

# Dynamic Pricing and Learning: Historical Origins, Current Research, and New Directions

Arnoud V. den Boer

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

a.v.denboer@utwente.nl

March 17, 2015

## **Abstract**

The topic of dynamic pricing and learning has received a considerable amount of attention in recent years, from different scientific communities. We survey these literature streams: we provide a brief introduction to the historical origins of quantitative research on pricing and demand estimation, point to different subfields in the area of dynamic pricing, and provide an in-depth overview of the available literature on dynamic pricing and learning. Our focus is on the operations research and management science literature, but we also discuss relevant contributions from marketing, economics, econometrics, and computer science. We discuss relations with methodologically related research areas, and identify directions for future research.

*Keywords:* dynamic pricing; demand estimation; learning

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Historical Origins of Pricing and Demand Estimation</b>	<b>4</b>
2.1	Demand functions in pricing problems . . . . .	5
2.2	Demand estimation . . . . .	5
2.3	Practical applicability . . . . .	5
<b>3</b>	<b>Dynamic Pricing</b>	<b>6</b>
3.1	Dynamic pricing with dynamic demand . . . . .	7
3.1.1	Demand depends on price-derivatives . . . . .	7
3.1.2	Demand depends on price history . . . . .	7
3.1.3	Demand depends on amount of sales . . . . .	8
3.2	Dynamic pricing with inventory effects . . . . .	8
3.2.1	Selling a fixed, finite inventory during a finite time period . . . . .	8
3.2.2	Jointly determining selling prices and inventory procurement . . . . .	10
<b>4</b>	<b>Dynamic Pricing and Learning</b>	<b>10</b>
4.1	No inventory restrictions . . . . .	11
4.1.1	Early work . . . . .	11
4.1.2	Bayesian approaches . . . . .	12
4.1.3	Non-Bayesian approaches . . . . .	15
4.2	Finite inventory . . . . .	17
4.2.1	Early work . . . . .	17
4.2.2	Bayesian approaches . . . . .	17
4.2.3	Non-Bayesian approaches . . . . .	19
4.3	Machine-learning approaches . . . . .	21
4.4	Joint pricing and inventory problems . . . . .	22
<b>5</b>	<b>Methodologically Related Areas</b>	<b>23</b>
<b>6</b>	<b>Extensions and New Directions</b>	<b>24</b>
6.1	Strategic consumer behavior . . . . .	24
6.2	Competition . . . . .	25
6.3	Time-varying market parameters . . . . .	26
6.4	Model misspecification . . . . .	27
<b>7</b>	<b>Conclusion</b>	<b>28</b>
	<b>References</b>	<b>28</b>

## 1 Introduction

Dynamic pricing is the study of determining optimal selling prices of products or services, in a setting where prices can easily and frequently be adjusted. This applies to vendors selling their products via Internet, or to brick-and-mortar stores that make use of digital price tags. In both cases, digital technology has made it possible to continuously adjust prices to changing circumstances, without any costs or efforts. Dynamic pricing techniques are nowadays widely used in various businesses, and in some cases considered to be an indispensable part of pricing policies.

Digital sales environments generally provide firms with an abundance of sales data. This data may contain important insights on consumer behavior, in particular on how consumers respond to different selling prices. Exploiting the knowledge contained in the data and applying this to dynamic pricing policies may provide key competitive advantages, and knowledge how this should be done is of highly practical relevance and theoretical interest. This consideration is a main driver of research on *dynamic pricing and learning*: the study of optimal dynamic pricing in an uncertain environment where characteristics of consumer behavior can be learned from accumulating sales data.

The literature on dynamic pricing and learning has grown fast in recent years, with contributions from different scientific communities: operations research and management science (OR/MS), marketing, computer science, and economics / econometrics. This survey aims at bringing together the literature from these different fields, and at highlighting some of the older (and sometimes forgotten) literature where many important results and ideas already can be found.

A few literature reviews on dynamic pricing and learning do already exist. [Araman and Caldentey \(2011\)](#) and [Aviv and Vulcano \(2012, Section 4\)](#) review in detail a number of recent studies, mostly from the OR/MS community; [Christ \(2011, Section 3.2.1\)](#) contains an elaborate discussion of a selection of demand learning studies; and [Chen and Chen \(2014\)](#) review recent research on multiple-product pricing, pricing with competition, and pricing with limited demand information. Our survey complements these publications by aiming at a larger scope, and, although our main focus is on the OR/MS literature, we also address relevant contributions from computer science, marketing, economics and econometrics.

**Content.** This survey reviews the literature on dynamic pricing with demand uncertainty. We discuss how this is embedded in the literature on dynamic pricing in general, but do not review all relevant research topics associated with dynamic pricing; for this we refer to [Bitran and Caldentey \(2003\)](#), [Elmaghraby and Keskinocak \(2003\)](#), [Talluri and van Ryzin \(2004\)](#), [Phillips \(2005\)](#), [Heching and Leung \(2005\)](#), [Gönsch et al. \(2009\)](#), [Rao \(2009\)](#), [Chenavaz et al. \(2011\)](#), [Deksnyte and Lydeka \(2012\)](#) and [Özer and Phillips \(2012\)](#). We focus on studies where selling price is a control variable; we only scarcely discuss learning in capacity-based revenue management ([Talluri and van](#)

---

<sup>1</sup>[Zhang et al. \(2011\)](#)

Ryzin, 2004) or learning in newsvendor/inventory control problems. Neither do we consider mechanism design (Harris and Raviv, 1981, Maskin and Riley, 1984) or auction theory with incomplete information (see e.g. Segal (2003), Babaioff et al. (2011) and the references therein), although there are some similarities with dynamic pricing and learning. Most of the studies that we discuss are written from an (online) retailer perspective; we do not consider social welfare optimization (Balcan et al., 2008, Chakraborty et al., 2013). We do not dive into specific details associated with particular application area such as pricing in queueing or telecommunication environments (Mutlu et al., 2012), road pricing (de Palma et al., 2006, Tsekeris and Voß, 2009, Lou, 2013), or electricity pricing (Garcia et al., 2005, Roozbehani et al., 2010), to name a few. We also neglect a recent stream of *empirical* studies that aims to explain the dynamic pricing strategies of sellers by fitting models to sales data (see e.g. Escobari (2012) for an example on prices of airline tickets, Sweeting (2012) on prices of Major League Baseball tickets, and Huang et al. (2012) on pricing for a used-car dealership). Finally, this survey focuses on studies where the *seller* learns about the *demand function*, and not on studies where *buyers* (or sellers) learn (typically about the quality of the product) (Bergemann and Välimäki, 1997, Caminal and Vives, 1999, Schlee, 2001, Bose et al., 2006, Bhalla, 2012, Ifrach et al., 2012, Ching et al., 2013).

**Methodology.** We used Google Scholar to find all relevant references that were available online October 1, 2014. We excluded double versions of the same papers, or conference papers that largely overlap with journal papers. For all papers that we found this way we looked on Google Scholar for relevant other work citing these papers. We also checked the websites of key authors in the field for relevant publications. We aimed for a comprehensive review of the dynamic-pricing-and-learning literature; for the other sections, on e.g. demand estimation or dynamic pricing under full information, we restrict to key papers and reviews.

**Organization of the paper.** This survey is organized as follows. We review in Section 2 some of the pioneering historical work on pricing and demand estimation, and discuss how this work initially was difficult to apply in practice. In Section 3 we sketch important references and developments for dynamic pricing in general, and in Section 4 we focus on the literature that specifically deals with dynamic pricing and learning. Connections between dynamic pricing & learning and related research areas are addressed in Section 5, and related new research directions are discussed in Section 6. The core of the paper is Section 4, while the other sections are supporting.

## 2 Historical Origins of Pricing and Demand Estimation

Dynamic pricing with learning can be regarded as the combined application of two research fields: (1) statistical learning, specifically applied to the problem of estimating demand functions, and (2) price optimization. Both these fields are already quite old, dating back more than a century. In this section we briefly describe the historical fundamentals out of which dynamic pricing and learning has emerged, by pointing to some key references on static pricing and estimating demand functions that have been important in the progress of the field.

## 2.1 Demand functions in pricing problems

[Cournot \(1838\)](#) is generally acknowledged as the first to use a mathematical function to describe the price-demand relation of products, and subsequently solve the mathematical problem of determining the optimal selling price. As vividly described by [Fisher \(1898\)](#), such an application of mathematical methods to study an economical problem was quite new and controversial at the time, and the work was neglected for several decades. Cournot showed that if  $F(p)$  denotes the demand as a function of price  $p$ , where  $F(p)$  is continuous, decreasing in  $p$ , and  $pF(p)$  is unimodal, then the price that maximizes the revenue  $pF(p)$  can be found by equating the derivative of  $pF(p)$  to zero. If  $F(p)$  is concave there is a unique solution, which is the optimal price (this is contained in Chapter IV of [Cournot \(1838\)](#)). In this way, Cournot was the first to solve a “static pricing” problem by mathematical methods.

## 2.2 Demand estimation

To make theoretical knowledge on optimal pricing theory applicable in practical problems, one needs to have an estimate of the demand function. The first known empirical work on demand curves is the so-called King-Davenant Law ([Davenant, 1699](#)) which relates the supply and price of corn (see [Creedy, 1986](#), for an exposition on the origins of this work). More advanced research on estimating demand curves, by means of statistical techniques such as correlation and linear regression, took off in the beginning of the 20<sup>th</sup> century. [Benini \(1907\)](#), [Gini \(1910\)](#) and [Lehfeldt \(1914\)](#) estimate demand curves for various goods as coffee, tea, salt, and wheat, using various curve-fitting methods. Further progress on methodology was made, among others, by [Moore \(1914, 1917\)](#), [Wright \(1928\)](#) and [Tinbergen \(1930\)](#); the monumental work of [Schultz \(1938\)](#) gives a thorough overview of the state-of-the-art of demand estimation in his time, accompanied by many examples. Further references and more information on the historical progress of demand estimation can be found in [Stigler \(1954, section II\)](#), [Stigler \(1962, particularly section iii\)](#), [Brown and Deaton \(1972\)](#), [Christ \(1985\)](#), and [Farebrother \(2006\)](#). A small sample of the many contemporary studies on demand estimation in different contexts is [Berry et al. \(1995\)](#), [McFadden and Train \(2000\)](#) and [Bajari and Benkard \(2005\)](#).

## 2.3 Practical applicability

Estimating demand curves of various products was in first instance not aimed at profit optimization of commercial firms, but rather used to support macro-economic theories on price, supply, and demand. Application of the developed methods in practical problems was initially far away. An illustrative quote is from [Hicks \(1935\)](#), who doubted the possibilities of applying the theory of monopoly pricing on practical problems, exactly because of the difficulty of estimating the demand curve:

It is evidently the opinion of some of the writers under discussion that the modern theory of monopoly is not only capable of throwing considerable light on the general

principles underlying an individualistic economic structure, but that it is also capable of extensive use in the analysis of particular practical economic problems, that is to say, in applied economics. Personally, I cannot but feel sceptical about this. [...] There does not seem to be any reason why a monopolist should not make a mistake in estimating the slope of the demand curve confronting him, and should maintain a certain output, thinking it was the position which maximized his profit, although he could actually have increased his profit by expanding or contracting. (Hicks, 1935, page 18,19).

Hawkins (1957) reviews some of the attempts made by commercial firms to estimate the demand for their products. Most of these attempts were not successful, and suffered from difficulties of obtaining sufficiently many data for reliable estimates, and of changes in the quality of the product and the prices of competitors. Even a very detailed study of General Motors on automobile demand ends, somewhat ironically, with:

The most important conclusion from these analyses of the elasticity of demand for automobiles with respect to price is that no exact answer to the question has been obtained. (Horner et al., 1939, p. 137).

In view of these quotations, it is rather remarkable that dynamic pricing and learning has nowadays found its way in practice; many applications have been reported in various branches such as airline companies, the hospitality sector, car rental, retail stores, internet advertisement, and many more. A main cause for this is the fact that nowadays historical sales data is typically digitally available. This significantly reduces the efforts required to obtain sufficiently accurate estimates of the demand function. In addition, whenever products are sold via the Internet or using digital price tags, the costs associated with adjusting the prices in response to updated information or changed circumstances are practically zero. In contrast, a price-change in the pre-digital era would often induce costs, for example because a new catalog had to be printed or price tags had to be replaced. For a detailed study on such price-adjustment costs we refer to Zbaracki et al. (2004) and the references therein. Slade (1998) and Netessine (2006) are two studies that consider dynamic pricing with costly price adjustments.

### 3 Dynamic Pricing

In this section we discuss the literature on dynamic pricing. Because of the huge amount of literature on this subject, we cannot give a complete overview of the field. Instead, we briefly describe some of the major research streams and key references, in order to provide a context in which one can position the literature on dynamic pricing with learning discussed in Section 4. For a more elaborate overview of the vast literature on dynamic pricing, we refer to the books Talluri and van Ryzin (2004), Phillips (2005), Rao (2009), Özer and Phillips (2012), and the reviews by Bitran and Caldentey (2003), Elmaghraby and Keskinocak (2003), Heching and Leung (2005), Gönsch et al. (2009), Seetharaman (2009), Chenavaz et al. (2011) and Deksnyte and Lydeka (2012).

The literature on dynamic pricing by a monopolist firm can roughly be classified as follows:

- Models where the demand function is dynamically changing over time.
- Models where the demand function is static, but where pricing dynamics are caused by the inventory level.

In the first class of models, reviewed in Section 3.1, the demand function changes according to changing circumstances: for example, the demand function may depend on the time-derivative of price, on the current inventory level, on the amount of cumulative sales, on the firm's pricing history, et cetera. In the second class of models, reviewed in Section 3.2, it is not the demand function itself that causes the price dynamics: a product offered in two different time periods against the same selling price is expected to generate the same amount of average demand. Instead, the price dynamics are caused by inventory effects, more concretely by changes in the marginal value of inventory. Naturally, it is also possible to study models that fall both in classes, where both the demand function is dynamically changing and the price dynamics are influenced by inventory effects; some of this literature is also reviewed in Section 3.2.

## 3.1 Dynamic pricing with dynamic demand

### 3.1.1 Demand depends on price-derivatives

Evans (1924) is one of the first to depart from the static pricing setting introduced by Cournot (1838). In a study on optimal monopoly pricing, he assumes that the (deterministic) demand is not only a function of price, but also of the time-derivative of price. This models the fact that buyers do not only consider the current selling price in their decision to buy a product, but also anticipated price changes. The purpose of the firm is to calculate a price function, on a continuous time interval, that maximizes the profit. The problem is solved explicitly using techniques from calculus of variations. Various extensions to this model are made by Evans (1925), Roos (1925, 1927a,b, 1934), Tintner (1937), and Smithies (1939). Thompson et al. (1971) study an extended version of the model of Evans (1924), where optimal production level, investment level, and output price have to be determined. Closely connected to this work is Simaan and Takayama (1976), who consider a model where supply is the control variable, and where the time-derivative of price is a function of the current supply and current price. Methods from control theory are used to derive properties of the optimal supply path.

### 3.1.2 Demand depends on price history

A related set of literature considers the effect of reference prices on the demand function. Reference prices are perceptions of customers about the price that the firm has been charging in the past; see Mazumdar et al. (2005) for a review on the subject. A difference between the reference price and the actual selling price influences the demand, and, as a result, each posted selling price does not only affect the current demand but also the future demand. Dynamic pricing models and properties of optimal pricing strategies in such settings are studied, among others, by Greenleaf

(1995), [Kopalle et al. \(1996\)](#), [Fibich et al. \(2003\)](#), [Heidhues and Köszegi \(2012\)](#), [Popescu and Wu \(2007\)](#) and [Ahn et al. \(2007\)](#).

### 3.1.3 Demand depends on amount of sales

Another stream of literature on dynamic pricing emerged from diffusion and adoption models for new products. A key reference is [Bass \(1969\)](#), and reviews of diffusion models are given by [Mahajan et al. \(1990\)](#), [Baptista \(1999\)](#), and [Meade and Islam \(2006\)](#). In these models, the demand for products does not only depend on the selling price, but also on the amount of cumulative sales. This allows modeling several phenomena related to market saturation, advertisement, word-of-mouth effects, and product diffusion. [Robinson and Lakhani \(1975\)](#) study dynamic pricing in such a model, and numerically compare the performance of several pricing policies with each other. Their work stimulated much further research on optimal dynamic pricing policies, see e.g. [Clarke et al. \(1982\)](#), [Kalish \(1983\)](#), [Clarke and Dolan \(1984\)](#), and the references therein. The models studied in these papers are deterministic, and somewhat related to the literature following [Evans \(1924\)](#): both types of pricing problems are solved by principles from optimal control theory, and the optimal pricing strategy is often given by the solution of a differential equation.

[Chen and Jain \(1992\)](#), [Raman and Chatterjee \(1995\)](#), and [Kamrad et al. \(2005\)](#) extend these models by incorporating randomness in the demand. In [Chen and Jain \(1992\)](#), the demand is determined by a finite-state Markov chain for which each state corresponds to a deterministic demand function that depends on price and cumulative sales. The optimal price path is characterized in terms of a stochastic differential equation, and compared with the optimal policy in a fully deterministic setting. [Raman and Chatterjee \(1995\)](#) model uncertainty by adding a Wiener process to the (known) deterministic component of the demand function. They characterize the pricing policy that maximizes discounted cumulative profit, and compare it with the optimal price path in the fully deterministic case. Under some specific assumptions, closed form solutions are derived. Similar models that incorporate demand uncertainty are analyzed by [Kamrad et al. \(2005\)](#). For various settings they provide closed-form solutions of the optimal pricing policies.

## 3.2 Dynamic pricing with inventory effects

There are two important research streams on dynamic pricing models where the dynamics of the optimal pricing policy are caused by the inventory level: (i) “revenue management” type of problems, where a finite amount of perishable inventory is sold during a finite time period, and (ii) joint pricing and inventory-procurement problems.

### 3.2.1 Selling a fixed, finite inventory during a finite time period

In this stream of literature, a firm is assumed to have a certain number of products at its disposal, which are sold during a finite time period. There is no replenishment: inventory that is unsold at the end of the selling horizon is lost, and cannot be transferred to another selling season. In



these problems, the dynamic nature of optimal prices is not caused by changes in the demand function, but by the fact that the marginal value of remaining inventory is changing over time. As a result, the optimal selling price is not a fixed quantity, but depends on the remaining amount of inventory and the remaining duration of the selling season.

[Kincaid and Darling \(1963\)](#) may be the first to characterize and analyze the optimal pricing policy in such a setting. A more recent key reference is [Gallego and van Ryzin \(1994\)](#). They consider a continuous-time setting where demand is modeled as a Poisson process, with arrival rate that depends on the posted selling price. The pricing problem is formulated as a stochastic optimal control problem, and the optimal solution is characterized using the Hamilton-Jacobi-Bellman equation. For the exponential demand function a closed-form solution is derived. Because the optimal policy changes prices continuously, which may be undesirable in practical applications, two heuristics are proposed: one based on a deterministic version of the problem, and one based on determining the optimal fixed price. Both these heuristics are shown to be asymptotically optimal as the expected amount of sales grows large or as the length of the time horizon converges to zero. The authors further propose price heuristics - and show their asymptotic optimality - for many extensions of the problem: a discrete instead of a continuous set of feasible prices, customers who arrive according to a compound Poisson process, a demand function that depends both on price and on the time elapsed since the start of the selling season, the presence of holding costs and a discount rate, the case where initial inventory is a decision variable, and a setting where resupply, cancellations and overbookings are possible.

Numerous other extensions and variations of the model by [Gallego and van Ryzin \(1994\)](#) have been studied: for example, settings with restrictions on the number of allowable prices or price changes ([Feng and Gallego, 1995](#), [Bitran and Mondschein, 1997](#), [Feng and Xiao, 2000a,b](#)), time-varying demand functions ([Feng and Gallego, 2000](#), [Zhao and Zheng, 2000](#)), and extensions to multiple stores ([Bitran et al., 1998](#)) or multiple products ([Gallego and van Ryzin, 1997](#), [Kleywegt, 2001](#)) that share the same finite set of resources.

The extension to multiple products can be formulated as a dynamic program, which typically is intractable due to the curse of dimensionality. A number of papers study heuristic solutions ([Gallego and van Ryzin, 1997](#), [Maglaras and Meissner, 2006](#), [Zhang and Cooper, 2009](#), [Dong et al., 2009](#), [Erdelyi and Topaloglu, 2011](#), [Zhang and Lu, 2013](#), [Jasin, 2014](#)), which are typically based on static-price policies or on deterministic approximations or decompositions of the original dynamic program. Another stream of literature focuses on deriving structural properties of the price optimization problem, for various models of consumer demand ([Akçay et al., 2010](#), [Suh and Aydin, 2011](#), [Li and Huh, 2011](#), [Gallego and Wang, 2014](#)).

Another important extension to [Gallego and van Ryzin \(1994\)](#) is strategically behaving customers: customers who, when arriving at the (online) store, do not immediately decide whether to buy the product, but instead wait for a while to anticipate possible decreases in the selling price. In contrast, so-called myopic customers instantly decide whether to buy the product at the moment they arrive at the store. In such settings, the demand at a certain moment depends on the past, present, and future selling prices. Dynamic pricing in view of strategic customers has received a considerable amount of research attention in recent years; a representative sample is [Su \(2007\)](#),

Aviv and Pazgal (2008), Elmaghraby et al. (2008), Liu and van Ryzin (2008), Levin et al. (2009), Cachon and Swinney (2009) and Su (2010). Reviews of these literature are given by Shen and Su (2007) and Gönsch et al. (2013). These studies typically have a game-theoretic flavor, since both the firm and the strategic customers have a decision problem to solve, with contradicting interests.

### 3.2.2 Jointly determining selling prices and inventory procurement

A main assumption of the literature discussed above is that the initial capacity level is fixed. In many situations in practice this is a natural condition: the number of seats in an aircraft, rooms in a hotel, tables in a restaurant, or seats in a concert hall are all fixed for a considerable time period, and modifications in the capacity occur at a completely different time scale than dynamic price changes. In many other settings, however, the initial capacity is a decision variable to the firm; in particular, when the firm can decide how many items of inventory should be produced or purchased. Pricing and inventory management can then be considered as a simultaneous optimization problem.

This research field bridges the gap between the pricing and inventory management literature. Many different settings and models are subject to study, with different types of production, holding and ordering costs, different replenishment policies (periodic or continuous), finite or infinite production capacity, different models for the demand function, et cetera. Extensive reviews of the literature on simultaneous optimization of price and inventory decisions can be found in Eliashberg and Steinberg (1993), Federgruen and Heching (1999), Elmaghraby and Keskinocak (2003, Section 4.1), Yano and Gilbert (2004), Chan et al. (2004), Chen and Simchi-Levi (2012), and Simchi-Levi et al. (2014).

## 4 Dynamic Pricing and Learning

In the static monopoly pricing problem considered by Cournot (1838), the demand function is deterministic and completely known to the firm. These assumptions are somewhat unrealistic in practice, and eventually it was realized that demand should be modeled as a random variable. One of the first to pursue this direction is Mills (1959), who assumes that the demand is the sum of a random term with zero mean and a deterministic function of price. He studies how a monopolist firm that sells finitely many products in a single time period should optimally set its production level and selling price. Further extensions of this model and properties of pricing problems with random demand are studied by Karlin and Carr (1962), Nevins (1966), Zabel (1970), Baron (1970, 1971), Sandmo (1971) and Leland (1972). An important research question in these studies is how the firm's optimal decisions are influenced by the demand uncertainty and by the firm's attitude towards risk (risk-neutral, risk-averse, or risk-preferred).

The papers mentioned above model demand as a random variable, but still assume that the expected demand (as a function of the selling price) is completely known by the firm. This makes these models somewhat unrealistic and not usable in practice. The common goal of the literature on dynamic pricing and learning is to develop pricing policies that take the intrinsic uncertainty

about the relation between price and expected demand into account.

In the next two sections we discuss the literature on dynamic pricing and learning. Section 4.1 considers the literature on the problem of a price-setting firm with infinite inventory and unknown demand function. This basically is the monopoly pricing problem described in Section 2.1, with uncertainty about the demand function. The full-information case of this problem is static; the price dynamics are completely caused by the fact that the firm learns about the price-demand relation through accumulating sales data. Section 4.2 discusses literature on pricing policies for firms selling a fixed, finite amount of inventory, with unknown demand function. For this problem, the full-information case is already dynamic by itself, as discussed in Section 3.2, and the learning aspect of the problem provides an additional source of the price dynamics.

## 4.1 No inventory restrictions

### 4.1.1 Early work

**Uppsala Econometrics Seminar.** The first analytical work on dynamic monopoly pricing with unknown demand curve seems to have been presented on August 2, 1954, at the 16th European meeting of the Econometric Society (Hansen, 1955), by F. Billström, H. Laadi, and S.A.O. Thore, with contributions from L.O. Friberg, O. Johansson, and H.O.A. Wold. A mimeographed report Billström et al. (1954) of the presented work has not been published in a journal, but an English reprint has appeared in Billström and Thore (1964) and Thore (1964). These two works discuss the problem of a monopolist facing a linear demand curve that depends on two unknown parameters. Thore (1964) proposes to use a dynamic pricing rule that satisfies  $\text{sign}(p_t - p_{t-1}) = \text{sign}((p_{t-1} - p_{t-2})(r_{t-1} - r_{t-2}))$ , where  $p_t$ ,  $r_t$  denote the price and revenue in period  $t$ . This models the following intuition: if a previous price increase led to an increase in revenue, the price will again be increased; otherwise it will be decreased. Similarly, if a previous price decrease led to an increase in revenue, the price will again be decreased; otherwise, it will be increased. In addition, Thore (1964) proposes to let the magnitude of the price adjustment depend on the difference between the last two revenue observations. He specifies two pricing rules in detail, and analyzes convergence properties of the resulting dynamical systems. Billström and Thore (1964) perform simulation experiments for one of these pricing rules, both in a deterministic demand setting and in a setting where a normally distributed disturbance term is added to the demand. They also extend the model to incorporate inventory replenishment, and provide a rule of thumb for the optimal choice of a constant appearing in the pricing rule.

**Subsequent work.** These studies emerging from the Uppsala Econometrics Seminar have not received much research attention in subsequent years. Clower (1959) studies a monopolist firm facing a linear, deterministic demand function whose parameters may change over time. He discusses several price-adjustment mechanisms that may be applied by the firm to adapt its prices to changing situations. Baumol and Quandt (1964) propose rules of thumb for the monopolist pricing problem, and assess their performance by a set of numerical experiments. In their Appendix A they propose a pricing rule equal to one of the rules proposed by Thore (1964), although they are

apparently unaware of that work. The authors investigate some convergence and stability properties of the resulting dynamical system, both in a discrete-time and continuous-time framework. Baetge et al. (1977) extend the simulation results of Billström and Thore (1964) to non-linear demand curves, and study the optimal choice of a constant appearing in the pricing rules. A final study in this line of research is from Witt (1986). He studies a model where a monopolist has to decide on price, output level in the current period and maximum output in the next period. Three decision rules are compared with each other via a computer simulation. In addition, their performance is compared with a laboratory experiment, where test subjects had to determine their optimal pricing strategy.

#### 4.1.2 Bayesian approaches

Several authors study the dynamic pricing and learning problem within a Bayesian framework.

**Work by Aoki.** One of the first is Aoki (1973), who applies methods from stochastic adaptive control theory. He considers a setting where the demand function depends on unknown parameters, which are learned by the decision maker in a Bayesian fashion. The purpose is to minimize (a function of) the excess demand. He shows how the optimal Bayesian policy can, in theory, be computed via dynamic programming, but that in many situations no closed-form analytical expression of the solution exists. He then proposes two approximation policies. In the first, certainty equivalent pricing (CEP), at each decision moment the price is chosen that would be optimal if the current parameter estimates were correct. In the second, called an approximation under static price expectation, the firm acts at each decision moments as if the chosen price will be maintained throughout the remainder of the selling period. Aoki (1974) shows that the prices generated by (two variants of) CEP and by the static-price-expectation approximation both converge a.s. to the optimal price.

**Variations and extensions to Aoki (1973).** An early study along the same lines of Aoki (1973) is Chong and Cheng (1975). They assume a linear demand function with two unknown parameters and normally distributed disturbance terms, and formulate the optimal pricing problem as a Bayesian dynamic program. They show that certainty equivalent pricing is the optimal policy in case the slope of the demand function is known. For the case that both slope and intercept are unknown, they propose three approximations to the optimal policy: the approximation under static price expectation from Aoki (1973), and two heuristics based on approximations of the value function. Their second approximation selects at each decision epoch the price that maximizes the difference between the expected profit and a term proportional to the covariance of the parameter estimates; this reflects the exploration-exploitation trade-off seen in many later work on optimization under uncertainty.

Other studies closely related to Aoki (1973) are Nguyen (1984, 1997), Wruck (1989), Lobo and Boyd (2003), Chhabra and Das (2011), Qu et al. (2013), and Kwon et al. (2012). Nguyen (1984) considers a quantity-setting monopolist firm facing random demand in multiple periods, where the demand function depends on an unknown parameter which is learned by the firm in a Bayesian fashion. Structural (monotonicity) properties of the optimal policy are derived, and its performance is

compared with a myopic one-period policy. [Nguyen \(1997, Section 5\)](#) discusses similar questions in the context of a price-setting monopolist. [Wruck \(1989\)](#) considers optimal pricing of durable and non-durable goods in a two-period model. The support of the uniformly distributed willingness-to-pay distribution is learned by Bayesian updating a uniform prior, and the optimal price policy is determined by solving a dynamic program. [Lobo and Boyd \(2003\)](#) consider the same setting as [Chong and Cheng \(1975\)](#), and compare via simulation the performance of four pricing policies with each other. [Chhabra and Das \(2011\)](#) study the finite-time performance of standard multi-armed bandit policies and of a policy that (possibly incorrectly) assumes a linear demand function whose unknown parameters are learned by Bayesian updating of a Beta prior. [Qu et al. \(2013\)](#) assume Bernoulli distributed demand with expectation a logit function of price, with a normal prior on the unknown parameters. Because this distributional form hampers exact calculation of posterior distributions, the authors discuss how to calculate an approximation. They propose a Bayesian-greedy price policy which cannot be computed exactly either, and show how an approximation can be calculated. A numerical study compares the pricing policies with a few alternatives. [Kwon et al. \(2012\)](#) study optimal markdown pricing in an infinite-horizon setting with discounted rewards, where the decision variables are the initial price, the markdown price, and the time of markdown. Cumulative demand is modeled as a Brownian motion with unknown drift which is either high or low and which is learned via Bayesian updating. The authors determine the optimal time of markdown, characterize the corresponding optimal initial and markdown prices, and prove a few monotonicity properties.

**Finite action set.** Some literature simplifies the problem by allowing only a finite set of feasible prices. This transforms the pricing-and-learning problem to a Bayesian multi-armed bandit problem. [Leloup and Deveaux \(2001\)](#) show for Bernoulli distributed arms that approximations to Gittins-index policies circumvent the computational problems associated with solving the full Bayesian dynamic program. [Wang \(2007\)](#) allows two feasible prices, assumes compound-Poisson distributed rewards from a certain parametric family, and approximate the Gittins index. [Cope \(2007\)](#) considers a general Dirichlet prior on the discretized reservation-price distribution of customers, and develops approximations to the intractable Bayesian dynamic program that are closely related to Gittins-index policies. He shows that his pricing heuristics converge to the optimal price under an average-reward criterion, and argues that their performance do not suffer much from a misspecified prior distribution.

**Incomplete learning and remedies.** A common theme in the references mentioned above is that it is often intractable to compute the optimal Bayesian policy, and that therefore approximations are necessary. [Rothschild \(1974\)](#) points to a more fundamental problem of the Bayesian framework. He assumes that there are only two prices the firm can choose, with demand for each price Bernoulli distributed with unknown mean. The dynamic pricing problem is thus viewed as a two-armed bandit problem. The optimal Bayesian policy can be computed via the corresponding dynamic programming formulation. The key result of [Rothschild \(1974\)](#) is that, under the optimal Bayesian strategy, with positive probability the price sequence converges to a price that (with hindsight) is not the optimal price. [McLennan \(1984\)](#) derives a similar conclusion in a related setting: the set of admissible prices is continuous, and the relation between price and expected demand is one of two known linear demand curves. It turns out that, under an optimal

Bayesian policy, the sequence of prices may converge with positive probability to a price different from the optimal price. This work is extended by [Harrison et al. \(2012\)](#), who show that in several instances a myopic Bayesian policy may lead to such “incomplete learning”. They propose two modifications of the myopic Bayesian policy that avoid incomplete learning, and prove bounds on their performance. [Afèche and Ata \(2013\)](#) derive incomplete-learning results of similar flavor in the context of pricing different types of customers in an M/M/1 queue. [Cheung et al. \(2013\)](#) extend [Harrison et al. \(2012\)](#) to the case with  $k \in \mathbb{N}$  unknown demand functions, and where in addition at most  $m$  price changes are allowed. They propose a pricing policy, show that it achieves  $\text{Regret}(T) = O(\log^{(m)} T)$  where  $\log^{(m)}$  denotes the  $m$ -times iterated logarithm, and prove that any policy has regret  $\Omega(\log^{(m)} T)$ <sup>2</sup>. Similar results are shown to hold in a continuous-time setting. [Keskin \(2014\)](#) models the cumulative deviation from the expected demand at an incumbent price as the sum of a Wiener process and a drift-term proportional to the difference between selling price and incumbent price. The unknown drift-coefficient is learned by Bayesian updating of the parameters of a Gaussian prior. The author explains why a myopic policy induces incomplete learning, and characterizes the optimal learning-policy as the solution to a partial differential equation (PDE). Based on this policy, he proposes a simple pricing rule that deviates from the myopic price proportionally to the squared coefficient of variation of the posterior belief on the optimal price, and that does not require solving a PDE. Numerical illustrations suggest that the performance of this heuristic is close to optimal.

**Risk-averse pricing.** All studies mentioned above assume that the firm is risk-neutral and optimizes the expected revenue. [Sun and Abbas \(2013\)](#) depart from this assumption by studying the optimal price in a Bayesian dynamic-pricing-and-learning problem with a risk-averse seller. [Choi et al. \(2012\)](#) study a family of simple pricing policies, in a Bayesian setting, based on separating the finite time horizon in an exploration and exploitation phase. They calculate the optimal risk-averse price with respect to a number of risk measures, and provide numerical examples.

**Economics and econometrics literature.** The economics and econometrics literature also contains several studies on pricing and Bayesian learning. [Prescott \(1972\)](#), [Grossman et al. \(1977\)](#), [Mirman et al. \(1993a\)](#) consider two-period models and study the necessity and effects of price experimentation. [Trefler \(1993\)](#) focuses on the direction of experimentation, and applies his results on several pricing problems. [Rustichini and Wolinsky \(1995\)](#) and [Keller and Rady \(1999\)](#) consider a setting where the market environment changes in a Markovian fashion between two known demand functions, and study properties of optimal experimentation. [Balvers and Cosimano \(1990\)](#) consider a dynamic pricing model where the coefficients of a linear demand model change over time, and discuss the implications of active learning. [Willems \(2014\)](#) aims to explain observed discreteness in price data. The author considers a model where the expected demand depends linearly on price via two time-varying parameters with known expectation. The author elaborates a Bayesian learning approach via dynamic programming, and discusses the differences between active and passive learning. [Easley and Kiefer \(1988\)](#), [Kiefer and Nyarko \(1989\)](#), [Aghion et al. \(1991\)](#) are concerned with Bayesian learning in general stochastic control problems with uncertainty. They study the possible limits of Bayesian belief vectors, and show that in some cases these limits may

---

<sup>2</sup>We use  $f(T) = O(g(T))$  to denote  $\sup_{T \in \mathbb{N}} f(T)/g(T) < \infty$ ,  $f(T) = o(g(t))$  to denote  $\limsup_{T \rightarrow \infty} f(T)/g(T) = 0$ , and  $f(T) = \Omega(g(T))$  to denote  $g(T) = O(f(T))$ , when  $f$  and  $g$  are functions on  $\mathbb{N}$ .

differ from the true value. This implies that active experimentation is necessary to obtain strongly consistent control policies.

**Related studies on optimal market design.** Finally, we mention the studies [Manning \(1979\)](#) and [Venezia \(1984\)](#) on optimal design of market research. [Manning \(1979\)](#) considers a monopolist firm facing a finite number of customers. By doing market research, the firm can ask  $n$  potential customers about their demand at some price  $p$ . Such market research is not for free, and the main question of the paper is to determine the optimal amount of market research. This setting is closely related to pricing rules that split the selling season in two parts (e.g. the first pricing rule proposed by [Witt \(1986\)](#)): in the first phase, price experimentation takes place in order to learn the unknown parameters, and in the second phase of the selling season, the myopic price is used. [Venezia \(1984\)](#) considers a linear demand model with unknown parameters, one of which behaves like a random walk. The firm learns about these parameters using Bayes' rule. In addition, the firm can learn the true current value of this random walk by performing market research (which costs money). The author shows how the optimal market-research policy can be obtained from a dynamic program.

#### 4.1.3 Non-Bayesian approaches

**Policies without performance bounds.** Despite the disadvantages of the Bayesian framework outlined above (computational intractability of the optimal solution, the results by [Rothschild \(1974\)](#) and [McLennan \(1984\)](#) on incomplete learning), it has taken several decades before research on pricing policies in a non-Bayesian setting took off. An early exception is [Aoki \(1974\)](#), who proposes a pricing scheme based on stochastic approximation in a non-Bayesian framework. He proves that the prices converge almost surely to the optimal price, and compares the policy with Bayesian pricing schemes introduced in [Aoki \(1973\)](#). More recently, [Carvalho and Puterman \(2005a,b\)](#) and [Morales-Enciso and Branke \(2012\)](#) propose several pricing heuristics based on approximations of an underlying finite-horizon dynamic program whose states contain historical price/demand observations (reminiscent of the dynamic-programming approximations developed by [Aoki \(1973, 1974\)](#)).

**Seminal paper with performance bounds.** A disadvantage of the many pricing heuristics that have been proposed in the literature, both in a Bayesian and a non-Bayesian setting, is that a qualitative statement of their performance is often missing. In many studies the performance of pricing policies is only evaluated numerically, without any analytical results. This changes with the groundbreaking work of [Kleinberg and Leighton \(2003\)](#), who quantify the performance of a pricing policy by  $\text{Regret}(T)$ : the expected loss in  $T$  time periods incurred by not choosing optimal prices. They consider a setting where buyers arrive sequentially to the firm, and buy only if their willingness-to-pay (WtP) exceeds the posted price. Under some additional assumptions, they show that if the WtP of the individual buyers is an i.i.d. sample of a common distribution, then there is no pricing policy that achieves  $\text{Regret}(T) = o(\sqrt{T})$ ; in addition, there is a pricing policy that achieves  $\text{Regret}(T) = O(\sqrt{T \log T})$ . In an adversarial or worst-case setting, where the WtP of individual buyers is not assumed to be i.i.d., they show that no pricing policy can achieve  $\text{Regret}(T) = o(T^{2/3})$ , and that there is a pricing policy with  $\text{Regret}(T) = O(T^{2/3}(\log T)^{1/3})$ .



**Parametric approaches.** Le Guen (2008), Broder and Rusmevichientong (2012), den Boer and Zwart (2014a), den Boer (2014a), and Keskin and Zeevi (2014) take a parametric approach, using classical maximum (quasi)likelihood or least-squares estimators to estimate unknown parameters of the demand function. Le Guen (2008) considers a multi-product setting with linear demand, assuming a particular structure on the unknown parameter matrices. He shows that certainty equivalent pricing augmented with price experimentation at predetermined time intervals leads to prices converging to the true optimal price, and proposes a pricing heuristic for non-linear demand functions. Assuming a single product setting with Bernoulli distributed demand, Broder and Rusmevichientong (2012) show a  $\sqrt{T}$  lower bound on the regret, using information-theoretic inequalities and techniques found in Besbes and Zeevi (2011), and show that a pricing policy that strictly separates exploration from exploitation achieves  $O(\sqrt{T})$  regret. This rate can be improved to  $O(\log T)$  if the demand-function is such that there are no “uninformative prices”: prices at which the expected demand given a (erroneous) parameter estimate is equal to the true expected demand (the existence of such prices plays an important role in the best achievable growth rate of the regret; cf. Harrison et al. (2012)). den Boer and Zwart (2014a) consider an extended class of generalized linear single-product demand models, and show in an example that certainty equivalent pricing is not strongly consistent. They propose a pricing policy that always chooses the price closest to the certainty equivalent price that guarantees a minimum amount of price dispersion. This price dispersion, measured by the sample variance of the selling prices, guarantees convergence of the prices to the optimal price, and leads to  $\text{Regret}(T) = O(T^{1/2+\delta})$ , for arbitrary small  $\delta > 0$ . den Boer (2014a) extends this policy to multiple products, attaining  $\text{Regret}(T) = O(\sqrt{T \log T})$  for so-called canonical link functions and  $\text{Regret}(T) = O(T^{2/3})$  for general link functions. Keskin and Zeevi (2014) assume a linear demand function, show a  $\sqrt{T}$  lower bound on the regret using proof techniques different from Broder and Rusmevichientong (2012), and generalize Broder and Rusmevichientong (2012) and den Boer and Zwart (2014a) by providing sufficient conditions for any pricing policy to achieve  $\text{Regret}(T) = O(\sqrt{T} \log T)$ . Assuming that the mean demand is exactly known at a certain price, they show that  $\text{Regret}(T) = O(\log T)$  is attainable. Both these results are extended to the multiple-product setting, focusing on a class of so-called orthogonal pricing policies.

**Robust optimization.** Eren and Maglaras (2010) study dynamic pricing in a robust optimization setting. They show that if an infinite number of goods can be sold during a finite time interval, it is optimal to use a price-skimming strategy. They also study settings where learning of the demand function occurs, but under the rather strong assumption that observed demand realizations are without noise. Bergemann and Schlag (2008, 2011) and Handel et al. (2013) also consider pricing in a robust framework, but their setting is static instead of dynamic. Handel and Misra (2014) consider a two-period model where a monopolist sets prices based on a set of demand curves feasible with acquired sales data. The authors describe and analyze the optimal two-period pricing policy that minimizes a dynamic version of the minimax regret, and investigate how customer preferences influence the difference between dynamic and static prices.

**Finite action set.** If the demand model is assumed to lie in a finite set of known demand functions, the dynamic-pricing-and-learning problem can be regarded as a multi-armed bandit problem with dependent arms. This viewpoint is taken by Tehrani et al. (2012), who develop a



pricing policy based on the likelihood-ratio test, and show that its regret is bounded assuming that there are no uninformative prices. Their work can be viewed as a non-Bayesian counterpart to [Harrison et al. \(2012\)](#) and [Cheung et al. \(2013\)](#).

**Variants.** Pricing without demand information in a queueing model is studied by [Haviv and Randhawa \(2014\)](#). They consider the problem of pricing delay-sensitive customers in an unobservable M/M/1 queue. The purpose of the paper is to study the impact of ignoring arrival-rate information on the optimal pricing strategy. The authors find that a policy that ignores this information performs surprisingly well, and in some cases can still capture 99% of the optimal revenue.

Finally, we mention [Jia et al. \(2013\)](#), who consider dynamic pricing and learning in an electricity market, where the goal is to steer the expected demand to a desired level. This particular objective is reminiscent of the multi-period control problem discussed in [Taylor \(1974\)](#) and [Anderson and Taylor \(1976\)](#). A stochastic-approximation type policy inspired by [Lai and Robbins \(1982\)](#) is shown to achieve  $O(\log T)$  regret, and in addition it is shown that no policy can achieve sub-logarithmic regret. The fact that logarithmic instead of  $\sqrt{T}$  regret can be achieved is caused by subtle differences between dynamic pricing and this multi-period control problem, which are further discussed in [den Boer and Zwart \(2014a, Remark 1\)](#).

## 4.2 Finite inventory

We here discuss the literature on dynamic pricing and learning in the presence of finite inventory that cannot be replenished. Most of the studies assume a finite selling season, corresponding to the models discussed in Section 3.2.1. Some studies assume an infinite time horizon, and consider the objective of maximizing total discounted reward.

### 4.2.1 Early work

[Lazear \(1986\)](#) considers a simple model where a firm sells one item during at most two periods. In the first period a number of customers visit the store; if none of them buys the item, the firm adapts its prior belief on the value of the product, updates the selling price, and tries to sell the item in the second period. The author shows that the expected profit increases by having two selling periods instead of one. He extends his model in several directions, notably by allowing strategic behavior of customers who may postpone their purchase decision if they anticipate a price decrease. [Sass \(1988\)](#) extends the model of Lazear and studies the relation between the optimal price strategy and the number of potential buyers.

### 4.2.2 Bayesian approaches

**Unknown arrival rate, known willingness-to-pay.** [Aviv and Pazgal \(2005b\)](#) start a research stream on Bayesian learning in dynamic pricing with finite inventory. Customers arrive according to a Poisson process with unknown arrival rate, and purchase a product with (known) probability

$\exp(-p)$ , where  $p$  is the current selling price. The unknown arrival rate is learned via Bayesian updates of a Gamma prior. The authors characterize the optimal continuous-time pricing policy by means of a differential equation. Because this equation does not always admit an explicit solution, three pricing heuristics are proposed: certainty equivalent pricing (CEP), a fixed price policy, and a naive pricing policy that ignores uncertainty on the market. Numerical experiments suggest that CEP performs quite well. [Lin \(2006\)](#) considers a similar setting, allowing for general willingness-to-pay distributions. He proposes a pricing policy where the seller sets the price based on repeatedly updated estimates of the demand distribution, and evaluates its performance via simulations.

[Araman and Caldentey \(2009\)](#), [Farias and van Roy \(2010\)](#) and [Mason and Välimäki \(2011\)](#) study the infinite-horizon discounted-reward case. [Araman and Caldentey \(2009\)](#) assume a known willingness-to-pay distribution and a two-point prior distribution on the unknown arrival rate, propose a pricing heuristic based on an asymptotic approximation of the value function of the underlying optimal-control problem, and compare its performance numerically with CEP, static pricing, and a two-price policy. In a similar setting [Farias and van Roy \(2010\)](#) take a finite mixture of gamma distributions as prior, propose a pricing heuristic called decay balancing, and show numerically that it frequently outperforms CEP and the heuristics of [Araman and Caldentey \(2009\)](#). They further show that the expected discounted reward obtained from decay balancing is at least one third of the optimal reward, and discuss an extension to multiple stores. [Mason and Välimäki \(2011\)](#) assume that only a single item is sold, with either high or low customer arrival-rate which is learned in a Bayesian fashion. The authors study structural properties of the optimal price policy and compare it to policies that neglect learning.

[Avramidis \(2013\)](#) observes that the number of arrivals or the number of sales is sufficient to compute the posterior arrival-rate distribution, given *any* prior. This means that the setting considered by [Aviv and Pazgal \(2005b\)](#), [Lin \(2006\)](#), [Araman and Caldentey \(2009\)](#) and [Farias and van Roy \(2010\)](#) can be resolved without imposing a specific family of priors (Gamma, two-point discrete, finite mixture of Gammas).

**Unknown arrival rate and unknown willingness-to-pay.** [Chen and Wang \(1999\)](#) consider pricing of a single asset in an infinite horizon with discounted rewards, and assume that the willingness-to-pay distribution is unknown but equal to one of two known distributions. The authors formulate a Bayesian dynamic program and prove that the optimal prices decline over time if the hazard rates of these distributions can be ranked uniformly; a counterexample shows that this condition cannot be relaxed to first-order stochastic dominance. [Sen and Zhang \(2009\)](#) extend the model of [Aviv and Pazgal \(2005b\)](#) by assuming that the purchase probabilities of arriving customers are not known to the firm. They assume that the demand distribution is an element of a finite known set, and consider a discrete-time setting with Bayesian learning and a gamma prior on the arrival rate. The optimal pricing policy can be explicitly calculated, and, in an extensive computational study, its performance is compared with both a full information setting and a setting where no learning occurs.

**Partially Observable Markov Decision Processes.** [Aviv and Pazgal \(2005a\)](#) consider a Markov-modulated demand environment, modelled by a Markov chain where each of the finite

states corresponds to a different known demand function, and where the state of the system is learned in a Bayesian fashion. The optimal pricing problem is formulated as a Partially Observable Markov Decision Process, that turns out to be computationally intractable. Various approximate solutions are proposed that rely on modifying the information structure of the problem, and their performance is evaluated in a numerical study. [Chen \(2014\)](#) considers a similar partially observable Markov decision process. In his setting, the seller estimates the willingness-to-pay distribution of customers based on two-sided censored observations. Three near-optimal price heuristics are proposed and their performance is assessed by numerical experiments. For exponentially or Weibull distributed demand, more refined results on the behavior of the heuristics are obtained.

#### 4.2.3 Non-Bayesian approaches

**Asymptotic regime where inventory grows large.** The optimal pricing problem studied by [Gallego and van Ryzin \(1994\)](#) does often not admit an explicit solution, and the authors therefore consider an asymptotic regime where both the demand and the level of inventory grow large. They prove that the optimal revenue obtained in this asymptotic regime serves as an upper bound for the optimal expected revenue of the original problem, and show that the optimal asymptotic pricing policy is to use a static price throughout the sales horizon. This static price is the maximum of the unconstrained optimal price (the revenue-maximizing price in the case of infinite inventory) and of the clearance price (the price that induces a stock-out precisely at the end of the sales horizon).

[Besbes and Zeevi \(2009\)](#) initiate a stream of literature that attempts to learn this optimal static price in an incomplete-information setting. For both a parametric and non-parametric setting they develop pricing policies, based on the idea of dividing the sales horizon into an “exploration” phase during which the demand function is learned and an “exploitation” phase during which the perceived optimal price is used. To establish performance bounds, they consider a sequence of problems indexed by  $n \in \mathbb{N}$ , where the  $n$ -th problem has initial inventory  $nx$  and demand function  $n\lambda$ , for some  $x > 0$  and some function  $\lambda$ . They prove an  $O(n^{-1/4}(\log n)^{1/2})$  upper bound on the relative regret of this policy in the non-parametric setting, an  $O(n^{-1/3}(\log n)^{1/2})$  bound in the parametric setting, and an  $O(n^{-1/2}(\log \log n)(\log n)^{1/2})$  bound in case the demand function is known up to a single unknown parameter. These upper bounds are complemented by results showing that all policies have relative regret  $\Omega(n^{-1/2})$ . In a non-parametric setting, [Wang et al. \(2014\)](#) improve these upper bounds by developing a policy with relative regret  $O(n^{-1/2}(\log n)^{4.5})$ ; apart from the logarithmic term, this policy thus achieves the asymptotically optimal regret rate. The pricing policy is based on the idea of iterative price experimentation in a shrinking series of intervals that with high probability contain the optimal price. [Lei et al. \(2014\)](#) provide a further improvement by even removing the logarithmic terms in the upper bound: they propose and analyze three pricing algorithms based on ideas from bisection search and stochastic approximation, and show that one of these algorithms achieves  $O(n^{-1/2})$  relative regret.

[Besbes and Zeevi \(2012\)](#) extend [Besbes and Zeevi \(2009\)](#) to a setting where multiple products share the same finite resources. They show that a policy which separates “exploration” and “exploitation” achieves relative regret  $O(n^{-1/(3+d)}(\log n)^{1/2})$ , where  $d$  is the number of products. This is improved to  $O(n^{-1/(3+d/s)}(\log n)^{1/2})$  if one assumes additional smoothness conditions on

the demand function, including  $s$ -times differentiability, and to  $O(n^{-1/3}(\log n)^{1/2})$  if the set of feasible prices is discrete and finite; this last setting is accompanied by an  $\Omega(n^{-1/2})$  lower bound on the relative regret of any policy. Another variant of the problem is studied by [Besbes and Maglaras \(2012\)](#), who study the situation where certain financial milestone constraints in terms of sales and revenues targets are imposed. They formulate a pricing policy that periodically updates its pricing decisions in order to the most stringent financial constraint, and show an  $O(n^{-1/2}(\log n)^{1/2})$  bound on the relative regret. [Avramidis \(2013\)](#) modifies the type of policy proposed by [Besbes and Zeevi \(2009\)](#) by estimating both arrival rate and purchase probabilities, and by using the solution of the finite-time Markov decision problem in the exploitation phase instead of the solution that corresponds to the asymptotic regime. Numerical experiments suggest that these modifications lead to lower regret.

**Alternative asymptotic regimes.** The high-volume regime of the studies mentioned above is not applicable for settings with relatively low inventory. This motivates [den Boer and Zwart \(2014c\)](#) to study a setting with multiple, consecutive, finite selling seasons and finite inventories. They show in a parametric framework that this setting satisfies a certain endogenous-learning property, which implies that price experimentation is hardly necessary to eventually learn the optimal prices. The authors prove that the (cumulative) regret of a small modification of certainty equivalent pricing is  $O(\log^2 T)$  after  $T$  selling seasons, and that any policy has regret  $\Omega(\log T)$ .

In [Babaioff et al. \(2014\)](#), a firm tries to sell  $k$  items to  $n$  potential customers. The authors propose an online pricing-and-learning policy that does not require parametric assumptions on the demand distribution, and that exploits the fact that the problem is closely related to multi-armed bandit problems (despite the finite inventory, which means that the optimal price depends on the current inventory level). They show a regret bound of  $O((k \log n)^{2/3})$ , and provide a matching lower bound on the regret. If the ratio  $k/n$  is sufficiently small this bound is improved to  $O(\sqrt{k} \log n)$ .

[Bertsimas and Perakis \(2006\)](#) consider pricing in a least-squares-learning setting with a single selling season. They formulate a dynamic program that describes the optimal pricing policy but is computationally intractable, and propose approximate solutions based on state-space reductions. For an extended model with competitors and slowly varying parameters, estimation methods and price policies are discussed. No performance bounds on pricing policies are provided.

**Robust optimization.** A number of studies take a robust approach, where the demand function is not learned over time but assumed to lie in some known uncertainty set. [Lim and Shanthikumar \(2007\)](#) and [Thiele \(2006\)](#) study this in a single-product setting, and [Lim et al. \(2008\)](#), [Thiele \(2009\)](#) in a multi-product setting. With these robust approaches no learning takes place, despite the accumulation of sales data. [Cohen et al. \(2013\)](#) develop an approach to dynamic pricing and learning that attempts to bridge this gap between robust and data-driven approaches, by sampling different scenarios from historical sales data. A robust extension of [Araman and Caldentey \(2009\)](#) and [Farias and van Roy \(2010\)](#), where finite inventory is sold during an infinite time horizon, is studied by [Li et al. \(2009\)](#). Another approach that does not rely on historical demand data is [Xiong et al. \(2010\)](#) (see also [Li and Li, 2010](#), [Li et al., 2013](#), [Coggun et al., 2014](#)). They model demand uncertainty using fuzzy set theory, propose different fuzzy programming models, and present an algorithm based on fuzzy simulation and a genetic algorithm to solve these problems.

[Dziecichowicz et al. \(2010\)](#) do not focus on determining optimal prices, but instead study the optimal timing of markdowns. The authors derive a robust optimization problem to determine the optimal markdown policy, which in some special cases can be solved exactly and in other cases can be approximated by a mixed-integer problem. The results are also extended to multiple products. Finally, [Ferrer et al. \(2012\)](#) assume that demand is a non-random function of price, and lies in an uncertainty set of demand functions known to the decision maker. The authors introduce a measure of risk aversion, and study the relation between risk-aversion and properties of the optimal pricing policy, both theoretically and by numerical simulations.

**Variants.** [Gallego and Talebian \(2012\)](#) consider a setting where a finite number of products is offered in multiple versions during a finite sales horizon. Demand is modeled by a customer choice model. The different product-versions share an unknown “core value”, which is estimated by maximum likelihood estimation. The possible time-varying arrival rate of customers is learned in a Bayesian fashion. The authors develop a pricing rule in a rolling horizon framework, and illustrate its behavior by a computational study. Somewhat related to this is [Berg and Ehtamo \(2012\)](#), where a firm sells different versions of a product to different customer segments. The utility functions of each segment are partly unknown, and learned by the firm using stochastic gradient methods or variants thereof.

### 4.3 Machine-learning approaches

A considerable stream of literature on dynamic pricing and learning has emerged from the computer science community. In general, the focus of these papers is not to provide a mathematical analysis of the performance of pricing policies, but rather to design a realistic model for electronic markets and subsequently apply machine learning techniques. An advantage of this approach is that one can model many phenomena that influence the demand, such as competition, fluctuating demand, and strategic buyer behavior. A drawback is that these models are often too complex to analyze analytically, and insights on the behavior of various pricing strategies can only be obtained by performing numerical experiments.

Machine-learning techniques that have been applied to dynamic pricing problems include evolutionary algorithms ([Shakya et al., 2009](#), [Ramezani et al., 2011](#)), particle swarm optimization ([Mullen et al., 2006](#)), reinforcement learning and Q-learning ([Kutschinski et al., 2003](#), [Raju et al., 2006](#), [Könönen, 2006](#), [Chinthalapati et al., 2006](#), [Schwind, 2007](#), [Cheng, 2008](#), [Han et al., 2008](#), [Vengerov, 2008](#), [Cheng, 2009](#), [Jintian and Lei, 2009](#), [Han, 2010](#), [Collins and Thomas, 2012, 2013](#), [Dogan and Güner, 2013](#), [Rana and Oliveira, 2014, 2015](#)), simulated annealing ([Xia and Dube, 2007](#)), Markov chain Monte Carlo methods ([Chung et al., 2012](#)), the aggregating algorithm ([Levin et al., 2009](#)) by [Vovk \(1990\)](#), goal-directed and derivative-following strategies in simulation ([DiMicco et al., 2003](#), [Vázquez-Gallo et al., 2014](#)), neural networks ([Brooks et al., 1999](#), [Kong, 2004](#), [Ghose and Tran, 2009](#), [Shakya et al., 2012](#), [Liu and Wang, 2013](#)), and direct search methods ([Brooks et al., 1999, 2001](#), [Kephart et al., 2001](#), [Brooks et al., 2002](#)).

These papers all use very different models, methods, assumptions and performance metrics. This makes it hard to compare different papers with each other, which we therefore do not attempt.

## 4.4 Joint pricing and inventory problems

A few studies consider the problem of simultaneously determining an optimal pricing and inventory replenishment policy while learning about the demand.

**Parametric approaches.** Most of them consider learning in a Bayesian framework. [Subrahmanyam and Shoemaker \(1996\)](#) assume that the unknown demand function lies in a finite known set of demand functions, and is learned in a Bayesian fashion. The optimal policy is determined by solving a dynamic program. Several numerical experiments are provided to offer insight in the properties of the pricing policy. [Bitran and Wadhwa \(1996\)](#) and [Bisi and Dada \(2007\)](#) study a similar type of problem, where an unknown parameter is learned in a Bayesian manner, and where the optimal decisions are determined by solving a dynamic program. [Bitran and Wadhwa \(1996\)](#) perform extensive computational experiments, and [Bisi and Dada \(2007\)](#) derive several structural properties of the optimal policy. [Zhang and Chen \(2006\)](#) consider Bayesian learning of a component of the demand distribution that does not depend on the selling price, and show that the finite-horizon expected discounted profit is maximized by a so-called base stock list price policy ([Porteus, 1990](#), [Federgruen and Heching, 1999](#)). Motivated by industrial practice of fashion retailers, [Choi \(2007\)](#) considers a scenario where the seller can order inventory and change selling prices at two distinct stages in the selling season. Information obtained from the first stage is used in a Bayesian manner to determine optimal decisions in the second stage. The author formulates a dynamic program, proves several structural properties, and carries out numerical experiments to illustrate his results. [Gao et al. \(2010\)](#) formulate and solve a two-period two-products pricing and inventory problem with Bayesian learning of unknown demand parameters. [Forghani et al. \(2013\)](#) allow for a single price change, formulate the optimal control problem, and show numerical examples. In [Boyacı and Özer \(2010\)](#), a manufacturer estimates the optimal capacity decision from advance sales information obtained prior to the regular selling season. Formal learning of demand parameters is not considered, but the electronic companion to the paper elaborates upon an extension to Bayesian learning.

**Non-parametric and robust approaches.** [Burnetas and Smith \(2000\)](#) consider a joint pricing and inventory problem in a non-parametric setting. They propose an adaptive stochastic-approximation policy, and show that the expected profit per period converges to the optimal profit under complete information. A robust approach to the dynamic pricing and inventory control problem with multiple products is studied by [Adida and Perakis \(2006\)](#). The focus of that paper is the formulation of the robust optimization problem, and to study its complexity properties. Related is the work of [Petruzzi and Dada \(2002\)](#). These authors assume that there is no demand noise, which means that the unknown parameters that determine the demand function are completely known once a demand realization is observed that does not lead to stock-out. [Adida and Perakis \(2010a\)](#) discuss several robust and stochastic optimization approaches to joint pricing and procurement under demand uncertainty, and compare these approaches with each other in a numerical study. [Mahmoudzadeh et al. \(2013\)](#) formulate a robust control problem for joint pricing and production in a hybrid manufacturing/remanufacturing system, where the coefficients of a linear demand function are unknown. [Arasteh et al. \(2014\)](#) consider a similar setting for a joint pricing-and-inventory problem.

**Variants.** Lariviere and Porteus (1995) consider the situation of a manufacturer that sells to a retailer. The manufacturer decides on a wholesale price offered to the retailer, and the retailer has to choose an optimal inventory replenishment policy. Both learn about a parameterized demand function in a Bayesian fashion. Properties of the optimal policy, both for the manufacturer and the retailer, are studied. Gaul and Azizi (2010) assume that a product is sold in different stores. The problem is to determine optimal prices in a finite number of periods, as well as to decide if and how inventory should be reallocated between stores. Parameters of the demand function are learned by Bayesian updating, and numerical experiments are provided to illustrate the method.

## 5 Methodologically Related Areas

Dynamic pricing under uncertainty is closely related to multi-armed bandit problems. This is a class of problems that capture many essential features of optimization problems under uncertainty, including the well-known exploration-exploitation trade-off: the decision maker should properly balance the two objectives of maximizing instant reward (exploitation of current knowledge) and learning the unknown properties of the system (exploration). This trade-off between learning and instant optimization is also frequently observed in dynamic pricing problems. The literature on multi-armed bandit problems is large; some key references are Thompson (1933), Robbins (1952), Lai and Robbins (1985), Gittins (1989), Auer et al. (2002); see further Vermorel and Mohri (2005), Cesa-Bianchi and Lugosi (2006), and Powell (2010). If in a dynamic pricing problem the number of admissible selling prices is finite, the problem can be modeled as a classical multi-armed bandit problem. This approach is e.g. taken by Rothschild (1974), Xia and Dube (2007), and Cope (2007). If the set of admissible selling prices is a continuum, then the dynamic pricing problem is closely related to the continuum-armed bandit problem. This problem has recently been studied, among others, by Kleinberg (2005), Auer et al. (2007), Cope (2009), Wang et al. (2009), Goldenshluger and Zeevi (2009), Rusmevichientong and Tsitsiklis (2010), Filippi et al. (2010), Abbasi-Yadkori et al. (2011), Yu and Mannor (2011), Slivkins (2011), Perchet and Rigollet (2013), Combes and Proutiere (2014). Pricing and learning in a time-varying market, as discussed in Section 6.3, is related to the non-stationary multi-armed bandit problem (Garivier and Moulines, 2011, Besbes et al., 2014).

Another important area related to dynamic pricing and learning is the study of convergence rates of statistical estimates. Lai and Wei (1982) study how the speed of convergence of least-squares linear regression estimates depend on the amount of dispersion in the explanatory variables. Their results are applied in several dynamic pricing problems with linear demand functions, such as Le Guen (2008) and Cooper et al. (2014). Similarly, results on the convergence rate of maximum-likelihood estimators, as in Borovkov (1998) and den Boer and Zwart (2014b), are crucial in the analysis of pricing policies by Besbes and Zeevi (2011), Broder and Rusmevichientong (2012) and den Boer and Zwart (2014a).

Dynamic pricing and learning can also be put in a general framework of stochastic control problems with parametric uncertainty. For various dynamic economic models, such problems have been considered by Easley and Kiefer (1988), Kiefer and Nyarko (1989), Marcet and Sargent (1989),



Wieland (2000a,b), Beck and Wieland (2002), Han et al. (2006), amongst many others. Kendrick et al. (2008) review some of this literature, focusing on continuous-time optimal control problems with linear system equations, quadratic cost functions, and Gaussian additive noise terms.

Clearly, many sequential optimization problems in operations research that are studied in an online-learning setting have methodological similarities with dynamic pricing and learning: inventory control (Scarf, 1959, Azoury, 1985, Huh et al., 2011, is just a small sample of the vast literature), online advertisement (e.g. Nazerzadeh et al., 2013), resource allocation (e.g. Ciocan and Farias, 2012), assortment planning (e.g. Caro and Gallien, 2007, Rusmevichientong et al., 2010, Sauré and Zeevi, 2013), product launch and product exit optimization (e.g. Hitsch, 2006), et cetera. These problems can also be studied in conjunction with optimal pricing. For example, Talebian et al. (2014) study simultaneous pricing and assortment optimization with Bayesian demand learning, and Section 4.4 lists a number of studies that combine pricing and inventory control in an incomplete-information setting.

## 6 Extensions and New Directions

Most of the literature discussed above studies dynamic pricing and learning for a monopolist firm that sells a single product to nonstrategic customers in a stationary market environment. Here we review literature on a number of extensions to this setting: strategic customer behavior, competition, time-varying markets, and model misspecification. Note that the first three of these extensions can in some sense be regarded as extending the literature of Section 3.1, on dynamic pricing with dynamic demand, to an incomplete-information framework.

### 6.1 Strategic consumer behavior

The importance of incorporating the effect of strategically behaving customers on the dynamic pricing policy has repeatedly been recognized; cf. the papers on dynamic pricing with strategic customer behavior mentioned in Section 3.2.1. The study of dynamic pricing with both strategic customer behavior and demand learning, however, is not that well developed.

**Learning of the willingness-to-pay distribution.** Studies in which the demand distribution is not assumed to be known are Loginova and Taylor (2008), Levina et al. (2009), Weaver and Moon (2011) and Caldentey et al. (2013), and Lazear (1986) which is discussed in Section 4.2.1. In Loginova and Taylor (2008), the seller learns the willingness-to-pay in a stylized model with a single customer whose product valuation is either “high” or “low”. The authors characterize game-theoretic equilibria and study the effect of strategic behavior of the buyer. Levina et al. (2009) consider a seller who learns about a complicated demand process that incorporates, among other things, strategic customer behavior. This behavior is described using a game-theoretic consumer choice model. The focus of the paper is on properties and numerical performance of an online-learning algorithm proposed by the authors. Weaver and Moon (2011) and Caldentey et al. (2013) take a robust-optimization approach. Inspired by sales of food and agriculture products, Weaver



and Moon (2011) formulate a robust pricing model with multiple products, multiple customer segments, and linear demand function, and numerically compare its performance with two “price-assurance policies”: a policy that offers refund to consumers if the price falls below their purchase price, and a policy that ensures that price will not decline sufficiently to induce refunds. Caldentey et al. (2013) consider a robust formulation of the seller’s pricing problem for a discrete-time infinite-inventory setting with unknown valuations and arrival times of customers, and characterize the pricing policies that minimize the worst-case regret, both in case of myopic or strategic customers. The authors also analyze the limit of the regret and optimal pricing strategies as the length of the discrete time periods goes to zero.

**No learning of the willingness-to-pay distribution.** In Chade and de Serio (2002), Ifrach et al. (2013) and Yu et al. (2013), the unknown quality of the product is learned. This introduces a game with strategic behavior of customers, who may postpone their purchase in order to retrieve more information, induce a markdown, or influence the belief of the seller about the quality of product. In Liu and van Ryzin (2011) consumers learn about the capacity of the seller in a repeated two-period selling season, and in Swinney (2011) consumers learn their valuations while the firm learns the market size. Mersereau and Zhang (2012) considers a firm that learns the relative amount of strategic customers. None of these studies involve learning the demand or willingness-to-pay distribution (given the quality of the product).

## 6.2 Competition

Several studies address the issue of incorporating competition in dynamic pricing and learning. This is challenging, because one generally does not know how competitors behave, and their behavior influences a firm’s own optimal pricing policy. The literature references in this section are not comprehensive, but meant to give a flavor of available approaches and results.

**Parametric approaches.** The already mentioned paper by Bertsimas and Perakis (2006) considers a least-squares learning setting with finite inventories, multiple competitors, and linear demand functions with slowly varying parameters. The authors propose an estimation and price-optimization scheme, and illustrate the method by numerical examples. The related study Kachani et al. (2007) considers multiple quantity-setting firms selling the same type of products with finite inventories. The price-demand relation is linear, with unknown parameters that may slowly vary over time. The authors formulate the optimization problem as a mathematical program with equilibrium constraints, provide some computational results, and derive a closed-form expression of a limiting Nash equilibrium in case of two firms and a selling horizon of two time periods. Kwon et al. (2009) formulate a finite-time finite-capacity oligopolistic pricing problem in the language of differential variational inequalities. They consider a demand model where the time-derivative of demand is proportional to the difference between a firm’s price and a moving average of prices used in the past. Unknown model parameters are estimated with a Kalman filter, which is recalculated at several moments in the sales horizon. The authors propose an algorithm to solve the resulting differential variational inequalities, and illustrate their method by a numerical example. A similar setting is considered by Li et al. (2010) and Chung et al. (2012), who incorporate a non-parametric

functional-coefficient autoregressive time series in their demand model. Duopoly models are studied, among many others, by [Coughlan and Mantrala \(1994\)](#) in a Bayesian framework and by [Choi and Jagpal \(2004\)](#) in a setting with risk-averse firms.

**Robust approaches.** A robust optimization approach to dynamic pricing in an oligopolistic environment with demand uncertainty is taken by [Perakis and Sood \(2006\)](#) (see also the closely related study [Friesz et al., 2012](#)). They consider a multi-period setting with fixed, finite inventory levels, and study Nash equilibrium price policies. [Adida and Perakis \(2010b\)](#) consider a joint pricing-and-inventory control problem with duopolistic competition and demand uncertainty, and formulate it as a robust optimization problem. The authors discuss existence and uniqueness of a Nash equilibrium and address various computational aspects of the problem.

**Economics and econometrics literature.** A large stream of economics literature is concerned with the long-term behavior of price adjustment processes in oligopolistic settings, cf. [Cyert and DeGroot \(1970\)](#), [Kirman \(1975\)](#), [Aghion et al. \(1993\)](#), [Mirman et al. \(1993b\)](#), [Fishman and Gandal \(1994\)](#), [Harrington \(1995\)](#), [Bergemann and Valimaki \(1996\)](#), [Gallego \(1998\)](#), [Alepuz and Urbano \(1999\)](#), [Rassenti et al. \(2000\)](#), [Belleflamme and Bloch \(2001\)](#), [Schinkel et al. \(2002\)](#), [Keller and Rady \(2003\)](#), [Dimitrova and Schlee \(2003\)](#), [Tuinstra \(2004\)](#). One usually assumes that firms are using a certain specific learning scheme, and then studies whether the selling prices converge to a Nash equilibrium. While these papers are interesting from the perspective of understanding price formation, application in business practice is hampered by the fact that one generally does not know which learning scheme competing firms are using. A remarkable early study on pricing and learning in a competitive environment is [Barta and Varaiva \(1976\)](#), who study convergence properties of various stochastic-approximation based price-adjustment rules.

**Computer science literature.** Contributions from the computer science community to dynamic pricing with demand uncertainty in a competitive environment include [Greenwald and Kephart \(1999\)](#), [Dasgupta and Das \(2000\)](#), [Tesauro and Kephart \(2002\)](#), [Kutschinski et al. \(2003\)](#), [Könönen \(2006\)](#), [Jumadinova and Dasgupta \(2008, 2010\)](#).

### 6.3 Time-varying market parameters

Recently several studies have appeared that depart from the strong assumption of unchanging market conditions. We here only list papers that include mathematical performance analysis of the proposed pricing rules.

[Besbes and Sauré \(2014\)](#) study the classical finite-inventory finite selling-season pricing problem in a setting with deterministic demand functions. The seller anticipates that the initial demand function  $d$  may at some moment  $\tau$  change to  $d^\theta \in \{d^1, \dots, d^K\}$ , where  $(\tau, \theta)$  are random variables with known distributions and where  $d^1, \dots, d^K$  are known functions. The authors formulate the optimal pricing problem and derive conditions that induce monotonicity properties of the optimal pricing strategy. [Chen and Farias \(2013\)](#) also study pricing in a finite-inventory setting where the market process is fluctuating over time. They propose a pricing heuristic which is based on the idea of frequently re-calculating the estimated optimal fixed price. The authors prove bounds on

the performance loss relative to an optimal clairvoyant pricing strategy in an asymptotic regime where inventory and demand grows large.

In an infinite-inventory setting, [Besbes and Zeevi \(2011\)](#) study a pricing problem where customers' willingness-to-pay (WtP) distribution changes at some unknown point in time. The WtP distribution before and after the change is assumed to be known, only the time of change is unknown to the seller. Lower bounds on the worst-case regret are derived, and pricing strategies are developed that achieve the order of these bounds. [Keskin and Zeevi \(2013\)](#) consider a linear demand function with time-varying parameters, which are estimated with weighted least squares. They measure the amount of change in the parameter process by the quadratic variation over  $T$  periods, and assume it is bounded by  $BT^\nu$ , for some known  $\nu \in [0, 1]$  and unknown  $B \geq 0$ . An  $\Omega(T^{(2+\nu)/3})$  lower bound on the regret attainable by any policy is proven, and a pricing policy that separates the time horizon into pure exploration and exploitation periods is shown to have regret  $O(T^{(2+\nu)/3})$ . More refined expressions are derived in case  $B$  is known. If only "bursty" changes occur, the authors show that a regret of  $O(T^{1/2} \log T)$  can be achieved. [den Boer \(2014b\)](#) considers a similar set-up with linear demand and weighted least-squares estimation, but assumes that only the size of the market (the intercept) is unknown. A myopic pricing policy is considered, and upper bounds on the long-term average regret are provided. The methodology is illustrated by applying it to pricing in the Bass model, and to pricing in a competitive environment.

A particular structural form of the evolution of the market process is assumed by [Wang et al. \(2013\)](#) and [Chakravarty et al. \(2014\)](#). [Wang et al. \(2013\)](#) assume that the demand intensity over the life-cycle of an information good behaves as the density function of a Weibull distribution, whose unknown parameters are learned in a Bayesian fashion using Monte-Carlo simulation techniques. [Chakravarty et al. \(2014\)](#) propose a simulation-based algorithm to calculate the optimal price path for a product-diffusion model introduced by [Robinson and Lakhani \(1975\)](#).

## 6.4 Model misspecification

Studies that assume a parametric form of the demand functions face the risk of model misspecification. [Besbes and Zeevi \(2014\)](#) show in a single-product setting that the loss incurred by such misspecification may in some settings be not that significant. Another form of misspecification is incorrectly assuming that there are no competitors present. [Schinkel et al. \(2002\)](#), [Tuinstra \(2004\)](#), [Bischi et al. \(2004, 2007\)](#), [Isler and Imhof \(2008\)](#), [Cooper et al. \(2014\)](#), [Anufriev et al. \(2013\)](#) study the effect of this error on the resulting equilibria in various (linear) models, elaborating on earlier work by Kirman ([Kirman, 1975, 1983, 1995](#), [Brousseau and Kirman, 1992](#)). In an airline revenue-management setting, [Cooper et al. \(2006\)](#) show that incorrectly assuming high-fare and low-fare class passengers behave independently can be detrimental for the firm's revenue.

## 7 Conclusion

Dynamic pricing with incomplete demand information is a topic that has received considerable research attention in recent years. Different scientific communities have studied characteristics of pricing policies, usually with different aims in mind: the operations research / management science literature typically focuses on finding an optimal pricing policy from the perspective of a seller, and on proving optimality properties in tractable models; the economics literature is generally more concerned with explaining price behavior and price formation observed in markets; and the computer science literature is typically not afraid to consider complicated demand models that are not tractable for mathematical analysis, but that can be handled using appropriate machine learning techniques. In spite of these different aims, there is much in overlap in the studied demand models, proposed pricing policies, and the techniques deployed to analyze the behavior of these policies. With this survey we aim to provide a comprehensive overview of these studies, and facilitate further research on this broad and lively research topic.

## Acknowledgments

Part of this work was written while the author was affiliated with Centrum Wiskunde & Informatica (CWI), Eindhoven University of Technology, and University of Amsterdam. The author is supported by an NWO Veni grant. This survey is a considerable extension of the literature review in chapter 2 of [den Boer \(2013\)](#). The careful reading and constructive suggestions of the editor and the referees have improved the exposition of the paper, and is kindly acknowledged.

## References

- Y. Abbasi-Yadkori, D. Pal, and C. Szepesvári. Improved algorithms for linear stochastic bandits. In J. Shawe-Taylor, R. S. Zemel, P. Bartlett, F. C. N. Pereira, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 24*, pages 2312–2320, 2011.
- E. Adida and G. Perakis. A robust optimization approach to dynamic pricing and inventory control with no backorders. *Mathematical Programming, Series B*, 107(1):97–129, 2006.
- E. Adida and G. Perakis. Dynamic pricing and inventory control: robust vs. stochastic uncertainty models - a computational study. *Annals of Operations Research*, 181(1):125–157, 2010a.
- E. Adida and G. Perakis. Dynamic pricing and inventory control: uncertainty and competition. *Operations Research*, 58(2):289–302, 2010b.
- P. Afèche and B. Ata. Bayesian dynamic pricing in queueing systems with unknown delay cost characteristics. *Manufacturing & Service Operations Management*, 15(2):292–304, 2013.
- P. Aghion, P. Bolton, C. Harris, and B. Jullien. Optimal learning by experimentation. *Review of Economic Studies*, 58(4):621–654, 1991.
- P. Aghion, M. Paz Espinosa, and B. Jullien. Dynamic duopoly with learning through market experimentation. *Economic Theory*, 3(3):517–539, 1993.
- H. S. Ahn, M. Gümüş, and P. Kaminsky. Pricing and manufacturing decisions when demand is a function of prices in multiple periods. *Operations Research*, 55(6):1039–1057, 2007.

- Y. Akçay, H. P. Natarajan, and S. H. Xu. Joint dynamic pricing of multiple perishable products under consumer choice. *Management Science*, 56(8):1345–1361, 2010.
- M. D. Alepuz and A. Urbano. Duopoly experimentation: Cournot competition. *Mathematical Social Sciences*, 37(2):165–188, 1999.
- T. W. Anderson and J. B. Taylor. Some experimental results on the statistical properties of least squares estimates in control problems. *Econometrica*, 44(6):1289–1302, 1976.
- M. Anufriev, D. Kopányi, and J. Tuinstra. Learning cycles in Bertrand competition with differentiated commodities and competing learning rules. *Journal of Economic Dynamics and Control*, 37(12):2562–2581, 2013.
- M. Aoki. On a dual control approach to the pricing policies of a trading specialist. In R. Conti and A. Ruberti, editors, *5th Conference on Optimization Techniques Part II*, volume 4 of *Lecture Notes in Computer Science*, chapter 24, pages 272–282. Springer, Berlin, Heidelberg, 1973.
- M. Aoki. On some price adjustment schemes. *Annals of Economic and Social Measurement*, 3(1):95–115, 1974.
- V. F. Araman and R. Caldentey. Dynamic pricing for nonperishable products with demand learning. *Operations Research*, 57(5):1169–1188, 2009.
- V. F. Araman and R. Caldentey. Revenue management with incomplete demand information. In J. J. Cochran, editor, *Encyclopedia of Operations Research*. Wiley, 2011.
- A. Arasteh, A. Aliahmadi, and A. Norang. Inventory policies and dynamic pricing under possibility and rivals. *The International Journal of Advanced Manufacturing Technology*, 72(5-8):627–642, 2014.
- P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multi-armed bandit problem. *Machine Learning*, 47(2):235–256, 2002.
- P. Auer, R. Ortner, and C. Szepesvári. Improved rates for the stochastic continuum-armed bandit problem. In N. Bshouty and C. Gentile, editors, *Learning Theory*, volume 4539 of *Lecture Notes in Computer Science*, pages 454–468. Springer, Berlin, Heidelberg, 2007.
- Y. Aviv and A. Pazgal. A partially observed Markov decision process for dynamic pricing. *Management Science*, 51(9):1400–1416, 2005a.
- Y. Aviv and A. Pazgal. Pricing of short life-cycle products through active learning. Working paper, Olin School of Business, Washington University, St. Louis. Available at [https://www.researchgate.net/publication/228674522\\_Dynamic\\_pricing\\_of\\_short\\_life-cycle\\_products\\_through\\_active\\_learning](https://www.researchgate.net/publication/228674522_Dynamic_pricing_of_short_life-cycle_products_through_active_learning), 2005b.
- Y. Aviv and A. Pazgal. Optimal pricing of seasonal products in the presence of forward-looking consumers. *Manufacturing & Service Operations Management*, 10(3):339–359, 2008.
- Y. Aviv and G. Vulcano. Dynamic list pricing. In Ö. Özer and R. Phillips, editors, *The Oxford Handbook of Pricing Management*, chapter 23. Oxford University Press, London, 2012.
- A. N. Avramidis. Learning in revenue management: Exploiting estimation of arrival rate and price response. Working paper, School of Mathematics, University of Southampton. Available at <http://www.personal.soton.ac.uk/aa1w07/v35.pdf>, 2013.
- K. S. Azoury. Bayes solution to dynamic inventory models under unknown demand distribution. *Management Science*, 31(9):1150–1160, 1985.
- M. Babaioff, L. Blumrosen, S. Dughmi, and Y. Singer. Posting prices with unknown distributions. In B. Chazelle, editor, *Innovations in Computer Science - ICS 2010, Tsinghua University, Beijing, China, January 7-9, 2011. Proceedings*, pages 166–178. Tsinghua University Press, 2011.
- M. Babaioff, S. Dughmi, R. Kleinberg, and A. Slivkins. Dynamic pricing with limited supply. *ACM Transactions on Economics and Computation*, Forthcoming, 2014.

- J. Baetge, G. Bolenz, W. Ballwieser, R. Hömberg, and P. Wullers. Dynamic price policies in monopolistic competition. In J. Rose and C. Bilciu, editors, *Modern Trends in Cybernetics and Systems. Proceedings of the Third International Congress of Cybernetics and Systems, Bucharest, Romania, August 25-29, 1975*, volume 1, pages 401–417. Springer Verlag, 1977.
- P. Bajari and C. L. Benkard. Demand estimation with heterogeneous consumers and unobserved product characteristics: a hedonic approach. *Journal of Political Economy*, 113(6):1239–1276, 2005.
- M. F. Balcan, A. Blum, and Y. Mansour. Item pricing for revenue maximization. In *Proceedings of the 9th ACM Conference on Electronic Commerce, EC '08*, pages 50–59, New York, NY, USA, 2008. ACM.
- R. J. Balvers and T. F. Cosimano. Actively learning about demand and the dynamics of price adjustment. *The Economic Journal*, 100(402):882–898, 1990.
- R. Baptista. The diffusion of process innovations: a selective review. *International Journal of the Economics of Business*, 6(1):107–129, 1999.
- D. P. Baron. Price uncertainty, utility, and industry equilibrium in pure competition. *International Economic Review*, 11(3):463–480, 1970.
- D. P. Baron. Demand uncertainty in imperfect competition. *International Economic Review*, 12(2):196–208, 1971.
- S. Barta and P. Varaiva. Microeconomics: stochastic models of price adjustment. *Annals of Economic and Social Measurement*, 5(3):267–281, 1976.
- F. M. Bass. A new product growth for model consumer durables. *Management Science*, 15(5):215–227, 1969.
- W. J. Baumol and R. E. Quandt. Rules of thumb and optimally imperfect decisions. *The American Economic Review*, 54(2):23–46, 1964.
- G. W. Beck and V. Wieland. Learning and control in a changing economic environment. *Journal of Economic Dynamics and Control*, 26(9-10):1359–1377, 2002.
- P. Belleflamme and F. Bloch. Price and quantity experimentation: a synthesis. *International Journal of Industrial Organization*, 19(10):1563–1582, 2001.
- R. Benini. Sull’uso delle formole empiriche a nell’economia applicata. *Giornale degli economisti*, 2nd series, 35:1053–1063, 1907.
- K. Berg and H. Ehtamo. Continuous learning methods in two-buyer pricing problem. *Mathematical Methods of Operations Research*, 75(1):287–304, 2012.
- D. Bergemann and K. H. Schlag. Pricing without priors. *Journal of the European Economic Association*, 6(2-3):560–569, 2008.
- D. Bergemann and K. H. Schlag. Robust monopoly pricing. *Journal of Economic Theory*, 146(6):2527–2543, 2011.
- D. Bergemann and J. Valimäki. Market experimentation and pricing. Cowles Foundation Discussion Paper 1122, Cowles Foundation for Research in Economics at Yale University, <http://cowles.econ.yale.edu/P/cd/d11a/d1122.pdf>, 1996.
- D. Bergemann and J. Välimäki. Market diffusion with two-sided learning. *The RAND Journal of Economics*, 28(4):773–795, 1997.
- S. Berry, J. Levinsohn, and A. Pakes. Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890, 1995.
- D. Bertsimas and G. Perakis. Dynamic pricing: a learning approach. In *Mathematical and Computational Models for Congestion Charging*, pages 45–79. Springer, New York, 2006.

- O. Besbes and C. Maglaras. Dynamic pricing with financial milestones: feedback-form policies. *Management Science*, 58(9):1715–1731, 2012.
- O. Besbes and D. Sauré. Dynamic pricing strategies in the presence of demand shocks. *Manufacturing & Service Operations Management*, Forthcoming, 2014.
- O. Besbes and A. Zeevi. Dynamic pricing without knowing the demand function: risk bounds and near-optimal algorithms. *Operations Research*, 57(6):1407–1420, 2009.
- O. Besbes and A. Zeevi. On the minimax complexity of pricing in a changing environment. *Operations Research*, 59(1):66–79, 2011.
- O. Besbes and A. Zeevi. Blind network revenue management. *Operations Research*, 60(6):1537–1550, 2012.
- O. Besbes and A. Zeevi. On the (surprising) sufficiency of linear models for dynamic pricing with demand learning. *Management Science*, Forthcoming, 2014.
- O. Besbes, Y. Gur, and A. Zeevi. Non-stationary stochastic optimization. Working paper, Columbia Business School, <http://ssrn.com/abstract=2296012>, 2014.
- M. Bhalla. Dynamic pricing under social learning with strategic consumers. Working paper No. 367, Indian Institute of Management, Bangalore. <http://ssrn.com/abstract=2114054>, 2012.
- F. Billström and S. A. O. Thore. Simulation experiments with dynamic price strategies in monopoly theory. In H. O. A. Wold, editor, *Econometric Model Building. Essays on the Causal Chain Approach*, pages 297–321. North Holland Publishing Co., Amsterdam, 1964.
- F. Billström, H. Laadi, S. A. O. Thore, L. O. Friberg, and O. Johansson. Some experiments with dynamization of monopoly models. Mimeo. 16th Meeting of the Econometric Society in Uppsala, 1954.
- G. I. Bischi, C. Chiarella, and M. Kopel. The long run outcomes and global dynamics of a duopoly game with misspecified demand functions. *International Game Theory Review*, 06(03):343–379, 2004.
- G. I. Bischi, A. K. Naimzada, and L. Sbragia. Oligopoly games with local monopolistic approximation. *Journal of Economic Behavior & Organization*, 62(3):371–388, 2007.
- A. Bisi and M. Dada. Dynamic learning, pricing, and ordering by a censored newsvendor. *Naval Research Logistics*, 54(4):448–461, 2007.
- G. Bitran and R. Caldentey. An overview of pricing models for revenue management. *Manufacturing & Service Operations Management*, 5(3):203–230, 2003.
- G. Bitran, R. Caldentey, and S. Mondschein. Coordinating clearance markdown sales of seasonal products in retail chains. *Operations Research*, 46(5):609–624, 1998.
- G. R. Bitran and S. V. Mondschein. Periodic pricing of seasonal products in retailing. *Management Science*, 43(1):64–79, 1997.
- G. R. Bitran and H. K. Wadhwa. A methodology for demand learning with an application to the optimal pricing of seasonal products. Working paper, WP #3898-96, MIT Sloan School of Management. Available at <http://dspace.mit.edu/handle/1721.1/2619>, 1996.
- A. Borovkov. *Mathematical Statistics*. Gordon and Breach Science Publishers, Amsterdam, 1998.
- S. Bose, G. Orosel, M. Ottaviani, and L. Vesterlund. Dynamic monopoly pricing and herding. *The RAND Journal of Economics*, 37(4):910–928, 2006.
- T. Boyacı and Ö. Özer. Information acquisition for capacity planning via pricing and advance selling: when to stop and act? *Operations Research*, 58(5):1328–1349, 2010.
- J. Broder and P. Rusmevichientong. Dynamic pricing under a general parametric choice model. *Operations Research*, 60(4):965–980, 2012.



- C. H. Brooks, S. Fay, R. Das, J. K. MacKie-Mason, J. O. Kephart, and E. H. Durfee. Automated strategy searches in an electronic goods market: learning and complex price schedules. In *Proceedings of the 1st ACM conference on Electronic commerce*, pages 31–40. ACM, 1999.
- C. H. Brooks, R. Das, J. O. Kephart, J. K. MacKie-Mason, and R. S. Gazzale. Information bundling in a dynamic environment. In *Proceedings of the IJCAI-01 Workshop on Economic Agents, Models, and Mechanisms, Seattle, WA*, volume 205, 2001.
- C. H. Brooks, R. S. Gazzale, R. Das, J. O. Kephart, J. K. MacKie-Mason, and E. H. Durfee. Model selection in an information economy: choosing what to learn. *Computational Intelligence*, 18(4):566–582, 2002.
- V. Brousseau and A. Kirman. Apparent convergence of learning processes in mis-specified games. In B. Dutta, D. Mookherjee, T. Parthasarathy, T. E. S. Raghavan, D. Ray, and S. Tijs, editors, *Game Theory and Economic Applications*, pages 303–331. Springer-Verlag, Berlin, 1992.
- A. Brown and A. Deaton. Surveys in applied economics: models of consumer behaviour. *The Economic Journal*, 82(328):1145–1236, 1972.
- A. N. Burnetas and C. E. Smith. Adaptive ordering and pricing for perishable products. *Operations Research*, 48(3):436–443, 2000.
- G. P. Cachon and R. Swinney. Purchasing, pricing, and quick response in the presence of strategic consumers. *Management Science*, 55(3):497–511, 2009.
- R. Caldentey, Y. Liu, and I. Lobel. Intertemporal pricing without priors. Working paper, New York University (NYU), <http://ssrn.com/abstract=2357083>, 2013.
- R. Caminal and X. Vives. Price dynamics and consumer learning. *Journal of Economics & Management Strategy*, 8(1):95–131, 1999.
- F. Caro and J. Gallien. Dynamic assortment with demand learning for seasonal consumer goods. *Management Science*, 53(2):276–292, 2007.
- A. X. Carvalho and M. L. Puterman. Learning and pricing in an internet environment with binomial demand. *Journal of Revenue and Pricing Management*, 3(4):320–336, 2005a.
- A. X. Carvalho and M. L. Puterman. Dynamic optimization and learning: how should a manager set prices when the demand function is unknown? Technical report, Instituto de Pesquisa Economica Aplicada - IPEA, Brasilia, Discussion Papers 1117. Available at [http://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td\\_1117.pdf](http://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td_1117.pdf), 2005b.
- N. Cesa-Bianchi and G. Lugosi. *Prediction, learning, and games*. Cambridge University Press, New York, 2006.
- H. Chade and V. V. de Serio. Pricing, learning, and strategic behavior in a single-sale model. *Economic Theory*, 19(2):333–353, 2002.
- T. Chakraborty, Z. Huang, and S. Khanna. Dynamic and nonuniform pricing strategies for revenue maximization. *SIAM Journal on Computing*, 42(6):2424–2451, 2013.
- S. Chakravarty, S. Padakandla, and S. Bhatnagar. A simulation-based algorithm for optimal pricing policy under demand uncertainty. *International Transactions in Operational Research*, 21(5):737–760, 2014.
- L. Chan, Z. Shen, and D. Simchi-Levi. Coordination of pricing and inventory decisions: a survey and classification. In D. Simchi-Levi, D. Wu, and Z. Shen, editors, *Handbook of Quantitative Supply Chain Analysis: Modeling in the E-Business Era*, pages 335–392. Springer, New York, 2004.
- L. Chen. Bayesian dynamic pricing with two-sided censored customer willingness-to-pay data. Working paper, The Fuqua School of Business, Duke University. Available at <https://faculty.fuqua.duke.edu/~lc91/More/>, 2014.



- M. Chen and Z. L. Chen. Recent developments in dynamic pricing research: multiple products, competition, and limited demand information. *Production and Operations Management*, Online pre-publication, 2014.
- X. Chen and D. Simchi-Levi. Joint pricing and inventory management. In Ö. Özer and R. Phillips, editors, *The Oxford Handbook of Pricing Management*, chapter 30. Oxford University Press, London, 2012.
- Y. Chen and V. F. Farias. Simple policies for dynamic pricing with imperfect forecasts. *Operations Research*, 61(3):612–624, 2013.
- Y. Chen and R. Wang. Learning buyers’ valuation distribution in posted-price selling. *Economic Theory*, 14(2):417–428, 1999.
- Y. M. Chen and D. C. Jain. Dynamic monopoly pricing under a Poisson-type uncertain demand. *The Journal of Business*, 65(4):593–614, 1992.
- R. Chenavaz, L. P. Carrier, L. Etienne, and C. Paraschiv. Dynamic pricing in management science. *Journal of Economics Studies and Research*, 2011, 2011.
- Y. Cheng. Dynamic pricing decision for perishable goods: a Q-learning approach. In *4th International Conference on Wireless Communications, Networking and Mobile Computing, 2008. WiCOM’08.*, pages 1–5. IEEE, 2008.
- Y. Cheng. Real time demand learning-based Q-learning approach for dynamic pricing in E-retailing setting. In *International Symposium on Information Engineering and Electronic Commerce IEEEC ’09*, pages 594–598, 2009.
- W. C. Cheung, D. Simchi-Levi, and H. Wang. Dynamic pricing and demand learning with limited price experimentation. Working paper, MIT. Available at <http://web.mit.edu/wanghe/www/research.html>, 2013.
- M. Chhabra and S. Das. Learning the demand curve in posted-price digital goods auctions. In *The 10th International Conference on Autonomous Agents and Multiagent Systems - Volume 1, AAMAS ’11*, pages 63–70, Richland, SC, 2011. International Foundation for Autonomous Agents and Multiagent Systems.
- A. T. Ching, T. Erdem, and M. P. Keane. Invited paper - learning models: an assessment of progress, challenges, and new developments. *Marketing Science*, 32(6):913–938, 2013.
- V. L. R. Chinthalapati, N. Yadati, and R. Karumanchi. Learning dynamic prices in MultiSeller electronic retail markets with price sensitive customers, stochastic demands, and inventory replenishments. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 36(1):92–106, 2006.
- S. Chan Choi and S. Jagpal. Duopoly pricing under risk aversion and parameter uncertainty. *Journal of Product & Brand Management*, 13(5):359–368, 2004.
- T. M. Choi. Pre-season stocking and pricing decisions for fashion retailers with multiple information updating. *International Journal of Production Economics*, 106(1):146–170, 2007.
- T. M. Choi, P. S. Chow, and T. Xiao. Electronic price-testing scheme for fashion retailing with information updating. *International Journal of Production Economics*, 140(1):396–406, 2012.
- C. Y. Chong and D. Cheng. Multistage pricing under uncertain demand. *Annals of Economic and Social Measurement*, 4(2):311–323, 1975.
- C. F. Christ. Early progress in estimating quantitative economic relationships in America. *The American Economic Review*, 75(6):39–52, 1985.
- S. Christ. *Operationalizing Dynamic Pricing Models: Bayesian Demand Forecasting and Customer Choice Modeling for Low Cost Carriers*. PhD thesis, Augsburg University, 2011.
- B. D. Chung, J. Li, T. Yao, C. Kwon, and T. L. Friesz. Demand learning and dynamic pricing under competition in a state-space framework. *Engineering Management, IEEE Transactions on*, 59(2):240–249, 2012.

- D. F. Ciocan and V. Farias. Model predictive control for dynamic resource allocation. *Mathematics of Operations Research*, 37(3):501–525, 2012.
- D. G. Clarke and R. J. Dolan. A simulation analysis of alternative pricing strategies for dynamic environments. *The Journal of Business*, 57(1):S179–S200, 1984.
- F. Clarke, M. N. Darrough, and J. Heineke. Optimal pricing policy in the presence of experience effects. *The Journal of Business*, 55(4):517–530, 1982.
- R. W. Clower. Some theory of an ignorant monopolist. *The Economic Journal*, 69(276):705–716, 1959.
- Ö. Coşgun, Y. Ekinici, and S. Yanık. Fuzzy rule-based demand forecasting for dynamic pricing of a maritime company. *Knowledge-Based Systems*, 70:88–96, 2014.
- M. Cohen, R. Lobel, and G. Perakis. Dynamic pricing through data sampling. Working paper, Massachusetts Institute of Technology. Available at <http://ssrn.com/abstract=2376667>, 2013.
- A. Collins and L. Thomas. Comparing reinforcement learning approaches for solving game theoretic models: a dynamic airline pricing game example. *Journal of the Operational Research Society*, 63(1):1165–1173, 2012.
- A. Collins and L. Thomas. Learning competitive dynamic airline pricing under different customer models. *Journal of Revenue and Pricing Management*, 12(5):416–430, 2013.
- R. Combes and A. Proutiere. Unimodal bandits: Regret lower bounds and optimal algorithms. In *Proceedings of the 31th International Conference on Machine Learning, ICML 2014, Beijing, China, 21-26 June 2014*, pages 521–529, 2014.
- W. L. Cooper, T. Homem de Mello, and A. J. Kleywegt. Models of the Spiral-Down effect in revenue management. *Operations Research*, 54(5):968–987, 2006.
- W. L. Cooper, T. Homem de Mello, and A. J. Kleywegt. Learning and pricing with models that do not explicitly incorporate competition. Working paper, University of Minnesota. Available at <http://www.menet.umn.edu/~billcoop/competitive2014.pdf>, 2014.
- E. Cope. Bayesian strategies for dynamic pricing in e-commerce. *Naval Research Logistics*, 54(3):265–281, 2007.
- E. W. Cope. Regret and convergence bounds for a class of continuum-armed bandit problems. *IEEE Transactions on Automatic Control*, 54(6):1243–1253, 2009.
- A. T. Coughlan and M. K. Mantrala. Dynamic competitive retail pricing behavior with uncertainty and learning. *Managerial and Decision Economics*, 15(1):3–20, 1994.
- A. A. Cournot. *Researches into the Mathematical Principles of the Theory of Wealth*. Translated in 1897 by N.T. Bacon, with a bibliography of Mathematical Economics by Irving Fisher. The Macmillan Company, New York, 1838.
- J. Creedy. On the King-Davenant "law" of demand. *Scottish Journal of Political Economy*, 33(3):193–212, 1986.
- R. M. Cyert and M. H. DeGroot. Bayesian analysis and duopoly theory. *Journal of Political Economy*, 78(5):1168–1184, 1970.
- P. Dasgupta and R. Das. Dynamic pricing with limited competitor information in a multi-agent economy. In P. Scheuermann and O. Etzion, editors, *Cooperative Information Systems*, volume 1901 of *Lecture Notes in Computer Science*, pages 299–310. Springer, Berlin, Heidelberg, 2000.
- C. Davenant. *An essay upon the probable methods of making a people gainers in the balance of trade*. James Knapton, London, 1699.
- A. de Palma, R. Lindsey, and S. Proost. Research challenges in modelling urban road pricing: an overview. *Transport Policy*, 13(2):97–105, 2006.

- I. Deksnys and Z. Lydeka. Dynamic pricing and its forming factors. *International Journal of Business and Social Science*, 3(23):213–220, 2012.
- A. V. den Boer. *Dynamic Pricing and Learning*. PhD thesis, Vrije Universiteit Amsterdam, 2013.
- A. V. den Boer. Dynamic pricing with multiple products and partially specified demand distribution. *Mathematics of Operations Research*, 39(3):863–888, 2014a.
- A. V. den Boer. Tracking the market: dynamic pricing and learning in a changing environment. Working paper, University of Twente. Available at <http://www.utwente.nl/ewi/sor/about/staff/boer/>, 2014b.
- A. V. den Boer and B. Zwart. Simultaneously learning and optimizing using controlled variance pricing. *Management Science*, 60(3):770–783, 2014a.
- A. V. den Boer and B. Zwart. Mean square convergence rates for maximum quasi-likelihood estimators. *Stochastic Systems*, 4:1–29, 2014b.
- A. V. den Boer and B. Zwart. Dynamic pricing and learning with finite inventories. Working paper, University of Twente. Available at <http://www.utwente.nl/ewi/sor/about/staff/boer/>, 2014c.
- J. M. DiMicco, P. Maes, and A. Greenwald. Learning curve: a simulation-based approach to dynamic pricing. *Electronic Commerce Research*, 3(3-4):245–276, 2003.
- M. Dimitrova and E. E. Schlee. Monopoly, competition and information acquisition. *International Journal of Industrial Organization*, 21(10):1623–1642, 2003.
- I. Dogan and A. R. Güner. A reinforcement learning approach to competitive ordering and pricing problem. *Expert Systems*, 37(16-17):8141–8161, 2013.
- L. Dong, P. Kouvelis, and Z. Tian. Dynamic pricing and inventory control of substitute products. *Manufacturing & Service Operations Management*, 11(2):317–339, 2009.
- M. Dziecichowicz, D. Caro, and A. Thiele. Robust timing of markdowns. Technical report (2010), Lehigh University, Department of Industrial and Systems Engineering, Bethlehem, PA. [http://www.optimization-online.org/DB\\_HTML/2010/11/2804.html](http://www.optimization-online.org/DB_HTML/2010/11/2804.html), 2010.
- D. Easley and N. M. Kiefer. Controlling a stochastic process with unknown parameters. *Econometrica*, 56(5):1045–1064, 1988.
- J. Eliashberg and R. Steinberg. Marketing-production joint decision-making. In J. Eliashberg and J. D. Lilien, editors, *Management Science in Marketing, Handbooks in Operations Research and Management Science*, volume 5, pages 827–880. North Holland Publishing Co., Amsterdam, 1993.
- W. Elmaghraby and P. Keskinocak. Dynamic pricing in the presence of inventory considerations: research overview, current practices, and future directions. *Management Science*, 49(10):1287–1309, 2003.
- W. Elmaghraby, A. Gülcü, and P. Keskinocak. Designing optimal pre-announced markdowns in the presence of rational customers with multi-unit demands. *Manufacturing & Service Operations Management*, 10(1):126–148, 2008.
- A. Erdelyi and H. Topaloglu. Using decomposition methods to solve pricing problems in network revenue management. *Journal of Revenue and Pricing Management*, 10(4):325–343, 2011.
- S. S. Eren and C. Maglaras. Monopoly pricing with limited demand information. *Journal of Revenue and Pricing Management*, 9:23–48, 2010.
- D. Escobari. Dynamic pricing, advance sales and aggregate demand learning in airlines. *The Journal of Industrial Economics*, 60(4):697–724, 2012.
- G. C. Evans. The dynamics of monopoly. *The American Mathematical Monthly*, 31(2):77–83, 1924.
- G. C. Evans. The mathematical theory of economics. *The American Mathematical Monthly*, 32(3):104–110, 1925.

- R. W. Farebrother. Early explorations in econometrics. In T. C. Mills and K. Patterson, editors, *Palgrave Handbook of Econometrics, Volume 1: Econometric Theory*. Palgrave MacMillan, Basingstoke, 2006.
- V. F. Farias and B. van Roy. Dynamic pricing with a prior on market response. *Operations Research*, 58(1):16–29, 2010.
- A. Federgruen and A. Heching. Combined pricing and inventory control under uncertainty. *Operations Research*, 47(3):454–475, 1999.
- Y. Feng and G. Gallego. Optimal starting times for end-of-season sales and optimal stopping times for promotional fares. *Management Science*, 41(8):1371–1391, 1995.
- Y. Feng and G. Gallego. Perishable asset revenue management with Markovian time dependent demand intensities. *Management Science*, 46(7):941–956, 2000.
- Y. Feng and B. Xiao. Optimal policies of yield management with multiple predetermined prices. *Operations Research*, 48(2):332–343, 2000a.
- Y. Feng and B. Xiao. A continuous-time yield management model with multiple prices and reversible price changes. *Management Science*, 46(5):644–657, 2000b.
- J. C. Ferrer, D. Oyarzún, and J. Vera. Risk averse retail pricing with robust demand forecasting. *International Journal of Production Economics*, 136(1):151–160, 2012.
- G. Fibich, A. Gavius, and O. Lowengart. Explicit solutions of optimization models and differential games with nonsmooth (asymmetric) reference-price effects. *Operations Research*, 51(5):721–734, 2003.
- S. Filippi, O. Cappe, A. Garivier, and C. Szepesvari. Parametric bandits: the generalized linear case. In J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, and A. Culotta, editors, *Advances in Neural Information Processing Systems 23*, pages 586–594. Curran Associates, Inc., 2010.
- I. Fisher. Cournot and mathematical economics. *The Quarterly Journal of Economics*, 12(2):119–138, 1898.
- A. Fishman and N. Gandal. Experimentation and learning with networks effects. *Economics Letters*, 44(1-2):103–108, 1994.
- K. Forghani, A. Mirzazadeh, and M. Rafiee. A price-dependent demand model in the single period inventory system with price adjustment. *Journal of Industrial Engineering*, 2013:1–9, 2013.
- T. L. Friesz, C. Kwon, T. I. Kim, L. Fan, and T. Yao. Competitive robust dynamic pricing in continuous time with fixed inventories. arXiv:1208.4374 [math.OC], 2012.
- A. G. Gallego. Oligopoly experimentation of learning with simulated markets. *Journal of Economic Behavior & Organization*, 35(3):333–355, 1998.
- G. Gallego and M. Talebian. Demand learning and dynamic pricing for multi-version products. *Journal of Revenue and Pricing Management*, 11:303–318, 2012.
- G. Gallego and G. van Ryzin. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science*, 40(8):999–1020, 1994.
- G. Gallego and G. van Ryzin. A multiproduct dynamic pricing problem and its applications to network yield management. *Operations Research*, 45(1):24–41, 1997.
- G. Gallego and R. Wang. Multiproduct price optimization and competition under the nested logit model with product-differentiated price sensitivities. *Operations Research*, 62(2):450–461, 2014.
- J.-j. Gao, Y. Zheng, and X. Mao. A joint decision model of dynamic pricing and dynamic inventory in an apparel supply chain with demand learning. In *Logistics Systems and Intelligent Management, 2010 International Conference on*, pages 692–696. IEEE, 2010.
- A. Garcia, E. Campos-Nañez, and J. Reitzes. Dynamic pricing and learning in electricity markets. *Operations Research*, 53(2):231–241, 2005.

- A. Garivier and E. Moulines. On upper-confidence bound policies for switching bandit problems algorithmic learning theory. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory. 22nd International Conference, ALT 2011, Espoo, Finland, October 5-7, 2011*, volume 6925 of *Lecture Notes in Computer Science*, chapter 16, pages 174–188. Springer, Berlin, Heidelberg, 2011.
- W. Gaul and A. D. Azizi. A demand learning data based approach to optimize revenues of a retail chain. In H. Locarek-Junge and C. Weihs, editors, *Classification as a Tool for Research*, Studies in Classification, Data Analysis, and Knowledge Organization, chapter 75, pages 683–691. Springer, Berlin, Heidelberg, 2010.
- T. K. Ghose and T. T. Tran. Dynamic pricing in electronic commerce using neural network. In G. Babin, P. Kropf, and M. Weiss, editors, *E-Technologies: Innovation in an Open World*, volume 26, pages 227–232. Springer, Berlin, Heidelberg, 2009.
- C. Gini. Prezzi e consumi. *Giornale degli Economisti, 3rd series*, 40:99–114, 235–249, 1910.
- J. C. Gittins. *Multi-armed bandit allocation indices*. Wiley-Interscience series in Systems and Optimization. Wiley, New York, 1989.
- A. Goldenshluger and A. Zeevi. Woodrooffe’s one-armed bandit problem revisited. *The Annals of Applied Probability*, 19(4):1603–1633, 2009.
- J. Gönsch, R. Klein, and C. Steinhardt. Dynamic pricing - State-of-the-Art. *Zeitschrift für Betriebswirtschaft, Ergänzungsheft 3 'Operations Research in der Betriebswirtschaft'*, pages 1–40, 2009.
- J. Gönsch, R. Klein, M. Neugebauer, and C. Steinhardt. Dynamic pricing with strategic customers. *Journal of Business Economics*, 83(5):505–549, 2013.
- E. A. Greenleaf. The impact of reference price effects on the profitability of price promotions. *Marketing Science*, 14(1):82–104, 1995.
- A. R. Greenwald and J. O. Kephart. Shopbots and pricebots. In A. Moukas, C. Sierra, and F. Ygge, editors, *Agent Mediated Electronic Commerce II*, pages 1–23. Springer, 1999.
- S. J. Grossman, R. E. Kihlstrom, and L. J. Mirman. A Bayesian approach to the production of information and learning by doing. *The Review of Economic Studies*, 44(3):533–547, 1977.
- J. Han, T. L. Lai, and V. Spivakovsky. Approximate policy optimization and adaptive control in regression models. *Computational Economics*, 27(4):433–452, 2006.
- W. Han. A dynamic pricing algorithm by Bayesian Q-learning. In *Computer Modeling and Simulation, 2010. ICCMS '10. Second International Conference on*, volume 2, pages 515–519, 2010.
- W. Han, L. Liu, and H. Zheng. Dynamic pricing by multiagent reinforcement learning. In *Electronic Commerce and Security, 2008 International Symposium on*, pages 226–229. IEEE, 2008.
- B. Handel and K. Misra. Robust new product pricing. Working paper, University of California, Berkeley. Available at <http://ssrn.com/abstract=2307991>, 2014.
- B. R. Handel, K. Misra, and J. W. Roberts. Robust firm pricing with panel data. *Journal of Econometrics*, 174(2):165–185, 2013.
- B. Hansen. Report of the Uppsala meeting, August 2-4, 1954. *Econometrica*, 23(2):198–216, 1955.
- J. E. Harrington. Experimentation and learning in a differentiated-products duopoly. *Journal of Economic Theory*, 66(1):275–288, 1995.
- M. Harris and A. Raviv. A theory of monopoly pricing schemes with demand uncertainty. *The American Economic Review*, 71(3):347–365, 1981.
- J. M. Harrison, N. B. Keskin, and A. Zeevi. Bayesian dynamic pricing policies: learning and earning under a binary prior distribution. *Management Science*, 58(3):570–586, 2012.

- M. Haviv and R. S. Randhawa. Pricing in queues without demand information. *Manufacturing & Service Operations Management*, 16(3):401–411, 2014.
- E. R. Hawkins. Methods of estimating demand. *Journal of Marketing*, 21(4):428–438, 1957.
- A. Heching and Y. T. Leung. Product pricing in the e-business era. In C. An and H. Fromm, editors, *Supply Chain Management on Demand*, chapter 4, pages 65–96. Springer, 2005.
- P. Heidhues and B. Köszegi. Regular prices and sales. *Theoretical Economics*, to appear, 2012.
- J. R. Hicks. Annual survey of economic theory: the theory of monopoly. *Econometrica*, 3(1):1–20, 1935.
- G. J. Hitsch. An empirical model of optimal dynamic product launch and exit under demand uncertainty. *Marketing Science*, 25(1):25–50, 2006.
- S. L. Horner, C. F. Roos, V. von Szeliski, A. T. Court, and S. M. DuBrul. *The Dynamics of Automobile Demand*. General Motors Corporation, New York, 1939.
- G. Huang, H. Luo, and J. Xia. Invest in information or wing it? A model of dynamic pricing with seller learning. Working paper, Yale School of Management. Available at [https://server1.tepper.cmu.edu/seminars/docs/Infovalue\\_Huang.pdf](https://server1.tepper.cmu.edu/seminars/docs/Infovalue_Huang.pdf), 2012.
- W. T. Huh, R. Levi, P. Rusmevichientong, and J. B. Orlin. Adaptive data-driven inventory control with censored demand based on Kaplan-Meier estimator. *Operations Research*, 59(4):929–941, 2011.
- B. Ifrach, C. Maglaras, and M. Scarsini. Monopoly pricing in the presence of social learning. NET Institute Working Paper No. 12-01, <http://ssrn.com/abstract=1957924>, 2012.
- B. Ifrach, C. Maglaras, and M. Scarsini. Bayesian social learning from consumer reviews. Working paper. Available at <http://ssrn.com/abstract=2293158>, 2013.
- K. Isler and H. Imhof. A game theoretic model for airline revenue management and competitive pricing. *Journal of Revenue and Pricing Management*, 7(4):384–396, 2008.
- S. Jasin. Reoptimization and self-adjusting price control for network revenue management. *Operations Research*, 62(5):1168–1178, 2014.
- L. Jia, Q. Zhao, and L. Tong. Retail pricing for stochastic demand with unknown parameters: an online machine learning approach. In *2013 Allerton Conference on Communication, Control and Computing, Monticello, IL*, 2013.
- W. Jintian and Z. Lei. Application of reinforcement learning in dynamic pricing algorithms. In *Automation and Logistics, 2009. ICAL '09. IEEE International Conference on*, pages 419–423. IEEE, 2009.
- J. Jumadinova and P. Dasgupta. Firefly-inspired synchronization for improved dynamic pricing in online markets. In *Self-Adaptive and Self-Organizing Systems, 2008. SASO '08. Second IEEE International Conference on*, pages 403–412. IEEE, 2008.
- J. Jumadinova and P. Dasgupta. Multi-attribute regret-based dynamic pricing. In W. Ketter, H. La Poutré, N. Sadeh, O. Shehory, and W. Walsh, editors, *Agent-Mediated Electronic Commerce and Trading Agent Design and Analysis*, volume 44 of *Lecture Notes in Business Information Processing*, pages 73–87. Springer Berlin Heidelberg, 2010.
- S. Kachani, G. Perakis, and C. Simon. Modeling the transient nature of dynamic pricing with demand learning in a competitive environment. In T. L. Friesz, editor, *Network Science, Nonlinear Science and Infrastructure Systems*, volume 102 of *International Series in Operations Research & Management Science*, pages 223–267. Springer Science+Business Media, New York, 2007.
- S. Kalish. Monopolist pricing with dynamic demand and production cost. *Marketing Science*, 2(2):135–159, 1983.
- B. Kamrad, S. S. Lele, A. Siddique, and R. J. Thomas. Innovation diffusion uncertainty, advertising and pricing policies. *European Journal of Operational Research*, 164(3):829–850, 2005.

- S. Karlin and C. R. Carr. Prices and optimal inventory policies. In K. Arrow, S. Karlin, and H. Scarf, editors, *Studies in Applied Probability and Management Science*, pages 159–172. Stanford University Press, Stanford, 1962.
- G. Keller and S. Rady. Optimal experimentation in a changing environment. *The Review of Economic Studies*, 66(3):475–507, 1999.
- G. Keller and S. Rady. Price dispersion and learning in a dynamic differentiated-goods duopoly. *The RAND Journal of Economics*, 34(1):138–165, 2003.
- D. A. Kendrick, H. M. Amman, and M. P. Tucci. Learning about learning in dynamic economic models. Discussion Paper 08-20, Tjalling C. Koopmans Research Institute, Utrecht School of Economics, 2008.
- J. O. Kephart, C. H. Brooks, R. Das, J. K. MacKie-Mason, R. Gazzale, and E. H. Durfee. Pricing information bundles in a dynamic environment. In *Proceedings of the 3rd ACM conference on Electronic Commerce*, pages 180–190. ACM, New York, 2001.
- N. B. Keskin. Optimal dynamic pricing with demand model uncertainty: a squared-coefficient-of-variation rule for learning and earning. Working paper, University of Chicago, Booth School of Business. Available at <http://ssrn.com/abstract=2487364>, 2014.
- N. B. Keskin and A. Zeevi. Chasing demand: Learning and earning in a changing environment. Working paper, University of Chicago, Booth School of Business. Available at <http://ssrn.com/abstract=2389750>, 2013.
- N. B. Keskin and A. Zeevi. Dynamic pricing with an unknown linear demand model: asymptotically optimal semi-myopic policies. *Operations Research*, 62(5):1142–1167, 2014.
- N. M. Kiefer and Y. Nyarko. Optimal control of an unknown linear process with learning. *International Economic Review*, 30(3):571–586, 1989.
- W. M. Kincaid and D. A. Darling. An inventory pricing problem. *Journal of Mathematical Analysis and Applications*, 7:183–208, 1963.
- A. Kirman. On mistaken beliefs and resultant equilibria. In R. Frydman and E. S. Phelps, editors, *Individual forecasting and aggregate outcomes*, pages 147–166. Cambridge University Press, New York, 1983.
- A. P. Kirman. Learning by firms about demand conditions. In R. H. Day and T. Graves, editors, *Adaptive Economics*, pages 137–156. Academic Press, New York, 1975.
- A. P. Kirman. Learning in oligopoly: Theory, simulation, and experimental evidence. In A. P. Kirman and M. Salmon, editors, *Learning and Rationality in Economics*, pages 127–178. Basil Blackwell, Cambridge, MA, 1995.
- R. Kleinberg. Nearly tight bounds for the continuum-armed bandit problem. In L. K. Saul, Y. Weiss, and L. Bottou, editors, *Advances in Neural Information Processing Systems 17*, pages 697–704. MIT Press, Cambridge, US, 2005.
- R. Kleinberg and T. Leighton. The value of knowing a demand curve: bounds on regret for online posted-price auctions. In *Proceedings of the 44th IEEE Symposium on Foundations of Computer Science*, pages 594–605, 2003.
- A. J. Kleywegt. An optimal control problem of dynamic pricing. Working paper, Georgia Institute of Technology, [http://www2.isye.gatech.edu/people/faculty/Anton\\_Kleywegt/OptControl.pdf](http://www2.isye.gatech.edu/people/faculty/Anton_Kleywegt/OptControl.pdf), 2001.
- D. Kong. One dynamic pricing strategy in agent economy using neural network based on online learning. In *WI '04 Proceedings of the 2004 IEEE/WIC/ACM International Conference on Web Intelligence*, pages 98–102, 2004.
- V. Könönen. Dynamic pricing based on asymmetric multiagent reinforcement learning. *International Journal of Intelligent Systems*, 21(1):73–98, 2006.



- P. K. Kopalle, A. G. Rao, and J. L. Assunção. Asymmetric reference price effects and dynamic pricing policies. *Marketing Science*, 15(1):60–85, 1996.
- E. Kutschinski, T. Uthmann, and D. Polani. Learning competitive pricing strategies by multi-agent reinforcement learning. *Journal of Economic Dynamics and Control*, 27(11-12):2207–2218, 2003.
- C. Kwon, T. L. Friesz, R. Mookherjee, T. Yao, and B. Feng. Non-cooperative competition among revenue maximizing service providers with demand learning. *European Journal of Operational Research*, 197(3): 981–996, 2009.
- H. D. Kwon, S. A. Lippman, and C. S. Tang. Optimal markdown pricing with demand learning. *Probability in the Engineering and Informational Sciences*, 26(1):77–104, 2012.
- T. L. Lai and H. Robbins. Iterated least squares in multiperiod control. *Advances in Applied Mathematics*, 3(1):50–73, 1982.
- T. L. Lai and H. Robbins. Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics*, 6(1):4–22, 1985.
- T. L. Lai and C. Z. Wei. Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. *The Annals of Statistics*, 10(1):154–166, 1982.
- M. A. Lariviere and E. L. Porteus. Manufacturer-retailer contracting under an unknown demand distribution. Working paper, Duke University / Stanford University, 1995.
- E. P. Lazear. Retail pricing and clearance sales. *The American Economic Review*, 76(1):14–32, 1986.
- T. Le Guen. Data-driven pricing. Master’s Thesis, Massachusetts Institute of Technology. Available at <http://hdl.handle.net/1721.1/45627>, 2008.
- R. A. Lehfeldt. The elasticity of demand for wheat. *The Economic Journal*, 24(94):212–217, 1914.
- Y. Lei, S. Jasin, and A. Sinha. Near-optimal bisection search for nonparametric dynamic pricing with inventory constraint. Working paper. University of Michigan, Ross School of Business Paper No. 1252. Available at <http://ssrn.com/abstract=2509425>, 2014.
- H. E. Leland. Theory of the firm facing uncertain demand. *The American Economic Review*, 62(3): 278–291, 1972.
- B. Leloup and L. Deveaux. Dynamic pricing on the internet: theory and simulations. *Electronic Commerce Research*, 1(3):265–276, 2001.
- Y. Levin, J. McGill, and M. Nediak. Dynamic pricing in the presence of strategic consumers and oligopolistic competition. *Management Science*, 55(1):32–46, 2009.
- T. Levina, Y. Levin, J. McGill, and M. Nediak. Dynamic pricing with online learning and strategic consumers: an application of the aggregating algorithm. *Operations Research*, 57(2):327–341, 2009.
- G. Li and W. Li. Dynamic pricing of perishable products with random fuzzy demand. In *2010 International Conference on Management Science & Engineering*, pages 191–199. IEEE, 2010.
- G. Li, Y. Xiong, and Z. Xiong. Robust dynamic pricing over infinite horizon in the presence of model uncertainty. *Asia-Pacific Journal of Operational Research*, 26(6):779–804, 2009.
- G. Li, Z. Xiong, Y. Zhou, and Y. Xiong. Dynamic pricing for perishable products with hybrid uncertainty in demand. *Applied Mathematics and Computation*, 219(20):10366–10377, 2013.
- H. Li and W. T. Huh. Pricing multiple products with the multinomial logit and nested logit models: concavity and implications. *Manufacturing & Service Operations Management*, 13(4):549–563, 2011.
- J. Li, T. Yao, and H. Gao. A revenue maximizing strategy based on Bayesian analysis of demand dynamics. In D. A. Field and T. J. Peters, editors, *Proceedings of the 2009 SIAM Conference on Mathematics for Industry*, pages 174–181, Philadelphia, PA, 2010. Society for Industrial and Applied Mathematics.



- A. E. B. Lim and J. G. Shanthikumar. Relative entropy, exponential utility, and robust dynamic pricing. *Operations Research*, 55(2):198–214, 2007.
- A. E. B. Lim, J. G. Shanthikumar, and T. Watewai. Robust multi-product pricing. Working paper, University of California, Berkeley. Available at <http://dx.doi.org/10.2139/ssrn.1078012>, 2008.
- K. Y. Lin. Dynamic pricing with real-time demand learning. *European Journal of Operational Research*, 174(1):522–538, 2006.
- G. Liu and H. Wang. An online sequential feed-forward network model for demand curve prediction. *Journal of Information & Computational Science*, 10(10):3063–3069, 2013.
- Q. Liu and G. J. van Ryzin. Strategic capacity rationing to induce early purchases. *Management Science*, 54(6):1115–1131, 2008.
- Q. Liu and G. J. van Ryzin. Strategic capacity rationing when customers learn. *Manufacturing & Service Operations Management*, 13(1):89–107, 2011.
- M. S. Lobo and S. Boyd. Pricing and learning with uncertain demand. Working paper, Fuqua School of Business, Duke University. Available at [http://www.stanford.edu/~boyd/papers/pdf/pric\\_learn\\_unc\\_dem.pdf](http://www.stanford.edu/~boyd/papers/pdf/pric_learn_unc_dem.pdf), 2003.
- O. Loginova and C. R. Taylor. Price experimentation with strategic buyers. *Review of Economic Design*, 12(3):165–187, 2008.
- Y. Lou. A unified framework of proactive self-learning dynamic pricing for high-occupancy/toll lanes. *Transportmetrica A: Transport Science*, 9(3):205–222, 2013.
- C. Maglaras and J. Meissner. Dynamic pricing strategies for multiproduct revenue management problems. *Manufacturing & Service Operations Management*, 8(2):136–148, 2006.
- V. Mahajan, E. Muller, and F. M. Bass. New product diffusion models in marketing: a review and directions for research. *Journal of Marketing*, 54(1):1–26, 1990.
- M. Mahmoudzadeh, S. J. Sadjadi, and S. Mansour. Robust optimal dynamic production/pricing policies in a closed-loop system. *Applied Mathematical Modelling*, 37(16-17):8141–8161, 2013.
- R. Manning. Market research by a monopolist: a Bayesian analysis. *Economica*, 46(183):301–306, 1979.
- A. Marcet and T. J. Sargent. Convergence of least squares learning mechanisms in self-referential linear stochastic models. *Journal of Economic Theory*, 48(2):337–368, 1989.
- E. Maskin and J. Riley. Monopoly with incomplete information. *The RAND Journal of Economics*, 15(2):171–196, 1984.
- R. Mason and J. Välimäki. Learning about the arrival of sales. *Journal of Economic Theory*, 146(4):1699–1711, 2011.
- T. Mazumdar, S. P. Raj, and I. Sinha. Reference price research: review and propositions. *Journal of Marketing*, 69(4):84–102, 2005.
- D. McFadden and K. Train. Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15(1):447–470, 2000.
- A. McLennan. Price dispersion and incomplete learning in the long run. *Journal of Economic Dynamic and Control*, 7(3):331–347, 1984.
- N. Meade and T. Islam. Modelling and forecasting the diffusion of innovation - a 25-year review. *International Journal of Forecasting*, 22(3):519–545, 2006.
- A. J. Mersereau and D. Zhang. Markdown pricing with unknown fraction of strategic customers. *Manufacturing & Service Operations Management*, 14(3):355–370, 2012.
- E. S. Mills. Uncertainty and price theory. *The Quarterly Journal of Economics*, 73(1):116–130, 1959.

- L. J. Mirman, L. Samuelson, and A. Urbano. Monopoly experimentation. *International Economic Review*, 34(3):549–563, 1993a.
- L. J. Mirman, L. Samuelson, and A. Urbano. Duopoly signal jamming. *Economic Theory*, 3(1):129–149, 1993b.
- H. L. Moore. *Economic Cycles; Their Law and Cause*. The Macmillan Company, New York, 1914.
- H. L. Moore. *Forecasting the Yield and the Price of Cotton*. The Macmillan Company, New York, 1917.
- S. Morales-Enciso and J. Branke. Revenue maximization through dynamic pricing under unknown market behaviour. In S. Ravizza and P. Holborn, editors, *3rd Student Conference on Operational Research (SCOR 2012)*, pages 11–20. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2012.
- P. B. Mullen, C. K. Monson, K. D. Seppi, and S. C. Warnick. Particle swarm optimization in dynamic pricing. In *2006 IEEE International Conference on Evolutionary Computation*, pages 1232–1239. IEEE, 2006.
- H. Mutlu, M. Alanyali, D. Starobinski, and A. Turhan. Online pricing of secondary spectrum access with unknown demand function. *Selected Areas in Communications, IEEE Journal on*, 30(11):2285–2294, 2012.
- H. Nazerzadeh, A. Saberi, and R. Vohra. Dynamic pay-per-action mechanisms and applications to online advertising. *Operations Research*, 61(1):98–111, 2013.
- S. Netessine. Dynamic pricing of inventory/capacity with infrequent price changes. *European Journal of Operational Research*, 174(1):553–580, 2006.
- A. J. Nevins. Some effects of uncertainty: simulation of a model of price. *The Quarterly Journal of Economics*, 80(1):73–87, 1966.
- D. Nguyen. The monopolistic firm, random demand, and Bayesian learning. *Operations Research*, 32(5):1038–1051, 1984.
- D. Nguyen. *Pricing under uncertainty*, volume 6 of *International Series in Quantitative Marketing*, chapter 2, pages 23–57. Springer, 1997.
- Ö. Özer and R. Phillips, editors. *The Oxford Handbook of Pricing Management*. Oxford University Press, London, 2012.
- G. Perakis and A. Sood. Competitive multi-period pricing for perishable products: a robust optimization approach. *Mathematical Programming*, 107(1):295–335, 2006.
- V. Perchet and P. Rigollet. The multi-armed bandit problem with covariates. *The Annals of Statistics*, 41(2):693–721, 2013.
- N. C. Petruzzi and M. Dada. Dynamic pricing and inventory control with learning. *Naval Research Logistics*, 49(3):303–325, 2002.
- R. Phillips. *Pricing and Revenue Optimization*. Stanford University Press, Stanford, CA, 2005.
- I. Popescu and Y. Wu. Dynamic pricing strategies with reference effects. *Operations Research*, 55(3):413–429, 2007.
- E. L. Porteus. Stochastic inventory theory. In D. P. Heyman and M. J. Sobel, editors, *Handbooks in Operations Research and Management Science*, volume 2, chapter 12, pages 605–652. Elsevier, Amsterdam, 1990.
- W. B. Powell. The knowledge gradient for optimal learning. In J. J. Cochran, editor, *Encyclopedia for Operations Research and Management Science*. Wiley, New York, 2010.
- E. C. Prescott. The multi-period control problem under uncertainty. *Econometrica*, 40(6):1043–1058, 1972.

- H. Qu, I. O. Ryzhov, and M. C. Fu. Learning logistic demand curves in business-to-business pricing. In R. Pasupathy, S. H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, editors, *Proceedings of the 2013 Winter Simulation Conference*. IEEE, 2013.
- C. V. L. Raju, Y. Narahari, and K. Ravikumar. Learning dynamic prices in electronic retail markets with customer segmentation. *Annals of Operations Research*, 143(1):59–75, 2006.
- K. Raman and R. Chatterjee. Optimal monopolist pricing under demand uncertainty in dynamic markets. *Management Science*, 41(1):144–162, 1995.
- S. Ramezani, P. A. N. Bosman, and H. La Poutre. Adaptive strategies for dynamic pricing agents. In *2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*, pages 323–328. IEEE, 2011.
- R. Rana and F. S. Oliveira. Real-time dynamic pricing in a non-stationary environment using model-free reinforcement learning. *Omega*, 47:116–126, 2014.
- R. Rana and F. S. Oliveira. Dynamic pricing policies for interdependent perishable products or services using reinforcement learning. *Expert Systems with Applications*, 42(1):426–436, 2015.
- V. R. Rao, editor. *Handbook of Pricing Research in Marketing*. Edward Elgar, Cheltenham, UK, 2009.
- S. Rassenti, S. S. Reynolds, V. L. Smith, and F. Szidarovszky. Adaptation and convergence of behavior in repeated experimental Cournot games. *Journal of Economic Behavior & Organization*, 41(2):117–146, 2000.
- H. Robbins. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, 58(5):527–535, 1952.
- B. Robinson and C. Lakhani. Dynamic price models for new-product planning. *Management Science*, 21(10):1113–1122, 1975.
- C. F. Roos. A mathematical theory of competition. *American Journal of Mathematics*, 47(3):163–175, 1925.
- C. F. Roos. A dynamic theory of economics. *Journal of Political Economy*, 35(5):632–656, 1927a.
- C. F. Roos. Dynamical economics. *Proceedings of the National Academy of Sciences of the United States of America*, 13(3):145–150, 1927b.
- C. F. Roos. *Dynamic Economics: Theoretical and Statistical Studies of Demand, Production, and Prices*. The Principia Press, Bloomington, 1934.
- M. Roozbehani, M. Dahleh, and S. Mitter. Dynamic pricing and stabilization of supply and demand in modern electric power grids. In *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, pages 543–548. Curran Associates, Inc., Redhook NY, USA, 2010.
- M. Rothschild. A two-armed bandit theory of market pricing. *Journal of Economic Theory*, 9(2):185–202, 1974.
- P. Rusmevichientong and J. N. Tsitsiklis. Linearly parameterized bandits. *Mathematics of Operations Research*, 35(2):395–411, 2010.
- P. Rusmevichientong, Z-J M. Shen, and D. B. Shmoys. Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations Research*, 58(6):1666–1680, 2010.
- A. Rustichini and A. Wolinsky. Learning about variable demand in the long run. *Journal of Economic Dynamics and Control*, 19(5-7):1283–1292, 1995.
- A. Sandmo. On the theory of the competitive firm under price uncertainty. *The American Economic Review*, 61(1):65–73, 1971.
- T. R. Sass. A note on optimal price cutting behavior under demand uncertainty. *The Review of Economics and Statistics*, 70(2):336–339, 1988.

- D. Sauré and A. Zeevi. Optimal dynamic assortment planning with demand learning. *Manufacturing & Service Operations Management*, 15(3):387–404, 2013.
- H. Scarf. Bayes solutions of the statistical inventory problem. *The Annals of Mathematical Statistics*, 30(2):490–508, 1959.
- M. P. Schinkel, J. Tuinstra, and D. Vermeulen. Convergence of Bayesian learning to general equilibrium in mis-specified models. *Journal of Mathematical Economics*, 38(4):483–508, 2002.
- E. E. Schlee. Buyer experimentation and introductory pricing. *Journal of Economic Behavior & Organization*, 44(3):347–362, 2001.
- H. Schultz. *The theory and measurement of demand*. University of Chicago Press, Chicago, 1938.
- M. Schwind. *Dynamic Pricing and Automated Resource Allocation for Complex Information Services*, volume 589 of *Lecture Notes in Economics and Mathematical Systems*. Springer Berlin Heidelberg, 2007.
- P. B. Seetharaman. Dynamic pricing. In V. R. Rao, editor, *Handbook of Pricing Research in Marketing*, chapter 17, pages 384–393. Edward Elgar, Cheltenham, UK, 2009.
- I. Segal. Optimal pricing mechanisms with unknown demand. *American Economic Review*, 93(3):509–529, 2003.
- A. Sen and A. X. Zhang. Style goods pricing with demand learning. *European Journal of Operational Research*, 196(3):1058–1075, 2009.
- S. Shakya, F. Oliveira, and G. Owusu. Analysing the effect of demand uncertainty in dynamic pricing with EAs. In M. Bramer, M. Petridis, and F. Coenen, editors, *Research and Development in Intelligent Systems XXV. Proceedings of AI-2008, the Twenty-eighth SGA International Conference on Innovative Techniques and Applications of Artificial Intelligence*, pages 77–90. Springer London, 2009.
- S. Shakya, M. Kern, G. Owusu, and C. M. Chin. Neural network demand models and evolutionary optimisers for dynamic pricing. *Knowledge-Based Systems*, 29:44–53, 2012.
- Z. M. Shen and X. Su. Customer behavior modeling in revenue management and auctions: a review and new research opportunities. *Production and Operations Management*, 16(6):713–728, 2007.
- M. A. Simaan and T. Takayama. Optimum monopolist control in a dynamic market. *IEEE Transactions on Systems, Man, and Cybernetics*, 6(12):799–807, 1976.
- D. Simchi-Levi, X. Chen, and J. Bramel. Integration of inventory and pricing. In *The Logic of Logistics*, Springer Series in Operations Research and Financial Engineering, pages 177–209. Springer New York, 2014.
- M. E. Slade. Optimal pricing with costly adjustment: evidence from retail-grocery prices. *The Review of Economic Studies*, 65(1):87–107, 1998.
- A. Slivkins. Contextual bandits with similarity information. arXiv:0907.3986v5 [cs.DS], 2011.
- A. Smithies. The maximization of profits over time with changing cost and demand functions. *Econometrica*, 7(4):312–318, 1939.
- G. J. Stigler. The early history of empirical studies of consumer behavior. *Journal of Political Economy*, 62(2):95–113, 1954.
- G. J. Stigler. Henry L. Moore and statistical economics. *Econometrica*, 30(1):1–21, 1962.
- X. Su. Intertemporal pricing with strategic customer behavior. *Management Science*, 53(5):726–741, 2007.
- X. Su. Optimal pricing with speculators and strategic consumers. *Management Science*, 56(1):25–40, 2010.

- S. Subrahmanyam and R. Shoemaker. Developing optimal pricing and inventory policies for retailers who face uncertain demand. *Journal of Retailing*, 72(1):7–30, 1996.
- M. Suh and G. Aydin. Dynamic pricing of substitutable products with limited inventories under logit demand. *IIE Transactions*, 43(5):323–331, 2011.
- Z. Sun and A. E. Abbas. Bayesian updating on price elasticity of uncertain demand. In *2013 IEEE International Systems Conference (SysCon)*, pages 222–228. IEEE, 2013.
- A. Sweeting. Dynamic pricing behavior in perishable goods markets: evidence from secondary markets for major league baseball tickets. *Journal of Political Economy*, 120(6):1133–1172, 2012.
- R. Swinney. Selling to strategic consumers when product value is uncertain: the value of matching supply and demand. *Management Science*, 57(10):1737–1751, 2011.
- M. Talebian, N. Boland, and M. Savelsbergh. Pricing to accelerate demand learning in dynamic assortment planning. *European Journal of Operational Research*, 237(2):555–565, 2014.
- K. T. Talluri and G. J. van Ryzin. *The Theory and Practice of Revenue Management*. Kluwer Academic Publishers, Boston, 2004.
- J. B. Taylor. Asymptotic properties of multiperiod control rules in the linear regression model. *International Economic Review*, 15(2):472–484, 1974.
- P. Tehrani, Y. Zhai, and Q. Zhao. Dynamic pricing under finite space demand uncertainty: a multi-armed bandit with dependent arms. arXiv:1206.5345 [cs.LG], 2012.
- G. Tesauro and J. O. Kephart. Pricing in agent economies using multi-agent Q-learning. *Autonomous Agents and Multi-Agent Systems*, 5(1):289–304, 2002.
- A. Thiele. Single-product pricing via robust optimization. Working paper, Lehigh University. Available at [http://www.optimization-online.org/DB\\_FILE/2006/01/1319.pdf](http://www.optimization-online.org/DB_FILE/2006/01/1319.pdf), 2006.
- A. Thiele. Multi-product pricing via robust optimisation. *Journal of Revenue and Pricing Management*, 8:67–80, 2009.
- R. G. Thompson, M. D. George, P. L. Brown, and M. S. Proctor. Optimal production, investment, and output price controls for a monopoly firm of the Evans’ type. *Econometrica*, 39(1):119–129, 1971.
- W. R. Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3-4):285–294, 1933.
- S. A. O. Thore. Price strategies of an “ignorant” monopolist. In H. O. A. Wold, editor, *Econometric model building. Essays on the Causal Chain Approach*. North Holland Publishing Co., Amsterdam, 1964.
- J. Tinbergen. Bestimmung und Deutung von Angebotskurven: Ein Beispiel. *Zeitschrift für Nationalökonomie*, 1(5):669–679, 1930.
- G. Tintner. Monopoly over time. *Econometrica*, 5(2):160–170, 1937.
- D. Treffer. The ignorant monopolist: optimal learning with endogenous information. *International Economic Review*, 34(3):565–581, 1993.
- T. Tsekeris and S. Voß. Design and evaluation of road pricing: state-of-the-art and methodological advances. *NETNOMICS: Economic Research and Electronic Networking*, 10(1):5–52, 2009.
- J. Tuinstra. A price adjustment process in a model of monopolistic competition. *International Game Theory Review*, 6(3):417–442, 2004.
- M. J. Vázquez-Gallo, M. Estévez, and S. Egidio. Active learning and dynamic pricing policies. *American Journal of Operations Research*, 4(2):90–100, 2014.
- I. Venezia. Optimal investments in market research. *European Journal of Operational Research*, 18(2):198–207, 1984.

- D. Vengerov. A gradient-based reinforcement learning approach to dynamic pricing in partially-observable environments. *Future Generation Computer Systems*, 24(7):687–693, 2008.
- J. Vermorel and M. Mohri. Multi-armed bandit algorithms and empirical evaluation. In *Proceedings of the European Conference of Machine Learning, Porto, Portugal (October 3-7)*, volume 3720 of *Springer Lecture Notes in Computer Science*, pages 437–448, 2005.
- V. G. Vovk. Aggregating strategies. In *Proceedings of the Third Annual Workshop on Computational Learning Theory, COLT '90*, pages 371–386, San Francisco, 1990. Morgan Kaufmann Publishers.
- X. Wang. Dynamic pricing with a Poisson bandit model. *Sequential Analysis*, 26(4):355–365, 2007.
- Y. Wang, J. Y. Audibert, and R. Munos. Algorithms for infinitely many-armed bandits. In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, editors, *Advances in Neural Information Processing Systems 21*, pages 1729–1736. Curran Associates, Inc., 2009.
- Z. Wang, S. K. Mukhopadhyay, and D. Q. Yao. Bayesian pricing strategy for information goods. *International Journal of Operational Research*, 17(4):399–416, 2013.
- Z. Wang, S. Deng, and Y. Ye. Close the gaps: a learning-while-doing algorithm for a class of single-product revenue management problems. *Operations Research*, 62(2):318–331, 2014.
- R. D. Weaver and Y. Moon. Pricing perishables. In G. Schiefer and U. Rickert, editors, *2011 International European Forum, February 14-18, 2011, Innsbruck-Igls, Austria*, pages 212–231, 2011.
- V. Wieland. Learning by doing and the value of optimal experimentation. *Journal of Economic Dynamics and Control*, 24(4):501–534, 2000a.
- V. Wieland. Monetary policy, parameter uncertainty and optimal learning. *Journal of Monetary Economics*, 46:199–228, 2000b.
- T. Willems. Actively learning by pricing: a model of an experimenting seller. Working paper. Available at <https://sites.google.com/site/twillems85/>, 2014.
- U. Witt. How can complex economical behavior be investigated? The example of the ignorant monopolist revisited. *Behavioral Science*, 31(3):173–188, 1986.
- P. G. Wright. *The tariff on animal and vegetable oils*. The Macmillan Company, New York, 1928.
- E. G. Wruck. *Dynamic pricing implications of uncertainty about demand*. PhD thesis, Cornell University, 1989.
- C. H. Xia and P. Dube. Dynamic pricing in e-services under demand uncertainty. *Production and Operations Management*, 16(6):701–712, 2007.
- Y. Xiong, G. Li, and K. J. Fernandes. Dynamic pricing model and algorithm for perishable products with fuzzy demand. *Applied Stochastic Models in Business and Industry*, 26(6):758–774, 2010.
- C. A. Yano and S. M. Gilbert. Coordinated pricing and production/procurement decisions: a review managing business interfaces. In A. K. Chakravarty and J. Eliashberg, editors, *Managing Business Interfaces*, volume 16 of *International Series in Quantitative Marketing*, chapter 3, pages 65–103. Springer, New York, 2004.
- J. Y. Yu and S. Mannor. Unimodal bandits. In L. Getoor and T. Scheffer, editors, *Proceedings of the 28th International Conference on Machine Learning, Bellevue, Washington, USA*, pages 41–48. Omnipress, 2011.
- M. Yu, L. Debo, and R. Kapuscinski. Strategic waiting for consumer-generated quality information: dynamic pricing of new experience goods. Working paper 13-78, University of Chicago, Booth School of Business. Available at <http://ssrn.com/abstract=2334423>, 2013.
- E. Zabel. Monopoly and uncertainty. *The Review of Economic Studies*, 37(2):205–219, 1970.

- M. J. Zbaracki, M. Ritson, D. Levy, S. Dutta, and M. Bergen. Managerial and customer costs of price adjustment: direct evidence from industrial markets. *The Review of Economics and Statistics*, 86(2): 514–533, 2004.
- D. Zhang and W. L. Cooper. Pricing substitutable flights in airline revenue management. *European Journal of Operational Research*, 197(3):848–861, 2009.
- D. Zhang and Z. Lu. Assessing the value of dynamic pricing in network revenue management. *INFORMS Journal on Computing*, 25(1):102–115, 2013.
- J. L. Zhang and J. Chen. Bayesian solution to pricing and inventory control under unknown demand distribution. *Operations Research Letters*, 34(5):517–524, 2006.
- Y. Zhang, F. Y. L. Chin, and H. F. Ting. Competitive algorithms for online pricing. In B. Fu and D. Z. Du, editors, *Computing and Combinatorics*, volume 6842 of *Lecture Notes in Computer Science*, pages 391–401. Springer Berlin Heidelberg, 2011.
- W. Zhao and Y. S. Zheng. Optimal dynamic pricing for perishable assets with nonhomogeneous demand. *Management Science*, 46(3):375–388, 2000.