Optimal Markdown Pricing: Implications of Inventory Display Formats in the Presence of Strategic Customers

Rui Yin, Yossi Aviv, Amit Pazgal and Christopher S. Tang*

February 2009

Abstract

We propose a game-theoretical model of a retailer who sells a limited inventory of a product over a finite selling season by using one of two inventory display formats: Display All (DA) and Display One (DO). Under DA, the retailer displays all available units so that each arriving customer has perfect information about the actual inventory level. Under DO, the retailer displays only one unit at a time so that each customer knows about product availability but not the actual inventory level. Recent research suggests that when faced with strategic consumers, the retailer could increase expected profits by making an upfront commitment to a price path. We focus on such pricing strategies in this paper, and study the potential benefit of DO compared to DA, and its effectiveness in mitigating the adverse impact of strategic consumer behavior. We find support for our hypothesis that the DO format could potentially create an increased sense of shortage risk, and hence it is better than the DA format. However, while potentially beneficial, a move from DA to DO is typically very far from eliminating the adverse impact of strategic consumer behavior. We observe that, generally, it is not important for a retailer to modify the level of inventory when moving from a DA to a DO format; a change in the display format, along with an appropriate price modification, is typically sufficient. Interestingly, across all scenarios in which a change in inventory is significantly beneficial, we observed that only one of the following two actions takes place: either the premium price is increased along with a reduction in inventory, or inventory is increased along with premium price reduction. We find that the marginal benefit of DO can vary dramatically as a function of the per-unit cost to the retailer. In particular, when the retailer's per-unit cost is relatively high, but not too high to make sales unprofitable or to justify exclusive sales to high-valuation customers only, the benefits of DO appear to be at their highest level, and could reach up to 20% increase in profit. Finally, we demonstrate that by moving from DA to DO, while keeping the price path unchanged, the volatility of the retailer's profit decreases.

Key words: Retailing, Dynamic Pricing, Game Theory Applications, Marketing-Operations Interface, Strategic Customers, Revenue Management, Inventory Display.

*Corresponding author: Yin. Yin: W.P. Carey School of Business, Arizona State University, Tempe, AZ 85287, email: rui.yin@asu.edu; Aviv: The Olin Business School, Washington University, St. Louis, MO 63130, email address: aviv@wustl.edu; Pazgal: The Jones Graduate School of Management, Rice University, Houston, TX 77252, email: Amit.Pazgal@rice.edu; Tang: UCLA Anderson School, UCLA, 110 Westwood Plaza, Los Angeles, CA 90095, email address: ctang@anderson.ucla.edu.

1 Introduction

Many retailers use post-season clearance sales as a reactive response to dispose of unsold items at the end of a selling season. However, some retailers proactively pre-announce their price-markdown schedules at the beginning of the selling season. A well-known example of pre-announced markdown pricing strategy has been adopted by Filene's Basement since 1908. At the Filene's Basement Boston store, most unsold items after 2, 4 and 6 weeks will be sold at 25%, 50% and 75% off the regular price, respectively. After 2 months, Filene's Basement donates all unsold items to charity; see Bell and Starr (1998) for more details. Lands' End Overstocks uses a similar pre-announced markdown pricing strategy to sell their leftover inventory via its "On the Counter" website. Other stores such as Dress for Less[1], Tuesday Morning discount stores, Wanamaker discount department store in Philadelphia, etc., have adopted a similar pre-announced markdown pricing strategy. Also, TKTS ticket booths in New York City and London offer pre-announced discount tickets for the same day Broadway shows.

Pre-announced markdown pricing strategy is intended to segment customers with different product valuations so that high (low) valuation customers will purchase the product at the regular (clearance) price; see Pashigian and Bowen (1991) and Smith and Achabal (1998) for comprehensive discussions. However, one of the drawbacks of pre-announced pricing schemes is that they cannot segment customers completely because they often lead to "*strategic waiting*:" a phenomenon in which some high valuation customers may postpone their purchases by waiting for the post-season clearance price, even when there is a risk of not getting the product due to stockout at the end of the selling season (c.f., Phillips (2005) and Fisher (2006)).

Given that strategic waiting has detrimental effect on revenues, some retailers have considered various sales mechanisms to discourage high valuation customers from waiting to the post-season clearance price. Besides a corporate level strategy that calls for no markdown pricing (see discussion of such strategies in Aviv and Pazgal (2007), Cachon and Swinney (2008), and Su and Zhang (2008a)), many companies in the retail industry offer internal price-matching guarantees, in which a retailer promises to reimburse a customer for the difference between the current purchase price

[1] Current web address: www.d4less.net/discount.htm.

and any lower price the retailer might offer within a fixed future time period (see, e.g., Debo et al. (2008)). But there also exist some *operational strategies* for counteracting strategic consumer behavior. For example, a retailer may limit the supply of the product so that high-valuation customers would face a higher risk of stockouts if they decide to wait for the clearance price (see, e.g., Liu and van Ryzin (2008)). Another possible strategy, which is the key theme of our paper, is to control strategic behavior via the *inventory display format*. Here, the underlying principle is that a display format that conceals inventory information can be used as a tool to influence customers' perceptions about the risk of stockouts if they decide to wait.

In this paper we consider two types of display formats that are commonly seen in retailing. The first is called the *"Display All"* (DA) format under which the retailer displays all available units so that customers have *perfect* information about the actual inventory level. For example, since 2005, Expedia.com provide their customers with perfect information about the exact number of plane tickets available at a particular price for a particular flight. Similarly, Filene's Basement and Benetton adopt a similar in-store display format by placing all available items on the sales floor. The second display format is called the *"Display One"* (DO) format. Here, the retailer displays only one item at a time so that customers have *imperfect* information about the current inventory level. For example, Lands' End Overstocks website provides each arriving customer with information about the availability but not the actual inventory level of each product. Similarly, various high-end stores such as the Bally store in Taiwan displays only a single unit of a product at a time.

In addition to influencing customers' perceptions about product availability, the DA and DO display formats may offer different advantages and impose certain operational consequences. For example, the DA format allows a brick-and-mortar retailer to utilize store space more effectively, and eliminate the need for additional stockrooms. In fact, this is a fundamental reason for stores such as Benetton and Seven-Eleven to adopt the DA format. However, when the display space in a store is limited, the DO format may be the natural choice that enables a retailer to present an assortment of different designs instead of multiple units of the same design. The DO format can also result in a less cluttered shop floor, hence infusing a sense of exclusivity in the product. As a consequence, a retailer may enjoy an increase in store traffic and higher customer valuations for the

product. Implementation of the DO format can also be more costly, due to the required handling involved in the replacement of a new unit after each sale. The store may also incur temporary lost sales if a unit is not replaced immediately after a sale is made. Because of these explicit and implicit costs, the DO format is usually observed in high-end exclusive stores (e.g., Van Cleef & Arpels, Cartier, Tiffany) that enjoy high profit margin and low store traffic. Obviously, when the DO format is implemented in an online retail environment, the aforementioned costs are irrelevant. This is because the difference between the web page display formats boils down to the amount of inventory information communicated to customers.

In this paper, we compare the display formats on the basis of the inventory information conveyed to customers. In particular, we deliberately exclude the aforementioned handling cost, merchandising, and storage-efficiency considerations from our analysis. Consequently, our models can be used as a basis for a more comprehensive examination of the impact of display formats on revenue performance. Our main hypothesis is that the DO display format induces a larger sense of product scarcity than the DA format, since inventory information is only partial. As a consequence, we expect high-valuation customers to feel more pressed to purchase the product at the premium price, rather than wait for a discount. Obviously, our purpose is to *scientifically* test the validity of this hypothesis, and to gauge the extent to which one display format is better or worse than the other. It is noteworthy that such scientific comparison is complex. First, the term "sense of scarcity" in our hypothesis is loose. Would a lack of accurate inventory level information always induce a stronger urge to purchase at premium prices? Consider, for instance, a situation in which a large burst of customers arrive to the store very early in the season and purchase the product at the premium price. Suppose that shortly afterwards, only a single unit is left in stock. In this case, the DA format could have been better for the seller, because customers would know that the actual inventory level is very low, rather than speculate that it could be higher. In other words, depending on the realization of customer arrivals to the store, one display format might end up (in retrospect) being better than the other. The second complication is that the two display formats lend themselves to different equilibria, both in terms of the optimal pricing and in terms of the customer purchasing behavior. As a result, it is hard to predict how the two equilibria compare with each other on the basis of intuitive arguments only.

To our knowledge, this paper is the first to examine the impact of display format on revenues under the presence of strategic consumers. We propose a stylized model of a retailer that seeks to maximize his expected profit by selling a product to a customer base that consists of two classes of customers: one with a high valuation, and the other with a low valuation. The retailer and the customers engage in a competitive situation, as follows. Initially, the retailer determines his level of inventory, and announces a regular price for the main part of the season, as well as a clearance price for the end of the season. Obviously, in order to determine his optimal inventory and pricing strategies, the retailer needs to anticipate the customers' purchasing behavior. Customers make their purchasing decisions strategically; i.e., they either purchase the product immediately upon their arrival to the store ("buy now") or postpone their purchase to the end of the season ("wait"). Because customers' decisions have potential effects on each other, they are engaged in a competitive subgame among themselves.

The remainder of the paper is organized as follows. We first review the related literature in §2. In §3 we describe our model preliminaries. The DA format and the DO format are analyzed in §4 and §5, respectively. In each of these sections, we first analyze the subgame among customers and determine their optimal purchasing strategy for any given announced price path. We then examine the retailer's best strategy, and present a method for calculating the retailer's expected profit. In §6 we propose two benchmark models, and a set of metrics for the analysis of the difference between the display formats. We report the results of an extensive numerical study in §7. On the basis of this study, we identify the conditions under which a change from DA to DO could be valuable to the retailer. We discuss possible model extensions in §8 and conclude in §9.

2 Literature Review

We now provide a brief review of the relevant literature[2]. Our paper belongs to the stream of management science literature that studies the effects of strategic consumer behavior in the context of revenue management. We provide below a review of a representative group of related research papers. To keep the review concise and effective, we split the presentation into two parts. We first review models that assume complete knowledge of inventory information, and then continue

[2] We refer the reader to the online addendum for a more detailed version of this literature review.

to discuss models that consider partial inventory information. The primary contribution of our research is in providing a bridge between the two types of models. Specifically, our paper compares the impact of two display formats (one that provides customers with complete information, and one that only provides partial information) on consumer purchasing behavior and retailer's pricing, inventory levels, and expected profits.

Elmaghraby et al. (2007) belongs to the first category of models. It considers a setting in which a seller uses a pre-announced markdown pricing mechanism, to sell a finite inventory of a product. The authors study the potential benefits of segmentation; namely, the difference between the seller's profit under the optimal markdown mechanism to that under the optimal single price. They also provide a detailed discussion on the design of profitable markdown mechanisms. Su (2007) presents a pricing control model in which consumers are infinitesimally small and arrive continuously according to a deterministic flow of constant rate. The customer population is heterogeneous along two dimensions: valuations, and degree of patience (vis-a-vis waiting). Su demonstrates that the heterogeneity in valuation and degree of patience jointly influence the structure of optimal pricing policies. Levin et al. (2005) propose a stochastic game-theoretical dynamic pricing model of a monopolistic that sells a product to a population of strategic consumers. They demonstrate that a company that ignores strategic consumer behavior may receive much lower total revenues than one that uses the strategic equilibrium pricing policy. They also show that when the initial capacity is a decision variable, it can be used together with the appropriate pricing policy to effectively reduce the impact of strategic consumer behavior.

The following papers consider models with imperfect information about the prevailing level of inventory. Aviv and Pazgal (2008) examine two types of markdown pricing policies: *announced* fixed-discount policies, where prices are predetermined and known in advance of the beginning of the season, and *dynamic* pricing policies, that prescribe clearance prices on the basis of the prevailing level of inventory at the time of discount. In their models, customers only know the *initial* level of inventory, but need to form a belief about the *current* inventory level. One of their findings is that while intuition may suggest that a flexible, information-dependent policies should be better for the retailer, pre-announced pricing policies can actually generate higher expected profits. Caldentey and Vulcano (2007) analyze a revenue management problem in which a seller operates an online

multi-unit auction. Consumers can get the product either from the auction or from an alternative list price channel. The problem is related to the topic of strategic consumer behavior in that customers need to decide whether to buy at the posted list price and get the item immediately and at no risk, or to join the auction and wait until its end, when the winners are revealed and the auction price is disclosed. Cachon and Swinney (2008) analyze a two-period model of inventory planning and dynamic pricing, to study the value of *quick response*[3] under strategic consumer behavior. In their model, the initial level of inventory is not announced to the customers, but can be revealed through the calculation of an equilibrium in the game between the retailer and the customers. They argue that when replenishment is feasible, the retailer should avoid committing to a price path over the season, even in the presence of strategic consumers. They suggest that a better approach is to be cautious with the initial quantity, and then markdown optimally. They find that the value of quick response is generally much greater in the presence of strategic consumers than without them. Furthermore, they argue that quick response may have a dramatically larger influence as a mechanism for mitigating strategic consumer behavior, than as a tool for matching of supply with demand. Su and Zhang (2008a) also utilize a rational-expectations model of a seller facing strategic consumers. In their model, the items are sold at a regular price throughout the season and all unsold items are sold at a reduced price at the end of the season. Unlike our models, customers do not form an explicit *time-dependent* belief about the prevailing level of inventory. Rather, they all assumed to form the *same* belief about the likelihood of product availability, which is shown to hold under rational expectations. Liu and van Ryzin (2008) propose a rational expectation model to investigate whether it is optimal for a firm to create rationing risk by deliberately understocking products. Customers are assumed to form the *same* belief about the fill rate (i.e., likelihood of obtaining the product), on the basis of rational expectations. Su and Zhang (2008b) use a rational expectation game-theoretical framework to study the effect of inventory information on consumer purchase behavior. They analyze two strategies for the seller: A *commitment* to a particular level of initial inventory, and providing customers with *guarantees* on product availability (in the form of ex-post compensation if the product is out of stock). They demonstrate that the seller can improve his profits by making a combination of inventory commitment and availability guarantees.

[3] A quick response setting is one in which the retailer is able to replenish stock during the season.

It is noteworthy that while our paper also focuses on the aspect of inventory-related information, we are handling a fundamentally different problem. In our paper, the initial level of inventory is known, and the key question is whether or not to share the information about the prevailing level of inventory.

3 The Base Model

Consider a retailer who orders Q units (a decision) at a unit cost c that is scheduled to be sold over a selling season that spans over $[0, T]$. In our model, we assume that the retailer can place a single order prior to the start of the season; the order will be received and become available for sale at time 0. In addition to deciding on the order quantity Q, the retailer needs to decide and pre-announce two prices at time 0: the *premium* price p_h (i.e., the selling price throughout the entire season), and the post-season clearance price p_l. Clearly, $p_h \geq p_l$. All units not sold at either prices can be salvaged at s per unit. The objective of the retailer is to maximize his expected profit by making three inter-related decisions: Q, p_h and p_l.

Strategic customers[4] arrive at the store according to a Poisson process with rate λ, where λ is independent of the announced prices p_h and p_l. Upon arrival, each customer must take his own valuation as well as the announced price path into consideration when making his purchase decision. To capture market heterogeneity, customers are classified into two classes according to their valuations. Specifically, all customers belong to class-0 have valuation of v_0 and all customers belong to class-1 have valuation of v_1, where $v_0 < v_1$. We assume that the arrival process can be described as a combination of two independent Poisson processes associated with class-0 and class-1 customers. Specifically, we let α_0 be the portion of class-0 customers in the market, and $\alpha_1 = 1 - \alpha_0$ be the complementary portion of class-1 customers. Throughout this paper, we assume that the set of parameters $\{\alpha_0, \alpha_1, v_0, v_1, \lambda, T, c, s\}$ is a common knowledge. Admittedly, this assumption is made for mathematical tractability, since an inclusion of parameter uncertainty (or even structural uncertainty) would make our analysis prohibitively complex. It could be possible to extend our work by adding some level of parameterized uncertainty. Such model, however, will

[4] To simplify our analysis, we assume that all arriving customers are strategic. In §8, we discuss the complexity that arise when dealing with a mixture of strategic and "myopic" customers.

need to be based on an even stronger assumption that the seller and customers know the statistical characterization of the game parameters. Another possible approach could be to develop a model of bounded rationality, in which the lack of parameter knowledge (or lack of sophistication) leads the customers and the seller to adopt certain types of heuristics.

To ensure some potential sales at the premium price and to enable the retailer to facilitate effective price discrimination, it suffices to consider the case when the premium price p_h satisfies the inequality: $v_0 \leq p_h \leq v_1$. Because $p_l \leq p_h$, we need to consider two settings: (i) $v_0 < p_l \leq p_h \leq v_1$, and (ii) $p_l \leq v_0 \leq p_h \leq v_1$. The first case corresponds to a setting in which the retailer posts prices that exclude class-0 customers. Such strategy can be desirable especially when the market consists primarily of high valuation customers or when v_1 is significantly larger than v_0. Obviously, if such exclusive-sales strategy is adopted, it is optimal for the retailer to set $p_l = p_h = v_1$. This way, class-1 customers will purchase the product at v_1, since they have no incentive to wait for the clearance price. Hence, regardless of the inventory display format, the retailer's expected profit can be expressed as[5]:

$$\Pi^1(Q) \doteq -(c-s)\,Q + (v_1 - s) \cdot N(Q, \alpha_1 \lambda T) \tag{1}$$

where $N(Q,\Lambda) \doteq \sum_x \min(Q,x) \cdot P_x(\Lambda) = Q - \sum_{x=0}^{Q-1}(Q-x)\,P_x(\Lambda)$, and $P_x(\Lambda)$ is the Poisson probability function (corresponding to a mean parameter of Λ). The retailer's optimal expected profit and order quantity can be determined by solving the "newsvendor" problem: $\Pi^1 \doteq \max_{Q\geq 0}\left\{\Pi^1(Q)\right\}$. The second case (i.e., when $p_l \leq v_0 \leq p_h \leq v_1$) reflects a setting in which the retailer chooses to target both customer classes. This case, which is considerably harder to analyze, is treated extensively in the remainder of this paper.

Let us examine the strategic purchasing behavior in the second setting. Essentially, each arriving customer needs to compare the surpluses associated with two options: "buy now" and "wait." As $p_l \leq v_0 \leq p_h \leq v_1$, it is clear that all class-0 will always wait for the clearance price, regardless of the display format. However, for a class-1 customer, the decision is less trivial. If he purchases the product immediately at the premium price p_h, then he will obtain a surplus $(v_1 - p_h)$. If he waits for the clearance price p_l, then his expected surplus could ideally be $(v_1 - p_l)$; nonetheless,

[5] The term $N(Q,\Lambda)$ is equal to the expected value $E[\min(X,Q)]$, for a random variable $X \sim$Poisson(Λ), and hence has the interpretation of the number of units sold in this context.

this customer needs to consider the possibility that an item will not be available to him, due to a stockout. To capture this aspect, we need to introduce the customer's perception of the likelihood of getting the product at the post-season clearance price p_l if he decides to wait and returns to the store at the end of the season. We refer to this likelihood as the *"perceived fill rate"*[6]; correspondingly, we shall refer to the complimentary probability as the *"perceived sense of scarcity."* The customer hence needs to assess the expected surplus ("perceived fill rate") $\times (v_1 - p_l)$. By comparing these two surpluses, it is clear that a class-1 customer will purchase the product at p_h if the fill rate is less than $(v_1 - p_h) / (v_1 - p_l)$, and will wait for the clearance price if the fill rate is greater than $(v_1 - p_h) / (v_1 - p_l)$. Therefore, the optimal purchasing behavior for each class-1 customer hinges upon the *individually-assessed* fill rates. As we shall see, the assessment of the fill-rate by an individual customer is a challenging task, as it depends on many factors including the inventory information available under the given display format, the customer's arrival time, and the purchasing behavior of all other customers.

4 The "Display All" (DA) Format Model

Under the DA format, each arriving customer has perfect information about the actual inventory level. We first study the customers' subgame, and then proceed to the retailer's problem.

4.1 Strategic Purchasing Under the Display All Format

Consider a given order quantity Q and a pre-announced price path that satisfies $p_l \leq v_0 \leq p_h \leq v_1$. Clearly, we anticipate class-0 to always wait for the clearance price p_l, since $v_0 \leq p_h$. For class-1, consider a customer who arrives at time t and observes k units available for sale. Associated with the state (k, t), let $H(k, t)$ be the perceived fill rate as defined earlier. Hence, this customer will buy immediately at p_h if $H(k, t) \leq \frac{v_1 - p_h}{v_1 - p_l}$; otherwise, the customer will wait for the clearance price p_l. To facilitate the analysis of the fill rates $H(k, t)$, we define the following auxiliary function:

$$\hat{H}_q(\Lambda) \doteq \sum_{x=0}^{q-1} P_x(\Lambda) + \sum_{x=q}^{\infty} \frac{q}{x+1} P_x(\Lambda) \tag{2}$$

[6] The existing literature uses different expressions to denote the perceived fill rates. For example, Aviv and Pazgal (2007) use the term "allocation probability."

The function $\hat{H}_q(\Lambda)$ represents the probability that a given class-1 customer will obtain a unit out of q available, if they need to be distributed among him and X additional customers, where $X \sim$Poisson(Λ). In other words, $\hat{H}_q(\Lambda) = E_X[q/\max(q, X+1)]$. (We assume allocation with equal probability, in case of shortage). The following proposition demonstraes that the function $\hat{H}_q$ is useful in constructing a lower-bound on the perceived fill rate.

Proposition 1 *Consider any given arbitrary purchasing policy (including randomized strategies). The following bound applies to the perceived fill rate $H(k,t)$:*

$$H(k,t) \geq \sum_{x=0}^{k-1} P_x(\alpha_1 \lambda (T-t)) \cdot \hat{H}_{k-x}(\alpha_0 \lambda T + \alpha_1 \lambda t).$$

(All proofs appear in the Appendix on the Online Addendum accompanying this paper.) The lower bound in the proposition has an intuitive interpretation, as it corresponds to the "worst case" scenario in which all class-1 customers arriving prior to time t and all class-0 customers will "wait" and all class-1 customers arriving after t will "buy now". This scenario yields a lower bound on the fill rate because it would generate the highest number of customers who wait and the lowest number of remaining units at the end of the season. Observe from the above proposition that the lower bound is increasing in k (see the Appendix). Hence, there exists a sufficiently large level of inventory (say k') so that the lower bound is large enough to ensure that $H(k,t) > (v_1 - p_h)/(v_1 - p_l)$, for all $k \geq k'$. Consequently, as k is sufficiently large, it is optimal for a class-1 customer to wait regardless of the purchasing behavior of other class-1 customers.

However, as the level of inventory k becomes lower, $H(k,t)$ becomes lower and each class-1 customer needs to take into account the behavior of other class-1 customers. To examine this formally, we introduce a sequence of values $\{\tau^*(k)\}$, defined as follows. For all $k = 1, 2, \ldots$:

$$\tau^*(k) = \begin{cases} 0 & \text{if } \frac{v_1 - p_h}{v_1 - p_l} < \hat{H}_k(\lambda T) \\ \text{See Eq. (4) below} & \text{if } \hat{H}_k(\lambda T) \leq \frac{v_1 - p_h}{v_1 - p_l} \leq \hat{H}_k(\alpha_0 \lambda T) \\ T & \text{if } \frac{v_1 - p_h}{v_1 - p_l} > \hat{H}_k(\alpha_0 \lambda T) \end{cases} \tag{3}$$

In the second case of (3), the value $\tau^*(k)$ is the unique solution to the equation

$$v_1 - p_h = (v_1 - p_l)\hat{H}_k(\lambda(T - \alpha_1 \tau^*(k))) \tag{4}$$

(For a proof of uniqueness, see Proposition 2 below.) Equation (4) is reminiscent of the equation $v_1 - p_h = (v_1 - p_l) H(k,t)$, which holds for times t (if exist) at which a class-1 customer is indifferent

between a "buy now" and a "wait" decision, if k units are available for sale at time t. Indeed, the values $\tau^*(k)$ will play an important role in defining the indifference points (or thresholds) for optimal purchasing policies in equilibrium.

Proposition 2 *The sequence $\{\tau^*(k)\}$ defined in (3) is unique and non-increasing in k. Furthermore, the values $\tau^*(k)$ are increasing in λ, v_1 and p_l, and decreasing in p_h.*

We are now ready to present a key structural property of the optimal purchasing strategy for class-1 customers. To this end, let us consider the following threshold-type policy that is based on the sequence of values $\{\tau^*(k)\}$ as presented in (3)-(4). This threshold policy can be described as follows: for any class-1 customer who arrives at time t and observes k_t units available for sale, he should "buy now" if $t < \tau^*(k_t)$ and "wait" if $t > \tau^*(k_t)$. (If $t = \tau^*(k_t) \in (0, T)$, the policy can prescribe either action at random.)

Theorem 1 *Under the DA format, an "always-wait" strategy for class-0 customers and a threshold-type policy (based on threshold values $\{\tau^*(k)\}$) for class-1 customers form the unique Nash equilibrium in the sub-game among customers.*

4.2 The Retailer's Problem Under the Display All Format

Anticipating customers' purchasing behavior in equilibrium as established in Theorem 1, the retailer needs to identify an order quantity Q and a pair of optimal prices (p_h, p_l) that maximize his expected profit. We first consider the regime $p_l \leq v_0 \leq p_h \leq v_1$. Let us define $f(j,k)$ to be equal to the retailer's expected profit gained during the interval $(\tau^*(j), T]$, given that k units are in stock at time $\tau^*(j)$; for notational convenience, let $\tau^*(Q+1) \equiv 0$. Then, the functions f satisfy the following recursive scheme:

Proposition 3 *Given the customers' purchasing behavior in equilibrium, the functions f satisfy*[7] *the following equations, for all $1 \leq k \leq j \leq Q+1$:*

$$f(j,k) = \begin{cases} \sum_{x=0}^{k-1} [xp_h + f(k, k-x)] \times P_x(\alpha_1 \lambda_1 (\tau^*(k) - \tau^*(j))) & \text{if } k < j \\ \quad + [\sum_{x=k}^{\infty} P_x(\alpha_1 \lambda_1 (\tau^*(k) - \tau^*(j)))] \times kp_h & \\ kp_l - (p_l - s) \sum_{x=0}^{k-1} (k-x) \times P_x(\lambda(T - \alpha_1 \tau^*(k))) & \text{if } k = j \end{cases}$$

[7] Clearly, the value of the functions f depend on the price path (p_h, p_l). However, the price parameters are omitted for brevity of exposition.

Using the recursive algorithm outlined in Proposition 3, we can compute the retailer's expected profit for any order quantity Q and price path (p_h, p_l). Specifically, let us use the notation $\pi^{DA}(Q, p_h, p_l) \doteq f(Q+1, Q) - cQ$ to denote the retailer's expected profit. The following proposition characterizes a property associated with the optimal price path associated with $\pi^{DA}(Q, p_h, p_l)$.

Proposition 4 *For any given order Q, and for any price path $\{p_l \leq v_0 \leq p_h \leq v_1\}$, there exists an optimal price path that satisfies the conditions: $p_l = v_0$ and $p_h \in \Omega_Q^{DA}$, where*

$$\Omega_Q^{DA} \doteq \left[v_0 \cdot \hat{H}_Q(\alpha_0 \lambda T) + v_1 \cdot \left(1 - \hat{H}_Q(\alpha_0 \lambda T)\right), v_0 \cdot \hat{H}_Q(\lambda T) + v_1 \cdot \left(1 - \hat{H}_Q(\lambda T)\right)\right] \subseteq [v_0, v_1]$$

Under the given price regime, Proposition 4 states that it is optimal to prescribe the clearance price $p_l = v_0$. For the optimal premium price p_h, the proposition establishes the range Ω_Q^{DA}. The upper bound of Ω_Q^{DA} is based on the price above which class-1 customers always wait, whereas the lower bound is based on the premium price below which all class-1 customers buy immediately.

In order to identify the optimal strategy for the retailer, we conducted a hierarchical search procedure, as follows. We sequentially examined a plausible set of values of the initial inventory level Q. Given each level of Q, we applied a standard line-search procedure to determine the optimal premium price by solving the following problem

$$\pi^{DA}(Q) \doteq \max_{p_h} \{\pi^{DA}(Q, p_h) : p_h \in \Omega_Q^{DA}\}$$

We then searched for the optimal quantity Q by solving the problem: $\pi^{DA} \doteq \max_{Q \geq 0} \left\{\pi^{DA}(Q)\right\}$. However, instead of searching all possible values of Q, we reduced our search effort by using Proposition 4 to construct an upper bound on the function $\pi^{DA}(Q)$ as follows.

Corollary 1 *For all $Q \geq 1$,*

$$\pi_-^{DA}(Q) \leq \pi^{DA}(Q) \leq \pi_+^{DA}(Q), \tag{5}$$

where

$$\begin{aligned}
\pi_-^{DA}(Q) &\doteq -(c-s)Q + (v_0 - s) \cdot N(Q, \lambda T) + (v_1 - v_0) \cdot \left(1 - \hat{H}_Q(\alpha_0 \lambda T)\right) \cdot N(Q, \alpha_1 \lambda T) \\
\pi_+^{DA}(Q) &\doteq -(c-s)Q + (v_0 - s) \cdot N(Q, \lambda T) + (v_1 - v_0) \cdot \left(1 - \hat{H}_Q(\lambda T)\right) \cdot N(Q, \alpha_1 \lambda T)
\end{aligned}$$

The verification of the lower bound is straightforward, since it represents the expected profit under the price path that prescribes the lowest value in Ω_Q^{DA} for the premium price and $p_l = v_0$. The upper bound is constructed by assuming $p_l = v_0$ and by assuming that all class-1 customers purchase the product at the highest possible premium price in the range Ω_Q^{DA}. The bounds in (5) lead to the following result.

Proposition 5 *Let* $\Delta_\pi^{DA}(Q) \doteq \pi^{DA}(Q+1) - \pi^{DA}(Q)$. *Then,*

$$\Delta_\pi^{DA}(Q) \leq -(c-s) + (v_0 - s) \cdot \left(1 - \sum_{x=0}^{Q} P_x(\lambda T)\right) + (v_1 - v_0) \cdot \left(1 - \hat{H}_{Q+1}(\lambda T)\right) \cdot \alpha_1 \lambda T \quad (6)$$

Moreover, this bound on $\Delta_\pi^{DA}(Q)$ *is decreasing in* Q.

The value of this proposition is that it allows us to employ a step-by-step (i.e., $Q = 0, 1, \ldots$) evaluation of the right-hand-side of (6), until we identify the first (smallest) Q for which this value is negative; say, a level $\bar{Q}$ such that $\Delta_\pi^{DA}(\bar{Q}) < 0$. This means that adding one more unit to the stock will certainly result in a *reduction* of the expected profit. This analysis is reminiscent of the "newsvendor model". Nonetheless, here, we are not using the actual value $\Delta_\pi^{DA}(Q)$, and so $\bar{Q}$ is not necessarily optimal. However, it is obvious that any $Q > \bar{Q}$ should be avoided, in view of the last part of the proposition. Therefore, we can limit our search to the set $Q \in \{0, \ldots, \bar{Q}\}$.

Finally, in order to identify the overall optimal expected profit associated with the DA format denoted hereafter by Π^{DA}, we need to compare the retailer's optimal expected profit associated with the second setting (i.e., π^{DA}) with the retailer's optimal expected profit associated with the first setting (i.e., $\Pi^1 \doteq \max_{Q \geq 0} \{\Pi^1(Q)\}$). Thus, $\Pi^{DA} \doteq \max\{\pi^{DA}, \Pi^1\}$. In the remainder of the paper, we use the notation $(Q^{DA}, p_h^{DA}, p_l^{DA})$ to express an optimal choice that leads to the performance Π^{DA}.

5 The "Display One" (DO) Format Model

Under the "Display One" (DO) format, the retailer displays only a single unit on the sales floor, and keeps the rest in a "storeroom." Suppose that, upon a sale, the retailer immediately retrieves a new unit from the storeroom and places it on display[8]. Let us focus initially on a given order

[8] A situation like this is common in an e-tailing environment. For example, Lands' End Overstocks' website provides each arriving customer about product availability but not the actual inventory level.

quantity Q, and a price path in the regime $p_l \leq v_0 \leq p_h \leq v_1$. Clearly, class-0 always wait for the clearance price p_l, whereas class-1 need to consider their perceived fill rates. Note that a class-1 customer only knows his arrival time (t) and whether or not the product is still in stock. Assuming an in-stock state, we denote the perceived fill rate by $\tilde{H}(t)$. The lack of perfect information about inventory turns the estimation of the fill rate into a technically challenging task. To overcome this modeling challenge, let us discuss three possible approaches.

The first approach has been examined by Yin and Tang (2006). In this approach, it is assumed that all arriving customers have identical beliefs that the current inventory level is equal to k with probability θ_k, where θ_k is given *exogenously* in the sense that the belief is independent of the retailer's decisions (Q, p_h and p_l) and of other customers' purchasing behavior. In this case, it is easy to verify that $\tilde{H}(t) = \sum_k \theta_k H(k,t)$, where $H(k,t)$ is as defined for the DA format.

In the second approach, one assumes that the perceived fill rate is constant so that $\tilde{H}(t) = \tilde{H}$, where $\tilde{H}$ is determined *endogenously*. For any given value of $\tilde{H}$, all class-1 customers will purchase the product at p_h if $\tilde{H} \leq (v_1 - p_h)/(v_1 - p_l)$, and wait for the clearance price p_l otherwise. Given this customer purchasing behavior, $\tilde{H}$ needs to be the resulting fill rate exactly. Thus, $\tilde{H}$ constitutes a "self fulfilling prophecy" for the fill rate. Since all consumers act in accordance to this belief, $\tilde{H}$ turns out to indeed be the "correct" actual fill rate. It could be argued that the consumers' strategy and beliefs are "closed" under rational expectations. However, while this approach has an interesting economic interpretation and it allows for reasonable tractability, it neglects the fact that the perceived fill rate should be time-dependent.

Because neither the first nor the second approach satisfactorily addresses the time-varying beliefs of the actual inventory level or the fill rate, we take a different tack. In this third approach, we assume that the initial order quantity Q is a common knowledge to all customers.[9] But to capture the fact that the perceived fill rate is time dependent, we let $\tilde{H}_Q(t)$ be the fill rate assessed by a customer who arrives at time t. Clearly, this likelihood depends on the purchasing behavior of other customers and is treated rigorously in the next section.

[9] Several online retailers (especially liquidators such as PacificGeek.com) provide consumers with their initial level of inventory for a particular item but do not update it during the selling season.

5.1 Strategic Purchasing Under the Display One Format

Consider a class-1 customer who arrives at time t and observes a unit on display (i.e., the product is still available for sale). The perceived fill rate $\tilde{H}_Q(t)$ maintains the following properties.

Proposition 6 *For any given (possibly randomized) time-dependent purchasing policy followed by class-1 customers, the perceived fill rate $\tilde{H}_Q(t)$ is continuous and non-decreasing in $t \in [0, T]$. Moreover, $\tilde{H}_Q(t)$ is strictly increasing in t at (and only at) times when the purchasing policy prescribes a "buy now" action with a non-zero probability.*

The significance of this proposition is that the expected surplus associated with a "wait" action is non-decreasing as a function of the arrival time. Therefore, if there is a point in time t that has $\tilde{H}_Q(t) > \frac{v_1 - p_h}{v_1 - p_l}$, then it is optimal for all class-1 customers who arrive on or after t to wait for the clearance price p_l. This implication is stated formally in the following theorem.

Theorem 2 *For any purchasing policy to be sustained in equilibrium, it must possess the following properties: all class-0 customers must wait for the clearance price p_l; and all class-1 customers must follow a threshold policy. Specifically, all class-1 customers arriving prior to a threshold τ should purchase the product at p_h and all class-1 customers arriving after the threshold τ should wait for the clearance price p_l.*

Theorem 2 suggests that it is sufficient to focus on threshold-type policies that can be characterized by a single threshold τ. Hence, in order to explicitly capture the impact of the threshold τ on the perceived fill rate, we substitute $\tilde{H}_Q(t)$ by $L_Q(t|\tau)$. The following proposition provides a closed-form expression for the latter function.

Proposition 7 *The probability function $L_Q(t|\tau)$ is given by*

$$L_Q(t|\tau) = \frac{\sum_{x=0}^{Q-1} \hat{H}_{Q-x}(\lambda(T - \alpha_1 \tau)) \times P_x(\alpha_1 \lambda \tau)}{\sum_{i=0}^{Q-1} P_i(\alpha_1 \lambda \cdot \min\{t, \tau\})} \tag{7}$$

for all $t \in [0, T]$. The function $L_Q(t|\tau)$ satisfies the following properties:

(i) *$L_Q(t|\tau)$ is continuous both in t and τ.*

(ii) *$L_Q(t|\tau)$ is strictly increasing in t on $[0, \tau)$ and constant on $[\tau, T]$.*

Equation (7) represents $L_Q(t|\tau)$ as a mix of conditional probabilities, with each one considering a specific number of class-1 customers that buy the product at premium price (anytime during the season)[10]. Of course, all of these conditional probabilities assume that the threshold policy (with τ) is followed by all class-1 customers. Part (i) of Proposition 7 implies that the function

$$L_Q(\tau) \doteq L_Q(\tau|\tau)$$

is continuous. Therefore, there always exists an equilibrium (say τ_Q) in the customers' game. In particular, the following values of τ represent equilibrium strategies: (i) $\tau_Q = T$, if $L_Q(T) < (v_1 - p_h)/(v_1 - p_l)$; (ii) $\tau_Q = 0$, if $L_Q(0) > (v_1 - p_h)/(v_1 - p_l)$; and (iii) any solution τ_Q (if exists) that satisfies the equation:

$$L_Q(\tau_Q) = (v_1 - p_h)/(v_1 - p_l). \tag{8}$$

As follows from our theoretical analysis below, there could be multiple equilibria in the customers' game; an important issue we will discuss shortly. Part (ii) of the proposition is reminiscent of a similar result for the DA format case: if at any time t it is strictly better for a customer to "wait," then all customers arriving after time t should "wait" too. We now provide further characterization of the function $L_Q(\tau)$.

Theorem 3 *The function $L_Q(\tau)$ satisfies the following properties:*

(i) *$L_1(\tau)$ is strictly increasing in τ.*

(ii) *For all $Q \geq 2$, the function $L_Q(\tau)$ is unimodal (quasi-convex) in τ. Moreover, the function $L_Q(\tau)$ attains its unique minimum in the range $(0, T]$, and it is strictly decreasing (increasing) for all values of τ below (above) that minimum point.*

(iii) *For every $Q \geq 1$ and $\tau \in [0, T]$, $L_Q(\tau)$ is decreasing in λ.*

From the perspective of game theory, Theorem 3 is key to the characterization of the equilibrium in the customers' subgame. It implies that, when $Q = 1$, there is always a unique equilibrium in the

[10] Clearly, under a threshold policy τ, all class-0 customers and only those class-1 customers arriving after τ will wait for the clearance price. As such, the total number of customers waiting for the clearance price is equal to a Poisson random variable with rate $\alpha_0\lambda T + \alpha_1\lambda(T - \tau) = \lambda(T - \alpha_1\tau)$. Therefore, if $Q - x$ units are available for sale, the fill rate for a focal customer is given by $\hat{H}_{Q-x}(\lambda(T - \alpha_1\tau))$.

game, and that the threshold in equilibrium is increasing in the ratio $(v_1 - p_h) / (v_1 - p_l)$. When $Q > 1$, the unimodality implies that there could be as many as three equilibrium points in the game. Specifically, this happens if $\min_{\tau \in [0,T]} L_Q(\tau) < (v_1 - p_h) / (v_1 - p_l) < \min\{L_Q(0), L_Q(T)\}$, in which case the three equilibrium points are $\tau = 0$ as well as the only two solutions to the equation $L_Q(\tau) = (v_1 - p_h) / (v_1 - p_l)$. The theorem also implies that if $(v_1 - p_h) / (v_1 - p_l) > L_Q(0)$, the game has a unique equilibrium strategy that has a positive threshold (i.e., $0 < \tau \leq T$). Clearly, for $(v_1 - p_h) / (v_1 - p_l) < \min_{\tau \in [0,T]} L_Q(\tau)$, the unique equilibrium is $\tau_Q = 0$ (i.e., everyone waits). The significance of part (iii) of the theorem will become clear shortly.

Given any game, a prediction of players' behavior which is not a Nash equilibrium cannot be commonly believed. Hence, when we study a game that has only one equilibrium, it must be the only rational prediction of players' behavior. However, when a game possesses multiple Nash equilibria the assumption that a particular one will be played relies on the existence of a process or mechanism that directs all the players to expect the same outcome. Schelling (1960) argues that when looking at the multiplicity of equilibria one should look at anything that focuses the players' attention on a particular equilibrium. Specifically, when an equilibrium offers the highest possible payoffs to all players (it is Pareto dominant among the equilibria) it is natural to assume that the players will focus on it. Indeed, as the following proposition shows, the game we analyze possesses such an equilibrium.

Proposition 8 *Suppose that $Q \geq 2$ and that $(v_1 - p_h) / (v_1 - p_l) < L_Q(0)$. Then, among all possible existing equilibrium strategies, $\tau_Q = 0$ strictly dominates the others (if any) in terms of the expected surplus gained by class-1 customers.*

Thus, under the conditions of Proposition 8, the subgame played by class-1 consumers is a coordination game possessing a single Pareto dominant equilibrium strategy profile; namely, when all class-1 consumers do not buy upon arrival but wait for the end-of-season sale, their equilibrium payoff is the highest possible. Since there is an equilibrium whose outcome is strictly preferred by every player to any other equilibrium outcome we can assume that the consumers will focus on it[11].

[11] For Further discussion, see Section 1.2.4 in Fudenberg and Tirole (1991).

Specifically, we shall consider the equilibrium that is uniquely defined as follows:

$$\tau_Q^* \doteq \begin{cases} 0 & \text{if } \frac{v_1-p_h}{v_1-p_l} \leq L_Q(0) \\ \text{The unique positive solution to Eq. (8)} & \text{if } L_Q(0) < \frac{v_1-p_h}{v_1-p_l} \leq \max\{L_Q(0), L_Q(T)\} \\ T & \text{if } \frac{v_1-p_h}{v_1-p_l} > \max\{L_Q(0), L_Q(T)\} \end{cases} \tag{9}$$

One of the immediate results that follow is similar to Proposition 2 for the DA case.

Proposition 9 *Under the DO format, the threshold τ_Q^* as defined in (9) is increasing in λ, v_1 and p_l, and decreasing in p_h.*

We next present a relationship between the threshold τ_Q^*, associated with the DO format, and the sequence of values $\{\tau^*(k) : k = 1, \ldots, Q\}$ associated with the DA format.

Proposition 10 *For any given level of inventory $Q \geq 1$, and a pair of prices (p_h, p_l) that satisfy the condition $p_l \leq v_0 \leq p_h \leq v_1$, we have*

$$\tau^*(Q) \leq \tau_Q^* \leq \tau^*(1)$$

Proposition 10 also suggests that by moving from a DA to DO while keeping the price path unchanged, the *volatility* of the retailer's profit under the DO model will be lower than that of the DA model. This is because the DO model appears to be less sensitive to the arrival time of the first class-1 customer than the DA model. To see this, note that if no arrival occurs prior to $\tau^*(Q)$, then no purchase at premium price will be made during the entire horizon under the DA model. However, purchases could still occur under the DO model until $\tau_Q^* \geq \tau^*(Q)$. Nonetheless, if sales are made early enough, and so the inventory declines, purchases under the DA model may still continue after τ_Q^*: the time at which class-1 customers begin to wait under the DO model.

5.2 The Retailer's Problem Under the Display One Format

Anticipating the threshold policy τ_Q^* adopted by all class-1 customers in equilibrium, it is easy to check that the retailer's expected profit can be expressed as:

$$\pi^{DO}(Q, p_h, p_l) = -(c-s)Q + (p_l - s) \cdot N(Q, \lambda T) + (p_h - p_l) \cdot N\left(Q, \alpha_1 \lambda \tau_Q^*(p_h, p_l)\right)$$

(Note that we explicitly describe $\tau_Q^*(p_h, p_l)$ to emphasize the dependency of the threshold on the price path) The retailer's expected profit $\pi^{DO}(Q, p_h, p_l)$ can be interpreted as follows. The first

term represents the loss incurred if no unit was sold during the season. The second term corresponds to the margin of $(p_l - s)$ for each unit that was sold during the season. The last term adds an additional margin of $(p_h - p_l)$ for each unit that was sold at the premium price p_h instead of the clearance price p_l.

Given the retailer's profit function $\pi^{DO}(Q, p_h, p_l)$, the following proposition characterizes a property associated with the optimal price path.

Proposition 11 *Consider a given finite level of initial inventory $Q > 0$, and suppose that the price path is limited to the range $\{p_l \leq v_0 \leq p_h \leq v_1\}$. Then, it is optimal to set $p_l = v_0$. Given $p_l = v_0$, the value of p_h can be practically limited to the range*

$$\Omega_{Q,\varepsilon}^{DO} \doteq \begin{cases} [L_Q(0)\, v_0 + (1 - L_Q(0))\, v_1 - \varepsilon, L_Q(0)\, v_0 + (1 - L_Q(0))\, v_1) & \text{if } L_Q(0) \geq L_Q(T) \\ [L_Q(T)\, v_0 + (1 - L_Q(T))\, v_1, L_Q(0)\, v_0 + (1 - L_Q(0))\, v_1) & \text{otherwise} \end{cases}$$

for any $\varepsilon > 0$. Specifically, while it is possible that no optimal value of p_h exists in this range, there always exists a price $p_h \in \Omega_{Q,\varepsilon}^{DO}$ that yields a performance that is arbitrarily close to optimal.

Similarly to Proposition 4 (the DA model), under the given price regime, it is optimal for the retailer to set the clearance price $p_l = v_0$. If the discount price was set lower than v_0, then by raising it to v_0, the retailer would gain in two ways: reducing the motivation of class-1 customers to wait, and obtaining higher revenues from those who wait. While, conceptually, this line of proof is not hard to follow, it relies on the fact that in equilibrium, the threshold value τ_Q^*, associated with the DO format, is non-decreasing in p_l. In other words, it relies heavily on the theoretical results presented in the previous subsection.

How do we choose a good premium price? Essentially, Proposition 11 shows that given the optimal choice of $p_l = v_0$, we can limit our attention to premium prices that satisfy the condition $L_Q(0) < (v_1 - p_h) / (v_1 - p_l) \leq \max\{L_Q(0), L_Q(T)\}$, virtually without loss of optimality. Note that this is the middle condition of (9), for which the threshold value τ_Q^* can be interpreted as the unique *indifference* point for class-1 customers; i.e., the single time in the season for which a class-1 customer is indifferent between a "buy now" and a "wait" decision. Once again, the bound provides us with computational convenience. For every value of $p_h \in \Omega_{Q,\varepsilon}^{DO}$, we can simply calculate the ratio $(v_1 - p_h) / (v_1 - v_0)$, and identify τ_Q^* via (8).

As in the DA case, we can determine the optimal premium price p_h by solving the following problem via a simple line-search.

$$\pi^{DO}(Q) \doteq \max\left\{\pi^{DO}(Q, p_h, p_l = v_0) : p_h \in \Omega^{DO}_{Q,\varepsilon}\right\}$$

for some ε value (we used $\varepsilon = 0.001v_1$). Using the results of Proposition 11, and following similar justifications as those explaining Corollary 1, we can use the boundary points of the range $\Omega^{DO}_{Q,\varepsilon}$ in order to obtain upper and lower bounds on $\pi^{DO}(Q)$. Specifically[12], we have

$$\pi^{DO}_{-}(Q) \leq \pi^{DO}(Q) < \pi^{DO}_{+}(Q) \tag{10}$$

where

$$\begin{aligned}\pi^{DO}_{-}(Q) \quad \doteq \quad & -(c-s)\,Q + (v_0 - s)\cdot N(Q, \lambda T) \\ & + (v_1 - v_0)\cdot\left(1 - \max\left\{L_Q(0), L_Q(T)\right\}\right)\cdot N(Q, \alpha_1\lambda T)\end{aligned}$$

and $\pi^{DO}_{+}(Q) = \pi^{DA}_{+}(Q)$. Since the upper bound in (10) is the same as that for the DA model, we can use the results of Proposition 5 to identify a maximal value of Q that needs to be considered for the purpose of solving the problem $\pi^{DO} \doteq \max_{Q\geq 0}\left\{\pi^{DO}(Q)\right\}$. Interestingly, we establish the following result.

Corollary 2 *For all $Q \geq 1$, $\pi^{DO}_{-}(Q) \geq \pi^{DA}_{-}(Q)$.*

This corollary is justified by the fact that $L_Q(T)$ is expressed (see definition) as a weighted average of the functions $\hat{H}_x(\alpha_0\lambda T)$ where $x = 1, \ldots, Q$, and that these functions are increasing in x. Thus, $\hat{H}_Q(\alpha_0\lambda T) \geq L_Q(T)$. Additionally, since $\hat{H}_Q(\cdot)$ is a decreasing function of its argument, it is straightforward to argue that $\hat{H}_Q(\alpha_0\lambda T) \geq H_Q(\lambda T) = L_Q(0)$. Obviously, since the values in the corollary are just lower bounds, the latter inequality does not imply that the optimal expected profits $\pi^{DO}(Q)$ and $\pi^{DA}(Q)$ maintain the same relationship. However, the inequality is at least consistent with our belief that the DO format can serve as a mechanism to place more pressure on class-1 customers (in terms of perceived risk of stockout), so they feel more inclined to purchase at premium prices.

[12] The lower bound in (5) may need to be corrected to $\pi^{DO}_{-}(Q) - \tilde{\varepsilon}$, depending on the value of ε that defines the set $\Omega^{DO}_{Q,\varepsilon}$. We avoid this non-crucial detail for the sake of brevity of exposition.

Finally, in order to identify the overall optimal expected profit associated with the DO format denoted by Π^{DO}, we need to compare the retailer's optimal expected profit associated with the second setting (i.e., π^{DO}) with the retailer's optimal expected profit associated with the first setting (i.e., $\Pi^1 \doteq \max_{Q\geq 0}\left\{\Pi^1(Q)\right\}$). Thus, $\Pi^{DO} \doteq \max\left\{\pi^{DO}, \Pi^1\right\}$. In the remainder of the paper, we use the notation $\left(Q^{DO}, p_h^{DO}, p_l^{DO}\right)$ to express an optimal choice that leads to the performance Π^{DO}.

6 Benchmark Models

In this section, we develop two benchmark models. In the first model, all customers are myopic (i.e., non-strategic). In the second model, the price path is constant so that $p_h = p_l$. First, consider the case when customers are myopic. In this case, regardless of the display format, a customer will purchase the product immediately upon arrival if his valuation is above the premium price p_h; otherwise, he will wait for the clearance price p_l. In the myopic model, it is optimal for the retailer to adopt the price path $\{p_h = v_1, p_l = v_0\}$. Additionally, it is easy to verify that the retailer's expected profit for a given initial order quantity Q is given by

$$\Pi^M(Q) \doteq -(c-s)\,Q + (v_0 - s)\cdot N(Q, \lambda T) + (v_1 - v_0)\cdot N(Q, \alpha_1 \lambda T) \tag{11}$$

The retailer can determine the optimal order quantity Q^M by solving the following problem: $\Pi^M \doteq \max_Q\left\{\Pi^M(Q)\right\}$. The solution to this "newsvendor" problem is given by the smallest value of Q that has $\Pi^M(Q+1) - \Pi^M(Q) \leq 0$, where

$$\Pi^M(Q+1) - \Pi^M(Q) = (v_1 - c) - (v_0 - s)\cdot \sum_{x=0}^{Q} P_x(\lambda T) - (v_1 - v_0)\cdot \sum_{x=0}^{Q} P_x(\alpha_1 \lambda T)$$

Notice that the retailer's expected profit Π^M is an *upper* bound on the optimal expected profit for the case of strategic customers under the DA and DO formats. This can be easily verified via the inequality

$$\Pi^M(Q) - \pi^{DA}(Q) \geq \Pi^M(Q) - \pi_+^{DA}(Q) = (v_1 - v_0)\cdot \hat{H}_Q(\lambda T)\cdot N(Q, \alpha_1 \lambda T) \geq 0,$$

and the fact that $\Pi^M(Q) \geq \Pi^1(Q)$. Hence, $\Pi^M(Q) \geq \max\left\{\pi^{DA}(Q), \Pi^1(Q)\right\} = \Pi^{DA}(Q)$. By maximizing the profit functions with respect to the initial order quantity, we have: $\Pi^M \geq \Pi^{DA}$. The exact same argument can be made for the DO model.

In our second benchmark model, the retailer charges a fixed-price throughout the entire season so that $p_h = p_l$. When the selling price is fixed, there is no incentive for customers to wait; hence, it is optimal for the retailer to set either $p_h = p_l = v_1$ or $p_h = p_l = v_0$. Notice that the retailer's expected profit is equal to $\Pi^1(Q)$ in the former case and equal to $-(c-s)Q + (v_0 - s) \cdot N(Q, \lambda T)$ in the latter case. Hence, the retailer's optimal expected profit is given by $\Pi^F \doteq \max_Q \left\{\Pi^F(Q)\right\}$, where $\Pi^F(Q) \doteq \max\left\{-(c-s)Q + (v_0 - s) \cdot N(Q, \lambda T), \Pi^1(Q)\right\}$. Because the fixed-price is a feasible price path in the DA and DO models, it is easy to show that

Proposition 12 *For all $Q \geq 1$,*

$$\Pi^F(Q) \leq \left\{\Pi^{DA}(Q), \Pi^{DO}(Q)\right\} \leq \Pi^M(Q). \tag{12}$$

Also, $\Pi^F \leq \left\{\Pi^{DA}, \Pi^{DO}\right\} \leq \Pi^M$.

7 Numerical Studies

We now present the details of a numerical study we conducted for comparing the retailer's optimal price path, initial order quantity, and expected profit under the two display formats.

7.1 The Parameter Space and Performance Metrics

Recall that our models are characterized by the parameters $\{\alpha_0, \alpha_1, v_0, v_1, \lambda, T, c, s\}$. Without loss of generality, we set $v_1 = 1$ ("changs of currency"), and $T = 1$ (time is now expressed as a portion of the sales horizon, and consequently λ is the average traffic during the entire horizon). As to the value of s, note that a simple shift-transformation (of the type $x := x - s$) of all cost, valuation and price parameters, transforms the problem into one with $s = 0$. Finally, in order to study the impact of costs on performance, we considered the scaled value $\bar{c} \doteq c/v_0$ instead of c. This transformation lends itself to a clearer interpretation of the numerical results. In summary, our numerical analysis focuses on the four parameters: $\{v_0, \alpha_0, \lambda, \bar{c}\}$. The core of our numerical experiments is spanned over $9 \times 9 \times 6 \times 5 = 2,430$ combinations of the following parameter values:

$$\begin{aligned} v_0 &\in \{0.1, 0.2, \ldots, 0.9\}, \alpha_0 \in \{0.1, 0.2, \ldots, 0.9\} \\ \lambda &\in \{1, 2, 5, 10, 15, 20\}, \bar{c} \in \{0.1, 0.3, 0.5, 0.7, 0.9\}. \end{aligned}$$

Below, we use the term "scenario" to denote each instance of parameter combination.

We now present some of the performance metrics used in our analyses. First, in order to examine the value of markdown pricing relative to fixed-price policies, we can use the benchmark measure Π^F. For instance, one can define the theoretical value of markdown pricing as the relative increase in expected profits if the seller moves from a fixed-price policy to a markdown policy. For the DA and DO model, this value can be measured by

$$\eta^{DA} \doteq \Pi^{DA}/\Pi^F - 1, \text{ and } \eta^{DO} \doteq \Pi^{DO}/\Pi^F - 1. \tag{13}$$

Hereafter, we will refer to η^{DA} and η^{DO} as *the benefits of markdown pricing.* Another measure of interest is the value of markdown pricing when customers are myopic, which can be defined as

$$\eta^M \doteq \Pi^M/\Pi^F - 1. \tag{14}$$

We shall refer to η^M as *the ideal benefits of markdown pricing.* Obviously, in view of (12), we have $0 \leq \left\{\eta^{DA}, \eta^{DO}\right\} \leq \eta^M$. Finally, in order to compare the retailer's profits under the two display formats, we propose the following measure, which we refer to as the *marginal benefit of DO*:

$$\delta \doteq \Pi^{DO}/\Pi^{DA} - 1.$$

When we report on the averages of the aforementioned measures, we use the notations $\bar{\eta}^{DA}$, $\bar{\eta}^{DO}$, $\bar{\eta}^M$, and $\bar{\delta}$. Additionally, when focusing on cases in which the inventory is exogenously given, the above measures should be interpreted accordingly (e.g., $\delta = \delta(Q) = \Pi^{DO}(Q)/\Pi^{DA}(Q) - 1$).

7.2 The Case of Fixed Inventory: A Revenue Maximization Problem

We initially considered settings in which the level of inventory (Q) is exogenously given. Since the procurement cost $cQ = \bar{c}v_0Q$ can be considered as sunk, the retailer's objective is to maximize expected *revenues.* In particular, we set $\bar{c} = 0$, which allows us to interpret the expected profit measures as expected revenues. Since $\bar{c}$ is kept fixed, we are left with the 486 combinations of the parameters $\{v_0, \alpha_0, \lambda\}$. For each of these parameter combinations, we considered five values of $Q \in \{1, 2, 5, 10, 15\}$, hence having $2,430$ scenarios in total. For each of these scenarios we assessed various performance metrics as discussed above.

Our first step was to explore the impact of the display format *alone*, keeping the price path fixed. Specifically, we considered the gap $\Pi^{DO}\left(Q, p_h^{DA}, p_l^{DA}\right)/\Pi^{DA}(Q) - 1$, which measures the gain in expected revenue when the retailer moves from DA to DO, but uses the optimal price path for the DA format. Given our speculation that the DO format increases the perceived level of shortage risk, we expect this gap to be non-negative. Indeed, confirming this intuition, we found that this gap was non-negative in all scenarios. However, we found that the average value of this gap is very small (approximately 0.05%), and its maximal level is 1.7% only. Next, we measured the marginal benefits of DO; namely, the potential benefit of a change in display format when the optimal price path is used in *each* format. Not surprisingly, we found this gap to be non-negative across all instances, in view of the fact that $\Pi^{DO}(Q) \geq \Pi^{DO}\left(Q, p_h^{DA}, p_l^{DA}\right)$. On average, the value of δ is 0.25%. More than 95% of the scenarios exhibited $\delta < 2\%$, and just about 2% exhibited $\delta > 3\%$. The maximal value of δ observed is 6.90%.

The above results suggest that when moving from DA to DO, a price modification (i.e., selecting the optimal price for the DO case) typically plays a larger role than merely changing the display format. As we expected, the optimal regular price (p_h) under the DO format was higher or equal to that under the DA format across all instances. We attribute this, again, to the retailer's ability to charge more due to the increased sense of scarcity that the DO format creates. Our results also demonstrate that the relative marginal benefits of DO, even at their highest levels, are not dramatic in magnitude. Nonetheless, often, even a small increase in revenues could have a high impact on expected profits; see, e.g., Marn and Rosiello (1992) and Marn et al. (2003).

Compared to our *fixed-price* benchmark case, the DA and DO models provide average benefits of $\bar{\eta}^{DA} = 1.37\%$ and $\bar{\eta}^{DO} = 1.64\%$, respectively. The maximal values of these benefits are both 46.5%. Furthermore, when looking at the ideal benefits of markdown pricing, we observe an average of $\bar{\eta}^{M} = 13.7\%$ and a maximum of 90.0%. The striking difference between $\bar{\eta}^{M}$ and $\bar{\eta}^{DO} - \bar{\eta}^{DA}$ leads us to conclude that a move from a DA to a DO format is very far from eliminating the adverse impact of strategic behavior on the retailer's ability to conduct effective price segmentation.

We studied the impact of the various parameters on the marginal benefits of DO. Table 1 below shows the distribution of the marginal benefit of the DO format across different values of v_0 and α_0. Note that when α_0 or v_0 are relatively small, the DO format performs the same as

DA, since it is optimal for the retailer to sell exclusively to class-1 customers (more specifically, $\Pi^{DA}(Q) = \Pi^{DO}(Q) = \Pi^{1}(Q)$). Obviously, when either α_0 or v_0 approaches 1, there is no potential

	α_0									
v_0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	**All**
0.1										**0.000%**
0.2									0.996%	**0.111%**
0.3								1.424%	0.676%	**0.233%**
0.4							1.021%	1.251%	0.469%	**0.305%**
0.5						0.790%	1.243%	0.973%	0.328%	**0.370%**
0.6					0.507%	0.908%	0.999%	0.714%	0.240%	**0.374%**
0.7				0.371%	0.565%	0.815%	0.759%	0.509%	0.156%	**0.353%**
0.8			0.247%	0.383%	0.556%	0.580%	0.497%	0.309%	0.093%	**0.296%**
0.9		0.118%	0.213%	0.267%	0.289%	0.300%	0.245%	0.142%	0.040%	**0.179%**
All	**0.000%**	**0.013%**	**0.051%**	**0.113%**	**0.213%**	**0.377%**	**0.529%**	**0.591%**	**0.333%**	**0.247%**

Table 1: The average relative marginal benefits of DO, δ, for different combinations of parameters α_0 and v_0. For convenience, cells with zero percentage points, are presented empty.

for segmentation, and consequently, DO performs the same as DA (in fact, the same as fixed-price policies). This table also indicates that the marginal benefits of DO tend to be higher when there is a sufficiently large portion of class-0 customers in the market so as to create a shortage risk for class-1 customers. This result supports our hypothesis that the DO format serves as a mechanism for increasing the sense of scarcity among class-1 customers.

Table 2 below shows the distribution of the relative marginal benefit of the DO format across different values of Q and λ. First, note that when $Q = 1$, the two display formats are identical,

	λ						
Q	1	2	5	10	15	20	**All**
1	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	**0.000%**
2	0.311%	0.737%	0.747%	0.273%	0.111%	0.052%	**0.372%**
5	0.006%	0.047%	0.675%	0.755%	0.407%	0.218%	**0.351%**
10	0.000%	0.000%	0.022%	0.505%	0.795%	0.571%	**0.316%**
15	0.000%	0.000%	0.000%	0.052%	0.422%	0.702%	**0.196%**
All	**0.063%**	**0.157%**	**0.289%**	**0.317%**	**0.347%**	**0.309%**	**0.247%**

Table 2: The average relative marginal benefits of DO, δ, for different combinations of Q and λ.

since knowing the status of inventory (in-stock or sold-out) automatically reveals the exact amount of inventory on hand (1 or 0, respectively). For $Q \geq 2$, the table suggests that the DO format offers noticeable benefits only when the level of λ lies within a medium range (relative to Q). We explain this phenomenon, as follows. As the store traffic λ increases significantly, the thresholds

values $\{\tau^*(k)\}$ and τ_Q^* increase towards T under both display formats; see Propositions 2 and 9. Hence, we expect a larger portion of class-1 customers to buy at premium price (instead of wait) under both display formats. Consequently, the marginal benefit of DO would vanish. If the traffic intensity λ is relatively low, then class-1 customers' perceived fill rates are high. As such, more class-1 customers will wait for the clearance price p_l, regardless of the display format.

7.3 The Case of Expected Profit Maximization

We now turn our attention to the case in which the retailer selects both the initial order quantity and price path so as to optimize his expected *profit*. In contrast to the previous section, the retailer can now select the initial order quantity Q. In addition to the typical role of setting capacity to meet demand, the retailer can use the quantity decision as a lever to influence customers' perception of product scarcity. In our numerical analysis, we considered all $2,430$ scenarios spanned by the set of combinations of the four parameter $\{v_0, \alpha_0, \lambda, \bar{c}\}$ as described in §7.1. For each scenario, we calculated the optimal profits under both display formats; namely, Π^{DA} and Π^{DO}. Additionally, we calculated the maximal profit that could be obtained under DO if the quantity Q^{DA}, optimal for the DA case, is ordered by the retailer; i.e., $\Pi^{DO}\left(Q^{DA}\right)$. Across all scenarios, we found that $\Pi^{DA} \doteq \Pi^{DA}\left(Q^{DA}\right) \leq \Pi^{DO}\left(Q^{DA}\right) \leq \Pi^{DO}$(the first inequality is consistent with our observations in §7.2, whereas the second inequality holds by virtue of optimality). Table 3 below provides summary statistics of the marginal benefits of DO (see last row), as well as its breakdown into the benefits that can be gained *without* modifying the level of inventory from that which is optimal under DA (first row of the table), as well as the *extra* benefits that can be gained if the retailer could select the optimal level of inventory under DO (second row of the table). Interestingly, we notice that,

	mean	median	95th percentile	99th percentile
$\Pi^{DO}\left(Q^{DA}\right)/\Pi^{DA} - 1$	0.80%	0.00%	5.06%	11.78%
$\left[\Pi^{DO} - \Pi^{DO}\left(Q^{DA}\right)\right]/\Pi^{DA}$	0.08%	0.00%	0.16%	1.69%
$\delta = \Pi^{DO}/\Pi^{DA} - 1$	0.88%	0.00%	5.64%	12.62%

Table 3: Summary statistics on the benefits of DO.

generally, it is not important for a retailer to modify the level of inventory if he decides to move from a DA to a DO format (no more than 0.16% benefits for 95% of scenarios). A change in the display format, along with an appropriate price modification, is typically sufficient.

Similarly to the previous section, we found that the marginal benefit of DO is the highest for medium values of v_0 and α_0 (see explanation, ibid.). For instance, consider Table 4 below, which shows the marginal benefits of DO for all 30 scenarios with $v_0 = 0.6$ and $\alpha_0 = 0.7$. This particular combination of v_0 and α_0 yields the highest average benefits, measuring 4.14%. Notice that in some

		Relative cost parameter (c-bar)					
		0.1	0.3	0.5	0.7	0.9	All
λ	1	**1.18%** (1.18%, 0.00%) 2 → 2	**0.00%** (0.00%, 0.00%) 1 → 1	**0.00%** (0.00%, 0.00%) 1 → 1	**0.00%** (0.00%, 0.00%) 0 → 0	**0.00%** (0.00%, 0.00%) 0 → 0	**0.24%** (0.24%, 0.00%)
	2	**0.54%** (0.54%, 0.00%) 4 → 4	**4.15%** (4.15%, 0.00%) 2 → 2	**7.32%** (7.32%, 0.00%) 2 → 2	**0.00%** (0.00%, 0.00%) 1 → 1	**0.00%** (0.00%, 0.00%) 1 → 1	**2.40%** (2.40%, 0.00%)
	5	**0.37%** (0.37%, 0.00%) 8 → 8	**1.94%** (1.94%, 0.00%) 6 → 6	**6.64%** (6.64%, 0.00%) 4 → 4	**12.91%** (12.91%, 0.00%) 3 → 3	**20.06%** (20.06%, 0.00%) 2 → 2	**8.38%** (8.38%, 0.00%)
	10	**0.28%** (0.28%, 0.00%) 14 → 14	**1.54%** (1.54%, 0.00%) 11 → 11	**4.19%** (4.19%, 0.00%) 9 → 9	**9.95%** (9.95%, 0.00%) 7 → 7	**17.09%** (17.09%, 0.00%) 4 → 4	**6.61%** (6.61%, 0.00%)
	15	**0.22%** (0.22%, 0.00%) 20 → 19	**1.32%** (0.98%, 0.34%) 17 → 16	**3.23%** (3.23%, 0.00%) 14 → 14	**8.39%** (8.39%, 0.00%) 11 → 11	**8.71%** (0.00%, 8.71%) 4* → 6	**4.37%** (2.56%, 1.81%)
	20	**0.23%** (0.17%, 0.06%) 26 → 25	**1.04%** (0.92%, 0.12%) 22 → 21	**2.75%** (2.69%, 0.06%) 19 → 18	**7.43%** (6.93%, 0.51%) 16 → 15	**2.67%** (0.00%, 2.67%) 6* → 8	**2.83%** (2.14%, 0.68%)
	All	**0.47%** (0.46%, 0.01%)	**1.67%** (1.59%, 0.08%)	**4.02%** (4.01%, 0.01%)	**6.45%** (6.36%, 0.08%)	**8.09%** (6.19%, 1.90%)	**4.14%** (3.72%, 0.42%)

Table 4: The marginal benefits of DO across all 30 individual scenarios with $v_0 = 0.6$ and $\alpha_0 = 0.7$. In each cell, the top value (boldface) represents the marginal benefit, $\Pi^{DO}/\Pi^{DA} - 1$. The breakdown of this benefit into $\Pi^{DO}(Q^{DA})/\Pi^{DA} - 1$ and $[\Pi^{DO} - \Pi^{DO}(Q^{DA})]/\Pi^{DA}$, respectively, is provided within brackets, The bottom numbers in each cell represent Q^{DA} and Q^{DO}, respectively. A '*' sign next to Q^{DA} indicates a situation in which it is beneficial to the retailer to sell exclusively to class-1 customers.

of these scenarios, the increase in profit gained by a move from an optimal DA to an optimal DO is impressive, reaching up to 20%. The table also indicates an intricate relationship between the marginal benefits of DO and the cost parameter $\bar{c}$. To appreciate the impact of this parameter, it is useful to consider the store traffic parameter, λ, simultaneously. Suppose, first, that λ is small; in this case, an increase in the cost of the product may lead the retailer not to sell it at all, or to just purchase a single unit (illustrated in the first two rows of table, as one moves from $\bar{c} = 0.1$ to $\bar{c} = 0.9$). Obviously, in these cases, there is no difference between the DA and DO formats. When the store traffic increases, the optimal level of inventory (Q^{DA} or Q^{DO}) may be at a medium level, for which the benefits of DO tend to be high (see previous section). Moreover, as $\bar{c}$ increases, the retailer becomes more heavily reliant on revenues gained by selling at premium

prices. Consequently, the potential of DO in maximizing profits increases. Such pattern can be seen in the middle rows of Table 4. When λ is high, and $\bar{c}$ increases towards v_0, it becomes less valuable to sell to class-0 customers, and hence more attractive to sell exclusively to class-1. Recall that the value of exclusive sales is in completely eliminating strategic behavior. An exclusive sales case is illustrated in the table, for $\lambda = 15$ and $\bar{c} = 0.9$. It happens, it this specific case, that a move to DO enables to retailer to increase the inventory from $Q^{DA} = 4$ to $Q^{DO} = 6$, and gain an increase of 8.71% in profit. Nonetheless, as $\bar{c}$ continues to approach 1, we anticipate the exclusive sales option to be optimal under both display formats, hence yielding no advantage to DO.

Finally, we study the nature of the *simultaneous* change in inventory and premium price, as one moves from DA to DO. We focus on the set of scenarios in which a move from DA to DO calls for a change in inventory (i.e., $Q^{DO} \neq Q^{DA}$). In particular, we limited our attention to scenarios in which the marginal benefits of DO are sufficiently significant (more than 2%), and when a change in inventory is sufficiently valuable ($\left[\Pi^{DO} - \Pi^{DO}\left(Q^{DA}\right)\right]/\Pi^{DA} > 1\%$). Figure 1 below depicts the distribution of the gaps $p_h^{DO} - p_h^{DA}$ and $Q^{DO} - Q^{DA}$ among the latter set of instances. Recall

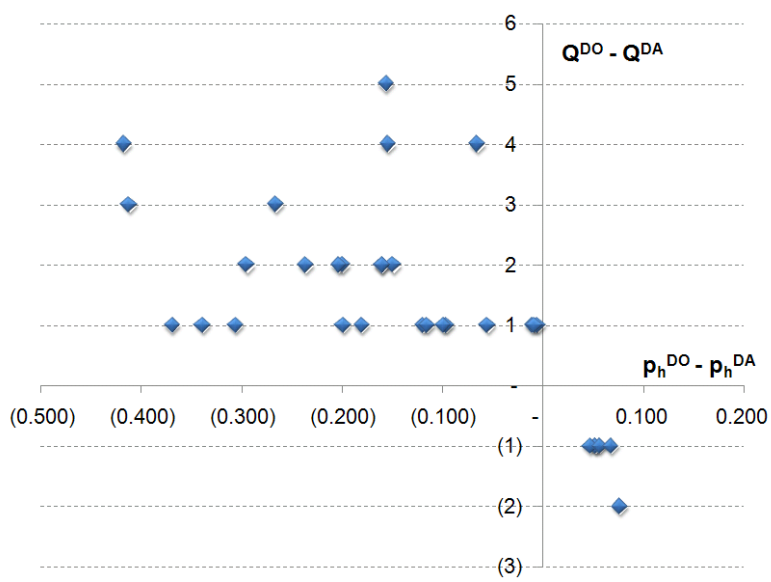

Figure 1: The joint distribution of premium price and inventory changes, when moving from a DA to a DO format. This figure includes all scenarios for which $\Pi^{DO}/\Pi^{DA} - 1 > 2\%$ and $\left[\Pi^{DO} - \Pi^{DO}\left(Q^{DA}\right)\right]/\Pi^{DA} > 1\%$.

that in the previous section we speculated that the customers' perception of product scarcity is

higher under DO than it is under DA. Consequently, we expect that when moving from DA to DO, a retailer may be able to increase his initial order quantity, but still maintain a sense of scarcity among class-1 customers. Alternatively, the retailer could charge a higher premium price, without driving class-1 customers to wait, to the same extent they would under the DA format. It is conceivable that the retailer could also simultaneously increase the quantity and the premium price. The data in Figure 1 confirms our intuition that a retailer should not simultaneously decrease the inventory and premium price. More interestingly, this figure shows that all instances reside in the upper-left or lower-right quarters of the chart. While this outcome is not necessarily guaranteed to hold in general, it is reasonable to expect situations in which one of the following two actions takes place: either the premium price is increased along with a *reduction* in inventory, or the opposite – inventory is increased along with premium price *reduction*.

8 Model Extensions, Discussion, and Ideas for Future Research

There are several interesting variants of our models that require relatively simple model extensions. For brevity of exposition, the reader is referred to the online addendum for technical details and managerial discussions. The first extension relaxes the assumption that, for each customer who decides to wait, he will return to the store when the clearance price, p_l, is posted. By considering uncertain customer returns, our model captures two countervailing forces as follows. When a customer is uncertain about his return, he is more inclined to purchase the product at the premium price. However, by taking into account that other customers are unsure about returning to the store, this customer is more inclined to wait for the post-season clearance sale. Our second extension deals with the case when returning to the store are costly to the customers. We establish that the purchasing threshold levels increase. In particular, class-1 customers are more inclined to buy at the premium price rather than wait. Interestingly, an increase in the cost of return influences the value of the DO format in two conflicting ways: As more customers buy at the premium price, the sense of scarcity generated by the DO format could possibly drive more customers to purchase at the premium price. However, as the costs of return increase, class-1 customers are pressed to buy at premium price anyway, regardless of the display format. In our third extension, we relax the assumption that the discount time takes place at the end of the season by allowing the retailer to

set the time of the discount at any point in time on or before the end of the season. As before, customers that arrive prior to the time of discount need to weigh the current surplus (gained by an immediate purchase) against the expected surplus that can be gained if they wait until the discount is offered. Our findings support our hypothesis that as the potential for segmentation increases (i.e., the discount is offered later in the season), the marginal value of DO increases due to the sense of scarcity it creates. In our fourth extension, we generalize our model by considering the case when customers receive non-negative (instead of zero) surpluses if they decide to wait and the items were sold out at the end of the season. For example, a customer that arrives to a jewelry store and is interested in a given luxury item, may end up buying an alternative item (and gain a positive utility) on the day of clearance sales, if the store runs out of stock. We argue[13] that, conceptually, such settings can be reflected in our models as the case in which the value $v_0 - c$ is narrow. Recall that under such circumstances, the potential benefit of DO could be dramatic; see Table 4.

Next, we propose several natural topics for further discussion and research. Consider first a situation in which a retailer is encountered with some mixture of strategic and myopic consumers. We believe that such kind of behavior heterogeneity will not have any significant influence on the spirit of our results. In other words, we expect the marginal value of DO to increase when there is a larger percentage of strategic customers among class-1 customers, together with a large potential for segmentation (i.e., a sufficiently large difference between v_1 and v_0, and a medium value of α_0). Next, suppose that market heterogeneity had to include more than two classes of customers. We make the following conjecture vis-a-vis the marginal benefits of the DO format. First, a retailer should identify the classes of customers he wishes to target. Obviously, classes with relatively low valuations should be *excluded*, particularly if their presence in the market is small. This is obviously true if deep markdown pricing that aims at such classes, might drive classes with higher valuations to wait. In the numerical study, we found that when the traffic intensity to the store is very high, relative to the number of units available, it makes a greater sense to target high-valuation classes exclusively. Therefore, in a multi-class setting, the retailer needs to identify the *classes* of customers to target. To this end, he needs to consider the variance in

[13] See online addendum for more details.

valuations and proportions of the different classes. The marginal benefits of DO appear to be the largest when there is a significant spread in valuations (between the targeted class with the highest valuation and the targeted class with the lowest valuation). Otherwise, high-valuation classes do not offer much potential for segmentation anyway, and hence the display format will hardly matter. The marginal benefits of DO are higher when the proportion of the high-valuation classes among the targeted customer base is at a medium level. If their proportion is very high, it makes sense to target these classes only in the first step mentioned above. If their proportion is too small, these customers' purchasing behavior will depend on the traffic intensity, but not that much on the display format (see detailed explanation in the numerical study section). Finally, for the DO format to be beneficial, the customers' arrival rate should be within a medium range, relative to the order quantity. The technical analysis of the latter subjects is prohibitively hard (see the online addendum, in which we explain the sources of complexity).

Finally, we identify three more research topics related to the exploration of *dynamic* pricing and display strategies. Our study has focused on settings in which the seller pre-announces the premium and end-of-season prices at the start of the season. In retail settings, sellers typically adopt a dynamic pricing strategy, rather than making a commitment to a markdown price path. Our main interest in studying pre-announced pricing strategies is driven by recent research (e.g., Aviv and Pazgal (2008)) which shows that under certain circumstances[14], a price commitment can actually work more effectively than dynamic pricing, in the presence of strategic customers. However, the study of the impact of display formats in dynamic pricing settings is still of significant interest. Based on our research results, together with the aforementioned lesson from Aviv and Pazgal (2008), we suspect that the marginal value of DO will be even more positive when sellers adopt dynamic pricing policies. Next, recall that we have assumed that customers form rational expectations about the purchasing decisions of every other potential customer. Furthermore, we assumed that consumers can perfectly estimate the likelihood of finding a product available if they return to the store at the end of the season. Hence, we consider a market with highly-sophisticated customers that are able to obtain information, calculate, and follow the equilibrium we predict. To complement this normative approach, one may choose to examine the aspect of strategic

[14] See also Cachon and Swinney (2008) for further discussion.

consumer behavior by exploring alternative descriptive models of the ways in which customers may *actually* behave. In particular, it is of interest to study settings in which customers exhibit *adaptive expectations* or *adaptive learning*. Liu and van Ryzin (2007), mentioned in §2, and Popescu and Wu (2007) use adaptive learning models to study markets with repeated interactions between a seller and customers. These recent papers and references therein can provide useful ideas on how to approach the study of the impact of display formats under adaptive expectations. Another interesting possibility is to explore a wider set of display format strategies. We consider two models only – the DA and DO formats. Obviously, variants of these display formats could include situations in which the retailer displays only a part of the inventory, or dynamically decide about the number of items to display on the basis of the time in the season and the actual inventory left in stock.

9 Concluding Remarks

We study two inventory display formats: Display All (DA) and Display One (DO). Under DA, the retailer displays all available units so that each arriving customer has perfect information about the actual inventory level. Under DO, the retailer displays only one unit at a time so that each customer knows about product availability but not the actual inventory level. For each display format and for any given ordering and pricing decisions, we showed that customers follow a threshold-type purchasing rule in equilibrium. Anticipating such behavior, we determined the retailer's profit function and proved certain properties of the optimal order quantity and optimal prices.

When the level of inventory is exogenously given, it is useful to consider the retailer's expected revenue. We found that a change from DA to DO, even without a change of the price path, can never worsen the retailer's expected revenue. This observation supports our hypothesis that the DO format increases the perceived level of shortage risk, and hence drives high-valuation customers to purchase the product at the premium price. When the optimal price path is used in each format, the advantage of using DO obviously increases. In particular, our results suggest that price modification typically plays a larger role than merely changing the display format. We observed that the marginal benefits of DO tend to be the largest when there is a sufficiently large spread in customers' valuations, the proportion of the high-valuation class in the population is at a modest level, and the customers' arrival rate is within a medium range relative to the initial inventory level.

Interestingly, we found that a move from DA to DO is very far from eliminating the adverse impact of strategic consumer behavior on the ability of the retailer to conduct effective price segmentation.

We also considered settings in which the retailer selects both inventory and pricing to maximize expected profit. In addition to the typical role of setting capacity to meet demand, the retailer can use inventory as a lever to influence customers' perception of product scarcity. Interestingly, we noticed that, generally, it is not important for a retailer to modify the level of inventory if he decides to move from a DA to a DO format; a change in the display format, along with an appropriate price modification, is typically sufficient. We found that under modest-to-high store traffic rates and sufficient spread in valuations, the marginal benefit of DO can vary dramatically as a function of the per-unit cost to the retailer. When the retailer's per-unit cost is relatively high, but not too high to make sales unprofitable or to justify exclusive sales to high-valuation customers only, the benefits of DO appear to be at their highest level, and could reach up to 20% increase in profit. This result is of particular interest, since high per-unit costs in our model can also reflect retail environments in which strategic customers weigh-in the possibility of alternative purchases if they wait for clearance sales and the item they wanted is out of stock.

Among all scenarios in which a move from DA to DO is profitable, we focused on those in which a change in inventory is sufficiently beneficial. For this set of scenarios, we studied the type of change in inventory and pricing decisions. Interestingly, we found that across all scenarios, one of two actions takes place: either the premium price is increased along with a *reduction* in inventory, or inventory is increased along with premium price *reduction*.

Our theoretical results demonstrate that by moving from DA to DO, while keeping the price path unchanged, the *volatility* of the retailer's profit decreases. This observation opens an interesting window for future research on settings with risk-prone or risk-averse retailers. Finally, we discussed several model extensions and proposed further ideas for future research.

Acknowledgement. The authors wish to thank the AE and anonymous referees for many valuable suggestions and comments on the earlier version of this work. The authors are also grateful to Sushil Bikhchandani for sharing his ideas throughout this research project.

References

[1] Aviv, Y., A. Pazgal. 2007. Optimal Pricing of Seasonal Products in the Presence of Forward-Looking Consumers. Forthcoming in Manufacturing & Service Operations Management.

[2] Bell, D.E., D. Starr. 1994. Filene's Basement. Harvard Business School Case 9-594-018.

[3] Cachon, G. P., R. Swinney. 2008. Purchasing, Pricing, and Quick Response in the Presence of Strategic Consumers. Forthcoming in Management Science.

[4] Caldentey, R., G. Vulcano. 2007. Online Auction and List Price Revenue Management. Management Science, 53 (5) 795-813.

[5] Debo, L., G. Lai, K. Sycara. 2008. Buy Now and Match Later: the Impact of Posterior Price Matching on Profit with Strategic Consumers. To appear in Manufacturing and Service Operations Management.

[6] Elmaghraby, W., A. Gulcu, P. Keskinocak. 2007. Designing Optimal Pre-announced Markdowns in the Presence of Rational Customers with Multi-unit Demands. Forthcoming in M&SOM.

[7] Fisher, M. L. 2006. Rocket Science Retailing. INFORMS Annual Meeting, Pittsburgh, Nov. 5-8.

[8] Fudenberg, D., J. Tirole. 1991. Game Theory. MIT Press, Cambridge, MA.

[9] Jerath, K., S. Netessine, S.K. Veeraraghavan. 2007. Last-Minute Selling and Opaque Selling. Working Paper, University of Pennsylvania.

[10] Levin, Y., J. McGill, M. Nediak. 2005. Optimal Dynamic Pricing of Perishable Items By a Monopolist Facing Strategic Consumers. Working Paper. Queen's University, Kingston, ON, Canada.

[11] Liu, Q., G. van Ryzin. 2008. Strategic Capacity Rationing to Induce Early Purchases. Management Science, 54 (6) 1115-1131.

[12] Liu, Q., G. van Ryzin. 2007. Strategic Capacity Rationing when Customers Learn. Working paper. Columbia Graduate School of Business.

[13] Marn, M. V., R. L. Rosiello. 1992. Managing Price, Gaining Profit. Harvard Business Review, September–October, pp. 84–93.

[14] Marn, M. V., E. V. Roegner, C. C. Zawada. 2003. The Power of Pricing. The McKinsey Quarterly, February, pp. 27-36.

[15] Phillips R. L. 2005. Pricing and Revenue Optimization, Stanford University Press.

[16] Popescu, I., Y. Wu. 2007. Dynamic Pricing Strategies with Reference Effects. Operations Research, 55 (3) 413-429

[17] Schelling, T.C. 1960. The Strategy of Conflict. Harvard University Press, Cambridge, MA.

[18] Su, X. 2007. Inter-temporal Pricing with Strategic Customer Behavior. Mgmt. Sci., 53 (5) 726-741.

[19] Su, X., F. Zhang. 2008a. Strategic Customer Behavior, Commitment and Supply Chain Performance. Forthcoming in Management Science.

[20] Su, X., F. Zhang. 2008b. On the Value of Commitment and Availability Guarantees When Selling to Strategic Consumers. Working paper.

[21] Yin, R., C.S. Tang. 2006. The Implications of Customer Purchasing Behavior and In-store Display Formats. Unpublished Manuscript. UCLA Anderson School of Management.