

Tracking the Market: Dynamic Pricing and Learning in a Changing Environment

Arnoud V. den Boer*

University of Twente, P.O. Box 217, 7500 AE Enschede

a.v.denboer@utwente.nl

July 2, 2013

Abstract

Dynamic pricing of commodities without knowing the exact relation between price and demand is a much-studied problem. Practically all existing studies assume that the parameters describing the market are constant during the selling period. This severely reduces their practical applicability, since, in reality, market characteristics may change all the time, without the firm always being aware of it. In the present paper we study dynamic pricing and learning in a changing market environment. We introduce a methodology that enables the price manager to hedge against changes in the market, and provide explicit upper bounds on the regret - a measure of the performance of the firm's pricing decisions. In addition, this methodology guides the selection of the optimal way to estimate the market process. We provide numerical examples from practically relevant situations to illustrate the methodology.

Keywords: dynamic pricing, learning, varying parameters

1 Introduction, Contributions, Literature

1.1 Introduction

Firms selling products or delivering services face the complex task of determining which selling price to charge to their customers. Generally, firms aim at choosing selling prices that maximize certain performance indicators, such as revenue, profit, or utilization rate. An intrinsic property of this decision problem is *lack of information*: the seller does not know how consumers respond to different selling prices, and thus does not know the optimal price. The problem of the firm is not merely about optimization, but also about learning the relation between price and market response.

The presence of digitally available and frequently updated sales data makes this problem essentially an *on-line* learning problem: after each sales occurrence, the firm can use the newly obtained sales

*Part of this research was done while the author was affiliated with Centrum Wiskunde & Informatica (CWI), Amsterdam, Eindhoven University of Technology, and University of Amsterdam.

data to update its knowledge (for example, via statistical estimation methods). If, in addition, selling prices can quickly be modified, without much costs or effort - as often is the case in web-based sales channels or in brick-and-mortar stores with digital price tags - the firm can immediately exploit its improved knowledge on consumer behavior by appropriately adapting the selling prices.

Optimal pricing policies for these type of problems have been researched extensively. Here we only list a sample of the recent OR/MS literature; for a more elaborate discussion, including relevant studies from the economics literature, we refer to den Boer and Zwart (2010).

Lobo and Boyd (2003), Carvalho and Puterman (2005a,b), Bertsimas and Perakis (2006), Besbes and Zeevi (2009), Broder and Rusmevichientong (2012), den Boer and Zwart (2010) and Harrison et al. (2011) are all studies that assume that the price-demand relation belongs to a parametric family, estimate the unknown parameters by classical estimation methods (such as linear regression or maximum likelihood estimation), and study optimal pricing policies. Similar approaches with Bayesian estimation methods, can be found in Lin (2006), Araman and Caldentey (2009), Farias and van Roy (2010) and Harrison et al. (2012). Robust or nonparametric approaches are taken by Kleinberg and Leighton (2003), Cope (2007), Lim and Shanthikumar (2007), Eren and Maglaras (2010) and Besbes and Zeevi (2009).

A main conclusion from this stream of literature is that, in general, firms should properly balance learning and instant optimization. That means that not always the price should be chosen that is optimal according to current parameter estimates, but some price variation should be induced to guarantee sufficient quality of future parameter estimates.

All these studies have the assumption in common that the relation between price and expected sales is stable during the time horizon under consideration: the unknown parameters that describe this relation do not change. This is a rather strong assumption, which makes these studies less applicable in practical situations. Markets are generally not stable, but may vary over time, without the seller immediately being aware of it (cf. Dolan and Jeuland (1981), Wildt and Winer (1983), and Section 2 of Elmaghraby and Keskinocak (2003)). These changes may have various causes: shifts in consumer tastes, competition (Wildt and Winer, 1983), appearance of technological innovations (Chen and Jain, 1992), market saturation and product diffusion effects related to the life cycle of a product (Bass, 1969, Dolan and Jeuland, 1981, Raman and Chatterjee, 1995), marketing and advertisement efforts (Horsky and Simon 1983), competitors entering or exiting the market, appearance of new sales channels, and many more.

Wildt and Winer (1983, page 365) argued already in 1983 that “constant-parameter models are not capable of adequately reflecting such changing market environments”. In fact, this issue has been known since longtime in the historical literature on statistical economics, as illustrated by the following quotation of Schultz (1925) on the law of demand:

“The validity of the theoretical law [of demand] is limited to a point in time. But in order to derive *concrete, statistical* laws our observations must be numerous; and in order to obtain the requisite number of observations, data covering a considerable period must be used. During the interval, however, important dynamic changes take place in the condition of the market. In the case of a commodity like sugar, the principal

dynamic changes that need be considered are the changes in our sugar-consuming habits, fluctuations in the purchasing power of money, and the increase of population.” (Schultz, 1925, page 409).

Although the literature on dynamic pricing and learning has increased rapidly in recent years, models with a varying market have hardly been considered. This motivates the current study of dynamic pricing and learning in a changing environment.

1.2 Contributions

In the present paper we study the problem of dynamic pricing and learning in a changing environment. We study the situation where a monopolist firm is selling a single type of product with unlimited inventory. We consider an additive demand model, where the expected demand for the product in a certain time period is the sum of a stochastic market process and a known function depending on the selling price. The characteristics of this stochastic process are unknown to the firm. Its value at a certain point in time may be estimated from accumulated sales data; however, since the market may be changing over time, estimation methods are needed that are designed for time-varying systems. We deploy two such estimators, namely estimation with a forgetting factor, and estimation based on a “sliding window” approach. For both estimators we derive an upper bound on the expected estimation error.

Next, we propose a simple, intuitive pricing policy: at each decision moment, the firm estimates the market process with one of the just mentioned estimators, and subsequently sets the next selling price equal to the price that would be optimal if the firm’s market estimate were correct. This is a so-called myopic or certainty equivalent policy: at each decision moment the firm acts as if being certain about its estimates. To measure the quality of this pricing policy, we define $\text{AverageRegret}(T)$, which measures the expected costs of not choosing optimal prices in the first T periods, and $\text{LongRunAverageRegret}$, which equals the limit superior of $\text{AverageRegret}(T)$ as T grows large. We derive upper bounds on $\text{AverageRegret}(T)$ and $\text{LongRunAverageRegret}$. These bounds are not only stated in terms of the variables associated with the used estimation method (the forgetting factor, or the size of the sliding window), but also in terms of a measure of the *impact* that market fluctuations have on the estimation error. Clearly, if the market is very unstable and inhibits very large and frequent fluctuations, the impact may become extremely large, which negatively affects the obtained revenue.

The novel, key idea of this study is that (i) this impact can be bounded, using assumptions on the market process that the firm makes a priori, (ii) the resulting upper bounds on $\text{AverageRegret}(T)$ and $\text{LongRunAverageRegret}$ can be used by the firm to determine the optimal estimator of the market (i.e. the optimal value of the forgetting factor or window size), (iii) this provides the firm explicit guarantees on the maximum expected revenue loss. This framework enables the firm to *hedge against change*: the firm is certain that the expected regret does not exceed a certain known value, provided the market process satisfies the posed assumptions. These assumptions may be very general and cover many important cases; for example, bounds on the probability that the market value changes in a certain period, bounds on the maximum difference between two

consecutive market values, or bounds on the maximum and minimum value that the market process may attain. We provide numerical examples to illustrate the methodology, in two practically relevant settings: in the first we make use of the well-known Bass model to model the diffusion of an innovative products; and in the second we consider an oligopoly where price changes by competitors causes occasional changes in the market. The application of our methodology on the Bass model makes this the first study that incorporates learning and pricing in this widely used product-diffusion model; thus far, only deterministic settings (Robinson and Lakhani, 1975, Dolan and Jeuland, 1981, Kalish, 1983), or random settings where no learning is present (Chen and Jain, 1992, Raman and Chatterjee, 1995, Kamrad et al., 2005) have been considered in the literature.

Summarizing, in one of the first studies on dynamic pricing and learning in a changing environment, our contributions are as follows.

- (i) We introduce a model of dynamic pricing and learning in a changing market environment, using a very generic description of the market process.
- (ii) We discuss two estimators of time-varying processes, and prove upper bounds on the estimation error.
- (iii) We propose a methodology that enables the decision maker to hedge against change. This results in explicit upper bounds on the regret, and guides the choice of the optimal estimator.
- (iv) We show the application of the methodology in several concrete cases, and offer numerical examples to illustrate its use and performance. These examples show that incorporating the changing nature of the market process can significantly improve a firm's revenue.

1.3 Comparison to Relevant Literature

The combination of dynamic pricing and learning in a changing market is a rather unexplored area. Besbes and Zeevi (2011) study a pricing problem where the willingness-to-pay (WtP) distribution of the customers changes at some unknown point in time. The WtP distribution before and after the change are assumed to be known, only the time of change is unknown to the seller. Lower bounds on the worst-case regret are derived, and pricing strategies are developed that achieve the order of these bounds. Chen and Jain (1992) consider optimal pricing policies in models where the demand not only depends on the selling price, but also on the cumulative amount of sales; in this way diffusion effects are modeled. In addition, the demand is influenced by an observable state variable, which models unpredictable events that change the demand function, and whose dynamics are driven by a Poisson process. Apart from these random events, the demand is fully deterministic and known to the firm, and learning by the firm is not considered. Hanssens et al. (2001) and Leeflang et al. (2009, Section 2.3) discuss several dynamic market models, as well as estimation methods, but do not integrate this with the problem of optimal dynamic pricing. A recent study that is closely related to our work is Besbes and Saure (2012). They consider dynamic pricing with finite selling season and finite inventory that cannot be replenished. The demand function is unknown and subject to abrupt changes. The authors focus on the trade-off between gaining revenue before and after the change-point, and derive in various settings structural

properties of the optimal price policy. A relevant study from the control literature is from Godoy et al. (2009). They consider an estimation problem in a linear system, where the parameters are subject to shock changes, and analyze the performance of a sliding-window linear regression method. A major assumption is that the controls are deterministic. This differs from pricing problems, where the prices (the controls) usually depend in a non-trivial way on all previously observed sales realizations. We also refer to recent work by Garivier and Moulines (2011) on multi-armed bandit problems with time-varying parameters. Two differences between their and our work are (i) they consider a discrete action set, whereas, in our setting, prices can be chosen from a continuum, and (ii) they restrict themselves to abruptly changing environments, whereas our analysis is more generic, including slowly changing environments.

1.4 Organization of the Paper

The rest of this paper is organized as follows. Section 2 introduces the model, discusses estimation methods for the market process, gives bounds on the estimation error, and provides a discussion on various model assumptions. Section 3 introduces the methodology for hedging against change: we formulate the pricing myopic policy and provide performance bounds in Section 3.1, we show in Section 3.2 how assumptions on the market process can be used to find the optimal estimator that minimizes these regret bounds, and provide in Section 3.3 three examples of the methodology. The results of two numerical studies are described in 4, and conclusions and directions for future research are discussed in Section 5. All mathematical proofs are contained in Section 6.

2 Model Primitives

2.1 Model Description

We consider a monopolist firm selling a single type of product. In each time period $t \in \mathbb{N}$, the firm decides on a selling price $p_t \in [p_l, p_h]$, where $0 \leq p_l < p_h < \infty$ denote the lowest and highest admissible price. After choosing the price, the seller observes demand d_t , which is a realization of the random variable $D_t(p_t)$. Conditional on the selling prices, the demand in different time periods is independent. The expected demand in period t , against a price p , is of the form

$$E[D_t(p)] = M(t) + g_t(p). \quad (1)$$

Here $(M(t))_{t \in \mathbb{N}}$ is a stochastic process called the *market process*, unobservable for the firm, and taking values in a (possibly infinite) interval $\mathcal{M} \subset \mathbb{R}$. Let \mathcal{F}_t be the σ -algebra generated by $d_1, p_1, M(1), \dots, d_t, p_t, M(t)$, \mathcal{F}_0 the trivial σ -algebra, and write $\epsilon_t = d_t - g_t(p_t) - M(t)$; then we assume that $M(t)$ and ϵ_t are \mathcal{F}_{t-1} -measurable, for all $t \in \mathbb{N}$. In addition we impose the following mild conditions on the moments of $M(t)$ and ϵ_t : there are positive constants σ_M and σ , such that

$$\sup_{t \in \mathbb{N}} E[M(t)^2 | \mathcal{F}_{t-1}] \leq \sigma_M^2 \text{ a.s.} \quad \text{and} \quad \sup_{t \in \mathbb{N}} E[\epsilon_t^2 | \mathcal{F}_{t-1}] \leq \sigma^2 \text{ a.s.} \quad (2)$$

The functions g_t in (1) model the dependence of expected demand on selling price. They are assumed to be known by the seller. After observing demand, the seller collects revenue $p_t d_t$, and proceeds to the next period. The purpose of the seller is to maximize expected revenue.

Let $r_t(p, M) = p \cdot (M + g_t(p))$ denote the expected revenue in period $t \in \mathbb{N}$, when the market process equals M and the selling price is set at p . The price that generates the highest amount of expected revenue, given that the current market equals M , is denoted by $p_t^*(M) = \arg \max_{p \in [p_l, p_h]} r_t(p, M)$.

We impose some mild conditions to ensure that this optimal price exists and is uniquely defined. In particular, we assume that for all admissible prices p and all $t \in \mathbb{N}$, $g_t(p)$ is decreasing in p , and twice continuously differentiable w.r.t. p , with first and second derivative denoted by $g_t'(p)$ and $g_t''(p)$. These two properties immediately carry over to the expected demand, and in fact are quite natural conditions for demand functions to hold. In addition, we assume that for all $M \in \mathcal{M}$ and all $t \in \mathbb{N}$ the revenue function $r_t(p, M)$ is unimodal with unique optimum $p_t^\#(M) \in \mathbb{R}$ satisfying $r_t'(p_t^\#(M), M) = 0$, and in addition

$$\sup \{r_t''(p_t^\#(M), M) \mid t \in \mathbb{N}, M \in \mathcal{M}, p_t^\#(M) \in [p_l, p_h]\} < 0, \quad (3)$$

where $r_t'(p, M)$ and $r_t''(p, M)$ denote the first and second derivative of $r_t(p, M)$ w.r.t. p .

The value of the market process and the corresponding optimal price are unknown to the seller. As a result, the decision maker might choose sub-optimal prices, which incurs a loss of revenue relative to someone who would know the market process and the optimal price. The goal of the seller is to determine a pricing policy that minimizes this loss of revenue. With a pricing policy we here mean a sequence of (possibly random) prices $(p_t)_{t \in \mathbb{N}}$ in $[p_l, p_h]$, where each price p_t may depend on all previously chosen prices p_1, \dots, p_{t-1} and demand realizations d_1, \dots, d_{t-1} .

To assess the quality of a pricing policy Φ , we define the following two quantities.

$$\text{AverageRegret}(\Phi, T) = \frac{1}{T-1} \sum_{t=2}^T E [r_t(p_t^*(M(t)), M(t)) - r_t(p_t, M(t))], \quad (4)$$

$$\text{LongRunAverageRegret}(\Phi) = \limsup_{T \rightarrow \infty} \text{AverageRegret}(\Phi, T). \quad (5)$$

Each term in the summand of (4) measures the expected revenue loss caused by not using the optimal price in period t . The expectation operator is because both p_t and $M(t)$ may be random variables. We start measuring the average regret from the second period. This simplifies several expressions that appear in further sections; in addition, in the first period, no data is available to estimate $M(1)$, and minimizing the instantaneous regret encountered in the first period is not possible. Furthermore, note that $\text{AverageRegret}(\Phi, T)$ and $\text{LongRunAverageRegret}(\Phi)$ are not observed by the seller, and thus can not directly be used to determine an optimal pricing policy.

2.2 Estimation of Market Process

Estimating the value of the market process gives vital information that is needed to determine the selling price. Since the market may change over time, the firm needs an estimation method that can handle such changes. In this section we describe two such methods: (I) estimation with forgetting factor, and (II) estimation with a sliding window.

(I) Estimation of $M(t)$ with forgetting factor. Let $\lambda \in [0, 1]$ be the forgetting factor, to be determined by the decision maker. The estimate $\hat{M}_\lambda(t)$, with forgetting factor λ , based on demand realizations d_1, \dots, d_t and prices p_1, \dots, p_t , is equal to

$$\hat{M}_\lambda(t) = \arg \min_{M \in \mathbb{R}} \sum_{i=1}^t (d_i - M - g_i(p_i))^2 \lambda^{t-i}. \quad (6)$$

The factor λ^{t-i} acts as a weight on the data $(p_i, d_i)_{1 \leq i \leq t}$. Data that lies further in the past gets a lower weight; data from the recent past receives more weight (unless $\lambda = 1$, in which case all available data gets equal weight, or $\lambda = 0$, in which case only the most recent observation is taken into account). This captures the idea that the longer ago data has been generated, the likelier it is that the corresponding value of the market process differs from its current value. Accordingly, data from longer ago is assigned a smaller weight than data from the more recent past. Whether this intuition is true depends of course on the specific characteristics of $M(t)$.

By differentiating the righthandside of (6) w.r.t. M , we obtain the following explicit expression for $\hat{M}_\lambda(t)$:

$$\hat{M}_\lambda(t) = \frac{\sum_{i=1}^t (d_i - g_i(p_i)) \lambda^{t-i}}{\sum_{i=1}^t \lambda^{t-i}}. \quad (7)$$

(II) Estimation of $M(t)$ with a sliding window. Let $N \in \mathbb{N}_{\geq 2} \cup \{\infty\}$ be the window size, determined by the decision maker. The estimate $\hat{M}_N(t)$, with sliding window size N , based on demand realizations d_1, \dots, d_t and prices p_1, \dots, p_t , is equal to

$$\hat{M}_N(t) = \arg \min_{M \in \mathbb{R}} \sum_{i=\max\{t-N+1, 1\}}^t (d_i - M - g_i(p_i))^2. \quad (8)$$

Here only data from the N most recent observations is used to form an estimate. All data that is generated longer than N time periods ago, is neglected (if $N = \infty$, then all available data is taken into account). Similar to the estimate with forgetting factor, the rationale behind the estimate $\hat{M}_N(t)$ is the idea that for data generated long ago, it is more likely that the corresponding market value differs from its current value. This is captured in the fact that only the N most recent observations are used to estimate $M(t)$. Whether this idea is correct depends again on the specifics of $M(t)$.

Differentiating the righthandside of (8) w.r.t. M , we obtain the following expression:

$$\hat{M}_N(t) = \frac{1}{\min\{N, t\}} \sum_{i=\max\{t-N+1, 1\}}^t (d_i - g_i(p_i)). \quad (9)$$

Remark 1. Both estimation methods (I) and (II) depend on a decision variable (λ resp. N) that can be interpreted as a measure for the responsiveness to changes in the market. A high value of λ resp. N means that much information from the historical data is used to form estimates; this is advantageous in case of a stable market, but disadvantageous in case of many or large recent changes in the market process. Similarly, a low value of λ resp. N implies that the estimate of $M(t)$ is mainly determined by recent data; naturally, this is more beneficial in a volatile market than in a stable market.

2.3 Impact Measure and Quality of Market Estimates

Market fluctuations influence the accuracy of the estimates $\hat{M}_\lambda(t)$ and $\hat{M}_N(t)$. The following quantities $I_\lambda(t)$ and $I_N(t)$ measure this *impact* of market variations on the estimates. Observe that this impact is not solely determined by the market process, but also by the choice of λ and N :

$$I_\lambda(t) = E \left[\left| \left(\frac{1-\lambda}{1-\lambda^t} \mathbf{1}(\lambda < 1) + \frac{1}{t} \mathbf{1}(\lambda = 1) \right) \sum_{i=1}^t (M(i) - M(t+1)) \lambda^{t-i} \right|^2 \right],$$

$$I_N(t) = E \left[\left| \frac{1}{\min\{N, t\}} \sum_{i=1+(t-N)^+}^t (M(i) - M(t+1)) \right|^2 \right].$$

The following proposition gives a bound on the expected estimation error of (I) and (II), in terms of λ , N , and the impact measures $I_\lambda(t)$ and $I_N(t)$.

Proposition 1. For all $t \in \mathbb{N}$,

$$E \left[\left| \hat{M}_\lambda(t) - M(t+1) \right|^2 \right] \leq 2\sigma^2 \left[\frac{(1-\lambda)(1+\lambda^t)}{(1+\lambda)(1-\lambda^t)} \mathbf{1}(\lambda < 1) + \frac{1}{t} \mathbf{1}(\lambda = 1) \right] + 2I_\lambda(t) \quad (10)$$

and

$$E \left[\left| \hat{M}_N(t) - M(t+1) \right|^2 \right] \leq 2 \frac{\sigma^2}{\min\{N, t\}} + 2I_N(t). \quad (11)$$

If the processes $(\epsilon_t)_{t \in \mathbb{N}}$ and $(M(t))_{t \in \mathbb{N}}$ are independent, then

$$E \left[\left| \hat{M}_\lambda(t) - M(t+1) \right|^2 \right] \leq \sigma^2 \left[\frac{(1-\lambda)(1+\lambda^t)}{(1+\lambda)(1-\lambda^t)} \mathbf{1}(\lambda < 1) + \frac{1}{t} \mathbf{1}(\lambda = 1) \right] + I_\lambda(t) \quad (12)$$

and

$$E \left[\left| \hat{M}_N(t) - M(t+1) \right|^2 \right] \leq \frac{\sigma^2}{\min\{N, t\}} + I_N(t), \quad (13)$$

with equality in (12), (13) if the disturbance terms are homoscedastic, i.e. $E[\epsilon_t^2 | \mathcal{F}_{t-1}] = \sigma^2$ for all $t \in \mathbb{N}$.

Remark 2. The first terms of the righthandsides of (10) - (13) are related to the natural fluctuations in demand. The lower these fluctuations, measured by σ^2 , the lower this part of the estimation error becomes. The second terms of the righthandsides of (10) - (13) relate to the impact that market fluctuations have on the quality of the estimate of $M(t)$. These terms are nonnegative, and equal zero if the market value does never change.

2.4 Discussion of Model Assumptions

Our demand model is of the additive form $E[D_t(p)] = M(t) + g_t(p)$, where $M(t)$ is unknown and $g_t(p)$ is known. The term $M(t)$ can be regarded as capturing various time-varying aspects of the true demand model, with possibly complex behavior that is not fully known or understood by the decision maker. If we would only assume that $M(t)$ lies in some known uncertainty \mathcal{M} , then a typical approach would be to optimize the price given the worst-case value of $M(t)$ in \mathcal{M} . A disadvantage of this robust optimization approach is that the accumulating observations of $(M(t))_{t \in \mathbb{N}}$ are not used by the firm to improve its price decisions. Our work distinguishes itself from the 'static' robust optimization approach by allowing some way learning or tracking the market process.

An alternative way of viewing the demand model is to regard $g_t(p)$ as the firm's local approximation of a more complex demand model, and $M(t)$ as the time-dependent deviation between this approximation and the true demand. In this way $M(t)$ may capture (unavoidable) model errors made by the firm.

Note that instead of an additive model one could also assume a multiplicative demand model, where the expected demand is the product of the two parts: $E[D_t(p)] = M(t) \cdot g_t(p)$. An advantage of a multiplicative model is that, under some additional assumptions, the aggregate demand in a time period may be explained in terms of the buying behavior of individual customers. For example, one could assume that individual customers have a willingness-to-pay (WtP) distribution $F(p)$: if the selling price equals p , a randomly selected customer buys a product with probability $1 - F(p)$. If there are $M(t)$ customers present, and their buying-decisions are mutually independent, then the expected aggregated demand $E[D(p)]$ has the multiplicative form $M(t) \cdot (1 - F(p))$. Such demand model can thus be explained in terms of the behavior of individual customers, but only using the strong assumptions that the customers behave independently and buy only a single product. In our setting it would be inappropriate to pose such strong assumptions on consumer behavior: we study how a seller can handle a volatile, unstable market while making only minor assumptions on its behavior.

Another motivation for the additive demand model is related to the optimal price. By differenti-

ating the revenue function w.r.t. p , one can easily show that the optimal price in a multiplicative demand model is the solution to the equation $pg'_t(p)/g_t(p) = -1$. This equation is independent of the market process, and as a result, the firm does not need to know or estimate the market in order to determine the optimal selling price. Intuitively it is clear that for many products such a model does not accurately reflect reality.

An important subclass of our demand model is the setting where $g_t(p) = g(p)$, i.e. the expected demand is the sum of a time-dependent part and a price-dependent time-homogeneous part. Such demand models appear frequently in the literature: for example, in models that incorporate competition (Puu, 1991, Tuinstra, 2004, Cooper et al., 2012), or models that capture market diffusion and saturation effects (Chen and Jain (1992, Section IV), Raman and Chatterjee (1995, Section 4.3), Kalish (1983, Section 3.3.1)). Some of the numerical examples in Section 4 apply our pricing policy to these two settings.

The assumptions on g_t and r_t are fairly standard conditions on demand and revenue functions, and ensure that the revenue function is locally strictly concave around the optimum. Clearly, if $p_t^\#(M)$ lies in the interval $[p_l, p_h]$ then $p_t^*(M) = p_t^\#(M)$, and if $p_t^\#(M) \notin [p_l, p_h]$, then $p_t^*(M)$ is the projection of $p_t^\#(M)$ on the interval $[p_l, p_h]$. It is not difficult to show that the conditions on g_t are satisfied for the linear demand model with $g_t(p) = -bp$ for some $b > 0$. For nonlinear demand functions with $g_t(p) = -bp^c$ for some $b > 0, c > 0, c \neq 1$, or $g_t(p) = -b \log(p)$ for some $b > 0$, the conditions are satisfied if the market process is bounded.

3 Hedging against Changes in the Market

In this section we show how a price manager can hedge against changes in the market. The key idea is that a simply myopic policy can be used (which means that one always chooses the price that is optimal according to current estimates of the market process), but that the parameter λ of the market estimator $\hat{M}_\lambda(t)$ (or N , for the estimator $\hat{M}_N(t)$) is chosen in a smart way.

As already alluded to in Section 2.2, the optimal value of λ or N depends on the nature of changes in the market process. If changes are frequent and/or large, λ and N should be chosen small, whereas in case of infrequent and small changes in the market one intuitively expects that λ should be chosen close to one, and N large.

Thus, in order to find a good choice of λ resp. N , the firm needs assumptions on the type of changes in the market that it is anticipating. Such assumptions can be translated into bounds on the behavior of the influence measures $I_\lambda(t)$, $I_N(t)$, which in turn lead to bounds on the regret of the myopic policy. These regret bounds depend on λ or N , and minimizing them leads to the optimal value of λ or N w.r.t. the assumptions on the market imposed by the firm.

The following two subsections elaborate this approach. Section 3.1 formulates the myopic policy and studies how the regret depends on the influence measures $I_\lambda(t)$, $I_N(t)$. Section 3.2 explains the methodology in more detail, and Section 3.3 provides three illustrative examples.

3.1 Performance Bounds for Myopic Policy

We consider the following simple, myopic pricing policy: at each decision moment the seller estimates the market value with one of the two estimation methods described in Section 2.2, and subsequently chooses the selling price that is optimal w.r.t. this estimate. In other words, the seller always acts as if the current estimate of the market is correct.

We denote this policy by Φ_λ if the market is estimated by method (I), with forgetting factor λ , and by Φ_N if the market is estimated by method (II), with sliding window of size N . The formal description of Φ_λ and Φ_N is as follows.

Myopic pricing policy Φ_λ / Φ_N

Initialization: Choose $\lambda \in [0, 1]$ or $N \in \mathbb{N}_{\geq 2} \cup \{\infty\}$.

Set $p_1 \in [p_l, p_h]$ arbitrarily.

For all $t \in \mathbb{N}$:

Estimation: Let $\hat{M}(\cdot)$ denote either $\hat{M}_\lambda(\cdot)$ (for policy Φ_λ) or $\hat{M}_N(\cdot)$ (for policy Φ_N).

Pricing: Set $p_{t+1} = p_{t+1}^*(\hat{M}(\cdot))$.

The following theorem provides upper bounds on the (long run) average regret for the myopic pricing policies, in terms of the influence measures $I_\lambda(t)$ and $I_N(t)$.

Theorem 1. *There is a $K_0 > 0$ such that for all $T \geq 2$,*

$$\begin{aligned} \text{AverageRegret}(\Phi_\lambda, T) &\leq 2K_0\sigma^2 \left[\frac{1-\lambda}{1+\lambda} + \frac{2}{T-1} \left(\frac{\lambda \log(\lambda) + (1-\lambda) \log(1-\lambda)}{(1+\lambda) \log(\lambda)} \right) \right] \mathbf{1}(\lambda < 1) \\ &\quad + 2K_0\sigma^2 \left[\frac{1 + \log(T-1)}{T-1} \right] \mathbf{1}(\lambda = 1) \\ &\quad + 2K_0 \frac{1}{T-1} \sum_{t=1}^{T-1} I_\lambda(t), \end{aligned}$$

and

$$\text{AverageRegret}(\Phi_N, T) \leq 2K_0\sigma^2 \left[\frac{\log(\min\{T-1, N\})}{T-1} + \frac{1}{\min\{N, T-1\}} \right] + \frac{2K_0}{T-1} \sum_{t=1}^{T-1} I_N(t).$$

Consequently,

$$\text{LongRunAverageRegret}(\Phi_\lambda) \leq 2K_0 \left[\sigma^2 \frac{1-\lambda}{1+\lambda} + \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T I_\lambda(t) \right], \quad (14)$$

for all $\lambda \in [0, 1]$, and

$$\text{LongRunAverageRegret}(\Phi_N) \leq 2K_0 \left[\sigma^2 \frac{1}{N} + \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T I_N(t) \right], \quad (15)$$

for all $N \in \mathbb{N}_{\geq 2} \cup \{\infty\}$, where we write $1/\infty = 0$.

The main idea of the proof is to show that there is a $K_0 > 0$ such that for any M and M' , the instantaneous regret in period t satisfies $r_t(p_t^*(M), M) - r_t(p_t^*(M'), M) \leq K_0(M - M')^2$. Subsequently we apply the bounds derived in Proposition 1.

Remark 3. By (12) and (13), if the processes $(\epsilon_t)_{t \in \mathbb{N}}$ and $(M(t))_{t \in \mathbb{N}}$ are independent, then all four inequalities of Theorem 1 are still valid if all righthandsides are divided by 2.

Remark 4. An explicit expression for K_0 is derived in the proof of Theorem 1. To obtain the most sharp bounds, one could also define K_0 directly as

$K_0 = \sup_{t \in \mathbb{N}} \inf_{M \neq M'} (r_t(p_t^*(M), M) - r_t(p_t^*(M'), M)) / (M - M')^2$. For the important special case of a stationary linear demand function, with $g_t(p) = g(p) = -bp$ for some $b > 0$ and $M(t) > 0$ for all $t \in \mathbb{N}$, it is not difficult to show $p_t^*(M) = \min\{\max\{M/(2b), p_l\}, p_h\}$ and $K_0 = 1/(4b)$.

Remark 5. In dynamic pricing and learning studies that assume a stable market, one often considers the asymptotic behavior of $\text{Regret}(\Phi, T) = (T - 1) \cdot \text{AverageRegret}(\Phi, T)$, where Φ denotes the pricing policy that is used. Typically one proves bounds on the growth rate of $\text{Regret}(\Phi, T)$ for a certain policy, e.g. $\text{Regret}(\Phi, T) = O(\sqrt{T})$ or $\text{Regret}(\Phi, T) = O(\log(T))$. A policy is considered 'good' if the speed of convergence of the regret is close the best achievable rate, cf. Broder and Rusmevichientong (2012), Harrison et al. (2011) and den Boer and Zwart (2010). In the setting with a changing market, a simple example makes clear that one cannot do better than $\text{Regret}(\Phi, T) = O(T)$ or $\text{AverageRegret}(\Phi, T) = O(1)$. Suppose $M(t)$ is a Markov process taking values in $\{M_1, M_2\} \in \mathbb{R}_+^2$, with $M_1 \neq M_2$, and suppose $P(M(t+1) = M_i | M(t) = M_j) = \frac{1}{2}$, for all $i, j \in \{1, 2\}$ and $t \in \mathbb{N}$. Let $g_t(p) = g(p) = -bp$ for some $b > 0$ and all $t \in \mathbb{N}$, and choose $[p_l, p_h]$ such that $p_t^\#(M_i) = M_i/(2b) \in (p_l, p_h)$, for $i = 1, 2$. Then for all $t \in \mathbb{N}$, the instantaneous regret incurred in period t satisfies

$$\begin{aligned} & E[r_t(p_t^*(M(t)), M(t)) - r_t(p_t, M(t))] \\ & \geq \inf_{p \in [p_l, p_h]} \left[\frac{1}{2}(r_t(p_t^*(M_1), M_1) - r_t(p, M_1)) + \frac{1}{2}(r_t(p_t^*(M_2), M_2) - r_t(p, M_2)) \right] \\ & \geq \frac{b}{2} \inf_{p \in [p_l, p_h]} \left[(p^*(M_1) - p)^2 + (p^*(M_2) - p)^2 \right] \\ & \geq \frac{b}{4} (p_t^*(M_1) - p_t^*(M_2))^2 \\ & \geq \frac{1}{16b} (M_1 - M_2)^2 > 0, \end{aligned}$$

which implies that no policy can achieve a sub-linear $\text{Regret}(\Phi, T) = o(T)$. In fact, any pricing policy achieves the optimal growth rate $\text{Regret}(\Phi, T) = O(T)$. Thus, the challenge of dynamic pricing and learning in such a changing environment is not to find a policy with optimal asymptotic growth rate, but rather to make the (long run) average regret as small as possible.

In view of the remark above, the question raises whether the bounds from Theorem 1 are sharp. The following proposition answers this question for the case of a linear stationary demand function with homoscedastic disturbance terms independent of the market process.

Proposition 2. Suppose $g_t(p) = g(p) = -bp$ for some $b > 0$ and all $t \in \mathbb{N}$, $E[\epsilon_t^2 | \mathcal{F}_{t-1}] = \sigma^2$ for

all $t \in \mathbb{N}$, the processes $(\epsilon_t)_{t \in \mathbb{N}}$ and $(M(t))_{t \in \mathbb{N}}$ are independent, and $M(t) \in [2bp_l, 2bp_h]$ a.s. for all $t \in \mathbb{N}$. Then, with $K_0 = 1/(4b)$,

$$\text{LongRunAverageRegret}(\Phi_\lambda) = K_0 \left[\sigma^2 \frac{1-\lambda}{1+\lambda} + \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T I_\lambda(t) \right], \quad (16)$$

for all $\lambda \in [0, 1]$, and

$$\text{LongRunAverageRegret}(\Phi_N) = K_0 \left[\sigma^2 \frac{1}{N} + \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T I_N(t) \right], \quad (17)$$

for all $N \in \mathbb{N}_{\geq 2} \cup \{\infty\}$, where we write $1/\infty = 0$.

3.2 Methodology for Hedging Against Changes

The bounds on the regret that we derive in Theorem 1 are stated in terms of the influence measures $I_\lambda(t)$ and $I_N(t)$. That means that the seller can get an explicit upper bound on the regret in terms of λ , N , if it can find upper bounds on the influence measures in terms of λ , N ; subsequently, an optimal choice of λ , N can be found by minimizing these upper bounds on the regret.

More precisely, the firm should translate its assumptions on the market process into (non-random) upper bounds on the terms $\frac{1}{T-1} \sum_{t=1}^{T-1} I_\lambda(t)$ and $\frac{1}{T-1} \sum_{t=1}^{T-1} I_N(t)$. By plugging these bounds into Theorem 1, it obtains bounds on $\text{AverageRegret}(\Phi_\lambda, T)$ and $\text{AverageRegret}(\Phi_N, T)$ in terms of λ and N . The optimal choices of λ and N are then determined by simply minimizing these bounds with respect to λ and N . In some cases an explicit expression for the optimal choice may exist, otherwise numerical methods are needed to determine the optimum.

The resulting optimal λ and N may depend on the length of the time horizon T . This may be undesirable to the firm, for instance because T is not known in advance, or because the time horizon is infinite. In this case it is more appropriate to minimize the $\text{LongRunAverageRegret}$. If the firm can translate its assumptions on the market process into upper bounds on the terms $\limsup_{T \rightarrow \infty} \frac{1}{T-1} \sum_{t=1}^{T-1} I_\lambda(t)$ and $\limsup_{T \rightarrow \infty} \frac{1}{T-1} \sum_{t=1}^{T-1} I_N(t)$, then these upper bounds can be plugged into (14) and (15), and the optimal λ and N can be determined by minimizing the resulting expression.

Remark 6. Observe that the optimal choices of λ and N are *independent* of the functions g_t . The relevant properties of g_t are captured by the constant K_0 , but its value does not influence the optimal λ and N . In a way this separates optimal estimation and optimal pricing: the first is determined by the impact of the market process, while only the latter involves the functions g_t . On the other hand, the variance of the demand distribution, related to σ^2 , does influence the optimal λ and N . In addition, note that by Remark 3, the factor 2 on the righthandsides of (14) and (15) can be removed if the processes $(\epsilon_t)_{t \in \mathbb{N}}$ and $(M(t))_{t \in \mathbb{N}}$ are independent. In practice, it may not always be known to the decision maker whether this condition is satisfied; but, fortunately, this does not influence the optimal choice of λ and N .

Remark 7. The above presented methodology of hedging against change has some similarities with robust optimization. There, one usually considers optimization problems whose optimal solutions depend on some parameters. These parameters are not known exactly by the decision maker, but assumed to lie in a certain “uncertainty set” which is known in advance. The optimal decision is then determined by optimizing against the worst case of the possible parameter values. An improvement of our methodology compared to robust optimization is that we allow for many different types of assumptions on the market process, as illustrated by the three examples described in Section 3.3. In contrast, robust optimization generally only assumes a setting of an uncertainty set. In addition, in robust optimization there is usually no learning of the unknown parameters, whereas our methodology allows using accumulating data to estimate the unknown process; in several instances this enables us to “track” the market process.

3.3 Examples

To illustrate the methodology, we look in more detail to three examples of assumptions on the market process: (i) bounds on the range of the market process, (ii) bounds on the maximum jump of the market process, and (iii) bounds on the probability that the market changes.

3.3.1 Bounds on the range of the market process.

In this section we consider the assumption that the market process is contained in a bounded interval.

Proposition 3. *If $\sup_{t \in \mathbb{N}} M(t) - \inf_{t \in \mathbb{N}} M(t) \leq d$ a.s., for some $d > 0$, then*

$$\text{LongRunAverageRegret}(\Phi_\lambda) \leq 2K_0 \left[\sigma^2 \frac{1-\lambda}{1+\lambda} + d^2 \right], \quad (18)$$

$$\text{LongRunAverageRegret}(\Phi_N) \leq 2K_0 \left[\sigma^2 \frac{1}{N} + d^2 \right], \quad (19)$$

for all $\lambda \in [0, 1]$, $N \in \mathbb{N}_{\geq 2} \cup \{\infty\}$, where we write $1/\infty = 0$.

The righthandsides of (18) and (19) are minimized by taking $\lambda = 1$ and $N = \infty$.

At first sight it may seem somewhat surprising that it is beneficial to take into account all available sales data to estimate the market, including ‘very old’ data. This can be explained by noting that in a period $t + 1$, all preceding values of the market $M(1), \dots, M(t)$ may differ by d from the current value $M(t + 1)$. In such a volatile market situation, it is best to ‘accept’ an unavoidable error caused by market fluctuations, and instead focus on minimizing the estimation error caused by natural fluctuations $\epsilon_1, \dots, \epsilon_t$ in the demand distribution. This is best done when all available data is taken into account; hence the optimality of choosing $\lambda = 1$ and $N = \infty$.

3.3.2 Bounds on one-step market changes.

In this section we consider the assumption that the one-step changes of the market process are bounded.

Proposition 4. *If $\sup_{t \in \mathbb{N}} |M(t) - M(t+1)| \leq d$ a.s., for some $d > 0$, then*

$$\text{LongRunAverageRegret}(\Phi_\lambda) \leq 2K_0 \left[\sigma^2 \frac{1-\lambda}{1+\lambda} + d^2 \frac{1}{(1-\lambda)^2} \right], \quad (20)$$

$$\text{LongRunAverageRegret}(\Phi_N) \leq 2K_0 \left[\sigma^2 \frac{1}{N} + \frac{1}{4} d^2 (N+1)^2 \right], \quad (21)$$

for all $\lambda \in [0, 1]$, $N \in \mathbb{N}_{\geq 2} \cup \{\infty\}$, where we write $1/\infty = 0$.

Consider the upper bound (20). The derivative of $\sigma^2 \frac{(1-\lambda)}{(1+\lambda)} + d^2(1-\lambda)^{-2}$ w.r.t. $\lambda \in (0, 1)$ is zero if and only if $(\sigma/d)^2(1-\lambda)^3 = (1+\lambda)^2$. Since $(1-\lambda)^3$ is decreasing and $(1+\lambda)^2$ is increasing in λ , we have the following possibilities:

1. $(\sigma/d)^2 \leq 1$. Then $\sigma^2 \frac{(1-\lambda)}{(1+\lambda)} + d^2(1-\lambda)^{-2}$ is increasing on $\lambda \in (0, 1)$, and the righthandside of (20) is minimized by taking $\lambda = 0$.
2. $(\sigma/d)^2 > 1$. Then there is a unique $\lambda^* \in (0, 1)$ that minimizes $\sigma^2 \frac{(1-\lambda)}{(1+\lambda)} + d^2(1-\lambda)^{-2}$. Although an explicit expression exists for λ^* , it is rather complicated, and it is not informative to state it here. The value of λ^* can be computed by solving a cubic equation.

Now consider the upper bound (21). The expression $\frac{\sigma^2}{N} + \frac{1}{4}d^2(N+1)^2$ on the righthandside of (21) is minimized by choosing N as the solution to $N^2(N+1) = 2(\sigma/d)^2$, which follows by taking the derivative w.r.t. N and some basic algebraic manipulations. It can easily be shown that there is a unique solution $N^* > 0$, at which the minimum is attained, and that $\frac{\sigma^2}{N} + c(N)$ is minimized by choosing N equal to either $\lfloor N^* \rfloor$ or $\lceil N^* \rceil$. If $(\sigma/d)^2 \leq 10/4$ then the optimal N equals 1, if $(\sigma/d)^2 > 10/4$ then the optimal N is strictly larger than 1. Figure 1 shows the relation between $(\sigma/d)^2$ and the values of λ^* , N^* that minimize the righthandside of (20), (21).

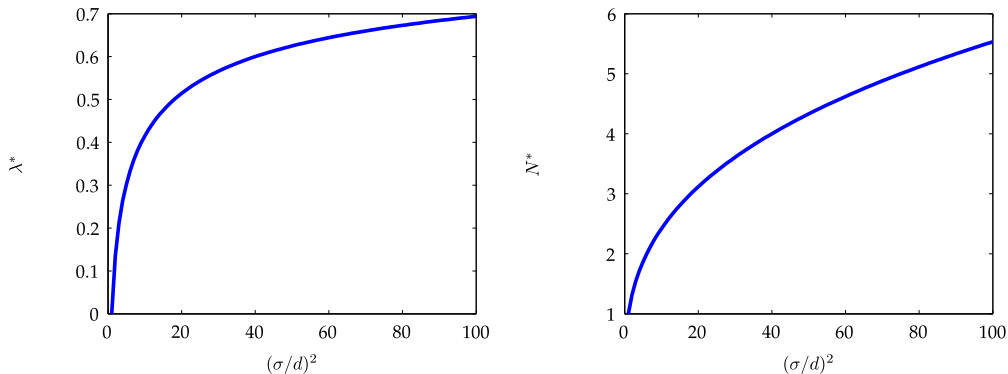


Figure 1: Relation between $(\sigma/d)^2$ and λ^* , N^* .

The quantity $(\sigma/d)^2$ serves as a proxy for the volatility of the market process $(M(t))_{t \in \mathbb{N}}$ relative to the variance of the disturbance terms $(\epsilon_t)_{t \in \mathbb{N}}$. Both for Φ_λ and Φ_N one can show that the optimal

choice of λ and N is monotone increasing in this quantity $(\sigma/d)^2$. The larger the volatility of the market compared to the variance of the disturbance terms, the fewer data should be used to estimate the market. If $(\sigma/d)^2$ is sufficiently small, then the market fluctuations are quite large relative to the variance of the disturbance terms, and it is optimal to take only the most recent data point into account to estimate the market.

3.3.3 Bounded jump probabilities for the market process.

In this section we consider assumptions on the maximum probability that the market value changes.

Proposition 5. *If $P(M(t+1) \neq M(t)) \leq \epsilon$ for all $t \in \mathbb{N}$ and some $\epsilon \geq 0$, and in addition $\sup_{t \in \mathbb{N}} M(t) - \inf_{t \in \mathbb{N}} M(t) \leq d$ for some $d > 0$, then*

$$\text{LongRunAverageRegret}(\Phi_\lambda) \leq 2K_0 \left[\sigma^2 \frac{1-\lambda}{1+\lambda} + d^2 \epsilon \frac{1}{(1-\lambda^2)} \right], \quad (22)$$

$$\text{LongRunAverageRegret}(\Phi_N) \leq 2K_0 \left[\sigma^2 \frac{1}{N} + d^2 \epsilon \frac{(N+1)(2N+1)}{6N} \right], \quad (23)$$

for all $\lambda \in [0, 1]$, $N \in \mathbb{N}_{\geq 2} \cup \{\infty\}$, where we write $1/\infty = 0$.

Consider the upper bound (22). The derivative of $\sigma^2 \frac{(1-\lambda)}{(1+\lambda)} + d^2 \epsilon (1-\lambda^2)^{-1}$ w.r.t. $\lambda \in (0, 1)$ is zero if and only if $\frac{\sigma^2}{d^2 \epsilon} (1-\lambda^2)^2 = \lambda(1+\lambda)^2$; this follows from basic algebraic manipulations. Since $(1-\lambda^2)^2$ is decreasing and $\lambda(1+\lambda)^2$ is increasing in λ , we have the following possibilities:

1. $\frac{\sigma^2}{d^2 \epsilon} \leq 1$. Then $\sigma^2 \frac{(1-\lambda)}{(1+\lambda)} + d^2 \epsilon (1-\lambda^2)^{-1}$ is increasing on $\lambda \in (0, 1)$, and the righthandside of (22) is minimized by $\lambda = 0$.
2. $\frac{\sigma^2}{d^2 \epsilon} > 1$. Then there is a unique $\lambda^* \in (0, 1)$ that minimizes $\sigma^2 \frac{(1-\lambda)}{(1+\lambda)} + d^2 \epsilon (1-\lambda^2)^{-1}$. It is the unique solution in $(0, 1)$ of the quartic equation $\frac{\sigma^2}{d^2 \epsilon} (1-\lambda^2)^2 = \lambda(1+\lambda)^2$, which can easily be solved numerically.

Now consider the upper bound (23). The expression $\frac{\sigma^2}{N} + d^2 \epsilon \frac{(N+1)(2N+1)}{6N}$ is minimized on \mathbb{R}_{++} by choosing $N^* = \sqrt{\frac{3\sigma^2}{d^2 \epsilon} + \frac{1}{2}}$, and the optimal N is equal to either $\lfloor N^* \rfloor$ or $\lceil N^* \rceil$. In addition, one can show that the optimal N equals 1 if $\frac{\sigma^2}{d^2 \epsilon} \leq \frac{1}{2}$, and is strictly larger than 1 if $\frac{\sigma^2}{d^2 \epsilon} > \frac{1}{2}$.

The quantity $\frac{\sigma^2}{d^2 \epsilon}$ serves as a proxy for the volatility of the market process $(M(t))_{t \in \mathbb{N}}$ relative to the variance of the disturbance terms $(\epsilon_t)_{t \in \mathbb{N}}$. The effect of $\frac{\sigma^2}{d^2 \epsilon}$ on λ^* and N^* is shown in Figure 2. It shows that the smaller the volatility of the market relative to natural fluctuations of demand (e.g. the larger $\frac{\sigma^2}{d^2 \epsilon}$), the more data should be taken into account to estimate the market process.

4 Numerical Illustration

In this section, we describe two numerical experiments that illustrate the method of hedging against changes outlined in Section 3. In the first we consider pricing with the Bass model for the market process. In the second we consider pricing in a setting with price-changing competitors.

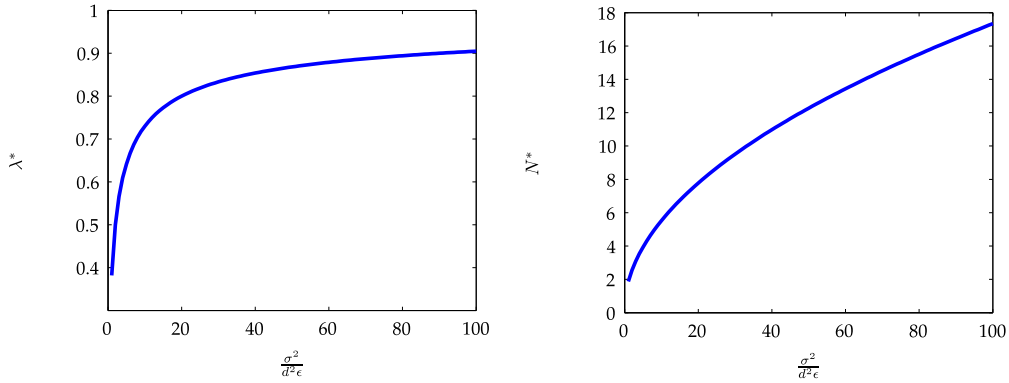


Figure 2: Relation between $\frac{\sigma^2}{\mu^2 \epsilon}$ and λ^* , N^*

4.1 Pricing with the Bass Model for the Market Process

The Bass model (Bass, 1969) is a widely-used model to describe the life-cycle or diffusion of an innovative product. An important property of this model is that the market process $M(t)$ is dependent on the realized cumulative sales up to time t .

Set-up:

The model for $M(t)$ is

$$M(t) = \max \left\{ 0, a + b \sum_{i=1}^{t-1} d_i + c \left(\sum_{i=1}^{t-1} d_i \right)^2 \right\},$$

cf. equation (4) of Dodds (1973). We choose $a = 33.6$, $c = -10^{-6}$ and $b = 0.0116$, and set $g_t(p) = g(p) = -p$ for all $t \in \mathbb{N}$, $p_l = 1$ and $p_h = 50$. Let $(\epsilon_t)_{t \in \mathbb{N}}$ be i.i.d. realizations of a standard normal distribution. The characteristic shape of the market that arises from this model is depicted in Figure 3. The solid lines denote a sample path of $M(t)$, the dashed lines a sample path of the estimates $\hat{M}_\lambda(t)$ and $\hat{M}_N(t)$.

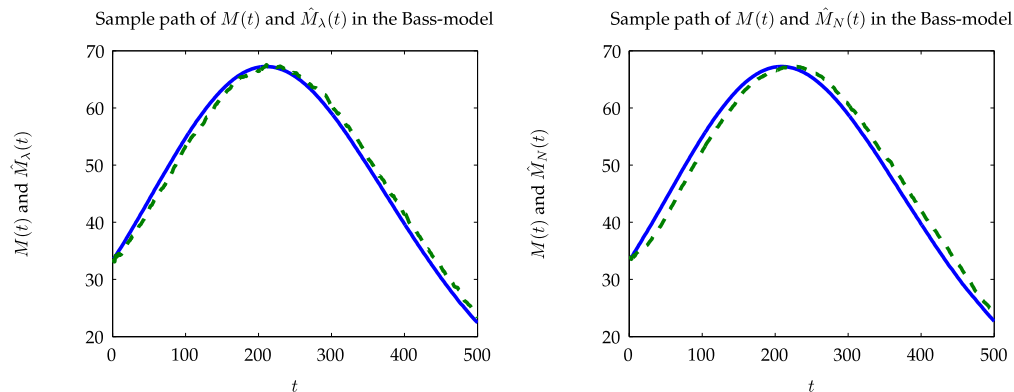


Figure 3: Sample path of $M(t)$ and $\hat{M}(t)$ in the Bass-model

For each $\lambda \in \{0.05, 0.10, 0.15, \dots, 0.90\}$ we run 1000 simulations of the policy Φ_λ , and for all $N \in \{2, 3, 4, \dots, 25\}$, we run 1000 simulations of Φ_N .

Results:

The solid lines in Figure 4 show the simulation-average of AverageRegret at $t = 500$ for both Φ_λ and Φ_N , at different values of λ . The dashed lines show the upper bounds $2K_0(\sigma^2\frac{1-\lambda}{1+\lambda} + c^{(I)}(\lambda))$ for Φ_λ , and $2K_0(\sigma^2/N + c^{(II)}(N))$ for Φ_N , where $c^{(I)}(\lambda)$ and $c^{(II)}(N)$ are as in Section 3.3.2, $\sigma^2 = 1$, $K_0 = 1/4$, and $d = 0.27$ (this was the largest observed value of $|M(t+1) - M(t)|$ over all t and all simulations. Of course, this quantity is in practice not observed by the seller, and a larger value of d just shifts the dashed lines upward in the figure).

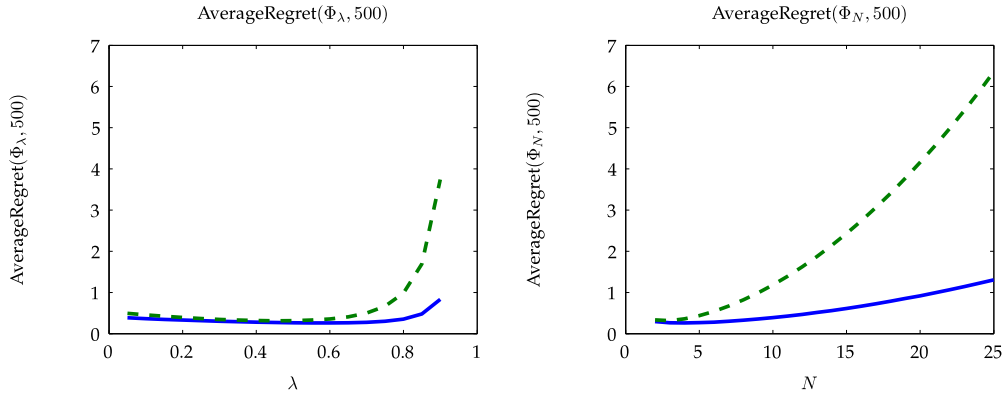


Figure 4: AverageRegret($\Phi_\lambda, 500$) and AverageRegret($\Phi_N, 500$) for the Bass model

The optimal value of λ according to our upper bound equals $\lambda = 0.45$, with a corresponding upper bound on the regret of 0.31. The simulation average of AverageRegret($\Phi_{0.45}, 500$) was equal to 0.27. The optimal value of λ according to the simulations, was $\lambda = 0.60$, with a simulation average of AverageRegret($\Phi_{0.60}, 500$) equal to 0.26.

The optimal value of N according to our upper bound equals $N = 3$, with a corresponding upper bound on the regret of 0.32. The simulation average of AverageRegret($\Phi_3, 500$) was equal to 0.27. The optimal value of N according to the simulations, was $N = 4$, with a simulation average of AverageRegret($\Phi_4, 500$) equal to 0.26.

Comparison to other methods:

Figure 3 shows that the range of values that the market process attains can be quite large. A robust optimization approach would give very conservative prices, and would lead to an average regret that is substantially larger than what is achieved by our pricing method. Neglecting the variability of $M(t)$ in the estimation step (by taking $\lambda = 1$ or $N = \infty$) is detrimental as well, as illustrated by Figure 4. Thus, in this scenario, taking into account the changing nature of the market process improves the performance of the firm significantly.

4.2 Pricing in the Presence of Price-Changing Competitors

Suppose the firm is acting in an environment where several competing companies are selling substitute products on the market. The firm knows that the competitors occasionally update their selling prices, but is not aware of the moments at which these changes occur.

In particular, consider the following case. The firm assumes that in each period, the probability that the market process changes because of the behavior of competitors, is not more than ϵ . If a change occurs, the maximum jump is assumed to be not more than d .

Set-up:

We choose $g_t(p) = g(p) = -p$ for all $t \in \mathbb{N}$, $p_l = 1$ and $p_h = 50$, and let $\epsilon = 0.02$, $d = 5$. At each period t a realization z_t of a uniformly distributed random variable on $[0, 1]$ is drawn. If $z_t \geq 0.02$ then $M(t) = M(t - 1)$; otherwise, $M(t)$ is drawn uniformly from the interval $[30, 35]$. Let $(\epsilon_t)_{t \in \mathbb{N}}$ be i.i.d. realizations of a standard normal distribution. (Note that these differ from the constant ϵ determined by the firm).

For each $\lambda \in \{0.10, 0.15, 0.20, \dots, 0.95\}$ we run 1000 simulations of the policy Φ_λ , and for all $N \in \{2, 3, 4, \dots, 25\}$, we run 1000 simulations of Φ_N .

Results:

The characteristic the shape of the market that arises from this model, is depicted in Figure 5. The solid lines denote a sample path of $M(t)$, the dashed lines a sample path of the estimates $\hat{M}_\lambda(t)$ and $\hat{M}_N(t)$.

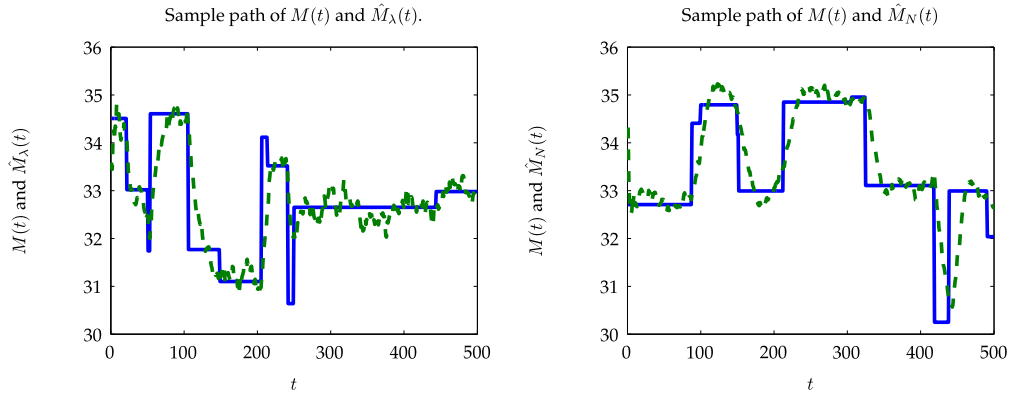


Figure 5: Sample path of $M(t)$ and $\hat{M}(t)$ in the model with price-changing competitors.

The solid lines in Figure 6 show the simulation average of AverageRegret at $t = 500$ for both Φ_λ and Φ_N , at different values of λ . The dashed lines show the upper bounds $K_0(\sigma^2 \frac{1-\lambda}{1+\lambda} + c^{(I)}(\lambda))$ for Φ_λ , and $K_0(\sigma^2/N + c^{(II)}(N))$ for Φ_N , where $c^{(I)}(\lambda)$ and $c^{(II)}(N)$ are as in Section 3.3.3, $\sigma^2 = 1$, $K_0 = 1/4$, $\epsilon = 0.02$, and $d = 5$. Note that $(\epsilon_t)_{t \in \mathbb{N}}$ and $(M(t))_{t \in \mathbb{N}}$ are here independent, and thus by Remark 6, the factor 2 in the righthandsides of (14) and (15) is not present.

The optimal value of λ according to our upper bound equals $\lambda = 0.50$, with a corresponding upper bound on the regret of 0.25. The simulation average of $\text{AverageRegret}(\Phi_{0.50}, 500)$ was equal to 0.11. The optimal value of λ according to the simulations, was $\lambda = 0.75$, with a simulation average of $\text{AverageRegret}(\Phi_{0.75}, 500)$ equal to 0.08.

The optimal value of N according to our upper bound equals $N = 3$, with a corresponding upper bound on the regret of 0.28. The simulation average of $\text{AverageRegret}(\Phi_3, 500)$ was equal to 0.12.

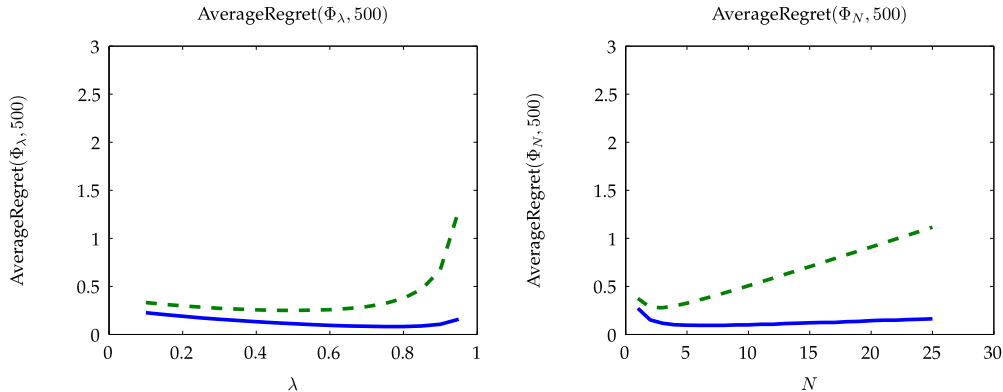


Figure 6: AverageRegret($\Phi_\lambda, 500$) and AverageRegret($\Phi_N, 500$) for experiment 3.

The optimal value of N according to the simulations, was $N = 6$, with a simulation average of AverageRegret($\Phi_6, 500$) equal to 0.09.

Comparison to other methods:

Figure 6 illustrates that taking into account all available data (i.e. $\lambda = 1$ or $N = \infty$) would lead to much larger regret than obtained at the optimal λ and N . Thus, similar to scenario (ii), taking into account the changing nature of the market process leads to a significant profit improvement. A robust maximin pricing policy would be to use

$$\arg \max_{p \in [1, 50]} \min_{M \in [30, 35]} p(M - p) = \arg \max_{p \in [1, 50]} p(30 - p) = 15$$

throughout the time horizon. This leads to an average regret of 1.1509, more than three times higher than the average regret of 0.3189 achieved by our method. Even assuming that $M(t)$ is fixed and equal to 32.5 (and using the corresponding optimal price $p = 16.75$ throughout the time horizon) would, in our simulations, lead to an average regret of 1.0745; still more than three times higher than what is achieved by our method.

5 Conclusion and Future Research

In this paper we study the problem of dynamic pricing and learning in a changing market environment. This is a major departure from the existing literature on dynamic pricing and learning, in which one practically always assumes that the market is stable. We consider a setting where the market process is modeled as a stochastic process, whose value is not directly observed by the firm. We discuss two suitable estimation methods, with a forgetting factor and with a sliding window, and prove bounds on the expected estimation errors. Subsequently we introduce a methodology that enables the firm to hedge against changes in the market. In particular, we show how assumptions on the market process, determined in advance by the firm, translate into upper bounds on the (long run) average regret, and, in addition, how these bounds can be used to derive the optimal forgetting factor or window size. We show in three concrete scenarios how

the methodology works, and provide numerical illustrations that show the good performance of the method in the Bass-market model and in a setting with price-changing competitors.

An important insight from our results is that taking into account the fluctuating nature of the market can significantly improve the pricing decisions of a firm.

Our results points to several interesting directions for future research. Related to the dynamic pricing model, an interesting extension would be to assume that both σ^2 and $g_t(p)$ are unknown, and have to be learned as well. To begin with, one could assume the parametric form $g_t(p) = g(p) = -bp$ for some $b > 0$. One step further is to consider the case that σ^2 and b themselves are also varying over time. Even for the bound functions $c^{(I)}(\lambda)$ and $c^{(II)}(N)$, information about their behavior might be derived from sales data, by estimating the impact $I_\lambda(t)$ and $I_N(t)$. An ad-hoc method to do so would be to replace all terms $M(i)$ in the definition of $I_\lambda(t)$ by their estimate $\hat{M}_i(t)$, and a similar procedure to estimate $I_N(t)$.

Finally, we believe that the methodology developed in this paper might be useful not only for the considered dynamic pricing problem, but also for other types of problems that involve simultaneous learning and optimizing in a changing environment. Two examples are stochastic inventory control problems (Huh and Rusmevichientong, 2009, Huh et al., 2011), or dynamic pricing with finite inventories (Besbes and Zeevi, 2009, den Boer and Zwart, 2011).

6 Proofs

Proof of Proposition 1

Equation (7) can be rewritten as

$$\hat{M}_\lambda(t) - M(t+1) = \frac{\sum_{i=1}^t \epsilon_i \lambda^{t-i}}{\sum_{i=1}^t \lambda^{t-i}} + \frac{\sum_{i=1}^t (M(i) - M(t+1)) \lambda^{t-i}}{\sum_{i=1}^t \lambda^{t-i}}.$$

Note that $(\sum_{i=1}^t \lambda^{t-i})^{-1} = (1-\lambda^t)^{-1}(1-\lambda)\mathbf{1}(\lambda < 1) + \frac{1}{t}\mathbf{1}(\lambda = 1)$ and $E[\epsilon_i \epsilon_j] = E[\epsilon_i E[\epsilon_j | \mathcal{F}_i]] = 0$ whenever $i < j$. As a result,

$$E \left[\left| \sum_{i=1}^t \epsilon_i \lambda^{t-i} \right|^2 \right] = \sum_{i=1}^t \lambda^{2(t-i)} E[\epsilon_i^2] \leq \sigma^2 \left(\frac{1-\lambda^{2t}}{1-\lambda^2} \mathbf{1}(\lambda < 1) + t \mathbf{1}(\lambda = 1) \right),$$

and (10) follows using $|a+b|^2 \leq 2a^2 + 2b^2$ for all $a, b \in \mathbb{R}$, and

$$\begin{aligned} & \left(\frac{1-\lambda^{2t}}{1-\lambda^2} \mathbf{1}(\lambda < 1) + t \mathbf{1}(\lambda = 1) \right) \left(\frac{1-\lambda}{1-\lambda^t} \mathbf{1}(\lambda < 1) + \frac{1}{t} \mathbf{1}(\lambda = 1) \right)^2 \\ &= \frac{1-\lambda}{1+\lambda} \frac{1+\lambda^t}{1-\lambda^t} \mathbf{1}(\lambda < 1) + \frac{1}{t} \mathbf{1}(\lambda = 1). \end{aligned}$$

If $(\epsilon_t)_{t \in \mathbb{N}}$ and $(M(t))_{t \in \mathbb{N}}$ are independent, then $E[\epsilon_i M(j)] = 0$ for all $i, j \in \mathbb{N}$, and (12) follows from

$$\begin{aligned} E \left[\left| \hat{M}_\lambda(t) - M(t+1) \right|^2 \right] &= E \left[\left| \frac{\sum_{i=1}^t \epsilon_i \lambda^{t-i}}{\sum_{i=1}^t \lambda^{t-i}} \right|^2 \right] + E \left[\left| \frac{\sum_{i=1}^t (M(i) - M(t+1)) \lambda^{t-i}}{\sum_{i=1}^t \lambda^{t-i}} \right|^2 \right] \\ &\leq \sigma^2 \left| \frac{\sum_{i=1}^t \lambda^{t-i}}{\sum_{i=1}^t \lambda^{t-i}} \right|^2 + E \left[\left| \frac{\sum_{i=1}^t (M(i) - M(t+1)) \lambda^{t-i}}{\sum_{i=1}^t \lambda^{t-i}} \right|^2 \right], \end{aligned}$$

with equality if $(\epsilon_t)_{t \in \mathbb{N}}$ is homoscedastic.

Similarly, equation (9) can be rewritten as

$$\hat{M}_N(t) - M(t+1) = \frac{1}{\min\{N, t\}} \left(\sum_{i=1+(t-N)^+}^t \epsilon_i + \sum_{i=1+(t-N)^+}^t (M(i) - M(t+1)) \right).$$

Equation (11) follows using $|a+b|^2 \leq 2a^2 + 2b^2$ for all $a, b \in \mathbb{R}$, and by noting

$$E \left[\left| \frac{1}{\min\{N, t\}} \sum_{i=1+(t-N)^+}^t \epsilon_i \right|^2 \right] = \frac{1}{\min\{N, t\}^2} \sum_{i=1+(t-N)^+}^t E[\epsilon_i^2] \leq \sigma^2 / \min\{N, t\}.$$

If $(\epsilon_t)_{t \in \mathbb{N}}$ and $(M(t))_{t \in \mathbb{N}}$ are independent, then $E[\epsilon_i M(j)] = 0$ for all $i, j \in \mathbb{N}$, and (13) follows

from

$$\begin{aligned}
& E \left[\left| \hat{M}_N(t) - M(t+1) \right|^2 \right] \\
&= E \left[\left| \frac{1}{\min\{N, t\}} \sum_{i=1+(t-N)^+}^t \epsilon_i \right|^2 \right] + E \left[\left| \frac{1}{\min\{N, t\}} \sum_{i=1+(t-N)^+}^t M(i) - M(t+1) \right|^2 \right] \\
&\leq \sigma^2 \left| \frac{1}{\min\{N, t\}} \sum_{i=1+(t-N)^+}^t 1 \right|^2 + E \left[\left| \frac{1}{\min\{N, t\}} \sum_{i=1+(t-N)^+}^t M(i) - M(t+1) \right|^2 \right],
\end{aligned}$$

with equality if $(\epsilon_t)_{t \in \mathbb{N}}$ is homoscedastic.

Proof of Theorem 1

We prove the theorem in two steps. In step 1, we show that there exists a $K_0 > 0$ such that for all $M \in \mathcal{M}$, $M' \in \mathbb{R}$ and for all $t \in \mathbb{N}$,

$$r_t(p_t^*(M), M) - r_t(p_t^*(M'), M) \leq K_0(M - M')^2. \quad (24)$$

In step 2 we apply this result with $M = M(t)$, $M' = \hat{M}_\lambda(t)$ or $M' = \hat{M}_N(t)$, to obtain the regret bounds.

Step 1. Fix an attainable value $M \in \mathcal{M}$ of the market process, fix $t \in \mathbb{N}$, and let $r_t'(p, M)$ and $r_t''(p, M)$ denote the first and second derivative of $r_t(p, M)$ w.r.t. p . Let $M' \in \mathbb{R}$.

Case 1: $p_t^*(M) = p_t^\#(M)$. Then by assumption $r_t'(p_t^*(M), M) = 0$, and a Taylor series expansion yields

$$r_t(p, M) = r_t(p_t^*(M), M) + \frac{1}{2} r_t''(\tilde{p}, M)(p - p_t^*(M))^2,$$

for some \tilde{p} on the line segment between p and $p_t^*(M)$. Let

$$K_t = \sup_{p \in [p_l, p_h]} |r_t''(p, M)| = \sup_{p \in [p_l, p_h]} |2g_t'(p) + g_t''(p)|,$$

and note that K_t is independent of M , and finite, because of the continuity of $g''(p)$. Then

$$r_t(p_t^*(M), M) - r_t(p, M) \leq \frac{K_t}{2} (p - p_t^*(M))^2 \quad \text{for all } p \in [p_l, p_h]. \quad (25)$$

Write $h_t(p) = -g_t(p) - pg_t'(p)$, and note that $r_t'(p, M) = M - h_t(p)$. By assumption, for each $M \in \mathcal{M}$ there is a unique $p_t^\#(M)$ such that $h_t(p) = M$, i.e. $p_t^\#(M) = h_t^{-1}(M)$ is well-defined. In addition, for all $M \in h_t([p_l, p_h]) = \{h_t(p) \mid p \in [p_l, p_h]\}$, we have $\frac{\partial}{\partial M} p_t^\#(M) = (h_t^{-1})'(M) = 1/h_t'(h_t^{-1}(M)) = -1/r_t''(p_t^*(M), M) > 0$. Thus, $p_t^\#(M)$ is continuous, differentiable, and monotone increasing on $M \in h_t([p_l, p_h])$. These properties imply the following: if there is an $M \in \mathcal{M}$ s.t. $p_t^\#(M) > p_h$, then there is an $M_h(t)$ s.t. $h_t^{-1}(M) > p_h$ whenever $M > M_h(t)$, $h_t^{-1}(M_h(t)) = p_h$,

and $h_t^{-1}(M) < p_h$ whenever $M < M_h(t)$. Similarly, if there is an $M \in \mathcal{M}$ s.t. $p_t^\#(M) < p_l$, then there is an $M_l(t) < M_h(t)$ s.t. $h_t^{-1}(M) > p_l$ whenever $M > M_l(t)$, $h_t^{-1}(M_l(t)) = p_l$, and $h_t^{-1}(M) < p_l$ whenever $M < M_l(t)$.

If $p_t^*(M') = p_t^\#(M')$, then a Taylor expansion yields

$$|p_t^*(M') - p_t^*(M)| = |h_t^{-1}(M') - h_t^{-1}(M)| \leq |M' - M|L_t,$$

where $L_t = \sup_{M \in h_t([p_l, p_h])} |(h_t^{-1})'(M)| = 1/\inf_{M \in h_t([p_l, p_h])} |r_t''(p_t^*(M), M)|$, which is finite by assumption.

If $p_t^*(M') < p_t^\#(M')$, then $p_t^*(M') = p_t^\#(M_h(t)) = p_h$, $M' > M_h(t)$, and

$$|p_t^*(M') - p_t^*(M)| = |p_t^\#(M_h(t)) - p_t^\#(M)| \leq |M_h(t) - M|L_t \leq |M' - M|L_t.$$

If $p_t^*(M') > p_t^\#(M')$, then $p_t^*(M') = p_t^\#(M_l(t)) = p_l$, $M' < M_l(t)$, and

$$|p_t^*(M') - p_t^*(M)| = |p_t^\#(M_l(t)) - p_t^\#(M)| \leq |M_l(t) - M|L_t \leq |M' - M|L_t.$$

It follows that $|p_t^*(M') - p_t^*(M)| \leq L_t|M' - M|$, and thus by (25) we have

$$r_t(p_t^*(M), M) - r_t(p_t^*(M'), M) \leq \frac{1}{2}K_t L_t^2 (M' - M)^2, \quad (26)$$

for all M' and all $M \in h_t([p_l, p_h])$.

Case 2: $p_t^*(M) \neq p_t^\#(M)$. Then $M \notin [M_l(t), M_h(t)]$. Suppose $M > M_h(t)$, the case $M < M_l(t)$ is treated likewise. If $M' > M_h(t)$ then $r_t(p_t^*(M), M) - r_t(p_t^*(M'), M) = 0$, suppose therefore $M' \leq M_h(t)$. We have

$$\begin{aligned} & r_t(p_t^*(M), M) - r_t(p_t^*(M'), M) = r_t(p_t^*(M_h(t)), M) - r_t(p_t^*(M'), M) \\ &= p_t^*(M_h(t))[M + g_t(p_t^*(M_h(t)))] - p_t^*(M')[M + g_t(p_t^*(M'))] \\ &= r_t(p_t^*(M_h(t)), M_h(t)) - r_t(p_t^*(M'), M_h(t)) + (p_t^*(M_h(t)) - p_t^*(M'))(M - M_h(t)) \\ &\leq \frac{1}{2}K_t L_t^2 (M' - M_h(t))^2 + L_t(M_h(t) - M')(M - M_h(t)) \\ &\leq \left(\frac{1}{2}K_t L_t^2 + \frac{1}{4}L_t \right) (M' - M)^2, \end{aligned}$$

where in the last inequality we use the fact $xy \leq \frac{1}{4}(x + y)^2$, $x, y \in \mathbb{R}$, with $x = M_h(t) - M'$, $y = M - M_h(t)$.

This completes the proof of (24), with $K_0 = \sup_{t \in \mathbb{N}} \frac{1}{2}K_t L_t^2 + \frac{1}{4}L_t$.

Step 2.

By Proposition 1, we obtain

$$\begin{aligned}
& \text{AverageRegret}(\Phi_\lambda, T) \\
&= \frac{1}{T-1} \sum_{t=1}^{T-1} E \left[r_t(p_t^*(M(t+1)), M(t+1)) - r_t(p_t^*(\hat{M}_\lambda(t)), M(t+1)) \right] \\
&\leq \frac{K_0}{T-1} \sum_{t=1}^{T-1} E \left[\left| \hat{M}_\lambda(t) - M(t+1) \right|^2 \right] \\
&\leq \frac{2K_0}{T-1} \sum_{t=1}^{T-1} \left[\sigma^2 \left[\frac{(1-\lambda)(1+\lambda^t)}{(1+\lambda)(1-\lambda^t)} \mathbf{1}(\lambda < 1) + \frac{1}{t} \mathbf{1}(\lambda = 1) \right] + I_\lambda(t) \right].
\end{aligned}$$

Since

$$\begin{aligned}
\sum_{t=1}^{T-1} \frac{\lambda^t}{1-\lambda^t} &= \frac{\lambda}{1-\lambda} + \sum_{t=2}^{T-1} \frac{\lambda^t}{1-\lambda^t} \leq \frac{\lambda}{1-\lambda} + \int_{t=1}^{T-2} \frac{\lambda^t}{1-\lambda^t} dt \\
&\leq \frac{\lambda}{1-\lambda} + \frac{-1}{\log(\lambda)} \int_{x=0}^{\lambda} \frac{1}{1-x} dx = \frac{\lambda}{1-\lambda} + \frac{\log(1-\lambda)}{\log(\lambda)},
\end{aligned}$$

we have for $\lambda < 1$,

$$\begin{aligned}
\frac{1}{T-1} \sum_{t=1}^{T-1} \frac{(1-\lambda)(1+\lambda^t)}{(1+\lambda)(1-\lambda^t)} &= \frac{1-\lambda}{1+\lambda} + \frac{2}{T-1} \frac{1-\lambda}{1+\lambda} \sum_{t=1}^{T-1} \frac{\lambda^t}{1-\lambda^t} \\
&\leq \frac{1-\lambda}{1+\lambda} + \frac{1}{T-1} \left(\frac{2\lambda}{1+\lambda} + 2 \frac{1-\lambda \log(1-\lambda)}{1+\lambda \log(\lambda)} \right), \tag{27}
\end{aligned}$$

and thus

$$\begin{aligned}
& \text{AverageRegret}(\Phi_\lambda, T) \\
&\leq 2K_0\sigma^2 \left[\frac{1-\lambda}{1+\lambda} + \frac{1}{T-1} \left(\frac{2\lambda}{1+\lambda} + 2 \frac{1-\lambda \log(1-\lambda)}{1+\lambda \log(\lambda)} \right) \right] \mathbf{1}(\lambda < 1) \\
&+ 2K_0\sigma^2 \left[\frac{1+\log(T-1)}{T-1} \right] \mathbf{1}(\lambda = 1) \\
&+ \frac{2K_0}{T-1} \sum_{t=1}^{T-1} I_\lambda(t).
\end{aligned}$$

In addition, we have

$$\begin{aligned}
& \text{AverageRegret}(\Phi_N, T) \\
&= \frac{1}{T-1} \sum_{t=1}^{T-1} E \left[r_t(p_t^*(M(t+1)), M(t+1)) - r_t(p_t^*(\hat{M}_N(t)), M(t+1)) \right] \\
&\leq \frac{K_0}{T-1} \sum_{t=1}^{T-1} E \left[\left| \hat{M}_N(t) - M(t+1) \right|^2 \right] \\
&\leq \frac{2K_0}{T-1} \sum_{t=1}^{T-1} \left[\frac{\sigma^2}{\min\{N, t\}} + I_N(t) \right] \\
&\leq 2K_0\sigma^2 \left[\frac{\log(\min\{T-1, N\})}{T-1} + \frac{1}{\min\{N, T-1\}} \right] + \frac{2K_0}{T-1} \sum_{t=1}^{T-1} I_N(t),
\end{aligned}$$

where we used

$$\begin{aligned}
\sum_{t=1}^{T-1} \frac{1}{\min\{N, t\}} &= \sum_{t=1}^N \frac{1}{t} + \sum_{t=N+1}^{T-1} \frac{1}{N} \leq 1 + \log(N) + \frac{T-1-N}{N} \quad \text{if } T-1 \geq N, \\
\sum_{t=1}^{T-1} \frac{1}{\min\{N, t\}} &= \sum_{t=1}^{T-1} \frac{1}{t} \leq 1 + \log(T-1) \quad \text{if } T-1 < N,
\end{aligned}$$

and thus

$$\sum_{t=1}^{T-1} \frac{1}{\min\{N, t\}} \leq \log(\min\{T-1, N\}) + \frac{T-1}{\min\{N, T-1\}}.$$

Proof of Proposition 2

The condition $M(t) \in [2bp_l, 2bp_h]$ a.s., for all $t \in \mathbb{N}$, implies $p^*(M) = M/(2b)$ for all attainable values of M , and $r(p^*(M), M) - r(p^*(M'), M) = (M - M')^2/(4b)$ for all attainable values of M and M' . By Proposition 1 we obtain

$$\begin{aligned}
& \text{LongRunAverageRegret}(\Phi_\lambda) \\
&= \limsup_{T \rightarrow \infty} \frac{1}{T-1} \sum_{t=1}^{T-1} E \left[r(p^*(M(t+1)), M(t+1)) - r(p^*(\hat{M}_\lambda(t)), M(t+1)) \right] \\
&= \limsup_{T \rightarrow \infty} \frac{K_0}{T-1} \sum_{t=1}^{T-1} E \left[\left| \hat{M}_\lambda(t) - M(t+1) \right|^2 \right] \\
&= \limsup_{T \rightarrow \infty} \frac{K_0}{T-1} \sum_{t=1}^{T-1} \left[\sigma^2 \left[\frac{(1-\lambda)(1+\lambda^t)}{(1+\lambda)(1-\lambda^t)} \mathbf{1}(\lambda < 1) + \frac{1}{t} \mathbf{1}(\lambda = 1) \right] + I_\lambda(t) \right] \\
&= K_0 \left[\sigma^2 \frac{(1-\lambda)}{(1+\lambda)} + \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T I_\lambda(t) \right],
\end{aligned}$$

and

$$\begin{aligned}
& \text{LongRunAverageRegret}(\Phi_N) \\
&= \limsup_{T \rightarrow \infty} \frac{1}{T-1} \sum_{t=1}^{T-1} E \left[r(p^*(M(t+1)), M(t+1)) - r(p^*(\hat{M}_N(t)), M(t+1)) \right] \\
&= \limsup_{T \rightarrow \infty} \frac{K_0}{T-1} \sum_{t=1}^{T-1} E \left[\left| \hat{M}_N(t) - M(t+1) \right|^2 \right] \\
&= \limsup_{T \rightarrow \infty} \frac{K_0}{T-1} \sum_{t=1}^{T-1} \left[\frac{\sigma^2}{\min\{N, t\}} + I_N(t) \right] \\
&= K_0 \left[\sigma^2 \frac{1}{N} + \limsup_{T \rightarrow \infty} \frac{1}{T-1} \sum_{t=1}^{T-1} I_N(t) \right].
\end{aligned}$$

6.1 Proof of Proposition 3

The assumption $\sup_{t \in \mathbb{N}} M(t) - \inf_{t \in \mathbb{N}} M(t) \leq d$ implies $\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T I_\lambda(t) \leq d^2$ and $\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T I_N(t) \leq d^2$. Together with Theorem 1 this proves the proposition.

6.2 Proof of Proposition 4

We show that the assumption $|M(t) - M(t+1)| \leq d$ a.s., for some $d \geq 0$ and all $t \in \mathbb{N}$, implies $\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T I_\lambda(t) \leq d^2(1-\lambda)^{-2}$ and $\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T I_N(t) \leq \frac{1}{4}d^2(N+1)^2$, for any $\lambda \in [0, 1)$ and $N \in \mathcal{N}_{\geq 2}$. Together with Theorem 1 this proves the proposition.

Let $\lambda \in [0, 1)$. Then

$$\begin{aligned}
\frac{1}{T-1} \sum_{t=1}^{T-1} I_\lambda(t) &= \frac{1}{T-1} \sum_{t=1}^{T-1} E \left[\left| \frac{1-\lambda}{1-\lambda^t} \sum_{i=1}^t (M(i) - M(t+1)) \lambda^{t-i} \right|^2 \right] \\
&\leq \frac{1}{T-1} \sum_{t=1}^{T-1} \frac{(1-\lambda)^2}{(1-\lambda^t)^2} \left| \sum_{i=1}^{t+1} d(t+1-i) \lambda^{t-i} \right|^2 \\
&= \frac{1}{T-1} \sum_{t=1}^{T-1} \frac{(1-\lambda)^{-2}}{(1-\lambda^t)^2} d^2 \left| -(t+1)(1-\lambda)\lambda^t + (1-\lambda^{t+1}) \right|^2 \\
&= \frac{1}{T-1} (1-\lambda)^{-2} d^2 \sum_{t=1}^{T-1} \left(1 - t\lambda^t \frac{1-\lambda}{1-\lambda^t} \right)^2,
\end{aligned}$$

from which it follows that

$$\limsup_{T \rightarrow \infty} \frac{1}{T-1} \sum_{t=1}^{T-1} I_\lambda(t) \leq d^2(1-\lambda)^{-2}.$$

Let $N \in \mathbb{N}_{\geq 2}$, then

$$\begin{aligned}
& \frac{1}{T-1} \sum_{t=1}^{T-1} I_N(t) \\
&= \frac{1}{T-1} \sum_{t=1}^{T-1} E \left[\left| \frac{1}{\min\{N, t\}} \sum_{i=1+(t-N)^+}^t (M(i) - M(t+1)) \right|^2 \right] \\
&\leq \frac{1}{T-1} \sum_{t=1}^{T-1} \left| \frac{1}{\min\{N, t\}} \sum_{i=1+(t-N)^+}^t d(t+1-i) \right|^2 \\
&= \frac{1}{T-1} \sum_{t=1}^{T-1} \left| \frac{d}{\min\{N, t\}} \sum_{j=1}^{\min\{N, t\}} j \right|^2 = \frac{1}{T-1} \sum_{t=1}^{T-1} \frac{1}{4} d^2 (\min\{N, t\} + 1)^2 \\
&= \frac{d^2}{4} \frac{1}{T-1} \sum_{t=1}^{T-1} (N+1)^2 + \frac{d^2}{4} \frac{1}{T-1} \sum_{t=1}^{\min\{T-1, N-1\}} [(t+1)^2 - (N+1)^2] \\
&= \frac{d^2}{4} (N+1)^2 + \frac{d^2}{4} \frac{1}{T-1} \left[-\min\{T-1, N-1\} (N+1)^2 + \sum_{t=2}^{\min\{T, N\}} t^2 \right] \\
&= \frac{d^2}{4} (N+1)^2 + \frac{d^2}{4} \frac{1}{T-1} \cdot \\
&\quad \left[(1 - \min\{T, N\}) (N+1)^2 - 1 + \min\{T, N\} (\min\{T, N\} + 1) (2 \min\{T, N\} + 1) / 6 \right],
\end{aligned}$$

where we used $\sum_{t=1}^N t^2 = N(N+1)(2N+1)/6$. After some algebraic manipulations, we derive that $\frac{1}{T-1} \sum_{t=1}^{T-1} I_N(t)$ can be upper bounded by

$$\begin{cases} \frac{1}{4} d^2 \frac{-1+T(T+1)(2T+1)/6}{T-1} & \text{if } T < N \\ \frac{1}{4} d^2 \left[(N+1)^2 + \frac{1}{T-1} N(-4N^2 - 3N + 7)/6 \right] & \text{if } T \geq N \end{cases} .$$

Taking $\limsup_{T \rightarrow \infty}$, we obtain

$$\limsup_{T \rightarrow \infty} \frac{1}{T-1} \sum_{t=1}^{T-1} I_N(t) \leq \frac{1}{4} d^2 (N+1)^2.$$

6.3 Proof of Proposition 5

We show that the assumptions $P(M(t+1) \neq M(t)) \leq \epsilon$ for all $t \in \mathbb{N}$ and some $\epsilon \geq 0$, and $\sup_{t \in \mathbb{N}} M(t) - \inf_{t \in \mathbb{N}} M(t) \leq d$ for some $d > 0$, imply $\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T I_\lambda(t) \leq d^2 \epsilon \frac{1}{(1-\lambda^2)}$ and $\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T I_N(t) \leq d^2 \epsilon \frac{(N+1)(2N+1)}{6N}$, for any $\lambda \in [0, 1)$ and $N \in \mathcal{N}_{\geq 2}$. Together with Theorem 1 this proves the proposition.

For $t \in \mathbb{N}$, define

$$X(t) = \min\{k \in \{1, \dots, t+1\} \mid M(k) = M(k+1) = \dots = M(t+1)\},$$

and note that $P(X(t) = k) \leq P(M(k-1) \neq M(k)) \leq \epsilon$. for all $k = 2, \dots, t+1$.

For $\lambda \in [0, 1)$, we have

$$\begin{aligned}
& E \left[\left| \sum_{i=1}^t (M(i) - M(t+1)) \lambda^{t-i} \right|^2 \right] \\
&= \sum_{k=1}^{t+1} E \left[\left| \sum_{i=1}^t (M(i) - M(t+1)) \lambda^{t-i} \right|^2 \mid X(t) = k \right] P(X(t) = k) \\
&\leq \sum_{k=2}^{t+1} E \left[\left| \sum_{i=1}^{k-1} (M(i) - M(t+1)) \lambda^{t-i} \right|^2 \mid X(t) = k \right] \epsilon \\
&\leq \sum_{k=2}^{t+1} d^2 \left| \sum_{i=1}^{k-1} \lambda^{t-i} \right|^2 \epsilon \\
&= d^2 \epsilon (1 - \lambda)^{-2} [(1 - \lambda^2)^{-1} (1 - \lambda^{2t}) - 2\lambda^t (1 - \lambda)^{-1} (1 - \lambda^t) + t\lambda^{2t}],
\end{aligned}$$

and thus

$$\begin{aligned}
& \limsup_{T \rightarrow \infty} \frac{1}{T-1} \sum_{t=1}^{T-1} I_\lambda(t) \\
&= \limsup_{T \rightarrow \infty} \frac{1}{T-1} \sum_{t=1}^{T-1} E \left[\left| \frac{1-\lambda}{1-\lambda^t} \sum_{i=1}^t (M(i) - M(t+1)) \lambda^{t-i} \right|^2 \right] \\
&\leq \limsup_{T \rightarrow \infty} \frac{1}{T-1} \sum_{t=1}^{T-1} \frac{1}{(1-\lambda^t)^2} d^2 \epsilon [(1 - \lambda^2)^{-1} (1 - \lambda^{2t}) - 2\lambda^t (1 - \lambda)^{-1} (1 - \lambda^t) + t\lambda^{2t}] \\
&= d^2 \epsilon (1 - \lambda^2)^{-1}.
\end{aligned}$$

Let $N \in \mathbb{N}_{\geq 2}$, then

$$\begin{aligned}
I_N(t) &= E \left[\left| \frac{1}{\min\{N, t\}} \sum_{i=1+(t-N)^+}^t (M(i) - M(t+1)) \right|^2 \right] \\
&= \sum_{k=1}^{t+1} E \left[\left| \frac{1}{\min\{N, t\}} \sum_{i=1+(t-N)^+}^t (M(i) - M(t+1)) \right|^2 \mid X(t) = k \right] P(X(t) = k) \\
&\leq \sum_{k=2}^{t+1} \left| \frac{1}{\min\{N, t\}} \sum_{i=1+(t-N)^+}^{k-1} d \right|^2 \epsilon = d^2 \epsilon \sum_{k=1+(t-N)^+}^t \left(\frac{k - (t-N)^+}{\min\{N, t\}} \right)^2 \\
&= d^2 \epsilon \frac{(\min\{N, t\} + 1)(2 \min\{N, t\} + 1)}{6 \min\{N, t\}},
\end{aligned}$$

and thus

$$\begin{aligned}
& \limsup_{T \rightarrow \infty} \frac{1}{T-1} \sum_{t=1}^{T-1} I_N(t) \\
& \leq \limsup_{T \rightarrow \infty} \frac{1}{T-1} \sum_{t=1}^{T-1} d^2 \epsilon \frac{(\min\{N, t\} + 1)(2 \min\{N, t\} + 1)}{6 \min\{N, t\}} \\
& = d^2 \epsilon \frac{(N+1)(2N+1)}{6N}.
\end{aligned}$$

Acknowledgment

We kindly thank Bert Zwart for reading and commenting on the manuscript.

References

- V. F. Araman and R. Caldentey. Dynamic pricing for nonperishable products with demand learning. *Operations Research*, 57(5):1169–1188, 2009.
- F. M. Bass. A new product growth for model consumer durables. *Management Science*, 15(5):215–227, 1969.
- D. Bertsimas and G. Perakis. Dynamic pricing: a learning approach. In *Mathematical and Computational Models for Congestion Charging*, pages 45–79. Springer, New York, 2006.
- O. Besbes and D. Saure. Dynamic pricing strategies in the presence of demand shocks. Working paper, 2012.
- O. Besbes and A. Zeevi. Dynamic pricing without knowing the demand function: risk bounds and near-optimal algorithms. *Operations Research*, 57(6):1407–1420, 2009.
- O. Besbes and A. Zeevi. On the minimax complexity of pricing in a changing environment. *Operations Research*, 59(1):66–79, 2011.
- J. Broder and P. Rusmevichientong. Dynamic pricing under a general parametric choice model. *Operations Research*, 60(4):965–980, 2012.
- A. X. Carvalho and M. L. Puterman. Learning and pricing in an internet environment with binomial demand. *Journal of Revenue and Pricing Management*, 3(4):320–336, 2005a.
- A. X. Carvalho and M. L. Puterman. Dynamic optimization and learning: How should a manager set prices when the demand function is unknown? Technical report, Instituto de Pesquisa Economica Aplicada - IPEA, Discussion Papers 1117, 2005b.
- Y. M. Chen and D. C. Jain. Dynamic monopoly pricing under a Poisson-type uncertain demand. *The Journal of Business*, 65(4):593–614, 1992.

- W. L. Cooper, T. Homem de Mello, and A. J. Kleywegt. Learning and pricing with models that do not explicitly incorporate competition. Working paper, 2012.
- E. Cope. Bayesian strategies for dynamic pricing in e-commerce. *Naval Research Logistics*, 54(3): 265–281, 2007.
- A. V. den Boer and B. Zwart. Simultaneously learning and optimizing using controlled variance pricing. *Management Science*, Accepted for publication, 2010.
- A. V. den Boer and B. Zwart. Dynamic pricing and learning with finite inventories. Submitted for publication, 2011.
- W. Dodds. An application of the Bass model in long-term new product forecasting. *Journal of Marketing Research*, 10(3):308–311, 1973.
- R. J. Dolan and A. P. Jeuland. Experience curves and dynamic demand models: implications for optimal pricing strategies. *Journal of Marketing*, 45(1):52–62, 1981.
- W. Elmaghraby and P. Keskinocak. Dynamic pricing in the presence of inventory considerations: research overview, current practices, and future directions. *Management Science*, 49(10):1287–1309, 2003.
- S. S. Eren and C. Maglaras. Monopoly pricing with limited demand information. *Journal of Revenue and Pricing Management*, 9:23–48, 2010.
- V. F. Farias and B. van Roy. Dynamic pricing with a prior on market response. *Operations Research*, 58(1):16–29, 2010.
- A. Garivier and E. Moulines. On Upper-Confidence bound policies for switching bandit problems algorithmic learning theory. volume 6925 of *Lecture Notes in Computer Science*, chapter 16, pages 174–188. Springer Berlin / Heidelberg, Berlin, Heidelberg, 2011.
- B. I. Godoy, G. C. Goodwin, J. C. Aguero, and A. J. Rojas. An algorithm for estimating time-varying commodity price models. In *Proceedings of the 48th IEEE Conference on Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference*, pages 1563–1568. IEEE, 2009.
- D. M. Hanssens, L. J. Parsons, and R. L. Schultz. *Market response models: econometric and time series analysis*. International series in quantitative marketing. Kluwer Academic Publishers, Boston, second edition, 2001.
- J. M. Harrison, N. B. Keskin, and A. Zeevi. Dynamic pricing with an unknown linear demand model: asymptotically optimal semi-myopic policies. Working paper, <http://faculty-gsb.stanford.edu/harrison/Documents/hkz-2.pdf>, 2011.
- J. M. Harrison, N. B. Keskin, and A. Zeevi. Bayesian dynamic pricing policies: learning and earning under a binary prior distribution. *Management Science*, 58(3):570–586, 2012.
- D. Horsky and L. S. Simon. Advertising and the diffusion of new products. *Marketing Science*, 2(1):1–17, 1983.

- W. T. Huh and P. Rusmevichientong. A nonparametric asymptotic analysis of inventory planning with censored demand. *Mathematics of Operations Research*, 34(1):103–123, 2009.
- W. T. Huh, R. Levi, P. Rusmevichientong, and J. B. Orlin. Adaptive Data-Driven inventory control with censored demand based on Kaplan-Meier estimator. *Operations Research*, 59(4): 929–941, 2011.
- S. Kalish. Monopolist pricing with dynamic demand and production cost. *Marketing Science*, 2(2):135–159, 1983.
- B. Kamrad, S. S. Lele, A. Siddique, and R. J. Thomas. Innovation diffusion uncertainty, advertising and pricing policies. *European Journal of Operational Research*, 164(3):829–850, 2005.
- R. Kleinberg and T. Leighton. The value of knowing a demand curve: bounds on regret for online posted-price auctions. In *Proceedings of the 44th IEEE Symposium on Foundations of Computer Science*, pages 594–605, 2003.
- P. S. H. Leeflang, T. H. A. Bijmolt, J. van Doorn, D. M. Hanssens, H. J. van Heerde, P. C. Verhoef, and J. E. Wieringa. Creating lift versus building the base: current trends in marketing dynamics. *International Journal of Research in Marketing*, 26(1):13–20, 2009.
- A. E. B. Lim and J. G. Shanthikumar. Relative entropy, exponential utility, and robust dynamic pricing. *Operations Research*, 55(2):198–214, 2007.
- K. Y. Lin. Dynamic pricing with real-time demand learning. *European Journal of Operational Research*, 174(1):522–538, 2006.
- M. S. Lobo and S. Boyd. Pricing and learning with uncertain demand. Working paper, http://www.stanford.edu/~boyd/papers/pdf/pric_learn_unc_dem.pdf, 2003.
- T. Puu. Chaos in duopoly pricing. *Chaos, Solitons & Fractals*, 1(6):573–581, 1991.
- K. Raman and R. Chatterjee. Optimal monopolist pricing under demand uncertainty in dynamic markets. *Management Science*, 41(1):144–162, 1995.
- B. Robinson and C. Lakhani. Dynamic price models for new-product planning. *Management Science*, 21(10):1113–1122, 1975.
- H. Schultz. The statistical law of demand as illustrated by the demand for sugar. *Journal of Political Economy*, 33(6):481–504, 1925.
- J. Tuinstra. A price adjustment process in a model of monopolistic competition. *International Game Theory Review*, 6(3):417–442, 2004.
- A. R. Wildt and R. S. Winer. Modeling and estimation in changing market environments. *The Journal of Business*, 56(3):365–388, 1983.