Dynamic Pricing; A Learning Approach

Dimitris Bertsimas and Georgia Perakis

Massachusetts Institute of Technology,

77 Massachusetts Avenue, Room E53-359.

Cambridge, MA 02139.

Phones: (617) 253-8277 (617)-253-4223

Email: georgiap@mit.edu dbertsim@mit.edu

August, 2001

Dynamic Pricing; A Learning Approach

Abstract

We present an optimization approach for jointly learning the demand as a function of price, and dynamically setting prices of products in an oligopoly environment in order to maximize expected revenue. The models we consider do not assume that the demand as a function of price is known in advance, but rather assume parametric families of demand functions that are learned over time. We first consider the noncompetitive case and present dynamic programming algorithms of increasing computational intensity with incomplete state information for jointly estimating the demand and setting prices as time evolves. Our computational results suggest that dynamic programming based methods outperform myopic policies often significantly. We then extend our analysis in a competitive environment with two firms. We introduce a more sophisticated model of demand learning, in which the price elasticities are slowly varying functions of time, and allows for increased flexibility in the modeling of the demand. We propose methods based on optimization for jointly estimating the Firm's own demand, its competitor's demand, and setting prices. In preliminary computational work, we found that optimization based pricing methods offer increased expected revenue for a firm independently of the policy the competitor firm is following.

1 Introduction

In this paper we study pricing mechanisms for firms competing for the same products in a dynamic environment. Pricing theory has been extensively studied by researchers from a variety of fields over the years. These fields include, among others, economics (see for example, [36]), marketing (see for example, [25]), revenue management (see for example, [27]) and telecommunications (see for example, [21], [22], [29], [32], [33]). In recent years, the rapid development of information technology, the Internet and E-commerce has had very strong influence on the development of pricing and revenue management.

The overall goal of this paper is to address the problem of setting prices for a firm in both noncompetitive and competitive environments, in which the demand as a function of price is not known, but is learned over time. A firm produces a number of products which require (and compete for in the competitive case) scarce resources. The products must be priced dynamically over a finite time horizon, and sold to the appropriate demand. Our research (contrasted with traditional revenue management) considers pricing decisions, and takes capacity as given.

Problem Characteristics

The pricing problem we will focus on in this paper has a number of characteristics:

(a) The demand as a function of price is *unknown* a priori and is learned over time. As a result, part of the model we develop in this paper deals with learning the demand as the firm acquires more information over time. That is, we exploit the fact that over time firms are able to acquire knowledge regarding demand behavior that can be utilized to improve profitability. Much of the current research does not consider this aspect but rather considers demand to be an exogenous stochastic process following a certain distribution. See [7], [8], [10], [11], [16], [17], [19], [29].

(b) Products are priced *dynamically* over a finite time horizon. This is an important aspect since the demand and the data of the problem evolve dynamically. There exists a great deal of research that does not consider the dynamic and the competitive aspects of the pricing problem jointly. An exception to this involves some work that applies differential game theory (see [1], [2], [9]).

(c) We explicitly allow competition in an *oligopolistic* market, that is, a market characterized by a few firms on the supply side, and a large number of buyers on the demand side. A key feature of such a market (in contrast to a monopoly) is that the profit one firm receives depends not just on the prices it sets, but also on the prices set by the competing firms. That is, there is no perfect competition in an oligopolistic market since decisions made by all the firms in the market impact the profits received by each firm. One can consider a cooperative oligopoly (where firms collude) or a noncooperative oligopoly. In this paper we focus on the latter. The theory of oligopoly dates back to the work of Augustin Cournot [12], [13], [14].

(d) We consider products that are *perishable*, that is, there is a finite horizon to sell the products, after which any unused capacity is lost. Moreover, *the marginal cost* of an extra unit of demand is relatively *small*. For this reason, our models in this paper ignore the cost component in the decision-making process and refer to revenue maximization rather than profit maximization.

Application Areas

There are many markets where the framework we consider in this paper applies. Examples include airline ticket pricing. In this market the products the consumers demand, are the origin-destination (O-D) pairs during a particular time window. The resources are the flight legs (more appropriately seats on a particular flight leg) which have limited capacity. There is a finite horizon to sell the products, after which any unused capacity is lost (perishable products). The airlines compete with one another for the product demand which is of stochastic nature. Other industries sharing the same features include the service industry (for example, hotels, car rentals, and cruise-lines), the retail industry (for example, department stores) and finally, pricing in an e-commerce environment. All these industries attempt to intelligently match capacity with demand via *revenue management*. A review of the huge literature in revenue management can be found in [27], [34] and [35].

Contributions

(a) We develop pricing mechanisms when there is incomplete demand information, by jointly setting prices and learning the firm's demand without assuming any knowledge of it in advance.

(b) We introduce a model of demand learning, in which the price elasticities are slowly varying functions of time. This model allows for increased flexibility in the modeling of the demand. We propose methods based on optimization for jointly estimating the Firm's own demand, its competitor's demand, and setting prices.

Structure

The remainder of this paper is organized as follows. In Section 2, we focus on the dynamic pricing problem in a non-competitive environment. We consider jointly the problem of demand estimation and pricing using ideas from dynamic programming with incomplete state information. We present an exact algorithm as well as several heuristic algorithms that are easy to implement and discuss the various resulting pricing policies. In Section 3, we extend our previous model to also incorporate the aspect of competition. We propose an optimization approach to perform the firm's own demand estimation, its competitor's price prediction and finally its own price setting. Finally, in Section 4, we conclude with conclusions and open questions.

2 Pricing in a Noncompetitive Environment

In this section we consider the dynamic pricing problem in a non-competitive environment. We focus on a market with a single product and a single firm with overall capacity c over a time horizon T. In the beginning of each period t, the firm knows the previous price and demand realizations, that is, $d_1, \ldots, d_{t-1}$ and $p_1, \ldots, p_{t-1}$. This is the data available to the firm. In this section, we assume that the firm's true demand is an unknown linear function of the form

$$d_t = \beta^0 + \beta^1 p_t + \epsilon_t,$$

that is, it depends on the current period prices p_t, unknown parameters β^0, β^1 and a random noise $\epsilon_t \sim N(0, \sigma^2)$. The firm's objectives are to estimate its demand dynamically and set prices in order to maximize its total expected revenue. Let $\mathcal{P} = [p_{\min}, p_{\max}]$ be the set of feasible prices.

This section is organized as follows. In Section 2.1 we present a demand estimation model. In Section 2.2, we consider the joint demand estimation and pricing problem through a dynamic programming formulation. Using ideas from dynamic programming with incomplete state information, we are able to reduce this dynamic programming formulation to an eight-dimensional one. Nevertheless, this formulation is still difficult to solve, and we propose an approximation that allows us to further reduce the problem to a five dimensional dynamic program. In Section 2.3 we separate the demand estimation from the pricing problem and consider several heuristic algorithms. In particular, we consider a one-dimensional dynamic programming heuristic as well as a myopic policy heuristic. To gain intuition, we find closed form solutions in the deterministic case. Finally, in Section 2.4 we consider some examples and offer insights.

2.1 Demand Estimation

As we mentioned at time t the firm has observed the previous price and demand realizations, that is, $d_1, \ldots, d_{t-1}$ and $p_1, \ldots, p_{t-1}$ and assumes a linear demand model $d_t = \beta^0 + \beta^1 p_t + \epsilon_t$, with $\epsilon_t \sim N(0, \sigma^2)$. The parameters β^0, β^1 and σ are unknown and are estimated as follows.

We denote by $\mathbf{x}_s = [1,\ p_s]'$ and by $\widehat{\beta}_s$ the vector of the parameter estimates at time s, $(\widehat{\beta}_s^0, \widehat{\beta}_s^1)$. We estimate this vector of the demand parameters through the solution of the least square problem,

$$\widehat{\beta}_t = \arg\min_{\mathbf{r} \in \Re^2} \sum_{s=1}^{t-1} (d_s - \mathbf{x}_s' \mathbf{r})^2, \qquad t = 3, \ldots, T. \tag{1}$$

Proposition 1 *: The least squares estimates (1) can be generated by the following iterative process*

$$\widehat{\beta}_t = \widehat{\beta}_{t-1} + \mathbf{H}_{t-1}^{-1}\mathbf{x}_{t-1}\left(d_{t-1} - \mathbf{x}'_{t-1}\widehat{\beta}_{t-1}\right), \qquad t = 3, \ldots, T$$

where $\widehat{\beta}_2$ is an arbitrary vector, and the matrices H_{t-1} are generated by

$$\mathbf{H}_{t-1} = \mathbf{H}_{t-2} + \mathbf{x}_{t-1}\mathbf{x}'_{t-1}, \qquad t = 3, \ldots, T,$$

with $\mathbf{H}_1 = \begin{bmatrix} 1 & p_1 \\ p_1 & p_1^2 \end{bmatrix}$. *Therefore,* $\mathbf{H}_{t-1} = \begin{bmatrix} t-1 & \sum_{s=1}^{t-1} p_s \\ \sum_{s=1}^{t-1} p_s & \sum_{s=1}^{t-1} p_s^2 \end{bmatrix}$.

Proof: The first order conditions of the least squares problem for $\widehat{\beta}_t$ and $\widehat{\beta}_{t-1}$ respectively, imply that

$$\sum_{s=1}^{t-1}\left(d_s - \mathbf{x}'_s\widehat{\beta}_t\right)\mathbf{x}_s = \mathbf{0} \tag{2}$$

$$\sum_{s=1}^{t-2}\left(d_s - \mathbf{x}'_s\widehat{\beta}_{t-1}\right)\mathbf{x}_s = \mathbf{0}. \tag{3}$$

If we write, $\widehat{\beta}_t = \widehat{\beta}_{t-1} + \mathbf{a}$, where $\mathbf{a}$ is some vector, it follows from (2) that

$$\sum_{s=1}^{t-1}\left(d_s - \mathbf{x}'_s\widehat{\beta}_{t-1} - \mathbf{x}'_s\mathbf{a}\right)\mathbf{x}_s = \mathbf{0}.$$

This in turn implies that,

$$\sum_{s=1}^{t-2}\left(d_s - \mathbf{x}'_s\widehat{\beta}_{t-1} - \mathbf{x}'_s\mathbf{a}\right)\mathbf{x}_s + \left(d_{t-1} - \mathbf{x}'_{t-1}\widehat{\beta}_{t-1} - \mathbf{x}'_{t-1}\mathbf{a}\right)\mathbf{x}_{t-1} = \mathbf{0}. \tag{4}$$

Subtracting (3) from (4) we obtain that

$$\sum_{s=1}^{t-1}\left(\mathbf{x}'_s\mathbf{a}\right)\mathbf{x}_s = \left(d_{t-1} - \mathbf{x}'_{t-1}\widehat{\beta}_{t-1}\right)\mathbf{x}_{t-1}.$$

Therefore, $\mathbf{a} = \mathbf{H}_{t-1}^{-1}\mathbf{x}_{t-1}\left(d_{t-1} - \mathbf{x}'_{t-1}\widehat{\beta}_{t-1}\right)$, with $\mathbf{H}_{t-1} = \sum_{s=1}^{t-1}\left(\mathbf{x}_s\mathbf{x}'_s\right) = \begin{bmatrix} t-1 & \sum_{s=1}^{t-1} p_s \\ \sum_{s=1}^{t-1} p_s & \sum_{s=1}^{t-1} p_s^2 \end{bmatrix}$.

■

Given $d_1, \ldots, d_{t-1}$ and $p_1, \ldots, p_{t-1}$, the least squares estimates are

$$\widehat{\beta}_t^1 = \frac{(t-1)\sum_{s=1}^{t-1} p_s d_s - \sum_{s=1}^{t-1} p_s \sum_{s=1}^{t-1} d_s}{(t-1)\sum_{s=1}^{t-1} p_s^2 - \left(\sum_{s=1}^{t-1} p_s\right)^2}, \qquad \widehat{\beta}_t^0 = \frac{\sum_{s=1}^{t-1} d_s}{t-1} - \widehat{\beta}_t^1 \frac{\sum_{s=1}^{t-1} p_s}{t-1}.$$

The matrix $\mathbf{H}_{t-1}$ is singular, and hence not invertible, when

$$t\sum_{s=1}^{t-1} p_s^2 = \left(\sum_{s=1}^{t-1} p_s\right)^2. \tag{5}$$

Notice that the only solution to the above equality is $p_1 = p_2 = \cdots = p_{t-1}$. If the matrix $\mathbf{H}_{t-1}$ is nonsingular, then the inverse is

$$\mathbf{H}_{t-1}^{-1} = \begin{bmatrix} \frac{\sum_{s=1}^{t-1} p_s^2}{(t-1)\sum_{s=1}^{t-1} p_s^2 - \left(\sum_{s=1}^{t-1} p_s\right)^2} & \frac{-\sum_{s=1}^{t-1} p_s}{(t-1)\sum_{s=1}^{t-1} p_s^2 - \left(\sum_{s=1}^{t-1} p_s\right)^2} \\ \frac{-\sum_{s=1}^{t-1} p_s}{(t-1)\sum_{s=1}^{t-1} p_s^2 - \left(\sum_{s=1}^{t-1} p_s\right)^2} & \frac{t-1}{(t-1)\sum_{s=1}^{t-1} p_s^2 - \left(\sum_{s=1}^{t-1} p_s\right)^2} \end{bmatrix}.$$

Therefore,

$$\mathbf{H}_{t-1}^{-1}\mathbf{x}_{t-1} = \begin{bmatrix} \frac{\sum_{s=1}^{t-1} p_s^2}{(t-1)\sum_{s=1}^{t-1} p_s^2 - \left(\sum_{s=1}^{t-1} p_s\right)^2} & \frac{-\sum_{s=1}^{t-1} p_s}{(t-1)\sum_{s=1}^{t-1} p_s^2 - \left(\sum_{s=1}^{t-1} p_s\right)^2} \\ \frac{-\sum_{s=1}^{t-1} p_s}{(t-1)\sum_{s=1}^{t-1} p_s^2 - \left(\sum_{s=1}^{t-1} p_s\right)^2} & \frac{t-1}{(t-1)\sum_{s=1}^{t-1} p_s^2 - \left(\sum_{s=1}^{t-1} p_s\right)^2} \end{bmatrix} \begin{bmatrix} 1 \\ p_{t-1} \end{bmatrix} = \begin{bmatrix} \frac{\sum_{s=1}^{t-2} p_s^2 - p_{t-1}\sum_{s=1}^{t-2} p_s}{(t-1)\sum_{s=1}^{t-1} p_s^2 - \left(\sum_{s=1}^{t-1} p_s\right)^2} \\ \frac{(t-2)p_{t-1} - \sum_{s=1}^{t-2} p_s}{(t-1)\sum_{s=1}^{t-1} p_s^2 - \left(\sum_{s=1}^{t-1} p_s\right)^2} \end{bmatrix}.$$

As a result, we can express the estimates of the demand parameters in period t in terms of earlier estimates as

$$\begin{bmatrix} \widehat{\beta}_t^0 \\ \widehat{\beta}_t^1 \end{bmatrix} = \begin{bmatrix} \widehat{\beta}_{t-1}^0 \\ \widehat{\beta}_{t-1}^1 \end{bmatrix} + \left(d_{t-1} - \widehat{\beta}_{t-1}^0 - \widehat{\beta}_{t-1}^1 p_{t-1}\right) \begin{bmatrix} \frac{\sum_{s=1}^{t-2} p_s^2 - p_{t-1}\sum_{s=1}^{t-2} p_s}{(t-1)\sum_{s=1}^{t-1} p_s^2 - \left(\sum_{s=1}^{t-1} p_s\right)^2} \\ \frac{(t-2)p_{t-1} - \sum_{s=1}^{t-2} p_s}{(t-1)\sum_{s=1}^{t-1} p_s^2 - \left(\sum_{s=1}^{t-1} p_s\right)^2} \end{bmatrix}.$$

The estimate for the variance σ at time t is given by

$$\widehat{\sigma}_t^2 \;=\; \sum_{\tau=1}^{t-1} \frac{\left(d_\tau - \widehat{\beta}_t^0 - \widehat{\beta}_t^1 p_\tau\right)^2}{t-3}.$$

Notice that the variance estimate is based on $t-1$ pieces of data, with two parameters already estimated from the data, hence there are $t-3$ degrees of freedom. Such an estimate is unbiased (see [30]).

2.2 An Eight-Dimensional DP for Determining Pricing Policies

The difficulty in coming up with a general framework for dynamically determining prices is that the parameters β^0 and β^1 of the true demand are not directly observable. What is observable though are the realizations of demand and price in the previous periods, that is, $d_1, ..., d_{t-1}$ and $p_1, ..., p_{t-1}$. This seems to suggest that ideas from dynamic programming with incomplete state information may be useful (see [3]). As a first step in this direction, during the current period t, we consider a dynamic program with state space $(d_1, ..., d_{t-1}, p_1, ..., p_{t-1}, c_t)$, control variable the current price p_t and randomness coming from the noise ϵ_t. We observe though that as time t increases, the dimension of the state space becomes huge and therefore, solving this dynamic programming formulation is not possible. In what follows we will illustrate that we can considerably reduce the high dimensionality of the state space.

First we introduce the notation, $\widehat{\beta}_{s,t} = \left(\widehat{\beta}_{s,t}^0,\ \widehat{\beta}_{s,t}^1\right)$, $s = t, \ldots, T$, which is the current time t estimate of the parameters for future times $s = t, \ldots, T$. Notice that $\widehat{\beta}_{t,t} = \widehat{\beta}_t$. Similarly to Proposition 1, we can update our least squares estimates through $\widehat{\beta}_{t+1,t} = \widehat{\beta}_{t,t} + \mathbf{H}_t^{-1}\mathbf{x}_t \left(\widehat{D}_t - \mathbf{x}_t'\widehat{\beta}_{t,t}\right)$. Notice that since in the beginning of period t demand d_t is not known, we replaced it with $\widehat{D}_t = \widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_t + \varepsilon_t$. As a result, vector $\widehat{\beta}_{t+1,t}$ is a random variable. A useful observation we need to make is that in order to calculate matrix $\mathbf{H}_t$ we need to keep track of the quantities $\sum_{\tau=1}^{t-1} p_\tau^2$ and $\sum_{\tau=1}^{t-1} p_\tau$. These will be as a result part of the state space in the new dynamic programming formulation.

It is natural to assume that the variance estimates change with time and do not remain constant in future periods. This is the case since the estimate of the variance will be affected by the prices. That is,

$$\begin{aligned}\varepsilon_s &\sim N\left(0,\ \widehat{\sigma}_s^2\right)\\ \widehat{\sigma}_s^2 &= \frac{\sum\limits_{\tau=1}^{s-1}\left(d_\tau-\widehat{\beta}_s^0-\widehat{\beta}_s^1 p_\tau\right)^2}{s-3}, \qquad s=t,\dots,T.\end{aligned}$$

This observation implies that we need to find a way to estimate the variance for the future periods from the current one. We denote by $\widehat{\sigma}^2_{t+1,\,t}$ the estimate (in the current period, t) of next period's variance.

Proposition 2 *: The estimate of next period's variance in the current period t is given by,*

$$\begin{aligned}\widehat{\sigma}^2_{t+1,\,t} = {} & \frac{\widehat{\sigma}_t^2(t-3)+2\widehat{\beta}_t^0\sum\limits_{s=1}^{t-1}d_s+2\widehat{\beta}_t^1\sum\limits_{s=1}^{t-1}d_s p_s-(t-1)\left(\widehat{\beta}_t^0\right)^2-2\widehat{\beta}_t^0\widehat{\beta}_t^1\sum\limits_{s=1}^{t-1}p_s-\left(\widehat{\beta}_t^1\right)^2\sum\limits_{s=1}^{t-1}p_s^2}{t-2}+\\ & \frac{\left(\widehat{\beta}_t^0\right)^2+\left(\widehat{\beta}_t^1 p_t\right)^2+\varepsilon_t^2+2\widehat{\beta}_t^0\widehat{\beta}_t^1 p_t+2\widehat{\beta}_t^0\varepsilon_t+2\widehat{\beta}_t^1 p_t\varepsilon_t-2\widehat{\beta}_{t+1}^0\sum\limits_{s=1}^{t-1}d_s}{t-2}+\\ & \frac{-2\widehat{\beta}_{t+1}^0\widehat{\beta}_t^0-2\widehat{\beta}_{t+1}^0\widehat{\beta}_t^1 p_t-2\widehat{\beta}_{t+1}^0\varepsilon_t-2\widehat{\beta}_{t+1}^1\sum\limits_{s=1}^{t-1}p_s d_s}{t-2}+\\ & \frac{-2\widehat{\beta}_{t+1}^1\widehat{\beta}_t^0 p_t-2\widehat{\beta}_{t+1}^1\widehat{\beta}_t^1 p_t^2-2\widehat{\beta}_{t+1}^1 p_t\varepsilon_t+t\left(\widehat{\beta}_{t+1}^0\right)^2}{t-2}+\\ & \frac{2\widehat{\beta}_{t+1}^0\widehat{\beta}_{t+1}^1\sum\limits_{s=1}^{t-1}p_s+2\widehat{\beta}_{t+1}^0\widehat{\beta}_{t+1}^1 p_t+\left(\widehat{\beta}_{t+1}^1\right)^2\sum\limits_{s=1}^{t-1}p_s^2+\left(\widehat{\beta}_{t+1}^1\right)^2 p_t^2}{t-2}.\end{aligned} \tag{6}$$

Proof: As a first step we relate quantities $\widehat{\sigma}_t^2=\frac{\sum_{s=1}^{t-1}\left(d_s-\widehat{\beta}_t^0-\widehat{\beta}_t^1 p_s\right)^2}{t-3}$ with $\widehat{\sigma}_{t+1}^2=\frac{\sum_{s=1}^{t}\left(d_s-\widehat{\beta}_{t+1}^0-\widehat{\beta}_{t+1}^1 p_s\right)^2}{t-2}$. By expanding the second equation and separating the period t terms from the previous period $t-1$ we obtain

$$\begin{aligned}\widehat{\sigma}_{t+1}^2 = {} & \frac{\sum\limits_{s=1}^{t-1}d_s^2+d_t^2-2\widehat{\beta}_{t+1}^0\sum\limits_{s=1}^{t-1}d_s-2\widehat{\beta}_{t+1}^0 d_t-2\widehat{\beta}_{t+1}^1\sum\limits_{s=1}^{t-1}p_s d_s-2\widehat{\beta}_{t+1}^1 p_t d_t}{t-2}+\\ & \frac{t\left(\widehat{\beta}_{t+1}^0\right)^2+2\widehat{\beta}_{t+1}^0\widehat{\beta}_{t+1}^1\sum\limits_{s=1}^{t-1}p_s+2\widehat{\beta}_{t+1}^0\widehat{\beta}_{t+1}^1 p_t+\left(\widehat{\beta}_{t+1}^1\right)^2\sum\limits_{s=1}^{t-1}p_s^2+\left(\widehat{\beta}_{t+1}^1\right)^2 p_t^2}{t-2}.\end{aligned} \tag{7}$$

Recall that $\widehat{\sigma}_t^2 = \frac{\sum_{s=1}^{t-1}\left(d_s - \widehat{\beta}_t^0 - \widehat{\beta}_t^1 p_s\right)^2}{t-3}$. This gives rise to,

$$\begin{aligned}\sum_{s=1}^{t-1} d_s^2 &= \widehat{\sigma}_t^2(t-3) + 2\widehat{\beta}_t^0 \sum_{s=1}^{t-1} d_s + 2\widehat{\beta}_t^1 \sum_{s=1}^{t-1} d_s p_s - (t-1)\left(\widehat{\beta}_t^0\right)^2 \\ &\quad -2\widehat{\beta}_t^0 \widehat{\beta}_t^1 \sum_{s=1}^{t-1} p_s - \left(\widehat{\beta}_t^1\right)^2 \sum_{s=1}^{t-1} p_s^2.\end{aligned} \tag{8}$$

We substitute (8) into (7) to obtain

$$\begin{aligned}\widehat{\sigma}_{t+1}^2 &= \frac{\widehat{\sigma}_t^2(t-3) + 2\widehat{\beta}_t^0 \sum_{s=1}^{t-1} d_s + 2\widehat{\beta}_t^1 \sum_{s=1}^{t-1} d_s p_s - (t-1)\left(\widehat{\beta}_t^0\right)^2 - 2\widehat{\beta}_t^0\widehat{\beta}_t^1 \sum_{s=1}^{t-1} p_s - \left(\widehat{\beta}_t^1\right)^2 \sum_{s=1}^{t-1} p_s^2}{t-2} + \\ &\quad \frac{d_t^2 - 2\widehat{\beta}_{t+1}^0 \sum_{s=1}^{t-1} d_s - 2\widehat{\beta}_{t+1}^0 d_t - 2\widehat{\beta}_{t+1}^1 \sum_{s=1}^{t-1} p_s d_s - 2\widehat{\beta}_{t+1}^1 p_t d_t + t\left(\widehat{\beta}_{t+1}^0\right)^2}{t-2} + \\ &\quad \frac{2\widehat{\beta}_{t+1}^0 \widehat{\beta}_{t+1}^1 \sum_{s=1}^{t-1} p_s + 2\widehat{\beta}_{t+1}^0 \widehat{\beta}_{t+1}^1 p_t + \left(\widehat{\beta}_{t+1}^1\right)^2 \sum_{s=1}^{t-1} p_s^2 + \left(\widehat{\beta}_{t+1}^1\right)^2 p_t^2}{t-2}.\end{aligned}$$

Nevertheless, in the beginning of period t, d_t is not known. Therefore, we replace in the previous equation, each occurrence of d_t with $\widehat{D}_t = \widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_t + \varepsilon_t$. This leads us to conclude that

$$\begin{aligned}\widehat{\sigma}_{t+1,\,t}^2 = {} & \frac{\widehat{\sigma}_t^2(t-3) + 2\widehat{\beta}_t^0 \sum_{s=1}^{t-1} d_s + 2\widehat{\beta}_t^1 \sum_{s=1}^{t-1} d_s p_s - (t-1)\left(\widehat{\beta}_t^0\right)^2 - 2\widehat{\beta}_t^0\widehat{\beta}_t^1 \sum_{s=1}^{t-1} p_s - \left(\widehat{\beta}_t^1\right)^2 \sum_{s=1}^{t-1} p_s^2}{t-2} + \\ & \frac{\left(\widehat{\beta}_t^0\right)^2 + \left(\widehat{\beta}_t^1 p_t\right)^2 + \varepsilon_t^2 + 2\widehat{\beta}_t^0\widehat{\beta}_t^1 p_t + 2\widehat{\beta}_t^0 \varepsilon_t + 2\widehat{\beta}_t^1 p_t \varepsilon_t - 2\widehat{\beta}_{t+1}^0 \sum_{s=1}^{t-1} d_s}{t-2} + \\ & \frac{-2\widehat{\beta}_{t+1}^0\widehat{\beta}_t^0 - 2\widehat{\beta}_{t+1}^0\widehat{\beta}_t^1 p_t - 2\widehat{\beta}_{t+1}^0 \varepsilon_t - 2\widehat{\beta}_{t+1}^1 \sum_{s=1}^{t-1} p_s d_s - 2\widehat{\beta}_{t+1}^1\widehat{\beta}_t^0 p_t - 2\widehat{\beta}_{t+1}^1\widehat{\beta}_t^1 p_t^2 - 2\widehat{\beta}_{t+1}^1 p_t \varepsilon_t + t\left(\widehat{\beta}_{t+1}^0\right)^2}{t-2} + \\ & \frac{2\widehat{\beta}_{t+1}^0 \widehat{\beta}_{t+1}^1 \sum_{s=1}^{t-1} p_s + 2\widehat{\beta}_{t+1}^0 \widehat{\beta}_{t+1}^1 p_t + \left(\widehat{\beta}_{t+1}^1\right)^2 \sum_{s=1}^{t-1} p_s^2 + \left(\widehat{\beta}_{t+1}^1\right)^2 p_t^2}{t-2}.\end{aligned}$$

■

This proposition suggests that in order to estimate the next period variance from the current one, we need to keep track of the following quantities

$$\widehat{\beta}_t^0,\ \widehat{\beta}_t^1,\ \sum_{\tau=1}^{t-1} p_\tau^2,\ \sum_{\tau=1}^{t-1} p_\tau,\ \sum_{\tau=1}^{t-1} p_\tau d_\tau,\ \sum_{\tau=1}^{t-1} d_\tau,\ \widehat{\sigma}_t^2.$$

This observation allows us to provide an eight-dimensional dynamic programming formulation with state space given by,

$$\left(c_s,\ \widehat{\beta}_s^0,\ \widehat{\beta}_s^1,\ \sum_{\tau=1}^{s-1} p_\tau^2,\ \sum_{\tau=1}^{s-1} p_\tau,\ \sum_{\tau=1}^{s-1} p_\tau d_\tau,\ \sum_{\tau=1}^{s-1} d_\tau,\ \widehat{\sigma}_s^2\right), \qquad s = t, \ldots, T.$$

We are now able to formulate the following dynamic program where the control is the price and the randomness is the noise.

An Eight-Dimensional DP Pricing Policy

$$J_T(c_T,\ \widehat{\beta}_T^0,\ \widehat{\beta}_T^1,\ \widehat{\sigma}_T^2) = \max_{p_T} E_{\varepsilon_T}\left[p_T \min\left\{\left(\widehat{\beta}_T^0 + \widehat{\beta}_T^1 p_T + \varepsilon_T\right)^+,\ c_T\right\}\right]$$

$$\begin{aligned}
for\ s &= \max\{3,\ t\}, \ldots, T-1 \\
&\quad J_s(c_s,\ \widehat{\beta}_s^0,\ \widehat{\beta}_s^1,\ \sum_{\tau=1}^{s-1} p_\tau^2,\ \sum_{\tau=1}^{s-1} p_\tau,\ \sum_{\tau=1}^{s-1} p_\tau d_\tau,\ \sum_{\tau=1}^{s-1} d_\tau,\ \widehat{\sigma}_s^2) \\
&= \max_{p_s} E_{\varepsilon_s}[p_s \min\left\{\left(\widehat{\beta}_s^0 + \widehat{\beta}_s^1 p_s + \varepsilon_s\right)^+,\ c_s\right\} \\
&\quad + J_{s+1}\left(\begin{array}{c}
c_s - \min\left\{\left(\widehat{\beta}_s^0 + \widehat{\beta}_s^1 p_s + \varepsilon_s\right)^+,\ c_s\right\}, \\
\widehat{\beta}_{s+1}^0,\ \widehat{\beta}_{s+1}^1, \\
\sum_{\tau=1}^{s-1} p_\tau^2 + p_s^2,\ \sum_{\tau=1}^{s-1} p_\tau + p_s, \\
\sum_{\tau=1}^{s-1} p_\tau d_\tau + p_s\left(\widehat{\beta}_s^0 + \widehat{\beta}_s^1 p_s + \varepsilon_s\right)^+ \\
\sum_{\tau=1}^{s-1} d_\tau + \left(\widehat{\beta}_s^0 + \widehat{\beta}_s^1 p_s + \varepsilon_s\right)^+, \\
\widehat{\sigma}_{s+1}^2
\end{array}\right)],
\end{aligned}$$

where

$$\begin{bmatrix} \widehat{\beta}^0_{s+1} \\ \widehat{\beta}^1_{s+1} \end{bmatrix} = \begin{bmatrix} \widehat{\beta}^0_{s} \\ \widehat{\beta}^1_{s} \end{bmatrix} + \varepsilon_s \begin{bmatrix} \dfrac{\sum_{\tau=1}^{s-1} p_\tau^2 - p_s \sum_{\tau=1}^{s-1} p_\tau}{s \sum_{\tau=1}^{s-1} p_\tau^2 + s p_s^2 - \left(\sum_{\tau=1}^{s-1} p_\tau + p_s \right)^2} \\ \dfrac{(s-1) p_s - \sum_{\tau=1}^{s-1} p_\tau}{s \sum_{\tau=1}^{s-1} p_\tau^2 + s p_s^2 - \left(\sum_{\tau=1}^{s-1} p_\tau + p_s \right)^2} \end{bmatrix},$$

with noise $\varepsilon_s \sim N(0, \widehat{\sigma}_s^2)$ and variance $\widehat{\sigma}_s^2$ given from the recursive formula in (6).

Notice that in the DP recursion s is ranging from $\max\{3,\, t\}$ to $T-1$. This is because in the expression for $\widehat{\sigma}^2_{s+2}$ we divide by $s-2$. Intuitively, we need at least three data points in order to estimate three parameters. When $t = 1$, the denominator in the expression for $\widehat{\sigma}^2_{t+1}$ should also be one, while when $t = 2$ the denominator can be chosen to be either one or two.

2.3 A Five-Dimensional DP for Determining Pricing Policies

Although the previous DP formulation is the correct framework for determining pricing policies, it has an eight-dimensional state space which makes the problem computationally intractable. For this reason we consider in this section an approximation that gives rise to a lower dimensional dynamic program that is computationally tractable. In particular, we relax the assumption that the noise at time t changes in time and is affected by future pricing decisions. In particular, we consider

$$\begin{aligned} \varepsilon_s &\sim N\left(0,\ \widehat{\sigma}_t^2\right), \qquad s = t, \ldots, T \\ \widehat{\sigma}_t^2 &= \frac{\sum_{\tau=1}^{t-1} \left(d_\tau - \widehat{\beta}^0_t - \widehat{\beta}^1_t p_\tau \right)^2}{t-3}. \end{aligned}$$

Moreover, as in the previous section

$$\widehat{\beta}_{t+1,t} = \widehat{\beta}_{t,t} + \mathbf{H}_t^{-1} \mathbf{x}_t \left(\widehat{d}_t - \mathbf{x}_t' \widehat{\beta}_{t,t} \right).$$

To calculate the matrix $\mathbf{H}_t$ we need to keep track of the quantities $\sum_{\tau=1}^{t-1} p_\tau^2$ and $\sum_{\tau=1}^{t-1} p_\tau$.

This gives rise to a dynamic programming formulation with state variables,

$$\left(c_s,\ \widehat{\beta}^0_s,\ \widehat{\beta}^1_s,\ \sum_{\tau=1}^{s-1} p_\tau^2,\ \sum_{\tau=1}^{s-1} p_\tau\right) \qquad s = t, \ldots, T. \tag{9}$$

A Five-Dimensional DP Pricing Policy

$$\begin{aligned} J_T\left(c_T,\ \widehat{\beta}^0_T,\ \widehat{\beta}^1_T\right) &= \max_{p_T \in \mathcal{P}} \mathrm{E}_{\varepsilon_T}\left[p_T \min\left\{\left(\widehat{\beta}^0_T + \widehat{\beta}^1_T p_T + \varepsilon_T\right)^+,\ c_T\right\}\right] \\ for\ s &= t, \ldots, T-1: \\ J_s\left(c_s,\ \widehat{\beta}^0_s,\ \widehat{\beta}^1_s,\ \sum_{\tau=1}^{s-1} p_\tau^2,\ \sum_{\tau=1}^{s-1} p_\tau\right) &= \max_{p_s \in \mathcal{P}} \mathrm{E}_{\varepsilon_s}[p_s \min\left\{\left(\widehat{\beta}^0_s + \widehat{\beta}^1_s p_s + \varepsilon_s\right)^+,\ c_s\right\} \\ &\quad + J_{s+1}\begin{pmatrix} c_s - \min\left\{\left(\widehat{\beta}^0_s + \widehat{\beta}^1_s p_s + \varepsilon_s\right)^+,\ c_s\right\}, \\ \widehat{\beta}^0_{s+1},\ \widehat{\beta}^1_{s+1}, \\ \sum_{\tau=1}^{s-1} p_\tau^2 + p_t^2,\ \sum_{\tau=1}^{s-1} p_\tau + p_t \end{pmatrix}], \end{aligned}$$

with

$$\begin{bmatrix} \widehat{\beta}^0_{s+1} \\ \widehat{\beta}^1_{s+1} \end{bmatrix} = \begin{bmatrix} \widehat{\beta}^0_s \\ \widehat{\beta}^1_s \end{bmatrix} + \varepsilon_s \begin{bmatrix} \dfrac{\sum_{\tau=1}^{s-1} p_\tau^2 - p_s \sum_{\tau=1}^{s-1} p_\tau}{s \sum_{\tau=1}^{s-1} p_\tau^2 + s p_s^2 - \left(\sum_{\tau=1}^{s-1} p_\tau + p_s\right)^2} \\ \dfrac{(s-1) p_s - \sum_{\tau=1}^{s-1} p_\tau}{s \sum_{\tau=1}^{s-1} p_\tau^2 + s p_s^2 - \left(\sum_{\tau=1}^{s-1} p_\tau + p_s\right)^2} \end{bmatrix}.$$

2.4 Pricing Heuristics

In the previous two subsections we considered two dynamic programming formulations for determining pricing policies. The first was an exact formulation with an eight-dimensional state space that was computationally intractable, while the second was an approximation with a five-dimensional state space that is more tractable. Nevertheless, although this latter approach is tractable it is still fairly complex to solve. Both of these formulations were based on the idea of performing jointly the demand estimation with the pricing problem. In this section, we consider two heuristics that are approximations but yet are computationally very easy to perform. They are based on the idea of separating the demand estimation from the pricing problem.

One-Dimensional DP Pricing Policy

In the beginning of period t, the firm computes the estimates $\widehat{\beta}_t^0$ and $\widehat{\beta}_t^1$ and solves a one-dimensional dynamic program assuming that these parameter estimates are valid over all future periods. That is, this heuristic approach ignores the fact that these estimates will in fact be affected by the current pricing decisions. In particular,

$$\begin{aligned}\widehat{d}_s &= \widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_s + \varepsilon_s, \quad s = t, \ldots, T \\ \varepsilon_s &\sim N\left(0,\ \widehat{\sigma}_t^2\right), \quad s = t, \ldots, T,\end{aligned}$$

with

$$\widehat{\sigma}_t^2 = \sum_{s=1}^{t-1} \frac{\left(d_s - \widehat{\beta}_t^0 - \widehat{\beta}_t^1 p_s\right)^2}{t-3}.$$

Subsequently, the firm solves the following dynamic program in the beginning of period t ($t = 1, \ldots, T$),

$$\begin{aligned}J_T(c_T) &= \max_{p_T \in \mathcal{P}} E_{\varepsilon_T}\left[p_T \min\left\{\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_T + \varepsilon_T\right)^+,\ c_T\right\}\right] \\ \text{for } s &= t, \ldots, T-1: \\ J_s(c_s) &= \max_{p_s \in \mathcal{P}} E_{\varepsilon_s}\left[\begin{array}{c} p_s \min\left\{\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_s + \varepsilon_s\right)^+,\ c_s\right\} + \\ J_{s+1}\left(c_s - \min\left\{\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_s + \varepsilon_s\right)^+,\ c_s\right\}\right)\end{array}\right].\end{aligned}$$

In this dynamic programming formulation the remaining capacity represents the state space, the prices are the controls and the randomness comes from the noise.

Deterministic One-Dimensional DP Policy

To gain some intuition, in what follows we examine the deterministic case (that is, when the noise $\varepsilon_s = 0$). As a result, after having computed the estimates $\widehat{\beta}_t^0$ and $\widehat{\beta}_t^1$, the firm solves the following DP in the beginning of period t ($t = 1, \ldots, T$),

$$\begin{aligned}J_T(c_T) &= \max_{p_T \in \mathcal{P}} p_T \min\left\{\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_T\right)^+,\ c_T\right\} \\ \text{for } s &= t, \ldots, T-1:\end{aligned}$$

$$
\begin{aligned}
J_s(c_s) &= \max_{p_s \in \mathcal{P}} p_s \min\left\{\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_s\right)^+,\ c_s\right\} \\
&\quad + J_{s+1}\left(c_s - \min\left\{\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_s\right)^+,\ c_s\right\}\right).
\end{aligned}
$$

This deterministic one-dimensional DP policy has a closed form solution. We establish its solution in two parts. Since the dynamic program is deterministic, an optimal solution is given by an open-loop policy (that is, we can solve for an optimal price path versus an optimal pricing policy, i.e. there is no dependence on the state). For the proofs that follow, we need to introduce the following definition.

Definition 1 *A price vector* $\boldsymbol{p} = (p_t, \ldots, p_T)'$ *leads to* **premature stock-out** *if*

$$
\sum_{s=t}^{T} \left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_s\right) > c_t.
$$

Lemma 1 *The optimal solution given by the one-dimensional DP is unique and satisfies* $p_t = \cdots = p_T$.

Proof: First we will show that any optimal solution must satisfy $p_t = \cdots = p_T$, then we will prove uniqueness. Suppose there exists an optimal solution $\boldsymbol{p}^*$ for which the above does not hold. Then at least two of the prices are different and at least one price is less than $p_{\max}$. Without loss of generality, assume that $p_t \neq p_{t+1}$ (the argument holds for any two prices). We will show that such a solution cannot be optimal. Next we will show that the optimal solution must satisfy,

$$
\sum_{s=t}^{T} \widehat{d}_s = \sum_{s=t}^{T} \left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_s^*\right) \leq c_t.
$$

This is true since otherwise we could increase at least one of the prices by a small amount (since at least one is strictly less than $p_{\max}$), and achieve greater revenue by selling the same number of units c_t at a slightly higher average price (contradicting the optimality of the solution). Therefore, the firm does not expect a premature stock-out and the optimal objective value is given by, $z^* = \sum_{s=t}^{T} p_s^* \left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_s^*\right)$. Notice that the revenue generated in periods t and $t+1$ is given by,

$$\begin{aligned} & p_t^* \left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_t^* \right) + p_{t+1}^* \left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_{t+1}^* \right) \\ = \quad & \widehat{\beta}_t^0 p_t^* + \widehat{\beta}_t^0 p_{t+1}^* + \widehat{\beta}_t^1 \left((p_t^*)^2 + (p_{t+1}^*)^2 \right). \end{aligned} \tag{10}$$

In what follows, consider setting price $\frac{p_t^* + p_{t+1}^*}{2}$ in periods t and $t+1$. Therefore, the revenue generated in periods t and $t+1$ is given by,

$$\widehat{\beta}_t^0 p_t^* + \widehat{\beta}_t^0 p_{t+1}^* + \frac{\widehat{\beta}_t^1}{2} \left(p_t^* + p_{t+1}^* \right)^2. \tag{11}$$

Comparing (11) with (10) we notice that the total revenue has been increased. This is a contradiction. Hence, any optimal solution must satisfy $p_t = \cdots = p_T$.

Next we demonstrate uniqueness. Suppose there exist two optimal solutions p^1 and p^2 of dimension $T-t+1$, where $p^1 = (p^1, \ldots, p^1)\,, \quad p^2 = (p^2, \ldots, p^2)\,.$

We consider three possibilities. First suppose that both price vectors lead to premature stock-out. The respective revenues are given by $c_t p^1$ and $c_t p^2$. Since $p^1 \neq p^2$, it follows that $c_t p^1 \neq c_t p^2$, (since $c_t > 0$). Therefore, it cannot be the case that both p^1 and p^2 are optimal (which is a contradiction).

Next suppose that exactly one price vector, say p^1, leads to premature stock-out. We know that for such a price vector to be optimal it must be the case that $p^1 = p_{\max}$, since otherwise we could increase p^1 by a small amount and improve the objective. Moreover, $p^2 < p_{\max}$ (since $p^2 \neq p^1$). Therefore p^2 also leads to premature stock out (contradicting the assumption that exactly one price vector leads to premature stock-out).

Finally suppose that neither price vector leads to premature stock-out. In this case, the respective revenue (objective) is given by,

$$z^1 = p^1 \left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p^1 \right) (T-t+1)\,, \quad z^2 = z^1 = p^2 \left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p^2 \right) (T-t+1)\,.$$

Consider the price vector p' (of dimension $T-t+1$) with each component given by, $\frac{p^1+p^2}{2}$. Since p^1 and p^2 do not lead to premature stock-out, neither does p'. In which case the revenue is given by,

$$z' = \frac{p^1+p^2}{2} \left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 \frac{p^1+p^2}{2} \right) (T-t+1)\,.$$

After some algebra (and since $z^2 = z^1$) we find that, $z' = z^1 - \frac{\widehat{\beta}_t^1}{4}\left(p^1 - p^2\right)^2 (T-t+1)$. Notice that $z' > z^1$. Therefore, p^1 and p^2 cannot be optimal (contradiction). Hence, the optimal solution is unique. ■

We use this result to prove the following theorem.

Theorem 2 *Under the assumption that $\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_{\max} > 0$ (that is, demand cannot be negative), in the deterministic case the one-dimensional DP offers the following closed form solution*

$$p_s^* = \max\left\{-\frac{\widehat{\beta}_t^0}{2\widehat{\beta}_t^1}, \frac{c_t - (T-t+1)\,\widehat{\beta}_t^0}{(T-t+1)\,\widehat{\beta}_t^1}\right\} \qquad s = t, \ldots, T.$$

However, if the above solution exceeds $p_{\max}$ then $p_s^* = p_{\max}$, while if the above solution is less than $p_{\min}$ then $p_s^* = p_{\min}$.

Proof: Consider the following price, $p^1 = \arg\max_{p \in \mathcal{P}} p\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p\right)$. Notice that since $\widehat{\beta}_t^1 < 0$ and the price set is continuous,

$$p^1 = \begin{cases} -\frac{\widehat{\beta}_t^0}{2\widehat{\beta}_t^1} & \text{if } \mathrm{p_{min}} \leq -\frac{\widehat{\beta}_\mathrm{t}^0}{2\widehat{\beta}_\mathrm{t}^1} \leq \mathrm{p_{max}} \\ p_{\min} & \text{if } -\frac{\widehat{\beta}_\mathrm{t}^0}{2\widehat{\beta}_\mathrm{t}^1} < \mathrm{p_{min}} \\ p_{\max} & \text{if } -\frac{\widehat{\beta}_\mathrm{t}^0}{2\widehat{\beta}_\mathrm{t}^1} > \mathrm{p_{max}} \end{cases}$$

The objective value z (total revenue) is the sum of each period's revenue. Letting z_s denote the revenue from period s, implies that $z_s \leq p^1\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p^1\right)$, for all $\mathrm{s} = \mathrm{t}, \ldots, \mathrm{T}$. Therefore, the total revenue is bounded, $z \leq p^1\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p^1\right)(T-t+1)$. We consider three cases:

CASE 1: Suppose that $\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p^1\right)(T-t+1) \leq c_t$. In this case the firm could set the price p^1 over each period and achieve revenue $p^1\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p^1\right)(T-t+1)$. Therefore, the objective's upper bound has been achieved and hence the solution $(p^1, \ldots, p^1)$ is optimal.

CASE 2: Suppose that $\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_{\max}\right)(T-t+1) > c_t$. In this case the solution $(p_{\max}, \ldots, p_{\max})$ has an associated objective value of $c_t p_{\max}$, which is clearly an upper bound on the objective. Therefore the solution $(p_{\max}, \ldots, p_{\max})$ is optimal.

CASE 3: Suppose that $\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p^1\right)(T-t+1) > c_t$ and $\left(\widehat{\beta}_\mathrm{t}^0 + \widehat{\beta}_\mathrm{t}^1 \mathrm{p_{max}}\right)(\mathrm{T}-\mathrm{t}+1) \leq \mathrm{c_t}$.

In this case the solution $(p^1,\ldots,p^1)$ cannot be optimal, since we could then increase at least one of the prices by a small amount ($p^1 < p_{\max}$), and achieve greater revenue by selling the same number of units c_t at a slightly higher average price. However, the previous lemma suggests that the unique optimal solution (of dimension $T-t+1$) has constant prices $p^* = (p^*,\ldots,p^*)$. Furthermore, we know that $\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p^*\right)(T-t+1) \le c_t$. Otherwise, as before, we could increase p^* by a small amount and achieve greater revenue by selling the same number of units c_t at a slightly higher price. Since,

$$\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p^1\right)(T-t+1) > c_t \quad \text{and} \quad \left(\widehat{\beta}_{\mathrm{t}}^0 + \widehat{\beta}_{\mathrm{t}}^1 \mathrm{p}_{\max}\right)(\mathrm{T}-\mathrm{t}+1) \le \mathrm{c}_{\mathrm{t}},$$

there exists a price p' such that $p^1 < p' \le p_{\max}$ and $\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p'\right)(T-t+1) = c_t$.

Intuitively, this is the price which will sell off exactly all of the firm's remaining inventory at the end of the horizon. Now consider the objective function as a function of the static price p. For $p_{\min} < p < p'$ the objective is given by $c_t p$ (since the firm stocks out before the end of the planning horizon) which is increasing in p. For $p^1 < p' \le p \le p_{\max}$ the objective is given by, $p\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p\right)(T-t+1)$. This is true because for these prices the firm does not stock out early, and each period's revenue is simply the product of price and demand. Now notice the above function is decreasing for all $p > p^1$. Furthermore, p' satisfies

$$p'\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p'\right)(T-t+1) = c_t p'.$$

We conclude that p' is the optimal solution in this case. Notice that solving for $p' = p^*$ one obtains,

$$p' = \frac{c_t - (T-t+1)\,\widehat{\beta}_t^0}{(T-t+1)\,\widehat{\beta}_t^1}.$$

■

We note that in the deterministic case the policies given by the one and five-dimensional DPs are equivalent. This follows since in the deterministic case $\varepsilon_s = 0$ and as a result, the future demand parameter estimates are not affected by the current pricing decision. Hence, $\begin{bmatrix} \widehat{\beta}_{s+1}^0 \\ \widehat{\beta}_{s+1}^1 \end{bmatrix} = \begin{bmatrix} \widehat{\beta}_s^0 \\ \widehat{\beta}_s^1 \end{bmatrix}$. Therefore, the five-dimensional DP can be reduced to the following three dimensional DP,

$$
\begin{aligned}
J_T\left(c_T,\, \widehat{\beta}_T^0,\, \widehat{\beta}_T^1\right) &= \max_{p_T \in \mathcal{P}} p_T \min\left\{\left(\widehat{\beta}_T^0 + \widehat{\beta}_T^1 p_T\right)^+,\ c_T\right\} \\
\text{for } s &= t, \ldots, T-1: \\
J_s\left(c_s,\, \widehat{\beta}_s^0,\, \widehat{\beta}_s^1\right) &= \max_{p_s \in \mathcal{P}} p_s \min\left\{\left(\widehat{\beta}_s^0 + \widehat{\beta}_s^1 p_s\right)^+,\ c_s\right\} \\
&\quad + J_{s+1}\left(c_s - \min\left\{\left(\widehat{\beta}_s^0 + \widehat{\beta}_s^1 p_s\right)^+,\ c_s\right\},\ \widehat{\beta}_s^0,\, \widehat{\beta}_s^1\right).
\end{aligned}
$$

Moreover, notice that the one-dimensional DP policy in the deterministic case is given by,

$$
\begin{aligned}
J_T(c_T) &= \max_{p_T \in \mathcal{P}} p_T \min\left\{\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_T\right)^+,\ c_T\right\} \\
\text{for } s &= t, \ldots, T-1: \\
J_s(c_s) &= \max_{p_s \in \mathcal{P}} p_s \min\left\{\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_s\right)^+,\ c_s\right\} \\
&\quad + J_{s+1}\left(c_s - \min\left\{\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_s\right)^+,\ c_s\right\}\right).
\end{aligned}
$$

When the firm uses the five-dimensional DP policy, since in the beginning of period t, $\left(\widehat{\beta}_s^0,\, \widehat{\beta}_s^1\right) = \left(\widehat{\beta}_t^0,\, \widehat{\beta}_t^1\right)$, for all $s = t, \ldots, T$, it follows, just like in the case of the one-dimensional DP policy, that the current parameter estimates are valid over all future periods. The DPs solved for both policies are in that case equivalent. The only difference is that the five-dimensional DP *explicitly* treats $\widehat{\beta}_t^0$ and $\widehat{\beta}_t^1$ as (constant) states while the one-dimensional DP *implicitly* treats $\widehat{\beta}_t^0$ and $\widehat{\beta}_t^1$ as (constant) states. This observation leads us to conclude that the two policies are equivalent.

The Myopic Pricing Policy

Finally, we introduce the last heuristic pricing policy, the myopic pricing policy. This policy maximizes the expected current period revenue over each period, without considering future implications of the pricing decisions. In period t $(t = 1, 2, \ldots, T)$

$$
p_t \in \arg\max_{p \in \mathcal{P}} p \mathrm{E}_{\varepsilon_t}\left[\min\left\{\left(\widehat{\beta}_t^0 + \widehat{\beta}_t^1 p + \widehat{\varepsilon}_t\right)^+,\ c_t\right\}\right],
$$

where $a^+ = \max(a,\ 0)$. Quantity c_t denotes the remaining capacity in the beginning of period t. Clearly the myopic policy is suboptimal since it does not take into account the number of periods left in the planning horizon. However, when capacity is sufficiently large the expected revenue obtained through the myopic and the one-dimensional DP policy become the same. This follows from the observation that when capacity is sufficiently large, both methods maximize current expected revenue. This myopic approach is optimal since the firm does not run the risk of stocking out before the end of the planning horizon that is, there are no future implications of the current pricing decision.

2.5 Computational Results

In the previous subsections we introduced dynamic pricing policies for revenue maximization with incomplete demand information based on DP (one, five and eight dimensional) as well as a myopic policy which we consider as a benchmark. We have implemented all methods except the eight-dimensional DP, which is outside today's computational capabilities.

We consider an example where true demand is given by $d_t = 60 - p_t + \varepsilon_t$, with $\varepsilon_t = 0$ initially and $\varepsilon_t \sim N(0, \sigma^2)$, where $\sigma = 4$. The prices belong in the set $\mathcal{P} = \{20,\ 21, \ldots,\ 40\}$, the total capacity is $c = 400$ and the time horizon is $T = 20$. As we discussed in the previous subsections we consider a linear model for estimating the demand, that is, $\widehat{d_t} = \widehat{\beta}_t^0 + \widehat{\beta}_t^1 p_t$.

We first assume a model of demand assuming that $\varepsilon_t = 0$, and we apply both the myopic and the one-dimensional DP policies, which is optimal in this case. In order to show the effect of demand learning we we plot in Figures 1 and 2 the least squares estimates of the intercept $\widehat{\beta}_t^0$ and the slope $\widehat{\beta}_t^1$. We notice that the estimates of the demand parameters indeed tend to the true demand parameters over time.

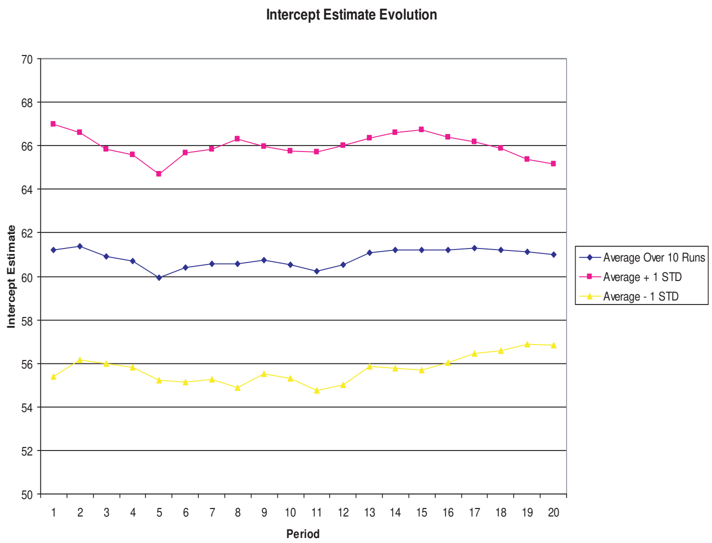

Figure 1: The estimate $\widehat{\beta}_t^0$.

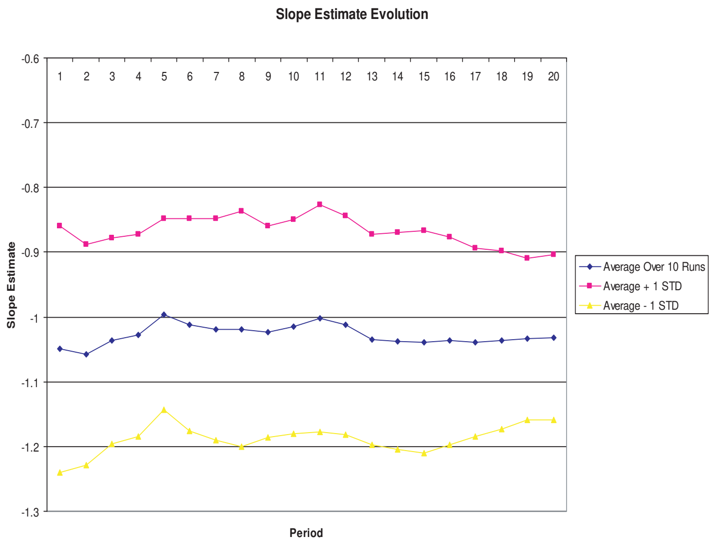

Figure 2: The estimate $\widehat{\beta}_t^1$.

In Table 1, we compare the total revenue and average price from the myopic and the one-dimensional DP policies, over 1,000 simulation runs. In general, as we mentioned earlier, for very large capacities both policies lead to the same revenue.

$T = 20$, $c = 400$	**Myopic**	**1-dim. DP**
Ave (Total Revenue)	$12,194$	$15,688$
Std (Total Revenue)	$1,162.9$	303.595
Ave(Ave Price)	30.9367	39.3595
Std (Ave Price)	2.8097	.6506

Table 1: Comparison of total revenue and average price for the myopic and the one-dimensional DP policies for $\varepsilon_t = 0$, over 1000 simulation runs with $T = 20$ and $c = 400$.

$T = 5$, $c = 125$	**Myopic**	**1-dim DP**	**5-Dim DP**
Ave.(Total Revenue)	$3,884.6$	$4,250.1$	$4,339.3$
Std (Total Revenue)	302.6	282.0	394.2
Ave.(Ave Price)	32.5	35.7	36.7
Std (Ave Price)	2.5	1.8	1.89

Table 2: Comparison of total revenue and average price for the myopic, the one-dimensional and five-dimensional DP policies for $\varepsilon_t \sim N(0, 16)$, over 1000 simulation runs with $T = 5$ and $c = 125$.

The results of Table 1 suggest that the one-dimensional DP outperforms the myopic policy significantly (by 28.65%). Moreover, the results become more dramatic as capacity drops.

We next consider the case that $\varepsilon_t \sim N(0, 16)$. In Table 2, we report the total revenue and average price from the myopic, one-dimensional DP and five-dimensional DP policies, over 1,000 simulation runs.

The results of Table 2 agree with intuition that the more computationally intensive methods lead to higher revenues. In particular, the one-dimensional DP policy outperforms the myopic policy (by 9.4%), and the five-dimensional DP policy outperforms the one-dimensional DP policy (by 2.09%). The results continue to hold for several values of the parameters we tested.

Overall, we feel that this example (as well as several others of similar nature) offers the following insights.

Insights:

1. All the methods we considered succeed in estimating accurately the demand parameters over time.

2. The class of DP policies outperforms the myopic policy. In addition, revenue increases with higher complexity of the DP method, that is the five-dimensianal DP policy outperforms the one-dimensional DP policy.

3 Pricing in a Competitive Environment

In this section, we study pricing under competition. In particular, we focus on a market with two firms competing for a single product in a dynamic environment, in which, the firm apart from trying to estimate its own demand, it also needs to predict its competitor's demand and pricing policy. Given the increased uncertainty, we use a more flexible model of demand, in which the firm considers that its own true demand as well as its competitor's demand have parameters that are time varying. Models of the type we consider in this section, were introduced in [5], and have nice asymptotic properties that we review shortly. Specifically, the firms have total capacity c_1 and c_2 respectively, over a finite time horizon T. In the beginning of each period t, Firm 1 knows the realizations of its own demand $d_{1,s}$, its own prices $p_{1,s}$ as well as its competitor's prices $p_{2,s}$, for $s = 1, \dots, t-1$. It does not directly observe, however, its competitor's demand.

We assume that each firm's true demand is an unknown linear function, where the true demand parameters are time varying, that is, for firm $k = 1, 2$ demand is of the form

$$d_{k,t} = \beta^0_{k,t} + \beta^1_{k,t} p_{1,t} + \beta^2_{k,t} p_{2,t} + \epsilon_{k,t},$$

where the coefficients $\beta^0_{k,t}, \beta^1_{k,t}, \beta^2_{k,t}$ vary slowly with time, i.e.,

$$|\beta^i_{k,t} - \beta^i_{k,t+1}| \le \delta_k(i), \qquad k = 1, 2;\ i = 0, 1, 2;\ t = 1, \dots, T-1.$$

This model assumes that demand for each firm $k = 1, 2$ depends on its own as well as its competitors current period prices $p_{1,t}$, $p_{2,t}$, unknown parameters $\beta^0_{k,t}$, $\beta^1_{k,t}$, $\beta^2_{k,t}$, and a random noise $\epsilon_{k,t} \sim N(0, \sigma^2_{k,t})$, $k = 1, 2$. The parameters $\delta_k(i)$, $i = 0, 1, 2$ are prespecified constants, called *volatility parameters*, and impose the condition that the coefficients $\beta^0_{k,t}, \beta^1_{k,t}, \beta^2_{k,t}$ are Lipschitz continuous. For example setting $\delta_k(i) = 0$, for some i, implies that the i^{th} parameter of the demand is constant in time (this is the usual regression condition).

Firm 1's objectives are to estimate its own demand, its competitor's reaction and finally, set its own prices dynamically in order to maximize its total expected revenue.

The results in [5] suggest that if the true demand is Lipschitz continuous, then the linear model of demand with time varying parameters we consider will indeed converge to the true demand. Moreover, the rate of convergence is faster than other alternative models. While we could use this model in the noncompetitive case of the previous section, it would lead to very high dimensional DPs that we could not solve exactly.

The remainder of this section is organized as follows. In Section 3.1, we present the firm's demand estimation model. In Section 3.2, we present a model that will allow the firm to predict its competitor's prices but also a model that the firm performs to set its own prices. Finally, in Section 3.3 we present some computaional results.

3.1 Demand Estimation

Each firm at time t estimates its own demand to be

$$\widehat{D}_{k,t} = \widehat{d}_{k,t} + \varepsilon_{k,t}, \qquad k = 1, 2$$

where $\widehat{d}_{k,t}$ is a point estimate of the current period demand and $\varepsilon_{k,t}$ is a random noise for firm $k = 1, 2$. The point estimate of the demand in current period t is given by $\widehat{d}_{1,t} = \widehat{\beta}^0_{1,t} + \widehat{\beta}^1_{1,t} p_{1,t} + \widehat{\beta}^2_{1,t} p_{2,t}$ and $\widehat{d}_{2,t} = \widehat{\beta}^0_{2,t} + \widehat{\beta}^1_{2,t} p_{1,t} + \widehat{\beta}^2_{2,t} p_{2,t}$. The parameter estimates are based on the price and demand realizations in the previous periods.

We assume that the parameter estimates $\widehat{\beta}^1_{1,t}$ and $\widehat{\beta}^2_{2,t}$ that describe how each firm's own price affects its own demand, are negative. This is a reasonable assumption since it states that the demand is decreasing in the firm's own price. Moreover, the parameter estimates $\widehat{\beta}^2_{1,t}$, $\widehat{\beta}^1_{2,t}$ are nonnegative, indicating that if the competitor sets for example, high prices they will increase the firm's own demand.

The firm makes the following distributional assumption on the random noise for each firm's demand,

$$\varepsilon_{k,t} \sim N(0, \widehat{\sigma}^2_{k,t}), \quad \text{where} \quad k = 1, 2,$$

and the demand variance estimated for each firm is,

$$\widehat{\sigma}^2_{1,t} = \frac{\sum_{\tau=1}^{t-1} \left(d_{1,\tau} - \widehat{\beta}^0_{1,t} - \widehat{\beta}^1_{1,t} p_{1,\tau} - \widehat{\beta}^2_{1,t} p_{2,\tau} \right)^2}{t-4}$$

$$\widehat{\sigma}^2_{2,t} = \frac{\sum_{\tau=1}^{t-1} \left(d_{2,\tau} - \widehat{\beta}^0_{2,t} - \widehat{\beta}^2_{2,t} p_{2,\tau} - \widehat{\beta}^1_{2,t} p_{1,\tau} \right)^2}{t-4}.$$

Notice that for the same reason as in the noncompetitive case, the variance estimates $\widehat{\sigma}^2_{k,t}$ for $k = 1, 2$, have $t - 4$ degrees of freedom.

For each firm $k = 1, 2$ we denote by $\widehat{\boldsymbol{\beta}}_k = (\widehat{\beta}_{k,1}, \widehat{\beta}_{k,2}, ..., \widehat{\beta}_{k,t-1})$, where $\widehat{\beta}_{k,t} = (\widehat{\beta}^0_{k,t}, \widehat{\beta}^1_{k,t}, \widehat{\beta}^2_{k,t})$.

In order to estimate its own demand Firm 1 solves the following problem.

$$\begin{aligned} \text{minimize}_{\widehat{\boldsymbol{\beta}}_1} & \sum_{\tau=1}^{t-1} |d_{1,\tau} - (\widehat{\beta}^0_{1,\tau} + \widehat{\beta}^1_{1,\tau} p_{1,\tau} + \widehat{\beta}^2_{1,\tau} p_{2,\tau})| \\ \text{subject to} \quad & |\widehat{\beta}^i_{1,\tau} - \widehat{\beta}^i_{1,\tau+1}| \le \delta_1(i), \quad i = 0, 1, 2, \;\; \tau = 1, 2, ..., t-2 \\ & \beta^1_{1,\tau} \le 0, \;\; \beta^2_{1,\tau} \ge 0. \end{aligned}$$

Note that we impose the constraint that the parameters are varying slowly with time. This is reflected in the numbers $\delta_1(i)$. Note that this problem can be transformed to a linear optimization model, which makes it attractive computationally.

Let $(\widehat{\beta}^i_{1,\tau})^*$, $i = 0, 1, 2$, $\tau = 1, ..., t-1$ be an optimal of this problem. Firm 1 would like now to make an estiamate for the parameters $(\widehat{\beta}^0_{1,t}, \widehat{\beta}^1_{1,t}, \widehat{\beta}^2_{1,t})$. We propose the estimate:

$$\widehat{\beta}^i_{1,t} = \frac{1}{N} \sum_{l=t-1-N}^{t-1} (\widehat{\beta}^i_{1,l})^*, \qquad i = 0, 1, 2,$$

that is the new estimate is an average of the estimates of the N previous periods. In particular, if we choose $N = 1$, the new estimate is equal to the estimate for the previous period.

3.2 Competitor's price prediction and own price setting

In order for Firm 1 to set its own prices in current period t, apart from estimating its own demand, it also needs to predict how its competitor (Firm 2) will react and set its prices in period t. The information available to Firm 1 at each time period, includes, apart from the realizations of its own demand, also the prices each firm has set in all the previous periods. We will assume that Firm 1 believes that its competitor is also setting prices optimally. In this case, Firm 1 is confronted with an *inverse* optimization problem. The reason for this is that Firm 1 tries to guess the parameters of its competitor's demand (by assuming it also belongs to a parametric family with unknown parameters) through an optimization problem that would exploit the actual observed competitor's prices. In what follows, we will distinguish between the uncapacitated and the capacitated versions of the problem.

Uncapacitated Case

As we mentioned, we assume that Firm 1 believes that Firm 2 is also a revenue maximizer and, as a result, solves the optimization problem,

$$\max_{p_{2,\tau}} \; p_{2,\tau}.(\widehat{\beta}^0_{2,\tau} + \widehat{\beta}^1_{2,\tau} p^1_{1,\tau} - \widehat{\beta}^2_{2,\tau} p_{2,\tau}), \qquad \tau = 1, ..., t.$$

This problem has a closed form solution of the form

$$\widehat{p}_{2,\tau} = \frac{\widehat{\beta}^0_{2,\tau} + \widehat{\beta}^1_{2,\tau} p^1_{1,\tau}}{-2\widehat{\beta}^2_{2,\tau}}, \qquad \tau = 1, ..., t.$$

Price $p^1_{1,\tau}$ denotes what Firm 1's estimate is of what Firm 2 believes for Firm 1's pricing. Examples of such estimates include: $p^1_{1,\tau} = p_{1,\tau}$, $p^1_{1,\tau} = p_{1,\tau-1}$, or an average of price realizations from several periods prior to period τ.

Firm 1 will then estimate the demand parameters for Firm 2 by solving the following optimization problem

$$\min_{\widehat{\boldsymbol{\beta}}_2} \; \sum_{\tau=1}^{t-1} \left| p_{2,\tau} - \frac{\widehat{\beta}^0_{2,\tau} + \widehat{\beta}^1_{2,\tau} p^1_{1,\tau}}{-2\widehat{\beta}^2_{2,\tau}} \right|$$

$$\text{subject to} \quad |\widehat{\beta}^i_{2,\tau} - \widehat{\beta}^i_{2,\tau+1}| \leq \delta_2(i), \qquad i = 0,1,2, \;\; \tau = 1,2,...,t-2,$$

$$\widehat{\beta}^1_{2,\tau} \geq 0, \;\; \widehat{\beta}^2_{2,\tau} \leq 0.$$

As in the model for estimating the current period demand for Firm 1, $\delta_2(i)$, $i = 0,1,2$, are *volatility parameters* that we assume to be prespecified constants. The solutions $(\widehat{\beta}^i_{2,\tau})^*$, $i = 0,1,2$, of this optimization model allow Firm 1 to estimate its competitor's current period demand by setting:

$$\widehat{\beta}_{2,t} = \frac{1}{N} \sum_{l=t-1-N}^{t-1} (\widehat{\beta}_{2,l})^*.$$

Myopic Own Price Setting Policy

After the previous analysis, Firm 1's own price setting problem follows easily. We assume that Firm 1 sets its prices by maximizing its current period t revenues. That is,

$$\max_{p_{1,t}} \; p_{1,t}.(\widehat{\beta}^0_{1,t} - \widehat{\beta}^1_{1,t} p_{1,t} + \widehat{\beta}^2_{1,t} \widehat{p}_{2,t}).$$

This optimization model uses the estimates of the parameters $\widehat{\beta}^i_{1,t}$, $i = 0,1,2$, that we described in Firm 1's own demand estimation problem, as well as the prediction of the competitor's price $\widehat{p}_{2,t} = \frac{\widehat{\beta}^0_{2,t} + \widehat{\beta}^1_{2,t} p^1_{1,t}}{-2\widehat{\beta}^2_{2,t}}$. Notice that this latter part also involves the estimates of the demand parameters $\widehat{\beta}^i_{2,t}$, $i = 0,1,2$ arising through the inverse optimization problem in the competitor's price prediction problem.

Capacitated Case

We assume that both firms face a total capacity c_1 and c_2 respectively that they need to allocate in the total time horizon. As before, Firm 1 makes the behavioral assumption that Firm 2 is also a revenue maximizer. Using the notation $x^+ = \max(0, x)$, the price prediction problem that Firm 1 solves for predicting its competitor's prices becomes

$$\hat{p}_{2,t} = \arg\max_{p \in \mathcal{P}_2} \; p \min \left\{ \left(\widehat{\beta}^0_{2,t} + \widehat{\beta}^2_{2,t} p_{2,t} + \widehat{\beta}^1_{2,t} p^1_{1,t} \right)^+, \; c_2 - \sum_{\tau=0}^{t-1} (\widehat{\beta}^0_{2,\tau} + \widehat{\beta}^2_{2,\tau} p_{2,\tau} + \widehat{\beta}^1_{2,\tau} p^1_{1,\tau})^+ \right\}$$

As in the uncapacitated case, $p^1_{1,\tau}$ denotes Firm 1's estimate of what Firm 2 assumes for Firm 1's own pricing. Examples include: $p^1_{1,\tau} = p_{1,\tau}$, or $p_{1,\tau-1}$, or considering an average of the prices Firm 1 sets in several previous periods. We can now estimate Firm 2's demand parameters through the following optimization model

$$\min \sum_{\tau=1}^{t-1} |p_{2,\tau} - \hat{p}_{2,\tau}|$$

$$\text{subject to} \quad |\widehat{\beta}^i_{2,\tau} - \widehat{\beta}^i_{2,\tau+1}| \leq \delta_2(i), \qquad i = 0, 1, 2, \ \tau = 1, 2, ..., t-2$$

$$\widehat{\beta}^1_{2,\tau} \geq 0, \ \widehat{\beta}^2_{2,\tau} \leq 0,$$

where $\hat{p}_{2,t} \in \arg\max_{p \in \mathcal{P}_2} \ p \min \left\{ \left(\widehat{\beta}^0_{2,t} + \widehat{\beta}^2_{2,t} p + \widehat{\beta}^1_{2,t} p^1_{1,t} \right)^+ , \ c_{2,t} \right\}$.

Let $(\widehat{\beta}^i_{2,\tau})^*$, $i = 0, 1, 2$, $\tau = 1, ..., t-1$ be optimal solutions to this optimization problem. As before, Firm 1 estimates its competitor's current period demand parameters as

$$\widehat{\beta}^i_{2,t} = \frac{1}{N} \sum_{l=t-1-N}^{t-1} (\widehat{\beta}^i_{2,l})^*, \qquad i = 0, 1, 2.$$

Myopic Own Price Setting Policy

After computing its own and its competitor's demand parameter estimates and establishing a prediction on the price of its competitor for the current period, Firm 1 is ready to set its own current period price. As in the uncapacitated case, Firm 1 solves the current period revenue maximization problem, that is,

$$p_{1,t} \in \arg\max_{p \in \mathcal{P}} \left[p \min \left\{ \left(\widehat{\beta}^0_{1,t} + \widehat{\beta}^1_{1,t} p + \widehat{\beta}^2_{1,t} \widehat{p}_{2,t} \right)^+ , c_{1,t} \right\} \right],$$

where $c_{1,t} = c_1 - \sum_{\tau=1}^{t-1} d_{1,\tau}$ is Firm 1's remaining capacity in period t. Moreover, the demand parameters $\widehat{\beta}^i_{1,t} = \frac{1}{N} \sum_{k=t-1-N}^{t-1} (\widehat{\beta}^i_{1,k})^*$, $\widehat{\beta}^i_{2,t} = \frac{1}{N} \sum_{l=t-1-N}^{t-1} (\widehat{\beta}^i_{2,l})^*$, $i = 0, 1, 2$, and finally, the estimates of the competitor's prices are $\hat{p}_{2,t} \in \arg\max_{p \in \mathcal{P}_2} \ p \min \left\{ \left(\widehat{\beta}^0_{2,t} + \widehat{\beta}^2_{2,t} p + \widehat{\beta}^1_{2,t} p^1_{1,t} \right)^+ , \ c_{2,t} \right\}$.

3.3 Computational Results

We consider two firms competing for one product. The true models of demand for the two firms respectively are as follows:

$$d_{1,t} = 50 - .05 p_{1,t} + .03 p_{2,t} + \varepsilon_{1,t}$$

Firm 1	Firm 2	1 Avg(Rev)	2 Avg(Rev)	1 Std(Rev)	2 Std(Rev)
Opt	**Rand**	$3,126,000$	$2,909,200$	$70,076$	$109,790$
Rand	**Rand**	$2,638,800$	$2,616,900$	$63,112$	$61,961$
Match	**Rand**	$2,602,700$	$2,603,200$	$117,470$	$123,070$
Opt	**Match**	$3,791,100$	$3,779,400$	$177,540$	$197,370$
Rand	**Match**	$2,603,200$	$2,602,700$	$123,070$	$117,470$
Opt	**Opt**	$3,757,700$	$3,804,700$	$70,577$	$129,530$
Rand	**Opt**	$2,909,200$	$3,126,000$	$109,790$	$70,076$
Match	**Opt**	$3,779,400$	$3,791,100$	$197,370$	$177,540$

Table 3: A comparison of revenues under random, matching, optimization based pricing policies.

$$d_{2,t} = 50 + .03p_{1,t} - .05p_{2,t} + \varepsilon_{2,t}$$

where the $\varepsilon_{1,t},\ \varepsilon_{2,t} \sim N(0,16)$. Moreover, the prices for both firms range in the sets $\mathcal{P}_1 = \mathcal{P}_2 = [100,\ 900]$, the time horizon is $T = 150$ and finally we assume that $p_{1,1} = p_{2,1} = 500$. Finally, we assume an uncapacitated setting.

We compare three pricing policies: (a) random pricing, (b) price matching, and (c) optimization based pricing using the methods we outlined in this section. A firm employing the random pricing policy chooses a price at random from the feasible price set. In particular, we consider a discrete uniform distribution over the set of integers [100, 900]. A firm employing the price matching policy sets, in the current period, the price its competitor set in the previous period. Finally, a firm employing optimization based pricing first solves the demand estimation problem in order to estimate its current period parameter estimates using linear programming, supposes its competitor will repeat its previous period pricing decision, and then uses myopic pricing in order to set its prices. In Table 3, we report the revenue from the three strategies, over 1000 simulation runs.

In order to obtain intuition from Table 3, we fix the strategy the competitor is using, and then see the effect on revenue of the policy followed by Firm 1. If Firm 2 is using the random pricing policy, it is clear that Firm 1 has a significant increase in revenue by using an optimization based policy. Similarly, if Firm 2 is using a matching policy, again the optimization based policy leads to significant improvements in revenue. Finally, if Firm 2 is using an optimization based policy, then the matching policy is slightly better than the optimization based policy. However, given that the margin is small and given the variability in the estimation process, it might still be possible for the optimization based policy to be stronger. It is thus fair to say, that at least in this example, no matter what policy Firm 2 is using, Firm 1 seems to be better off by using an optimization based policy.

4 Conclusions

We introduced models for dynamic pricing in an oligopolistic market. We first studied models in a noncompetitive environment in order to understand the effects of demand learning. By considering the framework of dynamic programming with incomplete state information for jointly estimating the demand and setting prices for a firm, we proposed increasingly more computationally intensive algorithms that outperform myopic policies. Our overall conclusion is that dynamic programming models based on incomplete information are effective in jointly estimating the demand and setting prices for a firm.

We then studied pricing in a competitive environment. We introduced a more sophisticated model of demand learning in which the price elasticity is a slowly varying function of time. This allows for increased flexibility in the modeling of the demand. We outlined methods based on optimization for jointly estimating the Firm's own demand, its competitor's demand, and setting prices. In preliminary computational work, we found that optimization based pricing methods offer increased revenue for a firm independently of the policy the competitor firm is following.

Acknowledgments

The first author would like to acknowledge the Singapore-MIT Alliance Program for supporting this research. The second author would like to acknowledge the PECASE Award DMI-9984339 from the National Science Foundation, the Charles Reed Faculty Initiative Fund, the New England University Transportation Research Grant and the Singapore-MIT Alliance Program for supporting this research. Both authors would also like to thank Marc Coumeri for performing some of the computations in this paper.

References

[1] Bagchi, A. 1984. *Stackleberg Differential Games in Economic Models*, Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, New York.

[2] Basar, T. 1986. *Dynamic Games and Applications in Economics*, Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, New York.

[3] Bertsekas, D. 1995. *Dynamic Programming and Optimal Control I*, Athena Scientific, MA.

[4] Bertsekas, D., and J. Tsitsiklis. 1996. *Neuro-Dynamic Programming*, Athena Scientific, MA.

[5] Bertsimas, D., Gamarnik, D., and J. Tsitsiklis. 1999. Estimation of Time-Varying Parameters in Statistical Models: An Optimization Approach, *Machine Learning*, **35**, 225-245.

[6] Bertsimas, D., and J. Tsitsiklis. 1997. *Introduction to Linear Optimization*, Athena Scientific, MA.

[7] Bitran, G., and S. Mondschein. 1997. Periodic Pricing of Seasonal Products in Retailing, *Management Science*, **43**(1), 64-79.

[8] Chan, LMA., Simchi-Levi, D., and Swann J. 2000. Flexible Pricing Strategies to Improve Supply Chain Performance, Working Paper.

[9] Dockner, E., and S. Jorgensen. 1988. Optimal Pricing Strategies for New Products in Dynamic Oligopolies, *Marketing Science*, **7**(4), 315-334.

[10] Federgruen, A., and A. Heching. 1997. Combined Pricing and Inventory Control Under Uncertainty, *Operations Research*, **47**(3), 454-475.

[11] Feng, Y., and G. Gallego. 1995. Optimal Starting Times for End-of-Season Sales and Optimal Stopping Times for Promotional Fares, *Management Science*, **41**(8), 1371-1391.

[12] Friedman, J.W. 1977. *Oligopoly and the Theory of Games*, North Holland, Amsterdam.

[13] Friedman, J.W. 1982. Oligopoly Theory in *Handbook of Mathematical Economics II* chapter 11, North Holland, Amsterdam.

[14] Friedman, J.W. 1983. *Oligopoly Theory*, Cambridge University Press, Cambridge.

[15] Fudenberg, D., and J. Tirole. 1986. *Dynamic Models of Oligopoly*. Harwood Academic, London.

[16] Gallego, G., and G. van Ryzin. 1994. Optimal Dynamic Pricing of Inventories with Stochastic Demand Over Finite Horizons, *Management Science*, **40**(8), 999-1020.

[17] Gallego, G., and G. van Ryzin. 1997. A Multiproduct Dynamic Pricing Problem and its Applications to Network Yield Management, *Operations Research*, **45**(1), 24-41.

[18] Gibbens, R.J., and Kelly F.P. 1998. Resource Pricing and the Evolution of Congestion Control. Working Paper.

[19] Gilbert, S. 2000. Coordination of Pricing and Multiple-Period Production Across Multiple Constant Priced Goods. *Management Science*, **46**(12), 1602-1616.

[20] Kalyanam, K. 1996. Pricing Decisions Under Demand Uncertainty: A Bayesian Mixture Model Approach. *Marketing Science*, **15**(3), 207-221.

[21] Kelly, F.P. 1994. On Tariffs, Policing and Admission Control for Multiservice Networks. *Operations Research Letters*, 15, 1-9.

[22] Kelly, F.P., Maulloo, A.K., and Tan, D.K.H. 1998. Rate Control for Communication Networks: Shadow Prices, Proportional Fairness and Stabilit. *Journal of the Operational Research Society*, 49, 237-252.

[23] Kopalle, P., Rao, A., and J. Assuncao. 1996. Asymmetric Reference Price Effects and Dynamic Pricing Policies. *Marketing Science*, **15**(1), 60-85.

[24] Kuhn, H. 1997. *Classics in Game Theory*, Princeton University Press, NJ.

[25] Lilien, G., Kotler, P., and K. Moorthy. 1992. *Marketing Models*, Prentice Hall, NJ.

[26] Mas-Colell, A., Whinston, M., and J. Green. 1995. *Microeconomic Theory*, Oxford University Press, New York.

[27] McGill, J., and G. Van Ryzin. 1999. Focused Issue on Yield Management in Transportation. *Transportation Science*, **33**(2).

[28] Nagurney, A. 1993. *Network Economics A Variational Inequality Approach*, Kluwer Academic Publishers, Boston.

[29] Paschalidis, I., and J. Tsitsiklis. 1998. Congestion-Dependent Pricing of Network Services. Technical Report.

[30] Rice, J. 1995. *Mathematical Statistics and Data Analysis,* Second Edition, Duxbury Press, California.

[31] Tirole, J., and E. Maskin. 1985. A Theory of Dynamic Oligopoly II: Price Competition, *MIT Working Papers.*

[32] Van Mieghen, J., Dada, M., 1999. Price vs Production Postponement. *Management Science*, 45, 12, 1631-1649.

[33] Van Mieghen, J. 1999. Differentiated Quality of Service: Price and Service Discrimination in Queueing Systems. Working Paper.

[34] Weatherford, L., and S. Bodily. 1992. A Taxonomy and Research Overview of Perishable Asset Revenue Management: Yield Management, Overbooking and Pricing. *Operations Research*, **40**(5), 831-844.

[35] Williamson, E. 1992. Airline Network Seat Inventory Control: Methodologies and Revenue Impacts. Ph.D. *Thesis Flight Transportation Lab, MIT.*

[36] Wilson, R. 1993. *Nonlinear Pricing*, Oxford University Press.