Revenue Management for Products with Unknown Quality and Observable Price and Sales History

Laurens Debo
Carnegie Mellon University

Nicola Secomandi
Carnegie Mellon University, ns7@andrew.cmu.edu
Revenue Management for Products with Unknown Quality and Observable Price and Sales History

L.G. Debo and N. Secomandi, Tepper School of Business, Carnegie Mellon University
Pittsburgh, PA, 15213, USA
March 2007

Abstract: This paper studies pricing strategies of a monopolist when customers are strategic and infer the product quality from historical sales and price paths. This behavior may be expected when selling a new, innovative product whose quality is unknown and customers have some, but, imperfect information about the product quality. The monopolist can sell a finite amount of inventory during a finite horizon. Its pricing policy controls the inventory (as in classical revenue management) as well as the customer learning. We derive the structure of the optimal pricing strategy and discuss the implications of strategic customer behavior.

Key words: Revenue Management, Inventory Management, Learning.

1 Introduction and Motivation

Any successful business strategy factors in the fundamental drivers of market uncertainty. Often, these are not well known to firms. Learning these drivers through the market’s reaction to prices helps with shaping current and future business strategies. In markets for new, innovative products, quality is one of the main drivers of uncertainty, not only for the firm, but also for the potential consumers. As a result, consumers also learn from historical sales and price trends. High sales at high prices are a strong indication to future customers that the new product’s quality is high and therefore increases their willingness to pay. We consider a tactical pricing problem that firms face when selling new, innovative products with substantial quality uncertainty during a finite season. High set-up costs make it economical for a firm to produce a large initial batch of products that can be sold before competitors introduce similar products. Our goal is to understand how historical prices and sales influence consumer’s purchasing behavior, and how a firm can control its pricing strategy as a function of the remaining inventory, time horizon and the realized sales over the historical price path.

An example can be found in the market for food supplements. Typically, many new products are introduced of which only a few become successful. When introducing a new product, a company produces one large opening batch at its own or manufacturing partner’s production facilities. There is a limited time window during which the firm can make revenues, before competitors (mass manufacturers) enter the market with similar products produced in large quantities at a much lower cost. The quality of a new food supplement cannot easily be communicated through, e.g., marketing or advertisement campaigns. Yet, some sports enthusiasts or health
conscious customers may have their own private assessment of the quality of new food supplements through previous experiences with similar products, or through consulting specialized product review information. As a result, customers complement their private information with other sources. One such source is the historical price and sales path. Strong sales at high prices may be an indication that the earlier customers have information that the product utility is high. Similarly, weak sales at low prices may be an indication that the product utility is low. Historical sales and price information may influence a future customer’s purchasing decision. Therefore, both the firm and consumers learn from the historical sales and price processes. Not only does dynamic pricing of the initial inventory aim at maximizing revenues during a finite sales season, it also impacts the customer’s quality assessment of new products.

In this paper, we seek a better understanding of the latter effect. To that end, we develop and analyze a stylized model in which a monopolist sells a batch of innovative products during a finite time horizon to customers that have private, but imperfect information about the product quality. Even though in practice customers may only know indirectly about the sales success of a product via, e.g., word-of-mouth, to demonstrate sharply the impact of historical prices and sales on the current customer purchasing behavior, we assume that these are fully observable to all customers. We also assume that customers are rational economic agents.

We address the following research questions: How does a monopolist impact the customers’ learning from historical prices and sales? How do inventory levels impact the learning behavior? How is the pricing policy different with a long vs. a short selling horizon? What are the typical optimal price paths that can be observed? How much expected revenue does the seller lose by ignoring learning?

Our main contribution to the dynamic pricing literature is to show how inventories and the remaining time horizon affect the structure of the optimal pricing policy when customers also learn about the product quality from historical sales and prices. In our setting, when the seller optimally experiments with a price that induces customers to learn, such a price should be based both on the opportunity cost of selling one unit of product and the effect that a successful or unsuccessful sales encounter (positive or negative learning outcome) has on the willingness to pay of future buyers. Thus, the opportunity cost is not always sufficient for making optimal pricing decisions, and, in the case of price experimentation, must be supplemented by the expected benefit of learning, a concept that we make precise in §5. We illustrate numerically the marked differences in optimal price paths when prices are computed with and without modeling learning, and estimate the likely potential benefit for the seller from explicitly modeling this feature. Here, we find that this benefit is substantial, a result that underscores the potential for the practical relevance of our work.

This paper is organized as follows. We review the related literatures in §2. We formalize a stylized model that captures the key characteristics of the problem formulated above in §3. We analyze the optimal pricing policy with ample initial inventory in §4 and with limited initial inventory in §5. We provide structural properties of the optimal pricing policy in §. In §6, we use numerical examples to illustrate the qualitative difference between
the optimal pricing policy with and without taking informational spillovers into account, and to quantify the potential benefit from modeling learning. We discuss our results and future research in §7.

2 Literature Review

Three streams of literature are relevant for the dynamic pricing problem described above. The first literature stream is referred to as the “herding” literature. The second literature stream is referred to as “revenue management.” In the third literature stream, authors study Bayesian learning in inventory context. Below, we discuss the key insights of these literature streams and the novelty of our contributions.

The first literature stream, initiated by Banerjee (1992) and Bikhchandani, Hirshleifer and Welch (1992), analyzes the equilibrium outcome when a sequence of individuals make decisions with incomplete information about the value of an asset. The asset can either have a negative or a positive value. Each individual has private, but inaccurate information about the asset value. Agents do not observe their predecessors’ private information, but do observe the outcome of their predecessors’ decisions (to buy the asset or not). The authors demonstrate that the influence of the observed decisions of the predecessors could be so strong that individuals ignore completely their own information and follow their predecessors’ decision. They refer to this phenomenon as an ‘informational cascade’. Informational cascades can be socially inefficient as agents can take the wrong decision, i.e., buying an asset with negative value, or, not buying a high value asset. The literature on informational cascades has largely focused on understanding the equilibrium outcome. Only a few papers study how firms can manipulate the formation of informational cascades. One obvious control variable that can be used to manipulate the formation of informational cascades are prices. When the price is very low (high), as long as the lowest realization of the true quality is above (below) the price, all (no) customers should buy. Extreme, high or low, prices reduce the informational externalities. Bose, Orosel and Vesterlund (2002) study the optimal pricing problem over an infinite horizon of a monopolist who sells an infinite quantity of a good to sequentially moving customers who have private information about the common, but uncertain, value of this good. They find that informational cascades occur with positive probability only if the monopolist is sufficiently patient and the true state of the world is known with high probability. Our model is similar to theirs, except that we consider a finite horizon and a finite inventory of products to be sold. This situation is more appropriate when the monopolist has strong economies of scale for the opening batch and only has a limited time window before competitors with similar, lower priced products enter the market, or, when the product life cycle is short.

The price-based revenue management literature studies optimal pricing strategies of a limited inventory for a retailer during a finite time horizon. We refer the reader to Talluri and van Ryzin (2004) for a comprehensive introduction to this literature. In this literature stream, customers are typically not considered to be strategic. In each period, the monopolist faces a stochastic price sensitive demand function that depends on the current price that is quoted. The main findings are that when the inventory is high or when the time until the end of
the horizon is short, the optimal price is low in order to clear any remaining inventory. With scarce inventory or ample remaining time the price is high. Strategic customer behavior has only recently been introduced in this literature. The closest to our problem is that studied by Popescu and Wu (2006), who study optimal dynamic pricing strategies when the customer demand is sensitive to a firm's pricing history. As the firm manipulates prices, customers form a reference price to which they anchor future prices. Customers have limited memory and cognitive biases. These authors find that optimal pricing policies in the long run induce a perception of monotonic prices whereby customers always perceive a discount or a surcharge. In their model, product quality is known to all customers. As a result, prices do not contain any quality-related information. Liu and van Ryzin (2005) study strategic behavior of consumers that optimally time their purchases in anticipation of future prices and availability. The resulting threat of shortages creates an incentive for customers to purchase early at higher current prices. Liu and van Ryzin study when it is profitable to create such shortages. For a review of related models with forward looking consumers see Elmaghraby and Keskinocak (2006). In these models, customers anticipate future prices (or discounts) and/or product availability with perfectly known quality. There are very few papers in this literature that study demand learning. Araman and Caldentey (2005) and Farias and Van Roy (2006) extend traditional dynamic pricing models to situations where the seller does not know the size of the market but does know the probability distribution of the customer valuations. Bitran and Wadhwa (1996) study a dynamic pricing model where the seller knows the parameters of the arrival process but does not know one of the parameters of the probability distribution that describes the customer's willingness to pay. Differently from these authors, in this paper the product quality is unknown to both the seller and customers. Customers observe privately some information related to the product quality, but complement this information with the realized historical sales and prices.

Finally, there is a related literature that studies how firms can set inventory levels and update parameters of the demand distribution from observing realized sales (Azoury, 1985, Lariviere and Porteus, 1999). In that literature stream, the firm updates an unknown parameter of the demand distribution using historical sales information. However, when the firm stocks out, the demand data is censored, which leads to a reduction of the learning capability. As a result, the firm may want to invest in excess inventory in order to enhance the learning. In that literature stream, the demand uncertainty is not linked to product quality uncertainty and customers do not behave strategically.

None of the above literatures addresses the key characteristics of the pricing problem studied in this paper: the herding literature does not consider a finite horizon and finite inventories. The revenue management literature does not take into account the informational spillovers that influence the strategic customer behavior. In this paper, we develop and analyze a model that addresses both these features.
3 The Model

Assume that the exact value of a product, $V_\omega$, depends on the state of the world, $\omega$, which is unknown. The prior distribution is: $\Pr(\omega = h) = \mu_0$ and $\Pr(\omega = l) = 1 - \mu_0$, where $0 < V_l < V_h$. During a season of $T$ periods, with probability $\lambda$, a customer arrives in each period. Each customer observes privately a signal, $s \in \{g, b\}$, that is correlated to $\omega$: $\Pr(s = g|\omega = h) = \Pr(s = b|\omega = l) = \alpha \in (\frac{1}{2}, 1)$. (If $\alpha = 1/2$ then the signal is not informative, if $\alpha = 1$ then a realized signal perfectly reveals the state of the world.) A manager starts with an initial inventory of $x_0$ units and can set the price $p_t$ in each period. A customer arriving in period $t$ decides whether to buy the product, $a_t = 1$, provided, that there is inventory left over, or not to buy the product, $a_t = 0$. The customer observes the history $H_t = \{(a_\tau, p_\tau), 0 \leq \tau \leq t - 1\}$ of actions taken by previously arriving customers and the history of prices charged by the manager. If the customer buys the product at price $p_t$, his utility is $\tilde{V} - p_t$, otherwise, his utility is 0. The game is characterized with $(\mu_0, V_h, V_t, \alpha, x_0, T)$. Note that in a certain period, the customer and the manager only observe the actions that previous customers have taken, not the realization of their signal.

Define $\Phi(\mu) = \mu \alpha + (1 - \mu)(1 - \alpha)$ as the probability that a good signal will be realized for a given prior belief, $\mu$. Define $\Psi^+(\mu) = \mu \alpha / \Phi(\mu)$ and $\Psi^-(\mu) = \mu (1 - \alpha) / [1 - \Phi(\mu)]$, which are the updated probabilities that the product value is high after observing a good and bad signal respectively, when the prior belief that the product is of high value is $\mu$. It can easily be shown that $\Psi^-(\mu) < \Psi^+(\mu)$. Suppose that the manager charges price $p$, and the prior belief is $\mu$, then the rational decision, $a^*$, of a customer observing $s$, provided that the inventory is strictly positive, $x > 0$, is:

$$a^*(x, \mu, g; p) = \begin{cases} 1, & V^+(\mu) \geq p \\ 0, & \text{o/w} \end{cases} \quad \text{and} \quad a^*(x, \mu, b; p) = \begin{cases} 1, & V^-(\mu) \geq p \\ 0, & \text{o/w} \end{cases} \quad (1)$$

with $V^-(\mu) = V_t + \Psi^-(\mu) (V_h - V_t)$ and $V^+(\mu) = V_t + \Psi^+(\mu) (V_h - V_t)$. It is immediate that the only prices that need to be considered in state $(\mu_t, x_t)$ are $V^-(\mu)$ and $V^+(\mu)$. Let $V_t(\mu_t, x_t)$ be the equilibrium profit of the manager in a subgame characterized by $(\mu_t, x_t)$. Define $V_T(\mu_T, x_T)$ := 0, $\forall (\mu_T, x_T)$. I.e. we assume that the firm does not receive any money for leftover inventory because the competition will take away all the residual demand. For $t < T$, and $\forall \mu_t$, let

$$V_t(\mu_t, x_t) = \begin{cases} \max\{v_t^+(\mu_t, x_t), v_t^-(\mu_t, x_t)\} & x_t > 0 \\ 0 & x_t = 0 \end{cases}$$

$$v_t^+(\mu_t, x_t) = \lambda \{\phi(\mu_t) [V^+(\mu_t) + \delta V_{t+1}(\Psi^+(\mu_t), x_t - 1)] + (1 - \phi(\mu_t))\delta V_{t+1}(\Psi^-(\mu_t), x_t)\} + (1 - \lambda)\delta V_{t+1}(\mu_t, x_t)$$

$$v_t^-(\mu_t, x_t) = \lambda [V^-(\mu_t) + \delta V_{t+1}(\mu_t, x_t - 1)] + (1 - \lambda)\delta V_{t+1}(\mu_t, x_t).$$

When charging a price $V^-(\mu)$, the customer buys irrespective of the realization of his signal. When charging price $V^+(\mu)$, with probability $\Phi(\mu)$, the customer buys the product (i.e. when he observes a good signal) and
with probability $1 - \Phi(\mu)$, the customer does not buy the product. If no customer arrives (with probability $1 - \lambda$), then, the inventory position and the prior belief that the product is of high quality do not change. If there is no inventory left over, the expected profits until the end of the horizon are zero.

4 Analysis: Optimal Pricing with Ample Inventory

In this section, we study the case in which there is enough inventory to cover the demand over the full horizon. As the demand per period is at most one, this means that the initial inventory satisfies $x_0 \geq T$. When the inventory is strictly positive, the optimal pricing in the last period $T$ depends on the market belief, $\mu_T$. In period $T$, the single period pricing policy is optimal.

We define $\Omega(\mu) := V^+(\mu)\Phi(\mu) - V^-(\mu)$ as the difference in expected revenues when the price is $V^+(\mu)$ and when the price is $V^-(\mu)$. Note that in the former case, a sale will be made with probability $\Phi(\mu)$. We also refer to $\Omega(\mu)$ as the per period additional expected revenue from charging a high price. Lemma 1 provides a condition for this quantity to be strictly positive, in which case charging $V^+(\mu_T)$ in period $T$ is optimal.

**Lemma 1 (Per period additional expected revenue)** If

$$\frac{V_h}{V_l} > \frac{(\alpha^2 - 3\alpha + 1)^2}{(\alpha + \alpha^2 - 1)^2}$$

for $\alpha > \frac{1}{2}\sqrt{5} - \frac{1}{2} = 0.61803$, then $\Omega(\mu) = 0$ has two roots, $\mu_-, \mu_+ \in [0, 1]$ and $\Omega(\mu) > 0$ for all $\mu \in (\mu_-, \mu_+)$, and $\Omega(\mu) < 0$ for all $\mu \in [0, 1] \setminus [\mu_-, \mu_+]$. Otherwise $\Omega(\mu) < 0$ for all $\mu \in [0, 1]$.

Figure 1 illustrates $V^+(\mu)$, $V^-(\mu)$ and $\Omega(\mu)$ for $V_h = 2$, $V_l = 1$, and $\alpha = 0.8$. Note that both prices are increasing in $\mu$ and that both prices become close when the belief $\mu$ is either very low or very high. In all cases when we refer to a high or a low price, the price is always with respect to the market belief, $\mu$. Thus, it may be that the low price with a high market belief is higher than the high price with a low market belief. We will occasionally refer to setting a high price, $V^+(\mu)$, as “experimenting” with the prices. When the belief is that the value of the product is low ($\mu \to 0$), then the customer does not buy the product for sure. Therefore, $\Omega(0) = -V_l < 0$. When the belief is that the value of the product is high ($\mu \to 1$), then the customer buys the product with probability $\alpha$. Therefore, $\Omega(1) = \alpha V_h - V_h < 0$. It follows that when the belief is very strong that the product is either of high quality or of low quality, charging a high price (w.r.t. the belief) is never optimal. It may be optimal to charge a high price when the belief is weak. According to Lemma 1 this is the case when either the ratio between $V_h$ and $V_l$ is high or $\alpha$ is high (since the right hand side of (2) decreases in $\alpha$ provided that $\alpha > 0.61803$). Therefore, this lemma also characterizes the optimal decision in the last period.

The following analysis is based on the following dynamic program, which is obtained by defining $V_t(\mu_t) :=
Figure 1: $V^+(\mu)$, $V^-(\mu)$ and $\Omega(\mu)$ for $V_h = 2$, $V_l = 1$ and $\alpha = 0.8$.

We characterize the optimal policy in Proposition 1.

Proposition 1 (Optimal policy with ample inventory)  
(a) The optimal price policy is determined by charging $V^+(\mu_t)$ if $\mu_t \in (\mu_{t-1}, \mu_t)$ and otherwise $V^-(\mu_t)$, where $\Delta V^\infty(\mu_t) > 0 \Leftrightarrow \mu_t \in (\mu_{t-1}, \mu_t)$, and $\Delta V^\infty(\mu_t) < 0 \Leftrightarrow \mu_t \notin (\mu_{t-1}, \mu_t)$, and $\Delta V^\infty(\mu_t) = 0 \Leftrightarrow \mu_t \in (\mu_{t-1}, \mu_t)$. 
(b) If there is a $t$ such that $(\mu_{t-1}, \mu_t)$ is not empty, as the remaining time until the end of the horizon increases, the range over which $V^+(\mu)$ is optimal increases:

$$\mu_{t-1} \leq \mu_t \leq \mu_{t-1}.$$ 

Figure 2 illustrates this structure for $\lambda = 0.8$, $V_h = 2$, $V_l = 1$, $x_0 = 4$, $T = 3$, $\alpha = 0.8$ and $\delta = 0.9$. Part (a) of Proposition 1 establishes that a high price will be charged when the belief is weak, i.e., neither close to 0 nor close to 1. Part (b) shows how the belief interval where a high price is charged changes over time: when there are fewer time periods left, the range of market beliefs for which $V^+(\mu)$ is charged decreases. Thus, it is optimal to charge a high price (i.e., experiment with prices) in the beginning of the horizon. This is intuitive as more periods are left during which revenues can be made. The optimal pricing policy generates the following price path: assume
that in a certain period the market belief is \( \mu_t \in (\mu_t, \bar{\mu}_t) \) and assume that condition (2) is not satisfied. Then, the firm charges a high price (with respect to the market belief). Depending on the sales realization, the belief (and thus the prices) may increase or decrease over time. However, the threshold beliefs \( (\mu_t, \bar{\mu}_t) \) also shrink over time, and will become the empty set in period \( T \) (due to condition (2) not being satisfied). As a result, there will exist a period, \( T' \), after which the price “freezes.” If the frozen belief, \( \mu_{T'} \), is higher than \( \bar{\mu}_{T'} \), then, for the remaining periods, a “relatively high” low price \( (V^- (\mu_{T'})) \) is charged. If the frozen belief, \( \mu_{T'} \), is lower than \( \bar{\mu}_{T'} \), then, for the remaining periods, a “relatively low” low price \( (V^- (\mu_{T'})) \) is charged. We can interpret this result as follows: the firm experiments during the initial period with the prices by tagging these to the market belief. As a result, market information is generated. If sales are successful, the market belief and the prices increase. Otherwise, the market belief and the prices decrease. In both cases, the firm stops experimenting and prevents further market learning by charging \( V^- \): As long as there is market learning, the prices risk falling due to, e.g., bad luck in sales. The firm stops market learning because it either wants to enjoy the high prices until the end of the horizon or because it wants to prevent a downward trend. We can consider the first strategy as a “skimming” strategy and the latter as a “penetration” strategy. As it will be more likely that sales are successful when the product is of high quality, a skimming strategy will emerge for high quality products and a penetration strategy for low quality product. However, a series of early “lucky strikes” in the beginning of the horizon of a low quality product may also lead to price increases and eventually a high price, as well as bad luck with initial sales may lead to price decreases and eventually a low price for high quality products.
Proposition 2 (Optimality of a low price in each period) Let \( \hat{W}_t^\infty (\mu_t) \) satisfy the following recursion:

\[
\hat{W}_t^\infty (\mu_t) = \hat{\Omega} (\mu_t) + \delta \left( \{ \hat{W}_{t+1}^\infty (\mu_{t+1}) \}^+ + \mathbb{E}_{\tilde{\mu}_{t+1}} \left[ \{ \hat{W}_{t+1}^\infty (\tilde{\mu}_{t+1}) \}^+ | \mu_t \right] \right)
\]

where \( \hat{W}_T^\infty (\mu_T) = \Omega (\mu_T) \) and

\[
\hat{\Omega} (\mu) = \Omega (\mu) + \delta \mathbb{E}_{\tilde{\mu}} [V^- (\tilde{\mu}) | \mu].
\]

If \( \hat{\Omega} (\mu) < 0 \) for all \( \mu \in [0, 1] \), then, for any \( t \), it is optimal to charge \( V^- (\mu_t) \), otherwise, when the horizon is long enough, there exists a \((\mu_t, \bar{\mu}_t) \) \( \neq \emptyset \) in which it is optimal to charge \( V^+ (\mu_t) \).

Proposition 2 provides a recursion for \( \hat{W}_t^\infty (\mu_t) \), which determines the price policy. Assume that \( \hat{W}_t^\infty (\mu_{-\infty}) < 0 \), i.e., at the beginning of an infinite horizon (starting at \( t = -\infty \) and ending at \( t = T \)), it is never optimal to experiment with high prices (i.e., to charge a high price w.r.t. the market belief). It that were the case, the recursion formula reduces to:

\[
\hat{W}_t^\infty (\mu_{-\infty}) = \hat{\Omega} (\mu_{-\infty}) + \delta \hat{W}_t^\infty (\mu_{-\infty})
\]

as the positive part of \( \hat{W}_t^\infty (\mu_{-\infty}) \) is zero. Therefore, \( \hat{W}_t^\infty (\mu_{-\infty}) = \hat{\Omega} (\mu_{-\infty}) / (1 - \delta) \). The latter is only consistent with the original assumption when \( \hat{\Omega} (\mu_{-\infty}) < 0 \). As a result, the sign of \( \hat{\Omega} (\mu) \) determines whether the firm will experiment with high prices or not. It is interesting to note that \( \delta \mathbb{E}_{\tilde{\mu}} [V^- (\tilde{\mu}) | \mu] \geq 0 \). Consider the case when \( \Omega (\mu) < 0 \) over \([0, 1]\). This implies that in the last period, the firm never experiments with its prices. It follows that when the firm cares more about the future (high \( \delta \)), or when the expected revenue from charging \( V^- (\mu) \) is high, the firm may find it optimal to experiment with high prices when the horizon is long enough as always charging a low price cannot be the optimal solution to the infinite horizon problem. Figure 3 displays the functions \( \Omega (\mu) \) and \( \hat{\Omega} (\mu) \) for the parameters of Figure 2. For these parameter settings, Figure 3 shows that in the last period the firm will never experiment with prices; however, if the horizon is long enough, the firm will experiment with prices when the market is the most “confused.”

So far, we have assumed that enough inventory is available to cover the maximum possible demand until the end of the horizon. Next, we study the optimal pricing policy with a limited supply of inventory.

5 Analysis: Optimal Pricing with Limited Inventory

In this section, we study the case in which there is not enough inventory to cover the full horizon. As the demand per period is at most one, the initial inventory satisfies \( x_0 < T \).

In Proposition 3, we characterize the first and second order properties of the optimal value function.

Proposition 3 (Optimal value function) If \( \alpha \in (1/2, 1] \) then \( V_t (\mu_t, x_t) \) is jointly increasing in \( (\mu_t, x_t) \), concave in \( x_t \) given \( \mu_t \), and convex in \( \mu_t \) given \( x_t \), \( \forall t \).

It is intuitive that when there is more inventory at hand, the expected future revenues are higher, as the ability to satisfy customer demand increases. Furthermore, as the market belief that the value of the product is
high increases, the expected future revenues are also higher as the firm can charge higher prices. The concavity of the value function in the inventory is a classical revenue management result; the marginal value of one extra unit of inventory decreases. For example, when more units are available than the remaining length of the selling horizon, the marginal value of one extra unit of inventory will be zero as (in our model) no more than one customer arrives per period. It is interesting that the optimal value function is convex in the market belief, which means that the marginal value of a small (infinitesimal) increase in the market belief increases as the market belief itself increases.

We now turn our attention to characterizing the structure of the optimal pricing policy as a function of the market belief and inventory level. We find it useful to provide an alternative, yet equivalent, formulation of our model. Define

$$W_t(\mu_t, x_t) := \Phi (\mu_t) V_{t+1} (\Psi^+ (\mu_t), x_t - 1) + (1 - \Phi (\mu_t)) V_{t+1} (\Psi^- (\mu_t), x_t) - V_{t+1} (\mu_t, x_t - 1).$$

Recalling that $V_T (\mu_T, x_T) = 0, \forall (\mu_T, x_T)$, and $V_T (\mu_T, 0) = 0$, we can rewrite the recursion for $x_t > 0$ as

$$V_t (\mu_t, x_t) = \lambda \left( V^- (\mu_t) + V_{t+1} (\mu_t, x_t - 1) + \{ \Omega (\mu_t) + \delta W_t (\mu_t, x_t) \}^+ \right) + (1 - \lambda) \delta V_{t+1} (\mu_t, x_t).$$

This reformulation makes it clear that it is optimal to charge a high price provided that $\hat{W}_t (\mu_t, x_t) \geq \Omega (\mu_t) + \delta W_t (\mu_t, x_t) > 0$. In Proposition 4, we show that we can characterize when this inequality holds by using no more than two critical market belief values that depend on the available inventory and the number of remaining time periods.
Proposition 4 (Optimal policy with limited inventory) (a) The optimal pricing policy is determined by charging $V^+ + \mu_t$ if $\mu_t \in \left( \mu_t(x_t), \pi_t(x_t) \right)$ and otherwise $V^- - \mu_t$, where $W_t(\mu_t, x_t) > 0$ if $\mu_t \in \left( \mu_t(x_t), \pi_t(x_t) \right)$, and $W_t(\mu_t, x_t) < 0$ if $\mu_t \notin \left( \mu_t(x_t), \pi_t(x_t) \right)$.

(b) For a given belief, $\mu_t$, $W_t(\mu_t, x_t)$ is decreasing in $x_t$.

(c) The optimal policy satisfies

$$
\mu_t(x_t) \leq \mu_{t+1}(x_t) \leq \pi_{t+1}(x_t) \leq \pi_t(x_t)
$$

and

$$
\mu_t(x_t - 1) \leq \mu_t(x_t) \leq \pi_t(x_t) \leq \pi_t(x_t - 1)
$$

Part (a) of Proposition 4 states that the optimal pricing policy has a two threshold structure: in a given period and with a given inventory level, the seller should charge a “high” price when the belief is in between the lower and higher belief thresholds, and a “low” price otherwise.

Part (b) implies that the following inequality holds:

$$
W_t(\mu_t, x_t) \geq W_t^\infty(\mu_t).
$$

The interpretation here is that when inventories are limited, the seller has an incentive to set high prices, which makes historical price decisions salient to future customers, and, therefore, impacts the market belief.

Part (c) characterizes the behavior of the two belief thresholds in the remaining time (respectively, inventory) for a given inventory (respectively, remaining time). Given an inventory level, when more periods remain before the end of the horizon, the range of beliefs for which a high price is charged is larger; in a given period, when there is more inventory, the range over which a high price is charged is narrower. These properties are appealing, because increasing the time available to sell a given amount of inventory makes the seller more prone to experiment, while increasing the amount of inventory to be sold in a given amount of time makes the seller less prone to experiment (recall that a sale occurs for sure if the seller charges a “low” price and a buyer does arrive in any given period).

Figure 4 illustrates the structure of the optimal pricing policy. The parameters are those used in Figure 2. We plot the value of $W_{T-3}(\mu_t, x_t)$ for different levels of inventory. When the inventory is larger than 4, the policy is exactly the same as on Figure 2, as at most four sales can be made between period $T - 3$ and period $T$. Notice that as inventory becomes scarcer, the region of beliefs over which the firm experiments with prices, i.e., where $V^+(\mu_t)$ is optimal, widens.

We now offer an interpretation of the optimal pricing decisions that brings to light a basic difference between our setting and traditional dynamic pricing. Since the seller can always sell a unit of product to an arriving customer at a “low” price, the seller’s choice between a “low” and a “high” price relies on evaluating the tradeoff
between a sure sale without learning, and a potential sale, which if successful causes customers to raise their valuations, but otherwise causes customers to lower them. Intuitively, one should optimally charge a low price provided that the net value of a sure sale (price minus opportunity cost) exceeds the sum of the expected value of selling at a high price and the expected benefit of learning induced by this price minus the expected opportunity cost of making a sale. We now formalize this basic intuition by using the concepts of net prices.

We interpret $V^-(\mu_t)$ and $V^+(\mu_t)$ as gross prices, and we introduce the concepts of opportunity cost and expected benefit from learning to arrive at the notions of net prices. With respect to $V^-(\mu_t)$, we define net price $\hat{V}_t^-(\mu_t, x_t)$ in stage $t$ and state $(\mu_t, x_t)$ as the price $V^-(\mu_t)$ minus the opportunity cost of selling one unit of inventory, which is the quantity

$$\hat{V}_t^-(\mu_t, x_t) = V_t^-(\mu_t) - OC_t(\mu_t, x_t)$$

The opportunity cost $OC_t(\mu_t, x_t)$ is the current expected value of the future net revenue foregone by selling one unit of inventory for sure in stage $t$ and state $(\mu_t, x_t)$. With respect to $V^+(\mu_t)$, we define net price $\hat{V}_t^+(\mu_t, x_t)$ in stage $t$ and state $(\mu_t, x_t)$ as the expected price $\phi(\mu_t)V^+(\mu_t)$ plus the expected benefit of learning minus the
Expected opportunity cost:
\[
\hat{V}_t^+ (\mu_t, x_t) = V_t^+ (\mu_t) + \text{EBL}_t (\mu_t, x_t) - \phi (\mu_t) \text{OC}_t (\mu_t, x_t)
\]
\[
\text{EBL}_t (\mu_t, x_t) = \frac{\phi (\mu_t) \delta [V_{t+1} (\Psi (\mu_t), x_t) - V_{t+1} (\mu_t, x_t)]}{[1 - \phi (\mu_t)] [V_{t+1} (\mu_t, x_t) - V_{t+1} (\Psi (\mu_t), x_t)]}.
\]

Here, when charging \(V^+ (\mu_t)\) in stage \(t\) and state \((\mu_t, x_t)\), conditional on there being a customer arrival, the seller expects to gain from a rise in the market belief if a sale occurs (positive learning outcome), and expects to lose from a drop in the market belief if a sale does not occur (negative learning outcome). The salient tradeoff of the pricing problem can be evaluated by comparing the net prices. It can easily be shown that it is optimal to charge \(V^+ (\mu_t)\) in stage \(t\) and state \((\mu_t, x_t)\) if and only if \(\hat{V}_t^+ (\mu_t, x_t) \geq \hat{V}_t^- (\mu_t, x_t)\), and that at most one net price can be negative in any stage and state.

### 6 Numerical Illustrations

In this section we illustrate numerically some of the structural properties established in §§4-5. We also estimate the potential benefit of our model by comparing its performance against the traditional price-based dynamic revenue management model that ignores the evolution of the belief. For completeness, this model is now presented.

**Dynamic pricing without learning.** The prior belief is \(\mu_0\). The price that the firm charges is either \(V^+ (\mu_0)\) or \(V^- (\mu_0)\) and does not change over time. Also, the firm’s belief that a customer will buy the product at a high (low) price is \(\Phi (\mu_0) (1 - \Phi (\mu_0))\) and does not change over time. Define \(V^R (x_T) := 0, \forall x_T\).

For \(t < T\), let the recursion is
\[
V_t^R (x_t) = \begin{cases} 
\max \{v_t^{R^+, +}(x_t), v_t^{R^-, +}(x_t)\} & x_t > 0 \\
0 & x_t = 0
\end{cases}
\]

\[
v_t^{R^+, +}(x_t) = \lambda \{\Phi (\mu_0) [V^+ + \delta V_{t+1}^R (x_t - 1)] + (1 - \Phi (\mu_0)) \delta V_{t+1}^R (x_t)\}
\]

\[
v_t^{R^-, +}(x_t) = \lambda [V^- + \delta V_{t+1}^R (x_t - 1)] + (1 - \lambda) \delta V_{t+1}^R (x_t).
\]

### 6.1 Ample inventory

We first consider the ample inventory case. The selling horizon ends in period \(T = 10\). The arrival probability, \(\lambda\), is 0.5. The initial belief, \(\mu_0\), is 0.5 and its strength, \(\alpha\), is 0.75. The low valuation, \(V_l\), is 0.5, the high valuation, \(V_h\), is 2. The discount factor, \(\delta\), is 1. The state of the world is \(h\) for the simulation. We solve the dynamic program with \(x_0 \geq T\) units of available inventory. Figure 5 illustrates the optimal prices in stages 8 and 9. The
"humps" in these plots denote high optimal prices. It is clear from this figure that the optimal price will freeze to $V^{-}(\mu_t)$ no later than stage $t = 9$. To illustrate this observation, we simulate price paths assuming that $\omega = h$, $\mu_0 = 0.5$, and $x_0 = 10$. We compare the price trajectory of the model with learning to the price trajectory of the model in which neither the firm nor the customers learn from the sales and price history.

Figure 6 displays some price trajectories: while prices may oscillate over time, they will eventually stabilize to a “low” price. In panels (a), (c), and (d) this happens before period 9, while in panel (b), this occurs in period 9. Also, on the top row, prices stabilize at a relatively high level. This is a skimming strategy. On the bottom row, prices stabilize at a low level, this is penetration strategy. Note that this is due to a series of back luck in initial sales, despite that the state of the world, $\omega$, is high. The no-sales at high prices make the market belief go down. It is interesting to note that the traditional dynamic pricing model employs a constant price in this case.

### 6.2 Limited inventory

We now consider the limited inventory case. We illustrate the possible price paths that can occur in equilibrium in the same setting as in the ample inventory case, except that the signal strength is 0.65, the initial inventory, $x_0$, is 6 units and the selling horizon ends in period $T = 14$. Thus, this is a fairly constrained problem with load factor (total expected demand divided by initial inventory) of 1.17.

**Optimal prices.** Figure 8 displays the optimal prices in time period 3, panel (a), and period 9, panel (b). As shown in part (c) of Proposition 4, given an inventory level, the humps shrink as the time-to-go decreases. Since charging a high optimal price means that the seller is experimenting with price in the hope to elicit valuable
Figure 6: Optimal price paths with and without learning in the ample inventory case.
information about the belief, this means that the seller optimally reduces this experimentation when the end of
the selling horizon approaches.

**Comparison of price trajectories with and without learning.** Figure 7 displays several price paths
obtained by applying the optimal decision rules of the models with and without learning on sample paths of
customer demands and signals. By convention we let the optimal price be \( V^+ (\mu) \) whenever \( x = 0 \). Panel (a).
This panel illustrates a case where the price path with learning is always above the price path without learning.
Consider the path with learning. In this case, two sales occur in the first two periods at a “high” price, which
causes the belief to be updated upward. In periods 3 and 4, no arrivals occur, so in period 5 it is optimal to
charge a “low” price. A sale occurs in this period without belief update. In period 6 there is no arrival, while in
period 7 a sale occurs at a “low” price. Since there are only 2 units of available inventory, it is optimal to charge
a high price in period 8, when a sale also occurs. The belief is updated upward, and it is optimal to charge a “high” price in period 9, when a customer arrives but no sale occurs. This causes the belief to be updated
downward, but it is still optimal to charge a “high” price both in period 10, when no sale occurs, and in period
11, when a sale occurs and inventory is depleted. So in this case, the optimal price path starts in the high-price
region, transitions down into the low-price regions, and then moves up into the high-price region again.

In the case without learning, the optimal initial price is low. Since a sale occurs in period 1, the optimal
price in period 2 is high. A sale occurs, and it is optimal to maintain a high price. No arrivals occur in periods
3 and 4, while in period 5 an arrival occurs, it is optimal to charge a high price, but no sale occurs. It is then
optimal to lower the price, no arrival occurs in period 6, while a sale occurs in period 7. It is still optimal to charge a low price in period 8, when another sale occurs. Inventory scarcity makes it optimal to raise the price
in period 9, when there is an arrival but no sale occurs. It is still optimal to charge a high price in period 10,
but no arrival occurs. Since time is scarce, it is optimal to drop the price in period 11, when a sale also occurs.
Inventory scarcity makes it optimal to raise price in period 12, when no sale occurs even if there is a customer
arrival. This causes price to drop in period 13, when there is no arrival. Finally, the last unit of inventory is sold
at a low price in period 14.

In this case, it is interesting to note that in period 5 the low price with learning is equal to the high price
without learning. Thus, a purchase occurs for sure in the learning case, while a purchase does not occur in the
case without learning, even if the arriving customer observes the same signal. This is because in the learning
case customers have updated their belief upward (to 0.7752 from the initial 0.50 level), and hence their effective
willingness to pay is higher than in the case without learning, where the belief is never updated. Thus, in the
case with learning the seller is able to successfully sustain charging higher prices toward the end of the horizon,
while this is not possible in the case without learning.

**Panel (b).** This panel displays a case where, aside from the initial period, these relationships are reversed. In
the learning case, it is optimal to charge a high price in period 1, when a customer arrives, but no sale occurs.
Figure 7: Optimal price paths with and without learning in the limited inventory case.
Hence, the belief is updated downward, and it is optimal to charge a low price in period 2. No arrivals occur in periods 2 and 3, and it is optimal to charge a low price in periods 4-6, with customers arriving and, consequently, sales occurring. Inventory scarcity makes it optimal to charge a high price in period 7, but no arrival occurs. Hence, in period 8 it is again optimal to charge a low price, a customer arrives and a sale is made. In period 9, it is optimal to charge a high price, but even if an arrival occurs there is no sale. The belief is updated downward, and in period 10 it is optimal to charge a low price. No arrival occurs, and in period 11 there is only 1 unit of inventory. The seller optimally charges a high price, but the arriving customer does not purchase. So the price drops in period 13, remains low in period 14 because no customer arrives in period 13, a sale occurs in period 14, and the inventory is depleted.

In the case without learning, the seller updates the optimal price upward in period 2, downward in period 3, and again upward in period 5. From this period onward, the optimal price remains the same, and all the available inventory is sold out by period 10. It is interesting to notice that in this case sales occur in the case without learning at prices that are higher than those charged in the case without learning, in which case sales do not occur. The reason is that learning causes the belief to decrease over time, which effectively cause the willingness to pay of customer to drop over time.

**Panel (c).** In the learning case, a purchase occurs in period 1 when the high price is charged and the belief is updated upward. In period 2, a high price is charged, a customer arrives, but no sale is made, and the belief is updated downward. In period 3, a high price is charged but no customer arrives. In period 4, a low price is charged and a sale occurs at a low price. The price is raised in period 5, when a sale is also made, so the belief increases. In period 6, a high price is charged but no arrival occurs, so in period 7 a low price is charged and a sale is made. In period 8, a sale occurs at the high price, and the belief is updated upward. In periods 9-11,
a high price is charged but no customer arrives. In period 12, there is a single unit of inventory, and the seller
decides to charge a low price. Inventory is depleted in period 13. The interesting feature of this panel is the
transition from a high to a low price in period 12. In the case without learning, the price and sale paths follow a
similar pattern to the case with learning. However, the seller is forced to drop the price is period 13 to a much
lower case to clear the last unit of inventory.

Panel (d). This panel illustrates a case where the two paths cross multiple times. In the learning case, no
sale occurs at the high price in period 1, so the belief is updated downward and a low price is charged in period
2. A sale occurs in this period. In each of periods 3-5 a sale occurs at a high price, and the belief is updated
upward. In period 6 no arrival occurs, while in period 7 a high price is charged, a customer arrives but no sale
occurs. Hence, the belief is updated downward. It remains optimal to charge a high price in periods 8 and 9,
when no customer arrives. In period 10, it is optimal to charge a low price and a sale occurs. In period 11, there
is a single unit of available inventory, and it is optimal to charge a high price. A sale occurs and the inventory is
depleted. In the case without learning, a sale occurs in period 1 at a low price, and from this period onward it
is optimal to charge a high price. The entire remaining inventory is sold by period 11. The interesting feature
of this comparison is that the belief path in the learning case is nonmonotone over time.

Panels (e) and (f). These panels show two cases where the price paths with and without learning are the
same for about half of the periods (ignoring the first period), but the paths with learning are above and below,
respectively, those without learning for the remaining periods. In panel (e), the two price paths are identical up
to period 7, when a sale occurs at a high price in both cases. In the learning case, the belief is updated upward,
and sales at high price are also made in periods 8 and 9. The inventory is depleted in period 9. The same
happens in the case without learning, but in this case the belief is not updated and the high prices charged by
the seller are lower than in the previous case. In panel (f), the price in the learning case oscillates between high
and low levels in the first four periods: when a high price is charged no arrival occurs, but when a low price is
charged a sale occurs. Then in period 5, a high price is charged, an arrival occurs, but no sale is made. Hence,
the belief is updated downward. In period 6 a sale occurs at the low price, in periods 6-8 a low price is charged
and a sale occurs in periods 6 and 8. In period 9 a high price is charged, a customer arrives, but no sale is made.
The belief is updated downward, and a low price is charged in periods 10-14, with sales occurring in periods 12
and 14. The inventory is exhausted in period 14. The interesting feature of panel (f) is that the seller effectively
shuts down learning in periods 10-14. The case without learning is similar, with the difference being that when a
purchase does not happen, the belief is not updated downward, so that the seller is able to charge higher prices
in the remaining periods.
Table 1: The benefit of making pricing decisions by modeling the belief evolution: % improvements of the optimal pricing policy relative to the value of the price-based traditional revenue management policy in stage 1 evaluated when customers do update their beliefs. Note that the averages are not averages of the percentage improvements. They are calculated on the total improvement across all the inventory levels.

<table>
<thead>
<tr>
<th>Initial Inventory</th>
<th>Initial Belief ($\mu_0$)</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>55.37</td>
<td>49.58</td>
<td>45.60</td>
<td>42.73</td>
<td>40.55</td>
<td>38.86</td>
<td>37.50</td>
<td>36.40</td>
<td>35.43</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>62.58</td>
<td>57.55</td>
<td>53.46</td>
<td>50.40</td>
<td>46.22</td>
<td>43.71</td>
<td>41.55</td>
<td>41.27</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>62.19</td>
<td>58.81</td>
<td>55.65</td>
<td>52.45</td>
<td>47.47</td>
<td>47.36</td>
<td>44.35</td>
<td>41.50</td>
<td>39.07</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>42.63</td>
<td>48.69</td>
<td>50.85</td>
<td>47.41</td>
<td>43.84</td>
<td>43.66</td>
<td>38.07</td>
<td>35.13</td>
<td>23.67</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>20.68</td>
<td>25.56</td>
<td>31.81</td>
<td>33.63</td>
<td>28.44</td>
<td>29.05</td>
<td>23.62</td>
<td>15.96</td>
<td>9.57</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>8.96</td>
<td>11.31</td>
<td>14.65</td>
<td>14.97</td>
<td>12.60</td>
<td>12.93</td>
<td>8.96</td>
<td>7.05</td>
<td>3.94</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>3.12</td>
<td>4.43</td>
<td>5.62</td>
<td>5.84</td>
<td>5.05</td>
<td>5.13</td>
<td>3.30</td>
<td>2.44</td>
<td>1.31</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>0.87</td>
<td>1.47</td>
<td>1.72</td>
<td>2.01</td>
<td>1.58</td>
<td>1.61</td>
<td>1.01</td>
<td>0.79</td>
<td>0.36</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>0.19</td>
<td>0.39</td>
<td>0.41</td>
<td>0.54</td>
<td>0.41</td>
<td>0.41</td>
<td>0.23</td>
<td>0.21</td>
<td>0.07</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.03</td>
<td>0.08</td>
<td>0.07</td>
<td>0.10</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>18.57</td>
<td>20.58</td>
<td>22.25</td>
<td>21.71</td>
<td>18.87</td>
<td>18.77</td>
<td>15.48</td>
<td>12.95</td>
<td>9.67</td>
</tr>
</tbody>
</table>

6.3 Benefit of Modeling Belief Evolution

In order to quantify the importance of modeling the belief evolution, we evaluate the optimal policy without learning, the optimal policy to dynamic program (3), when instead customers do form a belief about the product quality based on the historical sales and prices. We compare the expected revenue of this policy against the expected revenue of the optimal policy with learning, i.e., that characterized by Proposition 4. Table 6.3 records the improvement brought about by the optimal policy for different values of initial inventory and initial belief levels. Notice that for low initial inventory levels (one or two items) and strong belief that the product quality is bad, the improvements can be substantial. For medium inventory levels, the improvement decreases, and it is the highest when the market is the most confused (i.e., an initial belief around 0.5). For large initial inventory levels, the improvements are small. This table indicates that over a wide range of parameter values, ignoring strategic customer behavior may be expensive to a firm.

7 Conclusion and Discussion

In this paper, we have studied optimal pricing strategies of a monopolist that sells a finite quantity of a product during a finite horizon when the product quality is uncertain. Customers arrive sequentially and all have
private, imperfect information about the product quality. Customers complement their private information with information contained in the historical price and sales paths. As a result, not only does the optimal dynamic pricing policy trade off current and future expected profits, it also influences customer learning. The product quality can either be good or bad. The customer learning is captured in a second state variable (besides current inventory), the market belief, which is the probability that the product quality is high, based on all available (public) pricing and sales information. This extra dimension to the pricing policy may result in significantly different strategies than the case when no information can be derived from historical sales and price paths.

Our two main contributions are: (1) establishing and illustrating the structure of the optimal policy and (2) quantifying the benefit of the model relative to traditional price-based revenue management. We find that the optimal pricing strategy features a single threshold in inventory and two thresholds in the beliefs. For any given market belief, the firm can charge a high or a low price. Both prices increase as the market belief increases. The low price at a high market belief may be higher than the high price at a low market belief. High inventories make the firm price low and dispose of the inventory without changing the market belief. High uncertainty in the market belief and low inventories make the firm price high (relative to the market belief). In this case the monopolist experiments with its pricing policy (charges a high price) to probe the market belief, and the market belief is reviewed upward when a sale occurs, downward when no sale occurs. When the market belief is relatively strong (that the product quality is good or bad) the firm prices low and the market belief does not change anymore. With a strong belief that the product quality is good, the monopolist “freezes” the market belief, i.e., even though the price is low with respect to the market belief, it is relatively high. By charging that price, customers always buy, which creates high revenues for the monopolist and at the same time, prevents further learning. With a strong belief that the product quality is bad, the monopolist also charges a low price. That price prevents again the market from learning from no-sales and the price from sliding down further. We also find that a firm that does not take the market belief evolution into account can lose more than 5% of its potential profits for reasonable initial inventory levels.

Our model is different from those available in the revenue management literature, which does not model the evolution of belief on quality. Popescu and Wu (2006) do develop a model in which a reference price captures the historical price path but their model is grounded in prospect theory. Thus, they do not consider the information about the product quality contained in the historical price path. With our model, we complement this emerging stream of literature by considering rational agents that use Bayes’ rule to update the belief of the quality of a product based on the price and sales path, and find that optimal pricing policies can be significantly different when taking strategic customer behavior into account.

Differently from Bose, Orosel and Vesterlund (2002), we endogenize the economic cost of selling inventory, i.e., the opportunity cost, which is relevant in the limited inventory case. In other words, these authors model a make-to-order production process of a long lived product with constant marginal cost; instead we model a batch production process, i.e., with significant set-up costs, of a short lived product. This difference appears explicitly
in our interpretation of the net prices.

Our model is highly stylized. Assuming that customers are rational, enter the market one-by-one, and perfectly observe the historical price and sale paths may be strong assumptions, but allow us to obtain structural insights into the customer purchasing behavior and optimal price policy. As prices are public, their history are more easily observable than the history of sales. Yet, customers may learn about the latter via word-of-mouth communication. We believe that our model sets the stage for further research in which some of its assumptions are relaxed. Customers may not be able to observe perfectly the historical sale path. When during a period more than one customer can arrives, the aggregate sales at a given price will be relevant for those customers arriving in the future. It would be interesting to see how robust our insights are when customers only observe a noisy signal of the historical aggregate sales. A final interesting extension would be to consider different generations of customers with different private information about the product quality. It may be plausible to assume that more informed customers arrive in earlier periods than less informed customers.

8 References


M. A. Lariviere and E. L. Porteus, (1999), Stalking Information: Bayesian Inventory Management with Unobserved Lost Sales, Management Science, 45(3), 346 - 363.


9 Proofs [Available Upon Request From the Authors]