OPTIMAL PRICING AND PRODUCTION POLICIES OF A MAKE-TO-STOCK SYSTEM WITH FLUCTUATING DEMAND

JEAN-PHILIPPE GAYON

Laboratoire G-SCOP, Grenoble INP
38031 Grenoble Cedex 1, France
E-mail: jean-philippe.gayon@g-scop.fr

IŞILAY TALAY-DEĞIRMENCI

The Fuqua School of Business
Duke University
Durham, NC, 27708
E-mail: isilay.talay@duke.edu

FIKRI KARAESMEN AND E. LERZAN ÖRMECI

Department of Industrial Engineering
Koç University
34450, Sarıyer, İstanbul, Turkey
E-mail: fkaraesmen@ku.edu.tr; lormeci@ku.edu.tr

We study the effects of different pricing strategies available to a production–inventory system with capacitated supply, which operates in a fluctuating demand environment. The demand depends on the environment and on the offered price. For such systems, three plausible pricing strategies are investigated: static pricing, for which only one price is used at all times, environment-dependent pricing, for which price changes with the environment, and dynamic pricing, for which price depends on both the current environment and the stock level. The objective is to find an optimal replenishment and pricing policy under each of these strategies. This article presents some structural properties of optimal replenishment policies and a numerical study that compares the performances of these three pricing strategies.

1. INTRODUCTION

Dynamic pricing concerns the adjustment of prices to charge customers over time in order to maximize total revenues. It is by now a well-established strategy in service industries where typical examples include airline, hotel, and electric utilities management. These examples involve cases where the resources are perishable and nonrenewable. Recent overviews of this class of problems can be found in the review articles of McGill and Van Ryzin [20] and Bitran and Caldentey [2].

An important distinction of the above-mentioned service industries where dynamic pricing applications are already relatively mature is that their services cannot be provided in advance and stored. In recent years, there has been an increasing interest in dynamic pricing policies in industries where the sellers have the capability to store inventory. A recent review by Chan, Shen, Simchi-Levi, and Swann [5] presents several dynamic pricing examples from different industrial practices and emphasizes the need for employing pricing as a strategic tool. This motivates the need to assess the potential benefits that can be expected from dynamic pricing strategies in retail and production operations. Our focus in this article is on coordinated replenishment and pricing strategies for a production/inventory system that produces and sells a product with a long life cycle. In this class of problems, a seller faces a stochastic and price-sensitive demand and can replenish inventory with limited production capacity. The objective of the seller is to find a dynamic pricing and production policy to match the demand and the inventory in order to maximize his/her total profit.

Our model of the production/inventory system is based on a framework sometimes referred to as the make-to-stock queue. In this framework, the production capacity of the supply system is modeled by a single server that processes items one by one. Although simplified, this framework is known to yield useful insights in a number of interesting problems involving capacity/inventory interactions. The work of Buzacott and Shanthikumar [3] and Zipkin [32] comprise detailed expositions of such models. On the inventory side, we assume lost sales. A purchase can only take place if the product is available in stock at the time of demand arrival. Finally, on the demand side, we model the potential demand of customers as a Poisson process whose rate can fluctuate over time. The actual realized demand (i.e., purchase) depends not only on the current potential demand rate but also on the price offered and on product availability.

For a system operating in this environment, there are three plausible pricing strategies: static pricing (SP), for which a single price is used at all times, environment-dependent pricing (EDP), for which the price is allowed to change when the demand environment fluctuates, and dynamic pricing (DP), for which the price depends on both the current environment and the stock level. In this article, we use a Markov decision process framework to model this system as a make-to-stock queue operating under each of these strategies. Using this framework, we first show that optimal replenishment policies are environment-dependent base-stock-level policies for the three pricing strategies. For the DP strategy, we show that optimal dynamic prices are nonincreasing in the inventory level in each environment. For the SP and EDP

strategies, we show that, under certain conditions, optimal base-stock levels have a natural ordering in terms of the demand rates and prices. We also compare the performances of these three strategies by an extensive numerical study.

This article focuses on a quantitative comparison of the three different pricing policies. However, there might be other more qualitative considerations that play a role in the selection of a pricing policy. SP represents the traditional pricing since the price remains fixed over time, regardless of the changes in the environment and in the stock level. This type of policy is easy to implement. In addition, consumers might prefer the transparency of a known price that is not subject to any changes. At the other extreme, there is DP, which leads to frequent price changes, since even a change in the stock level might trigger a change in price. Therefore, DP might create negative consumer reactions. Moreover, its implementation requires sophisticated information systems that can accurately track sales and inventory data in real time and this can be extremely difficult especially if price changes require a physical operation such as a label change. EDP, on the other hand, allows the price to change only with the fluctuations in the potential demand rate, which may be relatively infrequent. Hence, the associated system changes the prices but not as frequently as the one with the DP does. As a result, this policy is somewhere in between static and dynamic policies regarding the implementation problems. One of the main objectives of this article is to assess the magnitude of the additional profit that might be gained by more complicated policies in order to help the policy selection decision.

The article is structured as follows. In Section 2, we provide a review of the related literature. Section 3 presents the model and the problem formulation. Section 4 establishes certain structural properties of the optimal dynamic and static policies. Section 5 comprises the numerical study. Finally, the conclusions are presented in Section 6.

2. LITERATURE REVIEW

Pricing models and integrated pricing and production decisions have been studied since the 1960s, and the excellent reviews of Yano and Gilbert [29], Chan et al. [5], and Elmaghraby and Keskinocak [11] provide a summary of the research articles on this area by focusing on different aspects of the problem. Our literature review focuses on the articles that consider stochastic demand and multiple replenishments over the planning horizon.

Most of the earlier articles consider periodic-review models. Zabel [30] studied first the multiplicative demand case for a single-period problem with convex production and linear inventory holding costs and lost sales. He demonstrated that there is a unique optimal policy for every initial stock level. Zabel [31] then extended those results to the multiperiod problem with both additive and multiplicative demand cases. He found that the firm should produce when the initial inventory is below a critical level and that the price is a decreasing function of the inventory. Thowsen [28] studied the case of price-sensitive nonstationary demand consisting of a general function of

price with an additive stochastic component. The unit production cost is linear, inventory holding and shortage costs are convex, and backlogging, partial-backlogging, and lost sales assumptions are considered along with the assumption of deteriorating inventory with a deterministic fraction. He conjectured that a period-dependent base-stock list price (y, p) policy, similar to critical number inventory policies, is optimal, and he specified the conditions for its optimality. Federgruen and Heching [12] examined a periodic-review model in which both the replenishment quantity and the price are decided at the beginning of each period. The demand is stochastic and the excess demand is backlogged. For a concave revenue function, convex inventory holding costs, and linear production costs, it is shown that a base-stock list price policy (y_t^*, p_t^*) is optimal for both average and discounted profit criteria in a finite or infinite horizon. In a given period, if the inventory level is below the base-stock level, it is increased to the base-stock level and the list price is charged. If the inventory level is above the base-stock level, then nothing is ordered and a price discount is offered. In addition, the optimal price is a nonincreasing function of the initial inventory level. Their results also extend to the case with production capacity limits. They conducted a numerical study based on data collected from a women's apparel retailer. In particular, they exhibited the magnitude of the benefits of DP under limited replenishment opportunities in a finite horizon context.

A number of articles consider the periodic-review model with fixed ordering costs (in addition to variable ordering costs). Thomas [27] considered a finite horizon model with price-sensitive random demand with backordering and proposed a variant of the well-known (s, S) policy that also includes a price parameter. He conjectured the optimality of this relatively simple (s, S, p) policy. This conjecture is proven in Chen and Simchi-Levi [9] for additive demand functions in a finite horizon setting. Chen and Simchi-Levi also showed the optimality of a slightly more complicated policy for general demand functions. The optimality proof is extended to the more challenging infinite horizon case in Chen and Simchi-Levi [10]. Feng and Chen [15] generalized the setting of Chen and Simchi-Levi and provide an alternative optimality proof as well as a computational procedure for calculating the optimal policy. There is also some recent work on the corresponding continuous-review model. Feng and Chen [14] presented results on optimal policy structure for a continuous-review system with a Poisson demand process. Chen and Simchi-Levi [8] generalized the optimal policy structure to a compound-renewal demand process. Chen, Wu, and Yao [7] also considered a continuous-review model in which the (price-sensitive) demand is modeled by a Brownian motion. They analyzed particular replenishment policies under different pricing structures. They showed through a combination of analytical results and numerical examples that the joint optimization of both decisions might result in significant profit improvement over the traditional way of performing sequential optimization. Moreover, they also found that DP results in a limited profit improvement over SP when both methods are optimally applied.

An important characteristic of a production/inventory system is the limited production capacity that induces endogenous and random lead times. Only a few of the

above-cited articles explicitly model capacity constraints or lead times. In the periodic-review case, Federgruen and Heching [12] discussed these extensions to their basic model, but their main focus was on the uncapacitated zero-lead-time case. In a recent article, Chan, Simchi-Levi, and Swann [6] considered a multiperiod horizon with limited capacity, lost sales (in case of stockout), and the possibility of discretionary sales (stock reservation for the future). Their computational results indicated that the benefit of DP tends to increase as the capacity becomes more constrained and as the demand seasonality increases.

The work by Li [19] seemed to be the only article whose focus is a production/inventory system modeled as a make-to-stock queue with joint pricing and replenishment considerations. He considered a continuous-review model with price-sensitive demand where the cumulative production and demand are modeled by Poisson processes with controllable intensities. The demand is a continuous function of the price. When there is demand in excess, sales are lost. The production and holding costs are linear and the demand intensity is controlled through the price. When the prices were set dynamically over time, Li showed that a base-stock policy is optimal. Moreover, it is shown that the optimal price is a nonincreasing function of the inventory level. Caldentey and Wein [4] considered a related problem for a make-to-stock queue, but in their model, price is modeled as an exogenous stochastic process and the decision is admission rather than pricing.

One important motivation for using DP might be the adjustment of prices to randomly changing demand environments. The pricing articles mentioned earlier do not explicitly model random fluctuations in the mean demand rate. A widely used approach to capture the effect of a fluctuating environment is defining the demand process to be driven by an exogenous Markov chain. The resulting demand process is called a Markov-modulated demand process. A number of articles investigated the effect of such demand processes on inventory systems without pricing considerations. Feldman [13] and Kalpakam and Arivarignan [17] are examples that investigate (s, S) inventory systems with Markov-modulated demand. Song and Zipkin [26] presented an inventory model that includes a fluctuating demand environment where the demand is a Markov-modulated Poisson process. The other components of the model are a fixed or stochastic order lead time, inventory holding and backorder costs, and a positive discount rate. They showed that without a fixed ordering cost, a world state-dependent base-stock policy is optimal, and that with a fixed ordering cost, a world state-dependent (s, S) policy is optimal. We use the Markov-modulated, price-sensitive Poisson demand to model the demand process in this study, and the demand process we use is similar to the one used by Song and Zipkin [26] except that we include price sensitivity. A number of articles (e.g., Beyer and Sethi [1], Sethi and Cheng [25], Özekici and Parlar [21]) considered similar models and proved the optimality of state-dependent base-stock policies in the absence of setup costs and the optimality of state-dependent (s, S)-type policies when there are fixed setup costs. It is important to note again that none of these articles considered pricing decisions, capacity limitations, and endogenous lead times.

To our knowledge, the work by Li [19] is the only article that investigates DP for a production/inventory system in a continuous-review setting. This article presents structural results on the optimal policy but does not attempt to assess the performance of dynamic pricing. We extend the model and the findings of Li in two major ways. First, we use a Markov-modulated demand process (with controllable intensities) instead of a Poisson demand process and present new structural results on the optimal DP policy in this setting. To this end, we also present alternative simple proofs of some of the intermediate properties in Li using the value iteration technique of dynamic programming as well as generalizing the results therein to a Markov-modulated Poisson demand process. Second, in addition to DP, we also investigate environment-dependent and static policies for a Markov-modulated Poisson demand model. This leads us to an objective assessment of the potential benefits of each of these policies.

3. MODEL FORMULATION

Consider a supplier that produces a single part at a single facility. The processing time is exponentially distributed with mean $1/\mu$ and the completed items are placed in a finished goods inventory. The unit variable production cost is c and the stock level is $X(t)$ at time t, where $X(t) \in \mathbb{N} = \{0, 1, \ldots\}$. We denote by h the induced inventory holding cost per unit time and h is assumed to be a convex function of the stock level.

The exogenous environment state representing the demand fluctuations evolves according to a continuous-time Markov chain with state space $E = \{1, \ldots, N\}$ and transition rates q_{ej} from state e to state $j \neq e$. We assume that this Markov chain is recurrent to avoid technicalities. The customers decide whether to purchase an item depending on the price p and on the state of the exogenous environment e. More precisely, the demands occur according to a Markov-modulated Poisson process (MMPP) with rate $\lambda_e(p)$. For all environment states, the set of allowable prices $\mathcal{P}$ is identical and it might be either discrete (finite or countably infinite) or continuous (uncountably infinite). When $\mathcal{P}$ is continuous, it is assumed to be a compact subset of the set of nonnegative real numbers $\mathbb{R}^+$. In this setting, the underlying process is Markovian, and we employ a Markov decision process framework to find optimal control policies for different pricing strategies. In addition, the analysis can be restricted to stationary Markovian policies since the optimal policy is known to belong to this class (Puterman [23]). Therefore, the current state of the system is exhaustively described, independent of time t, by the state variable (x, e), with x the stock level and e the environment state and (x, e) belongs to the state space $\mathbb{N} \times E$.

In our framework, we assume that the decision-makers always observe the state of the environment, which might not be possible in all settings. Whenever the environment changes are not observable, the system will realize a change in the environment through observing the demand for a while. The transition rates between environments are typically small when compared to the other transition rates, namely the production and demand rates. Hence, the decision-makers will deduce and adjust to

an environmental change in a relatively short time period. In any case, our results can be considered as an upper bound on the actual gain of the system.

For a fixed environment e, we impose several mild assumptions on the demand function. First, we assume that $\lambda_e(p)$ is decreasing in p and we denote by $p_e(\lambda)$ its inverse. One can then alternatively view the rate λ as the decision variable, which is more convenient to work with from an analytical perspective. Thus, the set of allowable demand rates is $\mathcal{L}_e = \lambda_e(\mathcal{P})$ in environment e. We also assume that when the set of prices $\mathcal{P}$ is continuous, p_e is a continuous function of λ. Second, $\lambda_e(p)$ is bounded by Λ_e, the potential demand rate in environment e, so that $\lambda_e(p) \leq \Lambda_e$ for all p and for all e. This reasonable assumption is necessary to uniformize the Markov decision process. Third, the revenue rate $r_e(\lambda) = \lambda p_e(\lambda)$ is bounded.

In this section, we present three models for the three different pricing strategies, namely static, environment dependent, and dynamic. We will represent our models as a combination of certain dynamic programming (DP) operators, where each operator corresponds to a certain event. Then we establish the structure of optimal policies by verifying that these operators satisfy certain properties. This methodology falls into the so-called event-based DP framework introduced by Koole [18]. The models for SP and EDP use the same set of operators, whereas the model for DP differs in the pricing operator. Hence, the results for the first two models follow from the same type of arguments, whereas the DP model requires additional consideration.

3.1. Static Pricing Problem

In the SP problem, the decision-maker has to choose a price p in $\mathcal{P}$ for the whole horizon, as well as a replenishment decision in each state (x, e), in order to maximize the average profit over an infinite horizon. The problem can be viewed in two steps.

In the first step, an optimal replenishment policy is identified for a given price p, which determines an arrival rate for each environment, $\lambda_e = \lambda_e(p)$. Hence, a single price p might specify different arrival rates in different environments, which are denoted by $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)$. Once p or $\boldsymbol{\lambda}$ are fixed, this problem can be formulated as a Markov decision process (MDP). We define the state of the system (x, e), where x is the inventory level and e is the environment. We let $v^{sp}(x, e; \boldsymbol{\lambda})$ be the relative value function of being in state (x, e) and let $g^{sp}(\boldsymbol{\lambda})$ (or $g^{sp}(p)$) denote the optimal average profit. Since the transition rate out of any state is finite, we use uniformization [24] and normalization ($\mu + \sum \Lambda_e + \sum_e \sum_{j \neq e} q_{ej} = 1$) to transform the continuous-time MDP into an equivalent discrete-time MDP. Then the optimality equations are as follows:

$$v^{sp}(x, e; \boldsymbol{\lambda}) + g^{sp}(\boldsymbol{\lambda}) = -h(x) + \mu T^1 v^{sp}(x, e; \boldsymbol{\lambda}) + T^2 v^{sp}(x, e; \boldsymbol{\lambda}) + T^{\lambda_e} v^{sp}(x, e; \boldsymbol{\lambda}),$$

where the operators T^1, T^2, and T^{λ} for any real-valued function $f(x, e)$ are defined as

$$T^1 f(x, e) = \max\{f(x, e), f(x + 1, e) - c\}, \tag{1}$$

$$T^2 f(x, e) = \sum_{j \neq e} q_{ej} f(x, j) + \left(\sum_{i \neq e} \Lambda_i + \sum_{i \neq e} \sum_{j \neq i} q_{ij} \right) f(x, e), \tag{2}$$

$$T^\lambda f(x, e) = \begin{cases} r_e(\lambda) + \lambda f(x - 1, e) + (\Lambda_e - \lambda) f(x, e), & \text{if } x > 0 \\ \Lambda_e f(x, e), & \text{if } x = 0. \end{cases}$$

The operators T^1, T^2, and T^λ respectively correspond to the production decision, the environment fluctuations as well as fictitious transitions, and the demand arrival. We also define the operator T^{sp} such that $v^{sp} + g^{sp}(\lambda) = T^{sp} v^{sp}$.

The second step is then to find an optimal static price p maximizing $g^{sp}(p)$ (or, equivalently, $g^{sp}(\lambda)$). There might be several optimal prices and we denote by p^{sp} a generic optimal static price. We define the overall optimal average profit g^{sp} of the SP problem as

$$g^{sp} = \max_p g^{sp}(p) = g^{sp}(p^{sp}).$$

3.2. Environment-Dependent Pricing Problem

The setting of the EDP problem is very close to the SP problem except that the decision-maker has to choose a price p_e in $\mathcal{P}$ for each environment e.

This problem can be viewed in two steps as well. In the first step, an optimal replenishment policy is identified for a given price vector $\mathbf{p} = (p_1, \ldots, p_N)$, where p_e denotes the offered price in environment e. We denote by $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)$ the equivalent demand vector, where $\lambda_e = \lambda_e(p_e)$. Again, we can formulate the first step as an MDP. Let $v^{edp}(x, e; \boldsymbol{\lambda})$ be the relative value function and $g^{edp}(\boldsymbol{\lambda})$ be the optimal average profit. Optimality equations are the same as in Section 3.1:

$$v^{edp}(x, e; \boldsymbol{\lambda}) + g^{edp}(\boldsymbol{\lambda}) = -h(x) + \mu T^1 v^{edp}(x, e; \boldsymbol{\lambda}) + T^2 v^{edp}(x, e; \boldsymbol{\lambda})$$
$$+ T^{\lambda_e} v^{edp}(x, e; \boldsymbol{\lambda}),$$

where the operator T^{edp} is defined as $v^{edp} + g^{edp}(\boldsymbol{\lambda}) = T^{edp} v^{edp}$.

The second step is to find an optimal price vector $\mathbf{p}^{edp} = (p_1^{edp}, \ldots, p_N^{edp})$, in the set of price vectors $\mathcal{P}^N$, maximizing $g^{edp}(\mathbf{p})$. There might be several optimal price vectors, and we denote by $\mathbf{p}^{edp}$ a generic optimal price vector. We finally define the overall optimal average profit g^{edp} by

$$g^{edp} = \max_{\mathbf{p}} g^{edp}(\mathbf{p}) = g^{edp}(\mathbf{p}^{edp}).$$

The MDP models for SP and EDP use the same set of operators. Hence, when we prove certain properties of operators T^1, T^2, and T^λ, optimal policies of both static and environment-dependent models will have the same structure. The difference between the two models stems from the second step, where optimal prices are determined. SP searches for a single price, whereas EDP needs to specify N prices, one for each environment.

3.3. Dynamic Pricing Problem

The DP problem is an extension of the model by Li [19], who analyzed a similar system operating in a nonfluctuating demand environment. Here, in addition to the demand environment, the price also depends on the inventory level and the decision-maker has to set a price for each state (x, e). This problem is different from the SP and EDP problem in the following way: Since both optimal replenishment and pricing policies depend on the current inventory level as well as the environment, both policies are determined as a result of an MDP.

We let $v^{dp}(x, e)$ be the relative value function of being in state (x, e) and g^{dp} denote the optimal average profit. The optimality equations are the same as in Section 3.1 except that the operator T^λ is replaced by the operator T:

$$v^{dp}(x, e) + g^{dp} = -h(x) + \mu T^1 v^{dp}(x, e) + T^2 v^{dp}(x, e) + T v^{dp}(x, e),$$

where the operator T corresponds to the demand rate decision (equivalently pricing decision) and is defined as

$$T f(x, e) = \max_{\lambda \in \mathcal{L}_e} \left\{ T^\lambda f(x, e) \right\}.$$

Then, similarly to SP and EDP strategies, we define the operator T^{dp} such that $v^{dp} + g^{dp} = T^{dp} v^{dp}$. As we mentioned earlier, the pricing operator for the DP problem is different than the one in SP and EDP problems. The operator T depends on the operator T^λ; however, it chooses the arrival rate that maximizes the value obtained from the operator T^λ in each state (x, e). Hence, the results for this problem will call for further investigation of the operator T.

In each state (x, e), there might be several optimal arrival demand rates and we denote by $\mathcal{L}(x, e)$ the set of optimal arrival rates:

$$\mathcal{L}(x, e) = \arg \max_\lambda T^\lambda v^{dp}(x, e).$$

The equivalent set of optimal prices $\mathcal{P}(x, e)$, in state (x, e), is then the image of $\mathcal{L}(x, e)$ by the function $p_e(\lambda)$. Note that $\mathcal{L}(x, e)$ and $\mathcal{P}(x, e)$ contain at least one element. This is obvious when $\mathcal{P}$ is discrete, and when $\mathcal{P}$ is compact, it follows from the fact that $T^\lambda f(x, e)$ is a continuous function, with respect to λ, defined on a compact set. We denote by $p^{dp}(x, e)$ a generic optimal price in state (x, e).

4. CHARACTERIZING THE STRUCTURE OF OPTIMAL POLICIES

In this section, we characterize the structure of optimal policies for the three pricing policies outlined in Section 3. We first show that for the three pricing strategies, there exists an optimal replenishment policy that can be defined as an environment-dependent base-stock policy. Then we will specialize on the DP strategy for which

optimal prices are proven to be monotone in the inventory level. Finally, we consider the EDP and SP strategies to show that the environment-dependent base-stock levels are monotone in the environment, under certain conditions on the fluctuating environment process.

In this article, we assume that the holding costs $h(x)$ as well as the production costs c do not change with the environment. However, it is easy to show that our results will still be valid when the assumptions of the model are generalized to include an environment-dependent inventory holding cost function ($h_e(x)$) and unit production cost (c_e) as long as $h_e(x)$ is convex and nondecreasing in x for all e.

4.1. Optimal Replenishment Policy

We will use the MDP formulations to prove that there exists an optimal environment-dependent base-stock policy under each of the pricing strategies. We first present the definition of an environment-dependent base-stock policy:

DEFINITION 1: *A replenishment policy that operates in a fluctuating demand environment is an environment-dependent base-stock policy if, for each environment state e, there exists a threshold, $s_e \geq 0$, such that the system produces if and only if the current inventory level is below this threshold.*

Now, we will argue that pricing strategy ϕ yields to an optimal environment-dependent base-stock policy if the value function v^ϕ is concave in the inventory level x for each environment e. Define the operator Δf as $\Delta f(x, e) = f(x + 1, e) - f(x, e)$, so that $\Delta v^\phi(x, e)$ represents the benefit of having an extra unit of item in the inventory in environment e. Moreover, from (1), it is optimal not to produce an item in state x if $\Delta v^\phi(x, e) < c$. The concavity of a function $f(x, e)$, on the other hand, is equivalent to having $\Delta f(x + 1, e) \leq \Delta f(x, e)$ for all (x, e). Hence, concavity of v^ϕ implies that the benefit of an extra item in the inventory is nonincreasing in x for each environment e. Then whenever this benefit drops below a certain value, more specifically below the production cost c, it will remain below c, which guarantees the existence of a base-stock level in each environment e. Our next result shows that the operators of the pricing strategies preserve concavity.

LEMMA 1: *For all pricing strategies $\phi = sp, edp, dp$, if $f(x, e)$ is concave with respect to x, then $T^\phi f$ is also concave with respect to x.*

A proof for Lemma 1 is given in Appendix A. Lemma 1 implies by the value iteration principle that the optimal value function v^ϕ is concave for all $\phi = sp, edp, dp$. As discussed earlier, the concavity of the optimal value function and (1) imply the existence of an optimal base-stock level in each environment e denoted by s_e^ϕ. Hence, we have our main result on optimal replenishment policies.

THEOREM 1: *For all pricing strategies $\phi = sp, edp, dp$, the optimal replenishment policy is an environment-dependent base-stock policy.*

4.2. **Monotonicity of Dynamic Prices**

In this subsection, we consider only the DP strategy. Consider the situation when the inventory level increases: Then the potential holding costs also increase. Hence, the system would like to sell the items faster, which can be achieved by decreasing the prices charged. This intuition holds for optimal pricing policies with a DP strategy, which is the main result of this subsection (its proof can be found in Appendix B).

THEOREM 2: *Let $x < y$, $p(x, e) \in \mathcal{P}(x, e)$, and $p(y, e) \in \mathcal{P}(y, e)$. Then $p(x, e) \geq p(y, e)$.*

Let us now assume that the revenue rate function, r_e, is strictly concave. Under this reasonable assumption, we can deduce some additional properties when the set of prices is continuous. Let p_e^* be the unique price that maximizes r_e. p_e^* represents the optimal price of a problem without inventory holding costs. Alternatively, as in Li [19], p_e^* can be interpreted as the optimal price in the corresponding deterministic problem (where there is no need to carry any inventory).

COROLLARY 1: *If the set of prices $\mathcal{P}$ is compact and the revenue rate r_e is strictly concave, we have the following properties. For all $(x, e) \in I\!N \times E$, we have the following:*

    C.1 *There exists a unique optimal price $p(x, e)$.*
    C.2 *If $x \leq s_e^{dp}$, $p(x, e) \geq p_e^*$. Otherwise $p(x, e) \leq p_e^*$.*

A proof for Corollary 1 is given in Appendix C. Corollary 1 states that the optimal dynamic price is higher than p_e^*, the optimal price in the corresponding deterministic system, unless the current inventory level is above the current base-stock level. When there is a single environment (i.e., stationary Poisson demand), the optimal dynamic price is always greater than or equal to p_e^* which coincides with the result of Li [19]. Another interesting point to note is that, similarly to Federgruen and Heching [12], a price discount (with respect to the optimal deterministic price) is offered whenever the inventory level is above the base-stock level. Moreover, as a consequence of Corollary 1, the price set at the base-stock level s_e^{dp}, in environment e, should be approximately equal to p_e^* [i.e., $p(s_e^{dp}, e) \approx p_e^*$].

4.3. **Monotonicity of Optimal Base-Stock Levels**

In this subsection, we show the monotonicity of optimal base-stock levels under certain conditions on the environment process for only SP and EDP strategies. We label the environments such that environment e represents a worse economical environment than environment $e + 1$ does, so that the environment labels increase as the economy in the corresponding environment improves. Then, we can view environment 1 as the worst possible economical environment and environment N as the best. With this interpretation, it is natural to assume that both the demand and revenue rates have an increasing order with respect to the environment labels; so that $\lambda_1 \leq \cdots \leq \lambda_N$ and

$\lambda_1 p_1 \leq \cdots \leq \lambda_N p_N$. Here, we note that under the SP strategy, the order of demand rates implies the order of revenue rates, and vice versa.

Now, we consider the meaning of transitions between the environments in this setting, for example, the transition from an environment e with $e < N - 1$ to environment N is more drastic when compared to the transition from environment $N - 1$ to environment N. Intuitively, more drastic changes in the state of the environment are rarer since they can occur as a result of more extraordinary events. The following condition is nothing but the mathematical statement of this intuition:

Condition 1:

    a. For all e and for all $j \leq e - 1$, $q_{ej} \geq q_{e+1,j}$.

    b. For all e and for all $j \geq e + 2$, $q_{e+1,j} \geq q_{ej}$.

In this setting (i.e., when both the demand and the revenue rates are in increasing order of environment labels), Condition 1 guarantees the following conclusion on the optimal value function:

LEMMA 2: *Assume that Condition 1 holds. For pricing strategies* $\phi = sp, edp$, *if* $\Delta f(x, e + 1) \geq \Delta f(x, e)$ *for all x and for all $e = 1, \ldots, N - 1$, then* $T^{\phi} \Delta f(x, e + 1) \geq T^{\phi} \Delta f(x, e)$.

A proof for Lemma 2 is given in Appendix D. Lemma 2 states that if the expected benefit of an additional product on inventory is more beneficial in state $e + 1$ than in state e, then operators T^{edp} and T^{sp} preserve this relation. Hence, Lemma 2, together with Theorem 1, establishes the monotonicity of optimal base-stock levels for SP and EDP strategies through the value iteration principle.

THEOREM 3: *Assume that Condition 1 holds. Then for pricing strategies* $\phi = sp, edp$, *the optimal base-stock levels in each environment e, s_e^{ϕ}, have the same ordering with the environment labels; that is, for all $e = 1, \ldots, N - 1$, $s_e^{\phi} \leq s_{e+1}^{\phi}$.*

Remark 1: Condition 1 is directly satisfied when the environment is represented by a birth-and-death process, since $q_{ej} = 0$ for $j \neq e - 1$ and $j \neq e + 1$. Therefore, the base-stock levels in each environment e for SP and EDP strategies (i.e., s_e^{edp} and s_e^{sp}) have the same ordering with the environment labels without any further conditions on the transition rates. In particular, a system with only two environments is a special case of a birth-and-death process, so Condition 1 is automatically satisfied for systems operating in an environment with two states. Then if $\lambda_1 \leq \lambda_2$ and $\lambda_1 p_1 \leq \lambda_2 p_2$, $s_1^{\phi} \leq s_2^{\phi}$ for $\phi = sp, edp$.

5. NUMERICAL RESULTS

In this section, we compare and contrast the properties of five strategies that are summarized in Table 1. The SP, EDP, and DP strategies have already been defined in

TABLE 1. Classification of Strategies

	Single Static Price	One Price per Environment e	One Price per State (x, e)
Single base-stock level	Static (S) strategy	Static base-stock (SB) strategy	Not considered
One base-stock level per environment e	Static pricing (SP) strategy	Environment-dependent pricing (EDP) strategy	Dynamic pricing (DP) strategy

Section 3. The static (S) strategy and the static base-stock (SB) strategy both follow a base-stock replenishment policy with a single base-stock level, independent of the environment state. In the S strategy, there is a single price for all the states of the system, whereas in the SB strategy, there is a specific price for each environment state.

We first present the details of the setup of the numerical experiments. Then we briefly investigate the single-environment system, after which we concentrate on the two-environment case.

5.1. Setup of the Numerical Experiments

We consider the linear demand function that is frequently used in the pricing literature. Let p be the price offered. Then we define the linear demand function and its associated revenue rate by

$$\lambda^{lin}(p) = \Lambda_e(1 - ap), \qquad p \in [0, 1/a], \tag{3}$$

$$r^{lin}(p) = \Lambda_e p(1 - ap), \tag{4}$$

where a is a positive real number. We note that r^{lin} is maximized at $p^{lin} = (2a)^{-1}$, regardless of the potential demand rate Λ_e. In our numerical study, we also considered the exponential demand function $\lambda^{exp}(p) = \Lambda_e e^{-ap}$. However, here we only report the results for the linear demand function, since our main conclusions are similar for both functions. Finally, we assume a standard linear holding cost function $h(x) = hx$.

Since we use the long-run average criterion, we restrict our attention to the recurrent states of the Markov chain generated by an optimal policy. Let $\mathcal{R}_\phi$ be the set of recurrent states when the optimal policy is employed under pricing strategy ϕ. The base-stock levels in each environment can be different from each other, so define $\bar{s}^\phi = \max_{e \in E}\{s_e^\phi\}$ as the maximum base-stock level. Then

$$\mathcal{R}^\phi = \{(x, e) : 0 \leq x \leq \bar{s}^\phi, \ 1 \leq e \leq N\}.$$

Now, we can also define the minimum and maximum dynamic prices in environment e by $\underline{p}_e^{dp} = p^{dp}(\bar{s}^{dp}, e)$ and $\bar{p}_e^{dp} = p^{dp}(1, e)$, respectively, by Theorem 2.

For a given problem, let g^ϕ be the optimal average profit using strategy ϕ ($\phi = s, sb, sp, edp, dp$). Notice that $g^s \leq g^{sb} \leq g^{edp} \leq g^{dp}$ and $g^s \leq g^{sp} \leq g^{edp} \leq g^{dp}$.

The computational procedure to evaluate g^s, g^{sb}, g^{sp}, g^{edp} and g^{dp} and their associated policies is summarized in Appendix E. We then define the relative profit gain for using policy ϕ instead of policy ϕ', $PG^{\phi,\phi'}$, by

$$PG^{\phi,\phi'} = \frac{g^\phi - g^{\phi'}}{g^{\phi'}}.$$

5.2. Single-Environment System

In this Subsection, we assume the system operates in a single environment with potential demand rate Λ. The structure of the optimal policy was investigated by Li [19], but no numerical experiment was carried out to investigate the benefit of a DP strategy versus a S strategy. Note that in this setting, S, SB, SP, and EDP strategies are equivalent; therefore, we will only consider the profit gain $PG = PG^{dp,s}$ in this subsection.

In a single-environment setting, any problem is described by the five parameters (Λ, a, μ, h, c). Now, we show how to reduce the problem to two parameters: the production rate μ and the holding cost h. The optimal policy depends neither on the time unit nor on the monetary unit. Therefore, we can set $a = 1$ and $\Lambda = 1$ in all the numerical experiments, without losing any generality. To illustrate this point, we observe that the problem (Λ, a, μ, h, c) and the problem $(1, 1, \mu/\Lambda, ha/\Lambda, ac)$ are equivalent. The profit gains and the production decisions are identical and the optimal prices of the first problem can be obtained as a times the optimal prices of the second one. Moreover, we take the unit production cost $c = 0$. This does not affect the generality of the results since the following two problems are equivalent: (I) a problem with a unit production cost of c ($c > 0$) and demand function $\lambda(p)$ and (II) a problem with zero production cost and demand function $\tilde{\lambda}(p) = \lambda(p + c)$.

In the following, we address the effects of holding costs, production rate and number of possible prices on the benefits of DP.

5.2.1. Comparison of pricing strategies. Figure 1 shows the relative profit gain (PG) as a function of the service rate, μ, for different holding costs, h. In all of the curves, the improvement due to DP manifests a nonmonotonic behavior in μ. In particular, the profit gain is low when the production capacity is extremely low or extremely high with respect to the potential demand rate. The profit gain peaks for intermediate levels of μ. When the production capacity is high, the replenishment policy is sufficient to manage the system and the impact of DP in this case is limited. When the production capacity is very low, the average inventory is also very low and the manager sells most of the time at the maximum price and the DP strategy and the S strategy are comparable.

Since the profit gain depends only on h and μ, we were able to evaluate the maximum profit gain that can be achieved. We obtained that the maximum profit gain, $\max_{h,\mu} PG^{dp,s}$, is 3.81% (realized when $\mu = 0.255$ and $h = 0.0123$). Apart from this extreme case, the main observation is that, in a single-environment system, when the static price is chosen effectively, the potential impact of DP is limited.
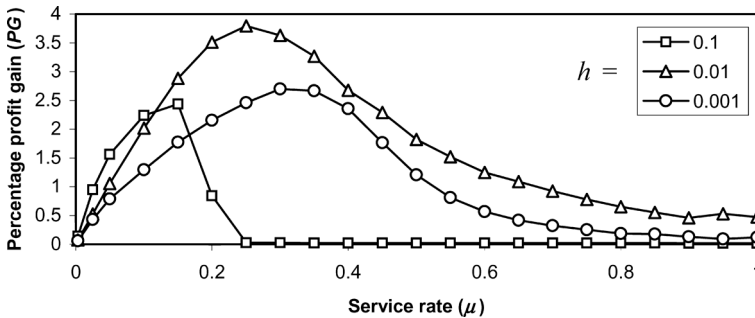
FIGURE 1. Percentage profit gain ($PG = PG^{dp,sp}$) as a function of the service rate μ and the holding cost h.

5.2.2. Impact of the price menu. In this Subsection, we address the following question: If the decision-maker prefers to use only a few prices in $\mathcal{P}$ for the DP strategy, how much can he or she expect to gain relative to the S policy? Table 2 shows the improvements due to DP when two prices, three prices, or an infinite number of prices can be chosen, for different service rates μ, with $h = 0.01$. The procedure to compute the optimal average cost with two and three prices is detailed in Appendix E. The maximum improvement is, obviously, obtained by continuous pricing, however, the percent improvements with two and three possible price values are quite high when compared to the maximum improvement. In fact, in the 20 experiments with varying service rates, $\mu \in \{0.05, 0.10, \ldots, 1.00\}$, the average benefit of obtained by a three-price menu is 92.5% of the average benefit of continuous pricing, whereas that of a two-price menu is 78.5%. Hence, we can reach the conclusion that if the values for possible prices are chosen effectively, three or even two different price values brings most of the benefit that can be expected from DP. This is a useful observation for the practice of DP since companies might prefer to have only a few different prices to reduce the potential negative customer reaction.

5.3. Two-Environment System

In this Subsection, we assume that the system operates in two environments: Low (L) and High (H) having potential demand rates $\Lambda_L = \Lambda - \epsilon$ and $\Lambda_H = \Lambda + \epsilon$,

TABLE 2. Relative Profit Gain due to Different Number of Prices in DP

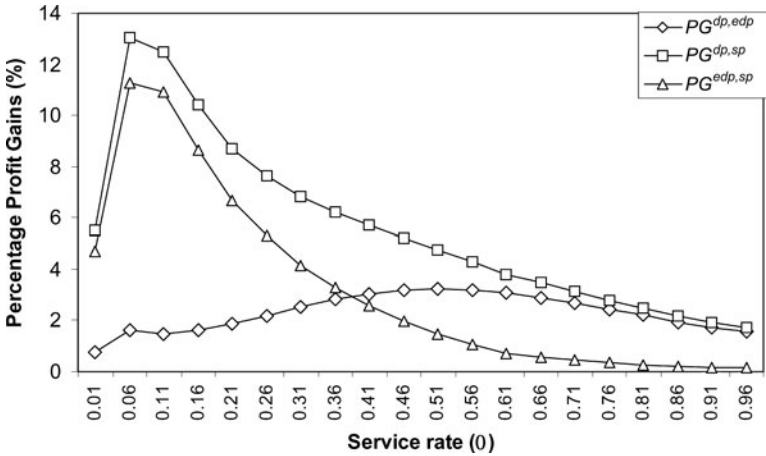| μ	Two Prices	Three Prices	Continuous
0.1	1.5%	1.9%	2.0%
0.3	2.7%	3.2%	3.6%
0.5	1.4%	1.7%	1.8%
0.7	0.7%	0.9%	0.9%
0.9	0.4%	0.4%	0.5%

FIGURE 2. Percentage profit gains ($PG^{dp,edp}$, $PG^{dp,sp}$, $PG^{edp,sp}$) as a function of the service rate μ for $\epsilon = 0.8$.

respectively, with $\epsilon \geq 0$. As ϵ increases, the fluctuations become more significant, and as Λ increases, the average potential demand increases. The case $\epsilon = 0$ is equivalent to a single-environment system with potential demand $\Lambda_H = \Lambda_L$. Moreover, we assume symmetric environment transition rates so that $q_{HL} = q_{LH} = q$. The focus of this Subsection is on the influence of the fluctuations represented by ϵ on the performance of the five strategies. More precisely, we compare the optimal average profits, the optimal base-stock levels, and the optimal prices.

The problem can be exhaustively described by the seven parameters $(\Lambda, a, \mu, h, c, \epsilon, q)$. As for the single-environment case, we set, without losing generality, $a = 1$, $\Lambda = 1$, and $c = 0$. Now, since q corresponds to the rate of changes in economic environments, it has to be much smaller than the average demand rate 1 and we set $q = 0.01$. We also have $\mu = 0.11$ in the following subsections, except for Figure 2, mainly because the relative profit gains are higher for this value, as Figure 2 will show. Finally, we fix $h = 0.01$, a reasonable value when the prices change in the interval $\mathcal{P} = (0, 1)$.

5.3.1. Comparison of Average Profits. Table 3 presents the profit gains for different values of ϵ. Recall that as ϵ increases, the fluctuations increase. We observe that optimal average profit for each pricing policy decreases with ϵ, which can be seen as the degrading effect of the fluctuating environment and the additional demand uncertainty it brings. The profit gain of SB, SP, EDP, and DP strategies, with respect to the S strategy, increases with ϵ, which confirms the capability of these policies to adjust the highly uncertain environments. We see that the ability to change the prices according to the environment brings more benefit than that to adjust the base-stock levels: a maximum of 12% for $PG^{sb,s}$ versus a maximum of 2.4% for

TABLE 3. Profit Gains for Different Demand Variabilities

ϵ	$PG^{sb,s}$	$PG^{sp,s}$	$PG^{edp,s}$	$PG^{dp,s}$
0	0.0%	0.0%	0.0%	2.2%
0.3	1.5%	0.0%	1.5%	3.8%
0.6	7.3%	0.5%	7.4%	10.0%
0.8	12.0%	2.4%	13.6%	15.2%

$PG^{sp,s}$. Moreover, if the base-stock levels are controlled in addition to the prices, the incremental benefit is very marginal, since the maximum of $PG^{edp,s}$ is only 13.6%, whereas it is 12% for $PG^{sb,s}$.

Now, we compare the performance of the main three strategies we considered: The benefit of the DP strategy with respect to the SP strategy can be significant in a two-environment system (sometimes >12%), contrary to a single-environment system (always <3.81%). We also observed that $\max_{\mu}\{PG^{dp,edp}\}$ is 3.38%, 2.89%, and 3.24% for $\epsilon = 0.3$ (with $\mu = 0.24$), $\epsilon = 0.6$ (with $\mu = 0.44$), and $\epsilon = 0.8$ (with $\mu = 0.5$), respectively. From these observations, we can also conclude that the DP strategy does not improve over the EDP strategy significantly. It appears that adjusting prices only when the environment changes is, most of the time, sufficient to obtain most of the benefit over the SP strategy. Moreover, in this numerical example, adjusting base-stock levels when the environment changes is not as beneficial as adjusting prices.

Now, we examine the effect of service rate on the relative profit gains. Since the differences between strategies are higher for higher values of ϵ, the relative profit gains are plotted as a function of service rate μ for $\epsilon = 0.8$ in Figure 2. All profit gains are nonmonotonic in μ. We observe that the SP strategy performs the worst when the processing capacity is relatively tight. As the processing rate increases, the excess processing capacity seems to compensate for the fluctuations so that the profit gains are relatively low (under 4%) when $\mu > 0.6$. Moreover, $PG^{dp,edp}$ is generally very low, with a maximum of 3.23%, pointing out again the effectiveness of EDP strategies. This observation supports our earlier conclusion on the comparable performance of EDP and DP strategies.

In other numerical results (not reported here), we have also explored the effects of the transition rate q between the economic environments. As the frequency of environment state transitions increase, the benefit of DP tends to decrease in general. One plausible explanation is that the pricing strategy cannot be applied long enough to properly gain all the possible benefit when the environment fluctuates frequently compared to the demand arrival rate. At the other extreme, when the environment fluctuations become very slow, there are two weakly coupled systems corresponding to each environment state and DP can be adjusted to each of these systems to reap the full benefits.

5.3.2. Observations on Optimal Base-Stock Levels and Prices.   In Section 4.3, under certain conditions on the MMPP we show that optimal base-stock

TABLE 4. Optimal Base-Stock Levels for Different ϵ

ϵ	s^s	s^{sb}	s_L^{sp}	s_H^{sp}	s_L^{edp}	s_H^{edp}	s_L^{dp}	s_H^{dp}
0	8	8	8	8	8	8	17	17
0.3	7	8	6	11	7	9	12	20
0.6	5	6	4	14	5	10	7	22
0.8	3	4	2	13	3	10	3	23

levels are ordered in the demand rates of the environments for strategies SP and EDP. This is confirmed in Table 4, which presents the optimal base-stock levels of the five pricing strategies for different values of ϵ. We see not only that $s_L^{sp} \leq s_H^{sp}$ and $s_L^{edp} \leq s_H^{edp}$ but also that $s_L^{dp} \leq s_H^{dp}$. This is quite intuitive, since the base-stock levels should be tuned to the demand rates. We also observe that $s_e^{\phi} \leq s_e^{dp}$, for $\phi = s, sb, sp, edp$ in all experiments, which is a natural consequence of the flexibility in the DP strategy: The system can store more inventory if it can sell easily when needed (by adjusting prices). Finally, as the fluctuations in demand (ϵ) increase, Λ_L decreases and Λ_H increases, which causes s_L^{ϕ} to decrease and s_H^{ϕ} to increase for all pricing strategies, $\phi = sp, edp, dp$, so that the gap between the base-stock levels in two environments increases. Finally, we observe that the SB strategy, while bringing most of the profit gain of the DP strategy, has lower base-stock levels than the average base-stock levels of the SP, EDP, and DP strategies.

5.3.3. Observations on Optimal Prices.
Table 5 presents optimal prices for all strategies for different ϵ. For prices of S and SB strategies, we observe that $p_L^{sb} \leq p^s \leq p_H^{sb}$. For prices of SP and EDP strategies, we observe that $p_L^{edp} \leq p^{sp} \leq p_H^{edp}$. For prices of EDP and DP strategies, we have $\underline{p}_e^{dp} \leq p_e^{edp} \leq \bar{p}_e^{dp}$ for both environments L and H.

Let Δp^{ϕ} be the gap between the highest and lowest price of strategy ϕ, in the recurrent region. Then $\Delta p^s = \Delta p^{sp} = 0$, $\Delta p^{sb} = p_H^{sb} - p_L^{sb}$, $\Delta p^{edp} = p_H^{edp} - p_L^{edp}$, and $\Delta p^{dp} = \bar{p}_H^{dp} - \underline{p}_L^{dp}$. As ϵ increases, p_L^{sb}, p_L^{edp}, and $\underline{p}_L^{dp}$ decrease while p_H^{sb}, p_H^{edp}, and $\bar{p}_H^{dp}$ increase to compensate the greater fluctuations of the demand. As a result, the price gap Δp^{ϕ} is increasing with ϵ for the DP, SB, and EDP strategies. Moreover, the gap is

TABLE 5. Optimal Prices for Different ϵ, Where $\bar{p}_e^d = \max\{p^{dp}(x, e)\}$, and $\underline{p}_e^{dp} = \min\{p^{dp}(x, e)\}$

ϵ	p^s	p_L^{sb}	p_H^{sb}	p^{sp}	p_L^{edp}	p_H^{edp}	$\underline{p}_L^{dp}$	$\bar{p}_L^{dp}$	$\underline{p}_H^{dp}$	$\bar{p}_H^{dp}$
0	0.79	0.79	0.79	0.79	0.79	0.79	0.50	0.85	0.50	0.85
0.3	0.78	0.74	0.82	0.78	0.74	0.82	0.42	0.82	0.51	0.87
0.6	0.74	0.65	0.83	0.75	0.65	0.84	0.33	0.75	0.51	0.88
0.8	0.75	0.55	0.84	0.78	0.57	0.84	0.19	0.65	0.51	0.99

much greater for a DP strategy than for an SB or an EDP strategy. For instance, when $\epsilon = 0.8$, we have $\Delta p^{dp} = 0.8$ while $\Delta p^{sb} = 0.29$ and $\Delta p^{edp} = 0.27$.

5.3.4. Overall Comparison of Strategies.   We observe in both Table 3 and Figure 2 that the EDP strategy achieves most of the benefits of a DP strategy. Moreover, the EDP strategy has the advantage of lower base-stock levels (see Table 4) and of smaller price differences (see Table 5). It is then safe to conclude that the EDP strategy is an excellent compromise, since it brings most of the benefit of the DP strategy while causing less reaction on the customer side with less variability in prices and requiring a reasonable storage space with less variability in the stock levels.

6. CONCLUSION

We show that for all three pricing strategies, optimal replenishment policies are characterized as environment-dependent base-stock policies. For DP, this generalizes the corresponding results in Li [19] to a fluctuating environment setting. For EDP and SP strategies, we have shown the optimality of environment-dependent base-stock policies. With EDP and SP, we assume that the arrival rate in each environment is fixed. Hence, the analysis of the MDP in their first step corresponds to the analysis of optimal replenishment policies for systems with lost sales that operate in a fluctuating environment. Consequently, the monotonicity of the base-stock levels with respect to environments under certain conditions is similar to the monotonicity properties of environment-dependent base-stock policies (see, e.g., Song and Zipkin [26]) in uncapacitated systems. To our knowledge, these are the first corresponding monotonicity results for a make-to-stock queue.

For the DP strategy, prices are shown to be decreasing in the inventory level for each environment e, which is again an extension of Li's results to systems operating in a fluctuating environment. This type of monotonicity agrees with the earlier results in different settings; for instance, Federgruen and Heching [12] presented corresponding results for a discrete-time finite horizon problem with nonstationary demand.

Within the context of revenue management, DP has received considerable attention (see, e.g., Caldentey and Wein [4]) in situations where an inventory has to be sold within a finite horizon without any replenishment opportunities. As articulated by Gallego and van Ryzin [16], there seem be two main reasons for employing DP. The first argues that it is a tool for handling statistical fluctuations in a random—but stationary—demand process. The second views it as an approach for responding to a shifting and possibly unpredictable demand function. Their analysis reveals that in the case of a stationary demand process, a static price policy is nearly optimal and that the benefits from DP are relatively small. At the same time, dynamic price adjustments, using a few prices, turn out to be well justified in the case of a time-varying demand process. The results in this article confirm these findings in a different setting. In particular, it is found that the profit improvement over SP by DP is modest, with a maximum that is less than 4% when the potential demand is a stationary

Poisson process. Although such a benefit might be considered high in certain settings, it should be considered a rather optimistic assessment since our model does not take into account certain intangible negative effects such as the burden of frequent price changes, negative customer reactions, and so forth.

On the other hand, similar to Gallego and van Ryzin [16], we observe that when the potential demand rate fluctuates randomly, there are significant improvements that could be achieved with respect to SP. The real value of adjusting the prices in our context appears to be smoothing the effects of demand rate fluctuations rather than smoothing the statistical fluctuations when the demand rate is constant. However, once again, full DP might not be the most appropriate means of achieving this end. It turns out that EDP policies that only adjust prices when the demand environment fluctuates are almost as effective as DP and mitigate its certain undesirable effects such as frequent price adjustments and negative consumer reaction. Extensions of this research that explicitly model consumer reactions to different pricing situations is an interesting path for future research.

Acknowledgments

References

1. Beyer, D. & Sethi, S.P. (1997). Average cost optimality in inventory models with markovian demands. *Journal of Optimization Theory Applications* 92: 497–526.
2. Bitran, G. & Caldentey, R. (2003). An overview of pricing models for revenue management. *Manufacturing and Service Operations Management* 5(3): 203–229.
3. Buzacott, J.A. & Shanthikumar, J.G. (1993). *Stochastic models of manufacturing systems*. Englewood Cliffs, NJ: Prentice Hall.
4. Caldentey, R. & Wein, L.M. (2006). Revenue management of a make-to-stock queue. *Operations Research* 54(5): 859–875.
5. Chan, L.M.A., Shen, Z.J.M., Simchi-Levi, D. & Swann, J. (2004). Coordination of pricing and inventory decisions: A survey and classification. In D. Simchi-Levi, S.D. Wu, & Z.J.M. Shen (eds.), *Handbook of quantitative supply chain analysis: Modeling in the e-business era*. Amsterdam: Kluwer, pp. 335–392.
6. Chan, L.M.A., Simchi-Levi, D. & Swann J. (2006). Dynamic pricing strategies for manufacturing with stochastic demand and discretionary sales. *Manufacturing and Service Operations Management* 8: 149–168.
7. Chen, H., Wu, O. & Yao, D.D. (2004). Optimal pricing and replenishment in a single-product inventory system. Working paper, Cheung Kong Graduate School of Business, University of British Columbia and Columbia University.
8. Chen, X. & Simchi-Levi, D. (2002). Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The continuous review model. Working paper, Operations Research Center, MIT, Cambridge, MA.
9. Chen, X. & Simchi-Levi, D. (2004). Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case. *Operations Research* 52: 887–896.
10. Chen, X. & Simchi-Levi, D. (2004). Coodinating inventory control and pricing strategies with random demand and fixed ordering cost: The infinite horizon case. *Mathematics of Operations Research* 29(3): 698–723.
11. Elmaghraby, W. & Keskinocak, P. (2003). Dynamic pricing in the presence of inventory considerations: Research overview, current practices and future directions. *Management Science* 49(10): 1287–1309.

12. Federgruen, A. & Heching, A. (1999). Combined pricing and inventory control under uncertainty. *Operations Research* 47(3): 454–475.
13. Feldman, R.M. (1978). A continuous review (*s*, *S*) inventory system in a random environment. *Journal of Applied Probability* 15: 654–659.
14. Feng, Y. & Chen, F.Y. (2003). Joint pricing and inventory control with setup costs and demand uncertainty. Working paper, The Chinese University of Hong Kong.
15. Feng, Y. & Chen, F. (2004). Optimality and optimization of a joint pricing and inventory-control policy for a periodic-review system. Working paper, The Chinese University of Hong Kong.
16. Gallego, G. & van Ryzin, G. (1994). Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science* 40(8): 999–1020.
17. Kalpakam, S. & Arivarignan, G. (1989). A lost sales system in a random environment. *Stochastic Analysis and Applications* 7: 367–385.
18. Koole, G. (1998). Structural results for the control of queueing systems using event-based dynamic programming. *Queueing Systems* 30: 323–339.
19. Li, L. (1988). A stochastic theory of the firm. *Mathematics of Operations Research* 13(3): 447–466.
20. McGill, J.I. & Van Ryzin, G.J. (1999). Revenue management: Research overview and prospects. *Transportation Science* 33(2): 233–256.
21. Özekici, S. & Parlar, M. (1999). Inventory models with unreliable suppliers in a random environment. *Annals of Operations Research* 91: 123–136.
22. Porteus, E.L. (2002). *Foundations of stochastic inventory theory*. Stanford, CA: Stanford University Press.
23. Puterman, M.L. (1994). *Markov decision processes*. New York: John Wiley and Sons.
24. Serfozo, R.F. (1979). An equivalence between continuous and discrete time markov decision processes. *Operations Research* 27(3): 616–620.
25. Sethi, S.P. & Cheng, F. (1997). Optimality of (*s*, *s*) policies in inventory models with markovian demand. *Operations Research* 45: 931–939.
26. Song, J.S. & Zipkin, P.H. (1993). Inventory control in a fluctuating demand environment. *Operations Research* 41: 351–370.
27. Thomas, L.J. (1974). Price and production decisions with random demand. *Operations Research* 22: 513–518.
28. Thowsen, G.T. (1975). A dynamic, nonstationary inventory problem for a Price/Quantity setting firm. *Naval Research Logistics* 22: 461–476.
29. Yano, C.A. & Gilbert, S.M. (2002). Coordinated pricing and production/procurement decisions: A review. In A. Chakravarty & J. Eliashbert (eds.), *Managing business interfaces: Marketing, engineering and manufacturing perspectives*. Amsterdam: Kluwer Academic.
30. Zabel, E. (1970). Monopoly and uncertainty. *The Review of Economic Studies* 37: 205–219.
31. Zabel, E. (1972). Multi-period monopoly under uncertainty. *Economic Theory* 5: 524–536.
32. Zipkin, P.H. (2000). *Foundations of inventory management*. New York: McGraw-Hill.

APPENDIX A
Proof of Lemma 1

Assume that f is concave or equivalently that Δf is decreasing in x. We can consider the terms separately: By assumption, $-h$ is concave. Concavity is preserved under the maximization operator (Porteus [22]); thus, $T^1 f$ is also concave. It is obvious that T^2 and T^λ also preserve concavity. Hence, T^{sp} and $T^{edp} f$ are concave as linear combinations of concave functions.

To prove that T^{dp} also preserves concavity, we must show that the operator T also preserves concavity. Koole [18] mentioned the fact that Tf is also concave. Let us provide a detailed proof.

We denote by $\mathcal{L}_f(x, e)$ the set of optimal arrival rates associated to value function f:

$$\mathcal{L}_f(x, e) = \arg\max_\lambda \{T^\lambda f(x, e)\}. \tag{A.1}$$

Let $\lambda(x-1, e) \in \mathcal{L}_f(x-1, e)$, $\lambda(x, e) \in \mathcal{L}_f(x, e)$, and $\lambda(x+1, e) \in \mathcal{L}_f(x+1, e)$. Assume that $x > 1$, then

$$\Delta Tf(x+1, e) - \Delta Tf(x, e)$$
$$= Tf(x+1, e) - Tf(x, e) - Tf(x, e) + Tf(x-1, e)$$
$$\leq T^{\lambda(x+1, e)}f(x+1, e) - T^{\lambda(x+1, e)}f(x, e) - T^{\lambda(x-1, e)}f(x, e) + T^{\lambda(x-1, e)}f(x-1, e) \tag{A.2}$$

$$= \lambda(x+1, e)\Delta f(x-1, e) + (\Lambda_e - \lambda(x+1, e))\Delta f(x, e)$$
$$\quad - \lambda(x-1, e)\Delta f(x-2) - (\Lambda_e - \lambda(x-1, e))\Delta f(x-1, e)$$
$$\leq \lambda(x+1, e)\Delta f(x-1, e) + (\Lambda_e - \lambda(x+1, e))\Delta f(x-1, e)$$
$$\quad - \lambda(x-1, e)\Delta f(x-1, e) - (\Lambda_e - \lambda(x-1, e))\Delta f(x-1, e) \tag{A.3}$$
$$= 0.$$

Inequality (A.2) follows from the definition of T and inequality (A.3) is a consequence of the concavity of f.

Assume now that $x = 1$.

$$\Delta Tf(1, e) - \Delta Tf(0, e)$$
$$= Tf(2, e) - Tf(1, e) - Tf(1, e) + Tf(0, e)$$
$$\leq T^{\lambda(2, e)}f(2, e) - T^{\lambda(2, e)}f(1, e) - T^{\Lambda_e}f(1, e) + \Lambda_e f(0, e)$$
$$= (\Lambda_e - \lambda(2, e))(\Delta f(1, e) - \Delta f(0, e)) - r_e(\Lambda_e)$$
$$\leq 0.$$

Thus, T_f is concave. Also, we conclude by linear combination that $T^{dp}f$ is concave.

APPENDIX B
Proof of Theorem 2

Let $x < y$, $\lambda(x, e) \in \mathcal{L}_f(x, e)$ (defined by (A.1)), and $\lambda(y, e) \in \mathcal{L}_f(y, e)$. We prove Theorem 2 by contradiction. Assume that $\lambda(x, e) > \lambda(y, e)$. Then

$$Tv^{dp}(y, e) = T^{\lambda(y, e)}v^{dp}(y, e)$$
$$= T^{\lambda(y, e)}v^{dp}(x, e) + \lambda(y, e)[\Delta v^{dp}(x-1, e) - \Delta v^{dp}(y-1, e)]$$
$$\quad + \Lambda_e(v^{dp}(y, e) - v^{dp}(x, e))$$
$$< T^{\lambda(x, e)}v^{dp}(x, e) + \lambda(x, e)[\Delta v^{dp}(x-1, e) - \Delta v^{dp}(y-1, e)]$$
$$\quad + \Lambda_e(v^{dp}(y, e) - v^{dp}(x, e))$$
$$= T^{\lambda(x, e)}v^{dp}(y, e).$$

The inequality $Tv^{dp}(y,e) < T^{\lambda(x,e)}v^{dp}(y,e)$ is contradictory to the definition of T. Therefore, the assumption $\lambda(x,e) > \lambda(y,e)$ is false and we conclude that $\lambda(x,e) \leq \lambda(y,e)$; thus, $p(x,e) \geq p(y,e)$.

APPENDIX C
Proof of Corollary 1

Assume that the revenue rate $r_e(\lambda)$ is strictly concave. Then $T^\lambda v^{dp}(x,e)$ is also strictly concave with respect to λ, which ensures the uniqueness of the maximizer of $T^\lambda v^{dp}(x,e)$ and implies the first part of Corollary 1. The unique maximizer, $\lambda(x,e)$, satisfies

$$r'_e(\lambda(x,e)) = \Delta v^{dp}(x-1,e) \tag{C.1}$$

and r'_e is decreasing by strict concavity of r_e. On the other hand, by definition of the optimal policy, we have

$$\begin{cases} \Delta v^{dp}(x-1,e) \geq 0 & \text{if } x \leq s_e^{dp}, \\ \Delta v^{dp}(x-1,e) \leq 0 & \text{if } x > s_e^{dp}. \end{cases} \tag{C.2}$$

We deduce, from (C.1) and (C.2) and monotonicity of r'_e that

$$\begin{cases} \lambda(x,e) \leq \lambda(p_e^*) & \text{if } x \leq s_e^{dp}, \\ \lambda(x,e) \geq \lambda(p_e^*) & \text{if } x > s_e^{dp}. \end{cases} \tag{C.3}$$

Finally, as $p_e(\lambda)$ is decreasing in λ, we obtain the second part of Corollary 1.

APPENDIX D
Proof of Lemma 2

Assume that environments are labeled such that $\lambda_e \leq \lambda_{e+1}$ and $\lambda_e p_e \leq \lambda_{e+1}p_{e+1}$ for all $e = 1, \ldots, N-1$, Condition 1 holds, and f satisfies $\Delta f(x,e+1) \geq \Delta f(x,e)$ for all x and for $1 \leq e \leq N-1$. Now, for $\phi = sp, edp$, we will show

$$\Delta T^\phi f(x,e+1) - \Delta T^\phi f(x,e) \geq 0.$$

We will consider each possible transition separately. First, note that the terms for inventory costs, h, cancel out so that they satisfy the inequality. For $x > 0$, the terms for the revenue $\lambda_e p_e$ also cancel out. However, for $x = 0$, the revenue terms in the above inequality become

$$\lambda_{e+1}p_{e+1} - 0 - \lambda_e p_e - 0 \geq 0,$$

since a revenue is obtained when the inventory level is 1 and no revenue is gained when $x = 0$. This inequality is true by our assumption on the increasing order of revenue rates in environment labels. On the other hand, the inequality remains the same for all common terms in the fictitious transitions of environments e and $e+1$ (due to q_{ij} (with $i \neq e$ and $i \neq e+1$), Λ_i (with $i \neq e$ and $i \neq e+1$), and $\Lambda_i - \lambda_i$ (with $i = e$ and $i = e+1$)) by the assumption on f. Hence, we

need to consider only the transitions due to a demand arrival, a replenishment decision, and an actual environment transition. For these transitions we will present the proof only for $x > 0$, since the case $x = 0$ is similar.

Consider the actual and fictitious demand rates in environments e and $e + 1$, so that in environment e, λ_e is the actual demand rate and λ_{e+1} is the fictitious, one and vice versa in $e + 1$:

$$
\begin{aligned}
\lambda_{e+1}&\Delta f(x-1,e+1) - \lambda_e \Delta f(x-1,e) + \lambda_e \Delta f(x,e+1) - \lambda_{e+1}\Delta f(x,e) \\
&= (\lambda_{e+1} - \lambda_e)\Delta f(x-1,e+1) + \lambda_e(\Delta f(x-1,e+1) - \Delta f(x-1,e)) \\
&\quad + \lambda_e \Delta f(x,e+1) - \lambda_{e+1}\Delta f(x,e) \\
&\geq (\lambda_{e+1} - \lambda_e)\Delta f(x,e+1) + \lambda_e(\Delta f(x-1,e+1) - \Delta f(x-1,e)) \\
&\quad + \lambda_e \Delta f(x,e+1) - \lambda_{e+1}\Delta f(x,e) \\
&= \lambda_{e+1}(\Delta f(x,e+1) - \Delta f(x,e)) + \lambda_e(\Delta f(x-1,e+1) - \Delta f(x-1,e)) \\
&\geq 0,
\end{aligned}
$$

where the first inequality holds since $\Delta f(x-1,e+1) \geq \Delta f(x,e+1)$ for all e by Lemma 1 and the second inequality holds by our assumption on f; the rest is some algebra.

Now, consider the transitions due to the replenishment decision, so we need to show that

$$
\begin{aligned}
\Delta T^0 f(x,e+1) - \Delta T^0 f(x,e) &= T^0 f(x+1,e+1) - T^0 f(x,e+1) - T^0 f(x+1,e) \\
&\quad + T^0 f(x,e) \geq 0.
\end{aligned}
$$

Let $a(x,e) = 1$ if it is optimal to replenish in state (x,e), and $a(x,e) = 0$ otherwise. Then we can write $T^0 f$ as

$$
T^0 f(x,e) = f(x + a(x,e), e) - a(x,e)c.
$$

We set $a = a(x,e+1)$ and $a' = a(x+1,e)$. Then

$$
\begin{aligned}
T^0 f(x+1,&e+1) - T^0 f(x,e+1) - T^0 f(x+1,e) + T^0 f(x,e) \\
&\geq f(x+1+a',e+1) - a'c - f(x+a,e+1) \\
&\quad + ac - f(x+1+a',e) + a'c + f(x+a,e) - ac \\
&= f(x+1+a',e+1) - f(x+a,e+1) - f(x+1+a',e) + f(x+a,e),
\end{aligned}
$$

where the inequality holds since T^0 is a maximizing operator so that action a and a' always perform worse than the optimal actions in states (x,e) and $(x+1,e+1)$, respectively. If $a = a'$, the above expression is positive by the assumption on f. Consider the two cases:

Case 1: $a = 1$ and $a' = 0$:

$$
f(x+1,e+1) - f(x+1,e+1) - f(x+1,e) + f(x+1,e) = 0.
$$

Case 2: $a = 0$ and $a' = 1$:

$$f(x + 2, e + 1) - f(x, e + 1) - f(x + 2, e) + f(x, e)$$
$$= \Delta f(x + 1, e + 1) + \Delta f(x, e + 1) - \Delta f(x + 1, e) - \Delta f(x, e)$$
$$\geq 0$$

by the assumption on f. Note that we add and subtract the terms $f(x + 1, e + 1)$ and $f(x + 1, e)$ in the equality.

Thus, the inequality holds for the replenishment decisions.

We finally consider the transitions due to environment, so we need to show

$$\sum_{j \neq e+1} q_{e+1,j} \Delta f(x, j) + \sum_{j \neq e} q_{e,j} \Delta f(x, e + 1) - \sum_{j \neq e} q_{e,j} \Delta f(x, j) - \sum_{j \neq e} q_{e+1,j} \Delta f(x, e) \geq 0.$$

The terms corresponding to $q_{e,e+1}$ and $q_{e+1,e}$ cancel out, and we analyze the rest in two groups. The first inequality we will show is

$$\sum_{j=1}^{e-1} q_{e,j}(\Delta f(x, e + 1) - \Delta f(x, j)) + \sum_{j=1}^{e-1} q_{e+1,j}(\Delta f(x, j) - \Delta f(x, e)) \geq 0.$$

Now, by Condition 1, we have $q_{e,j} \geq q_{e+1,j}$ for all $j = 1, 2, \ldots, e - 1$. Moreover, by our assumption on f, $\Delta f(x, e + 1) \geq \Delta f(x, j)$ for all $j = 1, 2, \ldots, e - 1$. Then we have

$$\sum_{j=1}^{e-1} q_{e,j}(\Delta f(x, e + 1) - \Delta f(x, j)) + \sum_{j=1}^{e-1} q_{e+1,j}(\Delta f(x, j) - \Delta f(x, e))$$

$$\geq q_{e+1,j} \left(\sum_{j=1}^{e-1}(\Delta f(x, e + 1) - \Delta f(x, j)) + \sum_{j=1}^{e-1}(\Delta f(x, j) - \Delta f(x, e)) \right)$$

$$= q_{e+1,j} \left(\sum_{j=1}^{e-1}(\Delta f(x, e + 1) - \Delta f(x, e)) \right)$$

$$\geq 0.$$

Now consider the second group:

$$\sum_{j=e+2}^{N} q_{e+1,j}(\Delta f(x, j) - \Delta f(x, e)) + \sum_{j=e+2}^{N} q_{e,j}(\Delta f(x, e + 1) - \Delta f(x, j))$$

$$\geq q_{e,j} \left(\sum_{j=e+2}^{N} (\Delta f(x, j) - \Delta f(x, e)) + \sum_{j=e+2}^{N} (\Delta f(x, e + 1) - \Delta f(x, j)) \right)$$

$$= q_{e,j} \left(\sum_{j=e+2}^{N} (\Delta f(x, e + 1) - \Delta f(x, e)) \right)$$

$$\geq 0,$$

where the first inequality is true since $q_{e+1,j} \geq q_{e,j}$ and $\Delta f(x, j) - \Delta f(x, e) \geq 0$ for all $j = e + 2, e + 3, \ldots, N$ by Condition 1 and by our assumption on f, and the second inequality holds by our assumption on f.

Thus, we have shown that the inequality given in Lemma 2 holds for all (x, e) with $x \geq 1$ whenever Condition 1 holds.

APPENDIX E
Computational Procedure

To evaluate g^{dp} and its associated policy, we solve the dynamic programs corresponding to each problem instance using the value iteration method. The value iteration algorithm is terminated only when a five-digit accuracy is achieved. The state space is truncated at $[0, m]$, where m is a positive integer. The size of the state space is increased until the average cost is no longer sensitive to the truncation level. In the dynamic program, we compute the unique price $p(x, e)$ that maximizes the function $T^\lambda v^{dp}$. Then we obtain the following in the case of a linear demand function:

$$p(x, e) = \begin{cases} 0 & \text{if } a\Delta v^{dp}(x - 1, e) < -1 \\ \frac{1}{2a}(1 + a\Delta v^{dp}(x - 1, e)) & \text{if } -1 \leq a\Delta v^{dp}(x - 1, e) \leq 1 \\ \frac{1}{a} & \text{if } a > 1. \end{cases} \tag{E.1}$$

The procedure to evaluate g^{edp} is as follows: We discretize the interval $[0, 1]$ with increments of 0.01. For each possible set of prices, the optimal expected average profit obtained by applying the optimal replenishment policy is calculated by solving the corresponding dynamic programming equations. Then the maximum expected average profit obtained from all these prices gives g^{edp}, and the corresponding price is the optimal environment-dependent prices p_L^{edp}, and p_H^{edp}. The procedure to evaluate g^s, g^{sb}, and g^{sp} is similar except that we imposed the prices and/or base-stock levels to be identical in the two environments.

The procedure to evaluate the optimal DP average cost with two prices, g^2, is as follows. Let $g^2(p_1, p_2)$ be the optimal DP average cost when the set of possible prices is $\{p_1, p_2\}$, with p_1 and p_2 two elements of $\mathcal{P}$. Using the same discretization, we compute $g^2 = \max_{p_1, p_2}\{g^2(p_1, p_2)\}$.

To evaluate g^3, the optimal DP average cost with three prices is very long since we have to run a dynamic program for 100^3 values. Hence, we have only computed a lower bound by setting the third price as the average of the two first prices. Let $g^2(p_1, p_2, p_3)$ be the optimal DP average cost when the set of possible prices is $\{p_1, p_2, p_3\}$ with p_1, p_2, and p_3 elements of $\mathcal{P}$. Then the optimization is performed only over two variables and $g^3 = \max_{p_1, p_2}\{g^3(p_1, p_2, (p_1 + p_2)/2)\}$.