Mathematical Methods of Operations Research
https://doi.org/10.1007/s00186-020-00704-y

ORIGINAL ARTICLE

A pricing problem with unknown arrival rate and price sensitivity

Athanassios N. Avramidis[1]

Received: 23 February 2019 / Revised: 20 January 2020

Abstract

We study a pricing problem with finite inventory and semi-parametric demand uncertainty. Demand is a price-dependent Poisson process whose mean is the product of buyers' arrival rate, which is a constant λ, and buyers' purchase probability $q(p)$, where p is the price. The seller observes arrivals and sales, and knows neither λ nor q. Based on a non-parametric maximum-likelihood estimator of (λ, q), we construct an estimator of mean demand and show that as the system size and number of prices grow, it is asymptotically more efficient than the maximum likelihood estimator based only on sale data. Based on this estimator, we develop a pricing algorithm paralleling (Besbes and Zeevi in Oper Res 57:1407–1420, 2009) and study its performance in an asymptotic regime similar to theirs: the initial inventory and the arrival rate grow proportionally to a scale parameter n. If q and its inverse function are Lipschitz continuous, then the worst-case regret is shown to be $O((\log n/n)^{1/4})$. A second model considered is the one in Besbes and Zeevi (2009, Section 4.2), where no arrivals are involved; we modify their algorithm and improve the worst-case regret to $O((\log n/n)^{1/4})$. In each setting, the regret order is the best known, and is obtained by refining their proof methods. We also prove an $\Omega(n^{-1/2})$ lower bound on the regret. Numerical comparisons to *state-of-the-art alternatives* indicate the effectiveness of our arrivals-based approach.

Keywords Estimation · Asymptotic efficiency · Exploration–exploitation · Regret · Asymptotic analysis

Mathematics Subject Classification 60K10 · 93E35 · 90B05 · 62G20

✉ Athanassios N. Avramidis
aa1w07@soton.ac.uk

[1] School of Mathematical Sciences, University of Southampton, Southampton SO17 1BJ, UK

Published online: 11 February 2020

1 Introduction

1.1 Background

Pricing and revenue management are important problems in many industries. Talluri and van Ryzin (2005) discuss instances of this problem that range over many industries, including fashion and retail, air travel, hospitality, and leisure. Early literature assumes the relationship between the mean demand and the price is known to the seller (Gallego and van Ryzin 1994). In practice, decision makers seldom have such knowledge. *Pricing and demand learning* is a stream of literature concerned with pricing under incomplete knowledge of the demand process, which is estimated. A standard model is that whenever the price is set at p, the demand is a Poisson process of rate $\Lambda(p)$, which is called the *demand function* (Besbes and Zeevi 2009, 2012; Wang et al. 2014).

Estimation methods are broadly divided into *parametric* and *non-parametric*. The former assume a certain functional form and carry *mis-specification risk*, while the latter make weaker assumptions and tend to alleviate this risk (Besbes and Zeevi 2009). For the single-product problem with finite inventory, prominent works are Besbes and Zeevi (2009) and Wang et al. (2014). In Besbes and Zeevi (2009), a *learning phase* during which certain *test prices* are applied allows estimating the demand function via the realized demand (sales). This estimate leads to a "good" fixed-price policy that is applied during the remainder of the selling horizon, the so-called *exploitation phase*. Performance is measured by the *worst-case regret* across a class of demand functions, where *regret* is the expected revenue loss relative to that achievable by acting optimally under full knowledge of the demand function.

The following setting is the main focus of this paper. Potential buyers arrive according to a Poisson process of rate λ, regardless of the price on offer; and whenever the price is p, an arriving customer purchases with probability $q(p)$ independently of everything else. The pair (λ, q) are unknown to the seller.

1.2 Overview of the main contributions

A high-level summary of our contribution is as follows:

1. Starting from an arrival rate λ and a purchase probability (function) $q(\cdot)$ we define a class of demand models such that whenever the price is p, the demand is a Poisson process of rate $\lambda q(p)$, which matches the concept of demand function in the standard model. The essential deviation from the standard model is that we postulate the existence of the scaling factor λ that is unknown in addition to $q(\cdot)$.
2. Assuming that both arrivals and sales are observed, we construct a maximum-likelihood estimator (MLE) of $(\lambda, \boldsymbol{q})$, where $\boldsymbol{q} = (q(p_1), \ldots, q(p_\kappa))$ for arbitrary (distinct) prices $p_1, \ldots, p_\kappa$. The product of estimated λ and $\boldsymbol{q}$ is our estimator of mean demand, and it is shown to be asymptotically more efficient than the maximum-likelihood estimator based only on sale data (*sales-only estimator*) (Besbes and Zeevi 2009; Wang et al. 2014; Lei et al. 2014). This development focuses primarily on mean-square estimation efficiency.

3. We work with a model class $\mathcal{L}$ whose essential requirement is that $q(\cdot)$ and its inverse are Lipschitz continuous. Besbes and Zeevi (2009, Section 4.2) work with a class $\mathcal{L}_{BZ}$ where the demand function and its inverse are Lipschitz continuous. Wang et al. (2014) and Lei et al. (2014) work with a smaller class, $\mathcal{L}_W$, where the demand function is twice differentiable. Class $\mathcal{L}$ is smaller than $\mathcal{L}_{BZ}$, but it includes models outside $\mathcal{L}_W$. Allowing an inventory constraint, we develop a policy based on the estimated λ and $\boldsymbol{q}$, and a counterpart based on the sales-only estimator. These policies' worst-case regret is shown to be $O(\log n/n)^{1/4}$, against $\mathcal{L}$ (Theorem 1) and against $\mathcal{L}_{BZ}$ (Theorem 2), respectively. The convergence rate improves slightly that in Besbes and Zeevi (2009), and is the best known in each case.
4. We provide a lower bound on regret (Theorem 3). It is closely related to existing ones under the standard model (i.e., without an arrival rate); specifically, Wang et al. (2014, Theorem 2), and to a lesser extent Besbes and Zeevi (2012, Theorem 2) (discrete finite price set) and Broder and Rusmevichientong (2012, Theorem 3.1) (discrete-time pricing). Our bound does not immediately follow from any of them because their set of admissible policies does not contain ours, due to our allowing the price to depend on the arrival history (in addition to the sale history).

The remainder of this paper Section 1.3 reviews related literature. Section 2 introduces our model and formulates the problem. Section 3 presents the estimation and pricing methods. Section 4 analyzes the estimation problem. Section 5 analyzes the (worst-case) regret. Section 5.3 contains the lower bound on regret. Section 6 compares numerically against alternative methods. Section 8 contains selected proofs.

1.3 Related literature

The literature on pricing strategies is vast. We refer to Bitran and Caldentey (2003) and Talluri and van Ryzin (2005) for comprehensive reviews, and to den Boer (2015) for a more recent survey. Gallego and van Ryzin (1994) characterizes optimal pricing policies and develops an upper bound to the optimal revenue via a full-information deterministic relaxation, all under the assumption that the demand function is known. More recent literature addresses pricing problems with unknown demand function. In many works, the demand process is known up to a finite number of parameters. In Lin (2006), Araman and Caldentey (2009), and Farias and Van Roy (2010), there is a single unknown parameter representing the market size. den Boer and Keskin (2019) allow a discontinuous demand function, but they restrict it parametrically in each continuity interval.

Well-studied is the case of a single product and no inventory constraint (Broder and Rusmevichientong 2012; den Boer and Zwart 2014; Besbes and Zeevi 2015). Besbes and Zeevi (2015) show that pricing algorithms based on a mis-specified linear model of the demand function can perform well, under conditions. Keskin and Zeevi (2014) provide general sufficient conditions for a pricing policy to achieve asymptotic regret optimality when the demand function is linear. The coordination of price and inventory decisions is a related area; Chen et al. (2019) employ non-parametric demand learning methods that bear resemblance to ours, but notable restrictions are that unsatisfied demand is backlogged and the replenishment lead time is negligible.

Closely related to our work are Besbes and Zeevi (2009), Wang et al. (2014) and Lei et al. (2014). Besbes and Zeevi (2009, Section 4.2) define a model class $\mathcal{L}_{BZ}$ where the demand function and its inverse are Lipschitz continuous, and achieve worst-case regret $O(n^{-1/4}(\log n)^{1/2})$ against this class; they also prove an $\Omega(n^{-1/2})$ lower bound on the regret of any admissible policy. Wang et al. (2014) and Lei et al. (2014) work with a smaller classes in which the demand function is smooth (twice-differentiable everywhere). Wang et al. (2014) iterate over shrinking price intervals, all of which contain the *optimal static price* under the deterministic relaxation (Gallego and van Ryzin 1994); the time spent learning and the number of test prices are controlled carefully as functions of the iteration count; the worst-case regret is $O(n^{-1/2}(\log n)^{4.5})$. Lei et al. (2014) develop iterative algorithms that also control carefully key parameters; their worst-case regret is $O(n^{-1/2})$. These methods first estimate the order relation between the unconstrained maximizer of the revenue-rate function and the *clearance price* (Gallego and van Ryzin 1994); this guides the onward choice of parameters. By excluding a non-differentiable demand function, these methods are not directly comparable to ours.

This paper has some connections to the continuum-armed bandit literature [e.g., Auer et al. (2002), Kleinberg (2004), Cope (2009)], but a clear distinguishing feature here is the presence of an inventory constraint. A recent work that addresses the inventory constraint is Babaioff et al. (2015).

2 Problem formulation and background

Model of Demand and Basic Assumptions A monopolist sells a single product. The selling horizon (period) is $T > 0$; after this point sales are discontinued, and any unsold products have no value. Product demand is as follows:

Assumption 1 (a) Customers arrive according to a Poisson process of rate λ. Whenever the price is p, an arriving customer purchases with probability $q(p) \in [0, 1]$ independently of everything else.

(b) The seller observes arrivals and sales throughout the selling period.

The *primitives* (λ, q) are unknown to the seller. The set of feasible prices is $[\underline{p}, \overline{p}] \cup p_\infty$, where $0 < \underline{p} < \overline{p} < \infty$, and $p_\infty > 0$ is a price that "turns off" demand, i.e., $q(p_\infty) = 0$. The *purchase probability* function $q(p)$ is assumed to be non-increasing and have an inverse. The *revenue rate* function per arrival, $r(p) := pq(p)$, is assumed to be concave. Clearly, the demand process generated through any pair (λ, q) is Poisson, with rate function $\Lambda(p) = \lambda q(p)$ whenever the price is p. Gallego and van Ryzin (1994), Besbes and Zeevi (2009), Wang et al. (2014) impose similar assumptions directly on the Poisson rate function (which is then called a *regular* demand function); we impose them on q instead. In analyzing the regret, we need:

Assumption 2 For some finite positive constants $\underline{\lambda}, \overline{\lambda}, \underline{M}, \overline{M}, m_a$:

(a) The arrival rate is bounded away from zero and infinity: $\underline{\lambda} \leq \lambda \leq \overline{\lambda}$.

(b) The purchase probability and its inverse are both Lipschitz functions. Specifically, for all $p, p' \in [\underline{p}, \overline{p}]$, we have $\underline{M}|p - p'| \leq |q(p) - q(p')| \leq \overline{M}|p - p'|$.
(c) The revenue rate can be made positive: $\max\{pq(p) : p \in [\underline{p}, \overline{p}]\} \geq m_a$.

Let $\mathcal{L} = \mathcal{L}(\overline{\lambda}, \underline{\lambda}, \underline{M}, \overline{M}, m)$ be the class of demand models satisfying Assumptions 1(a) and 2. Any bounded demand function Λ is representable in $\mathcal{L}$ as a pair $(\overline{\Lambda}, \Lambda(\cdot)/\overline{\Lambda})$ for any $\overline{\Lambda} \geq \sup_p \Lambda(p)$; moreover, smoothness properties of q and $\overline{\Lambda}q$ are identical.

Besbes and Zeevi (2009) (Assumption 1) define a model class $\mathcal{L}_{BZ}$ requiring Lipschitz continuity of the demand function and its inverse; Wang et al. (2014) define a smaller class $\mathcal{L}_W$ requiring, additionally, twice-differentiability.

Class $\mathcal{L}$ is strictly smaller than $\mathcal{L}_{BZ}$, and obeys no inclusion relation to $\mathcal{L}_W$. Indeed, if (λ, q) and (λ', q') satisfy $\lambda' \neq \lambda$ and $\lambda' q'(p) = \lambda q(p)$ for all p, and if these pairs cover the selling horizon non-trivially (each applying a positive time period), then they induce the time-invariant demand function $\lambda q(p)$; this model is not representable in $\mathcal{L}$, and yet is representable in $\mathcal{L}_{BZ}$ and $\mathcal{L}_W$ under suitable conditions on q, q'. On the other hand, a pair (λ, q) such that q is Lipschitz but not twice-differentiable everywhere induces a model with demand function λq that is inside $\mathcal{L}$ and outside $\mathcal{L}_W$, since λq inherits the non-differentiability of q.

Estimation problem In studying estimation efficiency, we need:

Assumption 3 No stock out occurs during the estimation (learning) phase.

This assumption, which is commonly made (Broder and Rusmevichientong 2012; den Boer and Zwart 2014; Besbes and Zeevi 2015; Keskin and Zeevi 2014), can probably be relaxed, or by-passed by showing that a stock-out is a rare event. On this point, in Lemma 1 we show that the probability of a stock-out vanishes fast when the system size increases and the fraction of time spent learning vanishes sufficiently fast. Since the efficiency analysis is lengthy even under Assumption 3, we do not attempt to relax this assumption.

The efficiency of the MLE-based demand estimator relative to the sales-only counterpart (based on the same learning time and number of test prices) is studied on a sequence of estimation problems indexed by n, with arrival rate λ_n, learning time unity, and number of test prices κ_n. We assume:

$$r_n := \frac{\lambda_n}{\kappa_n} \to \infty, \quad \kappa_n r_n e^{-r_n} \to 0 \quad \text{as } n \to \infty. \tag{1}$$

That is, r_n, the mean per-price-arrival-count, or less formally the "information per price" grows large. The second condition is aimed at keeping the bias (present in the MLE only) negligible relative to variance. In studying the efficiency in estimating the function $q(\cdot)$ (not merely $q(p)$ for fixed p) we need:

Assumption 4 The purchase-probability $q(\cdot)$ is continuous almost everywhere on $[\underline{p}, \overline{p}]$; that is, the set of its discontinuities has measure zero.

Pricing problem Let $x > 0$ denote the level of inventory at the start of the selling period. Let $(p_t : 0 \leq t \leq T)$ denote the *price process* (assumed to take values in $[\underline{p}, \overline{p}] \cup p_\infty$ and to have right-continuous paths with left limits). Let $N = (N(t) :$

$t \geq 0$) be a unit-rate Poisson process. Then, the cumulative demand up to time t has the representation $D_t = N\big(\lambda \int_0^t q(p_s)\mathrm{d}s\big)$ by Assumption 1(a). A process (p_t) is said to be *non-anticipating* if its value at any time t is only allowed to depend on past prices $\{p_u : u \in [0, t)\}$, past arrival counts $\{A_u : u \in [0, t)\}$, and past demand counts $\{D_u : u \in [0, t)\}$. In other words, the process (p_t) is adapted to the filtration $\mathcal{F}_t = \sigma(p_u^\pi, A_u, D_u : 0 \leq u < t)$. A *pricing policy* π is a method of constructing a non-anticipating price process, which we then denote $(p_t^\pi : 0 \leq t \leq T)$. In particular, $D_t^\pi := N\big(\lambda \int_0^t q(p_s^\pi)\mathrm{d}s\big)$ is the cumulative demand up to time t under the policy π. A policy π is said to be *admissible* if the induced price process p^π is non-anticipating and $\int_0^T \mathrm{d}D_s^\pi \leq x$ with probability 1. Let $\mathcal{P}$ denote the set of admissible policies. The seller's problem is: choose $\pi \in \mathcal{P}$ to maximize the mean revenue $J^\pi = J^\pi(x, T|\lambda, q) := \mathbb{E}_{\lambda,q} \int_0^T p_s^\pi \mathrm{d}D_s^\pi$, where $\mathbb{E}_{\lambda,q}$ denotes expectation with respect to (λ, q).

Full-information deterministic relaxation Here the stochastic elements above are replaced by their means, giving rise to the problem

$$J^D = J^D(x, T|\lambda, q) := \sup \int_0^T \lambda r(p_s)\mathrm{d}s$$
$$\text{s.t.} \int_0^T \lambda q(p_s)\mathrm{d}s \leq x, \quad p_s \in [\underline{p}, \overline{p}] \cup p_\infty \text{ for all } s. \tag{2}$$

In line with the literature, our motivation is that the optimal solution to (2) performs well for the original (stochastic) problem (Gallego and van Ryzin 1994, Theorem 3). The solution prices at $p^D := \max\{p^u, p^c\}$ (while inventory is positive), where $p^u := \arg\max_p\{pq(p)\}$ is the *unconstrained maximizer*; and $p^c := \arg\min_p |\lambda q(p) - x/T|$ is the *clearance price*. We define the *regret* of π as $1 - J^\pi/J^D$; this is conservative, since $\sup_{\pi \in \mathcal{P}} J^\pi(x, T|\lambda, q) \leq J^D(x, T|\lambda, q)$ (Gallego and van Ryzin 1994, [Theorem 1]).

Large-system performance analysis In the pricing problem, given a primitive $(\lambda, q) \in \mathcal{L}$, we consider a sequence of problems indexed by $n = 1, 2, \ldots$ so that in the n-th problem, the initial inventory is $x_n = nx$; the arrival rate is $\lambda_n = n\lambda$; the purchase-probability function q and the horizon T are fixed (same for all n). Our main objective is to study the regret as $n \to \infty$.

Model discussion Our assumption of existence of a price-independent arrival rate is quite restrictive; in e-commerce, for example, many customers are channeled to a store through a price-comparison service, and thus one would expect the arrival rate to increase as the price is lowered. On the other hand, it may be reasonable for some physical (brick and mortar) stores when customers exhibit loyalty, that is, the preference to shop at a particular store (e.g., a local grocery or convenience store). In this case, customers might arrive at a fairly constant rate regardless of price, and, once in store, their purchase decision could be price-dependent. For example, a convenience store selling a fixed inventory of beer (having a "use by" or "best by" date rendering the product worthless beyond that date), could postulate a price-independent arrival rate.

Our formulation deviates from Besbes and Zeevi (2009), Besbes and Zeevi (2012), Broder and Rusmevichientong (2012), Wang et al. (2014) in that the seller has additional information, namely the arrival counts. Consequently, correspondingly different is our notion of non-anticipation, and by consequence the notion of admissibility. Specifically, their definition of admissibility requires that the price process (p_t) is adapted to the filtration $\sigma(p_u^\pi, D_u : 0 \le u < t)$, which contains less information than our $\mathcal{F}_t$. The pricing policy we propose naturally involves the past arrival counts and therefore does not belong to the set of admissible policies they consider.

Notation Statements "$x := y$" and "$y =: x$" define x through y. When A is a set, $\mathbb{1}_{[A]}$ is the indicator function; A^c is the complement; and $|A|$ is the cardinality. When x is real, $\lfloor x \rfloor$ is its *floor*, $\lceil x \rceil$ is its *ceiling*, and x^+ is $\max\{0, x\}$. When a_n and b_n are nonnegative sequences, $a_n = O(b_n)$ means that a_n/b_n is bounded from above; $a_n = \Omega(b_n)$ means that a_n/b_n is bounded from below; $a_n \asymp b_n$ means that a_n/b_n is bounded from both above and below; $a_n \sim b_n$ means that $\lim_{n\to\infty} a_n/b_n = 1$; and $a_n = o(b_n)$ means that $\limsup_{n\to\infty} a_n/b_n = 0$. When X is a random variable, "$X \sim \cdot$" indicates the probability law of X, and "$X \Rightarrow \cdot$" denotes convergence in distribution.

3 Estimation and pricing methods

The estimation method in the presence of an arrival rate is as follows:

Method A (τ, κ) Let $p_1, \ldots, p_\kappa$ be any κ distinct prices in $[\underline{p}, \overline{p}]$, i.e., $p_i \neq p_j$ whenever $i \neq j$. Set the learning interval as $[0, \tau]$, and set $\Delta = \tau/\kappa$. For $i = 1, 2, \ldots, \kappa$, during the time interval $[(i-1)\Delta, i\Delta]$, price at p_i, and let A_i and S_i be the count of arrivals and sales, respectively. If a stock-out occurs before time τ, then price at p_∞ and stop sales. Put

$$\widehat{\lambda} = \frac{\sum_{i=1}^{\kappa} A_i}{\tau}, \quad \widehat{q}(p_i) = \frac{S_i}{A_i} \mathbb{1}_{[A_i > 0]}, \; i = 1, \ldots, \kappa, \tag{3}$$

as a joint estimator of λ and the $q(p_i)$.

The pricing policy in the presence of an arrival rate is as follows.

ALGORITHM AS or $\pi(\tau, \kappa)$.

Step 1 Initialization:

(a) Let the *learning time* be τ and the number of *test prices* be κ. Put $\Delta = \tau/\kappa$.
(b) Divide $[\underline{p}, \overline{p}]$ into κ equally spaced intervals. Let $\{p_i, i = 1, \ldots, \kappa\}$ be the midpoints (or left-endpoints or right-endpoints) of these intervals.

Step 2 Learning:

(a) For $i = 1, \ldots, \kappa$, and provided inventory is positive, apply price p_i during the time interval $[t_{i-1}, t_i]$, where $t_i := i\Delta$. If inventory is zero at any time, apply p_∞ until time T and stop.

(b) Let A_i and S_i be the number of arrivals and sales, respectively, during $[t_{i-1}, t_i)$. Compute

$$\widehat{\lambda} = \frac{\sum_{i=1}^{\kappa} A_i}{\tau}, \quad \widehat{q}(p_i) := \frac{S_i}{A_i} \mathbb{1}_{[A_i>0]} \tag{4}$$

Step 3 Compute $\widehat{p}^u = \arg\max_{1\le i\le\kappa}\{p_i\widehat{q}(p_i)\}$; $\widehat{p}^c = \arg\min_{1\le i\le\kappa}|\widehat{\lambda}\widehat{q}(p_i) - x/T|$; and set $\widehat{p} = \max\{\widehat{p}^u, \widehat{p}^c\}$.
Step 4 Pricing: On the interval $(\tau, T]$ apply price $\widehat{p}$ as long as inventory is positive. If inventory is zero at any time, apply p_∞ until time T and stop.

4 Results on estimation

4.1 Estimator of arrival rate and purchase probabilities

From Assumptions 1 and 3 there follow two properties:

P1. $A_1, \ldots, A_\kappa$ are independent Poisson$(\lambda\tau/\kappa)$ random variables.
P2. Given $\mathbf{A} := (A_1, A_2, \ldots, A_\kappa)$, the conditional law of $(S_1, S_2, \ldots S_\kappa)$ consists of the independent marginals $S_i \sim$ Binomial$(A_i, q(p_i))$ for all i.

The following result is elementary and thus given without proof.

Proposition 1 *Let Assumptions 1 and 3 hold. Let A_i and S_i be the count of arrivals and sales, respectively, when price p_i applies during Method $A(\tau, \kappa)$. Then, a maximum-likelihood estimator of $\big(\lambda, q(p_1), \ldots, q(p_\kappa)\big)$ is given in (3).*

4.2 Estimator of mean demand

Here we estimate the *demand vector* $(\lambda q(p_1), \ldots, \lambda q(p_\kappa))$ for finite κ. The invariance property of Maximum Likelihood Estimation (Bickel and Doksum 1977, Section 4.5) gives the following.

Proposition 2 *For any $p_1, \ldots, p_\kappa$, the vector $\big(\widehat{\lambda}\widehat{q}(p_1), \ldots, \widehat{\lambda}\widehat{q}(p_\kappa)\big)$, with $\widehat{\lambda}$ and $\widehat{q}(p_1), \ldots, \widehat{q}(p_\kappa)$ as in (3), is a maximum likelihood estimator of $\big(\lambda q(p_1), \ldots, \lambda q(p_\kappa)\big)$.*

We also refer to this estimator as the *arrivals-and-sales estimator*.

4.2.1 Bias and mean square error

For any price p, put $f(p) := q(p)[1 - q(p)]$ and define the *relative bias* $B := \mathbb{E}[\widehat{\lambda}\widehat{q}(p)]/(\lambda q(p)) - 1$ (shown below to not depend on p).

Proposition 3 *Let Assumptions 1 and 3 hold. Denote the i-th price applied during Method $A(\tau, \kappa)$ as $p = p_i$; denote A_i the associated arrival count, and let $\mathbf{A} = (A_1, \ldots, A_\kappa)$. Define $\sigma_1(p) := \mathbb{E}[\mathbb{V}ar(\widehat{\lambda}\widehat{q}(p)|\mathbf{A})]$ and $\sigma_2(p) := \mathbb{E}[\mathbb{E}^2[\widehat{\lambda}\widehat{q}(p)|\mathbf{A}]] -$*

$\lambda^2 q^2(p)(1+2B)$. *Let* $r := \lambda\tau/\kappa$*; and, with* $X_r \sim$ *Poisson*(r)*, let* $\rho := \mathbb{P}(X_r > 0) = 1 - e^{-r}$ *and* $h(r) := \mathbb{E}[X_r^{-1} \mathbb{1}_{[X_r>0]}]$. *Then*

$$B = -\frac{\kappa-1}{\kappa} e^{-r}, \quad \mathbb{E}[(\widehat{\lambda}\widehat{q}(p) - \lambda q(p))^2] = \sigma_1(p) + \sigma_2(p), \tag{5}$$

$$\sigma_1(p) = f(p)\tau^{-2} \left\{ \left[r^2(\kappa-1)^2 + r(\kappa-1) \right] h(r) + 2\rho r(\kappa-1) + r \right\}, \tag{6}$$

$$\mathbb{E}[\mathbb{E}^2[\widehat{\lambda}\widehat{q}(p)|\mathbf{A}]] = q^2(p)\tau^{-2} \left\{ \left[r^2(\kappa-1)^2 + r(\kappa-1) \right] \rho + 2r^2(\kappa-1) + r^2 + r \right\}. \tag{7}$$

Remark 1 For any estimator θ we have $\text{MSE}(\theta) = \mathbb{E}[\mathbb{V}ar(\theta|\mathbf{A})] + \mathbb{V}ar(\mathbb{E}[\theta|\mathbf{A}]) + \text{Bias}^2(\theta)$, where Bias and MSE denote bias and mean square error, respectively. Thus $\sigma_2(p) = \mathbb{V}ar(\mathbb{E}[\widehat{\lambda}\widehat{q}(p)|\mathbf{A}]) + [\lambda q(p)B]^2 \geq 0$; this term aggregates variance (of the conditional expectation) and bias squared.

Proof of Proposition 3 *Proof of* (5). Based on Properties P1 and P2 (Sect. 4.1), we have:

$$\mathbb{E}[\widehat{\lambda}\widehat{q}(p_i)|\mathbf{A}] = \widehat{\lambda}\mathbb{E}[\widehat{q}(p_i)|\mathbf{A}] = \widehat{\lambda} q(p_i)\mathbb{1}_{[A_i>0]}, \tag{8}$$

$$\begin{aligned} \mathbb{V}ar(\widehat{\lambda}\widehat{q}(p_i)|\mathbf{A}) &= \widehat{\lambda}^2 \mathbb{V}ar(\widehat{q}(p_i)|\mathbf{A}) \\ &= \widehat{\lambda}^2 A_i^{-1} q(p_i)(1-q(p_i))\mathbb{1}_{[A_i>0]}, \end{aligned} \tag{9}$$

$$\begin{aligned} \mathbb{E}[\widehat{\lambda}\widehat{q}(p_i)] = \mathbb{E}\left[\mathbb{E}[\widehat{\lambda}\widehat{q}(p_i)|\mathbf{A}]\right] &\overset{(a)}{=} q(p_i)\tau^{-1}\mathbb{E}\left[A_i \mathbb{1}_{[A_i>0]} + \sum_{1\le j\le\kappa: j\neq i} A_j \mathbb{1}_{[A_i>0]} \right] \\ &\overset{(b)}{=} q(p_i)\tau^{-1}\left[r + (\kappa-1)r\rho\right] \\ &= \lambda q(p_i)\left[1 - \frac{(\kappa-1)(1-\rho)}{\kappa}\right] \end{aligned}$$

where (8) and (9) follow immediately from Property P2; step (a) uses (8) and that $\widehat{\lambda} = \tau^{-1}\sum_i A_i$; step (b) uses: the independence of the A_i; $\mathbb{P}(A_i > 0) = \rho$; and $\mathbb{E}[A_i] = r$. This proves the left part of (5). For the right part of (5), put $X = \widehat{\lambda}\widehat{q}(p_i)$, $\mu = \lambda q(p_i)$, and $B = (\mathbb{E}[X] - \mu)/\mu$; now use the identities $\mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2] - \mu^2(1+2B)$; $\mathbb{E}[X^2] = \mathbb{E}[\mathbb{E}[X^2|\mathbf{A}]]$; and $\mathbb{E}[X^2|\mathbf{A}] = \mathbb{V}ar(X|\mathbf{A}) + \mathbb{E}^2[X|\mathbf{A}]$.

Proof of (6). In view of (9), we have $\sigma_1(p) = q(p_i)(1-q(p_i))\mathbb{E}[\widehat{\lambda}^2 A_i^{-1}\mathbb{1}_{[A_i>0]}]$. It suffices to verify that $\mathbb{E}[\widehat{\lambda}^2 A_i^{-1}\mathbb{1}_{[A_i>0]}]$ equals τ^{-2} times the expression in curly braces in (6); this follows immediately by writing $\widehat{\lambda}^2 = \tau^{-2}[(\sum_{j:j\neq i} A_j)^2 + 2A_i \sum_{j:j\neq i} A_j + A_i^2]$ and using Property P1 to resolve the expectation.

Proof of (7). Observe that $\mathbb{E}[\mathbb{E}^2[\widehat{\lambda}\widehat{q}(p)|\mathbf{A}]] = q^2(p)\mathbb{E}[\widehat{\lambda}^2\mathbb{1}_{[A_i>0]}]$ by (8). It suffices to verify that $\mathbb{E}[\widehat{\lambda}^2\mathbb{1}_{[A_i>0]}]$ equals τ^{-2} times the expression in curly braces in (7); this follows immediately by expanding $\widehat{\lambda}^2$ as above and using Property P1 to resolve the expectation. □

4.2.2 Asymptotic efficiency relative to the sales-only estimator

Consider the scale-n instance and any price p. Let $\widehat{\lambda_n q_n}(p)$ denote the MLE of mean demand, $\lambda_n q(p)$, and let $\widehat{\lambda q}_n(p)$ denote the sales-only estimator, each based on applying any κ_n prices over $1/\kappa_n$ time units each. In particular,

$$\widehat{\lambda q}_n(p) := \frac{S(p)}{1/\kappa_n},$$

where $S(p)$ is the sale count during a period of length $1/\kappa_n$ during which price p is applied; thus, $S(p)$ is Poisson$(r_n q(p))$.

To analyze the MLE, let $\mathbf{A}_n = (A_{1,n}, \ldots, A_{\kappa_n,n})$ denote the arrival counts observed under Method A $(1, \kappa_n)$, and define scale-n analogs of B, $\sigma_1(p)$ and $\sigma_2(p)$ (Proposition 3) as follows: $B_n := \mathbb{E}[\widehat{\lambda_n q_n}(p)]/[\lambda_n q(p)] - 1$; $\sigma_{1,n}(p) := \mathbb{E}[\mathbb{V}ar(\widehat{\lambda_n q_n}(p)|\mathbf{A}_n)]$; and $\sigma_{2,n}(p) := \mathbb{E}[\mathbb{E}^2[\widehat{\lambda_n q_n}(p)|\mathbf{A}_n]] - \lambda_n^2 q^2(p)(1 + 2B_n)$.

The scale n will affect the mean square error through the multiplicative factor $s_n := r_n \kappa_n^2 = \lambda_n \kappa_n = \lambda_n^2 r_n^{-1}$. Moreover, the set $\mathcal{I} := \{p : p \in [\underline{p}, \overline{p}], f(p) > 0\}$ will be needed.

Proposition 4 *Let* (1) *hold and let Assumption* 3 *hold for all* $n \geq \tilde{n}$ *and some finite* $\tilde{n}$.

(a) *We have*

$$\lim_{n\to\infty} \sup_{p\in\mathcal{I}} \left| \frac{\sigma_{1,n}(p)}{r_n \kappa_n^2} - f(p) \right| = 0, \quad \lim_{n\to\infty} \sup_{p\in\mathcal{I}} \left| \frac{\sigma_{2,n}(p)}{r_n \kappa_n} - q^2(p) \right| = 0 \qquad (10)$$

(b) *(Pointwise efficiency) If* $\lim_{n\to\infty} \kappa_n = \infty$, *then*

$$\lim_{n\to\infty} \sup_{p\in\mathcal{I}} \left| \frac{\mathbb{E}[(\widehat{\lambda_n q_n}(p) - \lambda_n q(p))^2]}{r_n \kappa_n^2} - f(p) \right| = 0, \qquad (11)$$

and for any price p *such that* $q(p) > 0$,

$$\lim_{n\to\infty} \frac{\mathbb{E}[(\widehat{\lambda q}_n(p) - \lambda_n q(p))^2]}{\mathbb{E}[(\widehat{\lambda_n q_n}(p) - \lambda_n q(p))^2]} = \begin{cases} [1 - q(p)]^{-1} & \text{if } q(p) < 1 \\ \infty & \text{if } q(p) = 1 \end{cases}$$

(c) *(Global efficiency) For each* $n \geq 1$, *let the two estimators be employed at the prices* $p_{i,n} = \underline{p} + (i - 1/2)\ell_n$, $i = 1, \ldots, \kappa_n$, *where* $\ell_n := (\overline{p} - \underline{p})/\kappa_n$ *and where* $\lim_{n\to\infty} \kappa_n = \infty$. *If Assumption* 4 *holds, then*

$$\lim_{n\to\infty} \frac{\mathbb{E}[\sum_{i=1}^{\kappa_n} (\widehat{\lambda q}_n(p_{i,n}) - \lambda_n q(p_{i,n}))^2]}{\mathbb{E}[\sum_{i=1}^{\kappa_n} (\widehat{\lambda_n q_n}(p_{i,n}) - \lambda_n q(p_{i,n}))^2]} = \frac{\int_{\mathcal{I}} q(p)\mathrm{d}p}{\int_{\mathcal{I}} f(p)\mathrm{d}p} =: \mu_\infty \geq 1. \qquad (12)$$

Proof of Proposition 4 All limits are meant as $n \to \infty$.

Proof of (a) Let $p \in \mathcal{I}$, i.e., $f(p) > 0$ (hence $q^2(p) > 0$). All terms in Proposition 3 and its proof are denoted by appending the scale n as a rightmost subscript. Observe:

$$\begin{aligned}\frac{\sigma_{1,n}(p)}{f(p)} &\overset{(a)}{=} \left[r_n^2(\kappa_n-1)^2 + r_n(\kappa_n-1)\right] h(r_n) + 2\rho_n r_n(\kappa_n-1) + r_n \\ &\overset{(b)}{=} \left[r_n^2(\kappa_n-1)^2 + r_n(\kappa_n-1)\right] r_n^{-1}(1+o(1)) + 2(1+o(1))r_n(\kappa_n-1) + r_n \\ &\overset{(c)}{=} r_n\kappa_n^2(1+o(1)).\end{aligned} \tag{13}$$

Step (a) is (6) re-arranged so that the right side does not depend on p. Step (b) uses: (i) $h(r_n) = r_n^{-1}(1+o(1))$, using Lemma 4 and $r_n \to \infty$; and (ii) $\rho_n = 1-e^{-r_n} = 1+o(1)$. Step (c) notes that $(\kappa_n-1)/(r_n\kappa_n^2) \to 0$, due to $r_n \to \infty$; this proves the first part of (a). Now observe:

$$\begin{aligned}\frac{\sigma_{2,n}(p)}{q^2(p)} &\overset{(a)}{=} \left[r_n^2(\kappa_n-1)^2 + r_n(\kappa_n-1)\right] \rho_n + 2r_n^2(\kappa_n-1) + r_n^2 + r_n - \lambda_n^2(1+2B_n) \\ &\overset{(b)}{=} \left[r_n^2\kappa_n^2 + r_n\kappa_n\right](1+o(1)) - \lambda_n^2(1+2B_n) \\ &\overset{(c)}{=} r_n\kappa_n\,(1+o(1) - 2r_n\kappa_n B_n) \\ &\overset{(d)}{=} r_n\kappa_n(1+o(1)).\end{aligned} \tag{14}$$

Step (a) is (7) re-arranged so that the right side does not depend on p. Step (b) uses that $\rho_n = 1+o(1)$ and collects terms (into $r_n^2\kappa_n^2$). Step (c) notes a cancellation (due to $r_n\kappa_n = \lambda_n$). Step (d) uses that $B_n = -e^{-r_n}(\kappa_n-1)/\kappa_n$ and thus $r_n\kappa_n B_n = o(1)$, by the second condition in (1). This proves the second part of (a).

Proof of (b). Result (11) follows immediately from part (a) and the assumption $\kappa_n \to \infty$ (i.e., $r_n\kappa_n$ vanishes relative to $r_n\kappa_n^2$). The sales-only estimator has $\mathbb{E}[(\widehat{\lambda q}_n(p) - \lambda_n q(p))^2] = \mathbb{V}ar(S(p))\kappa_n^2 = q(p)r_n\kappa_n^2$, and the second part of (b) follows from (11), since $f(p)/q(p) = 1-q(p)$.

Proof of (c). Recall that $s_n = r_n\kappa_n^2 = \lambda_n\kappa_n$. Put $f_n(p) := s_n^{-1}\mathbb{E}[(\widehat{\lambda}_n\widehat{q}_n(p) - \lambda_n q(p))^2]$, $I_n := \sum_{i=1}^{\kappa_n} f(p_{i,n})\ell_n$, and observe: $\left|\sum_{i=1}^{\kappa_n} f_n(p_{i,n})\ell_n - \int_{\mathcal{I}} f(p)\mathrm{d}p\right| \leq \left|\sum_{i=1}^{\kappa_n} f_n(p_{i,n})\ell_n - I_n\right| + \left|I_n - \int_{\mathcal{I}} f(p)\mathrm{d}p\right|$. Each (absolute difference) term on the right is arbitrarily small for all n sufficiently large. For the first term, this uses the uniform convergence of $f_n(\cdot)$ to $f(\cdot)$ shown in (11) (note $\sum_{i=1}^{\kappa_n} \ell_n = \overline{p} - \underline{p} < \infty$); for the second term, this uses that f is Riemann integrable on $[\underline{p}, \overline{p}]$ [recall that Riemann integrability is equivalent to almost-everywhere continuity (Rudin 1976, Theorem 11.33 (b))] and that the partition defined by the $p_{i,n}$ is arbitrarily fine for n large enough; that is, putting $p_{0,n} := \underline{p}$ and $p_{\kappa_n+1,n} := \overline{p}$, we have $\max_{1\leq i\leq\kappa_n+1}(p_{i,n} - p_{i-1,n}) = \ell_n \to 0$. This shows that $\lim_{n\to\infty}\sum_{i=1}^{\kappa_n} f_n(p_{i,n})\ell_n = \int_{\mathcal{I}} f(p)\mathrm{d}p$,, i.e., the denominator in (12) equals $s_n\ell_n^{-1}\int_{\mathcal{I}} f(p)\mathrm{d}p(1+o(1))$. An analogous argument shows that the enumerator equals $s_n\ell_n^{-1}\int_{\mathcal{I}} q(p)\mathrm{d}p(1+o(1))$, and the proof is complete. □

Remark 2 The condition $q(p) > 0$ in (b) excludes only the trivial case $q(p) = 0$; here, the mean square error is zero for both estimators.

Remark 3 To explain the efficiency gain intuitively, we first decompose the MSE of the sales-only estimator as was done for the MLE and then compare. As earlier, $\mathbf{A}_n = (A_{1,n}, \ldots, A_{\kappa_n,n})$ is the vector of price-specific arrival counts observed under Method A $(1, \kappa_n)$, which are independent Poisson(r_n). Without loss of generality, let $A_{1,n}$ and $S_{1,n}$ be the arrival and sale count associated to (any) target price p. For the sales-only σ_2, observe: $\mathbb{E}[\widehat{\lambda q}_n(p)|\mathbf{A}_n] = \mathbb{E}[S_{1,n}/(1/\kappa_n)|\mathbf{A}_n] \overset{(a)}{=} \mathbb{E}[S_{1,n}/(1/\kappa_n)|A_{1,n}] = \kappa_n A_{1,n} q(p)$, with step (a) due to the independence in Property P2. Thus (note the bias is zero): $\sigma^S_{2,n}(p) := \mathbb{V}ar(\mathbb{E}[\widehat{\lambda q}_n(p)|\mathbf{A}_n]) = \mathbb{V}ar(\kappa_n A_{1,n} q(p)) = r_n \kappa_n^2 q^2(p)$ which is asymptotically κ_n times larger than the MLE analog, in view of (14). To see this more intuitively, note that Assumption 1 implies that *all the arrival data* $\mathbf{A}_n$ contain information about the arrival rate; the sales-only estimator's conditional expectation given $\mathbf{A}_n$, that is $\kappa_n A_{1,n} q(p)$, uses just one of them ($A_{1,n}$ occurs over $1/\kappa_n$ time units, hence the scaling up by κ_n); in contrast, the MLE analog, that is $\widehat{\lambda}_n \mathbb{1}_{[A_{1,n}>0]} q(p)$, uses all of them, and due to this its variance is asymptotically κ_n times smaller (the MLE's bias is asymptotically negligible; see (14)). For the sales-only σ_1, observe: $\mathbb{V}ar(\widehat{\lambda q}_n(p)|\mathbf{A}_n) \overset{(a)}{=} \mathbb{V}ar(\widehat{\lambda q}_n(p)|A_{1,n}) = f(p) A_{1,n} \kappa_n^2$, with step (a) again due to the independence in Property P2; thus $\sigma^S_{1,n}(p) := \mathbb{E}[\mathbb{V}ar(\widehat{\lambda q}_n(p)|\mathbf{A}_n)] = f(p) r_n \kappa_n^2$, and $\sigma_{1,n}(p)/\sigma^S_{1,n}(p) \sim 1$, in view of (13); the MLE provides no benefit here as a consequence of the conditional independence of sale counts (at distinct prices) given $\mathbf{A}_n$ (Property P2).

Example 1 Let $q(p) = a - bp$, $p \in [0, 1]$, where $a > 0$, $b \geq 0$, and one requires $q(0) = a \leq 1$ and $q(1) = a - b \geq 0$. By (12), $\mu_\infty = \mu_\infty(a, b) = 1/[1 - c - b^2/(12c)] > 1$, where $c = a - b/2 > 0$. For example, $q(\underline{p}) = 1$ and $q(\overline{p}) = 0$ gives $\mu_\infty = 3$, while $q(\underline{p}) = 1$ and $q(\overline{p}) = 1/2$ gives $\mu_\infty = 9/2$.

5 Results on pricing

5.1 Regret upper bound in the presence of an arrival rate (the class $\mathcal{L}$)

For the n-th problem, we put $J^\pi_n := J^\pi(nx, T|n\lambda, q)$ for the expected revenue under policy π (Algorithm AS); and we put $J^D_n := J^D(nx, T|n\lambda, q)$, which is easily seen to be $nJ^D(x, T|\lambda, q)$. Our main result follows.

Theorem 1 *Define* $\pi_n := \pi(\tau_n, \kappa_n)$ *by Algorithm AS. If* $\tau_n \asymp (\log(n)/n)^{1/4}$ *and* $\kappa_n \asymp (n/\log(n))^{1/4}$*, then for some constant* K_0 *and some finite* $\underline{n}$*,*

$$\inf_{(\lambda,q)\in\mathcal{L}} \frac{J^{\pi_n}_n}{nJ^D} \geq 1 - K_0 \left(\frac{\log n}{n}\right)^{1/4} \quad for\ all\ n \geq \underline{n}. \tag{15}$$

For each n and $i = 1, \ldots, \kappa_n$, let $p = p_{i,n}$ denote the i-th price applied in the learning step of Algorithm $\pi(\tau_n, \kappa_n)$. Let $\overline{A}_n \sim$ Poisson$(\lambda n \tau_n)$; $A_{i,n} \sim$ Poisson(λr_n), where $r_n := n\tau_n/\kappa_n$; and $S_{i,n}|A_{i,n} \sim$ Binomial$(A_{i,n}, q(p))$ independently of $A_{i,n}$. The estimates of λ_n and $q(p)$ have the representation $\widehat{\lambda}_n = \mathbb{1}_{[\overline{A}_n<nx]}\overline{A}_n/(n\tau_n) +$

$\mathbb{1}_{[\overline{A}_n \geq nx]} Z$ and $\widehat{q}_n(p) = I S_{i,n} A_{i,n}^{-1} + (1-I) Z$ respectively, where $I = \mathbb{1}_{[\overline{A}_n < nx, A_{i,n} > 0]}$ and Z is a suitable random variable in each case.

A key intermediate result is Lemma 1 below; it bounds the probability of certain estimation errors and is based on the following.

Condition 1 *Let* $\lim_{n\to\infty} \kappa_n = \infty$. *Moreover assume: (a) For some* $\underline{c}_\tau, \overline{c}_\tau > 0$ *and* $0 < \psi_1 \leq \psi_2 < 1$, $\underline{c}_\tau n^{\psi_1} \leq n\tau_n \leq \overline{c}_\tau n^{\psi_2}$ *for all* $n \geq 1$*; and (b) For some* $\underline{c}_r > 0$ *and* $\beta > 0$, $r_n := n\tau_n/\kappa_n \geq \underline{c}_r n^\beta$ *for all* $n \geq 1$.

Lemma 1 *Fix* $\eta \geq 2$, *define* $r_n := n\tau_n/\kappa_n$, *and let Condition 1 hold. Let* $\delta_n = \delta_n(\lambda) = \left[4\eta\lambda \log n/(n\tau_n)\right]^{1/2}$*; and* $l_n = l_n(\lambda) = r_n(\lambda - \zeta_n)$, *where* $\zeta_n = \zeta_n(\lambda) = \left[2\eta\lambda \log(n)/r_n\right]^{1/2}$. *Put* $\mathcal{C} = [\underline{p}, \overline{p}]$. *Then*

(a) *For any finite* $\lambda > 0$, *there exists a finite* $n_0 = n_0(\lambda)$ *such that for all* $n \geq n_0(\lambda)$ *and for* $\epsilon_n = \epsilon_n(\lambda) = \left[\eta \log n/(2l_n)\right]^{1/2}$, *we have*

$$\inf_{p\in\mathcal{C}} \mathbb{P}\{|\widehat{\lambda}_n \widehat{q}_n(p) - \lambda q(p)| \leq \delta_n + \lambda\epsilon_n\} \geq 1 - C_1 n^{-\eta} \tag{16}$$

where $C_1 = C_1(\lambda) := 2C_0(\lambda) + 3$, *where* $C_0(\lambda) = \max\{1, [4\eta/(\lambda\beta e)]^{\eta/\beta}\}$.

(b) *Put* $C_1 = C_1(\underline{\lambda})$. *For any* $\alpha > 0$, *there exists a finite* $\underline{n} = \underline{n}(\alpha)$ *such that for all* $n \geq \underline{n}$,

$$\sup_{(\lambda,q)\in\mathcal{L}} \sup_{p\in\mathcal{C}} \mathbb{P}\left\{|\widehat{\lambda}_n \widehat{q}_n(p) - \lambda q(p)| > (1+\alpha)\left(\frac{\overline{\lambda}\eta \log n}{2r_n}\right)^{1/2}\right\} \leq C_1 n^{-\eta}, \tag{17}$$

$$\sup_{(\lambda,q)\in\mathcal{L}} \sup_{p\in\mathcal{C}} \mathbb{P}\left\{|\widehat{q}_n(p) - q(p)| > (1+\alpha)\left(\frac{\eta \log n}{2\underline{\lambda} r_n}\right)^{1/2}\right\} \leq C_1 n^{-\eta}. \tag{18}$$

Proof of Lemma 1 First, we construct $n_1(\lambda)$ and $n_2(\lambda)$ such that

$$l_n(\lambda) > 0 \quad \text{for all } n \geq n_1(\lambda) \ \text{ and } \ n\tau_n(\lambda + \delta_n) < nx \quad \text{for all } n \geq n_2(\lambda). \tag{19}$$

Since $r_n \geq \underline{c}_r n^\beta$, the condition $\log(n)/n^\beta < \underline{c}_r^{-1}\lambda/(2\eta)$ implies $l_n > 0$. Since $\log(n)/n^\beta$ is maximized at $x^* = e^{1/\beta}$ and decreases to zero for $x \in [x^*, \infty)$, we may set $n_1(\lambda) := \min\{n : n \geq e^{1/\beta},\ \log(n)/n^\beta < \underline{c}_r^{-1}\lambda/(2\eta)\} < \infty$. By simple calculation, $n_2(\lambda) := \lceil \{\overline{c}_\tau(\lambda + [4\eta\lambda/(e\psi_1 \underline{c}_\tau)]^{1/2})/x\}^{1/(1-\psi_2)} \rceil < \infty$ suffices.

Proof of (a). Let p denote the i-th price applied during learning, i.e., $p = p_{i,n}$ for some $i \in \{1, \ldots, \kappa_n\}$. Define the events

$$U_n = \left\{\left|\frac{\overline{A}_n}{n\tau_n} - \lambda\right| \leq \delta_n\right\},\ L_n = \{A_{i,n} \geq l_n\},\ D_n = \left\{\left|\frac{S_{i,n}}{A_{i,n}} - q(p)\right| \leq \epsilon_n\right\}. \tag{20}$$

Put $G_n := U_n \cap L_n \cap D_n$. Then

$$G_n \overset{(a)}{\subseteq} \{|\widehat{\lambda}_n - \lambda| \le \delta_n, |\widehat{q}_n(p) - q(p)| \le \epsilon_n\} \overset{(b)}{\subseteq} \{|\widehat{\lambda}_n \widehat{q}_n(p) - \lambda q(p)| \le \delta_n + \lambda\epsilon_n\} \quad (21)$$

for $n \ge n_0(\lambda)$, where $n_0(\lambda) := \max\{n_1(\lambda), n_2(\lambda)\} < \infty$. Step (a) above uses that $\overline{A}_n \le n\tau_n(\lambda + \delta_n) < nx$ and $l_n > 0$, which exclude a stock-out and a no-arrivals event, respectively; hence on the set G_n we have $|\widehat{\lambda}_n - \lambda| = |\overline{A}_n/(n\tau_n) - \lambda| \le \delta_n$ and $|\widehat{q}_n(p) - q(p)| = |S_{i,n}A_{i,n}^{-1} - q(p)| \le \epsilon_n$. Step (b) above uses the triangle inequality $|\widehat{\lambda}_n \widehat{q}_n(p) - \lambda q(p)| \le |\widehat{\lambda}_n - \lambda| + \lambda|\widehat{q}_n(p) - q(p)|$. Now:

$$\mathbb{P}({G_n}^c) \le \mathbb{P}({U_n}^c) + \mathbb{P}({L_n}^c) + \mathbb{P}(L_n \cap {D_n}^c), \quad (22)$$

$$\mathbb{P}({U_n}^c) \le 2C_0(\lambda)n^{-\eta} \quad \text{and} \quad \mathbb{P}({L_n}^c) \le n^{-\eta} \quad \text{for } n \ge 1, \quad (23)$$

$$\mathbb{P}(L_n \cap {D_n}^c) = \mathbb{P}\left\{A_{i,n} \ge l_n, \left|\frac{S_{i,n}}{A_{i,n}} - q(p)\right| > \epsilon_n\right\} \overset{(a)}{\le} 2e^{-2l_n\epsilon_n^2}$$
$$= 2n^{-\eta} \text{ for } n \ge n_1(\lambda) \quad (24)$$

where (23) uses Lemma 2 (Poisson large-deviation bound); and step (a) in (24) uses Lemma 3 (Hoeffding's inequality; the number of summands is $A_{i,n} \ge \max(l_n, 1)$ since $A_{i,n} \ge l_n > 0$ is integer-valued). From (22), (23), and (24), we obtain $\mathbb{P}({G_n}^c) \le (2C_0(\lambda) + 3)n^{-\eta}$, and note that the same $n_0(\lambda)$ in (21) suffices for any $p \in \mathcal{C}$. Thus, part (a) is proven.

Proof of (b). Letting $\theta_n := \sup_{(\lambda,q)\in\mathcal{L}} [\delta_n(\lambda) + \lambda\epsilon_n(\lambda)]$, (16) gives

$$\sup_{(\lambda,q)\in\mathcal{L}} \sup_{p\in\mathcal{C}} \mathbb{P}\left\{|\widehat{\lambda}_n \widehat{q}_n(p) - \lambda q(p)| > \theta_n\right\} \le \sup_{(\lambda,q)\in\mathcal{L}} C_1(\lambda)n^{-\eta} = C_1(\underline{\lambda})n^{-\eta}$$

for all $n \ge m$, where $m := \max\{\sup_{\lambda\in\mathcal{L}} n_1(\lambda), \sup_{\lambda\in\mathcal{L}} n_2(\lambda)\} = \max\{n_1(\underline{\lambda}), n_2(\overline{\lambda})\}$. Now

$$\frac{\theta_n}{\sqrt{\log n}} = \sup_{(\lambda,q)\in\mathcal{L}} \left[2\sqrt{\frac{\eta\lambda}{n\tau_n}} + \sqrt{\frac{\eta\lambda}{2r_n}}\left(1 - \sqrt{\frac{2\eta\log n}{\lambda r_n}}\right)^{-1/2}\right] \overset{(a)}{\sim} \sqrt{\frac{\overline{\lambda}\eta}{2r_n}}$$

where step (a) uses that the first term vanishes relative to the second one (since $n\tau_n/r_n = k_n \to \infty$); and that $(1 - \sqrt{(2\eta\log n)/(\lambda r_n)})^{-1/2} \sim 1$. This proves (17). Result (18) follows by an analogous argument. □

Proof of Theorem 1 By construction, there exist positive $\underline{c}_\tau, \overline{c}_\tau, \underline{c}_\kappa, \overline{c}_\kappa$ such that $\underline{c}_\tau(\log n/n)^{1/4} \le \tau_n \le \overline{c}_\tau(\log n/n)^{1/4}$ and $\underline{c}_\kappa(n/\log n)^{1/4} \le \kappa_n \le \overline{c}_\kappa(n/\log n)^{1/4}$ for all $n \ge 1$. Condition 1 is in force: in part (a) use $\psi_1 = 3/4$ and $\psi_2 > 3/4$; in part (b) use $\beta = 1/2$. Thus, Lemma 1 is in force. Let $\underline{n} = \underline{n}(\alpha)$ be as in Lemma 1(b). The proof now parallels that of Besbes and Zeevi (2009, Proposition 1), but the regret will be bounded by

$$u_n := \max\left\{\tau_n, \kappa_n^{-1}, \left(\frac{\log n}{r_n}\right)^{1/2}\right\}, \quad (25)$$

whose order $O((\log n/n)^{1/4})$ is smaller than theirs. Lemma 1 replaces Besbes and Zeevi (2009, Online Companion, Lemma 2). Put $X_n^{(L)} = \lambda n \tau_n \kappa_n^{-1} \sum_{i=1}^{\kappa_n} q(p_i)$ and $X_n^{(P)} = \lambda q(\widehat{p}) n(T-\tau_n)$; put $Y_n^{(L)} = N(X_n^{(L)})$, $Y_n^{(P)} = N(X_n^{(P)})$, and $Y_n = N(X_n^{(L)} + X_n^{(P)})$ ($Y_n^{(L)}$, $Y_n^{(P)}$, and Y_n are the would-be sales during learning, pricing, and overall, respectively, if inventory were infinite).

Step 1 The revenue achieved by π_n is bounded below by the revenue achieved in the pricing phase; during this phase, the number of units sold is $\min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\}$. It follows that

$$J_n^{\pi_n} \geq \mathbb{E}[\widehat{p} \min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\}]. \tag{26}$$

Step 2 We separate two cases: $\Lambda(\overline{p}) \leq x/T$ and $\Lambda(\overline{p}) > x/T$.

Case 1 $\Lambda(\overline{p}) \leq x/T$. We will show that $J_n^{\pi_n}$ is at least nJ^D minus an $O(u_n)$ term. Since $\min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\} \geq Y_n^{(P)} - (Y_n - nx)^+$, (26) gives

$$J_n^{\pi_n} \geq \mathbb{E}[\widehat{p} Y_n^{(P)}] - \overline{p}\mathbb{E}[(Y_n - nx)^+] = \mathbb{E}[r(\widehat{p})]\lambda n(T-\tau_n) - \overline{p}\mathbb{E}[(Y_n - nx)^+], \tag{27}$$

using $\widehat{p} \leq \overline{p}$ in the first step and $\mathbb{E}[\widehat{p} Y_n^{(P)}] = \mathbb{E}[r(\widehat{p})]\lambda n(T-\tau_n)$ in the second step. Lemma 7 shows that $\mathbb{E}[r(\widehat{p})] \geq r(p^D) - Ru_n - R_2/n^{\eta-1}$ for positive constants R, R_2 and for $n \geq \underline{n}$. Moreover, Lemma 9 shows that $\mathbb{E}[(Y_n - nx)^+] \leq K_E n u_n$ for some constant $K_E > 0$ and all $n \geq \underline{n}$. Thus

$$\begin{aligned} J_n^{\pi_n} &\geq \left[\lambda r(p^D) - Ru_n - \frac{R_2}{n^{\eta-1}}\right] n(T-\tau_n) - \overline{p} K_E n u_n \\ &\geq n\lambda r(p^D)T - K_1 n(u_n + \tau_n) = nJ^D - K_1 n(u_n+\tau_n) \text{ for all } n \geq \underline{n} \end{aligned} \tag{28}$$

for a constant K_1. By Assumption 2(c), $J^D \geq m^D$ for some $m^D > 0$, and thus

$$\frac{J_n^{\pi_n}}{J_n^D} = \frac{J_n^{\pi_n}}{nJ^D} \geq 1 - \frac{K_1}{J^D}(u_n+\tau_n) \geq 1 - \frac{K_1}{m^D}(u_n+\tau_n). \tag{29}$$

Case 2 $\Lambda(\overline{p}) > x/T$. Here, $p^D = \overline{p}$ and $J_n^D = n\overline{p}x$. Lemma 10 shows that $\min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\}$ (a lower bound on the quantity sold during the pricing phase) is close to nx and also $\widehat{p}$ is close to p^D, with high probability. Specifically, put $\mathcal{E} := \{\omega : \min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\} \geq nx - \tilde{K}_Y n u_n, |\widehat{p} - p^D| \leq K_c u_n\}$ where $\tilde{K}_Y$ and K_c are constants defined in Lemma 10. The mean revenue generated by π_n is bounded as follows:

$$\begin{aligned} J_n^{\pi_n} &\geq \mathbb{E}[\widehat{p} \min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\}] \\ &\overset{(a)}{\geq} \mathbb{E}[(p^D - K_c u_n) \min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\} | \mathcal{E}]\mathbb{P}(\mathcal{E}) \\ &\overset{(b)}{\geq} (p^D - K_c u_n)(nx - \tilde{K}_Y n u_n)\left(1 - \frac{C_2}{n^{\eta-1}}\right) \\ &\geq p^D nx - K_2 n u_n \quad \text{for all } n \geq \underline{n}, \end{aligned}$$

where both (a) and (b) follow from the definition of $\mathcal{E}$ and Lemma 10, and where C_2 and K_2 are suitable constants. We conclude that

$$\frac{J_n^\pi}{J_n^D} = \frac{J_n^\pi}{nJ^D} \geq 1 - \frac{K_2}{\underline{p}x} u_n \quad \text{for all } n \geq \underline{n}. \tag{30}$$

Step 3 The result (15) now follows from (29) and (30). □

Remark 4 We now provide nearly-explicit formulae for $n_1(\lambda)$ and $n_2(\lambda)$ in (19). In view of Theorem 1, it is optimal to set $\tau_n = c_\tau(\log n/n)^{1/4}$ and $\kappa_n \sim c_\kappa(n/\log n)^{1/4}$ for some positive c_τ, c_κ; for simplicity, we now ignore the integrality of κ_n and equate it to $c_\kappa(n/\log n)^{1/4}$. Substituting into δ_n and r_n, it is easy to see that $n_1(\lambda) := \min\{n : n \geq e,\ \log n/n < [(c_\kappa/c_\tau)^{-1}\lambda/(2\eta)]^2\}$ suffices. Moreover, since δ_n is decreasing for $n \geq e$, it is easy to see that $n_2(\lambda) := \min\{n : n \geq e, (\log n/n) < x/[c_\tau(\lambda + \delta_n)]^4\}$ suffices; a simpler but weaker formula for $n_2(\lambda)$ replaces δ_n by $\sup_n \delta_n = (4\eta\lambda c_\tau^{-1} e^{-3/4})^{1/2}$.

5.2 Besbes and Zeevi (2009) revisited: an improved convergence rate

We consider the setting in Besbes and Zeevi (2009). The demand function $\Lambda(p)$ is postulated directly, without reference to arrivals.

Theorem 2 *Let $\mathcal{L}_{BZ}$ be the demand class defined in Besbes and Zeevi (2009, Section 4.2), and let π_n' be given by Algorithm $\pi(\tau,\kappa)$ in Section 4.1 there, with $\tau = \tau_n$ and $\kappa = \kappa_n$. If $\tau_n \asymp (\log(n)/n)^{1/4}$ and $\kappa_n \asymp (n/\log(n))^{1/4}$, then there exists a constant K_0' such that for all $n \geq 1$,*

$$\inf_{\Lambda(\cdot)\in\mathcal{L}_{BZ}} \frac{J_n^{\pi_n'}}{nJ^D} \geq 1 - K_0'\left(\frac{\log n}{n}\right)^{1/4}.$$

Proof of Theorem 2 In analogy to Theorem 1, probability bounds are obtained for events defined via deviations proportional to a u_n defined as in (25) and such that $u_n \asymp (\log n/n)^{1/4}$. For example, to bound the error $|\widehat{p}^c - p^c|$, let the demand function $\Lambda(\cdot)$ satisfy $\underline{M}_\Lambda'|p_1 - p_2| \leq |\Lambda(p_1) - \Lambda(p_2)| \leq \overline{M}_\Lambda'|p_1 - p_2|$ for all p_1, p_2 in the price domain $[\underline{p}, \overline{p}]$; put $M = \sup_p \Lambda(p)$ and $K_c' = 4\underline{M}_\Lambda'^{-1}\max\{c_1, c_2\}$, where: $c_1 := \overline{M}_\Lambda'(\overline{p} - \underline{p})/2$ and $c_2 := 2(\eta M)^{1/2}$, for some $\eta \geq 2$. The proof proceeds as in Lemma 6; in step (e) the Poisson large-deviation bound (Lemma 2) replaces (18); we obtain, for all $n \geq 1$, $\mathbb{P}\{|\widehat{p}^c - p^c| > K_c'u_n\} \leq C_0(M)n^{-\eta+1}$. We proceed as follows: the constants multiplying u_n to form the deviation events in Lemmata 5, 6, 7, 8, 9, and 10 are increased, when necessary, by a factor no larger than $2\sqrt{2}$; these deviations are no less than a positive constant times $(\log n/r_n)^{1/2}$ (by $(\log n/r_n)^{1/2} \asymp (\log n/n)^{1/4}$); then, by Lemma 2, these deviation events have probability at most a constant times $n^{-\eta+1}$ for all n. The proof then closely parallels that of Theorem 1. □

In comparison to Besbes and Zeevi (2009, Proposition 1), τ_n is larger by a factor $(\log n)^{1/4}$; κ_n is smaller by a factor $(\log n)^{-1/4}$; and the order of the regret upper bound is smaller, by the factor $(\log n)^{-1/4}$.

5.3 Lower bound on regret

Example 2 Let $\underline{p} = 1/2$, $\overline{p} = 3/2$, $x = 2$ and $T = 1$. For any $\lambda_0 \in [\underline{\lambda}, \overline{\lambda}]$, define $\mathcal{M}$ as the family of demand models as in Assumption 1 with arrival rate λ_0 and purchase probability $q(p) = q(p; z) = 1/2 + p(1-z)$, where z is a parameter taking values in $Z = [\underline{z}, \overline{z}] = [1/3, 2/3]$.

For any z, $r(p; z) := pq(p; z)$ is the revenue rate per arrival, and $p^D(z)$ is the optimal price under z. For any z and any admissible policy π ($\pi \in \mathcal{P}$), the regret for the scale-n problem is abbreviated as $\mathcal{R}_n^\pi(z) := 1 - J_n^\pi(z)/(nJ^D(z))$, where $J_n^\pi(z)$ and $nJ^D(z)$ are the scale-n expected revenue under π, and that of the deterministic relaxation, respectively, under z.

Theorem 3 (Lower-bound example)

(a) *The family $\mathcal{M}$ is contained in $\mathcal{L}$, i.e., $\mathcal{M} \subseteq \mathcal{L}$.*

(b) *Let the n-th problem instance have arrival rate $\lambda_n = n\lambda_0$ for all $n \geq 1$. For any admissible pricing policy π and all $n \geq 1$,*

$$\sup_{z \in Z} \mathcal{R}_n^\pi(z) \geq \frac{K_3}{\sqrt{n}}$$

where K_3 is a positive constant that may depend on λ_0.

Proof of Theorem 3 The proof of (a) is a simple verification that we omit, so we now focus on proving (b). Consider an arbitrary admissible policy π and let ψ_t denote the associated price at time t. By its admissibility, the stochastic process $\{\psi_t\}$ is adapted to $\mathcal{F}_t$. Let z_1, z_2 be any elements of Z and let $t \in [0, T]$. For $i = 1, 2$, let $\mathbb{P}_{z_i}^{\pi(t)}$ denote the probability measure induced by π up to time t (i.e., describing the process $\{(A_s, D_s) : 0 \leq s \leq t\}$) when $z = z_i$ and $\mathbb{E}_{z_i}^{\pi(t)}$ denote the corresponding expectation. Putting

$$Q_{i,t} := q(\psi_t; z_i), \quad \xi_t := Q_{1,t}/Q_{2,t}, \quad I_t := 1 - \xi_t + \xi_t \log \xi_t$$

(these are stochastic processes adapted to $\mathcal{F}_t$), the Kullback-Leibler (KL) divergence between the two measures $\mathbb{P}_{z_1}^{\pi(t)}$ and $\mathbb{P}_{z_2}^{\pi(t)}$ is

$$\begin{aligned}
\mathcal{K}(\mathbb{P}_{z_1}^{\pi(t)}, \mathbb{P}_{z_2}^{\pi(t)}) &= \mathbb{E}_{z_1}^{\pi(t)}\left[\log \frac{\mathrm{d}\mathbb{P}_{z_1}^{\pi(t)}}{\mathrm{d}\mathbb{P}_{z_2}^{\pi(t)}}\right] \\
&= \mathbb{E}_{z_1}^{\pi(t)}\left[-\int_0^t n\lambda_0(Q_{1,s} - Q_{2,s})\mathrm{d}s + \int_0^t \log \xi_s \mathrm{d}D_s\right] \\
&= n\lambda_0 \mathbb{E}_{z_1}^{\pi(t)}\left[\int_0^t Q_{2,s}\,(1 - \xi_s + \xi_s \log \xi_s)\,\mathrm{d}s\right] \qquad (31)
\end{aligned}$$

where $\mathrm{d}\mathbb{P}_{z_1}^{\pi(t)}/\mathrm{d}\mathbb{P}_{z_2}^{\pi(t)}$ is the Radon–Nikodym derivative of $\mathbb{P}_{z_1}^{\pi(t)}$ with respect to $\mathbb{P}_{z_2}^{\pi(t)}$, and the last step uses that $\mathbb{E}_{z_1}^{\pi(t)} \int_0^t \log \xi_s \mathrm{d}D_s = \mathbb{E}_{z_1}^{\pi(t)} \int_0^t (\log \xi_s) n\lambda_0 Q_{1,s} \mathrm{d}s$; see Brémaud (1981) for background. When $t = T$, we write $\mathbb{P}_z^\pi$ instead of $\mathbb{P}_z^{\pi(t)}$.

Claim 1. For $z_1 = 1/2$, any $z_2 \in Z$, and any policy $\pi \in \mathcal{P}$ we have

$$\mathcal{K}(\mathbb{P}^{\pi}_{z_1}, \mathbb{P}^{\pi}_{z_2}) \leq 9n\lambda_0(z_1 - z_2)^2 \mathcal{R}^{\pi}_n(z_1). \tag{32}$$

Proof of Claim 1. Observe that $\xi_s = 1 + \epsilon$, where $\epsilon = \epsilon_s = (z_1 - z_2)(1 - \psi_s)Q^{-1}_{2,s}$. Note that $I_s = 1 - (1 + \epsilon) + (1 + \epsilon)\log(1 + \epsilon) \leq -\epsilon + (1 + \epsilon)\epsilon = \epsilon^2$, where the inequality uses $\underline{q} := \inf_z \inf_p q(p; z) = 1/6$; $|\epsilon| \leq |z_1 - z_2| \cdot |1 - \psi_s| \cdot \underline{q}^{-1} \leq (1/6) \cdot (1/2) \cdot 6 = 1/2$; and $\log(1 + \epsilon) \leq \epsilon$. Now we bound the integrand in (31): $Q_{2,s} I_s \leq Q_{2,s}\epsilon^2 = Q^{-1}_{2,s}(z_1 - z_2)^2(1 - \psi_s)^2 \leq \underline{q}^{-1}(z_1 - z_2)^2(1 - \psi_s)^2$; thus

$$\mathcal{K}(\mathbb{P}^{\pi}_{z_1}, \mathbb{P}^{\pi}_{z_2}) \leq n\lambda_0 \underline{q}^{-1}(z_1 - z_2)^2 \mathbb{E}^{\pi}_{z_1}\left[\int_0^T (1 - \psi_s)^2 \mathrm{d}s\right]. \tag{33}$$

Moreover, when z equals $z_1 = 1/2$, the optimal price is $p^D := p^D(z_1) = 1$, and

$$\mathcal{R}^{\pi}_n(z_1) \overset{(a)}{\geq} \frac{\mathbb{E}^{\pi}_{z_1} \int_0^T n\lambda_0[r(p^D; z_1) - r(\psi_s; z_1)]\mathrm{d}s}{\mathbb{E}^{\pi}_{z_1} \int_0^T n\lambda_0 r(p^D; z_1)\mathrm{d}s} \overset{(b)}{\geq} \frac{2}{3}\mathbb{E}^{\pi}_{z_1}\left[\int_0^T (1 - \psi_s)^2 \mathrm{d}s\right];$$

step (a) follows from the fact that the inventory constraint is being relaxed; step (b) uses that $r(p^D; z) - r(p; z) \geq (1/2)K(p^D - p)^2$ for any p and z, where $K := \inf_z \inf_p |r''(p; z)| = 2/3$; this follows from Taylor's theorem with order-two Lagrange remainder (since $r'(p^D; z) = 0$ and $r''(p; z) = -2z$ for all p and z, with primes denoting derivatives with respect to p); step (b) also notes that $r(p^D; z_1) = 1/2$. This and (33) prove Claim 1, which is analogous to Wang et al. (2014, Lemma 9). The proof onwards is fully analogous to that of Wang et al. (2014, Lemma 10 and Theorem 2) and omitted. □

Remark 5 The KL divergence in (31) involves a change of intensity for a Poisson process. Below we motivate this formula; note that Besbes and Zeevi (2012), Wang et al. (2014) start with a KL formula claimed from Brémaud (1981), but no specific reference is given. The fixed arrival rate λ_0 implies that the Radon-Nikodym derivative (likelihood ratio) *with respect to the arrival element in the sample space*, $\{A_s : 0 \leq s \leq t\}$, contributes nothing (the log-likelihood ratio is zero). The change of intensity is thus captured through the paths of the demand process only, $\{D_s : 0 \leq s \leq t\}$, with $q(\cdot; z_1)$ and $q(\cdot; z_2)$ appearing exactly as in these works, while the intensity of the base measure $\mathbb{P}^{\pi(t)}_{z_1}$ brings in λ_0 as multiplier.

Remark 6 In Claim 1, the constant 9 improves the constant of Wang et al. (2014, Lemma 9), which is 24.

Theorem 3 is parallel to Wang et al. (2014, Theorem 2) and Broder and Rusmevichientong (2012, Theorem 3.1); it replaces their demand function by our purchase probability function; their insights carry over here. Specifically, the functions q in $\mathcal{M}$ take the same value at the price $p = 1$: $q(1; z) = 1/2$ for all $z \in Z$. Such a price is called an *uninformative* price (Broder and Rusmevichientong 2012). Whenever one prices at this uninformative price, there is no gain in information about q. In order to learn the function q (i.e., the parameter z) and determine the optimal price, p^D, one must (for at least some time) set a price other than the uninformative one; on the other hand, when p^D and the uninformative price coincide (i.e., z equals $z_1 = 1/2$), pricing anywhere other than at p^D incurs revenue losses. This tension is reflected in the lower bound, which reflects lower bounds on error probabilities in hypothesis testing (Tsybakov 2009) that are fully analogous to those in Besbes and Zeevi (2012), Broder and Rusmevichientong (2012), Wang et al. (2014).

6 Numerical results

We compare policies that we index as follows: (1) our policy (Algorithm AS) (AS); (2) the policy in Besbes and Zeevi (2009, Section 4.1) (BZ); two variants of the policy in Wang et al. (2014, Section 7.1) with respect to the price interval that initializes their step 3: (3) the interval from step 2 (W0); (4) the interval $[\underline{p}, \overline{p}]$ (W1); and (5) policy BZ modified as in Theorem 2 (BZ-M).

In constructing test prices, the midpoints of relevant intervals are used; these work slightly better then left- or right-endpoints.

We revisit Wang et al. (2014, Table 1) so that regret numbers are comparable. Thus, we fix the initial inventory $x = 20$; selling horizon $T = 1$; and feasible price set $[\underline{p}, \overline{p}] = [0.1, 10]$. A pair (λ, q) (in $\mathcal{L}$) and a corresponding demand function $\Lambda(p) = \lambda q(p)$ are drawn randomly from one of two families:

- Linear: $\Lambda(p) = \lambda - \alpha p$ and $q(p) = 1 - (\alpha/\lambda)p$, with $\lambda \in [20, 30]$, $\alpha \in [2, 10]$.
- Exponential: $\Lambda(p) = \lambda e^{-\beta p}$ and $q(p) = e^{-\beta p}$, with $\lambda \in [40, 80]$, $\beta \in [1/3, 1]$.

The probability law is uniform for each parameter; to match their law in the linear case, uniform sampling applies to λ and α, not to α/λ.

Existing theory describes an optimal growth function for each policy parameter (phase duration, number of test prices, etc.); for example, for AS, $\tau_n \asymp f(n)$ where $f(n) = (\log n/n)^{1/4}$. Let $c > 0$ be any *scaling (constant), fixed for all* n. Note $f(cn) \sim c^{-1/4} f(n)$ as $n \to \infty$; for each parameter, a similar asymptotic applies, so that its growth *order* (and that of the regret) is unchanged when n is replaced by cn. We find that (mean-regret) sensitivity to c is substantial. Therefore, each parameter is set as $f(c_i x_n)$, where $c_i > 0$ is a variable; $x_n = 20n$ is the n-th initial inventory; and f is the optimal growth function for that parameter; see our theorems and Besbes and Zeevi (2009), Wang et al. (2014). Figure 1 shows the (estimated mean) regret as c_i varies relative to a reference value c_i^* found as the near-minimizer of the regret for $n = 10^6$ (for that case). We use: $c_1^* = c_5^* = 100$; $c_2^* = 5$; $c_3^* = c_4^* = 1$ (for linear family); and $c_3^* = c_4^* = 1/4$ (for exponential family). We see that suboptimal scaling affects policies W0 and W1 more than the others. The result $(c_1^*/c_2^*)^{-1/4} \approx 47\%$ indicates a learning-time ratio (policies AS to BZ) reduction relative to $c_1 = c_2$, and is consistent with policy AS having higher estimation efficiency. Table 1 shows the near-optimal regret ($c_i = c_i^*$); standard errors are $< 2\%$ for policies W0 and W1 for $n = 10^6$; and $< 1\%$ otherwise.

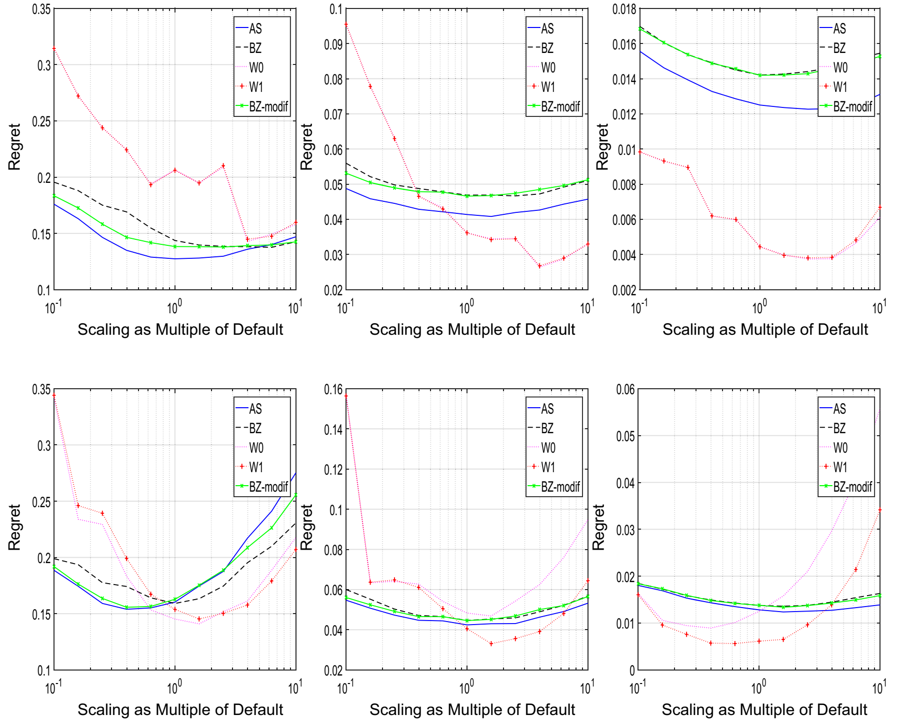

Fig. 1 Sensitivity of the mean regret to the scaling constant c_i for each policy i (linear and exponential family in top and bottom row, respectively) for $n \in \{10^2, 10^4, 10^6\}$ (from left to right). The x-axis is the ratio c_i/c_i^*; it spans [1/10, 10] in logarithmic scale

Table 1 Comparison of mean regret of the policies with near-optimal scaling ($c_i = c_i^*$)

Family	n	AS	BZ	W0	W1	BZ-M
Linear	10^2	0.1287	0.1423	0.2028	0.2030	0.1381
	10^3	0.0745	0.0828	0.1020	0.1021	0.0828
	10^4	0.0413	0.0466	0.0355	0.0354	0.0465
	10^5	0.0225	0.0260	0.0115	0.0116	0.0258
	10^6	0.0124	0.0142	0.0045	0.0045	0.0142
Exponential	10^2	0.1614	0.1549	0.1439	0.1525	0.1639
	10^3	0.0803	0.0831	0.1499	0.1479	0.0831
	10^4	0.0423	0.0446	0.0478	0.0406	0.0446
	10^5	0.0230	0.0243	0.0257	0.0178	0.0244
	10^6	0.0130	0.0137	0.0123	0.0061	0.0135

Our main findings follow. Policies W0 and W1 are superior for large systems, but they risk being inferior if the scaling is far from the optimal one (Fig. 1); we think the latter can happen in practice. With all policies scaled near-optimally, there is a system-size threshold n in $[10^3, 10^4]$ such that policy AS is (modestly) superior below it; as the size increases above this threshold, policy W1 dominates by a margin that increases. This point persists when we focus on policy W0 and take its performance from Wang et al. (2014) (they use near-optimal scaling). Policies BZ and BZ-M perform nearly identically, despite the latter's theoretically faster convergence.

7 Conclusion

Our estimation results apply independently of a pricing problem: start with a Poisson process N of rate λ on some finite interval and a probability $q(p)$, where p is a control. Setting the control at p thins the process N (each event of N is accepted with probability $q(p)$), inducing a thinned process N'. We showed that the empirical rate of N multiplied by the empirical thinning probabilities is a more efficient estimator of the rate of N', that is $\lambda q(p)$, than the empirical rate of N'. In the pricing problem, our regret upper bound improves that in Besbes and Zeevi (2009) by the factor $(\log n)^{-1/4}$ and is the result of refined bounding. Numerically, our method performs better than Wang et al. (2014) for systems that are not very large, and dominates Besbes and Zeevi (2009) regardless of system size; its superior estimation efficiency is behind this. In future work, the arrivals-and-sales estimator could replace the sales-only one commonly used in the literature; indeed, experience with the demand families in Sect. 6 suggests that this reduces the mean regret of the algorithm of Wang et al. (2014) (but not the convergence rate).

8 Proofs

Lemma 2 *Let $N(\cdot)$ be a unit-rate Poisson process. Suppose that $\lambda \in [0, M]$ and $r_n \geq n^\beta$ with $\beta > 0$. Let $\eta > 0$, $\epsilon_n = 2\eta^{1/2}M^{1/2}(\log n/r_n)^{1/2}$, and $C_P = C_P(M) := [4\eta/(M\beta e)]^{\eta/\beta}$. Define the events $\mathcal{A}_n := \{N(\lambda r_n) - \lambda r_n \geq r_n\epsilon_n\}$ and $\mathcal{B}_n := \{N(\lambda r_n) - \lambda r_n < -r_n\tilde{\epsilon}_n\}$, where $\tilde{\epsilon}_n := \epsilon_n/\sqrt{2}$. Then for all $n \geq 1$,*

$$\mathbb{P}(\mathcal{A}_n) \leq \begin{cases} n^{-\eta} & \text{if } \epsilon_n < M \\ C_P n^{-\eta} & \text{if } \epsilon_n \geq M \end{cases} \tag{34}$$

so $\mathbb{P}(\mathcal{A}_n) \leq C_0 n^{-\eta}$ for all $n \geq 1$, where $C_0 := \max\{1, C_P\}$. Moreover,

$$\mathbb{P}(\mathcal{B}_n) \leq n^{-\eta} \quad \text{for all } n \geq 1. \tag{35}$$

Remark 7 Lemma 2 parallels Besbes and Zeevi (2009, Online Companion, Lemma 2) but corrects their constant leading the $n^{-\eta}$ term.

Proof of Lemma 2 For *any nonnegative* sequence $\{\epsilon_n\}$,

$$\mathbb{P}\left\{\frac{N(\lambda r_n)}{r_n} - \lambda \geq \epsilon_n\right\} \overset{(a)}{\leq} e^{-r_n f_*(\lambda+\epsilon_n;\lambda)} \overset{(b)}{\leq} e^{-r_n f_*(M+\epsilon_n;M)}, \tag{36}$$

$$\mathbb{P}\left\{\frac{N(\lambda r_n)}{r_n} - \lambda < -\epsilon_n\right\} \overset{(a)}{\leq} e^{-r_n f_*(\lambda-\epsilon_n;\lambda)} \overset{(b)}{\leq} e^{-r_n f_*(M-\epsilon_n;M)}, \tag{37}$$

where $f_*(x;\lambda) := x\log(x/\lambda)+\lambda-x$, with $x \geq 0$, is the Fenchel-Legendre transform of the logarithmic moment-generating function of the Poisson(λ) law. Step (a) is Cramér's Theorem (Dembo and Zeitouni 1998, Theorem 2.2.3); step (b) notes the exponent's derivative in λ, i.e., $-r_n[\log(1+x/\lambda) - x/\lambda]$, with $x = \epsilon_n$ in (36) and with $x = -\epsilon_n$ in (37), is non-negative. Now a second-order Taylor expansion of $f_*(x;M)$ in x is used; note $f'_*(x;M) := \mathrm{d}f_*(x;M)/\mathrm{d}x = \log(x/M)$; $f''_*(x;M) := \mathrm{d}^2 f_*(x;M)/\mathrm{d}x^2 = 1/x$; and $f_*(M;M) = f'_*(M;M) = 0$. Thus, for some $\xi = \xi_n$ in $[0,\epsilon_n]$, $f_*(M+\epsilon_n;M) = [2(M+\xi)]^{-1}\epsilon_n^2$.

Proof of (34). *Case 1*: $\epsilon_n < M$. Since $\xi \in [0,\epsilon_n]$, we have $\xi < M$ and $[2(M+\xi)]^{-1} > 1/(4M)$; now (36) implies $\mathbb{P}(\mathcal{A}_n) \leq e^{-r_n\epsilon_n^2/(4M)}$; equating this to $n^{-\eta}$ and taking logarithms gives $-r_n\epsilon_n^2/(4M) = -\eta\log(n)$, i.e., $\epsilon_n = 2\eta^{1/2}M^{1/2}(\log n/r_n)^{1/2}$, as assumed. This proves (34) for $\epsilon_n < M$.

Case 2: $\epsilon_n \geq M$. We have $[2(M+\xi)]^{-1}\epsilon_n^2 \geq [2(M+\epsilon_n)]^{-1}\epsilon_n^2 \geq M/4$, where the first step uses that $\xi \leq \epsilon_n$ and the second step uses that the left side is minimized at $\epsilon_n = M$. Thus, (36) gives $\mathbb{P}(\mathcal{A}_n) \leq e^{-r_n M/4} \leq e^{-n^\beta M/4}$. A simple calculation shows that $e^{-n^\beta M/4} \leq cn^{-\eta}$ for $c = [4\eta/(M\beta e)]^{\eta/\beta}$ and all $n \geq 1$. This proves (34) for $\epsilon_n \geq M$.

Proof of (35). For $\tilde{\epsilon}_n \geq M$, the result holds trivially. For $\tilde{\epsilon}_n < M$, a Taylor expansion gives $f_*(M-\tilde{\epsilon}_n;M) = [2(M-\xi)]^{-1}\tilde{\epsilon}_n^2$ for some $\xi = \xi_n$ in $[0,\tilde{\epsilon}_n]$; note $0 < M-\xi \leq M$ and thus $[2(M-\xi)]^{-1}\tilde{\epsilon}_n^2 \geq [2M]^{-1}\tilde{\epsilon}_n^2$. Now (37) implies $\mathbb{P}(\mathcal{B}_n) \leq e^{-r_n\tilde{\epsilon}_n^2/(2M)}$, and the result follows. □

The following result is known as Hoeffding's inequality.

Lemma 3 *Let* $\{I_n\}$ *be independent Bernoulli(*q*) with* $q \in (0,1)$*, and let* $S_n := \sum_{i=1}^n I_i$*. Then for any nonnegative sequence* $\{\epsilon_n\}$ *and all* $n \geq 1$, $\max\{\mathbb{P}(S_n - nq \geq n\epsilon_n), \mathbb{P}(S_n - nq \leq -n\epsilon_n)\} \leq e^{-2n\epsilon_n^2}$.

8.1 Auxiliary results and their proofs

8.1.1 Results supporting section 4

Lemma 4 *Let* X_n *be a Poisson random variable with mean* n*, and let* $Z_n = (X_n/n)^{-1}\mathbb{1}_{[X_n>0]}$*. Then* $\lim_{n\to\infty}\mathbb{E}[Z_n] = 1$.

Proof We claim: (a) $Z_n \Rightarrow 1$ as $n \to \infty$; and (b) $\lim_{\alpha\to\infty}\sup_n \mathbb{E}[Z_n\mathbb{1}_{[|Z_n|\geq\alpha]}] = 0$; then, the result follows from Theorem 25.12 of Billingsley (1986). Condition (a) follows from: (i) $X_n/n \Rightarrow 1$; (ii) $Z_n = f(X_n/n)$, where $f(x) := x\mathbb{1}_{[x>0]}$ is continuous at $x = 1$; and (iii) the Continuous Mapping Theorem (Billingsley 1986, Theorem 29.2). To verify condition (b), note: $\mathbb{E}[Z_n \mathbb{1}_{[|Z_n|\ge\alpha]}] = \mathbb{E}[nX_n^{-1}\mathbb{1}_{[X_n\le n/\alpha]}] = n\sum_{k=1}^{\lfloor n/\alpha\rfloor} p(k;n)/k$, where the first step uses that $\{|Z_n| \ge \alpha\} = \{X_n \le n/\alpha\}$; and the second step puts $p(k;n) := \mathbb{P}(X_n = k) = e^{-n}n^k/k!$. The sum has at most n/α summands, and each of them is at most $\max_{1\le k\le n/\alpha} p(k,n) \le e^{-n}n^{n/\alpha}/\lfloor n/\alpha\rfloor! \sim c_\alpha^n(\alpha/(2\pi n))^{1/2} =: u_n(\alpha)$ where $c_\alpha := \left(\alpha^{1/\alpha}e^{-1+1/\alpha}\right)$ and the "$\sim$" step uses Stirling's formula. Thus $\lim_{\alpha\to\infty}\sup_n \mathbb{E}[Z_n\mathbb{1}_{[Z_n\ge\alpha]}]$ is at most $\lim_{\alpha\to\infty}\sup_n n\frac{n}{\alpha}u_n(\alpha) = 0$, which follows from $\lim_{\alpha\to\infty} c_\alpha = e^{-1} < 1$. □

8.1.2 Results supporting section 5.1

In this section, the notation and conditions used in Theorem 1 are in force; in particular, η satisfies $\eta \ge 2$; $r_n = n\tau_n/\kappa_n$; and u_n, defined in (25), satisfies

$$\underline{c}_u\left(\frac{\log n}{n}\right)^{1/4} \le u_n \le \overline{c}_u\left(\frac{\log n}{n}\right)^{1/4} \quad \text{for all } n \ge 1, \tag{38}$$

for some $\underline{c}_u > 0$ and $\overline{c}_u > 0$. By Assumption 2 and $\Lambda(p) := \lambda q(p)$,

$$\underline{M}_\Lambda|p_1 - p_2| \le |\Lambda(p_1) - \Lambda(p_2)| = \lambda|q(p_1) - q(p_2)| \le \overline{M}_\Lambda|p_1 - p_2| \tag{39}$$

for any $p_1,\ p_2, \lambda$, where $\overline{M}_\Lambda := \overline{\lambda M}$ and $\underline{M}_\Lambda := \underline{\lambda M}$. Moreover, the same assumption gives $|p_1q(p_1) - p_2q(p_2)| \le \overline{p}\overline{M}|p_1 - p_2| + \overline{q}|p_1 - p_2|$ for any $p_1,\ p_2$, where $\overline{q} := \sup_p q(p) = q(\underline{p})$. Thus $r(\cdot)$ is $\overline{M}_r$-Lipschitz with $\overline{M}_r := \overline{q} + \overline{M}\overline{p}$.

Lemma 5 (Unconstrained maximizer) *Let $\eta \ge 2$, $\alpha > 0$, and define $R_u := 4\max\{c_1, c_2\}$, where $c_1 := \overline{M}_r(\overline{p} - \underline{p})/2$, $c_2 := \overline{p}\eta^{1/2}(2\underline{\lambda})^{-1/2}(1+\alpha)$. Then $\mathbb{P}\{r(p^u) - r(\widehat{p}^u) \ge R_u u_n\} \le C_1/n^{\eta-1}$ for all $n \ge \underline{n}$, with C_1 and $\underline{n}$ as in Lemma* 1(b).

Proof of Lemma 5 Recall that p_i is a short form for $p_{i,n}$. Put $\widehat{q}(p_i) = \widehat{q}_n(p_i)$ and $\widehat{r}(p_i) = p_i\widehat{q}_n(p_i)$, and let j be the interval $(p_{j-1}, p_j]$ that contains p^u (we drop the dependence on n to lighten the notation). Now

$$\begin{aligned} r(p^u) - r(\widehat{p}^u) &= [r(p^u) - r(p_j)] + [r(p_j) - \widehat{r}(p_j)] \\ &\quad + [\widehat{r}(p_j) - \widehat{r}(\widehat{p}^u)] + [\widehat{r}(\widehat{p}^u) - r(\widehat{p}^u)] \\ &\le \overline{M}_r(\overline{p} - \underline{p})\kappa_n^{-1} + 2\max_{1\le i\le\kappa_n}|r(p_i) - \widehat{r}(p_i)|, \end{aligned} \tag{40}$$

since $|r(p^u) - r(p_j)| \le \overline{M}_r(\overline{p} - \underline{p})\kappa_n^{-1}$ (since $r(\cdot)$ is $\overline{M}_r$-Lipschitz and $|p^u - p_j| \le (\overline{p} - \underline{p})\kappa_n^{-1}$); $\widehat{r}(p_j) - \widehat{r}(\widehat{p}^u) \le 0$ (since $\widehat{p}^u = \arg\max_{1\le j\le\kappa_n} p_i\widehat{q}(p_i)$); and the other two terms' absolute value is at most $\max_{1\le i\le\kappa_n}|r(p_i) - \widehat{r}(p_i)|$. Now

$$\frac{R_u u_n}{2} - c_1\kappa_n^{-1} \ge c_2\left(\frac{\log n}{r_n}\right)^{1/2} \quad \text{for all } n \ge 1 \tag{41}$$

by construction of R_u. Now

$$
\begin{aligned}
\mathbb{P}\{r(p^u) - r(\widehat{p}^u) > R_u u_n\} &\overset{(a)}{\le} \mathbb{P}\left\{\overline{p} \max_{1\le i\le \kappa_n} |q(p_i) - \widehat{q}(p_i)| > \frac{R_u u_n}{2} - c_1\kappa_n^{-1}\right\} \\
&\overset{(b)}{\le} \sum_{i=1}^{\kappa_n} \mathbb{P}\left\{|q(p_i) - \widehat{q}(p_i)| > \frac{1}{\overline{p}}\left(\frac{R_u u_n}{2} - c_1\kappa_n^{-1}\right)\right\} \\
&\overset{(c)}{\le} \sum_{i=1}^{\kappa_n} \mathbb{P}\left\{|q(p_i) - \widehat{q}(p_i)| > \eta^{1/2}(1+\alpha)\left(\frac{\log n}{2\underline{\lambda} r_n}\right)^{1/2}\right\} \\
&\overset{(d)}{\le} \frac{C_1}{n^{\eta-1}} \quad \text{for all } n \ge \underline{n},
\end{aligned} \tag{42}
$$

where step (a) uses (40); step (b) uses a union bound; step (c) uses (41); and step (d) uses Lemma 1(b) and that $\kappa_n = o(n)$. □

Lemma 6 (Clearance price) *Let* $\eta \ge 2$, $\alpha > 0$, $K_c := 4\underline{M}_\Lambda^{-1} \max\{c_1, c_2(\alpha)\}$, *where* $c_1 := \overline{M}_\Lambda(\overline{p} - \underline{p})/2$ *and* $c_2 := c_2(\alpha) = (\eta\overline{\lambda}/2)^{1/2}(1+\alpha)$, *and* $\overline{M}_r := \overline{q} + \overline{M}\,\overline{p}$. *Let* C_1 *and* $\underline{n} = \underline{n}(\alpha)$ *be as in Lemma* 1(b). *For all* $n \ge \underline{n}$,

$$
\mathbb{P}\{|\widehat{p}^c - p^c| > K_c u_n\} \le \frac{C_1}{n^{\eta-1}}, \quad \mathbb{P}\{|r(\widehat{p}^c) - r(p^c)| > \overline{M}_r K_c u_n\} \le \frac{C_1}{n^{\eta-1}}. \tag{43}
$$

Proof of Lemma 6 Put $\widehat{\Lambda}(p_i) := \widehat{\lambda}_n \widehat{q}_n(p_i)$ and let j be the interval $(p_{j-1}, p_j]$ that contains p^c (we drop the dependence on n to lighten the notation). Now

$$
|\Lambda(\widehat{p}^c) - \Lambda(p^c)| \overset{(a)}{\le} \max_{1\le i\le\kappa_n} |\widehat{\Lambda}(p_i) - \Lambda(p_i)| + |\widehat{\Lambda}(\widehat{p}^c) - \Lambda(p^c)|, \quad \text{and} \tag{44}
$$

$$
|\widehat{\Lambda}(\widehat{p}^c) - \Lambda(p^c)| \overset{(b)}{\le} |\widehat{\Lambda}(p_j) - \Lambda(p^c)| \overset{(c)}{\le} \max_{1\le i\le\kappa_n} |\widehat{\Lambda}(p_i) - \Lambda(p_i)| + \overline{M}_\Lambda(\overline{p} - \underline{p})\kappa_n^{-1},
$$

where (a) uses the triangle inequality; (b) uses that $\widehat{p}^c = \arg\min_{1\le i\le\kappa_n} |\widehat{\Lambda}(p_i) - x/T| = \arg\min_{1\le i\le\kappa_n} |\widehat{\Lambda}(p_i) - \Lambda(p^c)|$; and (c) uses the triangle inequality; that $\Lambda(\cdot)$ is $\overline{M}_\Lambda$-Lipschitz (see (39)); and that $|p_j - p^c| \le (\overline{p} - \underline{p})\kappa_n^{-1}$. Now

$$
\begin{aligned}
\mathbb{P}\{|\widehat{p}^c - p^c| > K_c u_n\} &\overset{(a)}{\le} \mathbb{P}\{|\Lambda(\widehat{p}^c) - \Lambda(p^c)| > \underline{M}_\Lambda K_c u_n\} \\
&\overset{(b)}{\le} \mathbb{P}\left\{\max_{1\le i\le\kappa_n} |\widehat{\Lambda}(p_i) - \Lambda(p_i)| > \frac{\underline{M}_\Lambda K_c u_n}{2} - c_1\kappa_n^{-1}\right\} \\
&\overset{(c)}{\le} \mathbb{P}\left\{\max_{1\le i\le\kappa_n} |\widehat{\Lambda}(p_i) - \Lambda(p_i)| > (1+\alpha)\left(\frac{\eta\overline{\lambda}\log n}{2r_n}\right)^{1/2}\right\} \\
&\overset{(d)}{\le} \sum_{i=1}^{\kappa_n} \mathbb{P}\left\{|\widehat{\Lambda}(p_i) - \Lambda(p_i)| > (1+\alpha)\left(\frac{\eta\overline{\lambda}\log n}{2r_n}\right)^{1/2}\right\}
\end{aligned}
$$

$$\overset{(e)}{\leq} \frac{C_1}{n^{\eta-1}} \quad \text{for all } n \geq \underline{n}, \tag{45}$$

where step (a) uses (39); step (b) uses that $|\Lambda(\widehat{p}^c) - \Lambda(p^c)| \leq 2\max_{1\leq i\leq \kappa_n} |\widehat{\Lambda}(p_i) - \Lambda(p_i)| + \overline{M}_\Lambda(\overline{p} - \underline{p})/\kappa_n$, which follows from (44); step (c) uses that $K_c u_n/2 - c_1\kappa_n^{-1} \geq (\eta\overline{\lambda}/2)^{1/2}(1+\alpha)(\log n/r_n)^{1/2}$ (by construction of K_c); step (d) uses a union bound; and step (e) uses Lemma 1(b) and that $\kappa_n = o(n)$. The left part of (43) is proven, and the right part follows since $r(\cdot)$ is $\overline{M}_r$-Lipschitz. □

Lemma 7 (Revenue rate) *Let* $\alpha > 0$. *Put* $R = \max\{2R_u, \overline{M}_r^{-1} K_c, 2\overline{M}_r K_c\}$, *with* R_u *as in Lemma 5 and* K_c *as in Lemma 6. Let* C_1 *and* $\underline{n} = \underline{n}(\alpha)$ *be as in Lemma 1*(b). *Then* $\mathbb{E}[r(\widehat{p})] \geq r(p^D) - Ru_n - 2r(p^D)C_1/n^{\eta-1}$ *for all* $n \geq \underline{n}$.

Proof of Lemma 7 The proof is fully parallel to Besbes and Zeevi (2009, Electronic Companion, Lemma 4, Step 3), except that: Lemmata 5 and 6 replace their analogous results; our u_n in (25) and revenue-rate per arrival, $r(p) = pq(p)$, replace their u_n and revenue rate per time, $r(\cdot)$, respectively. □

Lemma 8 (Sales during learning) *For* $M := \overline{\lambda}\overline{q}$, $K_L = M + 2\eta^{1/2}M^{1/2}$, *and* C_0 *as in Lemma 2, we have* $\mathbb{P}(Y_n^{(L)} > K_L n u_n) \leq C_0/n^{\eta-1}$ *for all* $n \geq 1$ *and* $\eta \geq 2$.

Proof of Lemma 8 We have $\mathbb{P}(Y_n^{(L)} > K_L n u_n) = \mathbb{P}\big(\sum_{i=1}^{\kappa_n} N(\lambda q(p_i)r_n) > K_L n u_n\big) \leq \sum_{i=1}^{\kappa_n} \mathbb{P}\big(N(\lambda q(p_i)r_n) > K_L n u_n/\kappa_n\big)$, and thus

$$\begin{aligned}
\mathbb{P}\left(Y_n^{(P)} > K_L n u_n\right) &\overset{(a)}{\leq} \sum_{i=1}^{\kappa_n} \mathbb{P}\left\{N(\lambda q(p_i)r_n) - \lambda q(p_i)r_n > -Mr_n + K_L\frac{nu_n}{\kappa_n}\right\} \\
&\overset{(b)}{\leq} \sum_{i=1}^{\kappa_n} \mathbb{P}\left\{N(\lambda q(p_i)r_n) - \lambda q(p_i)r_n > 2\left(\frac{\eta M \log n}{r_n}\right)^{1/2}\right\} \\
&\overset{(c)}{\leq} \frac{C_0}{n^{\eta-1}},
\end{aligned}$$

where step (a) uses that $\lambda q(p_i) \leq M$; step (b) uses that $-Mr_n + K_L n u_n/\kappa_n = -Mn\tau_n/\kappa_n + K_L n u_n/\kappa_n \geq n/\kappa_n(K_L - M)u_n \geq 2\eta^{1/2}M^{1/2}(\log n/r_n)^{1/2}$, which uses that $u_n \geq \max\{\tau_n, (\log n/r_n)^{1/2}\}$; $n/\kappa_n \geq 1$; and the definition of K_L; and step (c) uses Lemma 2 and that $\kappa_n = o(n)$. □

Lemma 9 (Mean overshoot) *Let* $\Lambda(\overline{p}) \leq x/T$. *Take* $\underline{n} = \underline{n}(\alpha)$ *as in Lemma 1. Then* $\mathbb{E}[(Y_n - nx)^+] \leq K_E n u_n$ *for some positive* K_E *and all* $n \geq \underline{n}$.

Proof of Lemma 9 Let $K_\Lambda = \overline{M}_\Lambda K_c$, where K_c is defined in Lemma 6. Put $K_m = x + K_\Lambda T \sup_{n\geq 1} u_n < \infty$. Let $K_Y = 2\max\{K_L, K_\Lambda T + 2\underline{c}_u^{-1}\eta^{1/2}K_m^{1/2}e^{-1/4}\}$, with K_L as in Lemma 8, $\underline{c}_u$ as in (38), and $\eta \geq 2$. Since $Y_n = Y_n^{(L)} + Y_n^{(P)}$, we have

$$\begin{aligned}
\mathbb{P}(Y_n - nx > K_Y n u_n) &\leq \mathbb{P}(Y_n^{(L)} > K_Y n u_n/2) + \mathbb{P}(Y_n^{(P)} > nx + K_Y n u_n/2) \\
&\overset{(a)}{\leq} \frac{C_0}{n^{\eta-1}} + \mathbb{P}(Y_n^{(P)} > nx + K_Y n u_n/2) \quad \text{for } n \geq 1,
\end{aligned} \tag{46}$$

where step (a) uses Lemma 8 and $K_Y/2 \geq K_L$; and C_0 comes from Lemma 2. To bound the last term in (46), we note that $p^c - \widehat{p} = p^c - \max(\widehat{p}^u, \widehat{p}^c) \leq p^c - \widehat{p}^c$. By Lemma 6, $\mathbb{P}(p^c - \widehat{p} > K_c u_n) \leq \mathbb{P}(p^c - \widehat{p}^c > K_c u_n) \leq C_1 n^{-\eta+1}$ for all $n \geq \underline{n}$. This and the fact that $\Lambda()$ is decreasing and $\overline{M}_\Lambda$-Lipschitz imply

$$\mathbb{P}\{\Lambda(\widehat{p}) > \Lambda(p^c) + K_\Lambda u_n\} \leq \mathbb{P}\{p^c - \widehat{p} > K_c u_n\} \leq \frac{C_1}{n^{\eta-1}} \quad \text{for all} \quad n \geq \underline{n}. \tag{47}$$

Outside the event above, $\Lambda(\widehat{p})n(T - \tau_n)$ (the mean $Y_n^{(P)}$) is at most $n v_n$, where

$$v_n := [\Lambda(p^c) + K_\Lambda u_n](T - \tau_n) \leq x + K_\Lambda T u_n, \tag{48}$$

where the inequality uses $\Lambda(p^c) \leq x/T$, which follows from $\Lambda(\overline{p}) \leq x/T$. Now

$$\begin{aligned}
&\mathbb{P}\{Y_n^{(P)} > nx + K_Y n u_n/2\} \\
&\leq \mathbb{P}\{N(\Lambda(\widehat{p})n(T - \tau_n)) > nx + K_Y n u_n/2,\ \Lambda(\widehat{p}) \leq \Lambda(p^c) + K_\Lambda u_n\} \\
&\quad + \mathbb{P}\{\Lambda(\widehat{p}) > \Lambda(p^c) + K_\Lambda u_n\} \\
&\overset{(a)}{\leq} \mathbb{P}\{N((\Lambda(p^c) + K_\Lambda u_n)n(T - \tau_n)) > nx + K_Y n u_n/2\} + \frac{C_1}{n^{\eta-1}} \\
&= \mathbb{P}\{N(n v_n) - n v_n > nx + K_Y n u_n/2 - n v_n\} + \frac{C_1}{n^{\eta-1}} \overset{(b)}{\leq} \frac{C_0 + C_1}{n^{\eta-1}}
\end{aligned} \tag{49}$$

for all $n \geq \underline{n}$, where step (a) uses (47); and step (b) applies Lemma 2 with $r_n = n$ and $M = K_m \geq \sup_n v_n$; the lemma applies because $nx + K_Y n u_n/2 - n v_n \overset{(b1)}{\geq} nx + K_Y n u_n/2 - n(x + K_\Lambda T u_n) \overset{(b2)}{\geq} 2(\eta K_m n \log n)^{1/2}$, where step (b1) uses (48); and step (b2) uses that $u_n \geq \underline{c}_u (\log n/n)^{1/4}$ and $K_Y/2 - K_\Lambda T \geq 2\underline{c}_u^{-1} \eta^{1/2} K_m^{1/2} \sup_{n \geq 1} (\log n/n)^{1/4}$, by construction of K_Y. Now

$$\begin{aligned}
\mathbb{E}[(Y_n - nx)^+] &= \mathbb{E}[(Y_n - nx)^+ \mathbb{1}_{[Y_n - nx \leq K_Y n u_n]}] + \mathbb{E}[(Y_n - nx)^+ \mathbb{1}_{[Y_n - nx > K_Y n u_n]}] \\
&\leq K_Y n u_n + \mathbb{E}[Y_n \mathbb{1}_{[Y_n > nx + K_Y n u_n]}] \\
&\overset{(a)}{\leq} K_Y n u_n + (nx + K_Y n u_n + 1 + n\overline{\lambda}\overline{q}) \frac{2C_0 + C_1}{n^{\eta-1}} \leq K_E n u_n,
\end{aligned}$$

for all $n \geq \underline{n}$ and some constant K_E, where step (a) uses that for a Poisson random variable Z with mean μ, $\mathbb{E}[Z | Z > a] \leq a + 1 + \mu$ (Besbes and Zeevi 2009, Online Companion, Lemma 5); and it also uses $\mathbb{P}(Y_n - nx > K_Y n u_n) \leq (2C_0 + C_1)/n^{\eta-1}$, which follows from (46) and (49). □

Lemma 10 *Let* $\Lambda(\overline{p}) > x/T$. *Let* $\tilde{K}_Y = \max\{K_L, \Lambda(\overline{p}) + 2\underline{c}_u^{-1}[\eta \Lambda(\overline{p}) T]^{1/2} e^{-1/4}\}$, *with* K_L *as in Lemma 8,* $\underline{c}_u$ *as in (38), and* $\eta \geq 2$. *Put* $\mathcal{A} := \{\omega : \min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\} \geq nx - \tilde{K}_Y n u_n, |\widehat{p} - p^D| \leq K_c n u_n\}$, *with* K_c *as in Lemma 6. Then* $\mathbb{P}(\mathcal{A}) \geq 1 - (2C_0 + C_1)/n^{\eta-1}$ *for all* $n \geq \underline{n}$, *with* C_0 *as in Lemma 2 and* C_1, $\underline{n}$ *as in Lemma 1.*

Table 2 List of symbols in alphabetic order, English and Greek

Symbol	Description
A_i, $A_{i,n}$	Arrival count while testing at price p_i, $p_{i,n}$, respectively
$\overline{A}_n$	$= \sum_{i=1}^{\kappa_n} A_{i,n}$; aggregate arrival count during learning, n-th instance
$\mathbf{A}$, $\mathbf{A}_n$	$(A_1, \ldots, A_\kappa)$ and $(A_{1,n}, \ldots, A_{\kappa_n,n})$, resp.
B, B_n	Relative bias of $\widehat{\lambda q}(p)$ and $\widehat{\lambda}_n \widehat{q}_n(p)$, resp. (does not depend on p)
$f(p)$	$= q(p)[1 - q(p)]$
$h(r)$	$\mathbb{E}[X_r^{-1} \mathbb{1}_{[X_r>0]}]$, where $X_r \sim$ Poisson(r)
$\mathcal{I}$	The set of prices p such that $0 < q(p) < 1$, i.e., $f(p) > 0$
J^D	Opt. revenue in (2)
$J^D(z)$	Opt. revenue in the version of (2) with parameter z (Sect. 5.3)
J_n^π	Mean revenue under policy π, n-th instance
$J_n^\pi(z)$	Mean revenue under policy π and parameter z, n-th instance
$\mathcal{K}(\mathbb{P}_{z_1}^\pi, \mathbb{P}_{z_2}^\pi)$	Kullback–Leibler divergence between measures $\mathbb{P}_{z_1}^\pi$ and $\mathbb{P}_{z_2}^\pi$
ℓ_n	$= (\overline{p} - \underline{p})/\kappa_n$; price granularity, n-th instance (Proposition 4)
$\mathcal{L}$, $\mathcal{L}_{BZ}$	The class in Assumption 2 and in Besbes and Zeevi (2009), resp.
$\underline{M}$, $\overline{M}$	Satisfy $\underline{M}\|p_1 - p_2\| \le \|q(p_1) - q(p_2)\| \le \overline{M}\|p_1 - p_2\|$ for all p_1, p_2
$\underline{M}_\Lambda$, $\overline{M}_\Lambda$, $\overline{M}_r$	$\underline{\lambda M}$, $\overline{\lambda M}$, and $\overline{q} + \overline{M}\overline{p}$, resp., where $\overline{q} = \sup_p q(p) = q(\underline{p})$
$\mathcal{M}$	The class used for the regret lower bound (Sect. 5.3)
m_a, m^D	Satisfy $\max_p r(p) \ge m_a > 0$ and $J^D \ge m^D > 0$, resp.
$N(\cdot)$	A unit-rate Poisson process
$\mathbb{P}_z^\pi$	Prob. measure induced by policy π under param. z (Sect. 5.3)
$\mathcal{P}$	The set of admissible pricing policies
$[\underline{p}, \overline{p}]$	Closed interval of feasible prices
p_∞	Shut-off price: $q(p_\infty) = 0$
p^u, p^c, p^D	Prices via (2): unconstrained, constrained, and optimal price, resp.
$\widehat{p}^u$, $\widehat{p}^c$, $\widehat{p}$	Estimates of p^u, p^c, and p^D, resp.
p_i, $p_{i,n}$	i-th test price: original and n-th instance, resp.
$q(p)$	Probability of purchase, conditional on arrival, when price is p
$\widehat{q}(p)$, $\widehat{q}_n(p)$	Maximum-likelihood estimator of $q(p)$: original and n-th instance, resp.
$q(p; z)$	The version of $q(p)$ under parameter z
r	$= \lambda\tau/\kappa$; mean arrival count at any test price (Sect. 4.2.1)
r_n	$= \lambda_n \tau_n/\kappa_n$; mean arrival count at any test price, n-th instance (in Sect. 4.2.2, $r_n = \lambda_n/\kappa_n$, since $\tau_n = 1$)
$r(p)$	$= pq(p)$; revenue rate (per arrival) under price p
$r(p; z)$	$r(p; z) = pq(p; z)$; revenue rate under price p and param. z (Sect. 5.3)

Table 2 continued

Symbol	Description
$\mathcal{R}_n^{\pi}(z)$	Regret of policy π under parameter z, n-th instance (Sect. 5.3)
s_n	$= r_n \kappa_n^2$; multiplicative effect of n on mean square error (Sect. 4.2.2)
T	Length of selling horizon
u_n	$= \max\{\tau_n, \kappa_n^{-1}, (\log n / r_n)^{1/2}\}$; deviations in proving Theorems 1, 2
x	Initial inventory
$Y_n^{(L)}$, $Y_n^{(P)}$, Y_n	Sales, under π_n and $x = \infty$, while learning, pricing, and overall, resp.
Z	Domain of parameter z (Sect. 5.3)
κ, κ_n	Number of test prices: original and n-th instance, resp.
$\Lambda(p)$	$= \lambda q(p)$, the (Poisson) demand function associated to $(\lambda, q(\cdot))$
λ, λ_n	Rate of (Poisson) arrivals: original and n-th instance, resp.
$\underline{\lambda}, \overline{\lambda}$	Lower- and upper-bound on the arrival rate under $\mathcal{L}$, resp.
λ_0	A fixed arrival rate, used to specify the class $\mathcal{M}$ (Sect. 5.3)
$\widehat{\lambda}, \widehat{\lambda}_n$	Maximum-likelihood estimator of λ and λ_n, resp.
$\widehat{\lambda q}(p), \widehat{\lambda q}_n(p)$	Sales-only estimator of $\lambda q(p)$ and $\lambda_n q(p)$, resp.
π_n	The policy given by Algorithm AS, n-th instance
π_n'	The policy in Besbes and Zeevi (2009) and Theorem 2, n-th instance
$\sigma_1(p), \sigma_{1,n}(p)$	$\mathbb{E}[\mathbb{V}ar(\widehat{\lambda}\widehat{q}(p)\vert\mathbf{A})]$ and $\mathbb{E}[\mathbb{V}ar(\widehat{\lambda}_n\widehat{q}_n(p)\vert\mathbf{A}_n)]$, resp.
$\sigma_2(p)$	$= \mathbb{E}[\mathbb{E}^2[\widehat{\lambda}\widehat{q}(p)\vert\mathbf{A}]] - \lambda^2 q^2(p)(1 + 2B)$
	$= \mathbb{V}ar(\mathbb{E}[\widehat{\lambda}\widehat{q}(p)\vert\mathbf{A}]) + [\lambda q(p)B]^2$
$\sigma_{2,n}(p)$	$= \mathbb{E}[\mathbb{E}^2[\widehat{\lambda}_n\widehat{q}(p)\vert\mathbf{A}_n]] - \lambda_n^2 q^2(p)(1 + 2B_n)$; n-th instance of $\sigma_2(p)$
τ, τ_n	Duration of learning phase: original and n-th instance (Sect. 5), resp.

Proof of Lemma 10

$$
\begin{aligned}
\mathbb{P}(\mathcal{A}^c) &\overset{(a)}{\leq} \mathbb{P}(Y_n^{(P)} < nx - \tilde{K}_Y n u_n) + \mathbb{P}(Y_n^{(L)} > \tilde{K}_Y n u_n) + \mathbb{P}(|\widehat{p} - p^D| \leq K_c n u_n) \\
&\overset{(b)}{\leq} \mathbb{P}(Y_n^{(P)} < nx - \tilde{K}_Y n u_n) + \frac{C_0}{n^{\eta-1}} + \mathbb{P}(|\widehat{p}^c - p^c| \leq K_c n u_n) \\
&\overset{(c)}{\leq} \mathbb{P}(Y_n^{(P)} < nx - \tilde{K}_Y n u_n) + \frac{C_0}{n^{\eta-1}} + \frac{C_1}{n^{\eta-1}} \quad \text{for all } n \geq \underline{n},
\end{aligned}
\tag{50}
$$

where (a) uses a union bound; step (b) uses Lemma 8 and that $|\widehat{p} - p^D| \leq |\widehat{p}^c - p^c|$; and step (c) uses Lemma 6. Now, putting $\gamma_n := n\Lambda(\overline{p})(T - \tau_n)$,

$$
\begin{aligned}
\mathbb{P}\{Y_n^{(P)} < nx - \tilde{K}_Y n u_n\} &= \mathbb{P}\big\{N\big(\Lambda(\widehat{p})n(T - \tau_n)\big) < nx - \tilde{K}_Y n u_n\big\} \\
&\overset{(a)}{\leq} \mathbb{P}\big\{N\big(\gamma_n\big) < n\Lambda(\overline{p})T - \tilde{K}_Y n u_n\big\}
\end{aligned}
$$

$$\stackrel{(b)}{\leq} \frac{C_0}{n^{\eta-1}} \quad \text{for all } n \geq 1, \tag{51}$$

where step (a) uses that $\Lambda(\widehat{p}) \geq \Lambda(\overline{p})$ and $x \leq \Lambda(\overline{p})T$; and step (b) applies Lemma 2 with $r_n = n$ and $M = \Lambda(\overline{p})T$; the lemma applies because $n(\tilde{K}_Y u_n - \Lambda(\overline{p})\tau_n) \geq (\tilde{K}_Y - \Lambda(\overline{p}))\underline{c}_u n^{3/4} \log^{1/4} n \geq 2[\eta\Lambda(\overline{p})Tn \log n]^{1/2}$, where the last step uses $\tilde{K}_Y - \Lambda(\overline{p}) \geq 2\underline{c}_u^{-1}[\eta\Lambda(\overline{p})T]^{1/2} \sup_{n\geq 1}(\log n/n)^{1/4}$, by construction of $\tilde{K}_Y$. This proves (51); this and (50) give the result. □

9 List of notation

Table 2 lists the symbols that are essential to all formal statements (definitions, assumptions, conditions, claims, lemmata, propositions, theorems). Each symbol is accompanied by a description in English, or an expression, or definition, via other symbols in the table.

References

Araman VF, Caldentey R (2009) Dynamic pricing for nonperishable products with demand learning. Oper Res 57:1169–1188

Auer P, Cesa-Bianchi N, Fischer P (2002) Finite-time analysis of the multiarmed bandit problem. Mach Learn 47:235–256

Babaioff M, Dughmi S, Kleinberg R, Slivkins A (2015) Dynamic pricing with limited supply. ACM Trans Econ Comput 3(1):1–26

Besbes O, Zeevi A (2009) Dynamic pricing without knowing the demand function: risk bounds and near-optimal algorithms. Oper Res 57:1407–1420

Besbes O, Zeevi A (2012) Blind network revenue management. Oper Res 60:1537–1550

Besbes O, Zeevi A (2015) On the (surprising) sufficiency of linear models for dynamic pricing with demand learning. Manag Sci 61:723–739

Bickel PJ, Doksum KA (1977) Mathematical statistics: basic ideas and selected topics. Holden-Day Company, Oakland

Billingsley P (1986) Probability and measure, 2nd edn. Wiley, New York

Bitran G, Caldentey R (2003) An overview of pricing models for revenue management. Manuf Serv Oper Manag 5:203–229

Brémaud P (1981) Point processes and queues: martingale dynamics, vol 50. Springer, Berlin

Broder J, Rusmevichientong P (2012) Dynamic pricing under a general parametric choice model. Oper Res 60:965–980

Chen B, Chao X, Ahn H-S (2019) Coordinating pricing and inventory replenishment with nonparametric demand learning. Oper Res 67:1035–1052

Cope E (2009) Regret and convergence bounds for a class of continuum-armed bandit problems. IEEE Trans Autom Control 54:1243–1253

Dembo A, Zeitouni O (1998) Large deviation techniques and applications. Springer, Berlin
den Boer AV (2015) Dynamic pricing and learning: historical origins, current research, and new directions. Surv Oper Res Manag Sci 20:1–18
den Boer AV, Keskin NB (2019) Discontinuous demand functions: estimation and pricing. Manag Sci. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3003984
den Boer AV, Zwart B (2014) Simultaneously learning and optimizing using controlled variance pricing. Manag Sci 60:770–783
Farias VF, Van Roy B (2010) Dynamic pricing with a prior on market response. Oper Res 58:16–29
Gallego G, van Ryzin G (1994) Optimal dynamic pricing of inventories with stochastic demand over finite horizons. Manag Sci 40:999–1020
Keskin NB, Zeevi A (2014) Dynamic pricing with an unknown demand model: asymptotically optimal semi-myopic policies. Oper Res 62:1142–1167
Kleinberg RD (2004) Nearly tight bounds for the continuum-armed bandit problem. In: Proceedings of the 17th international conference on neural information processing systems, pp 697–704
Lei Y, Jasin S, Sinha A (2014) Near-optimal bisection search for nonparametric pricing with inventory constraint. https://deepblue.lib.umich.edu/handle/2027.42/108717
Lin KY (2006) Dynamic pricing with real-time demand learning. Eur J Oper Res 174:522–538
Rudin W (1976) Principles of mathematical analysis, 3rd edn. McGraw-Hill, New York
Talluri K, van Ryzin G (2005) Theory and practice of revenue management. Springer, Berlin
Tsybakov AB (2009) Introduction to nonparametric estimation. Springer, New York
Wang Z, Deng S, Ye Y (2014) Close the gaps: a learning-while-doing algorithm for a class of single-product revenue management problems. Oper Res 62:318–331