_t)$ is a convex function of x for all π_t.*

(b) The optimal policy takes the form of a state-dependent base-stock policy. There exists a sequence of nonnegative functions $\{y_t^*(\pi_t)\}$ such that it is optimal for the DM to order $\min\{y_t^*(\pi_t)-x_t,0\}$ *at the beginning of period t given inventory level x_t and belief π_t.*

We do not have closed-form expressions for the optimal policy, and given previous research it is unlikely that the optimal policy can be easily computed, much less simply expressed. We discuss the computability of the optimal policy in Section 4. However, as is often possible in finite horizon non-stationary inventory problems (e.g., see Morton and Pentico 1995), we are able to bound the optimal base-stock levels by easily computed myopic base-stock levels, which has the potential to reduce the search space for an optimal policy. The myopic policy is one in which the DM considers neither the evolution of future demand forecasts nor the carry-over of inventory between periods. The DM therefore treats each period as a single-period newsvendor problem. In our case, let $\Phi(\cdot|\pi_t)$ be the cumulative distribution function representing the DM's prediction of period t demand given belief π_t, i.e., $\Phi(d_t|\pi_t)=\int_0^{d_t}\phi(\xi|\pi_t)\mathrm{d}\xi$. Then, the base-stock level for period t under a myopic policy is given by $y_t^M(\pi_t)$ such that

$$\Phi(y_t^M(\pi_t)|\pi_t)=\begin{cases}\frac{p-c(1-\alpha)}{p+h} & ,t=1,\dots,T-1,\\ \frac{p-c}{p+h} & ,t=T.\end{cases} \tag{3}$$

The following proposition shows that this myopic policy upper-bounds the optimal policy. Proofs appear in an appendix unless otherwise indicated.

Proposition 2. *For all* $t = 1, 2, \ldots, T$, $y_t^M(\pi_t) \geq y_t^*(\pi_t)$.

We remark that both Proposition 1 and 2 extend to models with multiple potential change-points in both the past and future, as long as the timing of the potential change-points and their associated change priors and change probabilities are all known to the DM.

3.2. Monotonicity Properties of Optimal Base-Stock Levels

We explore in this subsection some monotonicity properties of the optimal base-stock levels with respect to demand history, the historical and change priors, and the change probability. Some definitions are needed here before we proceed.

Likelihood Ratio Order. Let $f(\cdot)$ and $g(\cdot)$ be two probability density functions. f is larger than g *in the likelihood ratio order*, denoted by $f \geq_{lr} g$, if for all $x_1 > x_2$, $f(x_1)/g(x_1) \geq f(x_2)/g(x_2)$.

Monotone Likelihood Ratio Property (MLRP). A distribution family $f(\cdot|\theta)$ with a parameter $\theta \in \Theta$ is said to have the *Monotone Likelihood Ratio Property (MLRP)* if $f(\cdot|\theta_1) \geq_{lr} f(\cdot|\theta_2)$ for all $\theta_1 \geq \theta_2$. Many common distributions, such as normal with known variance, binomial, Poisson, gamma, and Weibull, have MLRP (see Karlin and Rubin 1956).

Hereafter we assume that the demands are independent and from a distribution family $f(\cdot|\theta)$ with parameters $\theta \in \Theta$, and that $f(\cdot|\theta)$ has MLRP. The underlying implication of the MLRP assumption is that if a larger demand occurs, it becomes more likely that the underlying demand distribution $f(\cdot|\theta)$ has a higher θ parameter.

Scarf (1959) shows a monotonicity result in his setting with respect to the observed demand history. Specifically, the optimal base-stock level is increasing in the demand observation if the underlying demand process is stationary and the demand distribution $f(\cdot|\theta)$ is from the exponential family of the form $f(\xi|\theta) = \beta(\theta)e^{-\theta\xi}r(\xi)$ (with $r(\xi) = 0$ for $\xi < 0$). We can view our single change-point model as a variant of Scarf's model with MLRP demand and an initial prior being a mixture of distributions. The following proposition shows that we inherit Scarf's monotonicity result by generalizing his result to the case of MLRP demand.

Proposition 3. *Let* $y_t^*(\pi_t)$ *be the optimal base-stock level in period* t $(t = 1, \ldots, T)$ *given belief* π_t, *where* π_t $(t \geq 2)$ *is updated over* π_{t-1} *based on demand realization* d_{t-1}. *If the demand distribution family* $f(\cdot|\theta)$ *has MLRP, then the following hold:*
(a) $y_t^*(\pi_t) \leq y_t^*(\pi_t')$ *for* $\pi_t \leq_{lr} \pi_t'$;
(b) $y_t^*(\pi_t)$ *is increasing in* d_τ, *for all* $t \geq 2$, $\tau < t$.

Proposition 3(a) characterizes the behavior of the optimal base-stock level with respect to the DM's belief on the demand process. Intuitively, a larger (smaller) belief (in the likelihood ratio ordering) indicates a larger (smaller) demand parameter, which further implies a stochastically higher (lower) demand, which finally leads to a higher (lower) optimal base-stock level. Proposition 3(a) paves the way for establishing monotonicity properties of the optimal base-stock levels with respect to π^c, π^h, and γ in what follows. We use a closely related result when deriving the independentized lower bound in Section 4.1.2. Proposition 3(b) guarantees that it is always optimal to order more (less) in the next period if a higher (lower) demand is observed during the previous periods. We note that these results do not hinge on any specific assumptions on the initial belief π_1; it need not have a mixture form and can be any general distribution over the parameter space Θ.

As mentioned previously, our model is distinguished by its particular prior structure. The prior is a mixture of two distinct distributions. The following lemma establishes that this structure survives the DM's belief updating procedure.

LEMMA 1. *In the single potential change-point problem, let $\mathbf{d}_t = (d_1, \ldots, d_t)$ be any demand history up to period t, $t = 1, \ldots, T$. $\pi_t(\cdot|\mathbf{d}_{t-1})$ is then given by*

$$\pi_t(\theta|\mathbf{d}_{t-1}) = (1 - \gamma_t(\mathbf{d}_{t-1}))\pi_t^h(\theta|\mathbf{d}_{t-1}) + \gamma_t(\mathbf{d}_{t-1})\pi_t^c(\theta|\mathbf{d}_{t-1}),$$

where $\pi_t^h(\cdot|\mathbf{d}_{t-1})$ is updated over π^h based on $\mathbf{d}_{t-1}$, $\pi_t^c(\cdot|\mathbf{d}_{t-1})$ is updated over π^c based on $\mathbf{d}_{t-1}$, and $\gamma_t(\cdot)$ is a function of $\mathbf{d}_{t-1}$.

Lemma 1 shows that the belief updating procedure can be decomposed into two parts: one separately updates the beliefs conditioned on there being a change and on there being no change; the other updates the change probability. The belief is still in the form of a linear mixture distribution of those two updated beliefs, with the updated change probability as the weight. The following Proposition 4 uses this result to establish a relationship between the corresponding optimal base-stock levels. In Section 4 we will use the structure in Lemma 1 to derive an easily computed cost lower bound.

PROPOSITION 4. *In the single potential change-point problem, let $y_t^*(\pi_t)$ be the optimal base-stock level in period t $(t = 1, \ldots, T)$. Let $y_t^h(\pi_t^h)$ $(y_t^c(\pi_t^c))$ be the corresponding optimal base-stock level when the change probability $\gamma = 0$ (respectively, $\gamma = 1$). The following hold:*
(a) If $\pi^h \leq_{lr} \pi^c$, $y_t^h(\pi_t^h) \leq y_t^*(\pi_t) \leq y_t^c(\pi_t^c)$; otherwise if $\pi^c \leq_{lr} \pi^h$, $y_t^c(\pi_t^c) \leq y_t^*(\pi_t) \leq y_t^h(\pi_t^h)$;
(b) If $\pi^h \leq_{lr} \pi^c$, $y_t^*(\pi_t)$ is increasing in γ; otherwise if $\pi^c \leq_{lr} \pi^h$, $y_t^*(\pi_t)$ is decreasing in γ.

Proof. We only show the proofs of the first parts of (a) and (b). The proofs of the second parts follow from a straightforward modification.

It is easy to verify that if $\pi^h \leq_{lr} \pi^c$ and $\pi_1(\theta) = (1-\gamma)\pi^h(\theta) + \gamma\pi^c(\theta)$ for some $\gamma \in [0,1]$, then $\pi^h \leq_{lr} \pi_1 \leq_{lr} \pi^c$. Lemma 2(c) of Chen (2010) further guarantees that $\pi_t^h \leq_{lr} \pi_t \leq_{lr} \pi_t^c$ for all t. The first part of (a) follows directly from this result and from Proposition 3(a). Now define γ' such that $\gamma < \gamma' \leq 1$, and let $\pi_1'(\theta) = (1-\gamma')\pi^h(\theta) + \gamma'\pi^c(\theta)$. Because we can write π_1' as a convex combination of π^h and π^c, it follows that $\pi_1 \leq_{lr} \pi_1'$, and thus $\pi_t \leq_{lr} \pi_t'$ for all t. The desired result $y_t^*(\pi_t) \leq y_t^*(\pi_t')$ then follows from Proposition 3(a). $\square$

Proposition 4 provides sufficient conditions for the optimal base-stock levels of the single potential change-point problem to be bounded by those of the two degenerate problems — one with $\gamma = 0$ and the other with $\gamma = 1$. The result is intuitive: if an increase in demand is possible, the DM should order more than if the demand remains stable, and less than if the demand is guaranteed to increase. Moreover, the DM should order more as the change probability increases.

Proposition 4 may reduce the search space for optimal policies. It also motivates simple and computable heuristic ordering policies. In particular, for certain choices of π^h and π^c, the optimal solutions to the two degenerate problems can easily be computed by applying the dimensionality reduction technique in Scarf (1960) and Azoury (1985). A base-stock level in the form of a convex combination of these two solutions is an appealing heuristic policy. We have found such a policy to perform reasonably well, though we do not pursue it in the following section because it is outperformed by a related policy, which is greedy with respect to a convex combination of cost-to-go functions for the two degenerate problems.

4. Bounds and Policies

The usual approach to evaluate the performance of an inventory policy is to compare its expected cost with that of the optimal policy. However, the complexity of the Bayesian inventory control problem with potential change-points makes it intractable to compute optimal solutions. The dimensionality reduction technique in Scarf (1959) and Azoury (1985) is in general not applicable for our model with potential change-points. The conditions for applying the technique are:

1. Suppose that S_t is a sufficient statistic for demand observations up to period t. There is a function $q_t(S_t)$ such that $\phi(\xi|S_t) = (1/q_t(S_t))\psi_t(\xi/q_t(S_t))$, where $\psi_t(\cdot)$ is a probability density function that depends only on t;
2. The function $q_t(S_t)$ satisfies $q_{t+1}(S_t \circ d) = q_t(S_t)U_{t+1}(d/q_t(S_t))$ for some continuous real valued function U_{t+1} such that $\int_0^\infty U_{t+1}(u)\psi_t(u)\mathrm{d}u < \infty$, where $S_t \circ d$ denotes an update of S_t based on demand observation d.

However, since the beliefs in our problem are linear mixtures of distributions, there do not exist q_t functions that can serve as such scale parameters for the predictive demand distributions.

Therefore, it is computationally impractical to obtain the optimal policy or the optimal expected cost.

Treharne and Sox (2002) face a similar issue with an adaptive inventory control problem with similarities to our own. They point out the difficulty of computing an optimal policy even with an understanding of the policy structure, and they turn to heuristic policies. As an alternative approach, we develop lower bounds for the expected cost. Coupled with ordering heuristics derived from these bounds, we seek to bound the optimal cost as tightly as possible.

4.1. Bounds for Expected Cost

We develop two forms of lower bounds in this subsection that can be seen as applications of the "information relaxation" framework outlined by Brown et al. (2010). The first makes use of the decomposition of Lemma 1, while the second makes use of a novel relaxation we call the "independentized" problem.

4.1.1. The Mixture Lower Bound. Lemma 1 implies that the DM's belief in a period can be decomposed as a convex combination of the beliefs implied by two "degenerate" information structures in which a change is known to have occurred or known not to have occurred. If the degenerate problems are easily solved (e.g., if the historical prior π^h and change prior π^c satisfy the conditions of Azoury 1985), then the solutions can be easily employed to form an expected cost lower bound. Imagine an oracle who reveals to the DM whether or not a change has occurred. It is intuitive that the expected cost utilizing the oracle information would lower bound the true expected cost. This is the content of the following proposition.

PROPOSITION 5. *Let $\mathbf{d}_{t-1}$, $\pi_t(\cdot|\mathbf{d}_{t-1})$, $\pi_t^h(\cdot|\mathbf{d}_{t-1})$, $\pi_t^c(\cdot|\mathbf{d}_{t-1})$ and $\gamma_t(\mathbf{d}_{t-1})$ be defined as in Lemma 1. For all $t=1,\ldots,T$, define the mixture lower bound $LB_t^M(x_t|\mathbf{d}_{t-1})$ by*

$$LB_t^M(x_t|\pi_t(\cdot|\mathbf{d}_{t-1})) = (1-\gamma_t(\mathbf{d}_{t-1}))C_t(x|\pi_t^h(\cdot|\mathbf{d}_{t-1})) + \gamma_t(\mathbf{d}_{t-1})C_t(x|\pi_t^c(\cdot|\mathbf{d}_{t-1})),$$

then $LB_t^M(x_t|\pi_t(\cdot|\mathbf{d}_{t-1})) \le C_t(x_t|\pi_t(\cdot|\mathbf{d}_{t-1}))$.

Proof. The intuition behind the result is given above. The oracle information can be viewed as an information relaxation. Therefore, the proposition follows from Lemma 2.1 in Brown et al. (2010). □

4.1.2. The Independentized Lower Bound. When implementing an inventory policy with demand learning, the DM uses demand realizations in two ways: to calculate inventory positions and to update demand beliefs. We construct a lower bound using the notion of information relaxations (Brown et al. 2010) by relaxing only the information available to the DM for belief updating.

To motivate this, write as $\mathbf{D}_t = (\hat{D}_t, D_t)$ the DM's observation of demand in period t, where we artificially distinguish between the physical demand $\hat{D}_t$ that impacts inventory levels and the demand signal D_t that the DM uses to update his beliefs around θ. In the original problem, the physical demand and demand signal are one and the same and are therefore perfectly correlated. We write $\mathbf{D}_t^o$ for the original problem as $\mathbf{D}_t^o = (D_t, D_t)$. For the purpose of constructing a bound, we consider an "independentized" problem in which the physical demand and the demand signal are assumed to be independent of each other. We write $\mathbf{D}_t^\perp = (D_t^\perp, D_t)$ where both $D_t^\perp$ and D_t have a marginal density $\phi(\cdot|\pi_t)$, which is the predictive demand density implied by the belief π_t, but $D_t^\perp$ and D_t are independent, conditional on π_t.

Let $C_t(x_t|\pi_t)$ and $C_t^\perp(x_t|\pi_t)$ be the optimal expected costs of the original and the independentized problems, respectively, for periods $t, \ldots, T$ given initial inventory level x_t and belief π_t. Then we have

$$
\begin{aligned}
C_t(x_t|\pi_t) &= \min_{y \geq x_t} \left\{ c(y - x_t) + L(y|\pi_t) + \alpha \mathbb{E}_{\mathbf{D}_t^o = (D_t, D_t)|\pi_t} \left[C_{t+1}(y - D_t|\pi_t \circ D_t) \right] \right\}, \\
C_t^\perp(x_t|\pi_t) &= \min_{y \geq x_t} \left\{ c(y - x_t) + L(y|\pi_t) + \alpha \mathbb{E}_{\mathbf{D}_t^\perp = (D_t^\perp, D_t)|\pi_t} \left[C_{t+1}^\perp(y - D_t^\perp|\pi_t \circ D_t) \right] \right\},
\end{aligned}
$$

with terminal values $C_{T+1}(\cdot|\cdot) = C_{T+1}^\perp(\cdot|\cdot) = 0$.

With the notation above, we have the following proposition which shows that the optimal expected cost of the independentized problem serves as a lower bound for that of the original problem.

Proposition 6. *$C_t^\perp(x_t|\pi_t) \leq C_t(x_t|\pi_t)$ for all x_t, π_t, and $t = 1, \ldots, T$.*

The proof, in Appendix D, shows that the cost-to-go function, as a function of both the physical demand realization $d_t^\perp$ and demand signal realization d_t, is supermodular and that $\mathbf{D}_t^\perp = (D_t^\perp, D_t)$ is less than $\mathbf{D}_t^o = (D_t, D_t)$ in the supermodular ordering. High-level intuition is as follows. In the original problem, a small demand observation hurts the DM because it yields low revenues in the current period, but also because it implies a high end-of-period inventory position at the same time that demand forecasts are lowered. This combination of high inventory position and low demand forecast accentuates the possibility of inventory overage in the original problem. In the independentized problem, the correlation between high inventory positions and lowered demand forecasts is removed. In particular, high inventory positions and low demand forecasts are less likely to occur together.

Unfortunately, the independentized problem is not necessarily easier to solve than the original problem. To cope with this, we use the information relaxation approach proposed by Brown et al.

(2010) to construct a lower bound for the expected cost of the independentized problem. The basic idea is the following. At each decision point t we assume that an oracle reveals the entire future path of demand signals $(d_t, \ldots, d_T)$ to the DM. With this extra information and his current belief π_t, the DM is able to compute his future beliefs $\tilde{\pi}_{t+1}, \ldots, \tilde{\pi}_T$ recursively through

$$\tilde{\pi}_t = \pi_t \qquad \text{and} \qquad \tilde{\pi}_{u+1} = \tilde{\pi}_u \circ d_u, \quad \forall u = t, \ldots, T.$$

Let $\tilde{C}_t^{\perp}(x_t|\pi_t; (d_t, \ldots, d_T))$ be the optimal expected cost-to-go at period t given inventory level x_t, belief π_t and future demand signals $(d_t, \ldots, d_T)$. The independentized problem after relaxing future demand signals reduces to

$$\begin{aligned}\tilde{C}_t^{\perp}(x_t|\pi_t; (d_t, \ldots, d_T)) &= \tilde{C}_t^{\perp}(x_t|\tilde{\pi}_t, \ldots, \tilde{\pi}_T) \\ &= \min_{y \geq x_t} \left\{ c(y - x_t) + L(y|\tilde{\pi}_t) + \alpha \mathbb{E}_{D_t^{\perp}|\tilde{\pi}_t} \left[\tilde{C}_{t+1}^{\perp}(y - D_t^{\perp}|\tilde{\pi}_{t+1}, \ldots, \tilde{\pi}_T) \right] \right\},\end{aligned}$$

with $\tilde{C}_{T+1}^{\perp}(\cdot|\cdot) = 0$. This is in fact a stochastic inventory problem with *nonstationary*, *known* demand distributions, the solution to which can easily be obtained as the solution to a (fully observed) MDP with a one-dimensional state space. Because the oracle information is impermissible in the independentized problem, the optimal expected cost of the reduced problem will be lower than that of the independentized one.

We formally state the independentized lower bound as follows.

Proposition 7. *Let $(D_t, \ldots, D_T)$ denote the random demand signals in the independentized problem for periods $t, \ldots, T$. For all $t = 1, \ldots, T$, define the independentized lower bound $LB_t^I(x_t|\pi_t)$ by*

$$LB_t^I(x_t|\pi_t) = \mathbb{E}_{(D_t, \ldots, D_T)|\pi_t} \left[\tilde{C}_t^{\perp}(x_t|\pi_t; (D_t, \ldots, D_T)) \right],$$

then $LB_t^I(x_t|\pi_t) \leq C_t^{\perp}(x_t|\pi_t) \leq C_t(x_t|\pi_t)$.

Proof. The first inequality is an application of Lemma 2.1 in Brown et al. (2010). The second inequality follows from Proposition 6. □

We estimate $LB_t^I(x_t|\pi_t)$ in the numerical results using the following procedure. In an outer simulation, we randomly generate full demand signal paths $(d_1, \ldots, d_T)$ and calculate predictive demand distributions, $(\phi_1, \ldots, \phi_T)$, based on the generated demand signal paths. We then solve for each sequence of predictive demand distributions an inner optimization problem which is an inventory control problem with nonstationary, known demand distributions. These inner dynamic programming problems can be solved with straightforward backwards induction. The average of the resulting expected costs estimates the independentized lower bound.

To our knowledge, the "independentized" approach to bounding inventory problems with demand learning has not previously been used in the literature. An advantage of the approach over the mixture lower-bounding approach of Section 4.1.1 is that it requires efficient solutions only for inventory subproblems with known demand distributions, not for subproblems involving demand learning as required in Section 4.1.1. This widens its applicability. A drawback of the approach is that it is estimated via simulation. Due to estimation error, this means that technically we do not have a provable bound if it is based on a finite number of sample paths. In our numerical results, we estimate the bound based on a large number (100,000) of sample paths.

The approach may be useful for inventory problems involving demand learning beyond the one considered in this paper. It is clearly applicable for other generalizations of the Scarf (1959) model. Azoury (1985) shows that Scarf's model can be efficiently solved, but only for certain assumptions on the demand distribution. Without these assumptions, the optimal policy remains difficult to compute. In Section 5.2 we demonstrate that the independentized information relaxation is capable of meaningful bounds for the classic Scarf (1959) problem, for which we can generate the optimal costs for comparison.

4.1.3. Penalties. The information relaxation approach of Brown et al. (2010) also allows for the assignment of a penalty on each sample path, which potentially tightens the bound by penalizing the use of "impermissible" information in solving the inner problems. The lower bound for the optimal expected cost of the original problem is obtained by either simulation or analytical expression of the minimum expected value of the cost of the relaxed problem plus the penalty.

Unfortunately, we do not find computationally viable penalties for the two relaxations we have proposed. For the mixture lower bound, any natural penalty destroys the decomposition exploited by the information relaxation, and the inner problem becomes as difficult to solve as the original problem. For the independentized lower bound, limited-lookahead methods for computing penalties (as considered in Brown et al. (2010)) prove too time consuming to compute for the continuous prior and demand distributions we consider. As a result, in general we impose a zero penalty on our inner problems for computing the lower bounds. We leave further investigation of penalties for future work. Even with zero penalties, we see meaningfully tight bounds in our numerical results.

4.2. Heuristic Policies

We develop three heuristic policies for the single potential change-point problem: a myopic policy, a look-ahead policy based on the mixture lower bound, and a look-ahead policy based on the independentized lower bound. In Section 5 we evaluate these heuristics using the lower bounds in Section 4.1.

Myopic Policy. Each period the DM updates his belief based on the observed demand and then uses the single-period newsvendor solution as the base-stock level. This policy therefore forecasts demand using the potential change-point model but is not forward looking in its inventory optimization.

Look-Ahead Policy Based on Mixture Lower Bound (LA-M). This policy takes advantage of the mixture lower bound (LB^M) we have developed in the previous subsection. For each period t, the DM uses LB^M_{t+1} as an approximation for the optimal cost-to-go function in period $t+1$, $C_{t+1}(\cdot|\cdot)$, and solves the following problem:

$$C^M_t(x_t|\pi_t) = \min_{y \geq x_t} \left\{ c(y - x_t) + L(y|\pi_t) + \alpha \mathbb{E}_{D_t|\pi_t} \left[LB^M_{t+1}(y - D_t|\pi_t \circ D_t) \right] \right\}.$$

Of course, the LA-M policy is only implementable if the LB^M_{t+1} lower bound is simple to compute. Therefore, this policy is only attractive for instances in which the degenerate "change" (i.e., $\gamma = 1$) and "no change" (i.e., $\gamma = 0$) problems are easy to solve; e.g., when they conform to the assumptions of Scarf (1960) or Azoury (1985).

Look-Ahead Policy Based on Independentized Lower Bound (LA-I). This policy is very similar to the LA-M policy except that it uses the independentized lower bound LB^I_{t+1} instead of LB^M_{t+1} to approximate the optimal cost-to-go function for the next period. More specifically, in each period t the DM solves the following problem:

$$C^I_t(x_t|\pi_t) = \min_{y \geq x_t} \{ c(y - x_t) + L(y|\pi_t) + \alpha \mathbb{E}_{D_t|\pi_t} \left[LB^I_{t+1}(y - D_t|\pi_t \circ D_t) \right] \},$$

where $LB^I_{t+1}(\cdot|\cdot)$ is estimated using Monte-Carlo simulation as described in Section 4.1.2. This LA-I policy can be applied to the single change-point problem with any belief and demand distribution; however, the computation effort grows as more sample paths are used to estimate the LB^I lower bound.

5. Numerical Analysis

In this section we conduct numerical analyses to demonstrate the performance of the lower bounds and heuristics proposed in Section 4. Without loss of generality we normalize the purchasing cost c to zero and the unit holding cost h to one. We also assume no discounting ($\alpha = 1$) throughout the section.

We make use of the gamma-gamma conjugate pair as our model of demand in our numerical results. As we have pointed out in Section 4.1.1 and 4.2, the computability of the mixture lower bound and the LA-M policy requires a demand structure in which the degenerate problems are easy to solve. The gamma-gamma demand model meets this qualification and provides a unified framework to test our bounds and heuristics. We will first review the gamma-gamma demand model and its relevant properties in Section 5.1.

We will then test the independentized lower bound against Scarf (1960)'s Bayesian inventory problem with gamma-gamma demand in Section 5.2. Unlike the potential change-point problem, we are able to solve Scarf's problem optimally so as to understand the quality of the independentized lower bound.

Finally in Section 5.3, we will perform a comprehensive numerical study on bounds and heuristics for the potential change-point problem analyzed in Section 3 and 4.

5.1. The Gamma-Gamma Demand Model

The gamma-gamma demand model is a common one for the study of inventory management with demand learning (e.g., Azoury 1985, Scarf 1960, Chen 2010) because of its versatility and ease of updating. Assume that demand follows a gamma density with known shape parameter k and unknown scale parameter θ:

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

We assume an initial gamma prior with parameters (a, S) around the unknown scale parameter θ:

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

Given this information structure and demand observations $(d_1, \ldots, d_{t-1})$, it is well-known that sufficient statistics for Bayes updating are

$$a_t = a_{t-1} + k = a + k(t-1) \qquad \text{and} \qquad S_t = S_{t-1} + d_{t-1} = S + \sum_{i=1}^{t-1} d_i.$$

Furthermore, the updated distribution around θ at the beginning of period t is

$$\pi_t(\theta) = \pi(\theta|a_t, S_t) = \frac{S_t^{a_t} \theta^{a_t - 1} e^{-S_t\theta}}{\Gamma(a_t)},$$

and the predictive demand density can be written as

$$\phi(x|\pi_t) = \phi(x|a_t, S_t) = \frac{1}{S_t} \phi_t\left(\frac{x}{S_t}\right),$$

where $\phi_t(u) = \frac{\Gamma(a_t+k)}{\Gamma(a_t)\Gamma(k)} u^{k-1}(1+u)^{-(a_t+k)}$. A result of Scarf (1960), extended in Azoury (1985), is that the optimization (2) can be written as a one-dimensional dynamic program:

$$v_t(x) = \min_{y \geq x} \left\{ c(y-x) + L_t(y) + \alpha \int_0^\infty (1+u) v_{t+1}\left(\frac{y-u}{1+u}\right) \phi_t(u) \mathrm{d}u \right\}, t = 1, \ldots, T,$$

with $v_{T+1}(\cdot) = 0$. Let y_t^* denote its optimal solution for period t. Then we have

(i) $C_t(x|S_t) = S_t v_t(x/S_t)$,

(ii) $y_t^*(S_t) = S_t y_t^*$.

This property greatly simplifies calculation of our policies and bounds, in particular the mixture lower bound. Assuming that the change belief π_t^c for θ in period t is gamma with parameters (a_t^c, S_t^c) and that the no-change belief π_t^h is gamma with parameters (a_t^h, S_t^h), the mixture bound can be computed as

$$\begin{aligned} LB_t^M(x_t|\pi_t) = LB_t^M(x_t|\gamma_t, S_t^h, S_t^c) &= (1-\gamma_t)C_t(x|S_t^h) + \gamma_t C_t(x|S_t^c) \\ &= (1-\gamma_t)S_t^h v_t^h(x/S_t^h) + \gamma_t S_t^c v_t^c(x/S_t^c). \end{aligned}$$

5.2. Applying the Independentized Lower Bound to a Classical Problem

In this subsection we use the classic Bayesian inventory problem with gamma-gamma demand in Scarf (1960) to explore the behavior and quality of the independentized lower bound. Scarf's problem is closely related to the potential change-point problem in that it can be viewed as the degenerate problem for the mixture lower bound. In other words, it is a special case of the change-point problem with change probability equal to zero (or one).In addition, unlike the change-point problem, the optimal cost of the problem can easily be computed using the dimensionality reduction technique in Scarf (1960) and Azoury (1985). Therefore, it qualifies as a reasonable testbed for understanding the potential tightness of the independentized lower bound.

We use the gamma-gamma demand framework as described in Section 5.1, varying k such that $k = 1, 3$, and 5 ($k = 1$ indicates exponential demand). We consider two pairs of (a, S) parameters such that $(a, S) = (3, 10)$ and (6,20), indicating that the prior means (a/S) are the same at 0.3 and that the coefficients of variation ($\sqrt{1/a}$) of the priors are $1/\sqrt{3}$ and $1/\sqrt{6}$, respectively. We vary the unit shortage cost such that $p = 2, 4$, and 9, indicating critical fractiles (as defined in (3)) of 0.33, 0.80, and 0.90, respectively. To examine the effect of the length of the planning horizon T, we let $T = 5$ and 10. We follow a full factorial design, for $3 \times 2 \times 3 \times 2 = 36$ problem instances in total.

The results are shown in Table 1. Each lower bound LB^I is estimated by a Monte-Carlo simulation of 100,000 trials, and standard errors of the lower bound estimator are reported beside the LB^I estimates. The gap column shows the percentage deviation of LB^I from the optimal cost, which is computed by (Optimal Cost $-LB^I$)/Optimal Cost$\times$100%. Two of the instances have a negative gap due to estimation errors; in both instances the gap is not significantly different from zero at the 5% significance level.

We make two observations about these results. First, we are able to estimate the lower bounds precisely, resulting in standard errors no more than 0.5% of the optimal cost for each of the

Table 1 **The independentized lower bound for Scarf's Bayesian inventory control problem with gamma-gamma demand.**

Instance	k	p	a	S	T	Optimal Cost	LB^I	Std. Err.	Gap %
1	1	2	3	10	5	31.1420	31.0499	0.0500	0.30%
2					10	60.2223	60.0731	0.1237	0.25%
3			6	20	5	23.6196	23.6131	0.0151	0.03%
4					10	46.5628	46.5531	0.0406	0.02%
5		4	3	10	5	48.5762	48.2978	0.0803	0.57%
6					10	92.3798	92.1199	0.1877	0.28%
7			6	20	5	35.8211	35.7731	0.0226	0.13%
8					10	70.1277	70.0485	0.0600	0.11%
9		9	3	10	5	76.0456	75.3853	0.1215	0.87%
10					10	141.2961	140.1421	0.2852	0.82%
11			6	20	5	53.7740	53.6336	0.0339	0.26%
12					10	104.2665	104.0554	0.0898	0.20%
13	3	2	3	10	5	62.3703	61.9977	0.1209	0.60%
14					10	114.9831	114.2574	0.2630	0.63%
15			6	20	5	45.7556	45.7147	0.0396	0.09%
16					10	87.5032	87.5679	0.0956	-0.07%
17		4	3	10	5	92.9766	91.9282	0.1716	1.13%
18					10	167.1852	165.8111	0.3838	0.82%
19			6	20	5	66.0450	65.8697	0.0567	0.27%
20					10	124.7891	124.5628	0.1353	0.18%
21		9	3	10	5	139.7300	135.9056	0.2481	2.74%
22					10	242.7868	238.8543	0.5540	1.62%
23			6	20	5	94.4677	94.0393	0.0796	0.45%
24					10	175.7028	175.0220	0.1865	0.39%
25	5	2	3	10	5	85.8682	84.9443	0.1672	1.08%
26					10	153.9208	152.8968	0.5414	0.67%
27			6	20	5	61.9222	61.7811	0.0578	0.23%
28					10	116.0647	116.0949	0.1355	-0.03%
29		4	3	10	5	126.6605	123.7305	0.2335	2.31%
30					10	220.6594	217.9000	0.5122	1.25%
31			6	20	5	88.0404	87.6672	0.0810	0.42%
32					10	162.7306	162.4489	0.1872	0.17%
33		9	3	10	5	189.1255	181.6724	0.3620	3.94%
34					10	316.7586	309.0829	0.7130	2.42%
35			6	20	5	124.0675	122.9726	0.1118	0.88%
36					10	225.2149	224.5306	0.2572	0.30%

instances. Second, the independentized bounding method produces meaningful lower bounds for most of the instances. We find the average gap over the 36 instances to be 0.73% (negative gaps are truncated to zero), and smaller than 2% for 33 out of the 36 instances. We observe that the gap is relatively larger for larger critical ratios (i.e., for large p) and for large spread in the prior (i.e., small a). Based on its performance here, we include the independentized bound and the LA-I policy in our numerical study of the single potential change-point problem in the following subsection. We believe that this performance also shows the bound's potential for other versions of inventory management problems with demand learning for which optimal costs are difficult to obtain, although we do not study them further in this paper.

5.3. Bounds and Heuristics for the Change-Point Problem

In this subsection we numerically examine the performance of three heuristic policies – Myopic, LA-M and LA-I – for the single change-point problem introduced in Section 3 by comparing their expected costs with the lower bounds. In order to compute both the LB^M and LB^I lower bounds and their corresponding one-period look-ahead policies LA-M and LA-I, we assume a gamma-gamma conjugate demand structure. The demands are from a gamma distribution with parameters (k, θ). We only report the results for $k = 3$ here since we have observed results for $k = 1$ and 5 to be qualitatively similar. If the demand does not change at the beginning of the planning horizon, θ follows a gamma distribution with parameters (a^h, S^h); otherwise it follows a gamma distribution with parameters (a^c, S^c). We choose the shape parameters of the two prior distributions to be $a^h = 48$ and $a^c = 3$. We therefore have $a^h > a^c$, which implies that the DM is more uncertain about the demand distribution if the demand does change. This seems representative of practice, where the DM would have an accurate demand forecast based on an abundant demand history but would only have a coarse one following a potential demand shock. We fix $S^h = 160$ such that the no-change prior mean is $a^h/S^h = 48/160 = 0.3$. We vary S^c such that $S^c = 1, 5, 10, 15$ and 19, indicating extremely downward, downward, stationary, upward, and extremely upward potential changes in demand. We label the $S^c = 5, 10, 15$ cases as "moderate change" cases and the $S^c = 1$ and 19 cases as "extreme change" cases in which potential demand changes are quite large. We vary the initial change probability γ such that $\gamma = 0.2, 0.5$ and 0.8. The unit shortage cost p is set to be 4 and 9, indicating critical fractiles of 0.8 and 0.9, respectively. To examine the effect of the length of the planning horizon T, we let $T = 5$ and 10. Therefore, we have $5 \times 3 \times 2 \times 2 = 60$ instances in total in our full-factorial design.

For each instance, we compute the LB^M bound, estimate the LB^I using Monte-Carlo simulation with 100,000 demand signal paths, and estimate the expected costs of the Myopic, LA-M, and LA-I policies using simulations with an identical set of 10,000 demand paths. We also use the same set of demand paths to estimate the expected costs of two additional naïve policies – optimal policies as if $\gamma = 0$ (denoted by OPTNOCHG) and as if $\gamma = 1$ (denoted by OPTCHG) – as performance benchmarks. The OPTNOCHG policy would be adopted if the DM ignores the potential change-point and only uses the historical demand information for forecasting and inventory decisions. At the other extreme, the OPTCHG policy would be employed if the DM ditches all the historical demand information and starts fresh with a belief reflecting a change in demand.

Due to space limitations, we refer readers to the online supplement for detailed tables of results. We find the estimated independentized lower bound LB^I to be tighter than the mixture lower bound LB^M across all 60 instances, and therefore we use our estimated LB^I to evaluate optimality gaps of the various policies. We calculate the deviation of each policy's estimated expected costs from the estimated independentized lower bound LB^I (computed by $(\text{Cost} - LB^I)/LB^I \times 100\%$) for each instance. The percentage gaps, averaged over parameter levels, are summarized in Table 2 (for "moderate change" cases) and Table 3 (for "extreme change" cases).

We make several observations about the results in Table 2 and Table 3. First, for both moderate and extreme scenarios, myopic, LA-M and LA-I policies nearly always perform significantly better than the OPTCHG and OPTNOCHG policies. Intuitively, as change probability γ increases, the performance of OPTCHG gets better while that of OPTNOCHG gets worse. But even in their best instances (i.e., $\gamma = 0.2$ for OPTNOCHG and $\gamma = 0.8$ for OPTNOCHG), they yield larger gaps than the three heuristics. This highlights the danger of ignoring uncertainty around whether a demand change may or may not have happened.

Second, myopic, LA-M and LA-I policies have nearly the same performance under moderate scenarios, achieving average gaps of 1.16%, 1.15% and 1.16%, respectively. This suggests that the myopic policy may be an appealing choice except when extreme demand changes are possible, especially given its simplicity for implementation in practice. Other authors have found myopic policies to perform well in inventory contexts with demand learning (e.g., Lovejoy 1990). A managerial insight is that intelligent demand estimation may merit more attention than forward-looking optimization when a (moderate) demand shift may have recently occurred.

The myopic policy still performs reasonably well along with the two look-ahead heuristics when there has been a large potential increase in demand ($S^c = 19$). However, when there has been a potential extreme downward change in demand ($S^c = 1$), all the three heuristics exhibit larger gaps relative to the lower bound. The myopic policy yields an average gap of 15.74%. The LA-M and LA-I policy have much smaller average gaps (5.72% and 5.74%, respectively) than the myopic policy. This observation suggests that more sophisticated policies bring significant benefits over the myopic policy when relatively extreme changes are possible. The gaps discussed here reflect the deviation of the policies' expected costs only from the cost lower bound rather than the optimal cost. Therefore, these gaps are conservative in that they overestimate the optimality gaps.

Finally, although we have observed that LB^I is tighter than LB^M for all instances, the LA-M and LA-I policies (which approximate cost-to-go functions by LB^M and LB^I, respectively) have nearly the same performance under all scenarios. Since the LA-M policy requires less computing effort than the LA-I policy does, the LA-M policy is recommended over the LA-I policy when both are applicable. Recall, however, that the LA-M policy can only be efficiently computed for certain belief and demand models.

Table 2 Mean percentage gaps* for moderate change cases, averaged over parameter levels.

Parameter		Myopic	LA-M	LA-I	OPTCHG	OPTNOCHG
S^c	5	1.00%	0.99%	0.99%	5.79%	14.32%
	10	0.90%	0.97%	1.00%	3.41%	21.73%
	15	1.58%	1.50%	1.50%	7.67%	39.78%
p	4	0.74%	0.79%	0.81%	4.73%	19.83%
	9	1.58%	1.51%	1.51%	6.52%	30.72%
γ	0.2	0.83%	0.91%	0.90%	11.09%	11.89%
	0.5	1.53%	1.55%	1.58%	4.54%	26.49%
	0.8	1.12%	1.00%	1.00%	1.23%	37.45%
T	5	1.50%	1.50%	1.51%	6.95%	25.39%
	10	0.82%	0.81%	0.81%	4.29%	25.16%
Overall		1.16%	1.15%	1.16%	5.62%	25.27%

*Negative gaps are truncated to zero before averaging.

Table 3 Mean percentage gaps* for extreme change cases, averaged over parameter levels.

Parameter		Myopic	LA-M	LA-I	OPTCHG	OPTNOCHG
S^c	1	15.74%	5.72%	5.74%	35.06%	73.97%
	19	1.73%	1.56%	1.55%	12.09%	53.38%
p	4	6.79%	2.96%	2.97%	18.55%	61.74%
	9	10.68%	4.32%	4.32%	28.60%	65.60%
γ	0.2	5.16%	3.97%	3.96%	43.12%	21.92%
	0.5	11.74%	4.94%	4.96%	20.89%	53.49%
	0.8	9.30%	2.01%	2.02%	6.72%	115.60%
T	5	8.96%	4.32%	4.32%	29.43%	60.73%
	10	8.51%	2.97%	2.97%	17.72%	66.62%
Overall		8.73%	3.64%	3.64%	23.58%	63.67%

*Negative gaps are truncated to zero before averaging.

6. Parameter Estimation and Sensitivity

The demand model of Section 3 requires the specification of three inputs: a "no-change" or historical prior π^h, a change prior π^c, and a change probability γ. The no-change prior π^h, the forecast of demand in the absence of a potential change-point, can be estimated using established techniques applied to historical demand, and we do not elaborate on it here. However, in many contexts it may be less obvious how to estimate the parameters π^c and γ.

Selecting the change prior π^c entails predicting the direction and magnitude of a potential change. This represents a new demand regime for the firm by definition, but in many cases it may be a regime with past precedents. Imagine a retailer facing the entrance of a new competitor at one of its locations. It is likely to have faced similar entrances in the past at other locations. When potential change-points are driven by changes in the state of the economy (e.g., the example of women's clothing following the 2008 financial crisis, discussed in our introduction), financial markets (or forecasts thereof) may offer signals that can be used to inform demand forecasts (Osadchiy et al. 2013). If neither of these two approaches is applicable, a firm might generate π^c by inflating the variance of π^h and/or inflating or deflating its mean by percentages determined by expert opinions.

The change probability γ is particularly challenging to estimate because it is arguably most situation-specific and least amenable to estimation from historical data. Fortunately, we have found that the performance of our policies is relatively insensitive to mis-specification of γ. Figure 2 plots results from a numerical study similar to Section 5.3 except that we allow for misspecification of the change probability γ. The manager employs the LA-M heuristic, but computes forecasts and stocking decisions using a γ parameter that may differ from the parameter used to simulate the underlying demand process.

In contrast to our earlier development, we take a profit perspective here because changing the "true" value of γ changes expected demand, making a comparison between costs meaningless. Specifically, we translate expected costs into expected profit in the natural way, defining expected single-period profit as $\mathbb{E}\left[p\min\{y,D\}-c(y-x)-h(y-D)^+\right] = -\mathbb{E}\left[c(y-x)+h(y-D)^+ + p(D-y)^+\right] + p\mathbb{E}\left[D\right]$. The results in Figure 2 assume $c=0$, $h=1$, $p=9$, $T=5$, historical prior π^h given by a Gamma(48,160), and change prior π^c given by either Gamma(3,5) (i.e., "moderate decrease"), or Gamma(3,15) (i.e., "moderate increase"), where the parameters have the same interpretations as in Section 5.3. We have found consistent results across a broader set of instances.

We make a few observations about the expected profits in Figure 2. The expected profits naturally vary with the true underlying demand process, and the profits are always highest for each instance when the assumed γ matches the "true" one used to generate the demand data. What is striking, however, is that for a given problem instance the heuristic is relatively insensitive to the choice of change probability used for computing the policy, as long as this probability is chosen not too close to the extremes 0 and 1. In particular, the policy assuming $\gamma = 0.5$ exhibits fairly robust performance across different underlying demand processes.

We also observe the least variation in profits across instances for policies that assume demand will be low. That is, when the change prior indicates a possible downward change, the "flattest" profits are obtained by the policy assuming γ equal to 1.0 or 0.8. When the change prior indicates

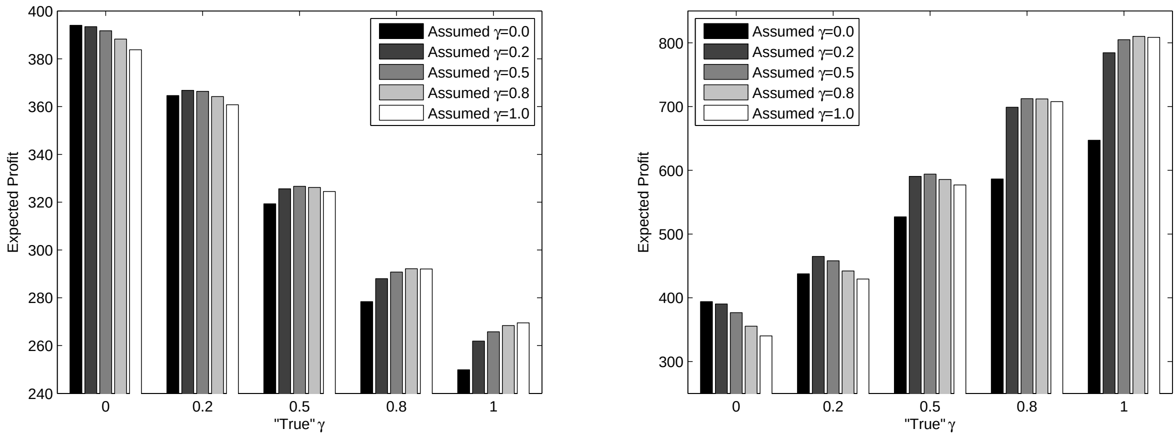


Figure 2 **Sensitivity to misspecification of the change probability γ when the change prior represents a moderate decrease (left) and moderate increase (right) in demand. The bars in each chart represent expected profits when the manager assumes $\gamma = 0.0$ (white), 0.2, 0.5, 0.8, and 1.0 (black).**

a possible upward change in demand, the flattest profits are obtained by the policy assuming γ equal to 0 or 0.2. This suggests that a conservative decision-maker worried about downside risk may wish to choose π^c and γ by erring on the side of underestimating demand.

Proposition 8 below formalizes this finding for a Bayesian repeated newsvendor setting in which there is no inventory carryover across periods. Consider a T-period Bayesian newsvendor problem with unit selling price r, unit purchasing cost $c < r$, and inventory that perishes at the end of each period with zero salvage value. As before, demands are i.i.d. with density $f(\cdot|\theta)$. Let $G(y|\pi) = \mathbb{E}_{D|\pi}[r\min\{y, D\}] - cy$ be the single-period expected profit given order quantity y and prior π. We denote by $\mathbf{y} = (y_1, \ldots, y_T)$ a non-anticipative inventory policy. In general, y_t may be a function of all the information that the DM has up to period t. Let $\mathbf{D}_t = (D_1, \ldots, D_t)$ be demand until period t. For any initial prior π, let $V_T(\pi)$ denote the optimal expected profit for this problem, i.e.,

$$V_T(\pi) = \max_{\mathbf{y}} \sum_{t=1}^{T} E[G(y_t|\pi \circ \mathbf{D}_{t-1})|\pi], \tag{4}$$

where $\pi \circ \mathbf{D}_{t-1}$ is the posterior updated based on demand history. Suppose that a conservative DM has a bounded set $\mathcal{P}$ that contains all candidate priors on θ, and there exists a "smallest" prior $\underline{\pi} \in \mathcal{P}$ such that $\underline{\pi} \leq_{lr} \pi$ for all $\pi \in \mathcal{P}$. The objective is to maximize the worst-case expected profit, which translates into a max-min version of problem (4):

$$R_T(\mathcal{P}) = \max_{\mathbf{y}} \min_{\pi \in \mathcal{P}} \sum_{t=1}^{T} E[G(y_t|\pi \circ \mathbf{D}_{t-1})|\pi].$$

Proposition 8. *Suppose that $f(\cdot|\theta)$ has MLRP. Then $R_T(\mathcal{P}) = V_T(\underline{\pi})$.*

The proposition says that the DM can obtain the optimal policy for the max-min problem by simply solving (4) for $\pi = \underline{\pi}$. We note that Proposition 8 is a fairly general statement about the choice of prior beliefs, and the intuition can be applied to the selection of π^c as well as γ.

To summarize this section, we have suggested a few ways for a manager to think about choosing the parameters π^c and γ. We show evidence that the results of our heuristics are relatively insensitive to the specification of the change probability, particularly if a change prior is chosen away from the extremes 0 and 1. We also find both analytically and numerically that a max-min formulation is solved by assuming the smallest change prior structure among a set of candidates. Therefore, a manager concerned about downside profit risk may choose to "play it safe" by erring on the side of underestimating demand.

7. Conclusions

While there is a rich and growing literature on inventory control with demand learning, the vast majority of this work assumes that demand follows a single, stationary demand model. We believe that our paper is novel in considering an adaptive inventory control problem following a potential change-point in demand. We formulate the problem as a Bayesian dynamic program and prove several fundamental structural results for the optimal state-dependent base-stock policy. Computing the optimal policy appears to be intractable, so we construct two novel lower bounds for the optimal expected cost, both based on the notion of information relaxations (Brown et al. 2010). One relies on decomposing the prior structure, whereas the other relies on relaxing the natural dependence between demand signals and inventory dynamics. The lower bounds are used both to generate heuristic policies and to evaluate their performance. We provide a discussion of estimation and sensitivity, including guidance on the selection of model parameters.

Our numerical study yields several insights on inventory management in uncertain demand environments. First, if a change in demand regime is suspected, managers can recover significant costs by accounting for this uncertainty. That is, a manager should remain wary of demand change-points. Second, a Bayesian myopic policy may be sufficiently good in many cases, suggesting that a manager may be justified in prioritizing demand estimation over forward-looking inventory optimization in these cases. Third, more sophisticated policies may be needed when extreme demand changes are possible. Fourth, a manager worried about profit downside may opt for lower demand estimates.

Several extensions of our model may be worth further exploration. First, we assume that demand is observable and backlogged; hence, the learning is passive in that the DM's observations are independent of ordering decisions. It would be interesting to examine the censored demand case. A conjecture is that the "stock more" result of Lariviere and Porteus (1999), Ding et al. (2002), and others may be accentuated in the presence of potential upward changes in demand. Second, most of our paper focuses on adaptive inventory control following a potential change-point in demand. Interesting questions also arise when a potential change-point is anticipated in the future (e.g., a planned marketing campaign). Third, we have assumed that the timing of potential change-points are known. Relaxing the model to allow for time uncertainty would also seem to be interesting and relevant to inventory management practice. Fourth, we believe the independentized bound we have developed may merit further investigation for other inventory models involving demand learning.

Acknowledgments
We thank David Brown for his insightful comments on an earlier version of the paper.

Appendix

A. Proof of Proposition 2

Proof. $y_t^*(\pi_t)$ is the solution to

$$H_t(x|\pi_t) = -p + (h+p)\Phi(x|\pi_t) + \alpha \mathbb{E}_{D_t|\pi_t}\left[\frac{\partial C_{t+1}}{\partial x}(x - D_t|\pi_t \circ D_t)\right] = -c.$$

For $t = T$, $\frac{\partial C_{T+1}}{\partial x}(\cdot|\cdot) = 0$, thus $H_T(y_T^*(\pi_T)|\pi_T) = -p + (h+p)\Phi(y_T^*(\pi_T)|\pi_T) = -c$, or $\Phi(y_T^*(\pi_T)|\pi_T) = \frac{p-c}{p+h} = \Phi(y_T^M|\pi_T)$, namely, $y_T^M(\pi_T) = y_T^*(\pi_T)$. One can show that due to the convexity of $C_{t+1}(\cdot|\cdot)$,

$$\mathbb{E}_{D_t|\pi_t}\left[\frac{\partial C_{t+1}}{\partial x}(x - D_t|\pi_t \circ D_t)\right] \geq -c,$$

therefore for $t = 1, \ldots, T-1$,

$$\begin{aligned} H_t(y_t^*(\pi_t)|\pi_t) &= -p + (h+p)\Phi(y_t^*(\pi_t)|\pi_t) + \alpha \mathbb{E}_{D_t|\pi_t}\left[\frac{\partial C_{t+1}}{\partial x}(x - D_t|\pi_t \circ D_t)\right] \\ &\geq -p + (h+p)\Phi(y_t^*(\pi_t)|\pi_t) - \alpha c. \end{aligned}$$

Since $y_t^*(\pi_t)$ satisfies $H_t(y_t^*(\pi_t)|\pi_t) = -c$, we have

$$\Phi(y_t^*(\pi_t)|\pi_t) \leq \frac{p - (1-\alpha)c}{p+h} = \Phi(y_t^M(\pi_t)|\pi_t),$$

namely, $y_t^*(\pi_t) \leq y_t^M(\pi_t)$. □

B. Proof of Proposition 3

Proof. We postpone the proof of part (a) until Appendix D, where Lemma 4 includes this result as a special case. Alternatively, part (a) can be proved directly with an extension of Theorem 2 in Scarf (1959) to all demand distribution families that have MLRP.

To prove part (b), we note that for any $d_\tau < d'_\tau$, Lemma 2 of Chen (2010) establishes that $\pi_t \leq_{lr} \pi'_t$. The result then follows from (a). □

C. Proof of Lemma 1

Proof. The proof is by induction. The lemma is true for $t=1$. Suppose it is true for some $t\geq 1$; that is

$$\pi_t(\cdot|\mathbf{d}_{t-1})=(1-\gamma_t(\mathbf{d}_{t-1}))\pi_t^h(\cdot|\mathbf{d}_{t-1})+\gamma_t(\mathbf{d}_{t-1})\pi_t^c(\cdot|\mathbf{d}_{t-1}).$$

Using Bayes rule, for $i\in\{h,c\}$, we have

$$\pi_{t+1}^i(\theta|\mathbf{d}_t)=\frac{\pi_t^i(\theta|\mathbf{d}_{t-1})f(d_t|\theta)}{\int_\Theta \pi_t^i(\omega|\mathbf{d}_{t-1})f(d_t|\omega)\mathrm{d}\omega},$$

and

$$\pi_{t+1}(\theta|\mathbf{d}_t)=\frac{[(1-\gamma_t(\mathbf{d}_{t-1}))\pi_t^h(\theta|\mathbf{d}_{t-1})+\gamma_t(\mathbf{d}_{t-1})\pi_t^c(\theta|\mathbf{d}_{t-1})]f(d_t|\theta)}{\int_\Theta[(1-\gamma_t(\mathbf{d}_{t-1}))\pi_t^h(\omega|\mathbf{d}_{t-1})+\gamma_t(\mathbf{d}_{t-1})\pi_t^c(\omega|\mathbf{d}_{t-1})]f(d_t|\theta)\mathrm{d}\omega}.$$

Write $I^i=\int_\Theta \pi_t^i(\omega|\mathbf{d}_{t-1})f(d_t|\omega)\mathrm{d}\omega$ for $i\in\{h,c\}$, then

$$\begin{aligned}\pi_{t+1}(\theta|\mathbf{d}_t)&=\frac{(1-\gamma_t(\mathbf{d}_{t-1}))\pi_t^h(\theta|\mathbf{d}_{t-1})f(d_t|\theta)}{(1-\gamma_t(\mathbf{d}_{t-1}))I^h+\gamma_t(\mathbf{d}_{t-1})I^c}+\frac{\gamma_t(\mathbf{d}_{t-1})\pi_t^c(\theta|\mathbf{d}_{t-1})f(d_t|\theta)}{(1-\gamma_t(\mathbf{d}_{t-1}))I^h+\gamma_t(\mathbf{d}_{t-1})I^c}\\&=\frac{(1-\gamma_t(\mathbf{d}_{t-1}))I^h}{(1-\gamma_t(\mathbf{d}_{t-1}))I^h+\gamma_t(\mathbf{d}_{t-1})I^c}\pi_{t+1}^h(\theta|\mathbf{d}_t)+\frac{\gamma_t(\mathbf{d}_{t-1})I^c}{(1-\gamma_t(\mathbf{d}_{t-1}))I^h+\gamma_t(\mathbf{d}_{t-1})I^c}\pi_{t+1}^c(\theta|\mathbf{d}_t).\end{aligned}$$

By defining

$$\gamma_{t+1}(\mathbf{d}_t)=\frac{\gamma_t(\mathbf{d}_{t-1})I^c}{(1-\gamma_t(\mathbf{d}_{t-1}))I^h+\gamma_t(\mathbf{d}_{t-1})I^c},$$

the lemma is true for $t+1$, which completes the induction. □

D. Proof of Proposition 6

For the purposes of this section, we consider an T-period generalized Bayesian inventory problem as described below. Let $\hat{C}_t(x|\pi)$ be the optimal expected cost for periods $t,\ldots,T$ given initial inventory level x and prior distribution π, where

$$\hat{C}_t(x|\pi)=\min_{y\geq x}\left\{c(y-x)+L(y|\pi)+\alpha\mathbb{E}_{\mathbf{D}_t=(\hat{D}_t,D_t)|\pi}\left[\hat{C}_{t+1}(y-\hat{D}_t|\pi\circ D_t)\right]\right\},$$

with terminal value $\hat{C}_{T+1}(\cdot|\cdot)=0$. We assume that $\hat{D}_t$ and D_t have the same marginal distribution induced by the prior π but their dependence is induced by some copula. We denote the minimizer of this expression by $\hat{y}_t^*(\pi)$. Note that the original and the independentized problems are both special cases of this formulation. In the original problem, $\hat{D}_t=D_t$, whereas in the independentized problem, $\hat{D}_t$ and D_t are independent with the same distribution induced by π.

The proof of Proposition 6 requires a few lemmas:

Lemma 2. *For all π and $t=1,\ldots,T+1$:*

(i) $\hat{C}_t(x|\pi)$ has a continuous derivative with respect to x, and is convex with respect to x;

(ii) The optimal policies are defined by single critical numbers $\hat{y}_t^*(\pi)\geq 0$;

(iii) $\hat{C}_t(x|\pi)$ has a continuous second derivative with respect to x at all points except perhaps $x=\hat{y}_t^*(\pi)$, at which point both the left and right hand second derivatives exist.

We omit the proof, as it is a minor modification of the one for Proposition 1.

Lemma 3. *Let $\mathbf{D}^i = (\hat{D}, D)|\pi^i$ be a random vector in which $\hat{D}$ and D have the same marginal predictive demand density $\phi(\cdot|\pi^i)$, for $i = 1, 2$, and suppose that $\mathbf{D}^1$ and $\mathbf{D}^2$ have a common copula. If $\pi^1 \leq_{lr} \pi^2$, then $\mathbf{D}^1 \leq_{st} \mathbf{D}^2$.*

Proof. Let $\hat{D}|\pi^i$ and $D|\pi^i$ denote random variables with density $\phi(\cdot|\pi^i)$ for $i = 1, 2$, then $\hat{D}|\pi^1 \leq_{st} \hat{D}|\pi^2$ and $D|\pi^1 \leq_{st} D|\pi^2$ (Lemma 2(d), Chen 2010). The lemma follows from Theorem 6.B.14 in Shaked and Shanthikumar (2007).

Lemma 4. *If $\pi^1 \leq_{lr} \pi^2$, the following hold for all x, π and $t = 1, \ldots, T+1$:*
(i) $\frac{\partial \hat{C}_t}{\partial x}(x|\pi^1) \geq \frac{\partial \hat{C}_t}{\partial x}(x|\pi^2)$;
(ii) $\hat{y}_t^*(\pi^1) \leq \hat{y}_t^*(\pi^2)$.

Proof. The proof is by induction. The lemma clearly holds when $t = T+1$ because $\hat{C}_{T+1}(\cdot|\cdot) = 0$. Suppose it is true for $t+1$. One can show that

$$\frac{\partial \hat{C}_t}{\partial x}(x|\pi) = \begin{cases} -c & , x < \hat{y}_t^*(\pi), \\ \hat{H}_t(x|\pi) & , x \geq \hat{y}_t^*(\pi), \end{cases}$$

where function $\hat{H}_t(\cdot|\pi)$ is defined by

$$\hat{H}_t(x|\pi) = -p + (h+p)\Phi(x|\pi) + \alpha \mathbb{E}_{(\hat{D}_t, D_t)|\pi}\left[\frac{\partial \hat{C}_{t+1}}{\partial x}(x - \hat{D}_t|\pi \circ D_t)\right].$$

For fixed $(\hat{d}_t, d_t)$, by the induction assumption, we have

$$\frac{\partial \hat{C}_{t+1}}{\partial x}(x - \hat{d}_t|\pi^1 \circ d_t) \geq \frac{\partial \hat{C}_{t+1}}{\partial x}(x - \hat{d}_t|\pi^2 \circ d_t), \tag{5}$$

since $\pi^1 \circ d_t \leq_{lr} \pi^2 \circ d_t$ (Lemma 2(c), Chen 2010). In addition, for $\hat{d}_t^1 \leq \hat{d}_t^2$, $d_t^1 \leq d_t^2$, we have

$$\frac{\partial \hat{C}_{t+1}}{\partial x}(x - \hat{d}_t^1|\pi^2 \circ d_t^1) \geq \frac{\partial \hat{C}_{t+1}}{\partial x}(x - \hat{d}_t^1|\pi^2 \circ d_t^2) \tag{6}$$

$$\geq \frac{\partial \hat{C}_{t+1}}{\partial x}(x - \hat{d}_t^2|\pi^2 \circ d_t^2), \tag{7}$$

where (6) follows from the induction assumption and that $\pi^2 \circ d_t^1 \leq_{lr} \pi^2 \circ d_t^2$ (Lemma 2(a), Chen 2010), and (7) from the convexity of $\hat{C}_{t+1}(\cdot|\pi^2 \circ d_t^2)$. Therefore, $\frac{\partial \hat{C}_{t+1}}{\partial x}(x - \hat{d}_t|\pi^2 \circ d_t)$ is decreasing in $(\hat{d}_t, d_t)$. In addition, $\pi^1 \leq_{lr} \pi^2$ together with Lemma 3 imply that

$$(\hat{D}_t, D_t)|\pi^1 \leq_{st} (\hat{D}_t, D_t)|\pi^2. \tag{8}$$

We thus have

$$\mathbb{E}_{(\hat{D}_t, D_t)|\pi^1}\left[\frac{\partial \hat{C}_{t+1}}{\partial x}(x - \hat{D}_t|\pi^1 \circ D_t)\right] \geq \mathbb{E}_{(\hat{D}_t, D_t)|\pi^1}\left[\frac{\partial \hat{C}_{t+1}}{\partial x}(x - \hat{D}_t|\pi^2 \circ D_t)\right] \tag{9}$$

$$\geq \mathbb{E}_{(\hat{D}_t, D_t)|\pi^2}\left[\frac{\partial \hat{C}_{t+1}}{\partial x}(x - \hat{D}_t|\pi^2 \circ D_t)\right], \tag{10}$$

where (9) results from (5), and (10) from (8) and the fact that $\frac{\partial \hat{C}_{t+1}}{\partial x}(x - \hat{d}_t|\pi^2 \circ d_t)$ is decreasing in $(\hat{d}_t, d_t)$ (Section 6.B.1, Shaked and Shanthikumar 2007). We conclude that $\hat{H}_t(x|\pi^1) \geq \hat{H}_t(x|\pi^2)$. Note that $\hat{y}_t^*(\pi)$ is the solution to the equation $H_t(x|\pi) = -c$. Also note that $H_t(x|\pi)$ is increasing in x. Hence,

$$\hat{H}_t(\hat{y}_t^*(\pi^1)|\pi^1) = -c = \hat{H}_t(\hat{y}_t^*(\pi^2)|\pi^2) \leq \hat{H}_t(\hat{y}_t^*(\pi^2)|\pi^1),$$

which indicates that $\hat{y}_t^*(\pi^1) \le \hat{y}_t^*(\pi^2)$.

It remains to show that $\frac{\partial \hat{C}_t}{\partial x}(x|\pi^1) \ge \frac{\partial \hat{C}_t}{\partial x}(x|\pi^2)$. Consider three cases:

(i) $x < \hat{y}_t^*(\pi^1)$. In this case, $\frac{\partial \hat{C}_t}{\partial x}(x|\pi^1) = \frac{\partial \hat{C}_t}{\partial x}(x|\pi^2) = -c$;

(ii) $\hat{y}_t^*(\pi^1) \le x < \hat{y}_t^*(\pi^2)$. In this case, $\frac{\partial \hat{C}_t}{\partial x}(x|\pi^1) = \hat{H}(x|\pi^1) \ge \hat{H}(\hat{y}_t^*(\pi^1)|\pi^1) = -c = \frac{\partial \hat{C}_t}{\partial x}(x|\pi^2)$;

(iii) $x > \hat{y}_t^*(\pi^2)$. In this case, $\frac{\partial \hat{C}_t}{\partial x}(x|\pi^1) = \hat{H}(x|\pi^1) \ge \hat{H}(x|\pi^2) = \frac{\partial \hat{C}_t}{\partial x}(x|\pi^2)$;

This completes the induction proof. □

With these lemmas established, we can now proceed to the proof of Proposition 6.

Proof of Proposition 6. The proof is by induction. The proposition is clearly true when $t = T+1$. Suppose for period $t+1$, $C_{t+1}^{\perp}(x|\pi) \le C_{t+1}(x|\pi)$ for all x, π.

Fix y. Consider function $K(d_t^{\perp}, d_t) = C_{t+1}^{\perp}(y - d_t^{\perp}|\pi \circ d_t)$. Taking the derivative with respect to $d_t^{\perp}$, we obtain

$$\frac{\partial K}{\partial d_t^{\perp}}(d_t^{\perp}, d_t) = -\frac{\partial C_{t+1}^{\perp}}{\partial (y - d_t^{\perp})}(y - d_t^{\perp}|\pi \circ d_t).$$

For $d_t^1 \le d_t^2$, Lemma 2 of Chen (2010) implies that $\pi \circ d_t^1 \le_{lr} \pi \circ d_t^2$. Lemma 4 therefore yields

$$\frac{\partial K}{\partial d_t^{\perp}}(d_t^{\perp}, d_t^1) \le \frac{\partial K}{\partial d_t^{\perp}}(d_t^{\perp}, d_t^2).$$

In other words, $K(\cdot,\cdot)$ has increasing differences in $(d_t^{\perp}, d_t)$. Thus, $K(\cdot,\cdot)$ is supermodular in $(d_t^{\perp}, d_t)$.

Let $F_t^o(\hat{d}_t, d_t)$ and $F_t^{\perp}(\hat{d}_t, d_t)$ be the distribution functions of the random vectors $\mathbf{D}_t^o = (D_t, D_t)|\pi$ and $\mathbf{D}_t^{\perp} = (D_t^{\perp}, D_t)|\pi$, respectively. Then we have

$$\begin{aligned}
F_t^{\perp}(\hat{d}_t, d_t) &= \mathbb{P}\left\{D_t^{\perp} \le \hat{d}_t, D_t \le d_t\right\} = \mathbb{P}\left\{D_t^{\perp} \le \hat{d}_t\right\}\mathbb{P}\left\{D_t \le d_t\right\} \\
&\le \min\left[\mathbb{P}\left\{D_t^{\perp} \le \hat{d}_t\right\}, \mathbb{P}\left\{D_t \le d_t\right\}\right] = \min\left[\mathbb{P}\left\{D_t \le \hat{d}_t\right\}, \mathbb{P}\left\{D_t \le d_t\right\}\right] \\
&= \mathbb{P}\left\{D_t \le \hat{d}, D_t \le d\right\} = F_o(\hat{d}, d).
\end{aligned}$$

Therefore, by (9.A.3) in Shaked and Shanthikumar (2007), $\mathbf{D}_t^o$ and $\mathbf{D}_t^{\perp}$ are ranked in the positive quadrant dependent (PQD) order: $\mathbf{D}_t^{\perp} = (D_t^{\perp}, D_t)|\pi \le_{PQD} (D_t, D_t)|\pi = \mathbf{D}_t^o$. By (9.A.18) in Shaked and Shanthikumar (2007), $\mathbf{D}_t^o$ and $\mathbf{D}_t^{\perp}$ are thus ranked in the supermodular order as follows:

$$\mathbf{D}_t^{\perp} = (D_t^{\perp}, D_t)|\pi \le_{sm} (D_t, D_t)|\pi = \mathbf{D}_t^o. \tag{11}$$

We finally have

$$\begin{aligned}
C_t^{\perp}(x|\pi) &= \min_{y \ge x}\left\{c(y-x) + L(y|\pi) + \alpha \mathbb{E}_{\mathbf{D}_t^{\perp} = (D_t^{\perp}, D_t)|\pi}\left[K(D_t^{\perp}, D_t)\right]\right\} \\
&\le \min_{y \ge x}\left\{c(y-x) + L(y|\pi) + \alpha \mathbb{E}_{\mathbf{D}_t^o = (D_t, D_t)|\pi}\left[K(D_t, D_t)\right]\right\} \\
&= \min_{y \ge x}\left\{c(y-x) + L(y|\pi) + \alpha \mathbb{E}_{\mathbf{D}_t^o = (D_t, D_t)|\pi}\left[C_{t+1}^{\perp}(y - D_t|\pi \circ D_t)\right]\right\} \\
&\le \min_{y \ge x}\left\{c(y-x) + L(y|\pi) + \alpha \mathbb{E}_{\mathbf{D}_t^o = (D_t, D_t)|\pi}\left[C_{t+1}(y - D_t|\pi \circ D_t)\right]\right\} \\
&= C_t(x|\pi),
\end{aligned}$$

where the first inequality follows from (11) and the definition of supermodular ordering, and the second follows from the induction assumption. This completes the proof. □

E. Proof of Proposition 8

Proof. Let $\mathbf{d}_T = (d_1, \ldots, d_T)$ and $\mathbf{d}'_T = (d'_1, \ldots, d'_T)$ be two demand paths such that $d_t \le d'_t$ for all t. Let $\mathbf{d}_t = (d_1, \ldots, d_t)$ and $\mathbf{d}'_t = (d'_1, \ldots, d'_t)$ be the subsequences of $\mathbf{d}_T$ and $\mathbf{d}'_T$ until period t, respectively. Let $\mathbf{y}(\pi) = (y_t(\pi \circ \mathbf{d}_{t-1}))$ denote the myopic (also the optimal) policy. It follows that $\underline{\pi} \circ \mathbf{d}_{t-1} \le_{lr} \pi \circ \mathbf{d}_{t-1}$ and hence

$$G(y_t(\underline{\pi} \circ \mathbf{d}_{t-1}) | \underline{\pi} \circ \mathbf{d}_{t-1}) \le G(y_t(\underline{\pi} \circ \mathbf{d}_{t-1}) | \pi \circ \mathbf{d}_{t-1}) \tag{12}$$

for all $\pi \in \mathcal{P}$ because $G(y|\pi) \le G(y|\pi')$ for all $\pi \le_{lr} \pi'$. To see this, note that $D|\pi \le_{st} D|\pi'$ for all $\pi \le_{lr} \pi'$ and that $\min\{y, d\}$ is an increasing function in d.

We also have $\underline{\pi} \circ \mathbf{d}_{t-1} \le_{lr} \underline{\pi} \circ \mathbf{d}'_{t-1} \le_{lr} \pi \circ \mathbf{d}'_{t-1}$ for all $\pi \in \mathcal{P}$, which implies that $y_t(\underline{\pi} \circ \mathbf{d}_{t-1}) \le y_t(\underline{\pi} \circ \mathbf{d}'_{t-1}) \le y_t(\pi \circ \mathbf{d}'_{t-1})$. Therefore, we have

$$G(y_t(\underline{\pi} \circ \mathbf{d}_{t-1}) | \pi \circ \mathbf{d}_{t-1}) \le G(y_t(\underline{\pi} \circ \mathbf{d}_{t-1}) | \pi \circ \mathbf{d}'_{t-1}) \le G(y_t(\underline{\pi} \circ \mathbf{d}'_{t-1}) | \pi \circ \mathbf{d}'_{t-1}), \tag{13}$$

where the first inequality follows from the fact that $\pi \circ \mathbf{d}_{t-1} \le_{lr} \pi \circ \mathbf{d}'_{t-1}$ and the second from the fact that $G(y|\pi \circ \mathbf{d}'_{t-1})$ is increasing over $y_t(\underline{\pi} \circ \mathbf{d}_{t-1}) \le y \le y_t(\pi \circ \mathbf{d}'_{t-1})$. As a consequence,

$$E[G(y_t(\underline{\pi} \circ \mathbf{d}_{t-1}) | \underline{\pi} \circ \mathbf{d}_{t-1}) | \underline{\pi}] \le E[G(y_t(\underline{\pi} \circ \mathbf{d}_{t-1}) | \pi \circ \mathbf{d}_{t-1}) | \underline{\pi}] \le E[G(y_t(\underline{\pi} \circ \mathbf{d}_{t-1}) | \pi \circ \mathbf{d}_{t-1}) | \pi]$$

for all $\pi \in \mathcal{P}$. The expectations are with respect to the random variable $\mathbf{d}_{t-1}$ over $\underline{\pi}$ and π in the first two expressions and the third expression, respectively. The first inequality directly follows from (12) whereas the second is due to (12) and that $\underline{\pi} \le_{lr} \pi$.

Denote by $\Pi_T(\mathbf{y}, \pi)$ the expected total profit when policy $\mathbf{y}$ is employed and π is used as the "true" prior for generating the Bayesian demand process. More specifically, $\Pi_T(\mathbf{y}, \pi) = \sum_{t=1}^{T} E[G(y_t | \pi \circ \mathbf{D}_{t-1}) | \pi]$. With this notation we can write $R_T(\mathcal{P}) = \max_{\mathbf{y}} \min_{\pi \in \mathcal{P}} \Pi_T(\mathbf{y}, \pi)$. Let $\pi^*(\mathbf{y}) = \arg\min_{\pi \in \mathcal{P}} \Pi_T(\mathbf{y}, \pi)$ for any policy $\mathbf{y}$, thus $\Pi_T(\mathbf{y}, \pi^*(\mathbf{y})) \le \Pi_T(\mathbf{y}, \underline{\pi})$ by definition. For policy $\mathbf{y}(\underline{\pi})$, it follows from the previous result that $\pi^*(\mathbf{y}(\underline{\pi})) = \underline{\pi}$, or $\Pi_T(\mathbf{y}(\underline{\pi}), \underline{\pi}) \le \Pi_T(\mathbf{y}(\underline{\pi}), \pi)$ for all $\pi \in \mathcal{P}$. Together with the fact that, for any policy $\mathbf{y}$, $\Pi_T(\mathbf{y}, \underline{\pi}) \le \Pi_T(\mathbf{y}(\underline{\pi}), \underline{\pi}) = V_T(\underline{\pi})$, we have $R_T(\mathcal{P}) = \max_{\mathbf{y}} \min_{\pi \in \mathcal{P}} \Pi_T(\mathbf{y}, \pi) = \max_{\mathbf{y}} \Pi_T(\mathbf{y}, \pi^*(\mathbf{y})) = \Pi_T(\mathbf{y}(\underline{\pi}), \underline{\pi}) = V_T(\underline{\pi})$. □

References

Azoury, Katy S. 1985. Bayes solution to dynamic inventory models under unknown demand distribution. *Management Sci.* **31**(9) 1150–1160.

Azoury, Katy S., Julia Miyaoka. 2009. Optimal policies and approximations for a bayesian linear regression inventory model. *Management Sci.* **55**(5) 813–826. doi:10.1287/mnsc.1080.0980.

Basseville, M., I. Nikiforov. 1993. *Detection of Abrupt Changes: Theory and Application*. Prentice-Hall, Englewood Cliffs, NJ.

Bensoussan, A., M. Çakanyıldırım, J. Minjárez-Sosa, A. Royal, S. Sethi. 2008. Inventory problems with partially observed demands and lost sales. *J. Optim. Theory Appl.* **136**(3) 321–340. 0022-3239.

Bensoussan, Alain, Metin Çakanyıldırım, Suresh P. Sethi. 2007. A multiperiod newsvendor problem with partially observed demand. *Math. Oper. Res.* **32**(2) 322–344.

Besbes, Omar, Assaf Zeevi. 2011. On the minimax complexity of pricing in a changing environment. *Oper. Res.* **59**(1) 66–79.

Brown, David B., James E. Smith, Peng Sun. 2010. Information relaxations and duality in stochastic dynamic programs. *Oper. Res.* **58**(4-Part-1) 785–801.

Chen, L., A. J. Mersereau. 2013. Analytics for operational visibility in the retail store: Censored demand and inventory record inaccuracy. Working paper, Duke University.

Chen, Li. 2010. Bounds and heuristics for optimal Bayesian inventory control with unobserved lost sales. *Oper. Res.* **58**(2) 396–413.

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

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

Graves, Stephen C. 1999. A single-item inventory model for a nonstationary demand process. *Manufacturing Service Oper. Management* **1**(1) 50–61.

Huh, Woonghee Tim, Retsef Levi, Paat Rusmevichientong, James B. Orlin. 2011. Adaptive data-driven inventory control with censored demand based on Kaplan-Meier estimator. *Oper. Res.* **59**(4) 929–941.

Iida, Tetsuo, Paul H. Zipkin. 2006. Approximate solutions of a dynamic forecast-inventory model. *Management Sci.* **8**(4) 407–425. doi:10.1287/msom.1060.0116.

Karlin, Samuel. 1960. Dynamic inventory policy with varying stochastic demands. *Management Sci.* **6**(3) 231–258.

Karlin, Samuel, H. Rubin. 1956. Distributions possessing a monotone likelihood ratio. *Journal of the American Statistical Association* **51**(276) 637–643.

Lai, Tze Leung. 1995. Sequential changepoint detection in quality control and dynamical systems. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**(4) 613–658.

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

Lovejoy, William S. 1990. Myopic policies for some inventory models with uncertain demand distributions. *Management Sci.* **36**(6) 724–738. doi:10.1287/mnsc.36.6.724.

Lu, Xiangwen, Jing-Sheng Song, Amelia Regan. 2006. Inventory planning with forecast updates: Approximate solutions and cost error bounds. *Oper. Res.* **54**(6) 1079–1097. doi:10.1287/opre.1060.0338.

Morton, Thomas E., David W. Pentico. 1995. The finite horizon nonstationary stochastic inventory problem: Near-myopic bounds, heuristics, testing. *Management Sci.* **41**(2) 334–343.

Osadchiy, Nikolay, Vishal Gaur, Sridhar Seshadri. 2013. Sales forecasting with financial indicators and experts' input. *Production and Operations Management* **22**(5) 1056–1076. doi:10.1111/poms.12022. URL `http://dx.doi.org/10.1111/poms.12022`.

Scarf, Herbert. 1959. Bayes solutions of the statistical inventory problem. *The Annals of Mathematical Statistics* **30**(2) 490–508.

Scarf, Herbert E. 1960. Some remarks on Bayes solutions to the inventory problem. *Naval Research Logistics Quarterly* **7**(4) 591–596.

Sethi, Suresh P., Feng Cheng. 1997. Optimality of (s, S) policies in inventory models with Markovian demand. *Oper. Res.* **45**(6) 931–939.

Shaked, Moshe, J. George Shanthikumar. 2007. *Stochastic Orders*. Springer, New York.

Song, Jing-Sheng, Paul Zipkin. 1993. Inventory control in a fluctuating demand environment. *Oper. Res.* **41**(2) 351–370.

Song, Jing-Sheng, Paul H. Zipkin. 1996. Managing inventory with the prospect of obsolescence. *Oper. Res.* **44**(1) 215–222.

Treharne, James T., Charles R. Sox. 2002. Adaptive inventory control for nonstationary demand and partial information. *Management Sci.* **48**(5) 607–624.