Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Dynamic pricing and inventory management with demand learning: A bayesian approach

Jue Liu, Zhan Pang, Linggang Qi

School of Management and Engineering, Nanjing University, Nanjing, China
Krannert School of Management, Purdue University, United States
College of Business, City University of Hong Kong, Hong Kong

Abstract

We consider a retail firm selling a durable product in a volatile market where the demand is price-sensitive and random but its distribution is unknown. The firm dynamically replenishes inventory and adjusts prices over time and learns about the demand distribution. Assuming that the demand model is of the multiplicative form and unmet demand is partially backlogged, we take the empirical Bayesian approach to formulate the problem as a stochastic dynamic program. We first identify a set of regularity conditions on demand models and show that the state-dependent base-stock list-price policy is optimal. We next employ the dimensionality reduction approach to separate the scale factor that captures observed demand information from the optimal profit function, which yields a normalized dynamic program that is more tractable. We also analyze the effect of demand learning on the optimal policy using the system without Bayesian update as a benchmark. We further extend our analysis to the case with unobserved lost sales and the case with additive demand.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

Firms operating in competitive markets often face significant volatility in demand and rapidly changing market environments. The on-going public health crisis of the COVID-19 pandemic is not only threatening the lives of human beings but also disrupting global supply chains and destroying the economy, exerting sudden and unprecedented challenges on business operations. Under various social distancing measures and regulations, consumers have to change their consumption and purchase behavior to adapt to the new situation. With little experience in such an extreme crisis, firms have to respond to the changing market conditions by dynamically adjusting their operational strategies while collecting data to learn the uncertainty.

Pricing has been used as one of the most important operational levers to navigate the COVID-19 crisis (Abdelnour et al., 2020). In some sectors, from groceries to medical supplies, demand has spiked to highest-ever levels, leading to significant shortages and putting upward pressures on prices, whereas in other sectors, from air travel to durable goods like automobiles, sharp drops in demand due to the reduced shopping behavior under social distancing measures led to excess inventories and drove down prices. The volatility of prices on Amazon, both in terms of frequency and magnitude of price adjustments, has been increased significantly since the outbreak of COVID-19 (Burns, 2020). For example, the price of a five-pound bag of Nishiki medium grain rice that was at around $10 has been changed up and down, hitting a peak of $59.99 on March 21 and then dropping to $20 on April 24 (Harrison et al., 2020).

The advances of information technologies and the dramatic growth of e-commerce have enabled firms to readily adopt the dynamic pricing strategies. Last two decades have witnessed the success of dynamic pricing in both hospitality industries (e.g., airline, hotel and car rental) and retailing (Elmaghraby and Keskinocak (2003) and Talluri and van Ryzin (2004)). Manufacturers such as Ford and Dell Computer have also integrated dynamic pricing with production and distribution strategies to improve supply chain efficiency and profitability (Biller et al., 2005). Amazon developed analytical tools that help sellers use dynamic pricing on Amazon marketplace where sellers can set pricing rules based on the Buy Box and inventory availability to automate pricing decisions. It was reported that Amazon changed product prices 2.5 million times a day with an average listed product changing prices every 10 min, fifty times more often than Walmart and Best Buy (Mehta et al., 2018).
Dynamic pricing, however, is a double-edged sword which, if not managed properly, can hurt customers and ruin the reputation. When Amazon first tested its dynamic pricing strategy on DVDs in 2000, customers who found out they were charged more than others for the same products got angry, forcing Amazon to apologize and issue refunds (ABC News, 2000). More recently, Amazon's dynamic pricing algorithm drove the price of a science textbook about fly from about $100 up to $24 million (Solon, 2011). In 2011, the British largest grocery retailer, Tesco, failed to revive its sales despite spending £500 million on price cuts and as a result its CEO quit (BBC, 2012).

The volatile and rapidly changing market environments require firms to actively learn the new demand while effectively integrating their pricing and inventory management strategies with demand learning. Bayesian decision theory provides a powerful statistical decision framework to combine learning and dynamic decision making (DeGroot, 1970). In this paper, we adopt the Bayesian approach to address this problem. More specifically, we study a single-product periodic-review inventory system with the objective of maximizing expected total discounted profit over a finite horizon. Ordering costs are proportional with order sizes, while inventory holding and stockout costs all depend on the end-of-period inventory level and shortfall. For generality, we allow partial backordering which embraces both the backordering case and the lost sales case. Demands in consecutive periods are independent and sensitive to prices. The demand model is specified as a multiplicative model which is specified as a product of a deterministic price-dependent function and a random perturbation. The deterministic component captures the demand dependence on the price with its curvature determining the price elasticity of demand. It can be constructed through the price-dependent purchasing probability under the willing-to-pay (WTP) distribution. In market research, the WTP distribution is typically estimated through pricing experiments (Bodea and Ferguson, 2012). For simplicity, we assume that the price-dependent component is known. The random perturbation represents the market size of which the distribution depends on an unknown parameter. The multiplicative form of the demand model has an appealing economic interpretation: The market size captures the maximum market potential and the price-dependent purchase probability captures the market share at the price (Bodea and Ferguson, 2012). We assume that the firm forms the prior belief on the distribution of the unknown parameter and updates its belief over time using the accumulative demand data according to the Bayes rule. We first formulate this problem as a Bayesian dynamic program with a two-dimensional state space, one for inventory level and the other one for a sufficient statistic which consists of historical demand information. The sufficient statistic can be either the time series of historical sales data or the updated demand distribution based on these sales data. We identify a set of regularity conditions on the demand model which allow us to show that the expected truncated demand is jointly concave and supermodular in the order-up-to level decision and the expected demand level. These desired properties ensure the optimality of a state-dependent base-stock list-price policy.

We then employ another set of conditions from the Bayesian inventory management literature to decompose the optimal profit-to-go function into two parts: a scale factor that depends on the demands accumulated over time and a profit-to-go function of a normalized model that is independent of the market size. The optimal order-up-to level is also decomposed as a product of the scale factor and the optimal order-up-to level of the normalized model, while the optimal pricing decision depends on the scaled inventory level. This approach reduces the dimensionality of the problem, which allows us to design efficient algorithms to compute the optimal policy.

To investigate the effect of demand learning on the optimal policy, we compare the optimal policy under demand learning to that without demand learning. In particular, we show that a myopic base-stock list-price policy is optimal for the stationary system without demand learning. Using the myopic policy as a benchmark, we show that demand learning may lead to lower future prices but the effect of learning on future inventory decisions depend on the realizations of random demand. Using a numerical example, we show that demand learning may lead to a lower base-stock level and lower optimal prices for high inventory levels in the current decision period. We also discuss the effect of demand variation on the optimal policy under demand learning.

We further discuss two extensions. The first extension is to generalize the model with complete demand information to the model with unobserved lost sales (i.e., censored demand). We show that the dimensionality reduction approach also applies. The second extension is to address the case with additive demand. We show that the optimality of base-stock list-price policy remains true under the regularity conditions but the dimensionality reduction approach no longer applies.

Our contributions are threefold. First, we bridge the inventory-pricing literature and the Bayesian inventory management literature to integrate the dynamic pricing and inventory control with demand learning in the Bayesian framework. Second, our model allows partial backordering which includes the complete backordering and lost-sales cases as special cases. There are limited results on the multiplicative demand in the lost-sales case in the literature. We add to this literature by identifying a set of demand regularity conditions that ensure the concavity of the expected sales revenue and hence the optimality of the state-dependent base-stock list-price policy. Third, we show that the dimensionality reduction approach applies to the Bayesian inventory-pricing model under the multiplicative demand, which simplifies the algorithm design and sheds new light into the dynamic pricing and inventory management problem. Furthermore, we extend the dimensionality reduction analysis to the case with unobserved lost sales.

The remainder of this paper is organized as follows. Section 2 reviews the related literature. Section 3 formulates the problem as a Bayesian dynamic program. Section 4 identifies the regularity conditions and characterizes the structure of the optimal policy. Section 5 introduces the dimensionality reduction approach. Section 6 discusses the effect of learning. Section 7 discusses two extensions: unobserved lost sales and additive demand. Section 8 summarizes the conclusions. All technical proofs are in the appendix.

2. Related literature

Our paper is related to two streams of dynamic inventory management literature: dynamic inventory-pricing problem and Bayesian inventory management.

There is a growing body of literature on dynamic pricing and inventory control problems; see Chen et al. (2012) for a review of recent developments. Most of inventor-pricing control models make the assumption that the demand distribution is known with certainty, which often leads to simple forms of policy structure that is relatively easy to calculate. For example, in a periodic-review backlog model without fixed order cost, Federgruen and Heching (1999) show that the base-stock list-price policy is optimal. That is, there is a (state-independent) base stock level and list price combination such that if the inventory level is below the base stock level, then an order is placed to increase the inventory level to the base stock level and the list price is charged; otherwise, nothing is ordered and an (inventory sensitive) price discount is offered. Their

J. Liu et al. / Computers and Operations Research 124 (2020) 105078
model is generalized by Chen and Simchi-Levi (2004) to the case with fixed cost and backorders, by Song et al. (2009) and Pang, 2011 to the case with fixed order cost and lost sales, by Pang et al. (2012) to the case with positive lead times, and by Chen et al. (2014) to perishable inventory systems for both backlog and lost-sales cases. We generalize this literature by incorporating demand learning with inventory-pricing decisions, accommodating both lost-sales and backlogging cases. It is notable that to our knowledge the only paper in this literature that addresses multiplicative demand and lost-sales is Song et al. (2009). Their analysis, however, is restricted to stationary settings while our model is nonstationary due to the nature of information update in learning.

Bayesian dynamic inventory models incorporate demand learning with the inventory decisions in the Bayesian framework. A Bayesian dynamic inventory problem can usually be modeled as a dynamic program with a multi-dimensional state space (see, e.g., Scarf, 1959). However, due to the curse of dimensionality, the computation of such a dynamic program is usually very difficult if not intractable. Assuming demand distribution is in the exponential family, Scarf (1960) introduces a state space reduction approach and shows that the solution of the Bayesian model with unknown demand distribution parameters can be easily obtained by solving a non-Bayesian model with known demand distribution parameters. Azoury (1985) extends Scarf’s results to the class of demand distributions including uniform, Weibull, and Gamma distributions. Miller (1986) shows that the reduction approach can be used in exponentially smoothed forecasts. Lovejoy (1990) provides a unified framework within which both the Bayesian model and exponential smooth model fall, and shows that myopic policies are optimal under some conditions. The key idea of the state reduction approach is to identify a class of predictive demand intensities which are scalable (see, e.g., §10 of Porteus (2002)) and then to show that the value functions can be expressed by a product of a value function of a normalized problem and a scaling function of the sufficient statistics. A common feature of all their models is that the demand is exogenous and only the inventory decision is considered. We generalize this literature by considering the joint pricing and inventory decisions.

Our paper attempts to generalize the Bayesian inventory models by incorporating pricing decisions, bridging the two streams of literature. There are a few papers along the same line. Subrahmanyan and Shoemaker (1996) propose a Bayesian dynamic program formulation to address the joint inventory and pricing decision problem in general retail situations but they do not provide structural analysis. Petruzzi and Dada (2002) study a pricing decision problem in general retail situations but they do not provide structural analysis. Petruzzi and Dada (2002) study a pricing decision problem in general retail situations but they do not provide structural analysis. Petruzzi and Dada (2002) study a pricing decision problem in general retail situations but they do not provide structural analysis. Petruzzi and Dada (2002) study a pricing decision problem in general retail situations but they do not provide structural analysis.

3. Problem formulation

Consider a firm selling a nonperishable product in a new market with price-sensitive random demand. Without knowing the demand distributions, the firm aims to integrate demand learning with the pricing and replenishment decisions to maximize the expected total discounted profit over a finite planning horizon of $T$ periods, indexed by $t = 1, \ldots, T$, with a discount factor $\alpha \in (0, 1]$. The sequence of events is as follows. At the beginning of each period, a belief on the demand distribution is drawn based on the past demand information, and a simultaneous decision is made regarding the size of a new replenishment order (if any) and the selling price. For each ordering decision, a variable ordering cost $c$ is incurred for each unit of inventory and the replenishment lead-time is zero. The objective of the system is to maximize the total discounted profit over the $T$ periods with a discount factor $\alpha \in (0, 1]$.

3.1. Demand model

Let $p_t \in \mathcal{P}$ be the retail price in period $t$ where $\mathcal{P} := [p_l, p_u]$ with $p_l \geq c$ where $c$ is the unit ordering cost. The demand function is assumed to be of the multiplicative form $D_t(p) = d_t(p)e_t$ where $d_t(p)$ is a deterministic real function of $p$ and $e_t$ is a continuous nonnegative random perturbation factor with an unknown distribution. The random perturbation $e_t$ captures the magnitude of the market size and is referred to as the demand scale factor. Following the Bayesian inventory literature (see, e.g., Azoury, 1985; Lariviere and Porteus, 1999; Petruzzi and Dada, 2002 and the references therein), we assume that the random perturbations are independently and identically distributed over time. The price-dependent function $d_t(p)$ models the customers’ sensitivity to price, and it is often expressed as the purchase probability at price $p$, $Pr(W_t > p)$, where $W_t$ is customer willing-to-pay (WTP). The WTP distribution is typically learned through econometric analysis or pricing experiments (Bodea and Ferguson, 2012). For simplicity, we assume that the WTP distribution, or equivalently, $d_t(p)$, is known and we will focus on learning the market size in this paper. Note that $d_t(p)$ can be time varying, which captures the time dependence of the demand model.

The multiplicative random perturbation is commonly adopted in the inventory-pricing literature (see, e.g., Karlin et al., 1962; Petruzzi and Dada, 1999; Elmaghraby and Keskinocak, 2003; Aydin and Porteus, 2008; Granot and Yin, 2008; Song et al., 2009 and references therein) as well as the revenue management literature (see, e.g., Gallego and van Ryzin, 1994; Talluri and van Ryzin, 2004; Araman and Caldentey, 2009 and references therein). The multiplicative randomness implies that the coefficient of variation of the demand is constant (Aydin and Porteus, 2008). As shown in Bodea and Ferguson (2012), a convenient way to empirically estimate the multiplicative model is to regress the natural logarithm of the sales against the price or its transformation, which ensures the estimated demand model to be nonnegative. Another common demand model in the literature is of the additive form $D_t(p) = d_t(p) + e_t$ (see, e.g., Petruzzi and Dada, 1999; Chen and Simchi-Levi, 2004; Pang, 2011). As to be shown later, the multiplicative demand provides scalability for the system with a family of conjugate distribution pairs, which allows us to provide a sharp characterization of the optimal policy whereas the additive
demand does not have the scalability. Hence, we restrict our analysis to multiplicative demand in our main analysis and discuss the case with additive demand in §7.2.

The price elasticity of the multiplicative demand is defined as 
\[ e_t(p) = -\frac{d p}{d r} \ln |p_t| / |p_r| \]
which can also be equivalently expressed as the derivative of the natural logarithm of the demand with respect to the natural logarithm of the price, i.e.,
\[ e_t(p) = -\frac{d \ln p_t}{d \ln p_r} \ln |p_t| / |p_r| \]. In particular, if the \( d_t(p) \) is of the exponential form, i.e., \( d_t(p) = e^{a-b p} \) with \( a > 0 \) and \( b > 0 \), then \( e_t(p) = b \), i.e., the price elasticity is a constant, and such a demand model is also called iso-elastic or constant-elasticity demand which is commonly used in practice to empirically estimate the average demand elasticity through a linear regression for \( \ln(D_t) \) against \( \ln(p_t) \) (Talluri and van Ryzin, 2004; Bodea and Ferguson, 2012).

Assume that \( d_t(p) \) is strictly decreasing in \( p \), which implies choosing a price \( p \in \mathcal{P} \), is equivalent to choosing a value \( d \in \mathcal{D}_t := [\mathcal{d}_t, \mathcal{d}_t] \) where \( \mathcal{d}_r = \mathcal{d}_r(\mathcal{p}_{r}) \) and \( \mathcal{d}_l = \mathcal{d}_r(\mathcal{p}_{r}) \) with \( \mathcal{d}_r > 0 \). Note that although in theory the price can be set to infinity at which the demand reduces to zero, in practice firms rarely take such an extreme action in retailing. Note that if \( E[\xi_t] = 1 \) then \( d_t(p) \) is equal to the expected demand level under the price \( p \). For convenience, we call \( d_t(p) \) the expected demand level in the following analysis. Letting \( P_t(d) \) be the inverse demand function, the price elasticity of demand is re-expressed in terms of \( d \) as
\[ e_t(d) = -\frac{\pi_t(d)}{\pi_t(d_t)} \ln |\pi_t(d_t)/\pi_t(d)| \]

3.2. Bayesian dynamic programming formulation

To learn the distribution of the random perturbation, the firm adopts a Bayesian framework (see §10 of Porteus (2002) and the references therein) to assume that the probability distribution of \( \xi_t \) is of the form \( \psi(\xi_t) \) with an unknown parameter \( \alpha_t \). Let \( \xi_t \) denote the sufficient statistic for the sample \( \xi_1, \xi_2, \ldots, \xi_t \) of the past demand observations at the beginning of period \( t \). The dynamics of the \( \xi_t \) are denoted by \( \mathcal{s}_{t-1} = \xi_t \). Let \( g(\omega|\xi_t) \) represent the prior probability density function of the unknown parameter \( \alpha_t \), which represents the belief about the underlying distribution of the \( \xi_t \) for any given statistic \( s \) in period \( t \). We assume that \( g(\omega|\xi_t) \) is chosen from a conjugate family of distributions such that the posterior distribution of \( \alpha_t \) (after updating using Bayes’ rule based on the sufficient statistic and the new demand observation) has the same distributional form as the prior. Then, given a statistic \( s \), the predictive demand density of period \( t \) is given by
\[ \phi_t(s|\xi_t) := \int_0^{\infty} \psi(\xi_t) g(\omega|\xi_t) d\omega \]

By Bayes’ rule, conditioned on the values of \( s \) and \( \alpha \), the posterior density on \( \alpha \)
\[ g(\omega|s, \alpha) = \frac{\psi(\xi_t) g(\omega|\xi_t) d\omega}{\int_0^{\infty} \psi(\xi_t) g(\omega|\xi_t) d\omega} \]

The conjugate prior assumption implies that \( g(\omega|s, \alpha) \) has the same functional form of \( g(\omega|\xi_t) \); see Porteus (2002, Chapter 10) for more discussions on the family of conjugate prior distributions.

Let \( x_t \) be the inventory level at the beginning of period \( t \), just before placing an order, and \( y_t \) be the inventory level after placing an order (i.e., order-up-to level) with \( y_t \leq x_t \). For each positive order, only the variable ordering cost \( c \) is incurred for each unit of inventory and the replenishment leadtime is zero. At the beginning of each period \( t \), the firm reviews its inventory level \( x_t \) and statistic \( s_t \) and then decides the order quantity \( y_t = x_t^+ \) and the price level \( p \in \mathcal{P} \) (or equivalently, the expected demand level \( d \in \mathcal{D}_r \)). The demand \( D_t \), or equivalently, \( x_t \), is realized by the end of this period. The excess inventory is carried over to the next period while the unmet demand may be partially backlogged or lost. The dynamics of inventory levels are
\[ x_{t+1} = \delta(y_t - D_t) + \lambda(y_t - D_t) + (1 - \lambda)(y_t - D_t)^+ = (y_t - D_t) + (1 - \lambda)(D_t - y_t)^+ \], \( \lambda \in [0, 1] \).

That is, if \( D_t > y_t \), a fraction \( \lambda \) of the unmet demand is backlogged while the others are lost. This treatment encompasses both the full backordering case (\( \lambda = 1 \)) and lost-sales case (\( \lambda = 0 \)). In particular, the lost-sales case is more common in retailing of consumer products. Clearly, \( \alpha \) is a piece-wise increasing convex function. Note that \( \delta(x) = g(x) \) for any real value \( x \) and constant \( k > 0 \), i.e., \( \alpha \) is homogenous of degree one.

Let \( \tilde{x}(x) := h^+ x^+ + h^- x^- \) denote the inventory holding cost, where \( h^+ \) and \( h^- \) represent the unit holding and shortage costs, respectively. In addition, we assume that \( h^+ > (1 - x/c) \) and \( h^- \geq x/c \), which implies that the shortage (backlog) cost is greater than the cost saving from delaying the purchasing to the next period, which removes the motivation to intentionally delay the purchase in face of the shortage. Hence, it suffices to require that the order-up-to level \( y_t \) is nonnegative.

The sales of each period, denoted by \( \gamma(D_t, y_t) \), is constrained (or truncated) by the inventory order-up-to level \( y_t \). We assume that all the lost sales are observable. The case with unobservable lost sales, i.e., censored demand, will be discussed later. Then we have
\[ \gamma(D_t, y_t) := \min(D_t, y_t) + \max(D_t - y_t, 0) = (1 - \lambda) \min(D_t, y_t) + \lambda D_t \]
where \( \min(D_t, y_t) \) is the sales truncated by the inventory level \( y_t \). It is easy to verify that \( \gamma(D_t, y_t) \) is also homogenous of degree one with respect to \( (D_t, y_t) \).

The problem can be formulated into a dynamic program with two state variables, \( (x, s) \), where \( x \) is the inventory level before ordering and \( s \) is the sufficient statistic up to current period. Let \( \bar{f}_t(x, s) \) denote the maximum expected present value of the profit over periods \( t \) through \( T \) starting with inventory level \( x \) and with a sufficient statistic \( s \) at the beginning of period \( t \). The Bellman equations, for \( t = 1, \ldots, T \), are:
\begin{align*}
\bar{f}_t(x, s) &= c x + \max \left\{ \Pi_t(y, d) := \pi_t(y, d) \right\}_{y, d \in \mathcal{D}_t} + \int_0^{\infty} \bar{f}_{t+1}(d - y|s; \epsilon) \phi_t(\epsilon|\pi_t) d\epsilon, \\
\pi_t(y, d) &= \int_0^{\infty} \Pi_t(y, d) d\epsilon \\
\Pi_t(y, d) &= \max \left\{ \Pi_t(y, d) := \pi_t(y, d) \right\}_{y, d \in \mathcal{D}_t} + \int_0^{\infty} \bar{f}_{t+1}(d - y|s; \epsilon) \phi_t(\epsilon|\pi_t) d\epsilon, \\
\end{align*}
where
\begin{align*}
\bar{f}_{T+1}(x, s) &= c x \quad \text{for all } t \quad \text{all } \Pi_t(y, d) = \pi_t(y, d) = 0 = x \\
\Pi_t(y, d) &= \max \left\{ \Pi_t(y, d) := \pi_t(y, d) \right\}_{y, d \in \mathcal{D}_t} + \int_0^{\infty} \bar{f}_{t+1}(d - y|s; \epsilon) \phi_t(\epsilon|\pi_t) d\epsilon, \\
\bar{f}_{T+1}(x, s) &= c x \quad \text{for all } t \quad \text{all } \Pi_t(y, d) = \pi_t(y, d) = 0 = x
\end{align*}

where
\[ \pi_t(y, d) = \int_0^{\infty} \Pi_t(y, d) [1 - \lambda - \lambda(y_t - d)] d\epsilon \phi_t(\epsilon|\pi_t) d\epsilon, \\
\Pi_t(y, d) = \max \left\{ \Pi_t(y, d) := \pi_t(y, d) \right\}_{y, d \in \mathcal{D}_t} + \int_0^{\infty} \bar{f}_{t+1}(d - y|s; \epsilon) \phi_t(\epsilon|\pi_t) d\epsilon, \\
\pi_t(y, d) = \int_0^{\infty} \Pi_t(y, d) [1 - \lambda - \lambda(y_t - d)] d\epsilon \phi_t(\epsilon|\pi_t) d\epsilon, \\
\Pi_t(y, d) = \max \left\{ \Pi_t(y, d) := \pi_t(y, d) \right\}_{y, d \in \mathcal{D}_t} + \int_0^{\infty} \bar{f}_{t+1}(d - y|s; \epsilon) \phi_t(\epsilon|\pi_t) d\epsilon, \\
\bar{f}_{T+1}(x, s) &= c x \quad \text{for all } t \quad \text{all } \Pi_t(y, d) = \pi_t(y, d) = 0 = x.
\]
(lexicographically largest) maximizer of $\Pi_t(y, d(s))$ and $p_t^*(s) = P_t(d_t^*(s))$.

4. Optimal policies

This section characterizes the structure of the optimal policy. It knows that the base-stock list-price policy is optimal for the back-ordering case (Federgruen and Heching, 1999; Chen and Simchi-Levi, 2004) and for the lost-sales case with additive demand (Pang, 2011). A sufficient condition for the optimality of the base-stock list-price policy is that the expected revenue function $R_t(y, d(s))$ is jointly concave in $(y, d)$ and that $f_t(x(s))$ is a non-increasing and concave function of $x$. Pang (2011) and Chen et al. (2014) identify a set of regularity conditions on the demand model under which $R_t(y, d(s))$ is jointly concave in $(y, d)$ when the demand is additive. In what follows, we first follow the approach of Pang (2011) and Chen et al. (2014) to identify a set of regularity conditions on the demand model and then characterize the optimal policy.

4.1. Regularity conditions

We first identify conditions under which the expected revenue $P_t(d) j d$ is concave in $d$. The first condition, proposed by Chen and Simchi-Levi (2004), is that $P_t(d) j d$ is concave in $d$:

\[(A1) \text{ For } t = 1, \ldots, T, (P_t(d) j d)^n = P_t^2(d) j d + 2P_t h(d) \leq 0 \text{ for all } d \in \mathcal{D}_t.\]

The economic interpretation of condition (A1) is that the marginal revenue contribution of the demand decreases in the demand level. The higher the demand level the less the marginal revenue contribution of increasing the expected demand level (i.e., reducing the price). An alternative condition, as identified by Pang (2011), is:

\[(A2) \text{ For } t = 1, \ldots, T, P_t^2(d) j d + P_t h(d) \leq 0 \text{ for all } d \in \mathcal{D}_t.\]

Clearly, (A2) implies that the expected revenue $P_t(d) j d$ is strictly concave in $d$, which is stronger than (A1), since $P_t(d)$ is strictly decreasing in $d$.

We next verify conditions (A1) and (A2), respectively, for some common demand models. The results are summarized in Table 1. See Pang (2011) and Chen et al. (2014) for a similar analysis for additive demand.

Table 1 shows that linear, log, logit and exponential models above all satisfy (A1) and (A2). For the iso-elastic demand, (A1) is satisfied if $b > 1$ while (A2) is always violated. Note that $d_t(p)$ and $d_t(p)/p$ are both convex in $p$ for the iso-elastic demand with $b > 1$ and hence the analysis of Federgruen and Heching (1999) and Lovejoy (2011) who require that $d_t(p)$ is concave in $p$ does not apply to the iso-elastic demand. For the power model, (A1) is satisfied if $b < 1$ and (A2) is satisfied if $d_t > 1/b$ and $d < 1$.

In the backlog case, the demand condition (A1) or (A2) ensures the optimality of base-stock list-price policy (see, e.g., Chen and Simchi-Levi, 2004). However, in the lost-sales case, (A1) or (A2) alone cannot guarantee the joint concavity of the expected truncated revenue $P_t(d) [\min(d, c), y]$ in terms of $d$ and $y$. For this reason, Federgruen and Heching (1999) argue that the base-stock list-price policy may fail to be optimal in the lost-sales setting. For the lost-sales case with additive demand $(D_t = d_t + c)$, Pang (2011) shows that $P_t(d) [\min(d, c), y]$ is jointly concave in $(y, d)$ if (A2) holds and the price elasticity of the lost-sales probability, defined as $\varepsilon_t(p) = -\frac{\partial \ln b_1^{d_t(p)}}{\partial \ln b}$, is greater than one. Note that the notation of the price elasticity of the lost-sales probability was first proposed by Lovejoy (2011) in a newsvendor setting to identify the conditions under which the expected profit is jointly concave in price and ordering decisions. Pang (2011) extends their analysis to dynamic models with additive demand and fixed ordering cost and identify conditions under which an $(s, S, p)$ policy is optimal.

We next extend the analysis of Pang (2011) to the case with multiplicative demand in the Bayesian learning model. With the multiplicative demand $D_t = d_t(p) e_t$, the probability of lost sales at an inventory level $y$ and expected demand level $d$ is $\Phi_t(y, d(s))$. Then the price elasticity of lost-sales probability (in terms of $p$) is

\[e_t^*(p, y) = -\frac{p}{\Phi_t(y, d_t(p)/p)} \frac{d}{dp} \Phi_t(y, d_t(p)/p) = p_t(y, d_t(p)/p) \varepsilon_t(p),\]

where $p_t(x(s)) = \frac{\lambda x(s)}{\lambda x(s) + \beta}$ is the generalized hazard rate. Using the inverse function $P_t(d)$, the price elasticity of lost-sales probability can be expressed in terms of $d$ as

\[e_t^*(d, y) = -\frac{P_t(d)}{\Phi_t(y, d)} \frac{\partial}{\partial d} \Phi_t(y, d) = p_t(y, d) \varepsilon_t(d).\]

Note that $p_t(y, d)/d < 0$ for $y = 0$, which implies that the price elasticity of lost-sales probability does not have a positive lower bound for any value $y$. That is, the analyses in Pang (2011) and Chen et al. (2014) do not apply immediately. To impose a uniform lower bound on the price elasticity of lost-sales probability as that of Pang (2011) and Chen et al. (2014), we need to identify a lower bound of $y_t(s)$ under the optimal policy, which implies that it suffices to restrict the order-up-to level decision to the value above such a lower bound.

Observe that the single-period expected profit $\pi_t(y, d(s))$ is concave in $y$ for any $d$ and $s$. Taking the first derivative of $\pi_t(y, d(s))$ with respect to $y$ yields

\[
\frac{\partial}{\partial y} \pi_t(y, d(s)) = \left[(1 - \lambda) (P_t(d) - xc) + h + h \right] \Phi_t(y, d(s))
- \left[(1 - \lambda) c + h \right].
\]

| Demand model | $d_t(p)$ | $P_t(d)$ | $P^*(d) j d + P^*(d)$ | $P^*(d) j d + 2P^*(d)$ |
|--------------|-----------|-----------|-------------------------|-------------------------|
| linear       | $a - bp$  | $(a - d)/b$ | $-1/b$                  | $-2/b$                  |
| log          | $\ln(a - bp)$ | $\frac{a - d}{a - b}$ | $-\frac{a - d}{a - b}$ | $-\frac{a - d}{a - b}$ |
| logit        | $\frac{a - d}{a}$ | $\frac{a - d}{a}$ | $\frac{a - d}{a}$ | $\frac{a - d}{a}$ |
| exponential  | $a^{1 - b}$ | $\frac{a}{a - 1}$ | $\frac{a}{a - 1}$ | $\frac{a}{a - 1}$ |
| iso-elastic  | $\frac{a}{a - d}$ | $\frac{a}{a - d}$ | $\frac{a}{a - d}$ | $\frac{a}{a - d}$ |
| power        | $\frac{a}{a - d}$ | $\frac{a}{a - d}$ | $\frac{a}{a - d}$ | $\frac{a}{a - d}$ |
Then the maximizer of $\pi_t(y, d[s])$ is $\hat{y}_t(d[s]) = dZ_t(d[s])$ where

$$Z_t(d[s]) = \Phi_1^{-1}\left( (1 - x)c + h - (1 - \lambda)(P_t(d) - xc) + h + h^\perp \right) |s\right).$$

The monotonicity of $P_t(d)$ implies that $Z_t(d[s])$ is strictly decreasing in $d \in D_t$ for any $\lambda < 1$.

Define $u_t = \min_{s > 0} y_t$ where $y_t$ is a uniform lower bound of $y_t(d[s])$ for any $d$ and $s$, defined as $y_t = \inf\{y_t(d[s])|s, d \in D_t\}$. Then $u_t < u_{t+1}$ for all $t$. The following lemma shows that $u_t$ is indeed a lower bound for the optimal order-up-to level $y_t(s)$.

**Lemma 1** (Bound). For all $t$ and $s$, the optimal order-up-to level $y_t(s)$ is greater than $u_t$.

Lemma 1 implies that it suffices to restrict the optimal order-up-to level decision to values above $u_t$. We are now ready to introduce the additional demand regularity conditions:\[B1\] $p_t(y/d(s)) \geq 1/\varepsilon_t(d)$ for all $d \in D_t$ and $y \geq u_t$,\[B2\] $p_t(y/d(s)) \geq 2/\varepsilon_t(d)$ for all $d \in D_t$ and $y \geq u_t$.

Conditions (B1) and (B2) imply that the price elasticity of the lost-sales probability is greater than one and two, respectively. Clearly, (B2) is stronger than (B1) as the former requires a greater price elasticity of the lost-sales probability. Since $u_t$ is a uniform lower bound of the optimal order-up-to level $y_t(s)$, conditions (B1) and (B2) apply to distributions with domains containing zero (e.g., Gamma distribution or Weibull distribution).

Note that

$$\frac{\partial y_t(d[s])}{\partial d} = -\frac{\Phi_1(y_t(d[s])|s)/\Phi_1(y_t(d[s])|s)}{\Phi_1(y_t(d[s])|s)} p_t(y_t(d[s])|s) - \frac{1}{\varepsilon_t(d)} p_t(d) + (h + h^\perp - (1 - \lambda)xc)/(1 - \lambda)$$

Since $h \geq xc$, we have $\frac{\partial y_t(d[s])}{\partial d} > 0$ if (B1) or (B2) holds, which implies that $y_t = \inf\{y_t(d[s])|s\}$.

**Remark 1.** In the newsvendor setting, Kocabiyikoglu and Popescu (2011) show that if $p_t(y/d(d[p]))/\varepsilon_t(p) \geq 1$, then the expected profit function is jointly concave and submodular in $(y, p)$. However, they do not impose any constraint on the value of $y$. As discussed above, $p_t(y/d(p)|s)/\varepsilon_t(p)$ approaches zero when $y$ approaches zero, which implies that when $y$ is sufficiently small, the condition $p_t(y/d(p)|s)/\varepsilon_t(p) \geq 1$ is violated. Our analysis fixes this problem by identifying the lower bound of the optimal order-up-to level.

4.2. Optimality of state-dependent base-stock list-price policy

We are now ready to employ the regularity conditions to characterize the structure of the optimal policy. The following theorem extends the analyses of Federgruen and Heching (1999) and Pang (2011) to the Bayesian formulation with multiplicative demand and lost sales appears to be much more complicated than that of Federgruen and Heching (1999) and Pang (2011).

Regularity conditions (i) and (ii) guarantee the joint concavity and supermodularity of the expected single-period profit $\pi_t(y, d)$. These properties in turn ensure the optimality of the base-stock list-price policies and the monotonicity of the pricing decision. Note that (A1) is weaker than (A2) but (B2) is stronger than (B1). So (i) is neither more general nor more restrictive than (ii). The dependence on the statistic $s$ captures the effect of demand learning from historical data. Intuitively, everything else being the same, the greater the demands realized in the previous periods, the higher the expected demand in the future and therefore the more to order.

In particular, when $\lambda = 1$, the model reduces to the fully backordering case. Following Chen and Simchi-Levi (2004), it suffices to impose the condition (A1) to ensure the optimality of the base-stock and list-price policy.

**Theorem 2** (Backordering case). Suppose (A1) holds for all $t$ and $s > 0$. Then a state-dependent base-stock list-price policy is optimal.

**Remark 2.** Our analysis uses the expected demand level and the order-up-to level as decision variables whereas Federgruen and Heching (1999) treat the price decision variable directly. Their analysis requires two conditions (a) $d_t(p)$ is concave in $p$ and (b) the expected holding/shortage cost $E[\varphi(y - D_t(p))]$ is jointly concave in $(p, y)$. Clearly, these two conditions are satisfied when $d_t(p)$ is linear. However, it is easy to verify that $d_t(p)$ is convex in $p$ for the exponential and iso-elastic demand models. More importantly, it remains a challenge to show whether (b) holds for more general nonlinear demand models. Therefore, it is more convenient to work on the inverse demand function instead of treating the price variable directly.

**Remark 3.** Our model is restricted to the class of conjugate prior distributions under which the learning process is to update the parameters of the posterior distributions, which provides strong tractability in both structural analysis and computation. But our analysis can be readily extended to address non-conjugate priors under which the posterior distributions are no longer of the same functional form of the prior distributions. Then the optimality of the state-dependent base-stock list-price policy in Theorem 1 remains true if the statistic state is replaced by the updated probability density function of $a_t$.

**Remark 4.** The expected single-period profit in (2) can be alternatively expressed

$$\pi_t(y, d[s]) = \tilde{R}_t(y, d[s]) - (1 - x)cy - \int_0^{\infty} \tilde{\varphi}_t(y - de|\varepsilon_t|s|de),$$

with

$$\tilde{R}_t(y, d[s]) = \int_0^\infty (P_t(d - xc) - \min(de, y) + \lambda de|\varepsilon_t|s|de).$$

Treating $P_t(d) = P_t(d - xc)$ as a modified inverse demand function, then the results of Theorem 1 remain true if we replace $e_t(d)$ by $-\frac{\partial e_t}{\partial P_t}(d)$, see Pang (2011) for a similar treatment.

5. Dimensionality reduction

Computing the optimal policies according to the Bayesian dynamic program (2) involves not only the inventory level but also the information state which could be multiple dimensions, which may require an extensive computational effort. Scarf (1960) and Azoury (1985) identify a class of scalable demand distributions and develop conditions under which the dimensionality of the problem is reduced: The solution of the original problem can be
obtained from the solution of a much simpler normalized problem that is independent of the scale (size) of the market. Following Azoury (1985), we impose two conditions.

(C1) There exists a positive real-valued function \( q_t(s) \) of the sufficient statistic \( s \) which is a scale parameter for \( \phi_t(\epsilon_t|s) \) such that

\[
\phi_t(\epsilon_t|s) = (1/q_t(s))\phi_t(\epsilon_t/q_t(s)),
\]

where \( \phi_t(\cdot) \) is a known probability density function for period \( t \) which is independent of \( s \), and

(C2) the function \( q_{t+1}(\cdot) \) satisfies

\[
q_{t+1}(s \circ t) = q_t(s)U_{t+1}(\epsilon_t/q_t(s)),
\]

where \( U_{t+1} \) is a known continuous function and satisfies

\[
\int_{\epsilon_t} q_{t+1}(\epsilon|\epsilon_t)\phi_t(\epsilon|\epsilon_t)d\epsilon_t < \infty.
\]

The following examples satisfy conditions (C1) and (C2).

**Example 1 (Gamma-Gamma Model).** Suppose \( \epsilon_t \) follows a Gamma distribution with \( \psi(\epsilon_t|\alpha_t) = \alpha_t e^{-\alpha_t \epsilon_t}/\Gamma(\alpha_t) \), where \( \alpha_t > 0 \) is the shape parameter and \( \psi(\cdot) \) is the unknown positive scale parameter. The Gamma distribution reduces to exponential distribution with rate \( \alpha_t \) if the scale parameter \( \alpha_t = 1 \). The mean and variance of the Gamma distribution are \( \alpha_t/\psi(\cdot) \) and \( \alpha_t^2/\psi(\cdot) \) respectively, so the coefficient of variation is \( \sqrt{\alpha_t} \). This implies that, given any initial sufficient statistic \( s_t = (a_t; \Lambda_t) \). Correspondingly, the mean and variance of the prior distribution are \( a_t/\Lambda_t \) and \( a_t^2/\Lambda_t^2 \) respectively. For \( \epsilon_t > 0 \) the predictive density is given by

\[
\phi_t(\epsilon_t|s_t) = \frac{\Gamma(\alpha_t+k_t)}{\Gamma(\alpha_t)}\frac{\alpha_t k_t e^{-\alpha_t \epsilon_t}}{(\alpha_t+k_t)^{\alpha_t+k_t}}
\]

where \( s_{t+1} = (a_{t+1}; \Lambda_t) \). We have \( a_t = a_t + (t-1)k_t \) and \( \Lambda_t = \Lambda_t + \sum_{t=1}^{t-1} \epsilon_t \) for all \( t > 1 \). Here, \( \alpha_t \) and \( \Lambda_t \) are the shape and scale parameters of the predictive distribution respectively. In addition, (C1) implies that \( \phi_t(\epsilon_t|s) = \phi_t(\epsilon_t|a_t, 1) \). In particular, for \( k = 1 \) and \( \alpha_t > 1 \), we have \( E(\epsilon_t) = \frac{\alpha_t k^2}{\alpha_t+k^2} \). As shown by Scarf (1960), the Gamma-Gamma model is scalable with \( q_t(s_t) = \Lambda_t \) and \( U_{t+1}(x) = (1 + x)^{\alpha_t} \).

Lariviere and Porteus (1999) note that the predictive demand distribution is (first-order) stochastically decreasing in the shape parameter \( \alpha_t \) and stochastically increasing in the scale parameter \( \Lambda_t \). The posterior density functions indicate that the shape parameter increases by one after each observation and can be illustrated as the accumulate number of observations. Since the coefficient of variation of both Weibull and Gamma distributions is strictly decreasing in the shape parameter, the relative uncertainty about the scale parameter decreases as the number of observations increases with the limit of the exact value of the scale parameter. The scale parameter of the prior also increases each period by the market size observation raised to 4th power for the Weibull prior and by the market size observation for the Gamma prior. For the Gamma-Gamma model, the limit for the mean of the prior is then \( \lim_{t \to \infty} a_t/\Lambda_t = 1/\lim_{t \to \infty} \sum_{t=1}^{t-1} \epsilon_t/t \) which is equal to true scale parameter (with probability one) of the demand distribution. Note that \( \sum_{t=1}^{t-1} \epsilon_t/t \) is an unbiased estimator of the prior mean.

We are now ready to apply the dimensionality reduction approach by applying conditions (C1) and (C2). The key idea is to first solve a normalized problem with nonstationary known demand distributions in which the scale parameter is set to one and then obtain the solution to the original problem by incorporating the scale parameter into the solution of the normalized problem.

Let \( v_{t+1}(x) = 0 \). Define functions \( \pi_t(r, d) \) recursively, for \( t = 1, \ldots, T \):

\[
v_t(x) = \max_{y \geq d \geq 0} \{ \Pi_t(y, d) := \pi_t(y, d) + \alpha \int_0^\infty v_{t+1}(\delta_t(y - d))U_{t+1}(\xi)\phi_t(\xi)d\xi \},
\]

where

\[
\pi_t(y, d) = R_t(y, d) - \alpha x \int_0^\infty \sigma_t(y, \phi_t(\xi))d\xi - \alpha x \int_0^\infty \phi_t(\xi)d\xi \phi_t(y, d)
\]

The above equalities define a normalized system where the scale parameter is always one. The underlying distributions of the normalized model can also be interpreted as the distribution of the unknown parameter of the prior is always updated as if the observation of the market size \( \epsilon_t \) were always zero and hence the scale parameter, which started from one at the beginning of the planning horizon, remains unchanged over time though the number of observations was increased.

Let \( y_t(x) \) and \( d_t(x) \) (and \( \pi_t(x) = P_t(d_t(x)) \) be the optimal solutions of the normalized system. Let \( y_t \) and \( d_t \) be the corresponding global maximizers of \( \Pi_t(y, d) \) and \( \pi_t \).

The following theorem shows the relationships between \( f_t(\cdot|s) \) and \( v_t(\cdot) \).

**Theorem 3 (Dimensionality Reduction).** If conditions (C1) and (C2) are satisfied, then for all \( t, x \) and \( s \)

\[
(a) f_t(x|s) = q_t(s) v_t(x/q_t(s)),
(b) y_t'(x|s) = q_t(s) y_t'(x/q_t(s)) = q_t(s) \max(y_t, x/q_t(s)) \text{ and } p_t'(x|s) = \min(p_t, p_t'(x/q_t(s))).
\]

In particular, \( y_t'(x) = q_t(s)y_t \) and \( p_t'(x) = p_t \).

**Theorem 3** shows that the optimal value functions to the original problem are equal to the product of the normalized profit functions and the scale function \( q_t(s) \). In other words, the value function of the original problem is decomposed into two factors where the scale function \( q_t(s) \) represents the scale (size) of the market and the normalized profit function \( v_t \) allows us to focus on the normalized market demand. Such a separation leads to a much simpler model.
The firm only needs to derive the optimal replenishment and pricing decisions based on the normalized model that is independent of the scale (size) of the market and then scales the solution based on the updated estimate of the market size.

More specifically, under the optimal policy, the optimal order-up-to level of the original problem, \( y'(x|s) \), can be derived by scaling the optimal order-up-to level of the normalized problem, \( y'_1(x) \), with the scale rate \( q(s) \), while the optimal pricing decision of the original problem is equal to the pricing decision of the normalized problem over the normalized stock level, \( x/q(s) \). In particular, \( q_0(s) = \Lambda_1 \) for the Gamma-Gamma Model and \( q(s) = \Lambda_1^{1/k} \) for the Weibull-Gamma model, where \( \Lambda_1 \) represents the accumulate sales over time. Hence, the more demands observed over time, the greater the value of the scale rate \( q(s) \) and therefore the greater the base stock level \( y'_1(s) \). That is, higher the demand and the more will be ordered. It is evident that the dimensionality reduction approach reduces the computational effort significantly.

Taking the Gamma-Gamma model as an example, we have \( s_{t+1} = (\Lambda_t + \epsilon_t, a_t, q_t + 1) \) for each period \( t \). When we compute the optimal policy according to the optimality Eqs. (2) of the decomposed formulation, then to evaluate the term \( \int f_{t-1}(y(y - d) - s) \phi_0(y)dy \), we have to assess the profit-to-go function \( f_t(y) = \int (\Lambda_t + \epsilon_t, a_t + 1))\) for all possible values of \( \epsilon_t \), which is not computationally tractable as \( \epsilon_t \) is a continuous variable with a support \((0, \infty)\). However, with the decomposed formulation (5), we only need to evaluate the single-dimensional normalized value function \( v_{n+1} \), to compute the optimal normalized policy and then employ the transformations introduced in Theorem 3 to obtain the optimal policy, which is clearly more tractable and efficient.

**Remark 5.** (Data-driven decision rule and operational statistics). The optimal Bayesian ordering decision rule, \( y'_1(x|s) = q_0(s)y'_1(x/q_0(s)) \), can be seen as a data-driven decision rule with the data contained in the statistic \( s \). Note that the functional forms of the scale factor \( q_0(s) \) and the optimal order-up-to decision of the normalized model \( y'_1(.) \) do not depend on the statistic \( s \) but on the shape parameter \( a \) (or the total number of observations). Such a data-driven decision rule is similar to the operational statistics proposed by Liyanage and Shanthikumar (2005) as an alternative approach to the conventional estimation-then-optimization statistical learning approach. Their approach aims to find a decision rule that maximizes the priori expectation of the performance uniformly for all possible values of the unknown distribution parameters under a class of structured decision rules. Liyanage and Shanthikumar (2005) identify a class of operational statistics that are homogenous of order one for newsvendor models and find that the optimal operational statistic of this class can be derived by solving a normalized model. Chu et al. (2008) propose a Bayesian approach to find the optimal operational statistic with a non-informative prior. In our model, the optimal ordering decision rule derived from the normalized model, \( y'_1(x) = y'_1(x/a_1, 1) \), plays the same role of the optimal decision rule of operational statistics as it maximizes the priori expectation (conditionally on the current statistic) for the class of operational statistical rules, expressed as \( y'_0(s, q(s)|s) \), that are scalable with the factor \( q(s) \) and \( s \) is updated as that of \( \Lambda_t \). The link between Bayesian decision rules and operational statistics sheds new light on more general operational statistics in dynamic settings.

6. Effect of Bayesian learning on optimal policy

This section aims to identify the effect of learning on the optimal decisions by comparing the optimal policies in the system without demand learning and the system with demand learning.

6.1. The system without demand learning

In the system without Bayesian demand learning, the replenishment decisions are made based on the initial belief at the beginning of the planning horizon. That is, the multiplicative random perturbation of demand in each period has the same density function \( \phi_1(y) \). Let \( g_t \) be the value-to-go function (corresponding to \( f_t \)) in period \( t \) and \( g_{t-1}(x|s) = 0 \). The optimality equations are

\[
g_t(x|s) = \max_{y \in \mathbb{A} \cap \mathbb{C}_j} \left\{ G_t(y, d) := \pi_t(y, d | s) \right\}
\]

\[
+ \alpha \int_0^\infty \left[ \left( y_d - d(y - d)|s \right) \phi_1(y|s) dy \right], t = 1, \ldots, T.
\]

Let \( y_t(x|s) \) and \( d_t(x|s) \) be the optimal solutions and \( \left( \hat{y}_t(s), \hat{d}_t(s) \right) \) the global optimal solution of the optimization problem in (6). In addition, let \( \hat{p}_t(x|s) = \hat{p}_t \left( \hat{d}_t(x|s) \right) \) and \( \hat{p}_t(s) = \hat{p}_t \left( \hat{d}_t(s) \right) \).

The following theorem reveals the structure of the optimal policy without demand learning.

**Theorem 4 (Optimal policy without learning).** Suppose one of the following two pairs of conditions holds: (i) \((A1)\) and \((B2)\), (ii) \((A2)\) and \((B1)\) for all \(t\).

(a) A base-stock list-price policy is optimal: for any \( t \) and \( s_t \), if the starting inventory level \( x < y_t(s_t) \) then it is optimal to order up to the base stock level \( y_t(s_t) \) and set the list price as \( p_t(s_t) \); if, otherwise, \( x \geq y_t(s_t) \) then it is optimal to place no order and set the price as \( \hat{p}_t(s_t) \). Moreover, the optimal price \( \hat{p}_t(x|s_t) \) is nonincreasing in \( x \) and satisfies \( \hat{p}_t(x|s_t) \leq \hat{p}_t(s_t) \) for \( x \geq y_t(s_t) \).

(b) In particular, if the system is stationary in the sense that all parameters are time invariant, a myopic base-stock list-price policy is optimal with the optimal myopic base-stock and list-price levels being \( (\hat{y}_t(s_t), \hat{p}_t(s_t)) \) such that \( p_t(s_t) = \hat{p}_t \left( d_t(x|s_t) \right) \) and \( (\hat{y}_t(s_t), \hat{d}_t(s_t)) = \arg \max_{y \in \mathbb{A} \cap \mathbb{C}_j} \left\{ \pi_t(y, d | s_t) \right\} \): If the inventory level \( x < \hat{y}_t(s_t) \), it is optimal to place an order up to \( \hat{y}_t(s_t) \) and set the price as \( \hat{p}_t(s_t) \); if, otherwise, \( x \geq \hat{y}_t(s_t) \), it is optimal to place no order and set the price as \( \hat{p}_t(x|s_t) = \hat{p}_t \left( d_t(x|s_t) \right) \) where \( d_t(x|s_t) = \arg \max_{y \in \mathbb{A} \cap \mathbb{C}_j} \pi_t(y, d | s_t) \).

Theorem 4 shows that the base-stock list-price policy is optimal for the system without Bayesian update. In particular, when the system is time invariant and the distribution of random demand perturbation remains as the initial one, a myopic base-stock list-price policy is optimal.

Using the dimensionality reduction approach in Section 5, under the assumptions \((C1)\) and \((C2)\), we can readily show that \( y'_0(s, d'_1(s)) = (q_0(s)y'_1(s)), \) with \( \hat{y}_t, \hat{d}_t = \arg \max_{y \in \mathbb{A} \cap \mathbb{C}_j} \pi_t(y, d | s) \). To gain sharper insights, we next derive a closed-form solution of the optimal myopic base-stock and list-price levels for a special case in the system without learning.

**Lemma 2 (Optimal myopic policy).** Consider a stationary backordering system (i.e., \( \lambda = 1 \)) with an exponential-Gamma perturbation distribution with \( \phi_1(\epsilon|s) = a_1 \Lambda_1^d / (\Lambda_1 + \epsilon)^{d+1}, a_1 > 1 \). The expected operating profit in period \( t \) can be expressed as
πc(y, d|s1) = d(P1(d) − xc + h′) − \frac{Λ1}{a1 − 1} − d[(1 − x)c + h′]y/d + (h′ + h′) \frac{Λ1}{a1 − 1} (1 + (y/d)/Λ1)^{−(a1−1)}.

The optimal myopic base-stock level is \( y^M(s_1) = \Lambda_1(\rho^{-1/\alpha_1} − 1)d^h(s_1) \) and the optimal myopic list price is \( p^M(s_1) = P_1(d^h(s_1)) \) with the expected demand level
\[
d^h(s_1) = \arg \max_{\text{d}(s_1)} \left[d(P_1(d) − xc + h′) − d(h′ + h′) \rho(a_1\rho^{-1/\alpha_1} − a_1 + 1)\right].
\]

In addition, \( d^h(s_1) (p^M(s_1)) \) is increasing (decreasing) in \( a_1 \) and independent of \( \Lambda_1 \) and \( y^M(s_1) \) is increasing in \( \Lambda_1 \).

Lemma 2 provides a closed-form solution to the optimal myopic base-stock list-price policy in a system without Bayesian learning. Recall that the predictive demand distribution is stochastically decreasing in the shape parameter \( \alpha_1 \) and stochastically increasing in the scale parameter \( \Lambda_1 \). Hence, intuitively, as the shape parameter becomes smaller or the scale parameter becomes greater, the demand becomes larger (stochastically), which drives a greater stock level and a higher price. We show that the optimal myopic list price is indeed decreasing in the shape parameter (but independent of the scale parameter) and the optimal myopic base-stock level is increasing in the scale parameter. However, it is unclear whether the optimal myopic base-stock level is decreasing in the shape parameter. Fig. 1 below provides a numerical example which shows that both the optimal myopic base-stock level and the optimal myopic pricing decisions tend to be lower as the shape parameter \( \alpha_1 \) becomes greater: the smaller the demand the less to order and the lower price.

6.2. Effect of demand learning

We next compare the myopic base-stock list-price policy in the system without learning, \( (y^M(s_1), p^M(s_1)) \), to that under demand learning, \( (y^L(s_1), p^L(s_1)) \), in a stationary system. For ease of interpretation, we restrict our discussion on a two-period setting (T = 2).

6.2.1. Effect of learning in period 2

In period 2, the objective function is \( π_2(y, d|s_2) \) where \( s_2 = (s_2, a_2) \) with \( A_2 = A_1 + ε_1 ≥ A_1 \) and \( a_2 = a_1 + 1 \). The preceding analysis implies that \( (y^L(s_2), p^L(s_2)) = (y^M(s_2), p^M(s_2)) \). Since \( p^M(s_1) \) is decreasing in \( a_1 \) and independent of \( A_1 \), we know that \( p^M(s_2) ≤ p^M(s_1) \), which implies that demand learning leads to a smaller list price in the last period. For the optimal base-stock levels, \( y^L(s_2) ≥ y^M(s_2) \) or equivalently, \( A_2(\rho^{-1/\alpha_2} − 1)d^M(s_2) ≥ A_1(\rho^{-1/\alpha_1} − 1)d^M(s_1) \), if and only if \( A_2 ≥ (\rho^{-1/\alpha_2} − 1)\rho(a_1\rho^{-1/\alpha_1} − a_1 + 1) \), which implies that the optimal base-stock level in the last period is determined by accumulated demand perturbations (observed market sizes) over the planning horizon: the greater the observed total market size over time the more will be ordered in the last period, and the optimal basestock level with demand learning should be greater than that without learning if and only if sufficient market sizes were observed.

6.2.2. Effect of learning in period 1

In period 1, pricing and inventory decisions are made taking into account the effect of learning on the future decisions. We can employ the dimensionality decomposition method to derive the optimal policy under Bayesian learning. Note that \( q_2(s_1) = Λ_1 \) and \( U_{s_1}(x) = 1 + x \). The optimal bases-stock and list-price levels can be computed by backward induction according to (5). Since \( y^L_2(1, a_2) \) is the optimal base-stock level in period 2, \( ν_2(x) \) remains constant at \( ν_2(y^L_2(1, a_2)) = η_2(y^L_2(1, a_2), a_2(1, a_2)) \) for \( x ≤ y^L_2(1, a_2) \) and then decreases in \( x \) for \( x > y^L_2(1, a_2) \). For any \( y ≤ y^L_2(1, a_2) \) and \( ε_1 > 0 \), we have \( (y − dε)/(1 + ε_1) < y^L_2(1, a_2) \) and hence \( ν_2((y − dε)/(1 + ε_1)) = ν_2(y^L_2(1, a_2)) \). Since \( y^L_2(1, a_2) = y^M(1, a_2) ≤ y^M(1, a_1) \) and \( y^L(1, a_1), d^L(1, a_1) \) is the global maximizer of \( π_c(y, d) \), we know that \( y^L_2(1, a_1) \) must be no less than \( y^L_2(1, a_2) \).

Note that the monotonicity of \( ν_2 \) implies that the term \( \int_0^{α_1} ν_2(δ(y − dε)/(1 + ε_1))dε \) is decreasing in \( y \) and increasing in \( d \), which provides a negative driving force for the inventory decision and a positive (negative) driving force for the expected demand level (price) compared to the myopic policy.

![Fig. 1. Optimal myopic policy for a stationary backlog system](image-url)
However, the supermodularity of the objective function $\Pi_1(y, d)$ implies that a greater value of $y$ will lead to a greater optimal value of $d$, and vice versa. Although it is difficult to analytically compare $(y^*_1(1, a_1), p^*_1(1, a_1))$ and $(y^T(1, a_1), p^T(1, a_1))$, we know that $d^*_t(x|k(1, a_1)) \supseteq d^T_t(x|k(1, a_1))$ for $x \geq \max\{y^*_1(1, a_1), y^T_1(M(1, a_1))\}$, which implies that when inventory level is sufficiently high the positive effect of demand learning on pricing decisions dominates the positive effect on ordering decisions.

To gain some insight, we provide a numerical example to compare $(y^M(1, a_1), p^M(1, a_1))$, $(y^*_1(1, a_1), p^*_1(1, a_1))$, and $(y^T_1(1, a_1), p^T_1(1, a_1))$. Fig. 2(a) shows that the optimal first-period basestock level under demand learning is slightly lower than the optimal myopic one and both of them are greater than the optimal normal-level under demand learning is slightly lower than the optimal normal-level under demand learning.

To compare optimal policies with and without demand learning (Fig. 2), we provide a numerical example to compare $(y^M(1, a_1), p^M(1, a_1))$, $(y^*_1(1, a_1), p^*_1(1, a_1))$, and $(y^T_1(1, a_1), p^T_1(1, a_1))$. Fig. 2(a) shows that the optimal first-period basestock level under demand learning is slightly lower than the optimal myopic one and both of them are greater than the optimal normal-level under demand learning.

The effect of demand learning on the optimal pricing and inventory control policy can also be impacted by the magnitude of the demand uncertainty which is typically measured by the coefficient of variation (CV) of the demand. It is known that for both Weibull and the gamma distributional specifications of the demand perturbation the CV does not depend on the unknown scale parameter $\omega$, which decreases in $\frac{1}{\sqrt{k}}$ for a given value of the scale parameter $k$. That is, the greater the value of the shape parameter $k$ the smaller the variation of the demand.

We next provide a numerical study on the effect of demand volatility on the optimal policy under a Gamma-Gamma model. We focus on the last period for simplicity. Figs. 3 and 4 compare the optimal normalized order-up-to levels and price levels for three levels of the shape parameter of demand ($k = 1, 2, 3$) with respect to the posterior distributions of $\omega$ with shape parameters $a_1 = 3$ and $a_2 = 10$, respectively. The CVs of the demand corresponding to different values of $k$ are $1/\sqrt{2}$ and $1/\sqrt{3}$, respectively, while the CVs of the posterior distribution of the unknown scale parameter $\omega$ are $1/\sqrt{3}$ and $1/\sqrt{10}$, respectively. Note that the greater the value of $a_2$ the greater the accuracy of the posterior distribution of $\omega$. From Figs. 3 and 4, we can observe that for both levels of $a_2$, as $k$ decreases, or equivalently, the demand volatility increases, the optimal normalized order-up-to levels (as well as the optimal base-stock level) decrease while the optimal list-price level increase. But when the inventory level is relatively high it is possible that a greater degree of demand variation may lead to a lower price for a larger value of $a_2$ (a more precise posterior distribution of $\omega$). These results imply that the ordering and pricing decisions can be used to leverage the profitability of the system. When the inventory level is relatively low, a higher level of demand volatility drives a higher price, which leads to a lower demand level, and a lower order-up-to level to match the demand and supply. When the inventory is relatively high and the posterior distribution is more precise, as the demand volatility is higher, instead of using a higher price to contain the demand level, it is better to set a lower price to speed up the inventory depletion.

6.3. Effect of demand variation

The effect of demand learning on the optimal pricing and inventory control policy can also be impacted by the magnitude of the demand uncertainty which is typically measured by the coefficient of variation (CV) of the demand. It is known that for both Weibull and the gamma distributional specifications of the demand perturbation the CV does not depend on the unknown scale parameter $\omega$, which decreases in $\frac{1}{\sqrt{k}}$ for a given value of the scale parameter $k$. That is, the greater the value of the shape parameter $k$ the smaller the variation of the demand. The CV of the posterior distribution of $\omega$ is $1/\sqrt{\omega} = 1/\sqrt{a_2 + k(t - 1)}$ which decreases in the number of observations $t - 1$.

7. Extensions

7.1. Unobserved lost sales (censored demand)

The preceding analysis is based on the assumption that unmet demand, backorders or lost sales, is fully observable. In this
subsection, we assume that unmet demand is lost and unobserved. This is common in retailing where consumers would purchase the product if it were available but may simply leave without identifying themselves when facing an empty shelf. In other words, the demand information in the presence of stock out is censored.

With censored demand (i.e., only sales are observed), the replenishment and pricing decisions interact with the demand learning process, which implies that the firm needs to take into account how its replenishment and pricing decisions affect the demand information. To address this problem, we need to modify the Bayesian dynamic programming formulation. For simplicity, we restrict our attention to the full lost-sales case (\( \lambda = 0 \)).

When the unmet demand is not observed and lost, the demand information is censored and hence \( C_t \) cannot be observed. By Bayes’ rule, given a sufficient statistic \( s \) and a censored demand information \( C_t; y = d(\cdot) \), the posterior density on \( x \) is updated as

\[
g(x | s, C_t; y = d(\cdot)) = \int_0^\infty w(y | x) g(x | s, d(\cdot)) dx \quad \text{if } d < y,
\]

\[
g(x | s, C_t; y = d(\cdot)) = \int_0^\infty w(y | x) g(x | s, d(\cdot)) dx \quad \text{if } d \geq y.
\]

Note that a conjugate prior for uncensored demand may no longer be a conjugate for the censored demand. One exception is the Weibull-Gamma family. With the Weibull-Gamma model, the posterior distribution conditioned on \( d_t \geq y \) is a Gamma distribution with \( s_{t+1} = (a_t, A_t + (y/d_t)^k) \).

Fig. 3. Effect of demand volatility under Gamma-Gamma model \( (\phi_t(x | \alpha_t, 1) = \frac{\Gamma(\gamma_t + k_t)}{\Gamma(\gamma_t) \Gamma(k_t)} x^{\gamma_t - 1} e^{-x/\gamma_t}, k_t \in \{1, 2, 3\}, \alpha_t = 3, P_t(d) = 1 - d \text{ for } d \in [0, 1], x = 1, c = 0, h^- = h^+ = 0.1) \).

Fig. 4. Effect of demand volatility under Gamma-Gamma model \( (\phi_t(x | \alpha_t, 1) = \frac{\Gamma(\gamma_t + k_t)}{\Gamma(\gamma_t) \Gamma(k_t)} x^{\gamma_t - 1} e^{-x/\gamma_t}, k_t \in \{1, 2, 3\}, \alpha_t = 10, P_t(d) = 1 - d \text{ for } d \in [0, 1], x = 1, c = 0, h^- = h^+ = 0.1) \).
Letting \( f_{T+1}(x|A, \Lambda) = 0 \), with Weibull-Gamma distributions, the optimality equation becomes

\[
f_t(x|A, \Lambda) = \max_{y, x, d_0} \left( \pi_t(y, d|a, \Lambda) + \alpha \int_0^\infty f_{T+1}(y - d - z|s) j_0 T d(z) \phi_z(z|a, \Lambda) dz \right).
\]

Adding the scale parameter \( \Lambda \) and Heching (1999), Chen and Simchi-Levi (2004), Chen et al. (2012) for the case with backorders and Pang (2011) for the case with lost sales. Our model can be readily modified to address the case with additive demand.

With the additive demand, the optimality equations, for \( t = 1, \ldots, T \), can be expressed as

\[
f_t(x|s) = \max_{y, x, d_0} \left\{ \pi_t(y, d|s) + \alpha \int_0^\infty f_{T+1}(y - d - z|s o c) j_0 T d(z) \phi_z(z|s) dz \right\},
\]

where \( f_{T+1}(s|s) = 0 \) and

\[
\pi_t(y, d|s) = R_t(y, d|s) - (1 - z) c T y + \alpha d T e \min(d + e, y) + z(d + e) \phi_z(z|s) \phi_z(z|s).
\]

Following the analysis of Pang (2011), we know that if (A2) holds and the price elasticity of the lost-sales probability for additive demand, expressed as \(- \frac{\partial \phi_z(z|s)}{\partial y} \), is greater than one then \( R_t(y, d|s) \) is jointly concave and supermodular in \((y, d)\). Using similar arguments of Pang (2011), one can readily show that the state-dependent base-stock list-price policy is optimal. Similar to Pang (2011), with additive demand, our analysis can be further extended to address the case with fixed order cost and show that the state-dependent \((s, S, p)\) policy is optimal.

It is natural to ask whether the dimensionality reduction approach applies to the model with additive demand; see Zhang and Chen (2006) for an attempt. It is notable that the preceding dimensionality reduction analysis relies on the scalability of the multiplicative demand. When the demand is of the multiplicative form, the sales of each period, \((1 - z) \min(d + e, y) + z(d + e)\), is scalable for any \( d \in \mathcal{D}_t \), which leads to the scalability of the revenue and profit functions. However, when the demand is of the additive form, scaling the random variable \( z_t \) requires scaling the decision variable \( d_t \) while the revenue function is nonlinear in \( d \) and therefore not scalable. As a result, the dimensionality reduction approach does not apply to the model with additive demand.

8. Concluding remarks

Operating businesses in the face of volatile market environments, especially the unprecedented global public health and economic crisis caused by COVID-19, firms need to redesign their operational strategies with all operational levers while actively learning the changing market conditions. Motivated by this, we study how a firm can coordinate its dynamic pricing and inventory management strategies while learning the demand (market size) dynamically using a Bayesian approach. We consider a price-sensitive demand model of the multiplicative form as a product of a deterministic price-dependent function and a random factor (market size) with unknown distribution. Unmet demand is partially backlogged/lost. Following Bayesian inventory management literature, we assume that the distribution of the market size depends on a random parameter and the firm forms a prior distribution on the random parameter at the beginning of the planning horizon.

Focusing on the class of conjugate prior distributions, we formulate the problem as a Bayesian dynamic program. We identify a set of regularity conditions under which the objective functions are jointly concave and supermodular in the order-up-to level and the expected demand level, which allows us to show the optimality of the state-dependent base-stock and list-price policy.

Furthermore, we employ a dimensionality reduction approach to decompose the profit-to-go functions as a product of a scale function that depends on the accumulate demands over time and the profit-to-go functions of the normalized problem, which significantly simplifies the computation and implementation of the model. We show that the optimal order-up-to levels can also be decomposed as the product of the scale function and the optimal order-up-to levels of the normalized problem. With the
decomposition, the firm can first compute the optimal policy for the normalized problem at the beginning of the planning horizon and then update the optimal inventory and pricing decisions with accumulate demand data as time goes by.

The dimensionality reduction approach also facilitate our discussion on the effect of demand learning on the optimal policy. Using the system without Bayesian update (i.e., the distribution based on the initial prior throughout the planning horizon), we show that demand learning may lower future prices while the future inventory decisions depend on the realized values of the random demand over time. Demand learning may lead to a lower base-stock level and a lower price for high inventory levels in the current period. We also discuss the effect of demand variation on the optimal policy under demand learning in a numerical example and observe that a greater degree of demand variation leads to lower optimal order-up-to levels and greater optimal list-price levels while the optimal price level may decrease in the level of demand variation as the inventory level becomes sufficiently high and the posterior distribution of the unknown scale parameter is more precise.

We further extend our structural analysis to the model with unobserved lost sales (i.e., censored demand), and show that the dimensionality reduction result also holds but the base-stock list-price policy may no longer be optimal. We also discuss the case with additive demand and show that the state-dependent base-stock list-price policy is also optimal but the dimensionality reduction does not apply due to the lack of scalability from the additive demand model.

Last but not least, we outline several limitations of our paper and future research directions. First, the classic Bayesian approach, though providing strong tractability in both structural analysis and computation, requires the firm to form the prior probability, which can be difficult in some situations. A potential alternative approach is the reinforcement learning (e.g., Q-learning) which requires little knowledge of the transition probability but focusing on the structure of the decision rules. Second, our Bayesian analysis focuses on learning the market size (random factor), assuming that the price-dependent factor is known for tractability. Note that the price-dependent factor can be explained as the purchasing probability (market share) of WTP at the given price. The distribution of WTP is typically estimated through econometric analysis or pricing experiments. It will be interesting to combine different learning approaches to further advance dynamic inventory models. Finally, the current paper focuses on theoretical developments of the empirical Bayesian inventory models though our model is data-driven by its nature. We will empirically examine our model with real-world data in the future.

A.2. Proof of Theorem 1

We first show that $R_t(y, d)(s)$ is jointly concave under conditions of (i) or (ii).

Note that $R_t(y, d)(s) = (1 - \lambda)P_t(y|d)(\text{min}(c, y)|s) + \lambda P_t(y|d)e|s]$. It is clear that (A1) or (A2) ensures that $P_t(y|d)$ is concave in $y$. Thus, it suffices to show that the expected truncated revenue (of lost-sales model), $R_t(y, d) = P_t(y|d)(\text{min}(c, y)|s)$, is jointly concave in $(y, d)$ for any $s$. To this end, we need to show that the Hessian matrix of $R_t(y, d)$ is negative semi-definite. It suffices to show that $\nabla^2 R_t(y, d)$ is concave in each component and the determinant of Hessian matrix, $\Delta(y, d)$, is nonnegative.

Taking the second-order derivative with respect to $d$ yields

$$\nabla^2 R_t(y, d) = -P_t(y|d)\Phi_t(y|d)/d/d < 0,$$

which indicates that $R_t(y, d)$ is concave in $y$.

Taking the cross–partial derivative with respect to $y$ and $d$ yields

$$\nabla^2 R_t(y, d) = P_t(y|d)\Phi_t(y|d)/d/d + P_t(y|d)\phi_t(y|d)/d/d^2$$

where the inequality is due to $\rho_t(y|d)e|d) > 1$ and $P_t(y|d)e|d) < 0$. Hence $R_t(y, d)$ is supermodular.

Pairing up the terms yields

$$\Delta(y, d) = \frac{d^2\Delta}{d\epsilon d\epsilon} < 0$$

Clearly, both terms of the RHS of (10) are nonnegative under assumptions (i) or (ii). Hence, the Hessian of $R_t(y, d)$ is negative semi-definite, implying that it is jointly concave in $(y, d)$ for any $s$. We are now ready to show the optimality of the base-stock list-price policies by induction.

Suppose that $f_{s+1}(x|s)$ is nonincreasing and concave in $x$ for any $s$. Note that $\delta(\cdot)$ is an increasing and convex function. Then $f_{s+1}(\delta(x)|s)$ is non-increasing and concave in $x$, which further implies that $f_{s+1}(\delta(y-de)|s)$ is jointly concave and supermodular in $(y, d)$ for any given $s$. Since concavity can be preserved under expectation operation, we know that $\int_{0}^{y} f_{s+1}(\delta(y-de)|s)\phi_s(e|s)de$ is jointly concave and supermodular in $(y, d)$ for any $s$. Similarly, the convexity of $\Delta_i$ implies that

Appendix A. Proofs of statements

A.1. Proof of Lemma 1

For $t < T$, we have $y_t(s) \geq u_t$ by the definition of $u_t$. For $t < T$, it suffices to show that $\Pi_t(y, d|s)$ is increasing in $y \leq u_t$ for any $s$ and $d \in \mathbb{D}_k$.

By induction. Suppose that for some $t \leq T$ we have $y_{t+1}(s) \geq u_{t+1}$, which implies that $f_{t+1}(y|s) = f_{t+1}(y_{t+1}(s)|s)$ (or, $\delta_{t+1}(y)|y| = 0$) for all $y \leq u_{t+1}$ and $s$. Note that $\delta(y-de) \leq y$ and $u_t < u_{t+1}$. Then, for any $y \leq u_t f_{t+1}(\delta(y-de)|s| = 0)$ must be a constant for any $d$ and $s$. and so is $\int_{0}^{u_t} f_{t+1}(\delta(y-de)|s|)\phi_s(e|s)de.$ Since $u_t \leq y_t(d|s) = \text{arg max}_{y} \pi_s(y, d|s)$ for any $d \in \mathbb{D}_k$ and $s$ and $\pi_s(y, d|s)$ is concave in $y$, $\pi_s(y, d|s)$ is increasing in $y$ for any $y \leq u_t$. Thus, for any $d$ and $s$, $\Pi_t(y, d|s)$ is increasing in $y$ for any $y \leq u_t$, which implies that $y_t(s) \geq u_t$. This completes the proof.
\[ f_0 \int \varphi(y - d)\phi_1(\epsilon|s)\epsilon\,d\epsilon \text{ is jointly convex and submodular in } (y,d) \text{ for any } s. \]

Since \( R_t(y,d|s) \) is jointly concave and supermodular in \((y,d), \Pi_t(y,d|s)\) is jointly concave and supermodular in \((y,d)\) for any given \( s \). Following the standard inductive arguments of Federgruen and Heching (1999) and Pang (2011), the desired results follow immediately.

### A.3. Proof of Theorem 3

We first prove part (a) by induction. Note that \( f_{t+1}(x|s) = 0 \) automatically satisfies the desired properties. Suppose \( f_t(x|s) = q_t(s)\Pi_{t-1}(x|q_t(s)) \) for some \( t \). We will show that \( f_1(x|s) = q_1(s)\Pi_0(x|q_1(s)) \).

First, by the definitions of \( \pi_t(y,d|s) \) and \( \pi_t(y,d) \), the assumption of \( \gamma_t \), and condition (C1), we have

\[
\pi_t(y,d|s) = \int_0^\infty \left[ \pi_t(d)q_t(s) \right] \frac{d\epsilon}{q_t(s)} - \alpha c q_t(s) \frac{d\epsilon}{q_t(s)} \frac{d\epsilon}{q_t(s)} - (1 - \alpha) c q_t(s) \frac{d\epsilon}{q_t(s)} - \pi_t(y,d|s)\gamma_t(y|q_t(s)) - \pi_t(y,d - d|s)\gamma_t(y|q_t(s))
\]

For any fixed \( s \), the optimal solution of (13) also maximizes the right-hand side of (12), i.e., \( \gamma_t(y|q_t(s)) = q_t(s)\pi_t(y,q_t(s),d) \) (which implies that \( p_t(x|s) = \pi_t(x|q_t(s)) \)). In particular, the state-dependent base-stock level \( y_t = q_t(s)_t \) and the optimal list price \( p_t(s) = p_t(S) \). This proves the part (b).

### A.4. Proof of Theorem 4

First, following similar arguments in the proof of Theorem 1, the optimality of the base-stock list-price policy and the monotonicity of the optimal policy follow immediately. In addition, both \( \pi_t(y,d|s) \) and \( G_t(y,d) \) are jointly concave and supermodular in \((y,d)\) and \( G_t(y) \) is decreasing in \( x \). Clearly, \( G_t(y,d(y)) \) is concave in \( y \).

We next consider the stationary system and prove the optimality of the myopic policy by induction.

Suppose that the optimal myopic base-stock and expected demand levels \( (y^M(s_1),d^M(s_1)) \) are optimal in period \( t+1 \), which implies that it is always optimal to order up to \( y^M(s_1) \) in period \( t+1 \) as the beginning inventory level is lower than \( y^M(s_1) \). As a result, the profit-to-go function \( g_{t-1}(y) \) is constant for \( y < y^M(s_1) \) and then decreasing for \( y > y^M(s_1) \). Hence, for \( y < y^M(s_1) \), \( g_{t-1}(y) \) is constant and so is \( \int_0^{g_{t-1}(\delta(y-d))\phi_1(\epsilon|s)} \epsilon \,d\epsilon \), which implies that both \( \pi_t(y,d|y(s_1)|s_1) \) and \( G_t(y,d|y(s_1)|s_1) \) are increasing in \( y \) for \( y < y^M(s_1) \) where \( d_t(y(s_1)) = \arg \max_{d_t(y(s_1))} \pi_t(y(s_1),d_t(y(s_1))|s_1) \).

In particular, \( \pi_t(y,d|y(s_1)|s_1) \) reaches the maximum when \( y = y^M(s_1) \), and we have \( d^M(s_1) = d_t(y^M(s_1)|s_1) \).

We next show that \( y^M(s_1) \) maximizes \( G_t^\dagger(y^M(s_1)|s_1) \). For any \( \eta > 0 \), on one hand, we have

\[ G_t(y^M(s_1) + \eta, d_t(y^M(s_1)|s_1)) - G_t(y^M(s_1) + \eta, d_t(y^M(s_1)|s_1)) \]

where the inequality is due to the facts that \( y^M(s_1), d_t(y^M(s_1)|s_1) \) is the global maximizer of \( \pi_t(\cdot, \cdot|s_1) \) and \( g_{t-1}(\cdot) \) and \( \delta(\cdot) \) is decreasing and increasing functions, respectively. On the other hand, we have

\[ \int_0^{g_{t-1}(\delta(y-d))\phi_1(\epsilon|s)} \epsilon \,d\epsilon \leq 0, \]

where for any fixed \( s \), the optimal solution of (13) also maximizes the right-hand side of (12), i.e., \( y^*(x|s) = q_t(s)\Pi_t(x|q_t(s)) = q_t(s) \max(y_t,x|q_t(s)) \) and \( d^t(x|s) = d_t(x|q_t(s)) \) which implies that \( p_t(x|s) = \Pi_t(x|q_t(s)) \). In particular, the state-dependent base-stock level \( y_t = q_t(s)_t \) and the optimal list price \( p_t(s) = p_t(S) \). This proves the part (b).
These results, together wish the joint concavity of \(G_t(\cdot, \cdot)\), imply
that\( \left(y^d(s_1), d^d(s_1)\right) = \arg\max_{y>0,0<d_1-s_1} G_t(y,d) \). Hence, the myopic base-stock list-policy is indeed optimal.

A.5. Proof of Lemma 2

Let \( z = y - d \). We have
\[
\pi(y, d; s_1) = \frac{(y - d)^{\alpha_1}}{(1 - \alpha_2,y - d)(s_1)} \frac{d(P_t(y - d + y) + h + n' + h y)}{d(y - d + y) + h + n' + h y} \frac{d(\alpha_1 y^{\alpha_1} - 1)}{d(y - d + y) + h + n' + h y} \int_z \phi_1(y, \alpha_1, \alpha_2, \alpha_3) \phi_1(y, \alpha_1, \alpha_2, \alpha_3)
\]
where the third equality is due to the integration by parts
\[
\int_z (e - z) \phi_1(y, \alpha_1, \alpha_2, \alpha_3) \phi_1(y, \alpha_1, \alpha_2, \alpha_3) \, \text{d}z
\]
which is equal to the above function is convex in \( \alpha_1 \) and decreasing in \( \alpha_1 \). For the optimal base-stock level \( y^d(s_1) = d^d(s_1) = \phi_1(s_1) \), since \( z^d(s_1) \) is increasing in \( \alpha_1 \) and decreasing in \( \alpha_1, y^d(s_1) \) is increasing in \( \alpha_1 \) but it is unclear whether it is decreasing in \( \alpha_1 \).

References

ABC News. September 29, 2000. Amazon error may end “dynamic pricing.” URL: http://abcnews.go.com/Technology/story?id=119399, accessed May 20, 2020.

Abdelnour, A., T. Babbitt, S. Moss., 2020. Pricing in a pandemic: navigating the COVID-19 crisis. (Mckinsey & Company) URL:https://www.mckinsey.com/business-functions/marketing-and-sales/out-insights/pricing-in-a-pandemic-navigating-the-covid-19-crisis-assessed May 20, 2020.

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

Azoury, K.S., 1985. Bayesian solution to dynamic inventory models under unknown demand distribution. Manage. Sci. 31 (9), 1150–1160.

BBC. 2012. Tesco’s UK chief executive quits. URL:https://www.bbc.com/news/business-17378409, accessed May 20, 2020.

Biller, S.L.M.A., Chan, D., Simchi-Levi, J. Swann, 2005. Dynamic pricing and the direct-to-customer model in the automotive industry. Electron. Commerce Res. 5 (2), 205–334.

Bodla, T., Ferguson, M., 2012. Pricing Segmentation and Analytics. Business Expert Press, New York.

BURNS, D. 2020. How COVID-19 affected U.S. consumer prices in March. (Reuters, April 10) URL:https://www.reuters.com/article/us-health-coronavirus-usa-prices/how-covid-19-affected-u-s-consumer-prices-in-march-idUSKCN21S20H, accessed May 20, 2020.

Chen, X., Pang, Z., Pan, L., 2014. Coordinating inventory control and pricing strategies for perishable products. Oper. Res. 62 (2), 284–300.

Chen, X., Simchi-Levi, D., 2004. Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: the finite horizon case. Oper. Res. 52 (6), 887–896.

Chen, X., Simchi-Levi, D., 2012. Pricing and inventory management. Phillips, P., Ozer, E. (Eds.), Oxford Handbook of Pricing Management, Oxford University Press, UK, pp. 784–822.

Cheung, C.C., Simchi-Levi, D., Wang, H., 2017. Dynamic pricing and demand learning with limited price experimentations. Oper. Res. 65 (6), 1722–1731.

Chu, L.Y., Shanthikumar, J.C, Shen, Z.-J.M., 2008. Solving operational statistics via a Bayesian analysis. Oper. Res. Lett. 36, 110–116.

DeGroot, M.H., 1970. Optimal Statistical Decisions. John Wiley & Sons Inc, Hoboken, New Jersey.

den Boer, A.V., 2015. Dynamic pricing and learning: historical origins, current research, and new directions. Survey Or. Res. Manage. Sci. 20 (1), 1–18.

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

Federgruen, A., Heching, A., 1999. Combined pricing and inventory control under uncertainty. Oper. Res. 47 (3), 454–475.

Gallego, G., van Ryzin, G., 1994. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. Manage. Sci. 40, 999–1020.

Granot, D., Yin, S., 2008. Price and order postponement in a decentralized newsvendor model with multiplicative and price-dependent demand. Oper. Res. 56 (1), 121–139.

Aydin, G., Porteus, E.L., 2008. Joint inventory and pricing decisions for assortment. Oper. Res. 56 (5), 1247–1255.

Harrison, S., Byrne, B., Williams, O., Smith, T., 2020. Why am I paying $60 for that bag of price on Amazon? (The Markup, April 28) URL:https://themarkup.org/ask-the-markup/2020/04/28/why-am-i-paying-60-for-that-bag-of-rice-on-amazon-com, accessed May 20, 2020.

Karlin, S., Carr, C.R., 1962. Prices and optimal inventory policy. In: Arrow, K.J., Lin, S.R., Scarf, H. (Eds.), Studies in Applied Probability and Management Science. Stanford University Press, Stanford, CA, pp. 159–172.

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

Liu, W., Shanthikumar, J.C, 2005. A practical inventory control policy using operational statistics. Oper. Res. Lett. 33, 341–348.

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

Kocabiyikoglu, A., Popescu, I., 2011. An elasticity approach on the newsvendor with price sensitive demand. Oper. Res. 59 (2), 301–312.

Mehta, N., Detroja, P., Agashe, A., 2018. Amazon changes prices on its product about every 10 minutes. (Business Insider, Aug. 10) URL:https://www.businessinsider.com/amazon-price-changes-2018-8, accessed May 20, 2020.

Miller, B., 1986. Scarf’s state reduction method, flexibility, and a dependent demand inventory model. Oper. Res. 34 (1), 83–90.

Pang, Z., 2011. Optimal dynamic pricing and inventory control with stock deterioration and partial backordering. Oper. Res. Lett. 39 (5), 375–379.

Pang, Z., Chen, F., Feng, Y., 2012. A note on the structure of joint inventory-pricing control with leadtimes. Oper. Res. 60 (3), 581–587.
Petruzzi, N.C., Dada, M., 1999. Pricing and the newsvendor problem: a review with extensions. Oper. Res. 47 (2), 183–194.

Petruzzi, N.C., Dada, M., 2002. Dynamic pricing and inventory control with learning. Naval Res. Logist. 49, 303–325.

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

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

Scarf, H.E., 1960. Some remarks on bayes solution of the statistical inventory problem. Naval Res. Logist. 7, 591–596.

Solon, O., 2011. How a book about flies came to be priced $24 million on Amazon. (Wired.com April 27) URL:https://www.wired.com/2011/04/amazon-flies-24-million/, accessed May 20, 2020.

Song, Y., Ray, S., Boyaci, T., 2009. Optimal dynamic joint inventory-pricing control for multiplicative demand with fixed order costs and lost sales. Oper. Res. 57 (1), 245–250.

Subrahmanyan, S., Shoemaker, R., 1996. Developing optimal pricing and inventory policies for retailers who face uncertain demand. J. Retail. 72, 7–30.

Talluri, K.T., van Ryzin, G., 2004. The Theory and Practice of Revenue Management. Kluwer Academic Publishers, Boston, MA.

Topkis, D.M., 1998. Supermodularity and Complementarity. Princeton University, Princeton, NJ.

Zhang, J.L., Chen, J., 2006. Bayesian solution to pricing and inventory control under unknown demand distribution. Oper. Res. Lett. 34 (5), 517–524.