POMS PRODUCTION AND OPERATIONS MANAGEMENT

PRODUCTION AND OPERATIONS MANAGEMENT POMS
Vol. 26, No. 2, February 2017, pp. 341–359
ISSN 1059-1478 | EISSN 1937-5956 | 17 | 2602 | 0341
DOI 10.1111/poms.12650

Bayesian Inventory Management with Potential Change-Points in Demand

Zhe (Frank) Wang
200 E. Dana Street, Mountain View, California 94041, USA, fzhewang@gmail.com

Adam J. Mersereau
Kenan-Flagler Business School, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3490, USA, ajm@unc.edu

We consider the inventory management problem of a firm reacting to potential change points in demand, which we define as known epochs at which the demand distribution may (or may not) abruptly change. Motivating 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 product redesign). In the periods following such a potential change point in demand, a manager is torn between using a possibly obsolete demand model estimated from a long data history and using a model estimated from a short, recent history. We formulate a Bayesian inventory problem just after a potential change point. We pursue heuristic policies coupled with cost lower bounds, including a new lower bounding approach to non-perishable Bayesian inventory problems that relaxes the dependence between physical demand and demand signals and that can be applied for a broad set of belief and demand distributions. Our numerical studies reveal small gaps between the costs implied by our heuristic solutions and our lower bounds. We also provide analytical and numerical sensitivity results suggesting that a manager worried about downside profit risk should err on the side of underestimating demand at a potential change point.

Key words: inventory theory and control; retailing; demand forecasting; dynamic programming
History: Received: February 2014; Accepted: September 2016 by Felipe Caro, after 2 revisions.

1. Introduction

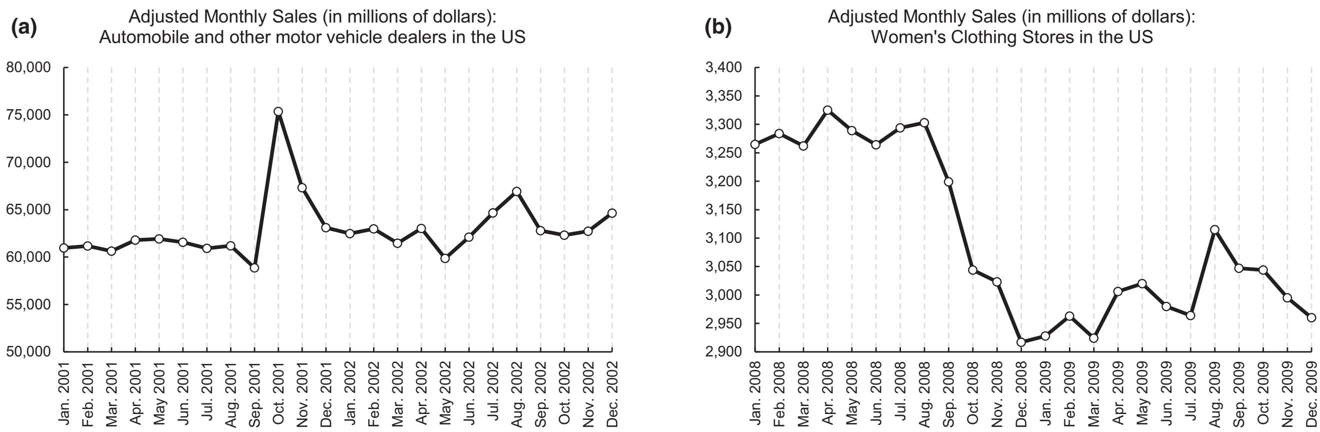
In most real-world inventory control problems, demand changes over time and the true underlying demand distribution is never fully known to the inventory manager. The manager makes dual use of historical demand data to populate the current demand distribution and also to detect fundamental changes in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

Figure 1 Examples of Potential Change Points in Demand Include (a) the Terrorism Events of September 2001 and (b) the Lehman Brothers Bankruptcy in September 2008



Source: U.S. Census Bureau.

the opening of a nearby attraction can all potentially usher in new demand regimes for a firm. The introduction of 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 trade-off 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 control 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 to understand the structure and behavior of the optimal policy, and we look for computationally tractable bounds and heuristics.

We model the evolution of the manager's belief on the demand process using a Bayesian framework, extending the model pioneered by Scarf (1959) to allow for an unknown demand parameter to be distributed according to a mixture of a "historical" prior distribution and a "change" prior distribution. We leverage the structure of our demand model to characterize the effects of observed demand and the manager's belief on the optimal (state-dependent) base-stock levels.

The optimal policy remains challenging to compute. Scarf (1959) and Azoury (1985) show that the stationary Bayesian inventory problem can be solved efficiently using a dimensionality reduction approach for particular assumptions on the prior and demand distributions, but these assumptions do not hold when the unknown parameter is described by a mixture of distributions. We pursue heuristic policies coupled with cost lower bounds specific to our setting. Our most sophisticated bounding approach is novel in its formulation of a "conditionally orthogonal" 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. In contrast to the dimensionality reduction approach of Scarf and Azoury, this approach can be applied for a broad set of belief and demand distributions. An extensive numerical analysis reveals that this bound and a look-ahead policy derived from it achieve small gaps. The numerical study also reveals that a myopic policy that accounts for potential change points (but that ignores future inventory dynamics) works well 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 study is organized as follows. We review related literature in section 2. In section 3, we formulate our Bayesian demand model and associated inventory control problem, and we present structural properties of the optimal inventory policy.

In section 4, we develop lower bounds for the optimal expected cost, and we introduce heuristic policies derived from these lower bounds. We numerically study these bounds and policies and measure their performance in section 5. In section 6, we discuss the estimation of model parameters and sensitivity to parameter misspecification. We conclude in section 7.

2. Literature Review

This study relates to the inventory control literature dealing with nonstationary and/or partially observed demand processes. For situations in which 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 in fluctuating demand environments. They assume that the demand distribution changes regime according to a known Markov chain and that the demand distribution in each regime is also fully known. Under these assumptions, they establish the optimality of state-dependent (s, S) policies. Sethi and Cheng (1997) show similar results in a generalized 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.

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 on inventory management problems when lost sales are unobserved and demand is therefore censored, assuming stationary demand (Bensoussan et al. 2007, 2008, Chen 2010, Chen and Plambeck 2008, Ding et al. 2002, Huh et al. 2011, Lariviere and Porteus 1999). Chen and Mersereau (2015) 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 shift in the past. This simplification yields structure that we exploit in deriving new results and bounds. Second, we model component demand distributions that are learned over time, whereas Treharne and Sox (2002) assume the demand distribution within each Markov state is known and fixed. We believe that our model brings distinct advantages in flexibility and parsimony. For further discussion, see section 3.2. While the bounds we develop in section 4 make use of results in Brown et al. (2010), we believe our "conditionally orthogonal" bound to be new.

In as much as our study considers a change in demand regime, it also relates to Besbes and Zeevi (2011), in which a decision-maker seeks to detect and exploit a potential change in customers' willingness-to-pay distribution through dynamic pricing.

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), Lai (1995), and the recent text Tartakovsky et al. (2015). This literature most commonly seeks to identify *when* a change occurs, focusing on the trade-off between detection delay and the risk of false alarm. Our interest is not in declaring when a change point occurs; rather, we formulate a dynamic optimization problem built on a stochastic model involving a potential change point. Our model specializes typical sequential change-point formulations in that we assume that the timing of our potential change point is known. In our model, a key unknown is whether or not the change actually occurs.

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, which we use in our algorithm development in section 4.

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 inventory position 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.

We assume the DM fully observes past demands without censoring, as does Scarf (1959). This assumption is driven in part by analytical tractability (as is our assumption of inventory backlogging), but we believe it is reasonable in practice when changes in demand are likely to impact a whole department, firm, or industry at the same time. This is the case, for example, for the September 2001 and Lehman Brothers bankruptcy contexts described in section 1. In such cases, the firm can use data across stock-keeping units to correct for demand censoring.

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 are independently drawn according to a density function $f(\cdot|\theta)$, where $\theta \in \Theta$ is an unknown parameter. The DM has a *historical prior* π^h (h stands for "history") on θ which reflects his prior knowledge of the demand parameter based on historical information. A *potential* change point occurs in period 1; thereafter, the DM is uncertain about whether the historical prior π^h continues to apply or whether the demand process has changed. The DM has a second prior distribution π^c on θ conditional on a change occurring (the superscript c stands for "change"). The *change probability* γ represents the DM's initial belief that a change has indeed occurred in period 1.

In practice, it is reasonable for the DM to estimate the historical prior using historical demand. However, it may be less obvious how to estimate the change prior π^c and the probability γ. We provide a full discussion of this in section 6, where we perform a sensitivity analysis and suggest robust choices for these parameters.

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,

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

We will show in section 3.4 that this update has a particular structure that enables our analysis. The predictive demand density in period t given belief π_t is defined by $\phi(\xi|\pi_t) = \int_\Theta f(\xi|\theta)\pi_t(\theta)d\theta$. 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 2 in section 3.4, 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 inventory position y after ordering and a demand realization d, the holding and shortage cost incurred in a single period is

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

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

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

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

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

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

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

3.2. Discussion of our Demand Model

Our choice of a mixture model as a prior distribution for the unknown demand parameter is driven by our interest, as discussed in section 1, in situations in which the DM has reason to believe a change in demand regime *may* have just occurred but is uncertain about whether a fundamental change has really transpired and, if so, about its extent. Our mixture model explicitly models this uncertainty. Such problems are most relevant and interesting in the few periods just after the potential change, and our choice of a parametric Bayesian model permits meaningful demand learning even with a few observations.

We use our model to illustrate numerically in Figure 2 the core demand learning tradeoff we seek to capture. The left panel of Figure 2 corresponds with a single demand path involving a change in the demand mean from 10 to 5 occuring at time $t = 0$, while the right panel corresponds to a demand path drawn from a stationary demand process with demand mean 10 units.[1] The gray curves show statistics of the predictive demand distribution under our mixture model. For comparison, we also plot predictive demand statistics for models that use the historical prior alone and the change prior alone from time $t = 0$ on.

In the left-hand plot, we see that the mixture model more quickly learns the changed demand mean compared with the model using the historical prior alone, while (as expected) not quite as quickly as the model that assumes a change definitely occurred. In the right-hand plot, we see that the coefficient of variation (CoV) for our mixture model jumps considerably less and stabilizes more quickly than the model that relies on the change prior alone. We conclude that our "mixture" model of demand learning achieves a robust balance of responsiveness (in the event a change actually occurs) and stability (in the event no change occurs).

Further testing (omitted to conserve space) shows the necessity of allowing for the component distributions (in particular, the change distribution) to be learned from data rather than fixed *a priori* in situations where the DM has uncertainty around the post-change demand parameter. Fixing and mis-specifying θ may prevent the mixture model from converging to the true mean and variance of demand.

Finally, we have considered alternative modeling approaches for modeling change points. Hypothesis testing-based approaches to change-point detection (e.g., Tartakovsky et al. 2015) do not naturally lead to forward-looking distributional forecasts that we require for multiperiod inventory control. Non-parametric methods (e.g., Huh et al. 2011) offer no concise state representation for use in a forward-looking dynamic optimization formulation.

3.3. Structure of the Optimal Policy

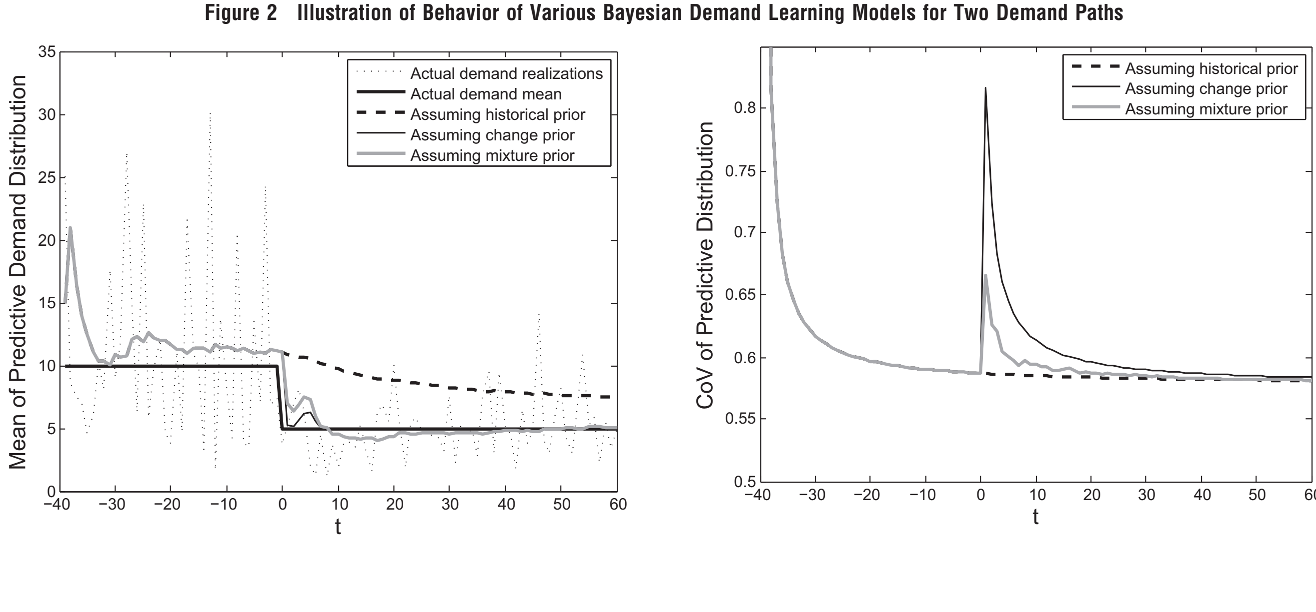
Although the demand process described in section 3.1 is complicated by the potential change points, it is still independent of the ordering decisions. Because of this, the cost functions are convex and a state-dependent base-stock policy is optimal. We state the following result for completeness, but 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 position 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,

Figure 2 Illustration of Behavior of Various Bayesian Demand Learning Models for Two Demand Paths



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 (see Theorem 9.4.2 of Zipkin 2000; also Karlin 1960, 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 across 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., $\Phi(d_t|\pi_t) = \int_0^{d_t} \phi(\xi|\pi_t)d\xi$. Then, the base-stock level for period t under a myopic policy is given by $y_t^M(\pi_t)$ such that

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

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 Propositions 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. Chen and Plambeck (2008) also show that a DM may want to stock less than the myopic inventory level when inventory is perishable, albeit in a different Bayesian inventory setting than ours (with stationary demand and censored observations).

3.4. 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.

3.4.1. 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 $d_1 > d_2, f(d_1)/g(d_1) \geq f(d_2)/g(d_2)$.

3.4.2. Monotone Likelihood Ratio Property. 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 conditionally orthogonal 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 require 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. We present an example in Appendix B showing the necessity of the MLRP assumption on $f(\cdot|\theta)$.

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.

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 as follows:

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 $\int_0^\infty U_{t+1}(u) \cdot \psi_t(u)du < \infty$, 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 to 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 lower bounds in this subsection. The first makes use of the decomposition of Lemma 1, while the second makes use of a novel relaxation we call the "conditionally orthogonal" problem. Both make use of the "information relaxation" framework outlined in Brown et al. (2010).

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. (Given that the DM is seeking to minimize cost, the additional information revealed by the oracle can only help achieve lower 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, and so the proposition follows from Lemma 2.1 in Brown et al. (2010). □

4.1.2. The Conditionally Orthogonal Lower Bound. In a Bayesian inventory problem like ours, demand realizations factor into both the update of the inventory position and the update of beliefs. We construct a lower-bounding problem by decoupling the inventory position update from the demand belief update and by revealing to the DM the demand signal available for the belief update. The latter step applies the notion of information relaxations (Brown et al. 2010).

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 positions 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 a "conditionally orthogonal" problem in which the physical demand and the demand signal are assumed independent of each other, conditional on currently available information. 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 orthogonal problems, respectively, for periods $t, \ldots, T$ given initial inventory position x_t and belief π_t. Then we have

$$C_t(x_t|\pi_t) = \min_{y \ge x_t}\Big\{c(y-x_t) + L(y|\pi_t) + \alpha\mathbb{E}_{\mathbf{D}_t^o=(D_t,D_t)|\pi_t}[C_{t+1}(y-D_t|\pi_t \circ D_t)]\Big\},$$

$$C_t^\perp(x_t|\pi_t) = \min_{y \ge x_t}\Big\{c(y-x_t) + L(y|\pi_t) + \alpha\mathbb{E}_{\mathbf{D}_t^\perp=(D_t^\perp,D_t)|\pi_t}[C_{t+1}^\perp(y-D_t^\perp|\pi_t \circ D_t)]\Big\},$$

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 conditionally orthogonal problem serves as a lower bound for that of the original problem.

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

The proof, in Appendix E, 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 conditionally orthogonal 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 conditionally orthogonal 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 conditionally orthogonal 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

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

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 position x_t, belief π_t and future demand signals $(d_t, \ldots, d_T)$. The conditionally orthogonal problem after relaxing future demand signals reduces to

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

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 conditionally orthogonal problem, the optimal expected cost of the reduced problem will be lower than that of the conditionally orthogonal one.

We formally state the conditionally orthogonal lower bound as follows.

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

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

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

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 $LB_t^O(x_t|\pi_t)$ in the numerical results using the following simulation 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 conditionally orthogonal lower bound.

To our knowledge, the "conditionally orthogonal" 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 signal paths.

In our numerical results, we estimate the bound based on a large number (100,000) of signal paths, which yields very small standard errors as reported in the electronic companion. Calculating LB^O for the instances we considered (which we will describe in detail in section 5.3), takes on average 0.124 seconds per signal path using MATLAB on a 3.2GHz personal computer. Our implementation is a naïve one in that we do not attempt to take advantage of parallel processing. We note that the bound is the mean of a quantity evaluated over independent signal paths, therefore the computation time grows approximately linearly in the number of signal paths N and the standard error shrinks as $1/\sqrt{N}$.

The approach may be useful for inventory problems involving demand learning beyond the one considered in this study. 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 under certain assumptions on the demand distribution. Without these assumptions, the optimal policy remains difficult to compute. In section 5.2 we demonstrate that the conditionally orthogonal 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 signal 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 conditionally orthogonal 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 conditionally orthogonal lower bound. In section 5, we evaluate these heuristics using the lower bounds in section 4.1.

4.2.1. 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.

4.2.2. 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:

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

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).

4.2.3. Look-Ahead Policy Based on Conditionally Orthogonal Lower Bound (LA-O). This policy is very similar to the LA-M policy except that it uses the conditionally orthogonal lower bound LB^O_{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:

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

where $LB^O_{t+1}(\cdot|\cdot)$ is estimated using simulation as described in section 4.1.2. This LA-O policy can be applied to the single change-point problem with any belief and demand distribution; however, the computational effort required grows with the number of signal paths used to estimate the LB^O 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 (α = 1) throughout the section. We have also run our experiments with discount factor α = 0.8 and found that the results do not change qualitatively.

We make use of the gamma-gamma conjugate pair as our model of demand in our numerical results. This demand structure is amenable to the dimensionality reduction technique of Scarf (1960) and Azoury (1985) for stationary versions of our problem. Given this demand structure we can therefore easily compute the degenerate problems required to compute the LB^M bound and the LA-M policy.

We will first review the gamma-gamma demand model and its relevant properties in section 5.1. We will then test the conditionally orthogonal 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 and compare our bound against the known optimal solution. Finally in section 5.3, we will perform a comprehensive numerical study on bounds and heuristics for the potential change-point problem analyzed in sections 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, Chen 2010, Scarf 1960) 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, ..., d_{t-1}), it is well-known that sufficient statistics for Bayes updating are

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

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

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

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:

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

with $v_{T+1}(\cdot) = 0$. Let y_t^* denote the optimal base-stock level for period t. Then we have

(i) $C_t(x|S_t) = S_t v_t(x/S_t)$,
(ii) $y_t^*(S_t) = S_t y_t^*$.

Property (i) 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 Conditionally Orthogonal Lower Bound to a Classical Problem

In this subsection, we use the classic Bayesian inventory problem with gamma-gamma demand from Scarf (1960) to explore the behavior and quality of the conditionally orthogonal lower bound. This problem is a special case of our change-point problem with change probability equal to zero (or one), and it can be solved using a dimension reduction technique, as previously discussed. Therefore, it qualifies as a reasonable testbed for understanding the potential tightness of the conditionally orthogonal lower bound.

We include a detailed description of the study and a table of results in the electronic companion to this study. We make two observations about the results, which cover 36 instances in a full factorial design. 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 instances. Second, the conditionally orthogonal 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 conducted a study of the instances in Table 1 of the electronic companion using summary tables and linear regression, and we found that the gap tends to increase with the shape parameter k of the gamma demand distribution, increase with the target service level (represented by p), increase with the prior CoV (measured by $1/\sqrt{a}$), and decrease with the time horizon T. Indeed, the instance with the largest gap (3.94%) is the one with "worst case" parameter values in light of these sensitivities.

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-O—for the single change-point problem introduced in section 3 by comparing their expected costs with the lower bounds.

We assume 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 $k = 5$ 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). The use of gamma distributions for both demand and the two prior distributions enables us to compute both the LB^M lower bound and

Table 1 Mean Percentage Gaps* for Moderate Change Cases, Averaged over Parameter Levels

Parameter		Myopic (%)	LA-M (%)	LA-O (%)	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.

its corresponding policy efficiently. 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 estimate the LB^O bound using 100,000 simulated demand signal paths to evaluate the expectation in Proposition 7. We also estimate the expected cost performance of the Myopic, LA-M, and LA-O policies by simulating their performance on an independent set of 10,000 demand sample paths. (The LA-O policy requires an evaluation of LB^O for each sample path, for which we use small simulations of 1000 signal paths.) We also use the same set of 10,000 sample paths used for the policy evaluation 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 ignored the potential change point and used only the historical demand information for forecasting and inventory decisions. The OPTCHG policy would be employed if the DM discards 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 electronic companion for detailed tables of results. We find the estimated conditionally orthogonal lower bound LB^O to be tighter than the mixture lower bound LB^M across all 60 instances, and so we use our estimated LB^O to bound optimality gaps of the various policies. We calculate the deviation of each policy's estimated expected costs from the estimated LB^O (computed by $(\text{Cost} - LB^O)/LB^O \times 100\%$) for each instance. The percent gaps, averaged over parameter levels, are summarized in Table 1 (for "moderate change" cases) and Table 2 (for "extreme change" cases). From these tables, we can infer the sensitivities of the gaps to various problem parameters; e.g., the gaps tend to grow with the service level (represented by p) and decrease with the time horizon T.

We make several observations about the results in Tables 1 and 2. First, for both moderate and extreme scenarios, myopic, LA-M and LA-O 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-O 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, 1992). 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% over these instances, while the LA-M and LA-O policies have much smaller average gaps (5.72% and 5.74%, respectively) than the myopic

Table 2 Mean Percentage Gaps* for Extreme Change Cases, Averaged over Parameter Levels

Parameter		Myopic (%)	LA-M (%)	LA-O (%)	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.

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 overestimate the true optimality gaps.

Finally, although we have observed that LB^O is tighter than LB^M for all instances, the LA-M and LA-O policies (which approximate cost-to-go functions by LB^M and LB^O, respectively) have nearly the same performance under all scenarios. Recall that the LA-M policy can only be efficiently computed for cases in which the "degenerate" problems referenced in section 4.1.1 can be solved easily. The LA-M policy is recommended for such cases; for other cases, the LA-O policy is likely to be more efficient to compute.

6. Parameter Estimation and Sensitivity

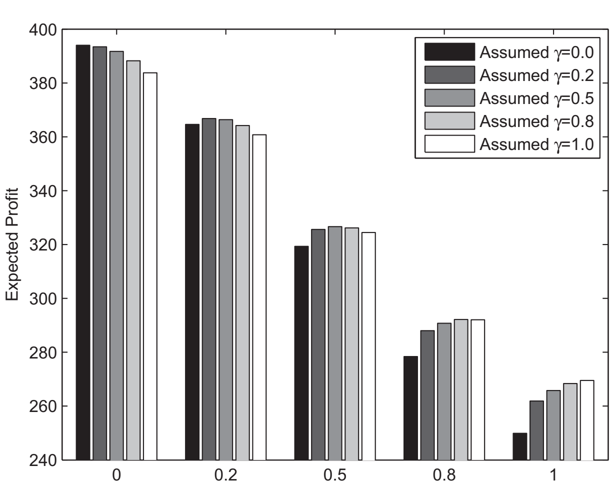
The demand model of section 3 requires the specification of three inputs: a "no-change" or historical prior π^h, a change prior π^c, and a change probability γ. The no-change prior π^h, the forecast of demand in the absence of a potential change point, can be estimated using established techniques applied 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.

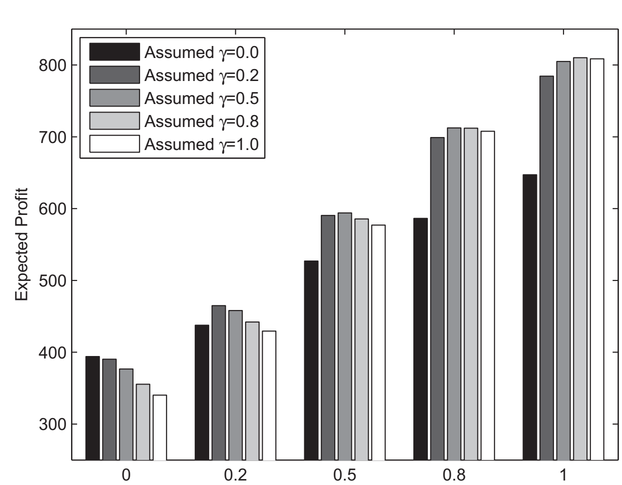
Selecting the change prior π^c entails predicting the direction and magnitude of a potential change. This represents a new demand regime for the firm by definition, but in many cases it may be a regime with past precedents. Imagine a retailer facing the entrance of a new competitor at one of its locations. 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 may provide signals that can 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 misspecification of γ. Figure 3 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 $\mathbb{E}[p\min\{y,D\} - c(y-x) - h(y-D)^+] = -\mathbb{E}[c(y-x) + h(y-D)^+ + p(D-y)^+] + p\mathbb{E}[D]$. The results in Figure 3 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.

Figure 3 **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 γ = 0.0 (black), 0.2, 0.5, 0.8, and 1.0 (white)**





We observe from Figure 3 that the profits are always highest for each instance when the assumed γ matches the "true" one used to generate the demand data. We also observe that the expected profit for each instance remains relatively flat as we move the assumed change probability γ from 0.2 to 0.5 to 0.8, and that assuming $\gamma = 0.5$ exhibits robust performance across all of the instances.

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 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 Equation (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

Our numerical study yields several insights on inventory management in uncertain demand environments. First, if a manager suspects a demand regime change, the manager is best served 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 merit further research. One important extension would be to the case with censored demand, in which the DM's future observations depend on current ordering decisions. 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, interesting questions arise when a potential change point is anticipated in the future (as opposed to the case we have considered where the potential change point is at a known time in the past). Third, it would seem relevant to inventory management practice to allow for uncertainty in the timing of potential change points in order to model demand shifts that occur for unobserved reasons. Fourth, we believe that the conditionally orthogonal bound idea may merit further investigation for other inventory models involving demand learning.

Acknowledgments

The authors thank the department editor Felipe Caro, an anonymous senior editor, and four anonymous reviewers for their constructive reviews of our study, David Brown for his insightful comments on an earlier version of the paper, and Li Chen, Vinayak Deshpande, Nur Sunar, and Serhan Ziya for helpful feedback. This paper represents research undertaken as part of the first author's Ph.D. dissertation at the University of North Carolina Kenan-Flagler Business School. The second author thanks the Sarah Graham Kenan Foundation for financial support.

Appendix A. Proof of Proposition 2

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

Appendix B. Proof of Proposition 3

We postpone the proof of part (a) until Appendix E, 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).

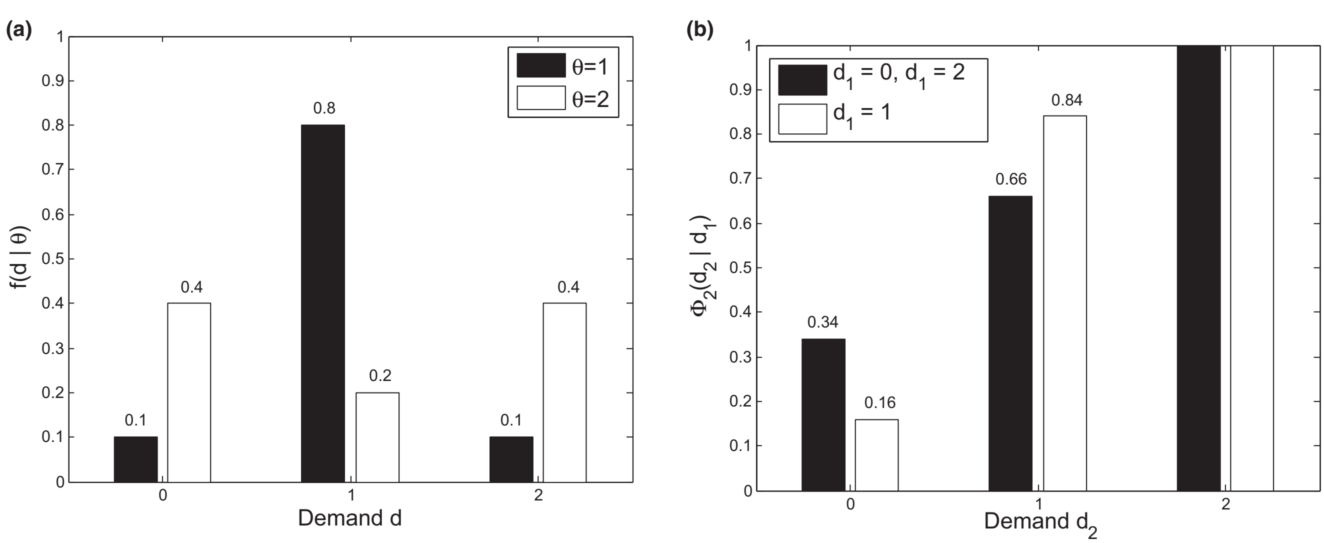
We provide an example showing that it is necessary for $f(\cdot|\theta)$ to have the MLRP property for part (b) to hold. Let the demand parameter θ take values in the set $\{1, 2\}$. Demand in each period is 0, 1, or 2 units, and the demand probability mass function is shown in Figure B1. Note that $f(\cdot|\theta)$ does not have the MLRP property; in particular, $\frac{f(0|\theta=2)}{f(0|\theta=1)} = \frac{f(2|\theta=2)}{f(2|\theta=1)} = 4 > 0.25 = \frac{f(1|\theta=2)}{f(1|\theta=1)}$. Now consider a two-period inventory problem. Figure B1 shows the predictive cumulative demand distribution given a uniform initial prior $\pi_1(\theta=1) = \pi_1(\theta=2) = 0.5$. Suppose we choose cost parameters such that the newsvendor critical ratio determining the period 2 base-stock level is 0.70. Then the optimal base-stock level in period 2 is two units if $d_1 = 0$ or $d_1 = 2$ but is one unit if $d_1 = 1$. Therefore, the optimal base-stock level is not increasing in d_1.

Appendix C. Proof of Lemma 1

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

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

Figure B1 An Example Showing that Proposition 3(b) May Not Hold if $f(\cdot|\theta)$ Does Not Have the Monotone Likelihood Ratio Property



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

$$\pi^i_{t+1}(\theta|\mathbf{d}_t) = \frac{\pi^i_t(\theta|\mathbf{d}_{t-1})f(d_t|\theta)}{\int_\Theta \pi^i_t(\omega|\mathbf{d}_{t-1})f(d_t|\omega)d\omega},$$

and

$$\pi_{t+1}(\theta|\mathbf{d}_t) = \frac{\begin{array}{c}[(1-\gamma_t(\mathbf{d}_{t-1}))\pi^h_t(\theta|\mathbf{d}_{t-1}) \\ +\gamma_t(\mathbf{d}_{t-1})\pi^c_t(\theta|\mathbf{d}_{t-1})]f(d_t|\theta)\end{array}}{\begin{array}{c}\int_\Theta[(1-\gamma_t(\mathbf{d}_{t-1}))\pi^h_t(\omega|\mathbf{d}_{t-1}) \\ +\gamma_t(\mathbf{d}_{t-1})\pi^c_t(\omega|\mathbf{d}_{t-1})]f(d_t|\theta)d\omega\end{array}}.$$

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

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

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. □

Appendix D. Proof of Proposition 4

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 $\pi^h_t \leq_{lr} \pi_t \leq_{lr} \pi^c_t$ 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). □

Appendix E. Proof of Proposition 6

For the purposes of this section, we consider a 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 position x and prior distribution π, where

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

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 conditionally orthogonal problems are both special cases of this formulation. In the original problem, $\hat{D}_t = D_t$, whereas in the conditionally orthogonal problem, $\hat{D}_t$ and D_t are independent with the same distribution induced by π.

The proof of Proposition 1 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

$$\frac{\partial \hat{C}_{t+1}}{\partial x}(x-\hat{d}_t|\pi^1\circ d_t) \geq \frac{\partial \hat{C}_{t+1}}{\partial x}(x-\hat{d}_t|\pi^2\circ d_t), \qquad \text{(E1)}$$

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

$$\frac{\partial \hat{C}_{t+1}}{\partial x}(x-\hat{d}_t^1|\pi^2\circ d_t^1) \geq \frac{\partial \hat{C}_{t+1}}{\partial x}(x-\hat{d}_t^1|\pi^2\circ d_t^2) \qquad \text{(E2)}$$

$$\geq \frac{\partial \hat{C}_{t+1}}{\partial x}(x-\hat{d}_t^2|\pi^2\circ d_t^2), \qquad \text{(E3)}$$

where Equation (E2) 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 Equation (E3) 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

$$(\hat{D}_t, D_t)|\pi^1 \leq_{st} (\hat{D}_t, D_t)|\pi^2. \qquad \text{(E4)}$$

We thus have

$$\mathbb{E}_{(\hat{D}_t,D_t)|\pi^1}\left[\frac{\partial \hat{C}_{t+1}}{\partial x}(x-\hat{D}_t|\pi^1\circ D_t)\right] \geq \mathbb{E}_{(\hat{D}_t,D_t)|\pi^1}\left[\frac{\partial \hat{C}_{t+1}}{\partial x}(x-\hat{D}_t|\pi^2\circ D_t)\right] \qquad \text{(E5)}$$

$$\geq \mathbb{E}_{(\hat{D}_t,D_t)|\pi^2}\left[\frac{\partial \hat{C}_{t+1}}{\partial x}(x-\hat{D}_t|\pi^2\circ D_t)\right], \qquad \text{(E6)}$$

where Equation (E5) results from (E1), and Equation (E6) from (E4) 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 $\hat{y}_t^*(\pi^1) \leq \hat{y}_t^*(\pi^2)$.

It remains to show that $\frac{\partial \hat{C}_t}{\partial x}(x|\pi^1) \geq \frac{\partial \hat{C}_t}{\partial x}(x|\pi^2)$. 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) $\hat{y}_t^*(\pi^1) \leq x < \hat{y}_t^*(\pi^2)$. In this case, $\frac{\partial \hat{C}_t}{\partial x}(x|\pi^1) = \hat{H}(x|\pi^1) \geq \hat{H}(\hat{y}_t^*(\pi^1)|\pi^1) = -c = \frac{\partial \hat{C}_t}{\partial x}(x|\pi^2)$;

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

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 $t + 1$, $C_{t+1}^{\perp}(x|\pi) \leq C_{t+1}(x|\pi)$ 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 $d_t^1 \leq d_t^2$, Lemma 2 of Chen (2010) implies that $\pi\circ d_t^1 \leq_{lr} \pi\circ d_t^2$. Lemma 4 therefore yields

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

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

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

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: $\mathbf{D}_t^{\perp} = (D_t^{\perp}, D_t)|\pi \leq_{PQD} (D_t, D_t)|\pi = \mathbf{D}_t^o$. 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:

$$\mathbf{D}_t^{\perp} = (D_t^{\perp}, D_t)|\pi \leq_{sm} (D_t, D_t)|\pi = \mathbf{D}_t^o. \qquad \text{(E7)}$$

We finally have

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

where the first inequality follows from Equation (E7) and the definition of supermodular ordering, and the second follows from the induction assumption. This completes the proof. □

Appendix F. Proof of Proposition 8

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

$$G(y_t(\underline{\pi}\circ\mathbf{d}_{t-1})|\underline{\pi}\circ\mathbf{d}_{t-1})\leq G(y_t(\underline{\pi}\circ\mathbf{d}_{t-1})|\pi\circ\mathbf{d}_{t-1}) \quad \text{(F1)}$$

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

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

$$
\begin{aligned}
&G(y_t(\underline{\pi}\circ\mathbf{d}_{t-1})|\pi\circ\mathbf{d}_{t-1}) \\
&\leq G(y_t(\underline{\pi}\circ\mathbf{d}_{t-1})|\pi\circ\mathbf{d}'_{t-1})\leq G(y_t(\underline{\pi}\circ\mathbf{d}'_{t-1})|\pi\circ\mathbf{d}'_{t-1}),
\end{aligned} \quad \text{(F2)}
$$

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

$$
\begin{aligned}
&E[G(y_t(\underline{\pi}\circ\mathbf{d}_{t-1})|\underline{\pi}\circ\mathbf{d}_{t-1})|\underline{\pi}] \\
&\leq E[G(y_t(\underline{\pi}\circ\mathbf{d}_{t-1})|\pi\circ\mathbf{d}_{t-1})|\underline{\pi}] \\
&\leq E[G(y_t(\underline{\pi}\circ\mathbf{d}_{t-1})|\pi\circ\mathbf{d}_{t-1})|\pi]
\end{aligned}
$$

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 (F1) whereas the second is due to (F1) and that $\underline{\pi}\leq_{lr}\pi$.

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

Note

[1]The specific instance is similar to those in section 5.3 (i.e., gamma demand distribution). Using notation to be introduced later, we assume an initial change prior of $\gamma_0=0.5$, we assume a known "shape" parameter $k=3$ for gamma demand, and we assume a "change" gamma prior with $a^c=3$, $S^c=5$. The "historical" gamma prior is generated based on the observations from time $t=-40$ to $t=0$ starting from $(a_{-40},S_{-40})=(3,10)$.

References

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

Azoury, K. S., J. 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. Ç̧akanyildirim, S. P. Sethi. 2007. A multiperiod newsvendor problem with partially observed demand. *Math. Oper. Res.* **32**(2): 322–344.

Bensoussan, A., M. Çakanyildirim, 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.

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

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

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

Chen, L., A. J. Mersereau. 2015. Analytics for operational visibility in the retail store: The cases of censored demand and inventory record inaccuracy. N. Agrawal, S. A. Smith, eds. *Retail Supply Chain Management*. Springer Science+Business Media, New York, 79–112.

Chen, L., E. L. Plambeck. 2008. Dynamic inventory management with learning about the demand distribution and substitution probability. *Manuf. Serv. Oper. Manag.* **10**(2): 236–256.

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

Graves, S. C. 1999. A single-item inventory model for a nonstationary demand process. *Manuf. Serv. Oper. Manag.* **1**(1): 50–61.

Huh, W. T., R. Levi, P. Rusmevichientong, J. B. Orlin. 2011. Adaptive data-driven inventory control with censored demand based on Kaplan-Meier estimator. *Oper. Res.* **59**(4): 929–941.

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

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

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

Lai, T. L. 1995. Sequential change-point detection in quality control and dynamical systems. *J. R. Stat. Soc. Ser. B Methodol.* **57**(4): 613–658.

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. doi:10.1287/mnsc.36.6.724.

Lovejoy, W. S. 1992. Stopped myopic policies in some inventory models with generalized demand processes. *Management Sci.* **38**(5): 688–707. doi:10.1287/mnsc.38.5.688.

Lu, X.J.-S. Song, A. 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, T. E., D. W. Pentico. 1995. The finite horizon nonstationary stochastic inventory problem: Near-myopic bounds, heuristics, testing. *Management Sci.* **41**(2): 334–343.

Osadchiy, N., V. Gaur, S. Seshadri. 2013. Sales forecasting with financial indicators and experts' input. *Prod. Oper. Manag.* **22**(5): 1056–1076. doi:10.1111/poms.12022.

Scarf, H. 1959. Bayes solutions of the statistical inventory problem. *Ann. Math. Stat.* **30**(2): 490–508.

Scarf, H. E. 1960. Some remarks on Bayes solutions to the inventory problem. *Nav. Res. Log. Q.* **7**(4): 591–596.

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

Shaked, M., J. G. Shanthikumar. 2007. *Stochastic Orders*. Springer, New York.

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

Song, J.-S., P. H. Zipkin. 1996. Managing inventory with the prospect of obsolescence. *Oper. Res.* **44**(1): 215–222.

Tartakovsky, A., I. Nikiforov, M. Basseville. 2015. *Sequential Analysis: Hypothesis Testing and Changepoint Detection*. CRC Press, Boca Raton, FL.

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

Zipkin, P. H. 2000. *Foundations of Inventory Management*. McGraw Hill, New York.

Supporting Information

Additional supporting information may be found in the online version of this article:

Appendix S1: Results for Section 5.2.
Appendix S2: Detailed Results for Section 5.3.