## Manufacturing & Service Operations Management

# Dynamic Inventory-Pricing Control Under Backorder: Demand Estimation and Policy Optimization

Qi Feng, Sirong Luo, Dan Zhang

Please scroll down for article—it is on subsequent pages

# Dynamic Inventory–Pricing Control Under Backorder: Demand Estimation and Policy Optimization

## Qi Feng

Krannert School of Management, Purdue University, West Lafayette, Indiana 47907, annabellefeng@purdue.edu

## Sirong Luo

School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai 200433, China; and
Key Laboratory of Mathematical Economics (SUFE), Ministry of Education, Shanghai 200433, China,
luo.sirong@mail.shufe.edu.cn

## Dan Zhang

Leeds School of Business, University of Colorado at Boulder, Boulder, Colorado 80309, dan.zhang@colorado.edu

Inventory-based dynamic pricing has become a common operations strategy in practice and has received considerable attention from the research community. From an implementation perspective, it is desirable to design a simple policy like a base-stock list-price (BSLP) policy. The existing research on this problem often imposes restrictive conditions to ensure the optimality of a BSLP policy, which limits its applicability in practice. In this paper, we analyze the dynamic inventory and pricing control problem in which the demand follows a generalized additive model (GAM). The GAM overcomes the limitations of several demand models commonly used in the literature, but introduces analytical challenges in analyzing the dynamic program. Via a variable transformation approach, we identify a new set of technical conditions under which a BSLP policy is optimal. These conditions are easy to verify because they depend only on the location and scale parameters of demand as functions of price and are independent of the cost parameters or the distribution of the random demand component. Moreover, although a BSLP policy is optimal under these conditions, the optimal price may not be monotone decreasing in the inventory level. We further demonstrate our results by applying a constrained maximum likelihood estimation procedure to simultaneously estimate the demand function and verify the optimality of a BSLP policy on a retail data set.

*Key words*: inventory–pricing; generalized additive models; dynamic programming
*History*: Received: September 6, 2010; accepted: June 16, 2013. Published online in *Articles in Advance* October 23, 2013.

## 1. Introduction

Dynamically adjusting product price based on inventory availability has become an important operations strategy for many firms, including Dell, Amazon, FairMarket, Land's End, J. C. Penney, MSN Auction, and Grainger (Elmaghraby and Keskinocak 2003, Chan et al. 2005). Firms have realized the inextricable connection between pricing and inventory decisions. On the one hand, demand shaping via dynamic pricing helps to mitigate the risk of under- or overstocking. On the other hand, frequent inventory replenishment allows for stabilizing the variations in price (Amihud and Mendelson 1983). A strategy that appropriately explores the inventory–price connection would greatly enhance the firms' ability to match supply with demand, leading to significant profit increase. The combined inventory and pricing planning has also received considerable attention in research community. Studies in this area have demonstrated major benefits derived from coordinating inventory and pricing decisions (e.g.,

Federgruen and Heching 1999, Chen and Simchi-Levi 2004a, Gimpl-Heersin et al. 2008).

Successful integration of inventory and pricing management requires careful considerations of implementation. It is desirable to design a simple policy that is easy to communicate, execute, and track. A plausible candidate is the base-stock list-price (BSLP) policy. Under this policy, one orders up to a base-stock level and charges a list price whenever the inventory level is below the base-stock level. When the inventory level is above the base-stock level, no order is issued, and the price charged depends on the inventory level. To be able to apply this policy, one needs to ensure that the business environment fits the conditions where the policy would yield a good performance. This requires that the underlying model used to compute the policy parameters is flexible enough to accommodate specific product characteristics and possibly changing demand patterns. Moreover, one should be able to easily verify based on

the estimates of model input whether or not a BSLP policy can be applied to a specific business case.

An appropriate price-dependent demand model is the central component for this problem. For a general demand model, however, the inventory–pricing problem has its technical challenge. As a result, most academic studies are restricted to either an additive demand model, in which the price only affects the location parameter of the demand distribution, or a multiplicative demand model, in which the price only affects the scale of the demand distribution (see a comprehensive review by Petruzzi and Dada 1999). The advantage of these demand models lies in their analytical tractability. However, from a statistical viewpoint, both demand models are *one-parameter* models, because only the location or scale parameter is estimated as a function of the price. Although additive and multiplicative models are often instrumental in identifying problem structures and generating insights for such problems, they are quite restrictive for demand estimation and may seriously limit the applicability of the models and results.

We use a *generalized additive model* (*GAM*) to capture the price–demand relationship. The GAM, widely used in the statistics literature, is a *two-parameter* model that considers the price impact on both the location and scale of the demand distribution. Hastie and Tibshirani (1990) provide the fundamental theory on the GAM. It has been applied to medical problems (e.g., Hastie and Herman 1990), financial risk screening (e.g., Liu and Cela 2007), and prediction of email spam (e.g., Hastie et al. 2001), among others.

Although the GAM has been used previously to derive analytical policy structure for inventory–pricing problems, it is mainly applied to single-period models (e.g., Young 1978). For multiperiod problems, one needs to impose restrictive conditions to ensure the proper behavior of the dynamic programming recursion and thus the optimality of a BSLP policy (e.g., Young 1978, Federgruen and Heching 1999). A major development from our analysis is to identify a new set of conditions for the optimality of a BSLP policy. In contrast to previous studies, the conditions derived from our analysis depend only on the location and scale parameters of the demand, but not on the cost structure and the distribution of random demand component. Moreover, because the location and scale parameters of the demand are scalar functions of the price, as long as one knows the functional form, it is straightforward to check the optimality conditions. Hence, our result enhances the applicability of the BSLP policy in practice.

An interesting observation from our analysis is that under the GAM, the optimal price may *not* be monotone with respect to the inventory level, even when a BSLP policy is optimal. This observation is in sharp contrast to a commonly held notion derived from previous studies—it is optimal to lower the price when the inventory level is high. Compared with those adopted in previous studies, the demand model used in our analysis allows the flexibility to capture the price interactions with both the location and scale of the demand without restricting the location–scale relationship. In particular, the location can have a higher or lower price elasticity than the scale. As a result, the demand is not necessarily decreasing in price for *every* realization, as is assumed in previous studies, leading to a possibly nonmonotone relation between the optimal price and the inventory level. The application of our model to a retail data set confirms that, indeed, such a situation can arise in practice.

Whereas there has been considerable analytical development of inventory–pricing problems, little has been done in using real data to verify the applicability of analytical results. To bridge this gap, we demonstrate how to apply our results to real data using statistical methods. Specifically, we formulate a constrained maximum likelihood estimation (CMLE) problem to fit our demand model to a retail data set. This formulation is general enough to incorporate various factors that affect the price–demand relationship, as well as capture nonstationarity in the demand parameters. Our data analysis verifies the optimality conditions for a BSLP policy and thus confirms the applicability of the policy. Moreover, we show that the benefit of using the GAM over using a one-parameter model can be significant.

The remainder of this paper is organized as follows. In §2, we review the related literature and discuss our contributions. In §3, we lay out the dynamic programming model. In §4, we transform the problem by redefining the decision variables, which allows us to characterize a set of easy-to-verify conditions for an optimal BSLP policy. We further demonstrate the application of these conditions on commonly used demand functions, as well as identify properties of the optimal pricing decision. In §5, we develop a demand estimation framework and apply it to a retail data set to demonstrate the value of the GAM over one-parameter demand models. This estimation framework also allows us to verify the conditions for an optimal BSLP policy. Finally, we conclude the study in §6 by summarizing our findings and pointing out future research directions. Proofs of formal results are relegated to the appendix.

## 2. Literature Review

There is a substantial literature on coordinating pricing and inventory decisions. Excellent surveys are provided by Elmaghraby and Keskinocak (2003), Yano

and Gilbert (2003), and Chan et al. (2005). This line of research dates back to Whitin (1955) with further developments on the single-period problem made by many authors. The discussions center around comparing the optimal price with the riskless price, the optimal price when no uncertainty is present, under risk neutrality (Mills 1959, Karlin and Carr 1962, Young 1978, Petruzzi and Dada 1999) or risk aversion (Agrawal and Seshadri 2000). These studies assume either an additive demand model, in which the price affects only the location of the demand, or a multiplicative demand model, in which the price affects only the scale of the demand. Young (1978) proposes a more general model, which involves a single random variable and the pricing decision affects both the location and the scale of the demand. Recent development on single-period models includes work by Yao et al. (2006), Kocabıyıkoğlu and Popescu (2011), and Lu and Simchi-Levi (2013). In particular, Kocabıyıkoğlu and Popescu (2011) introduce *the lost-sales elasticity* to characterize sufficient conditions under which the objective function exhibits desired properties and the problem admits a unique solution. Though the elasticity concept is useful for single-period lost-sales problem, it is not helpful to tackle multiperiod models. The technique needed for multiperiod models can be fundamentally different from those for single-period problems.

When extending the analysis from a single-period model to a multiperiod model, one finds the major challenge of preserving certain functional properties in the dynamic programming recursion (see, e.g., Polatoglu and Sahin 2000). Thowsen (1975) was the first to derive a BSLP policy in a multiperiod model. He considers a linear demand with an additive random component that follows $PF_2$ distribution and proves that the optimal price is weakly decreasing in the inventory. Young (1979) uses a more general demand function by extending his earlier work (Young 1978) to multiperiod model. He analyzes a stationary infinite horizon problem and assumes that the initial inventory is below the optimal base-stock level. Under this assumption, he obtains an optimal stationary BSLP policy. However, his analysis does not address the case when the initial inventory level is above the optimal base-stock level. Moreover, Young (1978, 1979) assumes that the random component in the demand follows $PF_2$ distribution. Federgruen and Heching (1999) revisit the model analyzed by Young (1979) but allow for a general demand noise distribution. They impose certain conditions on the scale and location parameters of the demand distribution, the realized demand as a function of the price, and the expected inventory holding and backorder cost function to ensure the optimality of a BSLP policy. As we demonstrate in our analysis, these conditions

can fail easily when the demand is a nonlinear function of the price. Recently, Chen and Zhang (2010) proved the optimality of a BSLP policy for the additive demand model under the assumptions that the expected revenue is quasi-concave in price, and the random demand component has a log-concave distribution. Our analysis suggests that these assumptions can be relaxed. We also adopt the general demand function used by Young (1978) and Federgruen and Heching (1999). We identify a new set of optimality conditions that depend only on the location and scale parameters of the demand. Because these parameters are deterministic, scalar functions of the price, it is easy to verify the optimality conditions.

There are several papers adopting an alternative demand model to generalize the additive and multiplicative models. Under this alternative model, there are two independent random components, one is additive to the mean demand, and the other is multiplicative to the mean demand. With this model, the inventory balancing equation is linear in the mean demand, which entails the joint convexity of the expected inventory cost and allows one to apply known techniques to derive the optimal policy. With this demand model, Chen and Simchi-Levi (2004a, b) and Huh and Janakiraman (2008) prove the optimality of an $(s, S, p)$ or $(s, S, A, P)$ policy when a fixed ordering cost is considered, and Feng (2010) and Feng and Shi (2012) derive a threshold policy for inventory–pricing decisions when supply capacities are uncertain. The major disadvantage of this demand model is that the price impact on the scale and location of the demand are inherently connected, which may restrict the applicability of the model. For example, the coefficient of variation for the demand is always weakly increasing in the price under this demand model. The demand model adopted in our study overcomes such limitation.

## 3. The Model

The firm faces a planning problem over $T$ periods. At the beginning of period $t$, $1 \leq t \leq T$, the (net) inventory level $x_t$ is reviewed. The firm needs to determine the order quantity $q_t$ to be received in period $t$ and the product price $p_t$ for an uncertain demand to be realized at the end of period $t$. The unmet demand is backordered, and the leftover inventory is carried over to the next period.

The firm pays $c_t$ dollars for each unit received. It also pays a surplus/shortage cost $H_t(\cdot)$ upon the demand realization. We assume that $H_t(\cdot) \geq 0$ is continuous and convex. (Thus, $H_t(\cdot)$ has at most a finite number of nondifferentiable points.) Also, $H_t(0) = 0$, $|H_t(x_1) - H_t(x_2)| \leq c^H |x_1 - x_2|$ for some positive and finite $c^H$, and $\lim_{x \to \pm\infty} H_t(x) = \infty$.

We assume that customers pay the current price $p_t$ in period $t$. This price may not be the same as the price quoted at the time when the customer actually obtains the product if the demand is backordered. In other words, the delay in filling the order does not affect the price charged. Instead, the manager pays a penalty cost for the backorder via $H_t(\cdot)$, which includes the loss of goodwill or compensation to customers. Such a pricing scheme is widely used in studying combined pricing and inventory decisions (e.g., Federgruen and Heching 1999, Chen and Simchi-Levi 2004a, Li and Zheng 2006). We consider a two-parameter demand function:

$$D_t(p_t, \epsilon_t) = \mu_t(p_t) + \sigma_t(p_t)\epsilon_t, \quad p_t \in [\underline{p}_t, \bar{p}_t], \quad (1)$$

where $\mu_t(\cdot) > 0$ and $\sigma_t(\cdot) \geq 0$ are deterministic functions of price $p_t$, and $\epsilon_t$ is a random variable with mean zero and standard deviation one. We assume that $\mu_t(\cdot)$ and $\sigma_t(\cdot)$ are finite, continuous, and twice differentiable; $\mu_t(p_t)$ is weakly decreasing in $p_t$, and $\sigma_t(p_t)$ is strictly monotone in $p_t$. In the statistics literature, this model is termed the *generalized additive model*. Since $\epsilon_t$ is a unit random variable, $\mu_t(p_t)$ and $\sigma_t(p_t)$ are, respectively, the location and scale parameters of the demand. To rule out unrealistic cases, we also assume $D_t(p_t, \epsilon_t) \geq 0$ for any $p_t \in [\underline{p}_t, \bar{p}_t]$ and $\underline{p}_t \geq c_t$.

The demand function defined in (1) provides more modeling flexibility than the one-parameter (additive or multiplicative) models. Note that when $\sigma_t(\cdot)/\mu_t(\cdot)$ is constant, (1) reduces to the multiplicative demand model. The assumption that $\sigma_t(\cdot)$ is strictly monotone leaves out the case of additive demand model, i.e., $\sigma_t(\cdot) = 0$. However, as our analysis unfolds, it is clear that a similar approach as the one developed in our derivation can be applied to handle this case (see our discussion in §4.4). The GAM covers most commonly used demand distributions including Exponential, Normal, Logistic, Weibull, Gamma, Laplace, Uniform, Cauchy, and Extreme Value distributions.

To formulate the dynamic pricing and inventory problem, let $V_t(x_t)$ denote the maximum expected profit given the inventory level $x_t$ at the beginning of period $t$. For $1 \leq t \leq T$, the optimality equation can be written as

$$V_t(x_t) = \max_{\substack{y_t \geq x_t, \\ \underline{p}_t \leq p_t \leq \bar{p}_t}} \left\{ \pi_t(x_t, y_t, p_t) \right.$$

$$\left. + \rho \mathsf{E}[V_{t+1}(y_t - D_t(p_t, \epsilon_t))] \right\}, \quad (2)$$

where $\pi_t(x_t, y_t, p_t) = p_t \mu_t(p_t) - c_t(y_t - x_t) - \mathsf{E}[H_t(y_t - D_t(p_t, \epsilon_t))]$ is the instantaneous profit obtained in period $t$ under the decision $(p_t, y_t)$, and $\rho \in [0, 1]$ is the discount factor. We also impose the terminal condition $V_{T+1}(x_{T+1}) = 0$ for all $x_{T+1}$.

## 4. The Optimality of a BSLP Policy

It is well known that the optimal policy for the problem defined in (2) can be very complex, because the objective function may have multiple local maxima in general. Our goal is to characterize general conditions under which the optimal policy for the problem defined in (2) exhibits a simple structure. Note that the one-period profit function $\pi_t(x_t, y_t, p_t)$ is separable in $x_t$ and $(y_t, p_t)$. Thus, the optimality of a BSLP policy can be easily established if $\pi_t(x_t, y_t, p_t)$ is jointly concave in $(y_t, p_t)$. Naturally, one would seek conditions under which the expected revenue $p_t \mathsf{E}[D_t(p_t, \epsilon_t)]$ is concave in $p_t$ and the expected inventory cost $\mathsf{E}[H_t(y_t - D_t(p_t, \epsilon_t))]$ is jointly convex in $(y_t, p_t)$. This is a common approach taken in the literature.

In their highly cited paper, Federgruen and Heching (1999) make three key assumptions to ensure the concavity of $\pi_t(x_t, y_t, p_t)$:

(FH-i) $\mu_t(p_t)$ and $\sigma_t(p_t)$ are weakly decreasing functions,

(FH-ii) $D_t(p_t, \epsilon_t)$ is weakly decreasing and concave in $p_t$ for any realization of $\epsilon_t$, and

(FH-iii) $\mathsf{E}[H_t(y_t - D_t(p_t, \epsilon_t))]$ is jointly convex in $(y_t, p_t)$.

Condition (FH-i) is generally easy to check. Condition (FH-ii) appears restrictive, because it is required for each realization of $\epsilon_t$. Since the random component $\epsilon_t$ can take either positive or negative values, (FH-ii) is satisfied if and only if $\sigma_t(p_t)$ is linear. Verifying condition (FH-iii) can be nontrivial because it involves integration, and the expression of the two-dimensional function can be complex. Also this condition is restrictive—it easily fails when the demand exhibits a nonlinear relation to the price. To see that, consider the following example with deterministic demand.

EXAMPLE 1 (DETERMINISTIC DEMAND). Suppose $D_t(p_t, \epsilon_t) = \mu_t(p_t)$ for $p_t \in [\underline{p}_t, \bar{p}_t]$ and $H_t(x) = h_t x^+ + b_t x^-$. To apply the result in Federgruen and Heching (1999) to this example, we must have that $\mu_t(p_t)$ is weakly decreasing and concave in $p_t$ so that (FH-i) and (FH-ii) hold, and that the inventory cost

$$\mathsf{E}[H_t(y_t - D_t(p_t, \epsilon_t))] = \begin{cases} h_t[y_t - \mu_t(p_t)] & \text{if } y_t \geq \mu_t(p_t), \\ b_t[\mu_t(p_t) - y_t] & \text{if } y_t < \mu_t(p_t), \end{cases}$$

is jointly convex in $(y_t, p_t)$ so that (FH-iii) holds. Note that these conditions are satisfied *only* when $\mu_t(p_t)$ is linear. In other words, the results obtained in Federgruen and Heching (1999) do not apply to commonly used nonlinear demand functions.

In what follows, we identify an alternative set of optimality conditions for a BSLP policy by transforming the decision variables of the problem. As our analysis unfolds, it becomes clear that these conditions

are more general and easier to verify than the ones imposed by Federgruen and Heching (1999).

### 4.1. Decision Variable Transformation and the BSLP Policy

In this section, we transform the problem defined in (2) into an equivalent optimization problem via a change of decision variables. Because $\sigma_t(\cdot)$ is strictly monotone, we can define

$$g_t(\tau_t) \equiv \sigma_t^{-1}(\tau_t) \quad \text{and} \quad z_t \equiv y_t - \mu_t(g_t(\tau_t)). \quad (3)$$

The function $g_t(\cdot)$ is the inverse of $\sigma_t(\cdot)$; i.e., $\tau_t = \sigma_t(p_t)$ implies $p_t = g_t(\tau_t)$. Thus, $\tau_t$ is a measure of the *risk level* involved in the demand function. Let $\bar{\tau}_t = \sup\{\sigma_t(p_t): p_t \in [\underline{p}_t, \bar{p}_t]\}$ and $\underline{\tau}_t = \inf\{\sigma_t(p_t): p_t \in [\underline{p}_t, \bar{p}_t]\}$. Because $\mu_t(p_t)$ and $\sigma_t(p_t)$ are monotone in $p_t$, we can express the mean demand as a function of $\tau_t$, i.e., $\mu_t(g_t(\tau_t)) = \mu_t(\sigma_t^{-1}(\tau_t)) = \mu_t(p_t)$. The quantity $z_t$ represents the *safety stock*, which is the amount of inventory prepared in addition to the average demand.

The problem defined in (2) can be transformed into one of finding an optimal safety stock $z_t$ and an optimal risk level $\tau_t$. Let $\bar{V}_t(x_t) = V_t(x_t) - c_t x_t$. Then, (2) can be rewritten as

$$\bar{V}_t(x_t) = \max_{(z_t, \tau_t) \in \Omega_t(x_t)} J_t(z_t, \tau_t), \quad (4)$$

where

$$J_t(z_t, \tau_t) = R_t(g_t(\tau_t)) - (c_t - \rho c_{t+1})z_t - \mathsf{E}[H_t(z_t - \tau_t \epsilon_t)] + \rho \mathsf{E}[\bar{V}_{t+1}(z_t - \tau_t \epsilon_t)], \quad (5)$$

$$\Omega_t(x_t) = \{(z_t, \tau_t): z_t + \mu_t(g_t(\tau_t)) \geq x_t, \ \tau_t \in [\underline{\tau}, \bar{\tau}]\}, \quad (6)$$

$$R_t(g) = (g - c_t)\mu_t(g). \quad (7)$$

In the above, $R_t(g_t(\tau_t))$ is the riskless profit function, and $\Omega_t(x_t)$ is the set of feasible decisions. Note that the expected inventory cost function $\mathsf{E}[H_t(z_t - \tau_t \epsilon_t)]$ is a composition of convex function and linear function of $(z_t, \tau_t)$ and is thus jointly convex in $(z_t, \tau_t)$. Consequently, establishing the joint concavity of $J_t(z_t, \tau_t)$ boils down to verifying the concavity of the function $R_t(g_t(\tau_t))$ and the convexity of the set $\Omega_t(x_t)$.

LEMMA 1 (CONVEX FEASIBLE SET). *If $\mu_t(g_t(\tau_t))$ is concave in $\tau_t$, $\Omega_t(x_t)$ is convex for each $x_t$.*

The convexity of the feasible set $\Omega_t(x_t)$ ensures the existence of an optimal solution for the optimization problem defined in (4)–(7). When $\Omega_t(x_t)$ is not convex, a common approach is to convexify the decision space by allowing for randomized policy (see, e.g., Beutler and Ross 1985, Feinberg 1994). With Lemma 1, we are now ready to derive an optimal BSLP policy for the transformed problem defined in (4)–(7).

THEOREM 2 (AN OPTIMAL BSLP POLICY). *For $t \in \{1, 2, \ldots, T\}$, when both $R_t(g_t(\cdot))$ and $\mu_t(g_t(\cdot))$ are concave, we have the following:*

(i) *The function $J_t(z_t, \tau_t)$ is jointly concave in $z_t$ and $\tau_t$.*

(ii) *The value function $\bar{V}_t(x_t)$ is nonincreasing and concave in $x_t$.*

(iii) *There exist a base-stock level $\hat{y}_t$ and a list price $\hat{p}_t$. If $x_t \leq \hat{y}_t$, it is optimal to order up to $\hat{y}_t$ and charge $\hat{p}_t$; otherwise, it is optimal not to order and there exists an optimal price $p_t^*(x_t)$.*

According to Theorem 2, when $R_t(g_t(\cdot))$ and $\mu_t(g_t(\cdot))$ are concave, the objective function and the value function of the DP are concave, which in turn leads to the optimal BSLP policy. Under this policy, it is optimal to order up to the base-stock level $\hat{y}_t$ and charge the list price $\hat{p}_t$ when the inventory level is below the base-stock level. When the inventory level is above the base-stock level, no order is issued, and the optimal price is a function of the inventory level.

Our next task is to identify sufficient conditions under which $R_t(g_t(\cdot))$ and $\mu_t(g_t(\cdot))$ are concave. This is done in §4.2. Another key feature of the BSLP policy described in Theorem 2 is that it does not necessarily require the optimal price $p_t^*(x_t)$ to be weakly decreasing in the inventory level $x_t$. Indeed, as we show in §4.3, it is possible that $p_t^*(x_t)$ is not monotone in $x_t$, which makes an interesting contrast to earlier findings (e.g., Thowsen 1975, Amihud and Mendelson 1983, Federgruen and Heching 1999).

### 4.2. Sufficient Conditions for an Optimal BSLP Policy

Direct verification of the concavity of $R_t(g_t(\tau_t))$ and $\mu_t(g_t(\tau_t))$ may not be straightforward because the inverse $g_t(\tau_t)$ of the scale parameter may not yield a closed form. The next theorem characterizes sufficient conditions that are expressed in terms of $\mu_t(p_t)$ and $\sigma_t(p_t)$ and are thus easy to verify.

THEOREM 3 (SUFFICIENT CONDITIONS). *The sufficient conditions for an optimal BSLP policy are*

$$\mu_t'(p_t)\frac{\sigma_t''(p_t)}{\sigma_t'(p_t)} \geq \mu_t''(p_t), \quad (8)$$

$$\mu_t(p_t)\frac{\sigma_t''(p_t)}{\sigma_t'(p_t)} \geq 2\mu_t'(p_t). \quad (9)$$

The conditions (8) and (9) ensure the concavity of $\mu_t(g_t(\tau_t))$ and $R_t(g_t(\tau_t))$, respectively. These conditions depend only on the one-dimensional functions $\mu_t(\cdot)$ and $\sigma_t(\cdot)$. As long as one knows the form of these functions, it is straightforward to check whether the above conditions hold or not. In sharp contrast to those derived from previous studies (e.g. Young 1978, Federgruen and Heching 1999, Kocabıyıkoğlu and Popescu 2011), the optimality conditions stated

above is *independent* of the cost parameters and the distribution of the random variable $\epsilon_t$.

To demonstrate the result in Theorem 3, we further consider two special cases of one-parameter demand models and several examples of commonly used demand functions.

Corollary 4 (One-Parameter Demand Models). *A BSLP policy is optimal when*

(i) (Multiplicative model) $\sigma_t(p_t)/\mu_t(p_t) = \eta$ *for* $p_t \in [\underline{p}_t, \bar{p}_t]$ *and* $2[\sigma_t'(p_t)]^2 \geq \sigma_t''(p_t)\sigma_t(p_t)$;

(ii) (Constant mean model) $\mu_t(p_t) = \mu_t$ *for* $p_t \in [\underline{p}_t, \bar{p}_t]$ *and* $\sigma_t(p_t)$ *is either decreasing concave or increasing convex.*

The examples below illustrate the conditions characterized in Corollary 4. As an important observation from these examples, a BSLP policy can be optimal when the demand function exhibits highly nonlinear response to price changes.

Example 2 (Multiplicative Demand). Suppose $D_t(p_t, \epsilon_t) = \sigma_t(p_t)(\eta + \epsilon_t)$ for $p_t \in [\underline{p}_t, \bar{p}_t]$. By Corollary 4, a BSLP policy is optimal in the following examples:

1. Linear-power function: $\sigma_t(p_t) = (\alpha - \beta p_t)^\gamma$, $\alpha > 0$, and $\beta > 0$.

2. Exponential function: $\sigma_t(p_t) = \alpha e^{-\beta p_t}$, $\alpha > 0$, and $\beta > 0$.

3. Isoelastic function: $\sigma_t(p_t) = \alpha p_t^{-\beta}$, $\alpha > 0$, and $\beta > 1$.

Example 3 (Constant Mean Demand). Suppose $D_t(p_t, \epsilon_t) = \mu_t + \sigma_t(p_t)\epsilon_t$ for $p_t \in [\underline{p}_t, \bar{p}_t]$. By Corollary 4, a BSLP policy is optimal in the following examples:

1. Linear-power function: $\sigma_t(p_t) = (\alpha - \beta p_t)^\gamma$, $\alpha > 0$, and $\beta > 0$.

2. Exponential function: $\sigma_t(p_t) = \alpha e^{-\beta p_t}$, $\alpha > 0$, and $\beta < 0$.

3. Isoelastic function: $\sigma_t(p_t) = \alpha p_t^{-\beta}$, $\alpha > 0$, and $-1 < \beta < 0$.

From Examples 2 and 3, we observe that our result can be applied to both concave (e.g., linear-power function with $\gamma < 1$ or isoelastic function with $-1 < \beta < 0$) and convex (e.g., linear-power function with $\gamma > 1$, exponential function, or isoelastic function with $\beta > 0$) demand functions commonly studied in the literature. Moreover, unlike the previous studies on inventory-based pricing models, our analysis allows the possibility that the standard deviation of the demand is increasing in the price (e.g., exponential function with $\beta < 0$ or isoelastic function with $-1 < \beta < 0$).

### 4.3. Monotone Property of the Optimal Price
The statement of the optimal BSLP policy in Theorem 2 is silent about whether the optimal price curve

$p_t^*(x_t)$ is monotone in the inventory level $x_t$. In this section, we first characterize conditions for a monotone price curve and then discuss cases where such monotonicity fails. For that, we need the following definition of the demand elasticity.

Definition 1 (Elasticity of Demand Function). The price elasticities of $\mu_t(p_t)$ and $\sigma_t(p_t)$ are defined, respectively, as $e_t^\mu(p_t) = -p_t\mu_t'(p_t)/\mu_t(p_t)$ and $e_t^\sigma(p_t) = -p_t\sigma_t'(p_t)/\sigma_t(p_t)$.

Theorem 5 (Monotonicity of Optimal Price). *The optimal price* $p_t^*(x_t)$, $x_t \geq \hat{y}_t$, *is weakly decreasing in* $x_t$ *when*

(i) $R_t(g_t(\tau_t))$ *and* $\mu_t(g_t(\tau_t))$ *are concave in* $\tau_t$, *and*

(ii) $\sigma_t(p_t)$ *is weakly decreasing in* $p_t$ *and* $e_t^\mu(p_t) \geq e_t^\sigma(p_t)$ *for* $p_t \in [\underline{p}_t, \bar{p}_t]$.

According to Theorem 5, a sufficient condition for a monotone price curve is that the location of the demand is more elastic to the price than the scale of the demand. In this case, the realized demand (for each value of $\epsilon_t$) is always weakly decreasing in price. As a result, it is optimal to set a low price to speed up inventory turnover and reduce holding cost when the inventory level is high. The next corollary suggests that such a situation easily arises in one-parameter demand models.

Corollary 6 (One-Parameter Demand Models). *The optimal price* $p_t^*(x_t)$, $x_t \geq \hat{y}_t$, *is weakly decreasing in* $x_t$ *when* $R_t(g_t(\tau_t))$ *and* $\mu_t(g_t(\tau_t))$ *are concave in* $\tau_t$, *and*

(i) (Multiplicative model) $\sigma_t(p_t)/\mu_t(p_t) = \eta$ *for* $p_t \in [\underline{p}_t, \bar{p}_t]$ *or*

(ii) (Constant mean model) $\mu_t(p_t) = \mu_t$ *for* $p_t \in [\underline{p}_t, \bar{p}_t]$ *and* $\sigma_t(p_t)$ *is weakly decreasing in* $p_t$.

When the price change leads to a higher response in the scale of the demand than in the location of the demand, the optimal price may not be monotone in the inventory level. This is demonstrated via an example in Figure 1. In this example, the price elasticity of $\mu(p)$ is 3 and that of $\sigma(p)$ is 5. We observe that the optimal price is increasing in the inventory level over the interval $[5, 43]$. This observation is in sharp contrast to previous studies on price–inventory problems, which always find that the optimal price is weakly decreasing in the inventory level. An exception is Feng (2010), who shows that the optimal price may not be monotone in the inventory level when the suppliers are unreliable. In a similar spirit to our result, her observation suggests that the price interaction with the demand variability has a significant impact on the optimal policy.

### 4.4. The Additive Demand Case
For the purpose of comparison, we include the analysis of additive demand model in this section. This

**Figure 1    Example in Which the Optimal Price Is Not Monotone in the Inventory Level**
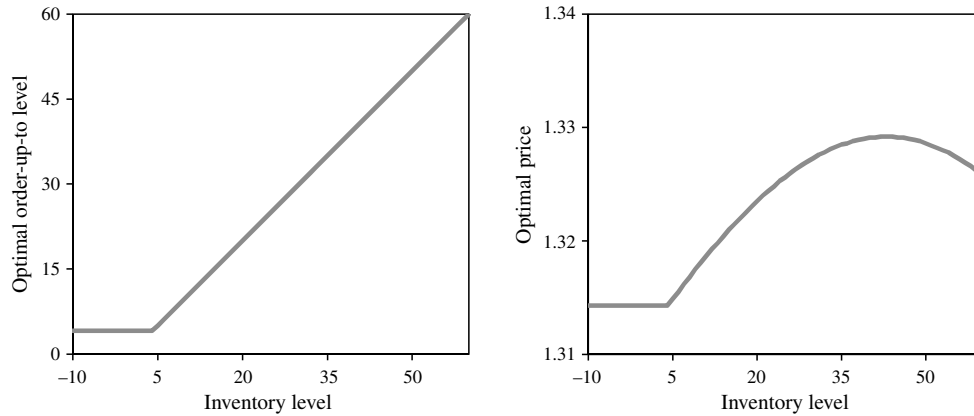


*Note.* $T = 1$, $c = 0.5$, $h = 0.05$, $b = 0.64$, $\mu(p) = 250 p^{-3}$, $\sigma(p) = 500 p^{-5}$, and $\epsilon \sim$ truncated $N(0, 1)$.

demand model violates the assumption that $\sigma_t(\cdot)$ is strictly monotone. However, this case can be analyzed using a method similar to the one presented above (i.e., by taking $s_t = \mu_t(p_t)$ and transforming the problem into one of deciding $s_t$ and $z_t$). This explains an observation from the next theorem that the condition derived for the additive demand model has a similar form to that for the multiplicative demand model in Corollary 4(i).

THEOREM 7 (ADDITIVE DEMAND MODELS). *Suppose* $\sigma_t(p_t) = \sigma_t$ *for* $p_t \in [\underline{p}_t, \bar{p}_t]$. *A BSLP policy is optimal and the optimal price* $p_t^*(x_t)$ *is weakly decreasing in* $x_t$ *when* $2[\mu_t'(p_t)]^2 \geq \mu_t''(p_t)\mu_t(p_t)$.

EXAMPLE 1 REVISITED (DETERMINISTIC DEMAND). Suppose $D_t(p_t, \epsilon_t) = \mu_t(p_t)$ for $p_t \in [\underline{p}_t, \bar{p}_t]$ and $H_t(x) = h_t x^+ + b_t x^-$. By Theorem 7, a BSLP policy is optimal in the following examples.

1. Linear-power function: $\mu_t(p_t) = (\alpha - \beta p_t)^\gamma$, $\alpha > 0$, and $\beta > 0$.

2. Exponential function: $\mu_t(p_t) = \alpha e^{-\beta p_t}$, $\alpha > 0$, and $\beta > 0$.

3. Isoelastic function: $\mu_t(p_t) = \alpha p_t^{-\beta}$, $\alpha > 0$, and $\beta > 1$.

## 5.    Statistical Estimation and Validation

In this section, we demonstrate how to estimate the demand parameters and verify the conditions derived in §4 using a retail data set. The objective is twofold. First, we quantify the benefit of using the two-parameter GAM over using one-parameter models. Second, we validate the technical conditions required for the optimality of a BSLP policy.

### 5.1.    Data

We use the data set from the Dominick's Finer Foods project conducted by the University of Chicago's Kilts Center. This database contains store-level pricing and sales data from Dominick's Finer Foods, one of the two biggest supermarket chains in the Chicago area. To demonstrate our results, we pick a soft drink product, 7UP, that has more than 100 different price points and seven-year sales data in 83 stores. Note that soft drink products typically exhibit a high brand loyalty (see, e.g., Cotterill et al. 1996, Dubé 2004, 2005), and thus the product demand is much more sensitive to its own price than to competing products' prices. For example, a regular 7UP drinker is unlikely to buy Coke just because the latter is on promotion. Moreover, it is commonly agreed in the retail industry that sales on soft drink products are highly price dependent. These product characteristics suggest that it is appropriate to apply our model to this product category.

When the data were collected, Dominick's Finer Foods priced products by 15 sales zones. Within each zone, a uniform price was set for all stores. We use zone-specific, instead of store-specific, explanatory variables to capture the heterogeneity among customers in different market areas. This is justified by the observation from Hoch et al. (1995), whose analysis on the same data set suggests that the variance in consumer characteristics across different market areas is substantially greater than that within each market area. After an initial variable selection using a generalized linear model (GLM), we narrowed our variables down to eight demographic variables, $\vec{X} = \{X_1, \ldots, X_8\}$, as listed in Table 1. Therefore, a sample observation contains values of the dependent variable (demand $D$) and the explanatory variables ($\vec{X}$ and $p$). The data set contains nearly 28,000 observations.

To remove the impact of outliers on the statistical modeling, we cap demand at its 99th percentile. We also remove samples with zero demand, because there is no price information for these observations. Histograms of demand at different price points and

**Table 1    Demographic Variables**

| | | |
|---|---|---|
| $X_1$ | EDUAT | % of college graduates |
| $X_2$ | RETIRED | % of retired |
| $X_3$ | HSIZEAVG | Average household size |
| $X_4$ | INCOME | Log of median income |
| $X_5$ | WORKWOM | % working women with full-time jobs |
| $X_6$ | HHVAL200 | % of households with value over \$200,000 |
| $X_7$ | MORTGAGE | % of households with mortgages |
| $X_8$ | SHOPINDX | Ability to shop (car and single family house) |

normality tests (e.g., Kolmogorov–Smirnov test) suggest that the demand distribution is approximately normal. Therefore, we set $\epsilon_t$ to be a normal random variable with mean zero and standard deviation one in our estimation. Note that the approach developed below can apply to any general distribution for $\epsilon_t$.

### 5.2. Generalized Additive Model with Two Parameters

The GAM, being a major development in statistics in the last two decades, extends the widely used generalized linear model. The GLM is considered an enhancement of the linear least square model. The latter applies only to normal distributions, whereas the former applies to the exponential family (e.g., Normal, Log-Normal, Gamma, Poisson, and Multinomial). The GLM estimates the mean through a linear or nonlinear link function and uses an associated nonlinear function of mean to estimate the variance. This method, to a certain degree, allows for fitting nonlinear and nonconstant variance (McCullagh and Nelder 1989). Use Gamma distribution as an example. The estimation of the mean $\mu$ is through a log-link function $\log(\mu) = \vec{\beta}\vec{X}$, where $\vec{X}$ is the vector of explanatory variables, and $\vec{\beta}$ is the corresponding coefficient vector. The estimation of variance is through an associated function $\nu(\mu) = \mu^2$. Different distributions have their own link functions and association functions (see, e.g., McCullagh and Nelder 1989). A disadvantage of this method lies in its restriction on the relationship between the mean and variance, which can be overcome by using a two-parameter model like the GAM. The GAM extends the GLM by using additive smoothing functions of explanatory variables, e.g., P-splines, cubic splines, or loess smoothing, to model the nonlinear relationship between the response variable and explanatory variables (e.g., Hastie and Tibshirani 1990, Hastie et al. 2001, Stasinopoulos and Rigby 2005). By employing the GAM, our model can take parameters, $\mu(p)$ and $\sigma(p)$, and a general distribution of the random component $\epsilon$.

To estimate our demand model, we use the following monotone link functions of $\mu$ and $\sigma$:

$$\nu_1(\mu) = \vec{\beta}_1\vec{X} + S_1(p) \quad \text{and} \quad \nu_2(\sigma) = \vec{\beta}_2\vec{X} + S_2(p), \quad (10)$$

where $\vec{X}$ is the vector of independent variables, which exhibit linear relations in the link functions of $\mu$ and $\sigma$; $\vec{\beta}_1$ and $\vec{\beta}_2$ are the corresponding coefficient vectors; and the additive functions $S_1(\cdot)$ and $S_2(\cdot)$ are smoothing functions of the price $p$. With the relation in (10), the mean and standard deviation as functions of $\{\vec{X}, p\}$ can be written as $\mu(p; \vec{X}) = \nu_1^{-1}[\vec{\beta}_1\vec{X} + S_1(p)]$ and $\sigma(p; \vec{X}) = \nu_2^{-1}[\vec{\beta}_2\vec{X} + S_2(p)]$, respectively.

### 5.3. Constrained Maximum Likelihood Estimation

Let $f(\cdot)$ denote the probability density function of the random demand component $\epsilon$. Our goal is to find the best fit for the functions $\mu(p; \vec{X})$ and $\sigma(p; \vec{X})$ based on the sample data. In the meantime, the estimated functions should satisfy the conditions in Theorem 3. For that, we formulate the CMLE problem as follows:

$$\max \sum_{j \in \mathcal{J}}\left[\log\left\{f\left(\frac{D_j - \mu(p; \vec{X}_j)}{\sigma(p; \vec{X}_j)}\right)\right\}\right] \quad (11)$$

$$\text{s.t.} \quad \mu'(p; \vec{X}_j)\sigma''(p; \vec{X}_j)$$
$$- \mu''(p; \vec{X}_j)\sigma'(p; \vec{X}_j) \leq 0, \quad \forall j, \quad (12)$$

$$2\mu'(p; \vec{X}_j)\sigma'(p; \vec{X}_j)$$
$$- \mu(p; \vec{X}_j)\sigma''(p; \vec{X}_j) \geq 0, \quad \forall j, \quad (13)$$

$$\mu'(p; \vec{X}_j) \leq 0, \sigma'(p; \vec{X}_j) \leq 0, \quad \forall j, \quad (14)$$

where $\{D_j, \vec{X}_j\}$ is the $j$th sample and $\mathcal{J}$ is the set of sample indices. The objective in (11) is defined for log-likelihood estimation. Constraints (12) and (13) correspond to conditions (8) and (9), respectively. By checking whether or not these constraints are binding, we can directly verify the optimality conditions for a BSLP policy from the data.

The optimization problem defined in (11)–(14) serves as a baseline model for the demand estimation. It can be used to estimate stationary demand parameters for a product that exhibits strong product-specific price elasticity but weak differentiation in its responses to price change across different sales zones. However, preliminary statistical analysis on our data set suggests a significant difference among sales zones and strong nonstationarity. Therefore, the above estimation model needs to be refined appropriately, which is discussed in detail in §§5.3.1 and 5.3.2.

**5.3.1. Zone-Specific Price Elasticity.** We find the isoelastic function a good fit for this data set. Moreover, the price elasticities across sales zones exhibit significant heterogeneity due to the different demographics characteristics. The zone-specific demand is expressed as

$$\mu_i(p; \vec{X}) = e^{\vec{\beta}_1\vec{X}}p^{-(\beta_\mu + \lambda_{\mu i})} \quad \text{and}$$
$$\sigma_i(p; \vec{X}) = e^{\vec{\beta}_2\vec{X}}p^{-(\beta_\sigma + \lambda_{\sigma i})}, \quad (15)$$

where $i \in \mathcal{I}$ is the zone index, and $\lambda_{\mu i}$ and $\lambda_{\sigma i}$ are the zone-specific price elasticities. In this model, the response of demand to price change is reflected by two factors: $\{\beta_\mu, \beta_\sigma\}$ captures the product-specific response, and $\{\lambda_{\mu i}, \lambda_{\sigma i}\}$ captures the effect of demographics on demand–price response.

We introduce the indicator $Z_i$, $i = 1, \ldots, 15$, as a dummy variable (i.e., $Z_i = 1$ if the sample is from zone $i$ and $Z_i = 0$ otherwise) and rewrite the CMLE problem defined in (11)–(14) as

$$\max\left\{-\sum_{j\in\mathcal{J}}\left[\vec{\beta}_2\vec{X} - \left(\beta_\sigma + \sum_{i\in\mathcal{I}}\lambda_{\sigma i}Z_i\right)\log(p)\right.\right.$$
$$\left.\left. + \frac{[D_j - e^{\vec{\beta}_1\vec{X} - (\beta_\mu + \sum_{i\in\mathcal{I}}\lambda_{\mu i}Z_i)\log(p)}]^2}{2(e^{\vec{\beta}_2\vec{X} - (\beta_\sigma + \sum_{i\in\mathcal{I}}\lambda_{\sigma i}Z_i)\log(p)})^2}\right]\right\} \quad (16)$$

$$\text{s.t.} \quad (\beta_\sigma + \lambda_{\sigma i}) - (\beta_\mu + \lambda_{\mu i}) \geq 0, \quad \forall i, \quad (17)$$

$$2(\beta_\mu + \lambda_{\mu i}) - 1 - (\beta_\sigma + \lambda_{\sigma i}) \geq 0, \quad \forall i, \quad (18)$$

$$(\beta_\mu + \lambda_{\mu i}) \geq 0, \quad \forall i. \quad (19)$$

Note that we shall set $Z_{15} = 1 - \sum_{j=1}^{14} Z_j$ and $\lambda_{15} = 0$ to avoid collinearity. Constraints (17) and (18) correspond to the concavity conditions in Theorem 3 for the isoelastic demand model. These constraints are imposed for each sales zone. Thus, the optimization leads to one demand function for each sales zone. The problem defined above allows for multizone estimation for a single product with stationary demand. In the next section, we further extend this formulation to account for nonstationarity.

We shall remark that in the case when one cannot identify a specific functional form for a good fit, a common method is to use cubic or quadratic spline to model $\mu(p; \vec{X})$ and $\sigma(p; \vec{X})$ by assuming both functions have the same basis functions and the same knots. Then, the constraints in the maximization problem reduce to the constraints on the coefficients of the basis functions.

**5.3.2. Nonstationarity.** Demands for retail products usually exhibit strong seasonality, which is also observed from the data set. For example, product consumptions typically rise in weeks with major holidays or events (e.g., Labor Day, Memorial Day, Superbowl weekend) and decrease dramatically in the subsequent weeks. One may use 52 dummy variables to model the seasonality to capture the difference across weeks. However, with only one product, such an approach typically leads to overfitting (Fok et al. 2007), because the coefficients of dummy variables would pick up most of the demand variance. To avoid the overfitting issue, we adopt the method described by Brockwell and Davis (2002). Specifically, we use a

week index $t$ ranging from 1 to 52 and a seasonality index $T = 52$. The time influence on the mean and standard deviation can be defined as

$$W_\mu(t) = \alpha_\mu t + \sum_{k=1}^{K}[\beta_{\mu k}^s \sin(2\pi kt/T) + \beta_{\mu k}^c \cos(2\pi kt/T)],$$

$$W_\sigma(t) = \alpha_\sigma t + \sum_{k=1}^{K}[\beta_{\sigma k}^s \sin(2\pi kt/T) + \beta_{\sigma k}^c \cos(2\pi kt/T)],$$

where $K$ is an integer. A large $K$ provides a good fit while making the model less parsimonious. One typically experiments with the data to find the best value of $K$. The terms $\alpha_\mu t$ and $\alpha_\sigma t$ capture the trend in demand parameters, and the sine and cosine terms capture the seasonality. Consequently, the demand parameters defined in (15) are modified to

$$\mu_i(p; \vec{X}) = e^{\vec{\beta}_1\vec{X} + W_\mu(t)}p^{-(\beta_\mu + \lambda_{\mu i})} \quad \text{and}$$
$$\sigma_i(p; \vec{X}) = e^{\vec{\beta}_2\vec{X} + W_\sigma(t)}p^{-(\beta_\sigma + \lambda_{\sigma i})}. \quad (20)$$

The CMLE problem becomes

$$\max\left\{-\sum_{j\in\mathcal{J}}\left[\vec{\beta}_2\vec{X} + W_\sigma(t) - \left(\beta_\sigma + \sum_{i\in\mathcal{I}}\lambda_{\sigma i}Z_i\right)\log(p)\right.\right.$$
$$\left.\left. + \frac{[D_j - e^{\vec{\beta}_1\vec{X} + W_\mu(t) - (\beta_\mu + \sum_{i\in\mathcal{I}}\lambda_{\mu i}Z_i)\log(p)}]^2}{2(e^{\vec{\beta}_2\vec{X} + W_\sigma(t) - (\beta_\sigma + \sum_{i\in\mathcal{I}}\lambda_{\sigma i}Z_i)\log(p)})^2}\right]\right\}, \quad (21)$$

subject to constraints (17)–(19).

**5.4. Results**
We apply the CMLE optimization to the data set to obtain the estimates in Table 2 for the three different demand models, namely, the GAM, an additive demand model (i.e., with $\sigma$ estimated as a constant), and a multiplicative demand model (i.e., with $\eta = \sigma(p; \vec{X})/\mu(p; \vec{X})$ estimated as a constant). We use the observations in the first six years as the training data set and the observations in the seventh year as the validation data set. The demand models are estimated from the training data set and validated using the validation data set. For each price, we score the validation data set, obtain predictions of the three demand models, and then calculate the sum of mean squared errors (SMSE) for each model. The SMSEs provide the performance comparison of three models.

From Table 2, we observe that both the product-specific and zone-specific price elasticities are significant. However, the magnitude of product-specific response is larger than that of zone-specific response. This is reasonable as most people have a similar "price expectation" for soft drink products.

It is straightforward to verify that the constraints in (17) and (18) are satisfied with strict inequality for 13 zones and they fail in zones 1 and 14. This suggests that the estimates are unconstrained maximizers

**Table 2  Parameter Estimates for One-Parameter and GAM Models**

| Variables | | GAM $\mu_t(p;\vec{X})$ | GAM $\sigma_t(p;\vec{X})$ | Additive $\mu_t(p;\vec{X})$ | Multiplicative $\sigma_t(p;\vec{X})$ |
|---|---|---|---|---|---|
| Intercept | | −0.207* | −2.519 | 0.913** | −0.592** |
| **Demographics** | | | | | |
| $\beta_1$ | EDUCAT | −0.974 | −0.842 | −0.615 | −1.013 |
| $\beta_2$ | RETIRED | N/A | 0.968 | N/A | N/A |
| $\beta_3$ | HSIZEAVG | 0.212 | 0.118 | N/A | 0.143 |
| $\beta_4$ | INCOME | 0.710 | 0.885 | 0.665 | 0.784 |
| $\beta_5$ | WORKWOM | −0.866 | −1.653 | −0.649 | −1.397 |
| $\beta_6$ | HHVAL200 | −0.613 | −0.944 | −0.953 | −0.745 |
| $\beta_7$ | MORTGAGE | −0.889 | −0.165** | −0.855 | −0.591 |
| $\beta_8$ | SHOPINDX | −0.958 | −1.187 | −0.775 | −1.050 |
| **Price elasticities** | | | | | |
| $\beta_\mu$ or $\beta_\sigma$ | Product-specific | 1.796 | 3.299 | 2.563 | 2.309 |
| $\lambda_1$ | Zone 1 | 1.251 | −0.253** | 0.822 | 0.673 |
| $\lambda_2$ | Zone 2 | 1.547 | 0.355 | 0.938 | 1.162 |
| $\lambda_3$ | Zone 3 | 1.726 | 0.269* | 0.846 | 1.226 |
| $\lambda_4$ | Zone 4 | 2.200 | 1.511 | 0.942 | 1.961 |
| $\lambda_5$ | Zone 5 | 0.813 | −0.604 | 0.689 | 0.294 |
| $\lambda_6$ | Zone 6 | 1.341 | 0.220* | 1.428 | 0.980 |
| $\lambda_7$ | Zone 7 | 1.081 | −0.031* | 0.907 | 0.726 |
| $\lambda_8$ | Zone 8 | 2.607 | 1.146 | 0.737 | 2.077 |
| $\lambda_9$ | Zone 9 | 2.545 | 1.598 | 0.877 | 2.236 |
| $\lambda_{10}$ | Zone 10 | 1.116 | 0.326** | 0.688 | 0.863 |
| $\lambda_{11}$ | Zone 11 | 0.924 | 0.349 | 0.810 | 0.741 |
| $\lambda_{12}$ | Zone 12 | 1.481 | 1.526 | 0.794 | 1.507 |
| $\lambda_{13}$ | Zone 13 | 0.296 | −0.117* | −0.329* | 0.181** |
| $\lambda_{14}$ | Zone 14 | 0.655 | 0.592 | 0.778 | 0.642 |
| **Trend and seasonality** | | | | | |
| $\alpha$ | WEEK $(t)$ | −0.001 | −0.0006 | −0.0026 | −0.0006 |
| $\beta_1^c$ | $\cos\left(\frac{2\pi t}{T}\right)$ | −0.132 | −0.364 | −0.098 | −0.238 |
| $\beta_1^s$ | $\sin\left(\frac{2\pi t}{T}\right)$ | 0.040 | −0.074 | 0.016 | −0.014 |
| $\beta_2^c$ | $\cos\left(\frac{4\pi t}{T}\right)$ | −0.061 | −0.018 | −0.030 | −0.016 |
| $\beta_2^s$ | $\sin\left(\frac{4\pi t}{T}\right)$ | −0.069 | −0.130 | 0.043 | −0.090 |
| $\beta_3^c$ | $\cos\left(\frac{6\pi t}{T}\right)$ | 0.055 | −0.108 | 0.115 | −0.036 |
| $\beta_3^s$ | $\sin\left(\frac{6\pi t}{T}\right)$ | −0.014 | −0.031 | 0.021 | −0.015 |
| **One-parameter model** | | | | | |
| $\sigma(p;\vec{X})=\sigma$ | | — | — | 120.000 | — |
| $\sigma(p;\vec{X})/\mu(p;\vec{X})=\eta$ | | — | — | — | 0.357 |
| Log-likelihood | | −113,523 | | −126,661 | −119,748 |
| SMSE in the 7th year | | 69,647,130 | | 82,945,416 | 72,388,138 |

*Notes.* All parameter estimates are significant at *p*-value 0.01 except where indicated otherwise with asterisks.
  $^*p \leq 0.1$; $^{**}p \leq 0.05$.

for the CMLE problem for the 13 zones, and thus a BSLP policy is optimal in each period. However, the conditions in Theorem 5 hold only for zone 1. In other words, the optimal price for other zones may not be monotone in the inventory level, depending on the cost structure.

Finally, we observe from Table 2 a significant difference between the GAM and the one-parameter demand models—the corresponding coefficients differ significantly. Also, the signs of the seasonal factors can be opposite. Overall, the GAM leads to a better fit than the one-parameter models—an 11.57% improvement over the additive demand model and a 5.48% improvement over the multiplicative demand model in the log-likelihood value. The benefit of this improvement is validated by the prediction accuracy computed from the validation data set. The GAM significantly reduces the prediction errors relative to the one-parameter models—a 19.09% reduction over the additive demand model and a 3.94% reduction over the multiplicative demand model in the SMSE value.

## 6.  Concluding Remarks

In this study, we revisit the classical dynamic inventory–pricing problem with backorder and advance the development of this problem in two important dimensions. First, our analysis leads to easy-to-verify conditions for the optimality of a BSLP policy. These conditions, more general than those identified in previous studies, allow for demand modeling flexibility and thus enhance the applicability of the BSLP policy. Moreover, we generalize the notion of BSLP policy by showing that the monotonicity of price as a function of the inventory is *not* required for the optimality of a BSLP policy in our model.

Second, we develop a constrained maximum likelihood estimation model to simultaneously estimate the demand model and verify our conditions. This framework incorporates various demographic characteristics of consumers, and is capable of handling multiple sales zones with nonstationary demands. The application of this framework to a retail data set allows us to verify the conditions obtained from our analytical derivations and confirm the applicability of a BSLP policy. The observation from the data also suggests a significant benefit of using the GAM over the one-parameter demand models.

Next we discuss the limitation of our study and point out future research directions. Our model assumes full backorder of the customer demand with the backorder cost independent of the product price. Extending the analysis to accommodate lost sales or partial backorder and allowing for price-dependent backorder penalty under the GAM can greatly enhance the applicability of the model. We consider only a single product in our model. As we have pointed out, our model applies to products with strong store and brand loyalty. For product categories that reveal strong substitution, one has to formulate a multiproduct model to account for the cross-product price elasticity. In view of the high product variety in many

industries, it is important to identify a simple-to-implement policy and explore the use of data for multiproduct planning.

## Acknowledgments

## Appendix. Proofs

PROOF OF LEMMA 1. For arbitrary values of $(z_t^1, \tau_t^1)$ and $(z_t^2, \tau_t^2)$ in $\Omega_t(x_t)$, if $\mu_t(g_t(\tau_t))$ is concave in $\tau_t$, for any given $\lambda \in [0, 1]$, we have

$$\lambda z_t^1 + (1-\lambda)z_t^2 + \mu_t(g_t(\lambda\tau_t^1 + (1-\lambda)\tau_t^2))$$
$$\geq \lambda z_t^1 + (1-\lambda)z_t^2 + \lambda\mu(g_t(\tau_t^1)) + (1-\lambda)\mu_t(g_t(\tau_t^2))$$
$$= \lambda(z_t^1 + \mu_t(g_t(\tau_t^1))) + (1-\lambda)(z_t^2 + \mu_t(g_t(\tau_t^2)))$$
$$\geq \lambda x_t + (1-\lambda)x_t = x_t.$$

The first inequality follows from the concavity of $\mu_t(g_t(\tau_t))$. The second inequality follows because both $(z_t^1, \tau_t^1)$ and $(z_t^2, \tau_t^2)$ are in $\Omega_t(x_t)$. Clearly, $\underline{\tau} \leq \lambda\tau_t^1 + (1-\lambda)\tau_t^2 \leq \bar{\tau}$, implying that $(\lambda z_t^1 + (1-\lambda)z_t^2, \lambda\tau_t^1 + (1-\lambda)\tau_t^2)$ is also in $\Omega_t(x_t)$. Therefore, $\Omega_t(x_t)$ is a convex set. □

PROOF OF THEOREM 2. We prove the results by induction. It is clear that $\bar{V}_{T+1}(x_{T+1}) = -c_t x_{T+1}$ is concave. Suppose that $\bar{V}_{t+1}(x_{t+1})$ is concave. Then, by assumption, $J_t(z_t, \tau_t)$ is jointly concave in $(z_t, \tau_t)$. It follows from Lemma 1 that $\bar{V}_t(x_t)$ is concave as concavity is preserved under maximization. Hence, we have part (i).

Since $J_t(z_t, \tau_t)$ is jointly concave, there exists a maximizer $(\hat{z}_t, \hat{\tau}_t)$ of $J_t$ over $z_t \in (-\infty, \infty)$ and $\tau_t \in [\underline{\tau}_t, \bar{\tau}_t]$. Define $\hat{y}_t = \hat{z}_t + \mu_t(g_t(\hat{\tau}_t))$ and $\hat{p}_t = \sigma_t^{-1}(\hat{\tau}_t)$. Note that both $\hat{y}_t$ and $\hat{p}_t$ are independent of $x_t$. We have two cases to consider:

*Case* 1: When $x_t \leq \hat{y}_t$, then $(\hat{z}_t, \hat{\tau}_t) \in \Omega_t(x_t)$, and thus the pair maximizes $J_t(z_t, \tau_t)$ over $\Omega_t(x_t)$. In this case, the optimal solution is $(z_t^*(x_t), \tau_t^*(x_t)) = (\hat{z}_t, \hat{\tau}_t)$ and it is optimal to set the order-up-to level to $y_t^*(x_t) = \hat{y}_t$ and charge a price $p_t^*(x_t) = \hat{p}_t$.

*Case* 2: When $x_t > \hat{y}_t$, then $(\hat{z}_t, \hat{\tau}_t) \notin \Omega_t(x_t)$. In this case, the optimal $(z_t^*(x_t), \tau_t^*(x_t))$ must satisfy the relation $z_t = x_t - \mu_t(g_t(\tau_t))$. Substituting this expression into the objective function, we obtain $J_t(x_t - \mu_t(g_t(\tau_t)), \tau_t)$. Let $z^0(\tau) \equiv \arg\max_z J_t(z, \tau)$. For any pairs $(x_t^1, \tau_t^1)$ and $(x_t^2, \tau_t^2)$ satisfying $x_t^i > z^0(\tau_t^i) + \mu_t(g_t(\tau_t^i))$, $i = 1, 2$, we have

$$J_t\left(\frac{x_t^1 + x_t^2}{2} - \mu_t\left(g_t\left(\frac{\tau_t^1 + \tau_t^2}{2}\right)\right), \frac{\tau_t^1 + \tau_t^2}{2}\right)$$
$$\geq J_t\left(\frac{x_t^1 + x_t^2}{2} - \frac{\mu_t(g_t(\tau_t^1)) + \mu_t(g_t(\tau_t^2))}{2}, \frac{\tau_t^1 + \tau_t^2}{2}\right)$$
$$= J_t\left(\frac{x_t^1 - \mu_t(g_t(\tau_t^1)) + x_t^2 - \mu_t(g_t(\tau_t^2))}{2}, \frac{\tau_t^1 + \tau_t^2}{2}\right)$$
$$\geq \frac{J_t(x_t^1 - \mu_t(g_t(\tau_t^1)), \tau_t^1) + J_t(x_t^2 - \mu_t(g_t(\tau_t^2)), \tau_t^2)}{2}.$$

The first inequality follows because $\mu_t(g_t(\tau_t))$ is concave, and $J_t(z_t, \tau_t)$ is nonincreasing for $z_t \geq z^0(\tau_t)$ under a fixed $\tau_t$. The second inequality directly follows from the joint concavity of $J_t(z_t, \tau_t)$. Thus, $J_t(x_t - \mu_t(g_t(\tau_t)), \tau_t)$ is jointly concave in $x_t$ and $\tau_t$. The optimal $\tau_t^*(x_t)$ is the maximizer of this function over $[\underline{\tau}_t, \bar{\tau}_t]$. Therefore, the optimal order-up-to level is $x_t$, and the optimal price is $p_t^*(x_t) = \sigma_t^{-1}(\tau_t^*(x_t))$.

Combining Cases 1 and 2, we obtain part (iii). Finally, from the above analysis, it is straightforward to see the preservation of concavity stated in part (ii). This completes the proof. □

PROOF OF THEOREM 3. Since $g_t'(\tau_t) = 1/\sigma_t'(p_t)$ and $g_t''(\tau_t) = -\sigma_t''(p_t)/[\sigma_t'(p_t)]^3$, (8) follows from

$$\frac{\partial^2 \mu_t(g_t(\tau_t))}{\partial \tau_t^2} = \mu_t''(g_t(\tau_t))[g_t'(\tau_t)]^2 + \mu_t'(g_t(\tau_t))g_t''(\tau_t)$$
$$= \frac{\mu_t''(p_t)}{[\sigma_t'(p_t)]^2} - \frac{\mu_t'(p_t)\sigma_t''(p_t)}{[\sigma_t'(p_t)]^3}.$$

To see (9), we note that $R_t(g_t(\tau_t)) = (g_t(\tau_t) - c_t)\mu_t(g_t(\tau_t))$. Taking the second-order derivative with $\tau_t$, we have

$$\frac{\partial^2 R_t(g_t(\tau_t))}{\partial \tau_t^2}$$
$$= \mu_t(g_t(\tau_t))g_t''(\tau_t) + 2\mu_t'(g_t(\tau_t))[g_t'(\tau_t)]^2$$
$$+ (g_t(\tau_t) - c_t)[\mu_t''(g_t(\tau_t))(g_t'(\tau_t))^2 + \mu_t'(g_t(\tau_t))g_t''(\tau_t)].$$

From the concavity of $\mu_t(g_t(\tau_t))$ and the proof of (8), the second term on the right-hand side must be nonpositive. We only need to check the sign of the first term; i.e.,

$$\phi(\tau_t) = \mu_t(g_t(\tau_t))g_t''(\tau_t) + 2\mu_t'(g_t(\tau_t))[g_t'(\tau_t)]^2$$
$$= -\frac{\sigma_t''(p_t)\mu_t(p_t) - 2\mu_t'(p_t)\sigma_t'(p_t)}{[\sigma_t'(p_t)]^3}.$$

The last equation uses $g_t'(\tau_t) = 1/\sigma_t'(p_t)$ and $g_t''(\tau_t) = -\sigma_t''(p_t)/[\sigma_t'(p_t)]^3$. Hence, the result follows. □

PROOF OF COROLLARY 4. Part (i) follows from Theorem 3 by setting $\sigma_t'(p_t) = \eta\mu_t'(p_t)$ and $\sigma_t''(p_t) = \eta\mu_t''(p_t)$. Part (ii) follows from Theorem 3 by setting $\mu_t'(p_t) = \mu_t''(p_t) = 0$. □

PROOF OF THEOREM 5. Note that $g_t(\tau_t)$ is weakly decreasing in $\tau_t$. Let $\mathcal{D}_t(\tau_t, \epsilon_t) = \mu_t(g_t(\tau_t)) + \tau_t\epsilon_t$. For each realization of $\epsilon_t$,

$$\frac{\partial \mathcal{D}_t(\tau_t, \epsilon_t)}{\partial \tau_t} = \mu_t'(g_t(\tau_t))g_t'(\tau_t) + \epsilon_t$$
$$= \frac{\mu_t'(p_t)}{\sigma_t'(p_t)} + \epsilon_t$$
$$\geq \frac{\mu_t'(p_t)}{\sigma_t'(p_t)} - \frac{\mu_t(p_t)}{\sigma_t(p_t)}.$$

The inequality follows from the assumption that the demand is nonnegative for any feasible price $p_t$. Since $\sigma_t'(p_t) < 0$, $\mathcal{D}_t(\tau_t, \epsilon_t)$ is increasing in $\tau_t$ when $e_t^\mu(p_t) \geq e_t^\sigma(p_t)$.

When $x_t \geq \hat{y}_t$, the objective function in (5) can be written as

$$J_t(x_t - \mu_t(g_t(\tau_t)), \tau_t) = R_t(g_t(\tau_t)) - (c_t - \rho c_{t+1})[x_t - \mu_t g_t(\tau_t)]$$
$$- \mathsf{E}[H_t(x_t - \mathcal{D}_t(\tau_t, \epsilon_t))]$$
$$+ \rho\mathsf{E}[\bar{V}_t(x_t - \mathcal{D}_t(\tau_t, \epsilon_t))].$$

For each realization of $\epsilon_t$, the right-hand side is supermodular in $(x_t, \tau_t)$. This implies that the optimal $\tau_t^*(x_t)$ is weakly increasing in $x_t$, and thus $p_t^*(x_t) = g_t(\tau_t^*(x_t))$ is weakly decreasing in $x_t$. $\quad\square$

PROOF OF COROLLARY 6. To see part (i), we note that $\sigma_t(p_t) = \eta\mu_t(p_t)$ is weakly decreasing in $p_t$ and $e_t^\mu(p_t) = e_t^\sigma(p_t)$. To show part (ii), we also observe from the proof of Theorem 5 that $\mathscr{D}_t(\tau_t, \epsilon_t)$ is weakly increasing in $\tau_t$ when $\mu_t(p_t) = \mu_t$, and hence the result follows. $\quad\square$

PROOF OF THEOREM 7. For the additive demand model, we perform a change of variable by defining $s_t = \mu_t(p_t)$. Consequently, the problem in (2) is transformed to one of finding the optimal $(s_t, y_t)$, which is a concave maximization problem as long as $\mu_t^{-1}(s_t)s_t$ is concave in $s_t$. This condition is equivalent to $2[\mu_t'(p_t)]^2 \geq \mu_t''(p_t)\mu_t(p_t)$. It is then easy to show that the optimal policy is a BSLP policy. The monotonicity of $p_t^*(x_t)$ follows from the observations that $\sigma_t(p_t)$ is constant, and thus $e_t^\mu(p_t) \geq 0 = e_t^\sigma(p_t)$. $\quad\square$

# References

Agrawal V, Seshadri S (2000) Impact of uncertainty and risk aversion on price and order quantity in the newsvendor problem. *Manufacturing Service Oper. Management* 2(4):410–422.

Amihud Y, Mendelson H (1983) Price smoothing and inventory. *Rev. Econom. Stud.* 50(1):87–98.

Beutler FJ, Ross KW (1985) Optimal policies for controlled Markov chains with a constraint. *J. Math. Anal. Appl.* 112:236–252.

Brockwell PJ, Davis RA (2002) *Introduction to Time Series and Forecasting* (Springer, New York).

Chan LMA, Shen ZJM, Simchi-Levi D, Swann J (2005) Coordination of pricing and inventory decisions: A survey and classification. Wu SD, Shen ZJM, eds. *Handbook of Quantitative Supply Chain Analysis: Modeling in the E-Business Era* (Kluwer Academic Publishers), 335–392.

Chen H, Zhang Z (2010) Joint pricing and inventory control with general additive demand. Working paper, University of British Columbia, British Columbia, Canada.

Chen X, Simchi-Levi D (2004a) 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 (2004b) Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The infinite horizon case. *Math. Oper. Res.* 29(3):698–723.

Cotterill RW, Franklin AW, Ma LY (1996) *Measuring Market Power Effects in Differentiated Product Industries: An Application to the Soft Drink Industry* (Department of Agricultural and Resource Economics, University of Connecticut, Storrs).

Dubé JP (2004) Multiple discreteness and product differentiation: Demand for carbonated soft drinks. *Marketing Sci.* 23(1):66–81.

Dubé JP (2005) Product differentiation and mergers in the carbonated soft drink industry. *J. Econom. Management Strategy* 14(4):879–904.

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

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

Feinberg EA (1994) Constrained semi-Markov decision processes with average rewards. *Math. Methods Oper. Res.* 39:257–288.

Feng Q (2010) Integrating dynamic pricing and replenishment decisions under supply capacity uncertainty. *Management Sci.* 56(12):2154–2172.

Feng Q, Shi R (2012) Sourcing from multiple suppliers for price-dependent demands. *Production Oper. Management* 21(3):547–563.

Fok D, Franses PH, Paap R (2007) Seasonality and non-linear price effects in scanner-data-based market-response models. *J. Econometrics* 138(1):231–251.

Gimpl-Heersin L, Rudlof C, Fleischmann M, Taudes A (2008) Integrating pricing and inventory control: Is it worth the effort? *Bus. Res.* 1(1):106–123.

Hastie T, Herman A (1990) An analysis of gestational age, neonatal size and neonatal death using nonparametric logistic regression. *J. Clinical Epidemiology* 43:1179–1190.

Hastie T, Tibshirani R (1990) *Generalized Additive Models* (Chapman and Hall, London).

Hastie T, Tibshirani R, Friedman J (2001) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer, New York).

Hoch SJ, Kim BD, Montgomery AL, Rossi PE (1995) Determinants of store-level price elasticity. *J. Marketing Res.* 32(1):17–29.

Huh WT, Janakiraman G (2008) Optimality results in inventory-pricing control: An alternate approach. *Oper. Res.* 56(3):783–790.

Karlin S, Carr CR (1962) Prices and optimal inventory policy. Arrow KJ, Karlin S, Scarf H, eds. *Studies in Applied Probability and Management Science* (Stanford University Press, Stanford, CA), 159–172.

Kocabıyıkoğlu A, Popescu I (2011) An elasticity approach to the newsvendor with price-sensitive demand. *Oper. Res.* 59(2):301–312.

Li Q, Zheng S (2006) Joint inventory replenishment and pricing control for systems with uncertain yield and demand. *Oper. Res.* 54(4):696–705.

Liu W, Cela J (2007) Improving credit scoring by generalized additive model. *SAS Global Forum 2007: Data Mining and Predictive Modeling, Orlando, FL.*

Lu Y, Simchi-Levi D (2013) On the unimodality of the profit function of the pricing newsvendor. *Production Oper. Management* 22(3):615–625.

McCullagh P, Nelder JA (1989) *Generalized Linear Models* (CRC Press, Boca Raton, FL).

Mills ES (1959) Uncertainty and price theory. *Quart. J. Econom.* 73(1):116–130.

Petruzzi N, Dada CM (1999) Pricing and the newsvendor problem: A review with extensions. *Oper. Res.* 47(2):183–194.

Polatoglu H, Sahin I (2000) Optimal procurement policies under price-dependent demand. *Internat. J. Production Econom.* 65(2):141–171.

Stasinopoulos M, Rigby R (2005) Generalized additive models for location scale and shape. *J. Royal Statist. Soc. Series C (Appl. Statist.)* 54(3):507–554.

Thowsen GT (1975) A dynamic, nonstationary inventory problem for a price/quantity setting firm. *Naval Res. Logist. Quart.* 22(3):461–476.

Whitin TM (1955) Inventory control and price theory. *Management Sci.* 2(1):61–80.

Yano CA, Gilbert SM (2003) Coordinated pricing and production/procurement decisions: A review. Chakravarty A, Eliashberg J, eds. *Managing business Interfaces: Marketing, Engineering and Manufacturing Perspectives* (Kluwer Academic Publishers, Boston), 65–103.

Yao L, Chen YF, Yan H (2006) The newsvendor problem with pricing. *Internat. J. Management Sci. Engrg. Management* 1(1):3–16.

Young L (1978) Price, inventory and the structure of uncertain demand. *New Zealand Oper. Res.* 6(2):157–177.

Young L (1979) Uncertainty, market structure, and resource allocation. *Oxford Econom. Papers* 31(1):47–59.