Intertemporal Price Discrimination with Time-Varying Valuations

Victor F. Araman

Olayan School of Business, American University of Beirut, Beirut, Lebanon, va03@aub.edu.lb

Bassam Fayad

Institut Mathématique de Jussieu, Paris, France, bassam.fayad@pmj-prj.fr

May 31, 2017

Abstract

A firm that sells a non perishable product considers intertemporal price discrimination in the objective of maximizing the long-run average revenue. Each period, a number of interested customers approach the firm and can either purchase on arrival, or remain in the *system* for a period of time. During this time, each customer's valuation changes following a discrete and homogenous Markov chain. Customers leave the system if they either purchase at some point, or their valuations reach an absorbing state v_0. We show that, in this context, cyclic strategies are optimal, or nearly optimal. When the pace of intertemporal pricing is constrained to be comparable to customers patience level, we have a good control on the cycle length and on the structure of the optimizing cyclic policies. We also obtain an algorithm that yields the optimal (or near optimal) cyclic solutions in polynomial time in the number of prices. We cast part of our results in a general framework of optimizing the long-run average revenues for a class of payoffs that we call *weakly coupled*, in which the revenue per period depends on a finite number of neighboring prices.

1 Introduction

It is difficult for a customer to estimate, before a purchase, the value that he can generate from a product. A product's value is usually subjective and is implied by the information available before the purchase occurs. When customers are interested in a product, they often go through a so-called, *purchasing decision process*, during which they get exposed to a stream of information that continuously affects their perceived value, and therefore, their willingness-to-pay (wtp). Consider for instance a customer who is looking into buying a brand new camera. The wtp could initially depend on the previous experience the customer had with that brand and the urgency to get the product. However, this valuation can change, for instance, according to the reviews, that such brand model is scoring, the specific features and the specs offered, the marketing campaigns the customer is being exposed to (e.g. targeted online advertising), the introduction of a new model, the availability and the prices of other competing products, and so on. Such information is received through time - starting from the moment the customer shows interest in the product - resulting in a dynamic update of the valuation.

Such behavior has been generally overlooked in the pricing literature where, often customers are assumed to have a reservation price (possibly drawn from some given distribution) constant throughout the sales process. In this work, we consider a continuous flux of customers interested in purchasing a product but who see their valuations changing through time. We try to shed some lights on the impact that such varying valuations have on the seller's optimal pricing policy. We consider for that a simple setting where a non perishable product is being sold throughout an infinite horizon during which the seller has *committed* to a sequence of prices. Heterogenous customers get interested by the product and approach the seller at different times. They could purchase on arrival or decide to remain in the system for a (random) period of time during which they will see their valuations stochastically changing following a discrete Markov chain. Customers' heterogeneity is modeled through the initial valuation each customer has of the product. The type of questions we are interested in are whether, in this context, there is an opportunity for price discrimination; and if there is one, how to go about characterizing it formally. We specifically examine how optimal and numerically efficient cyclic policies are, especially since such policies have been proven to be optimal in similar settings in the literature.

Intertemporal pricing is experienced in practice and often argued for as a way to price discriminate, and to take advantage of customers heterogeneity, scarcity of capacity, and demand stochasticity. In our case, we analyze this phenomenon in a specific context characterized by the following modelling choices: *i.)* We consider a continuous influx of customers. *ii.)* Customers who decide not to purchase on arrival remain interested in the product for a random amount of time (possibly very large) and during which their valuations change. Eventually, they either purchase or loose interest. *iii.)* We limit ourselves to the context of *non-strategic* consumers; where, customers would purchase as soon as their valuation reaches a value larger than the current price. We do discuss how our framework and results remain valid for strategic consumers. *iv.)* In this infinite horizon setting, we assume that the firm commits to its pricing policy at the start of the horizon and by doing so aims to maximize its long-run average revenue. *v.)* We disregard any inventory and cost related issues and assume the product is available throughout.

Our contribution is threefold.

1.) We suggest a simple, tractable and innovative model that incorporates time-varying valuations. We show that such natural consumer's behavior has a major impact on the pricing policy that the firm should adopt. In particular, overlooking such behavior may lead to a sizeable opportunity cost. Our analysis is undertaken in the context of intertemporal pricing in the presence of patient and myopic customers. In the extension section, the results are shown to hold also in the context of strategic customers.

2.) Despite the more complex structure of the current model, and in line with related work, we show that cyclic policies remain either optimal or near optimal. Moreover, we suggest an algorithm to obtain an optimal (or near optimal) policy that is polynomial in the number of prices available to the firm to select from. We introduce a control parameter M for the firm, that depicts the maximum number of price changes each customer witnesses during her purchasing process. Our general results apply to any value of M but are fully leveraged when we restrict M to take relatively small values. In particular, when customers can only see at most two different prices (i.e. $M = 1$), we show that the pricing policy is either a fixed price policy or cyclic and simple i.e. no one price is repeated during a cycle.

3.) We cast our results in a general framework of optimizing (on infinite sequences of discrete prices) the long-run average revenue. In this framework, we introduce two classes of payoffs that we denote by *affine* and *weakly coupled*. In the former, the revenue generated by one price is linear in the number of periods this price is set for. As for the latter, the revenue per period depends on a finite number of neighboring prices. Within this framework, we characterize the solution to the optimization problem and present an algorithm to obtain the solution in polynomial time in the number of prices available.

1.1 Literature Review

The literature on intertemporal pricing is quite extensive starting with the early work on durable goods (e.g., Coase (1972), Stockey (1981), Conlisk et al. (1984), Besanko and Winston (1990), Sobel (1981)). We mention a couple of main ideas of this stream of literature that is relevant to our analysis. First, price commitment policies in the presence of strategic consumers are an effective way to exercise market power (see Coase (1972) and Stockey (1981)). Moreover, in the latter work price discrimination is shown to be optimal when some level of heterogeneity exists. Customers in these models are present in the system from the start. On the other hand, Conlisk et al. (1984) show that price discrimination when prices are set dynamically can also be optimal. For that they consider a continuous flux of strategic consumers some of whom value the product at a high value and the rest at a low value. Customers remain in the system indefinitely as long as they did not purchase. Such pricing tactics are shown to be also cyclic and decreasing. More recently, the operations management literature have looked at different variants of intertemporal pricing models (see Bitran and Caldentey (2003) and Aviv and Vulcano (2012)). Often in this literature, consumers are myopic and arrive through time endowed with a valuation drawn from a given distribution. The customer purchases on arrival if the price is higher than his valuation, otherwise leaves the system. Some of this literature has also incorporated customers' strategic behavior (see the review of Shu and Su (2007)). Levin et al. (2009) and Levin et al. (2010) consider respectively a monopolistic and a set of firms selling to a set of differentiated customers segments and assume that all customers are in the system from the start. In both papers, they prove the existence of a unique subgame perfect equilibrium that determines the optimal dynamic pricing policy. Aviv and Pazgal (2008) consider forward looking consumers with declining valuations who arrive following a Poisson process and where the listed price changes only once. More closely related to our work is the paper of Su (2007) that considers a deterministic model where strategic consumers are differentiated not only through their valuation, but also, through their patience level. Optimal prices are shown to be monotone. The very recent work of Caldentey et al. (2016) considers an intertemporal pricing problem under minimax regret where both strategic and myopic patient-customers are considered. They develop a robust mechanism design to compute an optimal policy. Finally, two recent notes (Wang (2016) and Hu et al. (2016)) tackle the problem of intertemporal pricing in the presence of reference price effects in an infinite horizon setting with discounted revenues. Probably, besides that of Conlisk et al. (1984), the papers that are the most related to our model are the recent works of Besbes and Lobel (2015) and Liu and Cooper (2015). Their models are variants of Conlisk et al. (1984), but, like us, consider a seller who commits to the pricing policy at the start of the horizon and maximizes the long-run average revenue. In the case of Besbes and Lobel (2015), heterogenous consumers, through their valuations and their patience level, are strategic in a sense that they will purchase if the lowest price during their presence in the system is lower than their valuation. They prove that a cyclic pricing policy is optimal but, does not need to be monotonic. On the other hand, Liu and Cooper (2015) define patient customers in a myopic sense, i.e., where customers purchase as soon as the price drops below their willingness-to-pay. In their case, cyclic policies remain optimal but are shown to be decreasing monotone.

None of the above literature considers valuations that change through time. As a matter of fact, and to the best of our knowledge, it is only recently that a handful of papers have tackled such behavior (see, Garrett (2016), Deb (2014) and Gallego and Şahin (2010)). Garrett (2016) considers, similarly to us, an infinite horizon setting where customers arrive through time with valuations that stochastically change while facing a committed pricing path. There are though a number of differences with our work. First, the model is set in continuous time where both buyers and sellers have a common discount rate and where customers are homogeneous with a valuation process following a Markov process that continuously switches between only two values, Low and High. The seller, is optimizing on the set of pricing paths (on the entire positive line) in the presence of strategic consumers. The optimal pricing path is shown to be cyclic decreasing (à la Conlisk et al. (1984)). We, on the other hand, focus primarily on myopic but heterogenous consumers and where prices and valuations can take a finite number of values. Another recent paper that considers a change in valuation is that of Deb (2014). The model considers one unit of a good and one buyer who is present in the system from the start and has a valuation drawn from some distribution. At some random time, a stochastic shock occurs and cause the buyer's valuation to be drawn again (from the same initial distribution). Finally, we mention the paper of Gallego and Şahin (2010) that models changing valuations in the context of revenue management. They primarily look at a two-periods model with brief generalization to a multiperiod setting. In their case, the product is *consumed* and thus its value is realized in the last period but, customers have the choice to purchase it earlier. All customers interested in the product are present from the start.

1.2 Plan of the Paper

In Section 2, we introduce the modelling ingredients and state the main results of this work with respect to the optimality of cyclic policies, their structure and how efficiently these can be obtained. In Section 2.3, we give a motivating example depicting a Markovian valuation process, and discuss how a setting with changing valuations disrupts the typical structure used in the literature to prove that cyclic policies are optimal and tractable. In Section 3, we introduce the notion of weakly coupled payoffs that allows us to tackle the general case and to which we suggest an algorithm that can obtain the optimal cyclic policy in polynomial time in the number of prices and exponential in the coupling depth. In Section 4, we introduce the notion of $\bar{\tau}$-affine type payoffs. We first show that such payoffs induce a tractable solution of the optimization problem where the optimal policy is shown to be cyclic and simple. We finally prove in Section 5 that the payoff of the problem at hand is indeed weakly coupled. Moreover, when the prices change at a slow pace (compared to how valuations evolve), this same payoff is shown to be $\bar{\tau}$-affine. As we do so, we characterize the optimal payoff function in closed form. We devote Section 6 for some extensive numerical analysis that takes full advantage of the algorithm suggested earlier and shows the opportunity cost of neglecting customers changing valuations. This section also suggests some simple policies that behave numerically surprisingly well. Finally, in Section 7 we discuss some important extensions including how our framework can handle the case of strategic consumers. Appendix A, is reserved for some of the proofs that were postponed for clarity of the exposition. In Appendix B, we analyze the case where prices can only take two values and discuss, specifically, the existence of a reset time.

2 Model Description and Main Results

2.1 Model Description

We consider a stream of consumers interested in buying a non perishable product from a seller who commits at the beginning of the horizon to an infinite sequence of prices $\pi = (p_1, p_2, ...)$. A number N of these consumers arrive every period. At any time t, the new batch of arrivals have their respective valuations in a set Ω^*, where $\Omega := \{v_j : 0 \leq j \leq K\}$, with $v_0 < v_1 < v_2 ... < v_K$ and $\Omega^* = \Omega \setminus \{v_0\}$. We let γ_j be the proportion of N that has a valuation at arrival equals to v_j, with $\sum_{j=1}^{K} \gamma_j = 1$. Equivalently, $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_K)$ is the initial distribution of the valuations of arriving customers. Customers can buy upon arrival if they find the price suitable. If they don't, they remain for some time *in the system*. A customer in the system is a customer who is interested in the product and can decide to purchase at any point during this time. Consumers interested in a product see their valuation changing through time following a transient Markov chain with a given transition matrix Q known by the seller and taking values in Ω. Each entry q_{ij} of Q, is the probability that a customer with current valuation v_i sees her valuation changing to v_j in the next period. The consumer leaves the system if he purchases the product at some point, or if he reaches the absorbing state v_0, whichever occurs first. Once out of the system, the customer would not consider purchasing the product anymore. The absorbing state v_0 models the possibility that the customer might loose interest in the current product, purchase it (or a substitute) elsewhere, or simply loose patience; such absorbing state replaces the time discounting factor present in other infinite horizon problems. It is worth noting that a setting, where valuations remain constant throughout the purchasing process, is a special case of ours with Q being a diagonal matrix.

When it comes to making a purchasing decision, customers could be *myopic* or *strategic*. In the former case, customers purchase the product as soon as their valuation reach a value higher than the current price (e.g., Liu and Cooper (2015)). Strategic customers on the other hand, are those that make their purchasing decision, not only based on current price and valuation, but also taking into account future committed prices (e.g., Besbes and Lobel (2015)). In this paper, we restrict the main analysis to the case where customers are myopic. This case does not necessarily reduce the complexity of the problem as much as it allows us to strictly highlight the implications of changing valuations on the problem's formulation and the corresponding solution. We show in Section 7 how the main results extend to the strategic case. However, we do not offer a specific analytical model to depict this behavior as we believe this is beyond the scope of this work.

This work assumes that the prices set by the firm belong to a finite set. In our myopic setting, and without loss of generality, we assume that this set is exactly Ω^*. We denote by p_t the price set by the firm at time t, e.g., $p_t = v_k$ means that at time t the value of the price is set at v_k. We interchangeably use k and v_k to denote the value of the price, with $v_k \in \Omega^*$ and $k \in \mathbb{N}_K^* \equiv \mathbb{N} \cap [1, K]$. We denote by $\mathcal{P}$ the set of all possible pricing policies. Observe that the primitives of the problem are summarized through the triplet (γ, Q, Ω). Now, given such triplet, and a policy $\pi = (p_1, p_2, ...) \in \mathcal{P}$, we denote by $L(p_1, .., p_t)$ the payoff function, which is the expected revenues generated by the first t prices of π, and by $\mathcal{R}(\pi) = \limsup_{t\to\infty} \frac{1}{t} L(p_1, ..., p_t)$, the long-run average revenue generated by a policy π. We are looking to solve for

$$\sup_{\pi \in \mathcal{P}} \mathcal{R}(\pi). \tag{1}$$

We say that π^* is optimal if it is a solution to Equation (1). We say that a pricing policy π' is ε-optimal, for some positive ε, if

$$\sup_{\pi \in \mathcal{P}} \mathcal{R}(\pi) \leq \mathcal{R}(\pi') + \varepsilon.$$

Note that a detailed definition and formulation of the payoff function that takes also into account the state of the system will be given later.

We should also mention that the long-run average revenue type objective used in this work does not seem to be much of a restrictive setting, as some of the main results in this paper (e.g., Theorems 3 and 4) remain valid under an infinite horizon discounted revenue type payoff.

We introduce at this point two different regimes that govern a customer's behavior and the corresponding valuation process. For that, we introduce the notation Q_m, which represents the principal submatrix of Q on indices $\{1, ..., m\}$. We also denote by $\mathcal{V}_t(k)$ the current valuation of a customer who entered the system t periods earlier with an initial valuation v_k.

We say that customers have a *maximum* patience level, $\bar{\tau}$, if for all $k \in \mathbb{N}_K$, $\mathcal{V}_t(k)$ is a Markov chain governed by Q as long as $1 \leq t \leq \bar{\tau}$, and $\mathcal{V}_t(k) = v_0$, for $t \geq \bar{\tau} + 1$. With a slight abuse of notations, we write this assumption as follows

(MPL) $$\exists\ 0 < \bar{\tau} < \infty, \text{ such that } Q_k^{\bar{\tau}+1} \text{ is set at } 0, \text{ for } 1 \leq k \leq K.$$

This assumption forces the customer to exit at most after $\bar{\tau}+1$ periods since arrival. It reflects the fact that, once customers do not purchase on arrival, they set a (deterministic and finite) budget of time during which they learn about the product and possibly make a purchase, but beyond which they leave the system. During this maximum amount of time, the dynamics of the valuation process are governed by Q. Assumption (MPL) does not constrain in any way the (Markovian) matrix itself; it sets an *exogenous* upper limit on the duration of a customer in the system. This assumption is in line with the recent literature (see Besbes and Lobel (2015) and Liu and Cooper (2015)). Note also that the customer might reach the state v_0 and leave the system without purchasing in less than $\bar{\tau}$ time periods.

We say that customers have an *ε-bounded* patience level if for all $k \in \mathbb{N}_K$, and for all $t \geq 0$, $\mathcal{V}_t(k)$ is a Markov chain governed by Q and the following holds,

$$(\varepsilon\text{-}\mathbf{BPL}) \qquad \text{given some } \varepsilon > 0, \;\; \exists\; 0 < \bar{\tau} < \infty, \;\; \text{such that,} \;\; ||Q_{K-1}^{\bar{\tau}+1}|| < \varepsilon.$$

This assumption is very different in nature from (MPL) in the sense that it preserves the endogeneity of the exit option which is purely governed by the transition matrix, making $\bar{\tau}$ intrinsic to Q. Moreover, as opposed to (MPL) the duration the customer spends in the system under (ε-BPL) is only nearly bounded (by $\bar{\tau}$) and with a small probability a customer might spend a longer time in the system. A sufficient condition for (ε-BPL) to hold is that $||Q_{K-1}|| := 1-\nu < 1$, for some $\nu \in (0,1)$ which we assume to be satisfied in this regime, and where $||\cdot||$, is the infinite norm for matrices. This condition is equivalent to assume that every valuation state communicates with either state v_0 or/and v_K guaranteeing that customers exit the system with probability one. In the (MPL) case, ν can well be equal to one.

We end this discussion by noting that, besides the concept of changing valuations, the specific characteristics of the (ε-BPL) regime, namely, the endogeneity and the possible unboundedness of the patience level, represent also some of the main differentiating factors of this work.

2.2 Main Results

Definition 1 *(Cyclic policies)*
A policy $\pi = (k_1, k_2, \ldots)$ is called cyclic if there exists $n \geq 1$ such that $k_{j+n} = k_j$ for all $j \in \mathbb{N}$. When n is the minimal integer with the latter property, we say that $(k_1, k_2, .., k_n)$ is the cycle of π, n its size, and denote π by its cycle $\pi = (k_1, \ldots, k_n)$.
For any $t \geq 1$, a cyclic policy is said to be t-simple if any string of t consecutive prices appears at most once in its cycle.

A fixed-price policy, one where the same price is set indefinitely, can also be viewed as cyclic with $n = 1$.

We start by introducing a way to express any pricing policy π that will be convenient for our exposition. Any pricing policy $\pi \in \mathcal{P}$ can effectively be viewed as a sequence of *phases* of the form

$\pi = ((k_j, \tau_j) : j \geq 1), k_{j+1} \neq k_j$, where each $(k_j, \tau_j) \in \mathbb{N}_K^* \times \mathbb{N}$ is called a phase (of the pricing policy) with duration τ_j during which the price is continuously set at k_j. Specifically, k_1 is set during τ_1 periods, followed by $k_2 \neq k_1$ set during τ_2 periods, so on and so forth. For all $j \geq 1$, set $T_j := \tau_1 + \tau_2 + ... + \tau_j$.

Definition 2 *For any $\tau \geq 1$, we denote by*

- $\mathcal{P}_\tau^-$ *the set of policies $\pi = ((k_j, \tau_j) : j \geq 1)$ such that $\tau_j \leq \tau$ for all $j \geq 1$.*
- $\mathcal{P}_\tau^+$ *the set of policies $\pi = ((k_j, \tau_j) : j \geq 1)$ such that $\tau_j \geq \tau$ for all $j \geq 1$.*

Definition 3 (Simple cyclic τ-policies)

i.) For any $\tau \geq 1$, we denote by $\mathcal{A}_\tau$ the set of policies where each phase has a duration that is a multiple of τ. We can always write such policy as a sequence of τ-phases of the form $\pi = ((k_j, \tau) : j \geq 1)$, but where we do not suppose here that $k_{j+1} \neq k_j$. Policies in $\mathcal{A}_\tau$ will be called τ-policies.

ii.) For $M \in \mathbb{N}^*$, a cyclic policy in $\mathcal{A}_\tau$ is said to be M-simple if any string of M prices $(k_1, \ldots, k_M)$ corresponding to M consecutive τ-phases of the policy appears at most once during a cycle. For $M = 1$, we just say the policy is simple.

Note that, given the general definition of t-simple policies, we made a slight abuse of notations in the definition of M-simple τ-policies. In case $\tau = 1$, both definitions match.

2.2.1 A general upper-bound on the duration of the phases of optimal policies

Theorem 1 *If (ε-BPL) holds, then*

$$\sup_{\pi \in \mathcal{P}} \mathcal{R}(\pi) \leq \sup_{\pi \in \mathcal{P}_{\bar{\tau}}^-} \mathcal{R}(\pi) + \varepsilon.$$

In the case where (MPL) holds, the inequality is replaced by an equality and $\varepsilon = 0$.

Theorem 1 states that, unless it is a fixed-price policy, a (near) optimal policy must have the durations of all its phases upper-bounded by $\bar{\tau}$; where $\bar{\tau}$ is the maximum duration that (most) customers spend in the system. This property reduces greatly the search for an optimal policy on the *unconstrained* set $\mathcal{P}$, albeit the search set remains substantially large.

We see next to which extent the complexity of the search of the optimal policy in $\mathcal{P}_{\bar{\tau}}^-$ can be reduced, in general and under additional assumptions.

2.2.2 Controlled pricing pace, $\tilde{\tau}$-policies

Before stating the rest of our result, we start by introducing the notion of pricing pace that allows firms to regulate how frequently a price can change in their policies. We denote by $\tilde{\tau} \geq 1$ the minimal duration that the firms allows for each price in its pricing policy. The existence of a pricing pace seems quite natural. It reflects the fact that, in practice, prices change at a slower pace than that at which a customer's valuation evolves in the system. It could be due to multiple factors including the fact that changing prices are costly, whether from the logistical end or as a reflection of customers' dissatisfaction from frequent changes in prices. By setting such a minimum pace firms can control on how many different prices one customer could face during his lifespan in the system.
When the pricing pace is set to $\tilde{\tau}$ the firm's problem is reduced to a constrained optimization problem on $\mathcal{P}^+_{\tilde{\tau}}$. The case $\tilde{\tau} = 1$ corresponds to the general unconstrained problem.

A natural quantity that plays a crucial role in the search of the optimal strategy is then the parameter

$$M := \lceil \bar{\tau}/\tilde{\tau} \rceil.$$

Indeed, by choosing the pricing pace parameter $\tilde{\tau}$, the firm insures that customers witness no more than M price-changes during $\bar{\tau}$ periods spent in the system.

We address first a particularly important setting in this paper: the one where $M = 1$, that is $\tilde{\tau} \geq \bar{\tau}$.

Theorem 2 **(Case** $M = 1$**)** *For any $\tilde{\tau} \geq \bar{\tau}$, under (MPL) (resp., ($\varepsilon$-BPL)), there exists a simple cyclic $\tilde{\tau}$-policy that optimizes (resp., ε-optimal) $\mathcal{R}$ on $\mathcal{P}^+_{\tilde{\tau}}$.*

Recall that a cyclic simple $\tilde{\tau}$-policy $\pi = ((k_j, \tau_j) : j \geq 1)$, is a policy where $\tau_j = \tilde{\tau}$ for all $j \in \mathbb{N}$ and where there exists $n \geq 1$, with $k_{j+n} = k_j$ for all $j \in \mathbb{N}$; moreover, $k_i \neq k_l$ for all $i, l \in [1, n], i \neq l$. As a consequence, the cycle size of the optimal policy is at most $K\tilde{\tau}$.

Theorem 2 also implies that if the firm has the flexibility on setting $\tilde{\tau} \in [\bar{\tau}, +\infty)$, it is better off with $\tilde{\tau} = \bar{\tau}$ (which also includes fixed-price policies). Finall, recall that by setting $\tilde{\tau} = \bar{\tau}$ the firm makes sure that every customer typically experiences at most two consecutive phases. Interestingly, we observe numerically (see, Section 6), that under $\tilde{\tau} \geq \bar{\tau}$, the (near) optimal policy is either a fixed price policy or cyclic with two prices.

The case where $\bar{\tau} = \tilde{\tau} = 1$, is an interesting special case. In this setting, the firm is solving for the unconstrained optimization on $\mathcal{P}$ and can change price every period. On the other hand, customers who don't purchase on arrival remain for one additional period in the system and leave with high likelihood after that. In this case, Theorem 2 asserts again that a (near) optimal solution is cyclic and simple where each price is set for no more than one period during a cycle, hence, with a cycle's size being at most K.

The next result generalizes Theorem 2 to the case where $M > 1$ or equivalently, $\tilde{\tau} < \bar{\tau}$. Up to enlarging $\bar{\tau}$ if necessary, we will assume that $\bar{\tau} = M\tilde{\tau}$. We also restrict ourselves to policies in $\mathcal{A}_{\tilde{\tau}}$, i.e., where prices are set in multiples of $\tilde{\tau}$.

Theorem 3 **(Case** $M > 1$**)** *Let* $M > 1$*, and* $\bar{\tau} = M\tilde{\tau}$*. Then under (MPL) (resp., (*ε*-BPL)), there exists an* M*-simple cyclic* $\tilde{\tau}$*-policy that optimizes (resp.,* ε*-optimal)* $\mathcal{R}$ *on* $\mathcal{A}_{\tilde{\tau}}$*.*

Observe here that an M-simple strategy in $\mathcal{A}_{\tilde{\tau}}$ has a length at most $K^M\tilde{\tau}$.
The previous result is also true when $\tilde{\tau} = 1$ (i.e., $M = \bar{\tau}$). However, this result is mainly meaningful (computationally) when M is small. In fact, even in the case where $\tilde{\tau} \geq \bar{\tau}$ (i.e., $M = 1$), there are $\sum_{i=1}^{K} i!$ possible solutions that are cyclic and simple, and when $M > 1$ the number is clearly much larger. The next result tackles the complexity of finding the optimal cyclic policies of Theorems 2 and 3 by means of an adequate algorithm that will be presented in Section 3.2.

Theorem 4 *The optimal cyclic policies of Theorems 2 and 3 can be obtained, through an algorithm that requires* $\mathcal{O}(K^{4M})$ *elementary computations. This translates for the (*ε*-BPL) case into* $\mathcal{O}\big(\varepsilon^{-4\,\frac{\ln K}{\nu\tilde{\tau}}}\big)$*.*

In particular, when $\tilde{\tau} \geq \bar{\tau}$, the policy can be computed in $\mathcal{O}(K^4)$. Observe that the exponential in M curse of dimensionality is present even in the case of constant valuations (see Liu and Cooper (2015)). As a matter of fact, in the latter paper and despite a major reduction in the cycle size (in the order of $K + M + 1$), the complexity of the solution remained, as we understand it, exponential in both K and M; from that regard, our results help reducing the computational complexity in K despite the lack in our case of the so-called regenerative structure (see the discussion in section 2.3.2). We should stress though that our algorithm is generic, in the sense that it does not take into account any specific form of the matrix Q. One could expect achieving better results when Q has some special form. Finally, observe that the optimization problem under (ε-BPL) regime can still be solved in full generality (i.e. $\tilde{\tau} = 1$) in $\mathcal{O}(\varepsilon^{-4\ln K/\nu})$.

We end this section with a result that complements Theorems 2 and 3 which claims that if the pricing pace is too slow (i.e., $\tilde{\tau}$ large) the firm is better off setting a fixed-price policy.

Proposition 1 *Suppose that* $\nu > 0$*. For any* $\zeta > 0$*, there exists a threshold* $\tilde{\tau}_0$ *such that, if* $\tilde{\tau} \geq \tilde{\tau}_0$*, then a fixed-price policy is* ζ*-optimal on* $\mathcal{P}_{\tilde{\tau}}^{+}$*.*

The fixed-price policy of Proposition 1 is a strict optimum if and only if it strictly outperforms the other fixed-price policies.

2.3 Additional Notes

2.3.1 Customers Valuations as a Markov chain. A Motivating Example.

One main feature of the model we introduce in this paper is customers' changing valuations through time. We assume that these valuations follow a transient Markov chain with transition matrix Q. We discuss here one possible motivation for this choice.

One motivating factor of why customers valuations change through time is the active evaluation process that customers go through once they show interest in a product. During this evaluation phase, customers undergo a continuous processing of new information. We propose a simple setting that could depict a stylized way of how such process works. For that, suppose that the product's "real" value for a specific customer is a random variable $V \in \{V_L, V_H\}$. The customer knows the two possible values but will only know which one applies to his case once he makes the purchase (i.e., start using/consuming the product). The customer can decide to purchase as soon as he gets interested in the product. But, if he doesn't, he starts gathering some additional information and learns better the product's fit. As he does so, his valuation gets updated.

To model the learning dynamics of a specific customer, we assume that information is received following a unit-rate Poisson process $(N(t) : -\infty < t < \infty)$. One can think of it as the number of positive comments the product is receiving on some social network platform. We assume that $t = 0$ is the time at which this customer got interested in the product (and started following this information (Poisson) process). The customer then processes new information as it shows up, updates his prior based on whether this positive information is relevant for him and disregard it otherwise. Hence, his valuation is being impacted by a *thinned* Poisson process. Consistently with the possible values of V, and for simplicity, we assume that this thinned process has a random rate β that can only take two possible values, $\beta \in \{b_L, b_H\}$. Basically, if the rate of positive news he is finding relevant is $b = \beta_H$ (resp., $b = \beta_L$) then the customer's "real" valuation is $V = V_H$ (resp., $V = V_L$).

The customer actively learns the value of β and accordingly implies the value of V. Consequently, the valuation will jump upwards on positive-relevant news and will otherwise continuously decrease until the next positive-relevant news. We denote by $q_t = \mathbb{P}(\beta = b_H | \mathcal{F}_t)$ where $\mathcal{F}_t$ represents the entire information available in $(0, t)$. As the information is gathered the customer keeps updating the value of q_t through a Bayesian update starting with some prior q_0. At any time t, the customer's best estimate of the product's valuation -assuming he is still interested in purchasing the item- is the following

$$\tilde{V}_t := q_t \, V_H + (1 - q_t) \, V_L.$$

Effectively, the valuation at time t is given by $V_t = \tilde{V}_t \cdot Y_t$, where, Y_t is a binary variable that tracks whether the customer lost interest in the product or not e.g. $Y_t = I(\tau > t)$. We assume that τ is an exponential random variable independent of $(N(t) : t \geq 0)$; it models the patience of the customer or the maximum duration the customer is ready to wait to buy this particular good. It is not hard to prove that the process q_t is a continuous time, continuous state, bounded Markov process (see for instance, Peskir and Shiryaev (2000)), and thus the same holds for V_t. This process can then be discretized to match our assumption in this paper.

2.3.2 Cyclic Policies and Related Literature

Natural candidates for "good" policies are cyclic policies (which include fixed price policies). The literature has recognized that such policies are optimal in some variants of our current setting (see, Conlisk et al. (1984) and more recently, Ahn et al. (2007), Besbes and Lobel (2015), and Liu and Cooper (2015)). In order to prove that optimal policies are cyclic and tractable, the aforementioned papers relied all on showing that the system must eventually *reset* and for that they followed one of two approaches: *i.)* a regenerative point argument (e.g, Ahn et al. (2007), Besbes and Lobel (2015) and Liu and Cooper (2015)), or *ii.)* customers' accumulation type argument (see Conlisk et al. (1984)). We quickly review these two approaches.

i.) The regenerative argument is based on the following simple but powerful observation. When customers spend a bounded amount of time in the system, say $\bar{\tau} < \infty$, during which the valuations remain constant, and when the set of prices the firm selects from is finite, then *any* given policy $(p_t : t \geq 0)$ contains an infinite subsequence $(p_{t_k^*} : k \geq 1)$ where for every k

$$p_{t_k^*} \leq \min\{p_{t_k^*+1}, ..., p_{t_k^*+\bar{\tau}}\}.$$

This result is proved easily by contradiction and applies whether customers are myopic or strategic. On those times $(t_k^* : k \geq 1)$, also so-called regenerative points, the system resets in the sense that no customer that comes before this time would want to purchase at a later time. The duration between two regenerative points is proven to be bounded reducing the infinite horizon optimization to an optimization on cycles between two reset times.

ii.) The infinite accumulation argument (see, Conlisk et al. (1984)), relies on the fact that customers remain in the system indefinitely and leave only if they purchase. The optimality of a cyclic policy is obtained by showing that it is optimal for the firm to eventually set the lowest possible price and reset the system. Indeed, if such reset price is not set for some time, customers with low valuation start accumulating in the system. Eventually, this accumulation will generate in one period (under the reset price), more revenues than any higher price can generate during this same period.

When valuations change through time, then both approaches do not apply. Given a general transition matrix and a pricing policy, customers with lower valuations can possibly have higher valuations in the future and hence the regenerative points described above have no reason to exist even for the optimal policy. Moreover, the expected number of customers in the system, (obviously under (MPL) but also under (ε-BPL)), remains uniformly bounded, (see, Proposition 7), no matter the pricing policy adopted, preventing any accumulation of customers in the system. In Appendix B, in the case of $K = 2$, we obtain necessary and sufficient conditions under which the seller is better off setting a reset price at the end of the cycle, showing therefore that cyclic policies are strictly optimal in both regimes. When $K > 2$, this is not true anymore, where optimal or near optimal cyclic policies do not necessarily contain any reset price.

In conclusion, in the presence of stochastic valuations, the typical intertemporal pricing model fails to have the structure that allowed previous work to reduce the analysis to *decoupled* cycles. However, in our case, we are still able to show the cyclic behavior of the optimal policy despite the fact that these cycles are not decoupled. Moreover, by controlling on the depth of the coupled cycles we can prove that such policy is tractable and can be efficiently obtained. Unfortunately, such *weakly coupled* cycles fail to inherit any "nice" property (such as the so-called *reflection principle* as in Besbes and Lobel (2015) or monotonicity as in Liu and Cooper (2015)).

Next, we cast our main results in a general framework of optimizing the long-run average revenues for specific class of payoffs. We introduce first the notion of weakly coupled payoffs, in which the revenue per period depends on a finite number of neighboring prices. Next, we discuss affine payoffs where the expected revenues generated by one price are linear in the number of periods this price is set for.

3 Weakly coupled payoffs

In this section, we introduce a general framework to analyze the optimization problem formulated in (1). We note that this problem is one of a general open-loop optimization, over a finite set of prices, of the long-run average of some payoff function.

We identify below a general class of payoff functions for which the optimization problem is relatively tractable. This is the class of weakly coupled payoffs, to which we later show that the payoffs generated in the context of intertemporal pricing with changing valuations belong.
However, we believe that the general framework in which these payoffs are treated and the results we obtain are relevant beyond the current setting and thus we treat them separately, in this section, in a self contained way. We also mention that these results (and the corresponding main results of Theorem 3 and 4) are also valid under an infinite horizon discounted revenue type payoff.

Recall that we have to find $\pi \in \mathcal{P}$ that maximizes $\mathcal{R}(\pi) = \limsup_{t\to\infty} \frac{1}{t} L(p_1, ..., p_t)$. We expressed earlier any policy $\pi \in \mathcal{P}$ as a sequence of phases, $\pi = \big((k_i, \tau_i) : i \geq 1\big)$. In this section, it will be useful to adopt another way of viewing a policy $\pi \in \mathcal{P}$. Indeed, for any given fixed $\tau > 0$, we divide any infinite sequence $\pi = (p_1, p_2, ...)$ into finite chunks of τ consecutive prices of the form $w_1 = (p_1, p_2, ..., p_\tau)$, $w_2 = (p_{\tau+1}, p_2, ..., p_{2\tau})$, ... Without loss of generality, any policy $\pi \in \mathcal{P}$ can be written as an infinite string of such terms, $\pi = \big((w_n : n \geq 1) : w_n \in \mathbb{N}_K^\tau\big)$.

Definition 4 (Weakly coupled payoffs) *We say that a payoff (that generates a revenue function $\mathcal{R}$) is $(\varepsilon, \bar{\tau})$-weakly coupled for some $\bar{\tau}$ and ε positive, if for any $\tau \geq \bar{\tau}$, there exists a function $f : \mathbb{N}_K^\tau \times \mathbb{N}_K^\tau \to \mathbb{R}^+$, such that for any $\pi = \big((w_n : n \geq 1) : w_n \in \mathbb{N}_K^\tau\big) \in \mathcal{P}$, the following holds*

$$\left| \mathcal{R}(\pi) - \limsup_{n\to\infty} \frac{1}{n\tau} \sum_{i=1}^{n} f(w_i, w_{i+1}) \right| \leq \varepsilon.$$

A payoff is said to be $\bar{\tau}$-weakly coupled if it is $(\varepsilon, \bar{\tau})$-weakly coupled with $\varepsilon = 0$ and in this case the inequality in the definition is replaced with an equality.[1]

3.1 Optimizing weakly-coupled payoffs

Given some function $f : \mathbb{N}_K^\tau \times \mathbb{N}_K^\tau \to \mathbb{R}^+$, we denote by φ the long run-average of f, i.e. for any $\pi = (w_1, w_2, \ldots)$, where the w_i's belong to $\mathbb{N}_K^\tau$, we let

$$\varphi(\pi) = \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f(w_i, w_{i+1}). \tag{$*$}$$

We extend the definition of φ to finite strings. For a finite string $W = (w_1, ..., w_n)$, we denote by

$$\varphi(W) = \frac{1}{n} \sum_{i=1}^{n} f(w_i, w_{i+1}),$$

with the notation $w_{n+1} = w_1$. Note that $\varphi(W) = \varphi(\pi)$, where $\pi = (W, W, ...)$.

We are interested in finding the policy that maximizes φ.

Proposition 2 *For any $\tau \geq 1$, for any function $f : \mathbb{N}_K^\tau \times \mathbb{N}_K^\tau \to \mathbb{R}^+$, there exists a τ-simple cyclic $\pi^* \in \mathcal{P}$ that maximizes φ.*

Recall that τ-simple means that no $w \in \mathbb{N}_K^\tau$ appears more than once in the cycle of the policy.

Proof. Denote by $\mathcal{W}_n$ the set of strings of length less than n, that is $W \in \mathcal{W}_n$ if $W = (\omega_1, \ldots, \omega_l)$ with $\omega_i \in \mathbb{N}_K^\tau$ for all $i \leq l$ and $l \leq n$. We have that $\varphi(W) = \varphi(\pi(W))$, where $\pi(W) = (W, W, ...)$. Let $\bar{W}_n = \operatorname{argmax}_{W\in\mathcal{W}_n} \varphi(W)$. From the definition of φ it follows that for any $\pi \in \mathcal{P}$

$$\varphi(\pi) \leq \varphi(\bar{W}_n) + \frac{\max f}{n}.$$

Hence it is sufficient to prove that for any fixed n, the maximizer $\bar{W}_n$ can be taken to be simple. Suppose ω is such that $\bar{W}_n = (W, W')$ with W and W' starting with some $\omega \in \mathbb{N}_K^\tau$, and let N and N' be the sizes of W and W'. Then since

$$\varphi(\bar{W}_n) = \frac{1}{N + N'}(N\varphi(W) + N'\varphi(W'))$$

we get that $\varphi(W) = \varphi(W') = \varphi(\bar{W}_n)$. Continuing this procedure we reach a simple maximizer of φ on $\mathcal{W}_n$. From that we also conclude that

$$\varphi(\pi) \leq \varphi(\bar{W}),$$

where $\bar{W}$ is the maximizer of φ on all τ-simple cyclic policies. □

Let $M \geq 1$, and $\bar{\tau} = M\tilde{\tau}$. Recall also that a cyclic policy in $\mathcal{A}_{\tilde{\tau}}$ is said to be M-simple if a string of M consecutive $\tilde{\tau}$ phases with a given string of prices $(k_1, \ldots, k_M)$ appears at most once during a cycle. Since $\bar{\tau} = M\tilde{\tau}$ we have that a cyclic $\bar{\tau}$-simple policy (Definition 1) that lies in $\mathcal{A}_{\tilde{\tau}}$ is M-simple in $\mathcal{A}_{\tilde{\tau}}$ (Definition 3)

Therefore, the following immediate corollary of Proposition 2 (with $\tau = \bar{\tau}$) shows that Theorem 3 will follow if we just prove the property of weakly coupled for the payoffs generated in the context of (MPL) and (ε-BPL).

Corollary 1 *If a payoff L is $\bar{\tau}$-weakly coupled, then the maximum of $\mathcal{R}$ on $\mathcal{A}_{\tilde{\tau}}$ is reached by an M-simple cyclic $\tilde{\tau}$-policy. A similar statement with ε-optimal policies holds for $(\varepsilon, \bar{\tau})$-weakly coupled payoffs.*

3.2 Optimization Algorithm

The objective of this section is to present an algorithm to find the optimal τ-simple cyclic policies of Proposition 2.

For a finite string $W = (w_1, ..., w_n)$, we denote by

$$\tilde{\varphi}(W) = \frac{1}{n-1} \sum_{i=1}^{n-1} f(w_i, w_{i+1}).$$

We define the following map from $\mathbb{N}_K^\tau \times \mathbb{N}_K^\tau$ to the subsets of $\mathbb{N}_K^\tau$, denoted by $\sigma(\mathbb{N}_K^\tau)$,

$$\psi(w, w') = S; S = \{\bar{w} \in \mathbb{N}_K^\tau : \varphi(w, w', \bar{w}) = \max_{\tilde{w} \in \mathbb{N}_K^\tau} \varphi(w, \tilde{w}, \bar{w})\}.$$

A collection $\mathcal{W}_n$ of finite simple strings of the same size n is called admissible if for all $W \in \mathcal{W}_n$, it is of the form $(w_1, .., w_n)$ with the same $w_1 := a$; moreover, if $(W, W') \in {\mathcal{W}_n}^2$ such that $w_n = w'_n$, then $W = W'$. Set $\tilde{K} = K^\tau$.

Algorithm.

STEP 1. INITIALIZATION. Fix $(w_1, w_2) \in \mathbb{N}_K^\tau \times \mathbb{N}_K^\tau$, with $w_1 \neq w_2$, let $\mathcal{W}_2(w_1, w_2) = \{(w_1, w_2)\}$, $\mathcal{S}_2(w_1, w_2) = \emptyset$ and set $n = 2$

STEP 2. COMPUTATION. Given an admissible simple collection $\mathcal{W}_n$ with $n < \tilde{K}$ and $\#\mathcal{W}_n \leq \tilde{K}$, define an admissible collection $\mathcal{W}_{n+1} = \Psi(\mathcal{W}_n)$ in the following way

1. for each $W^i \in \mathcal{W}_n := (w_1^i, \ldots, w_n^i)$, consider $S^i = \psi(w_{n-1}^i, w_n^i) := \{s_1^i, \ldots, s_{J_i}^i\}$ where, $J_i \leq \tilde{K}$ depends on the pair (w_{n-1}^i, w_n^i). Consider all the strings
$$\{W_j^i = (w_1^i, \ldots, w_n^i, s_j^i) : 1 \leq j \leq J_i,\ 1 \leq i \leq \#\mathcal{W}_n\}.$$
2. Select from these strings those such that $s_j^i = w_1$. This set of strings (possibly empty) is denoted by $\mathcal{S}_{n+1}$
3. Eliminate from the remaining all the strings W_j^i's that are not simple
4. From those remaining, consider any two strings, such that $s_j^i = s_{j'}^{i'}$ for some (i, j) and (i', j'). If $\tilde{\varphi}(W_j^i) < \tilde{\varphi}(W_{j'}^{i'})$ or $\tilde{\varphi}(W_j^i) > \tilde{\varphi}(W_{j'}^{i'})$, then eliminate the string that has the smaller value of $\tilde{\varphi}$. If $\tilde{\varphi}(W_j^i) = \tilde{\varphi}(W_{j'}^{i'})$, we eliminate (randomly) one of them. We do that for all such strings. The remaining strings must define an admissible collection of simple strings which we denote by $\mathcal{W}_{n+1}$ and note that $\#\mathcal{W}_{n+1} \leq \tilde{K}$.

STEP 3. ITERATION. If $n = \tilde{K}$, STOP. Otherwise, set $n = n + 1$ and GO TO Step 2.

Given that each computation step generates simple strings, and by noticing that the first part of Step 2 can be done offline, we need at most $\mathcal{O}(\tilde{K}^2)$ computations to obtain $\mathcal{W}_n(w_1, w_2)$ and $\mathcal{S}_n(w_1, w_2)$, $n \leq \tilde{K}$. This algorithm takes as an input, w_1 and w_2 and hence, needs to be repeated for all possible values of w_1 and w_2.

Proposition 3 *If the cycle of the optimal strategy of Proposition 2 is $W := (w_1, w_2, \ldots, w_J)$, for some integer $J \geq 2$, then $W \in \mathcal{S}_J(w_1, w_2)$.*

As a consequence of Proposition 3, if we consider

$$\mathcal{S} := \bigcup_{(w_1,w_2),w_1 \neq w_2} \bigcup_{n \leq \tilde{K}} \mathcal{S}_n(w_1, w_2)$$

and the set of single value strategies

$$\mathcal{S}_0 := \bigcup_{w \in \mathbb{N}_K^\tau} \{(w)\}$$

then the optimal strategy W of Proposition 2 satisfies $W \in \mathcal{S}$. Note that the computation of $\mathcal{S}$ needs at most $\mathcal{O}(\tilde{K}^4)$ calculations of the type $\varphi(w, w', w'')$. As a corollary we obtain the following.

Corollary 2 *If a payoff is $\bar{\tau}$-weakly coupled, then the optimal M-simple cyclic $\tilde{\tau}$-policy of Corollary 1 can be obtained with $\mathcal{O}(\tilde{K}^4)$ elementary calculations.*

For the proof of our main results on inter-temporal pricing we retain from this section the following.

Conclusion. The conclusion of this section is that, in view of Corollaries 1 and 2, it suffices to prove the $\bar{\tau}$-weakly coupled property for the payoffs generated in the context of the (MPL) and (ε-BPL), (with $\bar{\tau}$ being the threshold defined respectively, by the (MPL) and (ε-BPL) regimes), to guarantee the validity of Theorems 3 and 4.

4 Affine Payoffs

We saw in the previous section that to prove our main theorems in the case $M > 1$ it will be sufficient to establish the adequate weakly coupled property on the payoffs under (MPL) and (ε-BPL).

To prove Theorems 1 and 2 we need a more specific property of the payoff that we call $\bar{\tau}$-affine and that we now introduce. When we later prove that our payoffs are $\bar{\tau}$-affine, such property will also help us characterize the payoff function in a closed form, which in turns allows to perform our algorithm and obtain the optimal policy.

4.1 Payoff pair

First we have to introduce more carefully the notion of payoff that is at play in an intertemporal pricing problem as ours.

We set $\pi = \big((k_j, \tau_j) : j \geq 1\big)$. We denote by $\boldsymbol{\theta}^j$ the vector which K entries, θ^j_m, $1 \leq m \leq K$, that measures the *expected* number of customers in the system with valuation m at time $T_j = \tau_1 + \ldots, \tau_j$, i.e., at the end of the j^{th} phase. In this work, we always assume $\boldsymbol{\theta}^j$ to remain bounded (this is implied for example, in the (ε-BPL) case, by the assumption $\nu > 0$ in our model setup), that is, there exists $\rho > 0$ such that $\boldsymbol{\theta}^{j-1} \in [0, \rho]^K$ (see, Proposition 7). Since we are interested in the optimization of the long run average revenue, we may assume that $\boldsymbol{\theta}^0 = 0$. The definition of the payoff involves then an operator $\boldsymbol{\Theta} : \mathcal{M} \equiv \mathbb{N}^*_K \times \mathbb{N} \times [0, \rho]^K \to [0, \rho]^K$ that yields $\boldsymbol{\theta}^j$ recursively:

$$\boldsymbol{\theta}^j = \boldsymbol{\Theta}\big(k_j, \tau_j | \, \boldsymbol{\theta}^{j-1}\big).$$

It is worth stressing that $\boldsymbol{\theta}^j$ accounts for both those customers that were in the system at time T_{j-1} and did not exit by time T_j, as well as those that arrived during the phase (k_j, τ_j) and did not exit by time T_j.

To complete the payoff's definition we further consider an operator $L : \mathcal{M} \to \mathbb{R}_+$, whereby given $\boldsymbol{\theta}^{j-1} \in [0, \rho]^K$,

$$L\big(k_j, \tau_j | \, \boldsymbol{\theta}^{j-1}\big)$$

is the total expected revenues generated during phase (k_j, τ_j). These expected revenues are the result of purchases that occur from either customers present in the system at time T_{j-1}, or, from those that arrive to the system during (k_j, τ_j). We stress that L does not account for revenues generated from any of these customers if the purchase occurs after time T_j. We drop the index j when we consider a general phase e.g. (k, τ) with a given expected state of the system at the beginning of the phase e.g. $\boldsymbol{\theta} \in [0, \rho]^K$. If the phase starts when no customers are in the system then we drop $\boldsymbol{\theta}$ from the notations of both operators.

A payoff is defined by a pair $(L, \boldsymbol{\Theta})$ as above.

4.2 Optimizing affine payoffs

Definition 5 *($\bar{\tau}$-affine and $(\varepsilon, \bar{\tau})$-affine Payoffs) A payoff $(L, \boldsymbol{\Theta})$ is said to be $\bar{\tau}$-affine if for any $k \in [0, K]$, there exists $\bar{\boldsymbol{\Theta}}_k \in [0, \rho]^K$, $A_k, B_k \in \mathbb{R}^+$, $\bar{B}'_k \in \mathbb{R}^K$ such that for any $\tau \geq \bar{\tau}$, for any $\boldsymbol{\theta} \in [0, \rho]^K$ we have that*

(A1) $\boldsymbol{\Theta}\big(k, \tau | \, \boldsymbol{\theta}\big) = \bar{\boldsymbol{\Theta}}_k$;

(A2) $L(k, \tau | \, \boldsymbol{\theta}) = A_k + B_k(\tau - \bar{\tau}) + \langle \boldsymbol{\theta}, \bar{B}'_k \rangle$.

$(\varepsilon, \bar{\tau})$-affine payoffs are defined similarly with equalities in $(A1)$ and $(A2)$ replaced by inequalities up to an error of ε.

The previous definition is interesting in the sense that it introduces and highlights the general property of the payoff function required for Theorems 1 and 2 to hold exactly. Propositions 6 and 5 are the analogues of these two theorems for payoffs that are $\bar{\tau}$-affine. What would remain to do in order to complete the proof of these theorems is to check that the payoffs of interest in this paper are indeed $\bar{\tau}$-affine for the right value of $\bar{\tau}$.

Before we state the main results of this section, we start with a simple proposition that first confirms the natural role that cyclic policies play in our setting and secondly, allow us in many instances, to restrict our analysis to cyclic policies.

Proposition 4 *Given a payoff $(L, \boldsymbol{\Theta})$, where $\boldsymbol{\Theta}$ is uniformly bounded, then for any $\zeta > 0$, there exists a cyclic policy that is ζ-optimal, where the cycle size is in the order of $\mathcal{O}(1/\zeta)$.*

Proof. See Appendix. □

The previous result that holds in broad generality (as it relies only on the boundedness of $\boldsymbol{\Theta}$) is to be compared with the much more tractable optimizations (Theorems 2–4) obtained also in great generality but for some special and relevant situations.

Proposition 5 *If the payoff $(L, \boldsymbol{\Theta})$ is $(\varepsilon, \bar{\tau})$-affine, then*

$$\sup_{\pi \in \mathcal{P}} \mathcal{R}(\pi) \leq \sup_{\pi \in \mathcal{P}_{\bar{\tau}}^{-}} \mathcal{R}(\pi) + \varepsilon.$$

In case the payoff is $\bar{\tau}$-affine then the latter holds with $\varepsilon = 0$.

Proof. We deal with the case of $\bar{\tau}$-affine, the $(\varepsilon, \bar{\tau})$-affine case being similar. From Proposition 4, we know that the boundedness of $\boldsymbol{\Theta}$ implies that optimal strategies can be arbitrarily well approximated by cyclic ones.
Fix now T arbitrary large and consider first the set of cyclic strategies of length T. We finish if we show that the maximum over these strategies is reached on a strategy in $\mathcal{P}_{\bar{\tau}}^{-}$. Fix an integer n and consider the set of strategies $\pi_n = ((k_i, \tau_i)..,(k_n : \tau_n))$ with $T_n = \sum_{i=1}^{n} \tau_i \leq T$. Suppose now that for some $\ell \in [1, n]$, all the parameters k_i, $i \leq n$ and $\tau_i, i \in [1, n] - \{\ell\}$ are fixed. For simplicity call $P = P(\{k_i\}_{i \in [1,n]}, \{\tau_i\}_{i \in [1,n]-\{\ell\}}, T)$, the set of all the corresponding strategies, where only the parameter $\tau_\ell \in [1, T - \sum_{i \in [1,n]-\{\ell\}} \tau_i]$ is left free. By what preceded, we finish if we show that either the maximum of $\mathcal{R}(\pi_n), \pi_n \in P$ is reached for $\tau_\ell \leq \bar{\tau}$, or the fixed price policy with price k_ℓ performs better than any $\pi_n \in P$. Indeed, from the definition of affine payoffs we get for any $\pi_n \in P$ and any $\tau_\ell \in [\bar{\tau}, T - \sum_{i \in [1,n]-\{\ell\}} \tau_i]$ that

$$\mathcal{R}(\pi_n) = \frac{1}{\sum_{i=1}^{n} \tau_i} (U + V(\tau_\ell - \bar{\tau})),$$

where U and V do not depend on the value of $\tau_\ell > \bar{\tau}$. We thus get that the function $\mathcal{R}(\pi_n)$ is a monotonous function of $\tau_\ell \in [\bar{\tau}, T - \sum_{i \in [1,n]-\{\ell\}} \tau_i]$. Should this function be decreasing, the maximum over $\pi_n \in P$ of $\mathcal{R}(\pi_n)$ would be reached for $\tau_\ell \leq \bar{\tau}$, while if the function is increasing then the fixed price policy with price k_ℓ (whose average revenue is V) would perform better than any $\pi_n \in P$. $\square$

Proposition 6 *If the payoff $(L, \mathbf{\Theta})$ is $\bar{\tau}$-affine (res. $(\varepsilon, \bar{\tau})$-affine), then if $\tilde{\tau} \geq \bar{\tau}$, there exists a simple cyclic $\tilde{\tau}$-policy that optimizes (resp., ε-optimal) $\mathcal{R}$ on $\mathcal{P}^+_{\tilde{\tau}}$.*

Proof. We deal with the case of $\bar{\tau}$-affine, the ε-affine case being similar. The same argument as that used in the proof of Proposition 5 implies that the optimum of $\mathcal{R}$ on $\mathcal{P}^+_{\tilde{\tau}}$ is reached on a fixed price policy or on a $\tilde{\tau}$-policy $\pi = ((k_j, \tau_j) : j \geq 1)$ with $\tau_j = \tilde{\tau}$ for all $j \in \mathbb{N}$. Observe also that $\bar{\tau}$-affine Payoffs clearly have the $\tilde{\tau}$-weakly coupled property in restriction to the space $\mathcal{A}_{\tilde{\tau}}$ of $\tilde{\tau}$-policies. Hence, the case $M = 1$ of Corollary 1 implies that the optimal $\tilde{\tau}$-policy can be taken to be simple cyclic. $\square$

As in the proof of Proposition 6, Theorem 4 in the case $M = 1$ follows from the conclusion of Section 3 if we prove the $\bar{\tau}$-affine property of the payoff generated by (MPL) and (ε-BPL).

We move now to the proof of Proposition 1 which claims that if $\nu > 0$, then there exists a threshold $\tilde{\tau}_0$ such that, if $\tilde{\tau} \geq \tilde{\tau}_0$, a fixed-price policy is optimal on $\mathcal{P}^+_{\tilde{\tau}}$.

Proof of Proposition 1. It suffices from Proposition 6 to consider simple cyclic policies in $\mathcal{A}_{\tilde{\tau}}$. For such a policy $\pi_c = (k_1, \ldots, k_n)$ (where each k_i is set for $\tilde{\tau}$ periods), we have from the definition of affine payoffs that

$$\mathcal{R}(\pi_c) = \frac{1}{n\tilde{\tau}} \sum_{i=1}^{n} A_i + B_i(\tilde{\tau} - \bar{\tau}) + T_{i,i+1}$$

where A_i and B_i are functions of k_i and $T_{i,i+1}$ is a function of (k_i, k_{i+1}) (and not of $\tilde{\tau}$) and where $k_{n+1} := k_1$. Observe also that B_i is the long run average revenue of the fixed price policy with price k_i.

As $\tilde{\tau} \to \infty$ it is clear that the fixed price policy with price k such that $B_k = \max_{k'} B_{k'}$ becomes an increasingly better approximation of the optimal policy in $\mathcal{A}_{\tilde{\tau}}$. In the generic case where $\max_{k'} B_{k'}$ is a strict maximum for $k' = k$ then as soon as $\tilde{\tau}_0$ is such that for any simple cyclic vector $(k_1, \ldots, k_n)$ we have

$$B_k(n\tau_0 + \bar{\tau} - \sum_{i=1}^{n} \frac{B_i}{B_k}(\tau_0 - \bar{\tau})) > \sum_{i=1}^{n} A_i + T_{i,i+1}$$

then the fixed-price policy with price k is strictly optimal on $\mathcal{P}^+_{\tilde{\tau}}$. $\square$

Finally, with respect to the proof of our main results we retain from this section the following.

Conclusion. The conclusion of this section is that the proofs of Theorems 1, 2 and Theorem 4 (in the case $M \leq 1$) will be completed once we show that the payoff pair $(L, \mathbf{\Theta})$ corresponding to our initial setting is $\bar{\tau}$-affine (resp., $(\varepsilon, \bar{\tau})$-affine), where $\bar{\tau}$ is the value defined respectively through the (MPL) and (ε-BPL) regimes.

5 Analysis and Proofs of the Main Results

In the previous section, we proved that the conclusions of our main results stated in Section 2.2 apply for some general classes of payoffs functions. It is our goal in this section to complete the proofs of these results and show that the payoffs at hand fit the framework of $\bar{\tau}$-weakly coupled and $\bar{\tau}$-affine payoffs. Our proof, that relies on the Markovian dynamics of the valuation process, allows us also to obtain closed form approximations and properties of both the expected revenues generated from every period and the expected number of customers present in the system at any point in time.

5.1 Problem Re-Formulation and Additional Results

The intertemporal pricing problem of interest in this work admits as a primitive a triplet (γ, Q, Ω). In this context, we consider the corresponding payoff pair $(L, \mathbf{\Theta})$, as defined in section 4.1. Without loss of generality, and due to the linearity of L in $\boldsymbol{\theta}$, we assume that $N = 1$. The next result confirms that $\mathbf{\Theta}$ is in our context uniformly upper bounded.

Proposition 7 *For any policy $\pi = \big((k_j, \tau_j) : j \geq 1\big) \in \mathcal{P}$, we have that for all $j \geq 1$ and $1 \leq m \leq K$,*

$$\theta_m^j = \mathbf{\Theta}_m(k_j, \tau_j |\, \boldsymbol{\theta}^{j-1}) \leq \rho < \infty,$$

with $\boldsymbol{\theta}^0 = \mathbf{0}$ and where $\rho = \min\{1/\nu, \bar{\tau}\}$.

Given, a pricing pace, $\tilde{\tau}$, and in light of the payoff pair introduced above, we can formulate the firm's optimization problem as follows,

$$\mathcal{R}^* := \sup_{\pi \in \mathcal{P}_{\tilde{\tau}}^+} \limsup_{n \to \infty} \frac{1}{T_n} \sum_{j=1}^{n} L(k_j, \tau_j |\, \boldsymbol{\theta}^{j-1}), \tag{2}$$

with $\boldsymbol{\theta}^0 = \mathbf{0}$ and $T_n = \sum_{i=1}^n \tau_i$. We let $\mathcal{R}(\pi) = \limsup_{n\to\infty} \frac{1}{T_n} \sum_{j=1}^n L(k_j, \tau_j |\, \boldsymbol{\theta}^{j-1})$, and call a solution optimal if it solves (2). In Section 5.2.1, we obtain an exact formulation of respectively L and Θ as a function of the entries of the transition matrix Q. These formulations can be handy to numerically compute the expected revenues generated from simple pricing policies such as fixed price policies. From the boundedness of $\mathbf{\Theta}$ we recall Proposition 4 which states that sufficiently long cyclic policies are near optimal. More specifically, it implies that cyclic policies of cycle size $\mathcal{O}(\rho/\varepsilon)$ are ε-optimal with respect to the unconstrained optimization and thus can be obtained in $\mathcal{O}(K^{1/\varepsilon})$.

We end this section by stating the two results that are required to close the loop and finalize the proofs of Theorems 1 to 4. Note that in the (ε-BPL) case $\bar{\tau}$ is defined so that $||Q_{K-1}^{\bar{\tau}}|| > \varepsilon$, which would imply that it is in the order of $\mathcal{O}(|\ln \varepsilon / \ln(1-\nu)|)$. Without loss of generality, we assume that in the (ε-BPL) regime, $\bar{\tau}$ is given by,

$$\bar{\tau}_\varepsilon = |\ln \varepsilon / \ln(1-\nu)|\,.$$

Proposition 8 *For any triplet $(Q, \boldsymbol{\gamma}, \Omega)$, if ($\varepsilon$-BPL) holds then for any $\varepsilon > 0$, the payoff $(L, \boldsymbol{\Theta})$ is ε-$\bar{\tau}$-affine, for any $\bar{\tau} \geq \bar{\tau}_\varepsilon$. Moreover, the parameters $A_k, B_k, \bar{B}'_k$ and $\bar{\Theta}_m(k)$'s making up the definition of a $\bar{\tau}$-affine payoff are given in closed form in Section 5.2.3.*

From the previous proposition, we conclude that as long as $\tilde{\tau} \geq \bar{\tau}_\varepsilon$, then Theorem 1, 2 and Theorem 4 (in the case $M \leq 1$) hold.

Proposition 9 *For any triplet $(\boldsymbol{\gamma}, Q, \Omega)$, if ($\varepsilon$-BPL) holds and $\tilde{\tau}$ is set such that $\bar{\tau}_\varepsilon = M\tilde{\tau}$ (i.e., $M > 1$), then for any $\varepsilon > 0$, the payoff $(L, \boldsymbol{\Theta})$ is $(\varepsilon, \bar{\tau})$-weakly coupled, for any $\bar{\tau} \geq \bar{\tau}_\varepsilon$.*

The previous propositions stated the result for the (ε-BPL) regime. Equivalent results hold under the (MPL) case with $\bar{\tau}_\varepsilon$ is replaced by a $\bar{\tau}$, and $(\varepsilon, \bar{\tau})$-affine (resp., -weakly coupled) replaced with $\bar{\tau}$-affine (resp., -weakly coupled).

5.2 Proofs of Propositions 8 and 9.

We start with Proposition 8. The main idea behind the proof is the following. The value of $\bar{\tau}_\varepsilon$, implied by (ε-BPL), guarantees that customers, present in the system at the start of a phase (k, τ) with $\tau \geq \bar{\tau}_\varepsilon$, would unlikely remain in the system by the end of it. As a result, the expected revenues generated *following* a long phase (k, τ) depend on (k, τ) and on the subsequent phases, but are independent of all the phases *preceding* (k, τ). This leads, as we show analytically in the next section, to an affine formulation of the payoff as stated in Definition 5

5.2.1 Preliminary Analysis

This section is devoted to some preliminary analysis and calculations. In particular, we characterize how the expected number of customers that are in the system evolve through time and we also obtain a closed form formulation of the payoff $(L, \boldsymbol{\Theta})$. Recall that $v_0 < v_1 < \ldots < v_K$.

We start by introducing new notations. We set the following

- $P_{i,j^+} = \sum_{l=j}^K q_{il}$. The probability to transfer from state i to any possible state $l \geq j$.
- $P_{i^-,j^+} = \left(\sum_{l=j}^K q_{1l}, \sum_{l=j}^K q_{2l}, \ldots, \sum_{l=j}^K q_{(i-1)l}\right)^T$. The column vector whose entries are the probabilities to go from a state strictly lower than i to all possible states higher than j.
- $U_{i,j^-} = \left(q_{i1}, \ldots, q_{i(j-1)}\right)$. The row vector whose entries are the probabilities to go from state i to a specific state strictly lower than j
- $Q_{i^-j^-} = \{q_{kl} : 1 \leq k < i, 1 \leq l < j\}$ is the matrix of transfer probabilities from states lower than i to states lower than j. It is a minor of the matrix Q. We denote by Q_m the square matrix $Q_{m^-m^-}$

Note first that "$-$" involves states strictly lower while "$+$" involves states larger or equal; secondly, all the above quantities involve transition probabilities between states in Ω^* and do not involve the absorbing state v_0. Recall also that the valuation process follows a transient Markov chain with an absorbing state at v_0 and obviously at v_K; hence its transition matrix can be written as follows

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ H & Q_{K-1} & H' \\ 0 & 0 & 1 \end{pmatrix},$$

where H (resp., H') is a column vector with dimension $K-1$, which entries are the probabilities to go from any state $v_k > v_0$ (resp., $v_k < v_K$) to v_0 (resp., v_K). Assuming that all the values of $H + H'$ are positive, we conclude that the non negative square matrix Q_{K-1} is such that the sum of each one of its raw is strictly smaller than 1. We denote by Λ the maximum among all these sums and $\nu = 1 - \Lambda$. The same holds for all Q_m's minor of the matrix Q_{K-1}. Hence, $||Q_{K-1}||_\infty \equiv \max_i \sum_j q_{ij} = \Lambda < 1$, where $||\cdot||_\infty$ is the infinity norm for matrices induced by the infinite vector norm (where for $X \in \mathbb{R}^m$, $||X||_\infty = \max_i |X_i|$). For $m < K$, we denote $||Q_m||_\infty = \Lambda_m$. We drop, from now on, the subscript infinite from the notation of the infinite norm.

Expected Number of Customers in the System.
We consider $\boldsymbol{\theta} = (\theta_1, ..., \theta_K) \in [0, \rho]^K$, where θ_m is again the expected number of individuals with valuation v_m who are currently in the system. Given an expected state of the system, $\boldsymbol{\theta}$, we recall that $\boldsymbol{\Theta}(k, \tau|\,\theta) = \big(\Theta_1(k, \tau|\boldsymbol{\theta}), ..., \Theta_K(k, \tau|\,\boldsymbol{\theta})\big)$, is the expected state of the system one phase later. Therefore, if $m \geq k$ then $\Theta_m(k, \tau|\,\boldsymbol{\theta}) \equiv 0$, otherwise,

$$\Theta_m(k, \tau|\,\boldsymbol{\theta}) = \left(\sum_{l=1}^{k-1} \theta_l\, U_{l,k^-} Q_k^\tau + \sum_{i=0}^{\tau-1} \sum_{l=1}^{k-1} \gamma_l\, U_{l,k^-} Q_k^i\right)_m + \gamma_m \tag{3}$$

We recognize that for τ large, the first terms in the previous sum have little effect on the outcome. We also introduce the quantity $\bar{\boldsymbol{\Theta}}$. For $m < k$, set

$$\bar{\Theta}_m(k) := \left(\sum_{i=0}^{\infty} \sum_{l=1}^{k-1} \gamma_l\, U_{l,k^-} Q_k^i\right)_m + \gamma_m < \infty. \tag{4}$$

We let $\bar{\Theta}_m(k) \equiv 0$, when $m \geq k$.

Expected Revenues.
We denote by $\mathcal{L}_{(k,\tau)}(m)$ the expected revenue generated in the next τ periods by one individual currently present in the system with valuation v_m, facing in the next τ periods the price v_k. It is

possible that the individual does not purchase during these τ periods; any revenues generated after τ are not included in $\mathcal{L}_{(k,\tau)}(m)$. Therefore,

$$\begin{aligned}\mathcal{L}_{(k,\tau)}(m) &= \big[P_{m,k^+} + U_{m,k^-}P_{k^-,k^+} + U_{m,k^-}Q_k P_{k^-,k^+} + U_{m,k^-}Q_k\,Q_k\,P_{k^-,k^+} + \dots\big]\cdot v_k \\ &= \big[P_{m,k^+} + \sum_{l=0}^{\tau-2} U_{m,k^-}\,Q_k^l\,P_{k^-,k^+}\big]\cdot v_k\end{aligned} \tag{5}$$

Given that $||Q_k|| := \Lambda_k < 1$ for $1 \leq k \leq K$, the previous formulation shows that if the duration τ of the phase is large enough, only the "first" terms of the sum would probably matter. We also denote by $\bar{\mathcal{L}}_{(k,\tau)}(m)$ the expected revenues generated from one individual who just arrived into the system with valuation v_m, and who will be facing the phase (k,τ). Therefore,

$$\bar{\mathcal{L}}_{(k,\tau)}(m) = \delta_{m\geq k}v_k + (1-\delta_{m\geq k})\mathcal{L}_{(k,\tau)}(m).$$

The first term depicts the revenues generated if at arrival the customer has a valuation larger or equal than the listed price k. Otherwise, the customer now is in the system and the expected revenues generated from the remaining period are given by $\mathcal{L}$ defined above.

Putting together some of these formulations we obtain a closed formulation of the expected long term revenues L.

Proposition 10 *Given a transition matrix, Q, a vector γ of initial proportions and a set Ω of valuations, the expected revenues generated from a set of consecutive phases $\big((k_j,\tau_j) : 1 \leq j \leq n)\big)$, is given by*

$$\begin{aligned}L(k_1,...,k_n:\tau_1,...,\tau_n) &= \sum_{j=1}^{n} L(k_j,\tau_j|\,\boldsymbol{\theta}^{j-1}) \\ &= \sum_{j=1}^{n}\sum_{m=1}^{K}\big[\gamma_m \sum_{\tau=1}^{\tau_j}\bar{\mathcal{L}}_{(k_j,\tau)}(m) + \theta_m^{j-1}\,\mathcal{L}_{(k_j,\tau_j)}(m)\big]\end{aligned} \tag{6}$$

with $\theta_m^j = \Theta_m(k_j,\tau_j|\,\boldsymbol{\theta}^{j-1})$ is given in closed form in (3) and $\boldsymbol{\theta}^0 = \mathbf{0}$.

Next we detail the proof of Proposition 8 which we divide in two parts (A1) and (A2) to parallel Definition 5, in the (ε-BPL) case.

5.2.2 Proof of Proposition 8 - (A1).

We denote by $\mathbf{e}^{\mathrm{T}} = (1, ..., 1)^{\mathrm{T}}$, with $\|\mathbf{e}^{\mathrm{T}}\| = 1$. We set $\Lambda = 1 - \nu$. From the formulation of $\boldsymbol{\Theta}$ in Equation 3 and that of $\bar{\boldsymbol{\Theta}}$ in Equation 4, we write that

$$
\begin{aligned}
\left\|\bar{\boldsymbol{\Theta}}(k) - \boldsymbol{\Theta}(k, \tau | \boldsymbol{\theta})\right\|_1 &= \left| -\Big(\sum_{l=1}^{K} \Theta_l \, U_{l,k^-}\Big) Q_k^{\tau} \, \mathbf{e}^{\mathrm{T}} + \Big(\sum_{l=1}^{k-1} \gamma_l \, U_{l,k^-}\Big) \sum_{i=\tau}^{\infty} Q_k^{i} \, \mathbf{e}^{\mathrm{T}} \right| \\
&\leq \max\left\{ \Big(\sum_{l=1}^{K} \Theta_l \, U_{l,k^-}\Big) Q_k^{\tau} \, \mathbf{e}^{\mathrm{T}}, \Big(\sum_{l=1}^{k-1} \gamma_l \, U_{l,k^-}\Big) \sum_{i=\tau}^{\infty} Q_k^{i} \, \mathbf{e}^{\mathrm{T}} \right\} \\
&\leq \max\left\{ \left\| \Big(\sum_{l=1}^{K} \Theta_l \, U_{l,k^-}\Big) \right\|_1 \|Q_k^{\tau} \, \mathbf{e}^{\mathrm{T}}\|, \left\| \Big(\sum_{l=1}^{k-1} \gamma_l \, U_{l,k^-}\Big) \right\|_1 \left\| \sum_{i=\tau}^{\infty} Q_k^{i} \, \mathbf{e}^{\mathrm{T}} \right\| \right\} \\
&\leq \max\left\{ \rho \, \Lambda \, \Lambda^{\tau}, \Lambda \sum_{i=\tau}^{\infty} \Lambda^{i} \right\} \\
&= \max\left\{ \frac{K \, \Lambda^{\tau+1}}{1 - \Lambda}, \frac{\Lambda^{\tau+1}}{1 - \Lambda} \right\} = \frac{K \, \Lambda^{\tau+1}}{1 - \Lambda}.
\end{aligned}
$$

In the second inequality we used the fact that $|X \, Y^{\mathrm{T}}| \leq \|X\|_1 \, \|Y^{\mathrm{T}}\|$. The last inequality is obtained as we observe that $\|U_{l,k^-}\|_1 = \sum_{i=1}^{k-1} q_{li} = \Lambda_l \leq \Lambda$, $||Q_k^{\tau}|| \leq \Lambda^{\tau}$ and $\sum_{t=\tau}^{\infty} \Lambda^t = \Lambda^{\tau}/(1 - \Lambda)$. In order to guarantee that term $\Lambda^{\tau}/(1 - \Lambda)$ is less than ε it is enough to take $\tau \geq \frac{\ln b \varepsilon}{\ln \Lambda}$ with $b = (1 - \Lambda)$. Note that $\ln b < 0$, hence, it suffices that $\tau \geq \bar{\tau}_{\varepsilon} \equiv \frac{\ln \varepsilon}{\ln \Lambda}$ which completes our proof. □

5.2.3 Proof of Proposition 8 - (A2).

We introduce some additional notations. Recall that $\mathcal{L}_{(k,\tau)}(m)$ (resp. $\bar{\mathcal{L}}_{(k,\tau)}(m)$) is the expected revenues generated in the next τ periods from one individual who is already in the system (resp. just arrived to the system) facing the price v_k for τ consecutive periods. Given a certain duration $\bar{\tau}$, we denote by

- $B_k'(m) = \mathcal{L}_{(k,\bar{\tau})}(m)$, is the expected revenues generated during the phase $(k, \bar{\tau})$ by an individual initially in state m,
- $B_k(m) = \bar{\mathcal{L}}_{(k,\bar{\tau})}(m)$, is the expected revenues generated during the phase $(k, \bar{\tau})$ by an individual arriving with valuation v_m,
- $A_k(m) = \sum_{\tau=1}^{\bar{\tau}} \bar{\mathcal{L}}_{(k,\tau)}(m)$
- $A_k := \sum_m \gamma_m A_k(m)$, is the aggregate expected revenues generated during the phase, $(k, \bar{\tau})$, by all the customers that arrived during this same phase,
- $B_k := \sum_m \gamma_m B_k(m)$, is the aggregate expected revenues generated during the phase, $(k, \bar{\tau})$, by all the customers that newly arrived at the beginning of this phase,

- $T_{k,k'} := \sum_m \bar{\Theta}_m(k) B'_m(k')$, is the expected revenues generated by $\bar{\mathbf{\Theta}}(k)$ during a phase $(k',\bar{\tau})$, where, $\bar{\mathbf{\Theta}}$ is given by (4) and is an approximation of the expected number of customers remaining in the system at the end of phase $(k,\bar{\tau})$.

We start with a simple observation.

Sublemma. Assume $\tau \geq \bar{\tau}$. Then

$$0 \leq \bar{\mathcal{L}}_{k,\tau}(m) - \bar{\mathcal{L}}_{k,\bar{\tau}}(m) \leq p_{k,\bar{\tau},m}\, v_K$$

with $p_{k,\bar{\tau},m}$ being the probability that the individual with valuation m does not exit from the system after having faced a price, k during $\bar{\tau}$ periods, i.e.

$$p_{k,\bar{\tau},m} = \sum_i \left(U_{m,k}\text{-}Q_k^{\bar{\tau}}\right)_i$$

The proof of the Sublemma is obvious. We move to show the inequality related to (A2) of Definition 5 in the case again of (ε-BPL) and give a closed form of the constants involved in that expression. We start with the quantity $L(k_1 : \tau_1|\boldsymbol{\theta})$ that can be written as the sum of three terms. First, the expected revenues generated from new customers in the last $\bar{\tau}_\varepsilon$ periods of the phase and that is exactly A_{k_1}. Secondly, the expected revenues generated from new comers that arrived in the first $\tau - \bar{\tau}_\varepsilon$ periods of the phase and that is exactly, $\sum_m \sum_{\tau=\bar{\tau}_\varepsilon+1}^{\tau_1} \gamma_m\, \bar{\mathcal{L}}_{(k_1,\tau)}(m)$. Finally, the revenues generated from customers that were in the system at the beginning of the phase $\boldsymbol{\theta}$, and that is $\sum_m \theta_m\, \mathcal{L}_{k_1,\tau_1}(m)$. We write that

$$\begin{aligned} L(k_1,\tau_1|\boldsymbol{\theta}) &= A_{k_1} + \sum_m \sum_{\tau=\bar{\tau}_\varepsilon+1}^{\tau_1} \gamma_m\, \bar{\mathcal{L}}_{(k_1,\tau)}(m) + \sum_m \theta_m\, \mathcal{L}_{k_1,\tau_1}(m) \\ &\leq A_{k_1} + \sum_m \sum_{\tau=\bar{\tau}_\varepsilon+1}^{\tau_1} \gamma_m (B_{k_1}(m) + p_{k_1,\bar{\tau}_\varepsilon,m}\, v_K) + \sum_m \theta_m\, B'_{k_1,m} + \sum_m \theta_m\, p_{k_1,\bar{\tau}_\varepsilon,m}\, v_K \\ &\leq A_{k_1} + B_{k_1}\,(\tau_1 - \bar{\tau}_\varepsilon) + \sum_m \theta_m\, B'_{k_1,m} + (\tau_1 - \bar{\tau}_\varepsilon)\,\varepsilon\, v_K + \rho\, v_K \varepsilon. \end{aligned}$$

The first inequality is due to i.) and to the observation that if $\tau' \geq \bar{\tau}_\varepsilon$, then $\bar{\mathcal{L}}_{k_1,\tau'}(m) \leq B_{k_1,m} + p_{k,\bar{\tau}_\varepsilon,m}\, v_K$ (resp., $\mathcal{L}_{k_1,\tau'}(m) \leq B'_{k_1,m} + p_{k,\bar{\tau}_\varepsilon,m}\, v_K$). The second inequality is due to the following bounds, $p_{k_1,\bar{\tau}_\varepsilon,m} \leq \varepsilon$, $\sum_m \gamma_m p_{k_1,\bar{\tau}_\varepsilon,m} \leq \varepsilon$ and $\sum_m \theta_m p_{k_1,\bar{\tau}_\varepsilon,m} \leq \rho\,\varepsilon$. This proves our result.

Note that if we consider a cyclic policy $(k_1,...,k_n : \tau_1,...,\tau_n)$ with $\tau_i \geq \bar{\tau}_\varepsilon$ for all $i \leq n$, then, from $(A1)$ we have that the expected revenues generated during (k_i,τ_i) are upper bounded by $L(k_i,\tau_i|\bar{\mathbf{\Theta}}(k_{i-1})) + \varepsilon$, and hence, by summing these terms and dividing by T_n while noticing that $n \leq T_n$ we obtain that

$$\begin{aligned} \mathcal{R}(k_1,...,k_n : \tau_1,...,\tau_n) &\leq \frac{1}{T_n} \sum_{i=1}^n L(k_i,\tau_i|\bar{\mathbf{\Theta}}(k_{i-1})) + \varepsilon \\ &\leq \frac{1}{T_n} \sum_{i=1}^n \Big(A_{k_i} + B_{k_i}\,(\tau_i - \bar{\tau}_\varepsilon) + \langle \bar{\Theta}_m(k_{i-1}), B'_{k_i} \rangle \Big) + (1 + \frac{\rho}{\bar{\tau}_\varepsilon})\, v_K\, \varepsilon. \end{aligned}$$

Having in mind Proposition 6, when we restrict ourselves to cyclic policies with $\tau_i \equiv \tilde{\tau}$, the previous inequality shows that payoffs that are $(\varepsilon, \tilde{\tau})$-affine are also $(\varepsilon, \tilde{\tau})$-weakly coupled.

5.2.4 Proof of Proposition 9

We move now to discuss the proof of Proposition 9. We focus on the (MPL) case. By writing $\pi = \big((w_n : n \geq 1) : w_n \in \mathbb{N}_K^{\bar{\tau}}\big)$, we can define the payoff $(L, \mathbf{\Theta})$ recursively, applied this time, - not on a phase (k, τ)) but, - on a string of prices $w \in \mathbb{N}_K^{\bar{\tau}}$. Assuming here that $\bar{\tau} = M\,\tilde{\tau}$, with $M > 1$, we observe that w is in fact a sequence of M consecutive phases $(k_i, \tilde{\tau})$ (with k_i possibly equal to k_{i+1}). Similarly to Proposition 8 - (A1), it is easily seen that for any $w \in \mathbb{N}_K^{\bar{\tau}}$, $\mathbf{\Theta}(w|\boldsymbol{\theta}) = \tilde{\mathbf{\Theta}}(w)$, for some given $\tilde{\mathbf{\Theta}}$, where $\boldsymbol{\theta}$ is the expected number of customer present in the system at the beginning of the string w. Recalling Proposition 10, we conclude that the long run average revenue of any finite string $(w_1, ..., w_n)$, $\varphi(w_1, ...w_n) = \frac{1}{n\,\bar{\tau}} \sum_{i=1}^{n} f(w_i, w_{i+1})$, with $w_{n+1} = w_1$, which implies that

$$\mathcal{R}(\pi) = \limsup_{n\to\infty} \frac{1}{n\bar{\tau}} \sum_{i=1}^{n} f(w_i, w_{i+1}).$$

As for the (ε-BPL) regime, the proof follows similar steps than those of Propositon 8 and is omitted. Moreover, using Equation (4) and Proposition 10, one can again obtain the function f in closed form.

6 Numerical Analysis

Our objective in this section is to understand the behavior of the solution to our optimization problem. We use our algorithm to obtain and analyze the optimal (or nearly optimal) cyclic policy for many choices of the matrix Q, while also varying $\bar{\tau}$, the valuation vector V and its size K. Without much loss of generality we set $\gamma = [1/K, \dots, 1/K]$. The parameters that we use in drawing different realizations of Q and V are the following:

i.) The parameter $\nu > 0$ represents the probability in any given period that a customer leaves the system by reaching the absorbing state v_0. Without a great loss of generality we often confine, for simplicity of notations, our numerical analysis to the case where ν is the same for any given state of the customer. The coefficients of the matrix Q_{K-2} are thus drawn randomly in $(0, 1)$ such that the sum on each row is equal to $1 - \nu$.

ii.) The parameter $e > 0$ is introduced to represent the average gap between consecutive valuations. That is, the vector $V = [v_1, \dots, v_K]$ is such that $v_1 = 1$ and $v_{i+1} = v_i\,(1 + e\,\theta_i)$ where θ_i is uniformly drawn in $(0, 1)$.

When the findings with respect to the (MPL) and (ε-BPL) cases are the same, we present these under the (MPL) setting. We later discuss the finding that is specific to the (ε-BPL) case. Hence, we let for now $\tilde{\tau} = \bar{\tau} = \tau$. For each set of parameters $(\bar{\tau}, \nu, e, K)$ we make $N \gg 1$ drawings of the different parameters to which we apply the analytic algorithm. We collected and summarized the statistics of the outcomes in the table below.

We call k-cyclic policy a cyclic and simple τ-policy of size $k\tau$ (i.e., with k different prices).

We say that a cyclic policy π outperforms π' by $a\%$ if $100(\bar{\mathcal{R}}(\pi) - \bar{\mathcal{R}}(\pi'))/\bar{\mathcal{R}}(\pi') = a$.

Set $\mathfrak{f}$ the percentage of times where the maximal policy performs better than all the fixed price policies, and $\mathfrak{d}$ denotes the average ratio (in percent) of how much the maximal policy outperforms the fixed price policy (this average is computed conditional on the fact that the maximal policy is not a fixed price one).

Set $\mathfrak{f}'$ the percentage of times where the maximal policy performs better than all the fixed price and two-cyclic, and $\mathfrak{d}'$ denotes the average ratio (in percent) of how much the maximal policy outperforms the maximum over the fixed price and two-cyclic (this average is computed conditional to the fact that the maximal policy has period larger or equal to three).

Finding I. Cyclic policies with cycle at most three are always optimal.

In the case $\tilde{\tau} = \bar{\tau}$ and (MPL) each customer sees at most two prices and this is true with probability larger than $1 - \varepsilon$ in the (ε-BPL) case. It is hence reasonable to conjecture that the optimal cyclic policies given by Theorem 2 actually have a short cycle size, even if K is large. This observation is confirmed by our numerical analysis that actually yields, in all the cases we considered, optimal policies that are k-cyclic with k smaller or equal than 3. Moreover, in all the cases we treated, the fixed price and two-cyclic policies were nearly optimal, with just very few instances where a three-cyclic policy outperformed them; and in those cases the performance ratio was less than 1%. Hence, our first finding is that for all values of the parameters and for all the instances considered, the optimal policies are k-cyclic with $k \leq 3$.

Finding II. Fixed price and two-cyclic policies are practically optimal.

Our second finding is that fixed price policies and period two-cyclic policies are almost optimal with very few cases where the optimal policy is of period three and in those rare events, it outperforms the best among the fixed price and two-cyclic policies by at most 1%. This is reflected in the two last columns of Table 1.

Finding III. As τ increases fixed price policies perform increasingly better compared to two-cyclic ones.

In line with our result of Propositions 1 we find that, as long as the probability to reach the absorbing state v_0 or v_K from all the other valuations is supposed to be strictly positive, then as $\bar{\tau}$ goes to ∞ the fixed price policies become the best. This is because the customers tend to see just one price as $\bar{\tau}$ goes to infinity. As shown by the fourth column of Table 1, the frequency at which two-cyclic policies outperform the fixed price policies is decreasing with τ. As shown by the fifth column of Table 1 we also see the outperforming rate of the two-cyclic policies $\mathfrak{d}$ is also a decreasing function of τ. Finally, the last line of these two columns confirm our result of Proposition 1, namely that as τ goes to infinity the optimal policy is a fixed price one.

Finding IV. As ν increases fixed price policies perform increasingly better compared to two-cyclic ones. (We also find that as the probability to reach the absorbing states v_0 or v_K, from all the other valuations, increases, the fixed price policies become the best. This is natural since when the latter probability grows, the customers spend less time in the system and it is clear for example that if each customer stays in the system for just one period, then the price that optimizes the revenue of a single period (without consideration of Θ) gives rise to the optimal policy that is the fixed price policy with this same price. As shown by the fourth column of Table 1, the frequency at which two-cyclic policies outperform the fixed price policies is decreasing with ν. As shown by the fifth column of the Table 1, we also see the outperforming rate of the two-cyclic policies $\mathfrak{d}$ is also a decreasing function of ν.

τ	ν	e	$\mathfrak{f}$	$\mathfrak{d}$	$\mathfrak{f}'$	$\mathfrak{d}'$
1	0.1	0.1	76	2.6	5	0.1
5	0.1	0.1	43	0.6	2	0
9	0.1	0.1	21	0.3	0	0
13	0.1	0.1	13	0	0	0
5	0.	0.1	67	3	5	0.3
5	0.1	0.1	54	0.6	4	0.1
5	0.2	0.1	16	0.1	0	0
5	0.3	0.1	1	0	0	0
5	0.1	0.05	45	0.3	2	0.1
5	0.1	0.15	43	0.7	3	0.1
5	0.1	0.25	29	1.2	2	0.2
5	0.1	0.35	34	1.7	2	0.2

Table 1. (MPL), $\tilde{\tau} = \bar{\tau} = \tau$, $K = 4$.

Finding V. Impact of disregarding changing valuations on revenues.

We illustrate these findings with a couple of specific examples. Consider first the following. Set $K = 4$ and $\bar{\tau} = \tilde{\tau} = 10$ under the (MPL) case and set the transition matrix Q close to the Identity,

$$Q = \begin{bmatrix} 0.9 & 0.05 & 0.03 & 0.02 \\ 0.05 & 0.8 & 0.1 & 0.05 \\ 0 & 0.05 & 0.95 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Note that $\nu(Q) = 0$ as is the case for the Identity matrix. Let $V = [1.0, 1.3, 1.45, 1.6]$. For this valuations vector and this transition matrix Q, the optimal policy the two-cyclic policy $[3, 2]$. If however the matrix is the Identity (i.e., the valuations are not changing), the optimal policy is given by $[2,1]$. Of course, if the seller disregards the impact of the changing valuations, and thus solves for the optimal policy using $Q' = \text{Id}$, she will end up setting the two-cyclic policy $[2,1]$ instead of $[3,2]$. The loss in profit of doing so is found numerically to be 4%. This result is quite robust. Indeed, the policy $[2,1]$ is in fact numerically found to be optimal for all diagonal matrices that are close to the Id. Such losses are typically obtained when one replaces the transition matrix Q by the identity matrix.

Now, let $V = [1.0, 1.1, 1.25, 1.4]$ and

$$Q = \begin{bmatrix} 0.7 & 0.1 & 0.05 & 0.07 \\ 0.05 & 0.7 & 0.07 & 0.07 \\ 0.07 & 0.07 & 0.7 & 0 \\ 0 & 0 & 0.05 & 0.8 \end{bmatrix}.$$

For any transition matrix that is diagonal, $Q' = Diag(d_1, d_2, d_3, d_4)$ where $d_i \in [0.6, 1]$, we find that the fixed price policy $[1]$ is optimal. On the other hand, when we consider Q to be the transition matrix, the policy $[3,2,1]$ is optimal and outperforms $[1]$ by 5%.

Finding VI. An example of three-cyclic optimal policy.

In the case of $\bar{\tau} = M\tilde{\tau}$, $M \geq 2$, one expects that optimal policies with longer cycles appear. However each customer sees $M+1$ prices in this context and it is reasonable to conjecture that the length of the cycles will be small if M is taken to be small. Because of the complexity K^{4M} we are bound to let M be small if we want to apply our algorithm.

In all the examples we have tested ($M \leq 4$) the optimal policies had a cyclic length of at most 3. However, as opposed to the case $M = 1$, we often got cases where a three-cyclic policy outperforms both two-cyclic and the fixed price policies by more than 5%. We illustrate this finding with the following example.

Consider the following (MPL) example with $\tilde{\tau} = 1, \tau = 3$ (that is $M = 3$). We take $K = 4$, $\bar{\tau} = 6$ and $\tilde{\tau} = 2$. We set $V = [1.0, 1.2, 1.4, 2]$ and $\gamma = [0.25, 0.25, 0.25, 0.25]$. We let

$$Q = \begin{bmatrix} 0.8 & 0.05 & 0.0 & 0 \\ 0.2 & 0.6 & 0.0 & 0.2 \\ 0.1 & 0.2 & 0.6 & 0 \\ 0.1 & 0.2 & 0.6 & 0 \end{bmatrix}.$$

We compute numerically the maximum fixed price policy which is given by $[1]$ with $\mathcal{R}([1]) = 1$. The optimal two-cyclic policy is given by $[3,1]$ and while $\mathcal{R}([3,1]) = 1.09$. Finally, the optimal policy is a three-cyclic policy and is given by $[4,3,1]$ with $\mathcal{R}([4,3,1]) = 1.14$. It is an example where the optimal policy outperforms the fixed price policies by 14% and the two-cyclic ones by 4%.

Finding VII. Tuning ε in the (ε-BPL) case.

Consider the (ε-BPL) case. Suppose we want to obtain the best possible *unconstrained* cyclic policy using our algorithm with some fixed M. We consider in what follows $M = 1$ ($\tilde{\tau} = \bar{\tau}$) but the same reasoning holds for any M. For that we fix K, Q, V, and $\boldsymbol{\gamma}$.

For every $\varepsilon > 0$ we define $\tau_\varepsilon = -\ln \varepsilon/\nu$. Recall that the error in the estimation of the revenue by our analytical formulas ($M = 1$) is less than ε. By setting $\bar{\tau} = \tau_\varepsilon$, we denote by $\bar{\mathcal{R}}$ the long-run average revenue generated by the algorithm that identifies a maximal policy that we also denote by π_ε. It is natural to define a lower bound on the performance of such maximal policy by setting $\mathcal{L}_\varepsilon := \bar{\mathcal{R}}(\pi_\varepsilon) - \varepsilon$.

Of course, in the case of fixed price policies we have that $\mathcal{R} = \bar{\mathcal{R}}$. To compare the performance of π_ε with the fixed price policies we then introduce the notation $\Delta_\varepsilon = \bar{\mathcal{R}}(\pi_\varepsilon) - \mathcal{R}_f^* - \varepsilon$ where $\mathcal{R}_f^*$ denotes the revenue of the best fixed price policy. Typically, as $\varepsilon \to 0$, fixed price policies are optimal that the algorithm would confirm. As ε becomes large, we may often find two-cyclic optimal policies π_ε. However as ε becomes too large, Δ_ε is inevitably negative which would imply that any non fixed price policy we find is not guaranteed to perform better than the fixed price ones once we account for the error ε of our approximations. As we vary $\varepsilon > 0$, we aim to maximize $\mathcal{L}(\pi_\varepsilon)$. We illustrate this approach using the following example. We let $V = [1.0, 1.2, 1.5, 2]$ and

$$Q = \begin{bmatrix} 0.65 & 0.05 & 0 & 0.05 \\ 0.05 & 0.65 & 0 & 0.05 \\ 0.05 & 0.05 & 0.6 & 0 \\ 0 & 0 & 0 & 0.6 \end{bmatrix}.$$

Note that $\nu(Q) = 0.3$. From Table 1, we observe that there is a broad range of values of ε, where a two-cyclic policy is strictly better than fixed-price policies. In particular, by setting $\varepsilon = 0.01$ the algorithm would generate a policy that will outperform the best fixed price policy by more than 3%, and represent our best approximation of the unconstrained optimization.

ε	τ_ε	π_ε	$\bar{\mathcal{R}}(\pi_\varepsilon)$	Δ_ε
10^{-13}	83	[4]	1.02	0
10^{-5}	32	[4,1]	1.03	0.02
10^{-3}	19	[4,1]	1.04	0.04
10^{-2}	12	[4,1]	1.05	0.05
10^{-1}	6	[4,1]	0.99	0
0.2	4	[4,1]	0.91	-0.1
0.4	2	[4,1]	0.73	-0.27

Table 2. An (ε-BPL) case, with $M = 1$.

7 Extensions

7.1 Asymptotic Analysis

In the previous section, we restricted our analysis to policies that satisfy assumptions (MPL) and (ε-BPL), both of which guarantee that the system has finite memory in a sense that all or most customers do not spend more than a finite amount of time in the system. In this section, we scale this amount of time by a parameter $\nu \to 0$ that represents the probability to exit the system. We also scale $\bar{\tau}$ by ν, defining $\bar{\tau}_\nu = \bar{\tau}/\nu$ for some fixed $\bar{\tau}$. We also assume that $\tilde{\tau}_\nu = \bar{\tau}_\nu$.

Recall that Proposition 1 stated that for a fixed $\nu > 0$, if $\tilde{\tau}$ is sufficiently large then a fixed price policy is optimal in $\mathcal{P}^+_{\tilde{\tau}}$. Our observation in the current section will be that as $\nu \to 0$, although $\tilde{\tau}_\nu \to \infty$, price discriminating policies may well outperfom fixed price policies. For that, we consider a sequence of intertemporal pricing problems, similar to the ones introduced above, that we parameterize by $\nu > 0$. The set of possible valuations Ω^ν remains the same equal to Ω. On the other hand, the transition matrix, Q^ν, is assumed to converge to the identity matrix $\mathcal{I}$ as $\nu \to 0$. For that and for each value of ν, we let $Q^\nu_K = \mathcal{I} - \nu\, W_K$ where $W_{K-1} = (w_{i,j} : 1 \le i, j \le K-1)$ is some invertible matrix and the column vectors H^ν and H'^ν (made respectively, of the transition probabilities of any state into state v_0 and v_K) are set in a way to guarantee that Q^ν remains a stochastic matrix (see, the representation of Q in 5.2.1). We denote by $\Lambda^\nu_k = ||Q^\nu_k|| = 1 - \alpha_k\, \nu$, for some fixed positive α_k. By assuming, without loss of generality, that $\min_k \alpha_k = 1$, we recover the same definition of ν introduced earlier in this paper. Finally, we recall that $\bar{\tau}^\nu$ is set to $\bar{\tau}/\nu$, for some constant value $\bar{\tau} \ge 1$.

In this setting, recalling Proposition 7 and its proof, the likelihood that customers exit the system because their valuations reached an absorbing state is decreasing to zero with ν. Theorefore, the expected number of customers remaining in the system is growing also in $1/\nu$.

We let W_k (corresponding to Q_k) be the restriction of W to the first $k-1$ rows and columns. We also define W_k^{-1} to be its inverse and denote by $Z_k = I_k - \exp(-\bar{\tau}\, W_k)$, where $\exp(-\bar{\tau}\, W_k)$ is the exponential matrix operator applied to $-\bar{\tau}\, W_k$.

We now state a lemma that gives the limits, as ν goes to zero, of the different components of the long-run average revenue. We kept the same notations for $\mathcal{L}, \bar{\mathcal{L}}$ and U, introduced earlier but to which we added a subscript ν.

Lemma 1 *We define* $P^W_{i^-,j^+} = \left(\sum_{l=j}^K w_{1l}, \sum_{l=j}^K w_{2l}, \dots, \sum_{l=j}^K w_{(i-1)l}\right)^T$. *For any* $1 \le k \le K$, *as* $\nu \to 0$, *the following limits hold*

$$i.)\quad \bar{\mathcal{L}}^\nu_{(k,\infty)}(m) \to \mathcal{L}^0_{(k,\infty)}(m) := \left(\delta_{m\ge k} + U^0_{m,k}\, W_k^{-1}\, P^W_{k^-,k^+}\right) v_k,$$

$$ii.)\quad \mathcal{L}^\nu_{(k,\bar{\tau}^\nu)}(m) \to \mathcal{L}^0_k(m) := \left(\delta_{m\ge k} + U^0_{m,k}\, Z_k\, W_k^{-1}\, P^W_{k^-,k^+}\right) v_k,$$

$$iii.)\quad \bar{\mathcal{L}}^\nu_{(k,\bar{\tau}^\nu)}(m) \to \mathcal{L}^0_k(m),$$

$$iv.)\quad \frac{A^\nu_k(m)}{\bar{\tau}^\nu} := \frac{1}{\bar{\tau}^\nu}\sum_{\tau=1}^{\bar{\tau}^\nu} \bar{\mathcal{L}}^\nu_{k,\tau}(m) \to \tilde{\mathcal{L}}^0_k(m) := \left(\delta_{m\ge k} + U^0_{m,k}\left(I_k - \frac{1}{\bar{\tau}}\, Z_k\, W_k^{-1}\right) W_k^{-1}\, P^W_{k^-,k^+}\right) v_k$$

$$v.)\quad \nu\, \bar{\mathbf{\Theta}}^\nu(k, \bar{\tau}^\nu) \to \bar{\mathbf{\Theta}}^0(k) \equiv \sum_{l=1}^{k-1} \gamma_l\, U^0_{l,k^-}\, Z_k\, W_k^{-1}$$

$$\textit{where,}\quad \bar{\mathbf{\Theta}}^\nu(k,\tau) \equiv \sum_{\tau=0}^{\bar{\tau}}\sum_{l=1}^{k-1} \gamma_l\, U^\nu_{l,k^-}\, (Q^\nu)^\tau \le \mathbf{\Theta}^\nu(k,\tau).$$

We denote by $\hat{B}^0_k = \sum_m \gamma_m \, \mathcal{L}^0_{(k,\infty)}(m)$ the expected long-run average revenue generated by the fixed price policy, v_k. We also denote by $A^0_k = \sum_m \gamma_m \, \tilde{\mathcal{L}}^0_k(m)$ and by $\bar{T}^0_{k,k'} = \frac{1}{\bar{\tau}} \sum_m \bar{\Theta}^0_m(k) \, \mathcal{L}^0_{k'}(m)$. We set

$$\zeta^* \equiv \sup_{(n,k_1,\ldots,k_n \,:\, k_i \neq k_j)} \frac{1}{n} \sum_{i=1}^{n} \left(A^0(k_i) + \bar{T}^0_{k_i,k_{i+1}} \right).$$

Note also that both ζ^* and the $\hat{B}^0_k$'s depend only on the primitives $\boldsymbol{\gamma}$, W, and Ω of the problem.

Proposition 11 *For any triplet $(\boldsymbol{\gamma}, W, \Omega)$ such that $\max_k \hat{B}^0_k < \zeta^*$, there exists $\nu_0 > 0$ so that for any $\nu \leq \nu_0$, there exists a simple cyclic $\bar{\tau}^\nu$-policy that strictly outperforms all fixed-price policies on $\mathcal{P}^+_{\bar{\tau}^\nu}$.*

Proof. It is enough to show that for ν small enough,

$$\zeta^* \leq \sup_{\pi \in \mathcal{A}_{\bar{\tau}^\nu}} \mathcal{R}^\nu(\pi).$$

This inequality results from the following. First, cyclic and simple policies for which the durations of all its phases are exactly equal to $\bar{\tau}^\nu$, belong to $\mathcal{A}_{\bar{\tau}^\nu}$. Consider any such policy, $(k_1, \ldots, k_n : \bar{\tau}^\nu, \ldots, \bar{\tau}^\nu)$. Its long-run average revenue is equal to $\sum_{i=1}^n \left(\frac{A^\nu_k}{n \, \bar{\tau}^\nu} + \frac{\bar{T}^\nu_{k_i,k_{i+1}}}{n \, \bar{\tau}^\nu} \right)$, where $\bar{T}^\nu_{k_i,k_{i+1}}$ is the revenues generated by $\boldsymbol{\Theta}^\nu(k_i, \bar{\tau}^\nu)$ during phase $(k_{i+1}, \bar{\tau}^\nu)$. Moreover, from Lemma 1 *ii.*) and *v.*), we conclude that for ν small enough

$$\bar{T}^0_{k_i,k_{i+1}} := \frac{1}{\bar{\tau}} \sum_m \gamma_m \bar{\Theta}^0_m(k) \, \mathcal{L}^0_{k'}(m) \leq \frac{1}{\bar{\tau}} \sum_m \gamma_m (\nu \, \Theta^\nu_m(k, \bar{\tau}^\nu)) \, \mathcal{L}^\nu_{k',\bar{\tau}^\nu}(m) = \frac{\bar{T}^\nu_{k_i,k_{i+1}}}{\bar{\tau}^\nu},$$

while from Lemma 1 *iv.*) and the definition of A^ν_k, we have that

$$A^\nu_{k_i} / \bar{\tau}^\nu \to A^0_{k_i},$$

as $\nu \to 0$. By putting these observations together we conclude that

$$\frac{1}{n} \sum_{i=1}^{n} \left(A^0(k_i) + \bar{T}^0_{k_i,k_{i+1}} \right) \leq \sup_{\pi \in \mathcal{A}_{\bar{\tau}^\nu}} \mathcal{R}^\nu(\pi).$$

By taking the sup on the LHS we obtain our result. □

Hence, any primitives of the problem $(\boldsymbol{\gamma}, W, \Omega)$ for which $\max_k \hat{B}^0_k < \zeta^*$ guarantee that price discriminating policies are optimal at the limit. Interestingly, our previous algorithm can again be applied to solve for ζ^* and check whether the condition of Proposition 11 holds or not.

7.2 Strategic Consumers

Strategic customers, once in the system, make their purchasing decision not only based on the current valuation and price, but also based on future prices. We assume in this discussion that customers adopt an up to $\bar{\tau}$ periods look-ahead policy, where $\bar{\tau}$ is still given by the (MPL) or (ε-BPL) assumption. It is beyond the scope of this paper to formulate a specific model of strategic behavior as we did in the myopic case, however, we are able to argue that similar results to Theorems 2 to 4 still hold in this case. To set the ideas clear, we start by a couple of examples of possible customers' strategic behavior.

Example 1. Customers present in the system purchase in a specific period if the current surplus is larger than all possible surplus in the future calculated using the current valuation. Suppose a customer is currently present in the system with valuation k who arrived t periods earlier and could remain in the system for at most another $\bar{\tau} + 1 - t$ periods. Such customer would decide to purchase the product at the current price k_1, if $k - k_1 \geq 0$ and $k_1 \leq \min\{k_2, ..., k_{\bar{\tau}+1-t}\}$, with $k_1, ..., k_{\bar{\tau}+1-t}$ are the next consecutive $\bar{\tau} + 1 - t$ prices the customer would face if she remains in the system. Basically, the customer would delay a purchase if she sees that the product is offered at a cheaper price during her possible "stay" in the system. In this example, we assume that the customer is not aware of the matrix Q and believes that her current valuation is the best prediction of her future valuation.

Example 2. Customers purchase in a specific period if the current surplus is larger than the expected value of the maximum surplus that can be generated in the next $\bar{\tau}$ periods. Consider a customer with current valuation k, present in the system. She would purchase at the current price k_1, if

$$k - k_1 \geq \mathbb{E}_k \max_{1 \leq s \leq \bar{\tau}} \{\mathcal{V}_s(k) - k_{s+1}\},$$

where k_1 is the current price, and where $k_2, ..., k_{\bar{\tau}+1}$ (resp., $\mathcal{V}_2(k), ..., \mathcal{V}_{\bar{\tau}+1}(k)$) are the next $\bar{\tau}$ prices (resp., valuations) the customer would face (resp., have) if she decides to remain in the system. The customer and the firm have here the same information with respect to the dynamics of the valuation process, governed by Q.

In such models, the dynamics of the valuation process is assumed to remain the same as before, the only difference is the purchasing decision criteria leading to a different payoff function.

However, we still have that the customers present in the system at the start of a phase (k, τ) with $\tau \geq \bar{\tau}$, would unlikely remain in the system by the end of it. As a result, the expected revenues generated *following* a long phase (k, τ) depend on (k, τ), but are independent of the phase *preceding* (k, τ). This leads, as in the case of myopic customers studied in detail in Section 5, to an affine formulation of the payoff similar to the one stated in Definition 5.

We only consider the (MPL) case. In order to handle the (ε-BPL) case we would need to know more about how customers are making their purchasing decision (for instance, whether we are in the case of Example 1 or Example 2). Once such analytical model is specified an equivalent result to Proposition 8 can be obtained. But otherwise, the rest of the analysis is similar to the (MPL) case.

The result is stated for $M = 1$, but again it can be generalized to any $M \geq 1$ in exactly the same way it was done for the myopic case.

We say that a customer is strategic if the decision to purchase at time t takes into account not only price p_t but also $p_{t+1}, ..., p_{t+\bar{\tau}}$.

Proposition 12 *Consider a triplet* $(\boldsymbol{\gamma}, Q, \Omega)$ *and assume that (MPL) holds with a given* $\bar{\tau}$ *so that* $\tilde{\tau} \geq \bar{\tau}$. *Moreover, assume that all arriving customers are strategic. Then, the results of Theorem 2 apply in this context and the corresponding cyclic policy can be obtained in* $\mathcal{O}(K^4)$.

The steps of the proof are the same than in the myopic case. First, we need to show that the payoff in this setting remains affine (an equivalent result to Proposition 8). Then, we apply Proposition 6 that reduces the optimization set to policies that are cyclic simple with the durations of their phases equal to $\tilde{\tau}$. Finally, we need to observe that the latter policies generate a payoff that is weakly coupled. The only difference with the myopic case is with respect to the first step where a more general definition of affine payoffs is required. For that, consider two consecutive phases (k, τ) and (k', τ') with $\tau, \tau' \geq \bar{\tau}$. In Definition 5, we replace (A1), with $\Theta(k, \tau|\boldsymbol{\theta}) = \bar{\Theta}(k, k') \in [0, \rho]^K$. As for (A2), the affine relationship remains the same except that we replace A_k by a positive constant $A_{k,k'}$. Basically, we allow these two constants to depend not only on the price k of the current phase, but also on the one of the following phase k'. Besides that, the rest of the proof follows the exact same approach than in the myopic case.

Notes

[1]Under an infinite horizon discounted revenue setting, the definition of weakly coupled payoff becomes $|\mathcal{R}(\pi) - \sum_{n=0}^{\infty} e^{-r\,n} f(w_n, w_{n+1})| \leq \varepsilon$, where $r > 0$ is the discount factor. Interestingly, all the results in Section 3 related to weakly coupled payoffs remain valid in this setting and that is also true for Theorems 3 and 4.

Appendix A

Proof of Proposition 4:

Let $\mathcal{R}_T^*$ be the maximum long-run average revenue generated by any cyclic policy with size exactly T (the number of such cyclic policies is finite and hence $\mathcal{R}_T^*$ exists and is finite.) Consider any policy $\pi \in \mathcal{P}$ and fix a duration T_0 that is multiple of T (say $T_0 = m\,T$). We truncate again this subpolicy into m consecutive time periods of duration T. Recall that the expected revenues generated during a time period of length T is the expected revenues generated by customers arriving during this period (which we denote by L_j, $1 \leq j \leq m$) plus the expected revenues from those who were in the system at the beginning of this time period. Given the uniform boundedness of $\boldsymbol{\Theta}$ by ρ, the latter can be upper bounded by ρv_K. We can write the following:

$$\frac{L(p_1, ..., p_{T_0})}{T_0} = \frac{\sum_{j=1}^m L_j}{T_0} + \frac{m\,\rho\,v_K}{m\,T} \leq \mathcal{R}_T^* + \frac{\rho\,v_K}{T}.$$

This is enough to prove that the lim sup on the LHS as $T_0 \to \infty$ remains upper bounded by the RHS. Indeed, if T_0 is not a multiple of T then the LHS can be upper bounded by the same term on the RHS plus a term of the form L_{m+1}/T_0 which is itself upper bounded by $T\,v_K/T_0$ which will vanish at the limit. $\square$

Proof of Proposition 7:

The expected number of individuals in a particular state, $k < K$ at time t, following any pricing policy, is less than the number of individuals that would have been in that state at time t, if the pricing policy would have been set at v_K all throughout t. Recall, that $\Theta_K = 0$ and $\boldsymbol{\theta}^0 = \mathbf{0}$, and thus from Equation 3 we have that

$$\begin{aligned} ||\boldsymbol{\Theta}(K,t)|| &\leq ||\sum_{i=0}^{t-1}\sum_{l=1}^{K-1} \gamma_l \, U_{l,K^-} Q_K^i|| + ||(\gamma_1, \gamma_2, ..., \gamma_{K-1}, 0)|| \\ &\leq \sum_{i=0}^{t-1}\sum_{l=1}^{K-1} \gamma_l \, ||U_{l,K^-}|| \, ||Q_K||^i + 1 \\ &\leq \Lambda \sum_{i=0}^{t-1} \Lambda^i + 1 \leq \frac{1}{1-\Lambda}. \end{aligned}$$

The second inequality is due to the fact that $\sum_{l=1}^{K-1} \gamma_l \, ||U_{l,K^-}|| \leq \Lambda \sum_l \gamma_l \leq \Lambda$. The bound above shows that each component of $\boldsymbol{\Theta}$ is bounded by $1/(1-\Lambda)$. Hence, $||\boldsymbol{\Theta}(K,t)||_1 \leq \rho = K/(1-\Lambda)$. If Q was doubly stochastic then the upper bound, ρ, can be reduced to $1/(1-\Lambda)$. □

Proof of Proposition 11:

The proof follows directly from Lemma 1 of which we give here a proof.

We write $\mathcal{L}_{(k,\tau)}(m) = [P_{m,k^+} + U_{m,k^-}(I - Q^{\tau-1})\,(I-Q_k)^{-1}\,P_{k^-,k^+}]\,v_k$. Recall that if $||Q_k|| < 1$ then $\sum_{t=0}^{t-2} Q_k = (I_k - Q_k^{t-1})(I_k - Q_k)^{-1}$. Now by scaling the system we notice that $P^{\nu}_{k^-,k^+} = -\nu\, P^{W}_{k^-,k^+}$, moreover, $P^{\nu}_{m,k^+} \to \delta_{m \geq k}$ and $U^{\nu}_{l,k^-} \to U^{0}_{l,k^-} < \infty$ as $\nu \to 0$. We can now write the following

$$\begin{aligned} \mathcal{L}^{\nu}_{(k,\infty)}(m) &= [P^{\nu}_{m,k^+} + U^{\nu}_{m,k^-}\,(\nu\, W_k)^{-1}\, P^{\nu}_{k^-,k^+}]\, v_k \\ &\to \mathcal{L}^{0}_{(k,\infty)}(m) \equiv [\delta_{m\geq k} - U^{0}_{m,k^-} W_k^{-1}\; P^{W}_{k^-,k^+}]\, v_k \end{aligned}$$

as $\nu \to 0$. Similarly, as we note that $(I - \nu\, W)^{\bar{\tau}^{\nu}} = \exp(\bar{\tau}/\nu\, \ln(I - \nu\, W)) \sim \exp(-\bar{\tau}\, W)$ as $\nu \to 0$, we have that

$$\begin{aligned} \mathcal{L}^{\nu}_{(k,\bar{\tau}^{\nu})}(m) &= [P^{\nu}_{m,k^+} + U^{\nu}_{m,k^-}(I_k - Q_k^{\bar{\tau}^{\nu}-1})\,(I_k - (I_k - \nu\, W_k))^{-1}\,(-\nu)\, P^{W}_{k^-,k^+}]\, v_k \\ &\to \mathcal{L}^{0}_{k}(m) \equiv [\delta_{m\geq k} + U^{0}_{m,k^-}(I_k - \exp(-\bar{\tau}\, W_k))\, W_k^{-1}\; P^{W}_{k^-,k^+}]\, v_k \end{aligned}$$

as $\nu \to 0$. Similarly, $\bar{\mathcal{L}}^{\nu}_{(k,\bar{\tau}^{\nu})}(m) \to \mathcal{L}^{0}_{k}(m)$ as $\nu \to 0$. Similar calculations than above show that if $||Q_k|| < 1$ then $\sum_{\tau=1}^{\bar{\tau}}(I_k - Q_k^{\tau-1}) = \bar{\tau}\, I_k - (I_k - Q_k^{\bar{\tau}})\,(I_k - Q_k)^{-1}$, and *iv.*) follows easily. As for *v.*), note from equation 3 that

$$\boldsymbol{\Theta}(k,\tau|\boldsymbol{\theta}) \geq \bar{\boldsymbol{\Theta}}(k,\tau) \equiv \sum_{l=1}^{k-1} \gamma_l\, U_{l,k^-} \sum_{i=0}^{\tau-1} Q_k^i = \sum_{l=1}^{k-1} \gamma_l\, U_{l,k^-}(I_k - Q_k^{\tau})(I_k - Q_k)^{-1}.$$

Moreover,

$$\nu\, \bar{\boldsymbol{\Theta}}^{\nu}(k,\bar{\tau}^{\nu}) \to \bar{\boldsymbol{\Theta}}^{0}(k) \equiv \sum_{l=1}^{k-1} \gamma_l\, U^{0}_{l,k^-}(I_k - \exp(-\bar{\tau}\, W_k))\, W_k^{-1}.$$

□

Appendix B

A Simpler Model. The case of $K = 2$

In this section, we consider the case where $K = 2$ and $\Omega^* = \{v_1, v_2\}$, with $v_0 \le v_1 \le v_2$. A seller facing a myopic customer would only consider a pricing policy $\mathbf{p}$ where at any time t, $p_t \in \{v_1, v_2\}$. We start by observing that the price v_1 is a reset price and that is, once set the system empties. Having this in mind, it is easy to see that the optimal policy is either a fixed price policy where for all t, $p_t = v_i$ $i = 1, 2$, or is cyclical of the form $\pi_c = \big((v_2, \tau), (v_1, 1)\big)$; the cycle starts with v_2 that is set for $0 < \tau < \infty$ consecutive periods and then v_1 is set once. Note that the extreme cases of $\tau \in \{0, \infty\}$ cover the cases where the prices are constant.

The arriving customers has valuation v_2 with probability γ, and have valuation v_1 with probability $\bar{\gamma} = 1 - \gamma$. The average revenue per period collected by the seller is $R(\tau)$, with

$$\mathcal{R}(\tau) = \frac{1}{\tau + 1}\left(\gamma v_2 \tau + v_1 + v_1 \frac{\bar{\gamma} q_{12}}{1 - q_{11}}\left(\tau - \frac{1 - q_{11}^{\tau}}{1 - q_{11}}\right) + v_1 \bar{\gamma} \frac{1 - q_{11}^{\tau}}{1 - q_{11}}(q_{11} + q_{12})\right)$$

The first two terms in parenthesis represent the revenues from customers that arrive and buy right away: those with valuation v_2 at price v_2 during τ periods, and everyone at price v_1 in the last period. The last two terms in parenthesis account for the customers that accumulate over time and end up buying at either v_1 or v_2. The expected number of customers buying in a particular period l, $1 \le l \le \tau$, at price v_2 is given by the expression

$$\bar{\gamma} q_{12} \sum_{n=0}^{l-2} q_{11}^{n} = \bar{\gamma} q_{12} \frac{1 - q_{11}^{l-1}}{1 - q_{11}},$$

i.e., these are all customers that arrive earlier with low valuation v_2, and for whom the valuation remains at v_2 and jumps to v_1 from at period $l - 1$. Summing over all l, $1 \le l \le \tau$, we get:

$$\sum_{l=1}^{\tau}\left(\bar{\gamma} q_{12} \frac{1 - q_{11}^{l-1}}{1 - q_{11}}\right) = \frac{\bar{\gamma} q_{12}}{1 - q_{11}}\left(\tau - \sum_{l=1}^{\tau} q_{11}^{l-1}\right) = \frac{\bar{\gamma} q_{12}}{1 - q_{11}}\left(\tau - \sum_{l=0}^{\tau-1} q_{11}^{l}\right) = \frac{\bar{\gamma} q_{12}}{1 - q_{11}}\left(\tau - \frac{1 - q_{11}^{\tau}}{1 - q_{11}}\right)$$

The expected number of customers buying in period $\tau + 1$ at price v_1 is given by those who arrived earlier with valuation v_1 and remained at v_1 until period τ, and then either stayed at v_1 or jumped to v_2, i.e.,

$$\bar{\gamma} \sum_{l=0}^{\tau-1} q_{11}^{l}(q_{11} + q_{12}) = \bar{\gamma}(q_{11} + q_{12}) \frac{1 - q_{11}^{\tau}}{1 - q_{11}}.$$

We next state our main result here where we show that under some conditions involving the primitives of the problem, the seller is better off implementing a strictly cyclic policy of the form $\pi_c = \big((v_2, \tau), (v_1, 1)\big)$ with τ finite.

Proposition 13 *We denote by* $\tau^* = \arg\max_{\tau \in \mathbb{N}}\{\mathcal{R}(\tau)\}$. *We have the following*

$$0 < \tau^* < \infty \qquad \textit{iff} \qquad C \ln q_{11} < B - v_1 < C,$$

where

$$B = \gamma\, v_2 + v_2 \bar{\gamma} \frac{q_{12}}{1 - q_{11}} = v_2\, \frac{\gamma q_{20} + q_{12}}{1 - q_{11}} > 0,$$

and

$$C = -v_2 \bar{\gamma} \frac{q_{12}}{(1 - q_{11})^2} + v_2 \bar{\gamma} \frac{q_{11} + q_{12}}{1 - q_{11}}.$$

If these conditions do not hold, then $\tau^* = 0$ *iff* $B - v_1 < C \ln\, q_{11}$*, and* $\tau^* = \infty$ *otherwise.*

Proof. Given the expressions of B and C, we rewrite the revenue function as follows

$$(\tau + 1)\, \mathcal{R}(\tau) = v_1 + B\, \tau + C\, (1 - q_{11}^{\tau}).$$

Note that $C > 0$ iff $v_2 < v_1 (1 + \frac{q_{11}}{q_{12}})\, (1 - q_{11})$ where the term $(1 + \frac{q_{11}}{q_{12}})\, (1 - q_{11}) \geq 1$. In order to prove that it is optimal for the seller to set p_t at v_1 and reset the system, it is enough to prove that τ^* the maximizer of $\mathcal{R}$ is finite i.e. $0 \leq \tau^* < \infty$. Simple calculations show that

$$\mathcal{R}'(\tau) = \frac{B - v_1 - C + C\, q_{11}^{\tau}\, (1 - (\tau + 1)\, \ln(q_{11}))}{(\tau + 1)^2}.$$

We denote by $\mathcal{N}(\tau)$ the numerator of $\mathcal{R}'(\tau)$. By taking the derivative with respect to τ of $g(\tau) = q_{11}^{\tau}\, (1 - (\tau + 1)\, \ln(q_{11}))$, we observe that this quantity is non-increasing in τ, from which we conclude that $\mathcal{N}(\tau)$ is itself monotone in τ. Therefore, the equation $\mathcal{R}'(\tau) = 0$ admits at most one solution. In light of that, we denote by $\hat{\tau}^*$ the supremum of $\mathcal{R}$ on the positive real line and conclude that a necessary and sufficient condition for $0 < \hat{\tau}^* < \infty$ is that

$$i.)\ \mathcal{N}(0) > 0 \quad \text{and} \quad ii.)\ \mathcal{N}(\infty) < 0.$$

Hence, by putting these two inequalities together we get that $-C(1 - \ln q_{11}) < B - v_1 - C$ and $B - v_1 - C < 0$. We finally get that

$$0 < \hat{\tau}^* < \infty \qquad \text{iff} \qquad C \ln\, q_{11} < B - v_1 < C.$$

Note that under these condition $\mathcal{R}(\tau) > \mathcal{R}(\infty) \equiv B$ for any $\tau \geq \hat{\tau}^*$ and $\mathcal{R}(\tau) > \mathcal{R}(0)$ for any $\tau \leq \hat{\tau}^*$ and hence $\tau^* := \arg\max_n \mathcal{R}(n) \in (0, \infty)$, which completes our proof.

References

Ahn, H., M. Gümüş, P. Kaminsky. 2007. Pricing and manufacturing decisions when demand is a function of prices in multiple periods. *Oper. Res.* **55**(6) 1169–1188.

Aviv, Y., A. Pazgal. 2008. Optimal pricing of seasonal products in the presence of forwardlooking consumers. *Manufacturing Service Oper. Management* **10** 339–359.

Aviv, Y., G. Vulcano. 2012. A partially observed markov decision process for dynamic pricing. *Chapter 23, in Handbook of Pricing Management, (Eds. O. Ozer and R. Phillips). Oxford University Press* 522–584.

Besanko, D., W. Winston. 1990. ptimal pricing skimming by a monopolist facing rational consumers. *Management Sci.* **36**(5) 555–567.

Besbes, O., I. Lobel. 2015. Intertemporal price discrimination: Structure and computation of optimal policies. *Management Sci.* **61**(1) 92–110.

Bitran, G., R. Caldentey. 2003. An overview of pricing models for revenue management. *Manufacturing Service Oper. Management* **5** 203–229.

Caldentey, R., Y. Liu, I. Lobel. 2016. Intertemporal pricing under minimax regret. *forthcoming, Oper. Res.* .

Coase, R. 1972. Durability and monopoly. *Journal of Law & Economics* **15**(1) 43–49.

Conlisk, J., E. Gerstner, J. Sobel. 1984. Cyclic pricing by a durable goods monopolist. *Quarterly Journal of Economics* **99** 489–505.

Deb, R. 2014. Intertemporal price discrimination with stochastic values. *Working Paper, University of Toronto* .

Gallego, G., Ö. Şahin. 2010. Revenue management with partially refundable fares. *Oper. Res.* **58**(4) 817–833.

Garrett, D. 2016. Intertemporal price discrimination: dynamic arrivals and changing values. *forthcoming American Economic Review* .

Hu, Z., X. Chen, P. Hu. 2016. Technical notedynamic pricing with gain-seeking reference price effects. *Oper. Res.* **64**(1) 150–157.

Levin, Y., J. McGill, M. Nediak. 2009. Dynamic pricing in the presence of strategic consumers and oligopolistic competition. *Management Sci.* **55**(1) 32–46.

Levin, Y., J. McGill, M. Nediak. 2010. Optimal dynamic pricing of perishable items by a monopolist facing strategic consumers. *Prod. Oper. Manage.* **19**(1) 40–60.

Liu, Y, W.L. Cooper. 2015. Optimal dynamic pricing with patient customers. *Oper. Res.* **63**(6) 1307–1319.

Shu, Z. M., X. Su. 2007. Customer behavior modeling in revenue management and auctions: A review and new research opportunities. *Production and Operations Management* **16** 713–728.

Sobel, J. 1981. Durable goods monopoly with entry of new consumers. *Econometrica* **59**(5) 11455–1485.

Stockey, N. 1981. Rational expectations and durable goods pricing. *The Bell Journal of Economics* **1**(12) 112–128.

Su, X. 2007. Intertemporal pricing with strategic customer behavior. *Management Sci.* **53**(5) 726.

Wang, Z. 2016. Technical noteintertemporal price discrimination via reference price effects. *Oper. Res.* **64**(2) 290–296.