Non-cooperative Competition Among Revenue Maximizing Service Providers with Demand Learning

Changhyun Kwon[1], Terry L. Friesz[1]*,
Reetabrata Mookherjee[2], Tao Yao[1] and Baichun Feng[1]

[1] : Department of Industrial and Manufacturing Engineering
The Pennsylvania State University
University Park, PA 16802, U.S.A.
[2] : Zilliant, Inc., Austin, TX 78704, U.S.A.

E-mail: chkwon@psu.edu, tfriesz@psu.edu,
Reetabrata.Mookherjee@zilliant.com, taoyao@psu.edu, buf118@psu.edu

Abstract

This paper recognizes that in many decision environments in which revenue optimization is attempted, an actual demand curve and its parameters are generally unobservable. Herein we describe the dynamics of demand as a continuous time differential equation based on an evolutionary game theory perspective. We then observe realized sales data to obtain estimates of parameters that govern the evolution of demand; these are refined on a discrete time scale The resulting model takes the form of a differential variational inequality. We present an algorithm based on a gap function for the differential variational inequality and report its numerical performance for an example revenue optimization problem.

Keywords : *Revenue management, Pricing, Demand learning, Differential games, Kalman filters*

1 Introduction

In rapidly growing literature on revenue management – see Talluri and van Ryzin (2004) and McGill and van Ryzin (1999) for comprehensive studies and survey – one of the most important issues is how to model service provider demand learning. Demand is usually represented as a function of price, explicitly and/or implicitly, and the root tactic upon which revenue management is based is to change prices dynamically to maximize immediate or short-run revenue. In this sense, the more accurate the model of demand employed in revenue optimization, the more revenue we can generate. Although demand may be viewed theoretically as the result of utility maximization, an actual demand curve and its parameters are generally unobservable in most markets. In this paper, we first describe the dynamics of demand as a differential equation based on an evolutionary game theory perspective and then observe the actual sales data to obtain estimates of parameters that govern the evolution of demand.

A dynamic non-zero sum evolutionary game among service providers who have no variable costs and hold no inventories is employed, in this paper, as a mathematical formulation, in particular a differential variational inequality. The providers also have fixed upper bounds on output as each has capacity constraints of available resources. We intend to provide numerical examples for the revenue management model we will introduce later in this paper with presenting effective and efficient numerical methods, which are specific for the problem structures.

---

*Communicating author - Tel : +1-814-863-2445, Fax : +1-814-863-4745

1.1 Revenue Management Model

The service providers of interest are in oligopolistic game theoretic competition according to a learning process that is similar to evolutionary game-theoretic dynamics and for which price changes are proportional to their signed excursion from a market clearing price. We stress that in this model firms are setting prices for their services while simultaneously determining the levels of demand they will serve. This is unusual in that, typically, firms in oligopolistic competition are modelled as setting either prices or output flows. The joint adjustment of prices and output is modelled here as determined by comparing current price to the price that would have cleared the market for the demand that has most recently been served. However, the service providers are unable to make this comparison until the current round of play is completed as knowledge of the total demand served by all competitors is required.

Kachani et al. (2004) put forward a revenue management model for service providers to address such joint pricing and demand learning in an oligopoly setting under fixed capacity constraints. The model they consider assumes a service provider's demand is a linear function of its price and other competitors' prices; each company learns to set their parameters over time, though the impact of a change in price on demand in one period does not automatically propagate to latter time periods. In our work we allow this impact to propagate to all the time periods down the line.

In this paper, we will only be considering a class of customers, so-called bargain-hunting buyers Dasgupta and Das (2000), who may represent the general public who are searching for personal or, to a limited extent, business services or products at the most competitive price; these buyers are willing to sacrifice some convenience for the sake of a lower price. Because the services and products are assumed to be homogeneous, if two sellers offer the same price, the tie is broken randomly. In other words, the consumer has no concept of brand preference.

1.2 Differential Variational Inequalities and Differential Games

The differential variational inequalities, or DVIs, are infinite-dimensional variational inequalities with a special structure, which involves ordinary differential equations, called the state dynamics in the optimal control theory. The importance and application of DVIs are rising in mechanics, mathematical economics, transportation research, and many other complex engineering systems. Recently, Pang and Stewart (2007) introduced DVIs formally, and found applications in ordinary differential equations with discontinuous right-hand sides, differential Nash games, and multi-rigid-body dynamics. Moreover, Friesz et al. (2006) used a DVI to model shippers' oligopolistic competition on networks. The DVI of our interest should not be confused with the differential variational inequality has been used in Aubin and Cellina (1984), which may be called a variational inequality of evolution, a name suggested by Pang and Stewart (2007).

The non-cooperative dynamic oligopolistic competition amongst the service providing agents may be placed in the form of a dynamic variational inequality wherein each individual agent develops for itself a distribution plan that is based on current and future knowledge of non-own price patterns. Each such agent is, therefore, described by an optimal control problem for which the control variables are the price of its various service classes; this optimal control problem is constrained by dynamics that describe how the agent's share of demand alters with time. The Cournot-Nash assumption that game agents are non-cooperative allows the optimal control problems for individual agents to be combined to obtain a single DVI.

1.3 Demand Learning, Model Parameter Estimation and the Kalman Filtering

Forecasting demand is crucial in pricing and planning for any firm in that the forecasts have huge impacts on the revenues. In revenue management literature, most of the research models demand as an exogenous stochastic process from a known distribution (Gallego and van Ryzin, 1994; Feng and Gallego, 1995). Such models are restrictive because (1) they depend largely on the complete knowledge of the demand characteristics before the pricing starts; (2) they do not incorporate any learning mechanisms which will improve the demand estimation as more information becomes available.

In this paper, the demand for each firm is governed by a dynamics controlled by prices set by its own and its competitors. However, the parameters in this demand dynamics are unknown. The demand forecasting is flexible in the sense that it is able to handle incomplete information about the demand function. Furthermore, the demand is learned over time and each firm can update its demand function as new information becomes

available. By using such kind of learning mechanism, the firm can better estimate the demand function thereby improving its profitability. Thus, less restriction is enforced in modeling the demand in this paper which makes the model more attractive in real world situations.

Demand learning has been studied extensively in many research areas. The typical approach to model the learning effect is the Bayesian learning technique. For the Bayesian learning approach, usually demand at any time can be model as a stochastic process following certain distribution with unknown parameter (or parameters). The unknown parameter has a known distribution which is called the prior distribution. Observed demand data are used to modify the belief about the unknown parameter based on Bayes' rule. In this approach, uncertainty in the parameter is resolved as more observations become available, and the distribution of the demand will approach its true distribution (Murray Jr and Silver, 1966; Eppen and Iyer, 1997; Bitran and Wadhwa, 1996). Recently, researchers have developed other learning mechanisms to resolve the demand uncertainty. To name a few, Bertsimas and Perakis (2006) develop a demand learning mechanism which estimates the parameters for the linear demand function by virtue of a least square method. Yelland and Lee (2003) use class II mixture models to capture both the uncertainty in model specification and demand changes in regime. They demonstrate that the class II mixture models are more efficient in forecasting the product demand for Sun Microsystems Inc. Lin (2006) proposes to use the real time demand data to improve the estimation of the customer arrival rate which in turn can be used to better predict the future demand distribution, and develops a variable-rate policy which is very immune to the changes in the customer arrival rate.

The learning demand approach proposed in this paper is Kalman Filter. As a state-space estimation method, the Kalman filtering has gained attention recently in economics and revenue management as one of the most successful forecasting methods. The Kalman filters, which is developed by Kalman (1960) and Kalman and Bucy (1961), originally to filter out system and sensor noise in electrical/mechanical systems, is based on a system of linear state dynamics, observation equations and normally distributed noise terms with mean zero. The Kalman filter provides *a prior* estimate for the model parameter, which is adjusted to *a posteriori* estimate by combining the observations. With this new model parameter, we repeatedly solve the dynamic pricing problem to maximize the revenue in the next time periods. For the detailed discussion and derivations of the Kalman filter dynamics, see Bryson and Ho (1975), or, for an introduction suitable for revenue management researchers, Talluri and van Ryzin (2004).

Among economics and revenue management literature, Balvers and Cosimano (1990) study a stochastic linear demand model in a dynamic pricing problem and obtain a dynamic programming formulation to maximize revenue. They use the Kalman filter to estimate the exact value of the intercept and elasticity in the stochastic linear demand model. Closely related, Carvalho and Puterman (2007) consider a log-linear demand model whose parameters are also stochastic, and test the model using the Monte Carlos simulation techniques. In addition, Xie et al. (1997) find an application of the Kalman filter in estimation of new product diffusion models, where they concluded the Kalman filter approach gives superior prediction performance compared to other estimation methods in such environment.

1.4 Numerical Methods for Solving DVIs

Among the algorithms to solve the variational inequality problems, decent methods with gap functions have special structures in themselves: a variational inequality problem can be converted to an equivalent optimization problem, whose objective function is always nonnegative and optimal objective function value is zero if and only if the optimal solution solves the original variational inequality problem. A number of algorithms in this class has been developed for the finite-dimensional variational inequalities, see, for example, Zhu and Marcotte (1994) and Yamashita et al. (1997). For the infinite-dimensional problems, Zhu and Marcotte (1998) and Konnov et al. (2002) extended descent methods using gap functions in Banach spaces and Hilbert spaces, respectively.

However, only a few number of numerical schemes for the DVIs has been reported in the literature despite of the increasing importance of the DVIs with applications in differential games. Pang and Stewart (2007) suggested two algorithms, namely, the time stepping method and the sequential linearization method. The former discretize the time horizon with finite differences for the state dynamics, and then constitute a discrete system with finite-dimensional variational inequalities. To ensure the convergence of the time stepping method, we need linearity with respect to control in state dynamics. To overcome this limitation,

the later scheme, the sequential linearization method, sequentially approximate the given DVI by a sub-DVI at each iterative state and control variable, and then solve each sub-DVI by available algorithms such as the time stepping method. To add a numerical method for solving DVIs and show how it works in differential games, we adopt the descent method using gap functions devised by Kwon and Friesz (2007) who extended gap functions for infinite-dimensional variational inequalities to obtain an equivalent optimal control problem.

1.5 Organization of the Paper

The remainder of the paper is organized as follows : Section 2 provides an introductory overview of DVIs. This is followed by a detailed exposition of our revenue management model of dynamic competition in Section 3. A model parameter estimation technique, which is called the Kalman filtering, is described for the revenue management model in Section 4. We provide a scenario for a firm and a numerical method to solve its optimal control problem without considering the game in Section 5. Section 6 shows how dynamic revenue management competition may be expressed as a DVI. In Section 7 we outline a descent method using gap functions for DVIs which is used in the next section to solve an example. Section 8 provides a detailed numerical example to show some interesting behaviors of the agents. Section 9 summarizes our findings and describes future research.

2 Abstract DVI

The finite- or infinite-dimensional variational inequality problem (VIP) is, for a compact and convex set U and function F, to find $u^* \in U$ such that

$$\langle F(u^*), u - u^* \rangle \geq 0 \qquad \forall u \in U$$

where $\langle \cdot, \cdot \rangle$ denotes the corresponding inner product. It is well-known that a VIP is closely related to an optimization problem. When a variable of an optimization problem has a representation of an ordinary differential equation, we call the problem an optimal control problem, which is closely related to a differential variational inequality problem we will introduce. We begin by letting

$$u \in \left(L^2 [t_0, t_f] \right)^m$$
$$x(u, t) = \arg \left\{ \frac{dy}{dt} = f(y, u, t), \; y(t_0) = y, \Gamma[y(t_f), t_f] = 0 \right\} \in \left(\mathcal{H}^1 [t_0, t_f] \right)^n \tag{1}$$

The entity $x(u, t)$ is to be interpreted as an operator that tells us the state variable x for each vector u and each time $t \in [t_0, t_f] \subset \mathbb{R}^1_+$; constraints on u are enforced separately. We assume that every control vector is constrained to lie in a set U, where U is defined so as to ensure the terminal conditions may be reached from the initial conditions intrinsic to (1). In light of the operator (1), the variational inequality of interest to us takes the form:

find $u^* \in U$ such that
$$\langle F(x(u^*, t), u^*, t), u - u^* \rangle \geq 0 \text{ for all } u \in U \tag{2}$$

where

$$U \subseteq \left(L^2 [t_0, t_f] \right)^m \tag{3}$$
$$x^0 \in \mathbb{R}^n \tag{4}$$
$$F : \left(\mathcal{H}^1 [t_0, t_f] \right)^n \times \left(L^2 [t_0, t_f] \right)^m \times \mathbb{R}^1_+ \longrightarrow \left(L^2 [t_0, t_f] \right)^m \tag{5}$$
$$f : \left(\mathcal{H}^1 [t_0, t_f] \right)^n \times \left(L^2 [t_0, t_f] \right)^m \times \mathbb{R}^1_+ \longrightarrow \left(L^2 [t_0, t_f] \right)^n \tag{6}$$
$$\Gamma : \left(\mathcal{H}^1 [t_0, t_f] \right)^n \times \Re^1_+ \longrightarrow \left(\mathcal{H}^1 [t_0, t_f] \right)^r \tag{7}$$

Note that $\left(L^2\left[t_0, t_f\right]\right)^m$ is the m-fold product of the space of square-integrable functions $L^2\left[t_0, t_f\right]$ and the inner product in (2) is defined by

$$\left\langle F\left(x\left(u^*, t\right), u^*, t\right), u - u^*\right\rangle \equiv \int_{t_0}^{t_f}\left[F\left(x\left(u^*, t\right), u^*, t\right)\right]^T\left(u - u^*\right) \geq 0$$

while $\left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n$ is the n-fold product of the Sobolev space $\mathcal{H}^1\left[t_0, t_f\right]$. We refer to (2) as a differential variational inequality and give it the symbolic name $DVI(F, f, U)$.

To analyze (2) we will rely on the following notion of regularity:

Definition 1 (Regularity) *We call $DVI(F, f, U)$ regular if:*

1. *$x\left(u, t\right):\left(L^2\left[t_0, t_f\right]\right)^m \times \mathbb{R}_{+}^1 \longrightarrow\left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n$ exists and is continuous and Gateaux-differentiable with respect to u[1];*

2. *$\Gamma\left(x, t\right)$: is continuously differentiable with respect to x;*

3. *$F(x, u, t)$ is continuous with respect to x and u;*

4. *$f\left(x, u, t\right)$ is convex and continuously differentiable with respect to x and u;*

5. *U is convex and compact; and*

6. *$x^0 \in \mathbb{R}^n$ is known and fixed.*

The motivation for this definition of regularity is to parallel as closely as possible those assumptions needed to analyze traditional optimal control problems from the point of view of infinite dimensional mathematical programming.

Furthermore, there is a fixed point form of $DVI(F, f, U)$. In particular we state the following result without proof (see Mookherjee, 2006):

Theorem 2 (Fixed Point Problem) *When regularity in the sense of Definition 1 holds, $DVI(F, f, U)$ is equivalent to the following fixed point problem:*

$$u = P_U\left[u - \alpha F\left(x(u), u, t\right)\right]$$

where $P_U\left[.\right]$ is the minimum norm projection onto $U \subseteq\left(L^2\left[t_0, \tau\right]\right)^m$ and $\alpha \in \mathbb{R}_{++}^1$.

We now state and prove the following existence result (for the proof, see Mookherjee, 2006):

Theorem 3 (Existence) *When the regularity in the sense of Definition 1 holds, $DVI(F, f, U)$ has a solution.*

We are now ready give the revenue management model of our interest.

3 Revenue Management Model

A set of service providers are competing in an oligopolistic setting, each with the objective of maximizing their revenue. These service providers have very high fixed costs compared to their relatively low variable or operating costs. Therefore, the providers focus only on maximizing their own revenue.

Each company provides a set of services where each service type is assumed to be homogeneous. For example, the difference between an economy class seat on Southwest and an economy class seat on Jet Blue is indiscernible by customers; the only differences that the customers perceive are the prices charged by the different service providers.

All service providers can set the price for each of their services. The price that they charge for each service in one time period will affect the demand that they receive for that service in the next time period. The price that the service provider charges is compared to the rolling average price of their competitors and the provider's demand is affected accordingly. The amount of service that each company can provide has an upper bound. The providers must therefore choose prices that create an amount of demand for their services that will maximize their revenue while ensuring that the demands do not exceed their capacities.

---

[1]This condition is guaranteed when the other regualrity conditions are satisifed. See Bressan and Piccoli (2005).

3.1 Basic Notation

We denote the set of revenue managing firms as $\mathcal{F}$, each of whom is providing a set of services $\mathcal{S}$. Continuous time is denoted by the scalar $t \in \mathbb{R}_+^1$, while t_0 is the finite initial time and $t_f \in \mathbb{R}_{++}^1$ the finite terminal time so that $t \in [t_0, t_f] \subset \mathbb{R}_+^1$.

Each firm $f \in \mathcal{F}$ controls price

$$\pi_i^f \in L^2[t_0, t_f]$$

corresponding to each service type $i \in \mathcal{S}$, where $L^2[t_0, t_f]$ is the space of square-integrable functions for the real time interval $[t_0, t_f] \in \mathbb{R}_+^1$. The control vector of each firm $f \in \mathcal{F}$ is

$$\pi^f \in \left(L^2[t_0, t_f]\right)^{|\mathcal{S}|}$$

which is the concatenation of

$$\pi \in \left(L^2[t_0, t_f]\right)^{|\mathcal{S}| \times |\mathcal{F}|},$$

the complete vector of controls.

We also let

$$D_i^f(\pi, t) : \left(L^2[t_0, t_f]\right)^{|\mathcal{S}| \times |\mathcal{F}|} \times \mathbb{R}_+^1 \longrightarrow \mathcal{H}^1[t_0, t_f]$$

denote the demand for service $i \in \mathcal{S}$ of firm $f \in \mathcal{F}$ and define the vector of all such demands for firm f to be

$$D^f \in \left(\mathcal{H}^1[t_0, t_f]\right)^{|\mathcal{S}|}$$

and all such demands for service $i \in \mathcal{S}$ of firm $f \in \mathcal{F}$

$$D \in \left(\mathcal{H}^1[t_0, t_f]\right)^{|\mathcal{S}|x|\mathcal{F}|}$$

We will use the notation

$$D^{-f} = (D_i^g : i \in \mathcal{S}, g \in \mathcal{F} - \{f\})$$

for the vector of all service levels provided by the competitors of the firm $f \in \mathcal{F}$.

3.2 Demand Dynamics

In evolutionary game theory the notion of comparing a moving average to the current state is used to develop ordinary differential equations describing learning processes; see Fudenberg and Levine (1999). To proceed, we first assume the qualities of services provided by different agents are homogeneous, hence the customers' decision is only dependent on their prices. One possible extension to consider service qualities in demand dynamics is defining a utility function dependent on both price and quality. If the quality levels are exogenous to the problem and remain constant within the time horizon, we can always recast the utility-based model to price-based model. Hence our formulation in this paper is readily applicable to consider quality levels.

The demand for the service offerings of firm $f \in \mathcal{F}$ evolve according to the following evolutionary game-theoretic dynamics:

$$\frac{dD_i^f}{dt} = \eta_i^f \cdot \left(\tilde{\pi}_i - \pi_i^f\right) \qquad \forall i \in \mathcal{S}, \ f \in \mathcal{F} \tag{8}$$

$$D_i^f(t_0) = K_{i,0}^f \qquad \qquad \forall i \in \mathcal{S}, \ f \in \mathcal{F} \tag{9}$$

where $\tilde{\pi}_i$ is the moving average price for service $i \in \mathcal{S}$ given by

$$\tilde{\pi}_i(t) = \frac{1}{|\mathcal{F}|(t - t_0)} \int_{t_0}^t \sum_{g \in \mathcal{F}} \pi_i^g(\tau) d\tau \quad \forall i \in \mathcal{S}$$

while $K_{i,0}^f \in \mathbb{R}_{++}^1$ and $\eta_i^f \in \mathbb{R}_{++}^1$ are exogenous parameters for each $i \in \mathcal{S}$ and $f \in \mathcal{F}$. The firms set the parameter η_i^f by analyzing the past demand data and the sensitivity of the demand with respect to price. The demand for a service type i of a firm f changes over time in accordance with the excess between the

6

firm's price and the moving average of all agents' prices for the particular service. The coefficient η_i^f controls how quickly demand reacts to price changes for each firm f and service type i. Some providers may specialize in certain services and may be able to adjust more quickly than their competitors.

In the literature, most assumes that observed demand in period t and $t+1$ are independent and demand at time t is influenced by price at time t (which is set at time $t-1$ prior observing the demand for period t), this is a stronger assumption and we know in reality there are some learning that takes place in terms of consumer's expectations and price-anticipation. In our model we captured this 'learning' behavior in a naive way: customers have some 'reference price' of a differentiated commodity which gets updates at the end of every period as they learn about the market condition and demand for service/goods is proportional to the price differential (tatonnement dynamics). This way demand at time t not only depends on price set at time t, but on the complete price trajectories $[0, t]$. In this sense, the dynamic equation (8) based on 'moving average price' of demand learning is more realistic. The 'replicator dynamics (see Hofbauer and Sigmund, 1998)', of which equation (8) is an instance, reflects a widely respect theoretical view or hypothesis that learning is the notion of comparison to the moving average.

These dynamics represent a learning mechanism for the firms. As stated here, the dynamics are reminiscent of replicator dynamics which are used in evolutionary games. The rate of growth of demand, can be viewed as the rate of growth of the firm f with respect to service type i. This growth follows the "basic tenet of Darwinism" and may be interpreted as the difference between the fitness (price) of the firm for the service and the rolling average fitness of all the agents for that service.

3.3 Constraints

There are positive upper and lower bounds, based on market regulations or knowledge of customer behavior, on service prices charged by firms. Thus we write

$$\pi_{\min,i}^f \leq \pi_i^f \leq \pi_{\max,i}^f \quad \forall i \in \mathcal{S}, \ f \in \mathcal{F}$$

where the $\pi_{\min,i}^f \in \mathbb{R}_{++}^1$ and $\pi_{\max,i}^f \in \mathbb{R}_{++}^1$ are known constants. Similarly, there will be an upper bound on the demand for services of each type by each firm as negative demand levels are meaningless; that is

$$D_i^f \geq 0 \quad \forall i \in \mathcal{S}, \ f \in \mathcal{F}$$

Let $\mathcal{R}$ be the set of resources that the firms can utilize to provide the services, $|\mathcal{R}|$, cardinality of the set $\mathcal{R}$, denotes number of resources at firm's disposal. Define the incidence matrix (Talluri and van Ryzin, 2004) $A = (a_{lm})$ as

$$a_{lm} = 1 \text{ if resource } l \text{ is used by the service type } m$$
$$= 0 \text{ otherwise}$$

Let C^f be firm f's capacity of resources. Joint resource-constraints for firm f are

$$0 \leq A D_i^f \leq C_i^f \quad \forall i \in \mathcal{S}, \ f \in \mathcal{F} \tag{10}$$

3.4 The Firm's Optimal Control Problem

Since revenue management firms have very little variable costs and high fixed costs, each firm's objective is to maximize revenue which in turn ensures the maximum profit since the variable costs are negligible. We further note that each firm $f \in \mathcal{F}$ faces the following problem: with the π^{-f} as exogenous inputs, solve the following optimal control problem:

$$\max_{\pi^f} J_f(\pi^f, \pi^{-f}, t) = \int_{t_0}^{t_f} e^{-\rho t} \left(\sum_{i \in \mathcal{S}} \pi_i^f \cdot D_i^f \right) dt - e^{-\rho t_0} \Psi_0^f \tag{11}$$

subject to

$$\frac{dD_i^f}{dt} = \eta_i^f \cdot \left(\tilde{\pi}_i - \pi_i^f \right) \quad \forall i \in \mathcal{S}, \ f \in \mathcal{F} \tag{12}$$

$$D_i^f(t_0) = K_{i,0}^f \quad \forall i \in \mathcal{S} \tag{13}$$

$$\pi_{\min,i}^f \le \pi_i^f \le \pi_{\max,i}^f \quad \forall i \in \mathcal{S} \tag{14}$$

$$0 \le AD_i^f \le C_i^f \quad \forall i \in \mathcal{S} \tag{15}$$

where Ψ_0^f is the fixed cost of production for firm f which can be dropped from the problem later, ρ is the nominal discount rate compounded continuously, and $\int_{t_0}^{t_f} e^{-\rho t} \left(\sum_{i \in \mathcal{S}} \pi_i^f \cdot D_i^f \right) dt$ is the net present value (NPV) of revenue. From the familiarity with these dynamics, we may restate them as: for all $f \in \mathcal{F}$

$$\frac{dD_i^f}{dt} = \eta_i^f \cdot \left(\frac{y_i}{|\mathcal{F}|(t - t_0)} - \pi_i^f \right) \quad \forall i \in \mathcal{S} \tag{16}$$

$$\frac{dy_i}{dt} = \sum_{g \in \mathcal{F}} \pi_i^g \quad \forall i \in \mathcal{S} \tag{17}$$

$$N_i^f(t_0) = K_{i,0}^f \quad \forall i \in \mathcal{S} \tag{18}$$

$$y_i(t_0) = 0 \quad \forall i \in \mathcal{S} \tag{19}$$

As a consequence we may rewrite the optimal control problem of firm $f \in \mathcal{F}$ as

$$\max_{\pi^f} J_f(\pi^f, \pi^{-f}, t) = \int_{t_0}^{t_f} e^{-\rho t} \left(\sum_{i \in \mathcal{S}} \pi_i^f \cdot D_i^f \right) dt \tag{20}$$

subject to

$$\frac{dD_i^f}{dt} = \eta_i^f \cdot \left(\frac{y_i}{|\mathcal{F}|(t - t_0)} - \pi_i^f \right) \quad \forall i \in \mathcal{S} \tag{21}$$

$$\frac{dy_i}{dt} = \sum_{g \in \mathcal{F}} \pi_i^g \quad \forall i \in \mathcal{S} \tag{22}$$

$$D_i^f(t_0) = K_{i,0}^f \quad \forall i \in \mathcal{S} \tag{23}$$

$$y_i(t_0) = 0 \quad \forall i \in \mathcal{S} \tag{24}$$

$$\pi_{\min,i}^f \le \pi_i^f \le \pi_{\max,i}^f \quad \forall i \in \mathcal{S} \tag{25}$$

$$0 \le AD_i^f \le C_i^f \quad \forall i \in \mathcal{S} \tag{26}$$

Consequently,

$$D(\pi) = \arg \left\{ \frac{dD_i^f}{dt} = \eta_i^f \cdot \left(\frac{y_i}{|\mathcal{F}|(t - t_0)} - \pi_i^f \right) \right.$$
$$\frac{dy_i}{dt} = \sum_{g \in \mathcal{F}} \pi_i^g; \quad D_i^f(t_0) = K_{i,0}^f$$
$$\left. 0 \le AD_i^f \le C_i^f \quad \forall f \in \mathcal{F}, i \in \mathcal{S} \right\} \tag{27}$$

where we implicitly assume that the dynamics have solutions for all feasible controls. In compact notation this problem can be expressed as: with the π^{-f} as exogenous inputs, compute π^f in order to solve the following optimal control problem:

$$\left. \begin{array}{ll} \max & J_f(\pi^f, \pi^{-f}, t) \\ \text{subject to} & \pi^f \in \Lambda_f \end{array} \right\} \forall f \in \mathcal{F} \tag{28}$$

where
$$\Lambda_f = \left\{\pi^f : (21), (22), (23), (24), (25) \text{ and } (26) \text{ hold}\right\}$$

4 Estimation of Model Parameter

So far, we have introduced a revenue management model for service providers wherein every parameter and dynamics are known to be deterministic. However, in reality, the model parameter are usually unknown to the modeler and the firm, and follows stochastic distributions. Let us suppose the sales season or the total planning time horizon of the service commodities is $[T_0, T_F]$, and we want to update model parameters L times within the season. Hence, the time horizon $[T_0, T_F]$ is divided by L sub-intervals as

$$[T_0, T_0 + \Delta T], \; [T_0 + \Delta T, T_0 + 2\Delta T], \; ..., \; [T_0 + (L-1)\Delta T, T_F]$$

where $\Delta T = (T_F - T_0)/L$. In the time moment at $t = T_0 + l\Delta T$, each firm updates model parameters based on the observation in $[T_0 + (l-1)\Delta T, t = T_0 + l\Delta T]$, for $l = 1, 2, ..., L-1$.

Assuming that the modelling error and observation error follows Normal distributions, in this section, we introduce a forecasting method for the model parameter, η_i^f's, so-called the Kalman filtering, which originated from estimating the state space in electric control systems. The Kalman filter method is to minimize the squared estimation error. The posterior estimates are updated as long as the new observations are available; therefore Kalman filter utilizes all the up-to-date information. For other methods, the estimation is conducted only once and ignores the rich information contained in the new observations. Also, Kalman filter is robust in a sense that it can handle inaccurate observations. Other methods usually assume that the observations are accurate.

For the brevity of the presentation, we drop the superscript f for each firm. Recall the dynamics for the demand is then

$$\frac{dD_i(t)}{dt} = \eta_i \left(\frac{\int_{t_0}^t \sum_{g \in \mathcal{F}} \pi_i^g(\tau) d\tau}{t - t_0} - \pi(t) \right) \tag{29}$$

The speed of change in demand, η_i, is fixed for solving the dynamic pricing problem, but it is stochastic and the real value of it is unknown. After one planning horizon is over, we want to update the model parameter, η_i, so that we would have more precise pricing policy for the next planning horizon. For this purpose, we assume we collect the data during the previous planning horizon. Although we decide the pricing policy in continuous time, most of collecting data activities occur in discrete time in practice. Let us denote such discrete time index by k, and assume we collect data K times in a planning horizon. With the vector notation $\eta = (\eta_i : i \in \mathcal{S})$, the dynamics of η is given by the following process

$$\eta(k+1) = \eta(k) + \xi(k)$$

where $\xi(k)$ is a random noise from a Normal distribution $N(0, Q)$. The matrix Q is known and called the process-noise covariance matrix.

The value of the parameter η can not be observed directly but only by the change of realized demand, which can be defined as

$$z(k) \equiv \Delta D(k) = \eta(k) \left(\frac{\sum_{j=0}^k \sum_{g \in \mathcal{F}} \pi_i^g(j)}{k\Delta k} - \pi(k) \right) \Delta k + \omega(k) \tag{30}$$

where

$$\Delta D(k) = D(k+1) - D(k)$$
$$D = (D_i : i \in \mathcal{S})$$
$$\Delta k \equiv \frac{t_f - t_0}{K}$$

and $\omega(k)$ is a random noise of observation from a normal distribution $N(0, R)$. The matrix R is known and called the measurement noise covariance matrix. The expression (30) is obtained by discretizing the state dynamics (29) of D.

9

Referring Section 12.6 of Bryson and Ho (1975), we obtain the Kalman filter dynamics

$$\hat{\eta}(k) = \bar{\eta}(k) + V(k)[z(k) - H(k)\bar{\eta}(k)]$$
$$\bar{\eta}(k+1) = \bar{\eta}(k)$$
$$P(k) = \left[M(k)^{-1} + H(k)^T R^{-1} H(k)\right]^{-1}$$
$$M(k+1) = P(k) + Q$$

where we defined

$$V(k) \equiv P(k) H(k) R^{-1}$$
$$H(k) \equiv \left(\frac{\sum_{j=0}^{k} \sum_{g \in \mathcal{F}} \pi_i^g(j)}{k \Delta k} - \pi(k)\right) \Delta k$$

and $\bar{\eta}(k)$ is *a priori* estimate of $\eta(k)$ before observation and $\hat{\eta}(k)$ is *a posteriori* estimate after observation. Hence, the process of the estimation is followed:

1. Observe the previous planning time horizon to obtain sequences of exercised optimal price $\{\pi(k)\}$ and observed change in demand $\{z(k)\}$.

2. Let the time index $k = 0$ and assume the initial values $P_0 = 1$ and $\bar{\eta}_0 = \eta$ (the value we used for the previous planning).

3. At index k, we forecast

$$\bar{\eta}(k+1) = \bar{\eta}(k)$$
$$M(k+1) = P(k) + Q$$

4. Based on the observation z_{k+1}, we update

$$\hat{\eta}(k) = \bar{\eta}(k) + V(k)[z(k) - H(k)\bar{\eta}(k)]$$

   where

$$H(k) \equiv \left(\frac{\sum_{j=0}^{k} \sum_{g \in \mathcal{F}} \pi_i^g(j)}{k \Delta k} - \pi(k)\right) \Delta k$$
$$P(k+1) = \left[M(k+1)^{-1} + H(k+1)^T R^{-1} H(k+1)\right]^{-1}$$
$$V(k+1) \equiv P(k+1) H(k+1) R^{-1}$$

5. Let $\bar{\eta}(k+1) = \hat{\eta}(k+1)$.

6. If $k = K$ stop; otherwise set $k = k+1$ and go to step 3.

When we finish the estimation process, we have $\bar{\eta}(K)$ at hand, which is the value of η which will be used in the planning for the next time horizon.

5   A Numerical Example of the Single Firm's Problem

Before we proceed to the game-theoretic model for the competition of service providers, we first describe the solution method, so-called the gradient projection algorithm, for the case when only one firm exists in the market. We obtain an optimal control problem, in which the criterion is nonlinear, the state dynamics are linear, and the control set is convex and compact with a state-space constraint (10). Although the gradient projection algorithm is very popular among optimal control researchers, it is not easy to find a written

statement of the algorithm in the revenue management literature. More information such as convergence and varieties of the method is found in Minoux (1986), Polak (1971) and Bryson Jr. (1999).

To continue discussion in general form of the problem, we penaltize the state-space constraint so that the criteria becomes

$$\max_{\pi^f} J_f(\pi^f, \pi^{-f}, t) = \int_{t_0}^{t_1} \left[e^{-\rho t} \left(\sum_{i \in \mathcal{S}} \pi_i^f \cdot D_i^f \right) - \frac{\mu}{2} \left\{ \max \left(0, AD^f - C^f \right)^2 + \min \left(0, D^f \right)^2 \right\} \right] dt$$

where μ is the penalty coefficient. With the above penalty functions, our problem is reduced to the following general class of optimal control problems:

$$\min J(u) = K\left[x(t_f), t_f\right] + \int_{t_0}^{t_f} f_0\left((x, u, t\right) dt \tag{31}$$

subject to

$$\frac{dx}{dt} = f(x, u, t) \tag{32}$$

$$x(t_0) = x^0 \tag{33}$$

$$u \in U \subset V \tag{34}$$

where x^0 is a known, fixed vector and both t_0 and t_f are fixed. In addition V is a Hilbert space, in particular, $V = \left(L^2\left[t_0, t_f\right]\right)^m$ which is a m-fold square-integrable space.

Note that

$$H(x, u, \lambda, t) = f_0(x, u, t) + \lambda^T f(x, u, t)$$

is the Hamiltonian for the unconstrained problem (31), (32) and (33).

Gradient Projection Algorithm

Step 0. Initialization. Set $k = 0$ and pick $u^0(t) \in \left(L^2\left[t_0, t_f\right]\right)^m$.

Step 1. Find State Trajectory. Using $u^k(t)$ solve the state initial value problem

$$\frac{dx}{dt} = f\left(x, u^0, t\right)$$

$$x(t_0) = x^0$$

and call the solution $x^k(t)$.

Step 2. Find Adjoint Trajectory. Using $u^k(t)$ and $x^k(t)$ solve the adjoint final value problem

$$(-1)\frac{d\lambda}{dt} = \frac{\partial H\left(x^k, u^k, \lambda, t\right)}{\partial x}$$

$$\lambda(t_f) = \frac{\partial K\left[x(t_f), t_f\right]}{\partial x}$$

and call the solution $\lambda^k(t)$.

Step 3. Find Gradient. Using $u^k(t)$, $x^k(t)$ and $\lambda^k(t)$ calculate

$$\nabla_u J(u^k) = \frac{\partial H\left(x^k, u^k, \lambda, t\right)}{\partial u}$$

$$= \frac{\partial f_0(x^k, u^k, t)}{\partial u} + \left(\lambda^k\right)^T \frac{\partial f\left(x^k, u^k, t\right)}{\partial u}$$

Step 4. Update and Apply Stopping Test. For a suitably small step size θ_k, update according to

$$u^{k+1} = P_U\left[u^k - \theta_k \nabla J\left(u^k\right)\right]$$

11

	Revenue
a priori	$119,990
observed	$145,180
a posteriori	$154,110

Table 1: A priori, observed and a posteriori revenue of the firm

where $P_U[\cdot]$ denotes the minimum norm projection onto U. If an appropriate stopping test is satisfied, declare

$$u^*(t) \approx u^{k+1}(t)$$

Otherwise set $k = k + 1$ and go to Step1.

5.1 A Numerical Example

Let us consider a market where a single service provider f is offering four services. The planning horizon for this problem is a month or 30 days. Each firm updates the model parameter once based on daily demand observations. The firm wants to maximize the revenue with the given parameters:

$$\eta = \begin{bmatrix} 0.1 \\ 0.08 \\ 0.12 \\ 0.09 \end{bmatrix}$$

$$\pi_{\min} = \begin{bmatrix} 30 \\ 40 \\ 60 \\ 130 \end{bmatrix}, \pi_{\max} = \begin{bmatrix} 85 \\ 135 \\ 180 \\ 205 \end{bmatrix}, A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 300 \\ 210 \\ 150 \\ 60 \\ 255 \end{bmatrix}$$

Once they have the optimal pricing policy, they exercise it for the planning horizon and observe what really happens in the market. The observation of the actual realized demand occurs every week so that the firm has 52 observations for the year. With this data, the firm learns the model parameter η with the process-noise covariance and the measurement noise covariance matrices are given as

$$Q = \begin{bmatrix} 0.3 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.5 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.2 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.3 \end{bmatrix}, R = \begin{bmatrix} 0.2 & 0.2 & 0.1 & 0.1 \\ 0.1 & 0.3 & 0.2 & 0.1 \\ 0.1 & 0.2 & 0.2 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.3 \end{bmatrix}$$

As the results are presented in Table 1, the observed revenue is higher than the *a priori* revenue estimation which is obtained by solving the RM model with the initial η. This means that the actual speed of change in demand dynamics is slower than the value used in planning. Hence, the firm could increase price of service without worrying the demand to be much decreased. With the Kalman filter dynamics, we adjust the parameter η as

$$\eta_{\text{adjusted}} = \begin{bmatrix} 0.0403 \\ 0.0436 \\ 0.0775 \\ 0.0471 \end{bmatrix}$$

which are less than the original value as we expected. With this new forecast, we perform another dynamic price planning for the next year.

Figure 1: *A priori* demand, observed demand and *a posteriori* demand for each service

Figure 2: *A priori* optimal price and *a posteriori* optimal price for each service

The demand and price for each service is presented in Figures 1 and 2, respectively. After the estimation, the service prices are set at higher level generating more revenue. We should note that the problem solving after estimation is *not* to follow the optimal pricing rule associated with the observed demand, but rather to obtain more accurate pricing rule with updated parameters. Although the revenue to be generated at the next planning could be less than planned in the reality, it would be even worse if we do not estimate the exact value of the model parameter.

6   DVI Formulation of the Competition

Each service provider is a Cournot-Nash agent that knows and employs the current instantaneous values of the decision variables of other firms to make its own non-cooperative decisions. Therefore (28) defines a set of coupled optimal control problems, one for each firm $f \in \mathcal{F}$. It is useful to note that (28) is an optimal control problem with fixed terminal time and fixed terminal state. Its Hamiltonian is

$$H_f \left(\pi^f; D^f; \lambda^f; \sigma, \alpha^f, \beta^f; \pi^{-f}; t \right) \equiv e^{-\rho t} \left(\sum_{i \in \mathcal{S}} \pi_i^f \cdot D_i^f \right) + \Phi_f \left(\pi^f; D^f; \lambda^f; \sigma^f; \alpha^f; \beta^f; \pi^{-f} \right) \qquad (35)$$

where

$$\Phi_f \left(\pi^f; D^f; \lambda^f; \sigma, \alpha^f, \beta^f; \pi^{-f} \right) = \left\{ \sum_{i \in \mathcal{S}} \lambda_i^f \left[\eta_i^f \cdot \left(\frac{y_i}{|\mathcal{F}| \, (t - t_0)} - \pi_i^f \right) \right] + \sum_{i \in \mathcal{S}} \sigma_i \left(\pi_i^f + \sum_{g \in \mathcal{F} \setminus f} \pi_i^g \right) \right.$$

$$\left. + \sum_{i \in \mathcal{S}} \alpha_i^f \left(-D_i^f \right) + \sum_{i \in \mathcal{S}} \beta_i^f \left(AD_i^f - C_i^f \right) \right\} \qquad (36)$$

while $\lambda_i^f \in \mathcal{H}^1 \left[t_0, t_f \right]$ is the adjoint variable for the dynamics associated with the firm f with service type i and $\lambda \in \left(\mathcal{H}^1 \left[t_0, t_f \right] \right)^{|\mathcal{F}| \times |\mathcal{S}|}$, $\sigma_i \in \mathbb{R}_+^1, \alpha_i^f \in \mathbb{R}_+^1$ and $\beta_i^f \in \mathbb{R}_+^1$ are the dual variables arising from the auxiliary state variables $(y_i, i \in \mathcal{S})$ and state space constraints where $\sigma \in \mathbb{R}_+^{|\mathcal{S}|}, \alpha^f \in \mathbb{R}_+^{|\mathcal{S}|}$ and $\beta^f \in \mathbb{R}_+^{|\mathcal{S}|}$. The instantaneous revenue for firm f is $\sum_{i \in \mathcal{S}} \pi_i^f \cdot D_i^f$. We assume in the balance of this paper that the game (28) is regular in the sense of Definition 1. Therefore, the maximum principle (Bryson and Ho, 1975) tells us that an optimal solution to (28) is a sextuplet $\left\{ \pi^{f*} (t), D^{f*} (t); \lambda^{f*} (t), \sigma^*, \alpha^{*f}, \beta^{*f} \right\}$ requiring that the nonlinear program

$$\max H_f \quad \text{s.t.} \quad \pi_{\min}^f \leq \pi^f \leq \pi_{\max}^f$$

be solved by every firm $f \in \mathcal{F}$ for every instant of time $t \in [t_0, t_f]$ where

$$\pi_{\min}^f = \left\{ \pi_i^f : i \in \mathcal{S} \right\}$$

$$\pi_{\max}^f = \left\{ \pi_i^f : i \in \mathcal{S} \right\}$$

Consequently, any optimal solution must satisfy at each time $t \in [t_0, t_f]$:

$$\pi^{f*} = \arg \left\{ \max_{\pi_{\min}^f \leq \pi^f \leq \pi_{\max}^f} H_f \left(\pi^f; D^f; \lambda^f; \sigma, \alpha^f, \beta^f; \pi^{-*f}; t \right) \right\} \qquad (37)$$

which in turn, by virtue of regularity, is equivalent to

$$\left[\nabla_{\pi^{*f}} H_f^*\right]^T \left(\pi^f - \pi^{*f}\right) \leq 0 \quad \text{for all} \ \left(\begin{array}{c} \pi_{\min}^f \\ \pi_{\min}^f \end{array} \right) \leq \left(\begin{array}{c} \pi^{*f} \\ \pi^f \end{array} \right) \leq \left(\begin{array}{c} \pi_{\max}^f \\ \pi_{\max}^f \end{array} \right) \tag{38}$$

where

$$H_f^* \equiv e^{-\rho t} \left(\sum_{i \in \mathcal{S}} \pi_i^{*f} \cdot D_i^{*f} \right) + \Phi_f^* \tag{39}$$

and

$$\Phi_f^* = \Phi_f \left(\pi^{*f}; D^{*f}; \lambda^{*f}; \sigma^{*f}; \alpha^{*f}; \beta^{*f}; \pi^{-f} \right) \tag{40}$$

From (35)

$$\left[\nabla_{\pi^{*f}} \left[e^{-\rho t} \left(\sum_{i \in \mathcal{S}} \pi_i^{*f} \cdot D_i^{*f} \right) \right] + \nabla_{\pi^{*f}} \Phi_f^* \right]^T \left(\pi^f - \pi^{*f} \right) \leq 0$$

$$\text{for all} \ \left(\begin{array}{c} \pi_{\min}^f \\ \pi_{\min}^f \end{array} \right) \leq \left(\begin{array}{c} \pi^{*f} \\ \pi^f \end{array} \right) \leq \left(\begin{array}{c} \pi_{\max}^f \\ \pi_{\max}^f \end{array} \right)$$

Further, adjoint dynamics and state dynamics govern

$$\frac{\partial H_f}{\partial D_i^f} = (-1) \frac{d\lambda_i^{f*}}{dt} \tag{41}$$

$$\frac{\partial H_f}{\partial \lambda_i^f} = \frac{dD_i^{f*}}{dt} \tag{42}$$

while due to absence of terminal time constraint, the transversality condition gives

$$\lambda^{f*}(t_f) = \gamma^T \frac{\partial \Gamma \left[D^{f*}(t_f), t_f \right]}{\partial D^{f*}(t_f)} = 0$$

which gives rise to a two point boundary value problem.

With this preceding background, we are now in a position to create a variational inequality for the non-cooperative competition amongst the firms. We consider the following DVI which has solutions that are Cournot-Nash equilibria for the game described above in which individual firms maximize their revenue in light of current information about their competitors:

$$\text{find} \ \pi^* \in \Omega \ \text{such that}$$

$$\int_{t_0}^{t_f} \left(\sum_{s \in \mathcal{S}} \sum_{f \in \mathcal{F}} \frac{\partial H_f^*}{\partial \pi_i^f} \left(\pi_i^f - \pi_i^{f*} \right) \right) dt \leq 0 \tag{43}$$

$$\text{for all} \ \pi \in \Omega = \prod_{f \in \mathcal{F}} \Omega_f$$

where H_f^* is defined in (39)-(40) and

$$\Omega_f = \left\{ \pi^f : \pi_{\min}^f \leq \pi^f \leq \pi_{\max}^f \right\}$$

This DVI is regular in the sense of Definition 1 and a convenient way of expressing the Cournot-Nash game that is our present interest.

7   A Descent Method for Solving DVI

When the regularity conditions given in Definition 1 holds, $DVI(F, f, U)$ is reduced to the class of variational inequalities in a Hilbert space $\mathcal{H}$ considered by Konnov et al. (2002); i.e. U is a non-empty closed and convex subset of a Hilbert space, and F is a continuously differentiable mapping of u. This allows us to analyze $DVI(F, f, U)$ by considering gap functions, which are defined by

Definition 4 *A function* $\varphi : U \longrightarrow \mathbb{R}_+$ *is called a gap function for* $DVI(F, f, U)$ *when the following statements hold:*

1. $\varphi(u) \geq 0$ *for all* $u \in U$

2. $\varphi(u) = 0$ *if and only if* u *is the solution of* $DVI(F, f, U)$

Let us consider a function

$$\varphi_\alpha(u) = \max_{v \in U} \Phi_\alpha(u, v) \tag{44}$$

where

$$\Phi_\alpha(u, v) = \langle F(x, u, t), u - v \rangle - \alpha\phi(u, v)$$

$$x(u, t) = \arg \left\{ \frac{dy}{dt} = f(y, u, t), \; y(t_0) = y^0 \right\} \in \left(\mathcal{H}^1[t_0, t_f] \right)^n$$

$$U \subseteq \left(L^2[t_0, t_f] \right)^m$$

and we assume ϕ is a function that satisfies following assumptions: (1) ϕ is continuously differentiable on $\left(L^2[t_0, t_f] \right)^{2m}$; (2) ϕ is non-negative on $\left(L^2[t_0, t_f] \right)^{2m}$; (3) $\phi(u, \cdot)$ is strongly convex with modulus $c > 0$ for any $u \in \left(L^2[t_0, t_f] \right)^m$; and (4) $\phi(u, v) = 0$ if and only if $u = v$. The maximization problem (44) has a unique solution since $\Phi_\alpha(u, v)$ is strongly convex in v and U is convex. Let the solution of (44) $v_\alpha(u)$ such that $\varphi_\alpha(u) = \Phi_\alpha(u, v_\alpha(u))$, then we have the following result from Kwon and Friesz (2007):

Lemma 5 *The function* $\varphi_\alpha(u)$ *defined by (44) is a gap function for* $DVI(F, f, U)$*, and* u *is the solution to* $DVI(F, f, U)$ *if and only if* $u = v_\alpha(u)$*.*

In addition to the regular gap functions, let us introduce a special class of gap functions, so-called D-gap functions, which are of the form

$$\psi_{\alpha\beta}(u) = \varphi_\alpha(u) - \varphi_\beta(u)$$

for $0 < \alpha < \beta$. While $\varphi_\alpha(u)$ is not differentiable in general, $\psi_{\alpha\beta}(u)$ is Gateaux-differentiable. Among many gap functions, Fukushima (1992) considered a case of gap function, the regularized gap function, and Konnov et al. (2002) extended it to Hilbert spaces; in particular,

$$\phi(u, v) = \frac{1}{2} \|v - u\|^2 \tag{45}$$

which satisfies the assumption on $\phi(\cdot)$, the strong convexity with modulus $c > 0$. Adopting Fukushima (1992)'s gap function, we obtain

$$\Phi_\alpha(u, v) = \langle F(x, u, t), u - v \rangle - \frac{\alpha}{2} \|v - u\|^2$$

The corresponding D-gap function becomes

$$\psi_{\alpha\beta}(u) = \langle F(x, u, t), v_\beta(u) - v_\alpha(u) \rangle - \frac{\alpha}{2} \|v_\alpha(u) - u\|^2 + \frac{\beta}{2} \|v_\beta(u) - u\|^2 \tag{46}$$

where we denote

$$v_\alpha(u) = \arg \max_{v \in U} \Phi_\alpha(u, v) \tag{47}$$

$$v_\beta(u) = \arg \max_{v \in U} \Phi_\beta(u, v) \tag{48}$$

15

With the D-gap function (46), $DVI(F, f, U)$ is equivalent to the following optimal control problem $OCP(\psi_{\alpha\beta}, f, U)$:

$$\min \psi_{\alpha\beta}(u)$$

$$= \langle F(x, u, t), v_\beta(u) - v_\alpha(u) \rangle - \frac{\alpha}{2} \|v_\alpha(u) - u\|^2 + \frac{\beta}{2} \|v_\beta(u) - u\|^2$$

$$= \int_{t_0}^{t_f} \left\{ F(x, u, t) [v_\beta(u) - v_\alpha(u)] - \frac{\alpha}{2} [v_\alpha(u) - u]^2 + \frac{\beta}{2} [v_\beta(u) - u]^2 \right\} dt \qquad (49)$$

subject to

$$\frac{dx}{dt} = f(x, u, t) \qquad (50)$$

$$x(0) = x_0 \qquad (51)$$

$$u \in U \qquad (52)$$

This is a Bolza form of standard optimal control problem, except the objective functional involves the maximizers of subproblems defined by (47) and (48), $v_\alpha(u)$ and $v_\beta(u)$.

Now we are interested in the gradient of the objective functional $\psi_{\alpha\beta}(u)$, which is, in fact, equivalent to the gradient of the corresponding Hamiltonian function in the theory of optimal control. Let us define the Hamiltonian function for $OCP(\psi_{\alpha\beta}, f, U)$

$$H(x, u, \lambda, t) = F(x, u, t) [v_\beta(u) - v_\alpha(u)] - \frac{\alpha}{2} [v_\alpha(u) - u]^2 + \frac{\beta}{2} [v_\beta(u) - u]^2 + \lambda f(x, u, t) \qquad (53)$$

The gradient of $\psi_{\alpha\beta}(u)$ is determined as following:

Theorem 6 *Suppose $F(x, u, t)$ is Lipschitz continuous on every bounded subset of $\left(L^2[t_0, t_f]\right)^m$. Then $\psi_{\alpha\beta}(u)$ is continuously differentiable in the sense of Gateaux, and*

$$\nabla \psi_{\alpha\beta}(u) = \frac{\partial}{\partial u} H(x, u, \lambda, t)$$

$$= \frac{\partial F(x, u, t)}{\partial u} [v_\beta(u) - v_\alpha(u)]$$

$$+ \alpha [v_\alpha(u) - u] - \beta [v_\beta(u) - u] + \lambda \frac{\partial f(x, u, t)}{\partial u}$$

We propose the following descent algorithm for $OCP(\psi_{\alpha\beta}, f, U)$, in which the main philosophy remains same as in usual optimal control problems except we need to solve the state dynamics, the adjoint dynamics and two sub-maximization problems for v_α and v_β to obtain the current information of the descent direction.

Step 0. Initialization. Choose $0 < \alpha < \beta$. Pick $u^k(t) \in U$ and set $k = 0$.

Step 1. Finding state variables. Solve the state dynamics

$$\frac{dx}{dt} = f(x, u^k, t)$$

$$x(0) = x_0$$

and call the solution $x^k(t)$.

Step 2. Finding adjoint variables. Solve the adjoint dynamics

$$-\frac{d\lambda}{dt} = \nabla_x H(x^k, u^k, \lambda, t)$$

$$= \frac{\partial F(x^k, u^k, t)}{\partial x} [v_\beta(u^k) - v_\alpha(u^k)] + \lambda \frac{\partial f(x^k, u^k, t)}{\partial x}$$

$$\lambda(t_f) = 0$$

and call the solution $\lambda^k(t)$, where

$$v_\alpha\left(u^k\right) = P_U\left[u^k - \frac{1}{\alpha}F\left(x^k, u^k, t\right)\right]$$

$$v_\beta\left(u^k\right) = P_U\left[u^k - \frac{1}{\beta}F\left(x^k, u^k, t\right)\right]$$

Step 3. Finding the gradient. Determine

$$\begin{aligned}
\nabla\psi_{\alpha\beta}^k(t) &= \nabla_u H\left(x^k, u^k, \lambda^k, t\right) \\
&= \frac{\partial F\left(x^k, u^k, t\right)}{\partial u}\left[v_\beta\left(u^k\right) - v_\alpha\left(u^k\right)\right] \\
&+ \alpha\left[v_\alpha\left(u^k\right) - u^k\right] - \beta\left[v_\beta\left(u^k\right) - u^k\right] + \lambda\frac{\partial f\left(x^k, u^k, t\right)}{\partial u}
\end{aligned}$$

Step 4. Updating the current control. For a suitably small step size

$$\theta_k \in \mathbb{R}_{++}^1$$

determine

$$u^{k+1}(t) = P_U\left[u^k(t) - \theta_k\nabla\psi_{\alpha\beta}^k(t)\right]$$

Step 5. Stopping Test. For $\epsilon \in \mathbb{R}_{++}^1$, a preset tolerance, stop if

$$\|\psi_{\alpha\beta}\left(u^{k+1}\right)\| < \epsilon$$

and declare

$$u^* \approx u^{k+1}$$

Otherwise set $k = k + 1$ and go to Step1.

The convergence result is stated as following without proof:

Theorem 7 (Convergence) *Suppose the functional $\psi_{\alpha\beta} : U \longrightarrow \Re_+$ is strongly convex with modulus $\rho > 0$ and $\nabla\psi_{\alpha\beta}(u)$ is defined and satisfies the Lipschitz condition*

$$\|\nabla\psi_{\alpha\beta}\left(u_1\right) - \nabla\psi_{\alpha\beta}\left(u_2\right)\| \le \delta\|u_1 - u_2\| \tag{54}$$

for all $u_1, u_2 \in U$. Then the decent algorithm converges to the minimum u^* of $\psi_{\alpha\beta}$ on U for step size choices

$$\theta \in \left(0, \frac{2\rho}{\delta^2}\right) \tag{55}$$

The proofs of theorems in this section and detail discussions for gap functions are found in Kwon and Friesz (2007).

8   A Numerical Example for the Competition with Model Parameter Update

Consider a market where two service providers $f = 1, 2$ are offering four services. The planning horizon for this problem is a month or 30 days. Each firm updates the model parameter once based on daily demand observations. The speed of change in demand η is assumed taking the values

	Service type, i	1	2	3	4
$\eta :$	Firm 1	0.10	0.08	0.12	0.09
	Firm 2	0.11	0.07	0.10	0.15

	Revenue, Firm 1	Revenue, Firm 2
before estimation	$65,060	$48,720
observed	$90,100	$71,870
after estimation	$396,900	$440,010

Table 2: A priori, observed and a posteriori revenue of the firm

Each firm's initial demand at time t_0 for services K_0 is shown below.

K_0 :

Service type, i	1	2	3	4
Firm 1	10.0	17.5	22.5	30.0
Firm 2	9.5	16.5	20.0	31.0

The maximum resource of available for each firm C is considered as

C :

Resource type, j	1	2	3	4	5
Firm 1	300	210	150	60	255
Firm 2	180	150	120	75	210

Assume the maximum price $\pi_{\max}$ and $\pi_{\min}$ are given as

$\pi_{\max}$:

Service type, i	1	2	3	4
Firm 1	85	135	180	205
Firm 2	75	108	185	210

and

$\pi_{\min}$:

Service type, i	1	2	3	4
Firm 1	30	40	60	130
Firm 2	45	50	65	115

Once they have the optimal pricing policy, they exercise it for the planning horizon and observe what really happens in the market. The observation of the actual realized demand occurs every week so that the firm has 52 observations for the year. With this data, the firm learns the model parameter η with the process-noise covariance and the measurement noise covariance matrices are given for both firms as

$$Q = \begin{bmatrix} 0.3 & 0.1 & 0.2 & 0.2 \\ 0.1 & 0.5 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.2 & 0.1 \\ 0.1 & 0.2 & 0.2 & 0.3 \end{bmatrix}, R = \begin{bmatrix} 0.2 & 0.2 & 0.1 & 0.1 \\ 0.1 & 0.3 & 0.2 & 0.1 \\ 0.1 & 0.2 & 0.2 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.3 \end{bmatrix}$$

As the results are presented in Table 2, the observed revenue is higher than the projected *a priori* best revenue estimation. This means that the actual speed of change in demand is slower than used in planning. Then the firm could increase price of service without worrying the demand to be much decreased. With the Kalman filter dynamics, we adjust the parameter η as

$$\eta_{\text{adjusted}} = \begin{bmatrix} 0.0403 \\ 0.0436 \\ 0.0775 \\ 0.0471 \end{bmatrix}$$

which are less than the original value as we expected. With this new forecast, we perform another dynamic price planning for the next year.

The demand trajectories for each service is presented in Figures 3 and 4 for firms 1 and 2, respectively. We observe that demand after estimation is less sensitive to the price than the demand before estimation

Figure 3: Firm 1's demand trajectories (before estimation, observed, after estimation) for different services

Figure 4: Firm 2's demand trajectories (before estimation, observed, after estimation) for different services

as the parameter value η is smaller. The trajectories of the price charged for each service is presented in Figures 5 and 6 for firms 1 and 2, respectively. The prices are relatively at higher levels in the simulation after estimation than before estimation. Although the revenue to be generated at the next year could be less than planned in the reality, we would have more accurate dynamic pricing policy, as the learning process repeats.

We note that the price will not reach a stationary state, in general. The optimal price trajectories we would obtain in each sub-time horizon are up-to-date reflections of customers' choice behavior, hence the price trajectories will depend on current market state which is uncertain. However, if there is no shock in market and we have a long enough time horizon, the price trajectory must reach a stationary state.

8.1 Computational Performance

The descent method described earlier converged after 10 iterations for this numerical example with the preset tolerance $\varepsilon = 10^{-4}$. In Figure 7, the objective values, which is the gap function value, at each iteration for a priori and a posteriori problem solving are presented. Even though the sizes of the gap function are very big at the first iterations for both problems, they decrease very rapidly towards zero, at which the equilibrium point exists. The run time for this example is less than 10 seconds using a desktop computer with a Intel Xeon processor and 4 GB RAM. The computer code for the descent method is written in MATLAB 7.0.

9 Concluding Remarks

This paper has provided a foundation for the evolutionary dynamic games with demand fluctuation and a forecast methodology for model parameters. In addition, numerical methods for optimal control problems and differential variational inequalities were described and shown to be effective in solving the revenue management problems we introduced. In the model and examples we have presented, it is assumed that the model parameters remain constant over the planning time horizon. At the end of the horizon, each firm estimates its own model parameter using the Kalman filter based on the collection of observations gathered in discrete times. The Kalman filtering has been proven to be one of the most effective and efficient tool to remove the uncertainty in system dynamics and observation errors in many applications. However, each firm does not update its parameter in real-time in the model developed in this paper. This motivates to study more sophisticated models in the future. One possible extension will be the feedback control model using the Kalman filter in real-time, although the assumption that the economic observations are made continuously may be under debate. When we consider a firm's optimal control problem, if the criterion is quadratic and the dynamics is linear, so-called LQ control problem, it is well-known that the LQ regulator works very well in most cases. However, the revenue management model usually does not have quadratic criterion, while the system dynamics are most likely linear in state and control. Therefore, it will be desired that numerical and/or analytical approximation techniques are developed.

Figure 5: Firm 1's price trajectories (before estimation, after estimation) for different services

Figure 6: Firm 2's price trajectories (before estimation, after estimation) for different services

Figure 7: Change in gap function values

References

Aubin JP, Cellina A, Differential Inclusions. Springer-Verlag 1984.

Balvers RJ, Cosimano TF. Actively learning about demand and the dynamics of price adjustment. The Economic Journal 1990;100; 882–898.

Bertsimas D, Perakis G, Dynamic pricing: A learning approach. In: Lawphongpanich S, Hearn DW, Smith MJ (Eds), Mathematical and Computational Models for Congestion Charging, pp. 45–79, Springer US 2006;.

Bitran GR, Wadhwa HK, A methodology for demand learning with an application to the optimal pricing of seasonal products 1996, M.I.T. Sloan School of Management, Working Paper #3896-96.

Bressan A, Piccoli B, An Introduction to the Mathematical Theory of Control. Lecture Notes for a course on Control Theory, Penn State 2005.

Bryson AE, Ho YC, Applied Optimal Control. Hemisphere Publishing Company 1975.

Bryson Jr AE, Dynamic Optimization. Addison Wesley 1999.

Carvalho AX, Puterman ML, Dynamic pricing and learning over short time horizons 2007, working paper, under revision.

Dasgupta P, Das R, Dynamic pricing with limited competitor information in a multi-agent economy. In: 7th. International Conference on Cooperative Systems, Springler-Verlag 2000; pp. 299 – 310.

Eppen G, Iyer A. Improved Fashion Buying with Bayesian Updates. Operations Research 1997;45(6); 805–819.

Feng Y, Gallego G. Optimal starting times for end-of-season sales and optimal stopping times for promotional fares. Management Science 1995;41; 1371–1391.

Friesz T, Rigdon M, Mookherjee R. Differential variational inequalities and shipper dynamic oligopolistic network competition. Transportation Research Part B: Methodological 2006;40(6); 480–503.

Fudenberg D, Levine DK, The Theory of Learning in Games, Second edition. The MIT Press 1999.

Fukushima M. Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. Mathematical Programming 1992;53; 99–110.

Gallego G, van Ryzin GJ. Optimal dynamic pricing of inventories with stochastic demand over finite horizon. Management Science 1994;40; 999–1020.

Hofbauer J, Sigmund K, Evolutionary Games and Replicator Dynamics. Cambridge University Press 1998.

Kachani S, Perakis G, Simon C, A transient model for joint pricing and demand learning under competition. presented in the 4th Annual INFORMS Revenue Management and Pricing Section Conference, MIT, Cambridge 2004.

Kalman R. A new approach to linear filtering and prediction problems. Journal of Basic Engineering 1960; 82(1); 35–45.

Kalman R, Bucy R. New results in linear filtering and prediction theory. Journal of Basic Engineering 1961; 83(3); 95–108.

Konnov IV, Kum S, Lee GM. On convergence of descent methods for variational inequalities in a Hilbert space. Mathematical Methods of Operations Research 2002;55; 371–382.

Kwon C, Friesz TL, A descent method for differential variational inequalities and applications in differential games 2007, the Pennsylvania State University, working paper.

Lin KY. Dynamic pricing with real-time demand learning. European Journal of Operational Research 2006; 174; 522–538.

McGill J, van Ryzin G. Revenue management: Research overview and prospects. Transportation Science 1999;33(2); 233–256.

Minoux M, Mathematical Programming: Theory and Algorithms. John Wiley and Sons 1986.

Mookherjee R, A Computable Theory of Dynamic Games and Its Applications. Ph.D. thesis, The Pennsylvania State University 2006.

Murray Jr G, Silver E. A Bayesian Analysis of the Style Goods Inventory Problem. Management Science 1966;12(11); 785–797.

Pang JS, Stewart D. Differential variational inequalities. Mathematical Programming 2007;forthcoming.

Polak E, Computational Methods in Optimization. Academic Press 1971.

Talluri KT, van Ryzin GJ, The Theory and Practice of Revenue Management. Springler 2004.

Xie J, Song XM, Sirbu M, Wang Q. Kalman filter estimation of new product diffusion models. Journal of Marketing Research 1997;34; 378–393.

Yamashita N, Taji K, Fukushima M. Unconstrained optimization reformulations of variational inequality problems. Journal of Optimization Theory and Applications 1997;92(3); 439–456.

Yelland PM, Lee E, Forecasting product sales with dynamic linear mixture models. Tech. rep., Sun Microsystems Laboratories, Technical Report TR-2003-122 2003.

Zhu DL, Marcotte P. An extended descent framework for variational inequalities. Journal of Optimization Theory and Applications 1994;80(2); 349–366.

Zhu DL, Marcotte P. Convergence properties of feasible descent methods for solving variational inequalities in Banach spaces. Computational Optimization and Applications 1998;10(1); 35–49.