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 it is asymptotically more efficient than the maximum likelihood estimator based only on sale data. Based on the same estimator, we develop a pricing algorithm paralleling Besbes and Zeevi (2009). 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 the proof methods of Besbes and Zeevi (2009). Numerical comparisons to the policies in Besbes and Zeevi (2009) and Wang et al. (2014) indicate the effectiveness of our arrivals-based approach.

Keywords estimation · asymptotic efficiency · exploration-exploitation · regret · asymptotic analysis.

Mathematics Subject Classification (2010) 60K10 · 93E35 · 90B05 · 62G20

A Pricing Problem with Unknown Arrival Rate and Price Sensitivity

Athanassios N. Avramidis

the date of receipt and acceptance should be inserted later

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.

Much of the earlier literature makes the critical assumption that the relationship between the mean demand rate and the price, often referred to as the *demand function* or *demand curve*, is known to the seller. In Gallego and van Ryzin (1994), the seller knows the demand function at the start of selling and designs optimal policies based on this knowledge. In practice, decision makers seldom have full knowledge about the demand model. Much of the recent literature addresses this issue by learning (estimating) the underlying demand model by price experimentation and observation of the realized demand. The established measure of performance is the *regret*: this is the expected revenue under-performance relative to an oracle that knows fully the demand function.

The absence of full information about the demand model introduces a tension between exploration (demand learning) and exploitation (revenue earning). The longer one spends learning the demand properties, the less time remains to exploit that knowledge and earn revenue; on the other hand, less time spent on demand learning results in higher uncertainty that could diminish the revenue earned during the exploitation phase. Besbes and Zeevi (2009)

Athanassios N. Avramidis
Faculty of Mathematical Sciences, University of Southampton, Southampton, SO17 1BJ, UK.
Tel: +44 2380 595136
E-mail: aa1w07@soton.ac.uk

formulate pricing problems under incomplete information, and highlight the key trade-offs.

Models of demand uncertainty can be broadly divided into *parametric* and *non-parametric*. A parametric model assumes a certain functional form and consequently carries *mis-specification risk*, which means the potential for revenue underperformance; this point is demonstrated in Besbes and Zeevi (2009). A nonparametric approach imposes less structure on the assumed functional form. It is therefore reasonable to expect that a non-parametric model mitigates the mis-specification risk. On the other hand, pricing with a mis-specified parametric (linear) demand model can perform well (Besbes and Zeevi 2015).

Prominent works that take a non-parametric approach, in a single-product setting, are Besbes and Zeevi (2009), Wang et al. (2014), and Lei et al. (2014). We summarize some key ideas in their simpler form, following Besbes and Zeevi (2009). During a *learning phase*, the demand function is estimated (learned) by observing the realized demand (sales) at certain *test prices*. Based on the estimated demand function, a "good" fixed-price policy is applied during the remainder of the selling horizon, the so-called *exploitation phase*. The revenue performance is measured in terms of the *worst-case regret* across a class of demand functions. Lower and upper bounds are derived on this regret. The tension between exploration and exploitation is resolved near-optimally in an asymptotic regime where both demand and inventory level (capacity) grow in proportion.

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 the demand function is their product, $\lambda q(\cdot)$. We introduce a class $\mathcal{D}$ by requiring that $q(\cdot)$ and its inverse are Lipschitz continuous.
2. Besbes and Zeevi (2009) work with a class $\mathcal{D}_{BZ}$ in which the demand function and its inverse are Lipschitz continuous, and Wang et al. (2014) work with a smaller class $\mathcal{D}_W$ in which the demand function is twice differentiable. Our class $\mathcal{D}$ is smaller than $\mathcal{D}_{BZ}$, but includes models that are outside $\mathcal{D}_W$.
3. Based on arrival and sale data (i.e., non-purchasing arrivals are also observed), we construct a maximum-likelihood estimator of $(\lambda, \boldsymbol{q})$, where $\boldsymbol{q} = (q(p_1), \ldots, q(p_\kappa))$ for some price set $\{p_1, \ldots, p_\kappa\}$. The product of the 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).

4. We develop a pricing policy based on the estimated λ and $\boldsymbol{q}$, and a counterpart policy based on the sales-only estimator. These policies' worst-case regret is shown to be $O(\log n/n)^{1/4}$, against $\mathcal{D}$ (Theorem 1) and against $\mathcal{D}_{BZ}$ (Theorem 2), respectively. The convergence is (slightly) faster than that obtained in Besbes and Zeevi (2009), and arises through refined bounds and slightly different asymptotics for the time spent learning and the number of test prices.
5. We correct the large-deviation bound for the sample mean of independent Poisson variables that was initiated in Besbes and Zeevi (2009, Online Companion, Lemma 2) and is commonly used thereafter, including Besbes and Zeevi (2012, Online Companion, Lemma 1) and Wang et al. (2014, Lemma 11). We only correct the constant, not the convergence rate.

1.3 Related Literature

The literature on pricing strategies is vast. We refer to Bitran and Caldentey (2003); Elmaghraby and Keskinocak (2003); Talluri and van Ryzin (2005) for comprehensive reviews on the subject. For a more recent survey, see den Boer (2015).

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. The majority of work adopts a parametric model: the demand process is known up to a finite number of parameters. In Lin (2006), Aviv and Pazgal (2005), Araman and Caldentey (2009), and Farias and Van Roy (2010), there is a single unknown parameter representing the market size. den Boer and Zwart (2015) assume a two-parameter demand model, and provide bounds on the regret. In den Boer and Keskin (2017), the demand function is allowed to have a number of discontinuities, but is still restricted parametrically inside each continuity interval.

Some studies consider a demand process that is adversarial to the seller (Lim and Shanthikumar 2007; Ball and Queyranne 2009). Such studies are not concerned with, and do not allow for, learning the demand model.

For the single-product setting without the inventory constraint, a stream of literature addresses demand learning and characterizes the regret (Broder and Rusmevichientong 2012; den Boer and Zwart 2014; Besbes and Zeevi 2015). Besbes and Zeevi (2015) shows 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 setting we study involves a Poisson demand process, where the demand function is unknown to the seller; this is a well-studied situation. Most closely related to the present paper are the works of Besbes and Zeevi (2009), Wang et al. (2014) and Lei et al. (2014). In these, a worst-case regret guarantee is provided against a class of demand functions that is significantly larger than typical families with one or two parameters (linear, exponential, logit, etc.). Besbes and Zeevi (2009) impose the weakest requirement on the demand function, only requiring it to be Lipschitz continuous ~~we call this class $\mathcal{D}_2$.~~ Wang et al. (2014) and Lei et al. (2014) impose stronger conditions, thus working with a smaller class of demand functions; ~~we call it $\mathcal{D}_3$.~~ Relative to these works, we additionally require the existence of a (constant) arrival rate; and we require the seller to observe arrivals that do not purchase (in addition to those that do purchase).

We now review these latter works and relate them to this paper. Under the full-information deterministic relaxation (Gallego and van Ryzin 1994), the optimal policy prices at the constant price $p^D = \max(p^u, p^c)$, where: (i) p^u, the *optimal unconstrained price*, maximizes the revenue rate function, that is, the function $p\lambda(p)$, where p is price and $\lambda(\cdot)$ is the demand function; and (ii) p^c, the *constrained price*, minimizes the absolute difference between the mean demand and the inventory at hand. One estimates the demand function, and thereby the price p^D, from the demand observed at certain test prices. The non-parametric algorithm of Besbes and Zeevi (2009, Section 4.1) involves relatively simple price testing and achieves a regret $O(n^{-1/4}(\log n)^{1/2})$. These authors also prove an $\Omega(n^{-1/2})$ lower bound on the regret of any admissible policy. Our worst-case regret upper bound against their class of demand functions improves theirs, and is the best known for this class (Theorem 2).

Wang et al. (2014) and Lei et al. (2014) work with a smaller class of demand models that requires smoothness (twice-differentiability) of the demand function. Wang et al. (2014) perform sequential price testing on a set of shrinking price intervals that contain p^D with high probability, and achieve regret $O(n^{-1/2}(\log n)^{4.5})$. Lei et al. (2014) achieve regret $O(n^{-1/2})$, which closes the rate gap against the lower bound. The main contrast between our work and theirs is that their worst-case regret convergence is faster, but the model class against which their worst-case guarantee applies does not fully include our class (neither class contains the other).

Our work is also related to the continuum-armed bandit literature (e.g., Auer et al. (2002); Kleinberg and Leighton (2003)). While our approach has some high-level connections with the bandit approach, the presence of an inventory constraint in our problems clearly distinguishes our work from theirs.

The Remainder of this Paper. Section 2 introduces our model and formulates the problem. Section 3 presents the estimation and pricing methods for our main model, in which non-purchasing arrivals are observable. Section 4 analyzes the estimation error and the estimation efficiency. Section 5 analyzes the (worst-case) regret in two settings: the main one, in which non-purchases are observed (§5.1); and another one, in which non-purchases need not be observed (§5.2). Section 6 presents a numerical comparison against alternative methods. Section 8 contains selected proofs.

2 Problem Formulation and Background

Model of Demand and Basic Assumptions. We consider a monopolist that sells a single product. The selling horizon is denoted $T > 0$, and after this point sales are discontinued, and any unsold products have no value. Product demand is modeled as follows:

Assumption 1 *(a) (Demand Model.) Customers arrive according to a Poisson process of rate λ, regardless of the price. Whenever the price is p, an arriving customer purchases with probability $q(p) \in [0, 1]$ independently of everything else.*
(b) (Seller's Information.) The seller observes the counting processes of arrivals and sales throughout the selling period.

We refer to the pair (λ, q) as the *primitives*. The seller knows nothing about these primitives, except for their membership in some (broad) class that we specify later.

In upper-bounding the regret, we need:

Assumption 2 *For some finite positive constants $\underline{\lambda}$, $\overline{\lambda}$, $\underline{M}$, $\overline{M}$, m_a, with $\underline{\lambda} < \overline{\lambda}$ and $\underline{M} \leq \overline{M}$:*
(a) The arrival rate is bounded away from zero and infinity: $\underline{\lambda} \leq \lambda \leq \overline{\lambda}$.
(b) The price sensitivity 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 per customer is positive: $\max\{pq(p) : p \in [\underline{p}, \overline{p}]\} \geq m_a$.

Let $\mathcal{D} = \mathcal{D}(\overline{\lambda}, \underline{\lambda}, \underline{M}, \overline{M}, m)$ be the class of demand models satisfying Assumptions 1(i) and 2. Any bounded demand function Λ is representable as the pair $(\overline{\Lambda}, \Lambda(\cdot)/\overline{\Lambda})$ (as in Assumption 1(a)), provided only $\overline{\Lambda} \geq \sup_p \Lambda(p)$.

In relevant literature, no reference to an arrival rate is made. Besbes and Zeevi (2009) (Assumption 1) define a class $\mathcal{D}_{BZ}$, the essential requirement being Lipschitz continuity of the demand function and its inverse; Wang et al. (2014) define a smaller class $\mathcal{D}_W$ in which the demand function is twice differentiable. Our Assumption 2 has Lipschitz-continuity requirements on the purchase probability instead. It follows that $\mathcal{D}$ is included in $\mathcal{D}_{BZ}$; moreover, the inclusion is strict, as we explain shortly.

The key restriction of $\mathcal{D}$ arises from the possibility that the primitives (arrival rate and purchase probability function) are time-shifting. To make this precise, consider two pairs (λ, q) and (λ', q') with $\lambda' \neq \lambda$ and $\lambda' q'(p) = \lambda q(p)$ for all p. If these pairs span (cover) the selling horizon, with each applying for a positive amount of time, then the induced demand function, $\Lambda(p) = \lambda q(p)$, is time-invariant. This model is not representable in $\mathcal{D}$, and yet would be representable in $\mathcal{D}_{BZ}$ and $\mathcal{D}_W$ under suitable smoothness conditions on q, q'.

On the other hand, a pair (λ, q), such that q is Lipschitz and not twice-differentiable everywhere, induces a model that is inside $\mathcal{D}$ but outside $\mathcal{D}_W$ (since the demand function λq inherits from q the non-differentiability). In summary, class $\mathcal{D}$ is strictly smaller than $\mathcal{D}_{BZ}$, and obeys no inclusion relation to $\mathcal{D}_W$.

The assumption of constant arrival rate may not be sensible if, for example, customers arrive to an e-commerce site via a price-comparison site; on the other hand, if customers tend to be loyal to the seller, then it may be reasonable.

Estimation problem. For the sole purpose of analyzing estimation error and 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), and which is valid when inventory is unlimited, does not appear to be crucial and can probably be relaxed. This intuition is based on the analysis of regret, where we show that the probability of a stock-out vanishes fast when the system size increases and the fraction of time spent learning vanishes at an appropriate (polynomial) rate (Proposition 5). However, given that the analysis of estimation efficiency is lengthy even when Assumption 3 holds, we do not attempt to relax it in this paper.

Our measure of global estimation efficiency accounts for two separate cases, allowing the set of feasible prices to be finite or an infinite continuum. To analyze the latter case, we require the following:

Assumption 4 *(a) The price lies in a continuum $[\underline{p}, \overline{p}]$ with $\underline{p} < \overline{p}$.*
(b) The purchase-probability $q(\cdot)$ is continuous almost everywhere on $[\underline{p}, \overline{p}]$; that is, it has at most countably many discontinuities.
(c) The set $\{p : q(p) \notin \{0, 1\}\}$ has positive Lebesgue measure.

Assumption 4(a) simply says that there is a continuum of feasible prices, which is bounded. Assumptions 4(b)-(c) are very mild, and only exclude functions $q(\cdot)$ unlikely to arise in applications.

Pricing Problem. The set of feasible prices is assumed to be $[\underline{p}, \overline{p}] \cup p_\infty$, 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 seller applies this price whenever the inventory is depleted. When we study the estimation of $(\lambda, q(\cdot))$ in isolation from the pricing problem (see §4), we will include the case where the set of feasible prices is finite. We assume the following: $q(p)$ is continuous and nonincreasing in the price p; and $q(\cdot)$ has an inverse denoted $q^{-1}(\cdot)$. A price of p induces a *revenue rate per arrival* $r(p) := pq(p)$ and a *revenue rate per unit of time* $r_t(p) := \lambda r(p) = p\Lambda(p)$, where $\Lambda(p) := \lambda q(p)$ is the *demand function resulting from* (λ, q).

We use π to denote a *(pricing) policy.* Any policy π induces a *price process*, which we denote $p^\pi := (p_t^\pi : 0 \leq t \leq T)$. We assume that the sample paths of p^π are right-continuous with left limits. Let $N = (N(t) : t \geq 0)$ denote a unit-rate Poisson process. The cumulative number of arrivals up to time t is then given by $N(\lambda t)$. Under a policy π, the cumulative demand up to time t has the representation

$$D_t^\pi = N\left(\lambda \int_0^t q(p_s^\pi)\mathrm{d}s\right),$$

which follows from Assumption 1(a) and the well-known result known as “thinning of a Poisson process”.

A policy π is said to be *non-anticipating* if the induced price process is only allowed to depend on past prices $\{p_u^\pi : u \in [0,t)\}$, past arrival counts $\{N_u : u \in [0,t)\}$, and past demand counts $\{D_u^\pi : u \in [0,t)\}$. (More formally, policy π is said to be non-anticipating if the induced price process $(p_t^\pi : 0 \leq t \leq T)$ is adapted to the filtration $\mathcal{F}_t = \sigma(p_u^\pi, N_u, D_u^\pi : 0 \leq u < t)$, where $\sigma(\cdot)$ denotes the sigma field generated by the indicated random elements).

Let $x > 0$ denote the inventory level (number of product units) at the start of the selling period. A pricing policy π is said to be *admissible* if the induced price process $(p_t^\pi : 0 \leq t \leq T)$ is such that $p_t^\pi \in [\underline{p}, \overline{p}] \cup p_\infty$ for all t and

$$\int_0^T \mathrm{d}D_s^\pi \leq x \quad \text{with probability 1.} \tag{1}$$

With $\mathcal{P}$ denoting the set of admissible pricing policies, the seller's problem can be stated as: choose $\pi \in \mathcal{P}$ to maximize the expected revenue

$$J^\pi := J^\pi(x, T|\lambda, q) := \mathbb{E}_{\lambda,q}\left[\int_0^T p_s^\pi \mathrm{d}D_s^\pi\right], \tag{2}$$

where the notation indicates that the expectation is with respect to the true demand process (λ, q).

Full-Information Deterministic Relaxation. The *full-information relaxation* (FIR) is the deterministic optimization problem where the stochastic elements in (1) and (2) are replaced by their means:

$$\begin{aligned} J^D := J^D(x,T|\lambda,q) := \sup \int_0^T r_t(p_s)\mathrm{d}s &= \sup \int_0^T \lambda r(p_s)\mathrm{d}s \\ \text{s.t.} \int_0^T \Lambda(p_s)\mathrm{d}s &= \int_0^T \lambda q(p_s)\mathrm{d}s \leq x \\ p_s &\in [\underline{p}, \overline{p}] \cup p_\infty \text{ for all } s \in [0,T]. \end{aligned} \tag{3}$$

Our method is derived from the optimal solution to this problem, which is (Besbes and Zeevi 2009):

$$p^u := \arg\max{}_p\{pq(p)\}, \quad p^c := \arg\min{}_p \left|\lambda q(p) - \frac{x}{T}\right|, \quad p^D := \max\{p^u, p^c\}. \tag{4}$$

Applying the fixed price p^D while inventory is positive performs well as the system size grows (Gallego and van Ryzin 1994, Theorem 3); our method, in line with Besbes and Zeevi (2009), is based on approximating p^D. In view of the relation $J^\pi(x,T|\lambda,q) \le J^D(x,T|\lambda,q)$ (Gallego and van Ryzin 1994, Theorem 1), we will bound the regret in relation to J^D rather than J^π.

Notation. Throughout the paper, the following notation is used. Definition is indicated by ":=" and "=:", with colon on the side being defined. The set of natural numbers is $\mathbb{N} := \{0,1,2,\ldots\}$. For a set A, $\mathbb{1}_{[A]}$ denotes the indicator function; A^c denotes the complement; and $|A|$ denotes the cardinality. For any real x, we write $\lfloor x \rfloor$ for the *floor*, the largest integer that is no larger than x; $\lceil x \rceil$ for the *ceiling*; and x^+ for the *positive part* $\max\{0,x\}$. With a_n and b_n being 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; and $a_n \sim b_n$ means that $\lim_{n\to\infty} a_n/b_n = 1$. With X being a random variable, we write "$X \sim$" to indicate the probability law of X.

3 Estimation and Pricing Methods

Our estimation naturally pursues the arrival rate λ, together with purchase probabilities $q_i = q(p_i)$ associated to a finite set of prices $\mathcal{P} = \{p_1,\ldots,p_\kappa\}$. Our estimation method is as follows:

Method A (τ,κ). Set the learning interval as $[0,\tau]$, and set $\Delta = \tau/\kappa$. During time $(i-1)\Delta$ to $i\Delta$, $i = 1,2,\ldots,\kappa$, price at p_i and record the arrival count, A_i, and the sale count, S_i. If a stock-out occurs before time τ, then price at p_∞ and stop sales. Put $A = \sum_{i=1}^{\kappa} A_i$ and put

$$\widehat{\lambda} = \frac{A}{\tau}, \quad \widehat{q}_i = \frac{S_i}{A_i}\mathbb{1}_{[A_i>0]},\ i = 1,\ldots,\kappa, \tag{5}$$

as a joint estimator of λ and the corresponding $q(p_i)$. It will be seen shortly that, assuming that no stock-out occurs during the estimation phase, this estimator is a maximum-likelihood one.

We now introduce a pricing policy (algorithm), aimed for the case when there exists an arrival rate and the seller observes both arrival and sale data. This is an adaptation of the non-parametric algorithm of Besbes and Zeevi (2009, Section 4.1) to this setting.

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

Step 1. Initialization:
(a) Set the learning interval to be $[0,\tau]$, and the number of prices to experiment with to be κ. Put $\Delta = \tau/\kappa$.
(b) Divide $[\underline{p},\overline{p}]$ into κ equally spaced intervals and let $\{p_i, i = 1,\ldots,\kappa\}$ be the mid-points (or the left endpoints or the right-endpoints) of these intervals.

Step 2. Learning (testing):

(a) On the interval $[0, \tau]$ apply price p_i from $t_{i-1} = (i-1)\Delta$ to $t_i = i\Delta$ for $i = 1, \ldots, \kappa$, as long as inventory is positive. If inventory is zero at any time, apply p_∞ until time T and stop.

(b) Let N_i and S_i be the number of arrivals and sales, respectively, during $[t_{i-1}, t_i)$. Let $N = \sum_{i=1}^{\kappa} N_i$. Compute

$$\widehat{\lambda} = \frac{N}{\tau}, \quad \widehat{q}_i := \frac{S_i}{N_i} \mathbb{1}_{[N_i > 0]} \tag{6}$$

Step 3. Optimization: Compute

$$\widehat{p}^u = \arg\max_{1 \le i \le \kappa} \{p_i \widehat{q}_i\}$$
$$\widehat{p}^c = \arg\min_{1 \le i \le \kappa} |\widehat{\lambda}\widehat{q}_i - x/T|$$

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.

Recalling the solution to the full-information relaxation (4), we see that the learning phase estimates the prices p^u, p^c and p^D, by $\widehat{p}^u$, $\widehat{p}^c$, and $\widehat{p}$, respectively. Tension exists in choosing τ and κ, paralleling Besbes and Zeevi (2009). Briefly, as τ increases, one expects better estimation accuracy but also larger revenue losses due to suboptimal pricing; it is reasonable to expect these losses are of order τ. The revenue function is only learned at a limited resolution of κ prices, and an error of order $1/\kappa$ is incurred in estimating p^u. This tension is made precise and resolved in a large-system analysis in §5.1.

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 the vector $\mathbf{A} := (A_1, A_2, \ldots, A_\kappa)$, the conditional law of $(S_1, S_2, \ldots S_\kappa)$ consists of independent marginals of the form $S_i \sim$ Binomial(A_i, q_i) for all $i = 1, \ldots, \kappa$.

The following result is elementary.

Proposition 1 *(Maximum Likelihood Estimator of the Arrival Rate and Price Sensitivity.) Let Assumptions 1 and 3 hold. Let Method* $A(\tau, \kappa)$ *produce data* $(A_1, S_1, \ldots, A_\kappa, S_\kappa)$*, where* A_i *and* S_i *are the count of arrivals and sales, respectively, while the offered price is* p_i*. Then, a maximum-likelihood estimator of* $\theta := (\lambda, q_1, \ldots, q_\kappa)$ *is given in (5).*

Proof of Proposition 1. In view of P1 and P2, the log-likelihood is

$$\ell(\theta; \mathcal{D}) \propto -\lambda\tau + A\log(\lambda) + \sum_{i=1}^{k}(S_i \log(q_i) + (A_i - S_i)\log(1 - q_i)).$$

It is now easy to see that $\widehat{\lambda}$ and $\widehat{q}_i$ in (5) are the unique solution to the first-order optimality conditions, $0 = \partial\ell/\partial\lambda = -\tau + A/\lambda$, and $0 = \partial\ell/\partial q_i = S_i/q_i - (A_i - S_i)/(1 - q_i)$. (If $A_i = 0$ for some i, then $\widehat{q}_i = 0$ is an MLE, but not uniquely so.) □

4.2 Estimator of Mean Demand

Here we are interested in estimating the *demand vector* $(\lambda q_1, \ldots, \lambda q_\kappa)$ for any finite κ. The invariance property of Maximum Likelihood Estimation (Bickel and Doksum 1977, Section 4.5) immediately gives the following.

Proposition 2 *(Maximum Likelihood Estimator of the Demand Vector.) Let $\mathcal{P} = \{p_1, \ldots, p_\kappa\}$ be any subset of the set of feasible prices. A maximum likelihood estimator of $(\lambda q_1, \ldots, \lambda q_\kappa)$ is $(\widehat{\lambda}\widehat{q}_1, \ldots, \widehat{\lambda}\widehat{q}_\kappa)$, where $\widehat{\lambda}$ and $\widehat{q}_1, \ldots, \widehat{q}_\kappa$ are given in (5).*

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

In Section 4.2.1 we study this estimator's bias and mean square error (MSE) locally, that is, at a single price. In Section 4.2.2 we study, in an asymptotic (large sample) regime, the local and global estimation efficiency relative to an estimator that uses sale counts only, which is in wide use (Besbes and Zeevi 2009, 2012; Wang et al. 2014; Lei et al. 2014).

4.2.1 Bias and Mean Square Error

We derive explicit expressions for the bias and mean square error (MSE) of the arrivals-and-sales estimator. To state the result, let X_λ denote a Poisson(λ) random variable, and put

$$\begin{aligned} \rho &:= \mathbb{P}(X_{\lambda\tau/\kappa} > 0) = 1 - \exp\big(-\frac{\lambda\tau}{\kappa}\big), \\ h(\lambda) &:= \mathbb{E}[X_\lambda^{-1}\mathbb{1}_{[X_\lambda > 0]}], \\ c_1 = c_{1,\lambda,\tau,\kappa} &:= h\left(\frac{\lambda\tau}{\kappa}\right) + 2\rho + (\kappa - 1)^{-1}. \end{aligned} \tag{7}$$

Relative bias is defined by the relation $B := \mathbb{E}[\widehat{\lambda}\widehat{q}_i]/(\lambda q_i) - 1$. An equivalent definition uses the relation $\mathbb{E}[\widehat{\lambda}\widehat{q}_i] := \lambda q_i(1 + B)$.

Proposition 3 *(Mean Square Error and Bias.) Under Assumptions 1 and 3, we have*

$$B = -\frac{\kappa - 1}{\kappa}\exp\left(-\frac{\lambda\tau}{\kappa}\right), \tag{8}$$

and

$$\mathbb{E}[(\widehat{\lambda}\widehat{q}_i - \lambda q_i)^2] = \sigma_1 + \sigma_2, \tag{9}$$

where

$$\sigma_1 := \mathbb{E}[\mathbb{V}ar(\widehat{\lambda}\widehat{q}_i|\mathbf{A})] = q_i(1-q_i)\left[\lambda^2\left(\frac{\kappa-1}{\kappa}\right)^2 h\left(\frac{\lambda\tau}{\kappa}\right) + \lambda\tau^{-1}\frac{\kappa-1}{\kappa}c_1\right], \tag{10}$$

$$\sigma_2 := \mathbb{E}[\mathbb{E}^2[\widehat{\lambda}\widehat{q}_i|\mathbf{A}]] - (\lambda q_i)^2(1+2B) = q_i^2\lambda\tau^{-1}\left\{1 + B\left[1-\lambda\tau(1+\kappa^{-1})\right]\right\}. \tag{11}$$

Proof of Proposition 3. *Proof of (8).* We use Properties P1 and P2 (Section 4.1). From P2, we obtain

$$\mathbb{E}[\widehat{\lambda}\widehat{q}_i|\mathbf{A}] = \widehat{\lambda}\mathbb{E}[\widehat{q}_i|\mathbf{A}] = \widehat{\lambda}q_i\mathbb{1}_{[A_i>0]}, \tag{12}$$

$$\mathbb{V}ar(\widehat{\lambda}\widehat{q}_i|\mathbf{A}) = \widehat{\lambda}^2\mathbb{V}ar(\widehat{q}_i|\mathbf{A}) = \widehat{\lambda}^2A_i^{-1}q_i(1-q_i)\mathbb{1}_{[A_i>0]}. \tag{13}$$

We have

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

where step (a) uses (12) and the fact that $\widehat{\lambda} = \tau^{-1}\sum_i A_i$; step (b) uses the independence of the A_i; the fact that $\mathbb{E}[A_i] = \lambda\tau/\kappa$; and the fact that $\mathbb{P}(A_i > 0) = 1 - \exp(-\lambda\tau/\kappa) = \rho$. In the last expression above, the term after unity in the square bracket is, by definition, the relative bias; that is, $B = -(\kappa - 1)/(1-\rho)/\kappa$.

Proof of (9). Write $X = \widehat{\lambda}\widehat{q}_i$, $\mu = \lambda q_i$, and $B = (\mathbb{E}[X]-\mu)/\mu$. The result is simply the identity $\mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2] - \mu^2(1+2B)$, where we further use the identities $\mathbb{E}[X^2] = \mathbb{E}[\mathbb{E}[X^2|\mathbf{A}]]$ and $\mathbb{E}[X^2|\mathbf{A}] = \mathbb{V}ar(X|\mathbf{A}) + \mathbb{E}^2[X|\mathbf{A}]$.

Proof of (10). The left side of (10) is, in view of (13),

$$\mathbb{E}[\mathbb{V}ar(\widehat{\lambda}\widehat{q}_i|\mathbf{A})] = q_i(1-q_i)\mathbb{E}[\widehat{\lambda}^2A_i^{-1}\mathbb{1}_{[A_i>0]}].$$

Expanding $\widehat{\lambda}^2$ as in

$$\widehat{\lambda}^2 = \tau^{-2}[(\sum_{j:j\ne i}A_j)^2 + 2A_i\sum_{j:j\ne i}A_j + A_i^2], \tag{14}$$

and using Property P1 in taking expectation, we find

$$\mathbb{E}[\widehat{\lambda}^2A_i^{-1}\mathbb{1}_{[A_i>0]}] = \tau^{-2}\left[\lambda\tau\frac{\kappa-1}{\kappa}\left(1+\lambda\tau\frac{\kappa-1}{\kappa}\right)h\left(\frac{\lambda\tau}{\kappa}\right) + 2\lambda\tau\frac{\kappa-1}{\kappa}\rho + \frac{\lambda\tau}{\kappa}\right].$$

A simple re-arrangement gives (10).

Proof of (11). Observe that $\mathbb{E}[\mathbb{E}^2[\widehat{\lambda}\widehat{q}_i|\mathbf{A}]] = q_i^2\mathbb{E}[\widehat{\lambda}^2\mathbb{1}_{[A_i>0]}]$ by (12). Expanding $\widehat{\lambda}^2$ as in (14) and using Property P1, we obtain

$$\mathbb{E}[\widehat{\lambda}^2\mathbb{1}_{[A_i>0]}] = \tau^{-2}\left[\rho\lambda\tau\frac{\kappa-1}{\kappa}\left(1+\lambda\tau\frac{\kappa-1}{\kappa}\right) + 2\lambda\tau\frac{\kappa-1}{\kappa}\frac{\lambda\tau}{\kappa} + \frac{\lambda\tau}{\kappa}\left(1+\frac{\lambda\tau}{\kappa}\right)\right].$$

A simple re-arrangement gives (11). □

4.2.2 Asymptotic Efficiency Relative to the Sales-Only Estimator

Here we study the asymptotic efficiency of the arrivals-and-sales estimator relative to the one that uses sale counts only. To define the latter, apply test price p_i over a period of length $\varDelta := \tau/\kappa$, and let S_i be the observed sale count. The *sales-only* estimator of the demand vector is defined as

$$\widehat{\lambda q}_i := \frac{S_i}{\varDelta}, \quad i = 1, \ldots, \kappa. \tag{15}$$

From the fact that S_i is a Poisson$(q_i\lambda\tau/\kappa)$ random variable, it follows that $\mathbb{E}[(\widehat{\lambda q}_i - \lambda q_i)^2] = \mathbb{V}ar[S_i/(\tau/\kappa)]] = q_i\lambda\kappa\tau^{-1}$.

The asymptotic efficiency is analyzed for a sequence of estimation problems. In problem n, the arrival rate is λ_n, and the learning time is τ_n. Depending on whether the set of feasible prices is finite or an infinite continuum, Method A is applied as follows. If the set of feasible prices is finite, $\{p_1, \ldots, p_\kappa\}$ say, then Method A(τ_n, κ_n) is applied at all prices in the feasible set (thus, $\kappa_n = \kappa$ for all n). If, instead, the set of feasible prices is an infinite continuum $[\underline{p}, \overline{p}]$ (with $\underline{p} < \overline{p}$), then Method A$(\tau_n, \kappa_n)$ is applied at the prices $p_i = \underline{p} + (i - 1/2)\ell_n$, $i = 1, \ldots, \kappa_n$, where $\ell_n = (\overline{p} - \underline{p})/\kappa_n$. (These prices are the midpoints of the partition of $[\underline{p}, \overline{p}]$ into κ_n equal intervals.) The MLE in (5) is denoted $(\widehat{\lambda}\widehat{q}_{1,n}, \ldots, \widehat{\lambda}\widehat{q}_{\kappa_n,n})$, and the sales-only estimator in (15) is denoted $(\widehat{\lambda q}_{1,n}, \ldots, \widehat{\lambda q}_{\kappa_n,n})$.

We now present our efficiency measures and describe the limit under which these are analyzed. *Global efficiency* is defined as

$$\mu_n := \frac{\mathbb{E}[\sum_{i=1}^{\kappa_n}(\widehat{\lambda q}_{i,n} - \lambda_n q_i)^2]}{\mathbb{E}[\sum_{i=1}^{\kappa_n}(\widehat{\lambda}\widehat{q}_{i,n} - \lambda_n q_i)^2]}. \tag{16}$$

Local efficiency is a simpler measure, where each sum above is replaced by a term corresponding to a single price. The limit we consider requires:

$$r_n := \frac{\lambda_n\tau_n}{\kappa_n} \to \infty, \quad \lambda_n\tau_n \exp\left(-\frac{\lambda_n\tau_n}{\kappa_n}\right) = \kappa_n r_n \exp(-r_n) \to 0 \quad \text{as } n \to \infty. \tag{17}$$

Less formally, the above says the following. First, the mean arrival count at any price, $r_n := \lambda_n\tau_n/\kappa_n$ grows large (thus the mean arrival count during learning, $\lambda_n\tau_n$, also grows large). In the finite-price-set case, $\kappa_n = \kappa$ is a bounded sequence, so the second condition follows from the first. In the infinite-price-set case we will need $\kappa_n \to \infty$ to "learn" a demand function defined on a continuum; here the second condition is non-trivial: it requires the growth of "information" r_n to dominate the growth of the price granularity κ_n.

Proposition 4 *(Asymptotic Efficiency Relative to the Sales-Only Estimator.) Consider a sequence of problems as in (17). Let Assumption 3 hold for all* $n \geq n_0$*, for some finite* n_0*.*

(a) For any price p_i, and with $q_i := q(p_i) \in (0, 1]$, we have

$$\lim_{n\to\infty} \frac{\mathbb{E}[(\widehat{\lambda q}_{i,n} - \lambda_n q_i)^2]}{\mathbb{E}[(\widehat{\lambda}\widehat{q}_{i,n} - \lambda_n q_i)^2]} = \frac{1}{1-q_i}. \tag{18}$$

(b) Let the set of feasible prices be $\{p_1, \ldots, p_\kappa\}$, for some finite κ. Put $q_i = q(p_i)$. Suppose $q_i \notin \{0,\ 1\}$ for all i. Then

$$\lim_{n\to\infty} \mu_n = \frac{\sum_{i=1}^{\kappa} q_i}{\sum_{i=1}^{\kappa} q_i(1-q_i)} =: \mu_\infty \geq 1. \tag{19}$$

(c) Let Assumption 4 hold and let $\kappa_n \to \infty$ as $n \to \infty$. Then

$$\lim_{n\to\infty} \mu_n = \frac{\int_{[\underline{p},\overline{p}]} q(p)\mathrm{d}p}{\int_{[\underline{p},\overline{p}]\cap\{p:q(p)<1\}} q(p)[1-q(p)]\mathrm{d}p} =: \mu_\infty \geq 1. \tag{20}$$

Proof of Proposition 4. *Proof of (18).* The MSE of the arrivals-and-sales estimator (the denominator in (18)) was analyzed in Proposition 3. Terms seen there are now denoted by appending the scale index n as an additional (rightmost) subscript; for example, B_n denotes the scale-n version of B in (8), that is, $B_n = -[(\kappa_n - 1)/\kappa_n]\exp(-\lambda_n\tau_n/\kappa_n)$; $\sigma_{1,n}$ denotes the scale-n version of σ_1 in (10); and so on for all terms there. All limits below are with respect to $n \to \infty$.

We need the asymptotic of $h(n)$ in (7) as $n \to \infty$; the result is

$$\lim_{n\to\infty} \frac{h(n)}{n^{-1}} = 1, \tag{21}$$

which is proven in Lemma 3, and which we write as $h(n) \sim n^{-1}$.

From (9) we have $\mathbb{E}[(\widehat{\lambda}\widehat{q}_{i,n} - \lambda_n q_i)^2] = \sigma_{1,n} + \sigma_{2,n}$, and we now derive the asymptotics of these terms. Equation (10) states that $\sigma_{1,n}$ is $q_i(1-q_i)$ times (the term in square brackets)

$$\begin{aligned}
&\lambda_n^2 \left(\frac{\kappa_n - 1}{\kappa_n}\right)^2 h\left(\frac{\lambda_n\tau_n}{\kappa_n}\right) + \lambda_n\tau_n^{-1}\frac{\kappa_n-1}{\kappa_n}c_{1,n} \\
&\overset{(a)}{\sim} \lambda_n^2 \left(\frac{\kappa_n - 1}{\kappa_n}\right)^2 \left(\frac{\lambda_n\tau_n}{\kappa_n}\right)^{-1} + \lambda_n\tau_n^{-1}\frac{\kappa_n-1}{\kappa_n}\left(2 + \frac{1}{\kappa_n - 1}\right) \\
&= \lambda_n\tau_n^{-1}\kappa_n
\end{aligned}$$

where in step (a) we observed: (i) $h(\lambda_n\tau_n/\kappa_n) \sim (\lambda_n\tau_n/\kappa_n)^{-1}$ (due to $\lambda_n\tau_n/\kappa_n \to \infty$ and (21)); and (ii) $c_{1,n} = \left[h(\lambda_n\tau_n/\kappa_n) + 2\rho_n + (\kappa_n - 1)^{-1}\right] \sim 2 + (\kappa_n - 1)^{-1}$ (this follows from (i), the fact that $\lambda_n\tau_n/\kappa_n \to \infty$, and the fact that $\lim_{n\to\infty}\rho_n = \lim_{n\to\infty}[1-\exp(-\lambda_n\tau_n/\kappa_n)] = 1$). Thus, $\sigma_{1,n} \sim q_i(1-q_i)\lambda_n\tau_n^{-1}\kappa_n$.

Now consider (11). Note that $\left[1 - \lambda_n\tau_n(1+\kappa_n^{-1})\right] \sim -\lambda_n\tau_n$ (since $\lambda_n\tau_n \to \infty$, and the remainder term is bounded in n). Now, noting that $\lim_{n\to\infty}\lambda_n\tau_n B_n = 0$ by the second condition in (17), we obtain $\sigma_{2,n} \sim q_i^2\lambda_n\tau_n^{-1}$.

If $q_i < 1$, then $\sigma_{2,n}$ is negligible: $\sigma_{2,n}/\sigma_{1,n} \to 0$, since $\kappa_n \to \infty$. In contrast, if $q_i = 1$, then $\sigma_{1,n}$ is zero. In summary,

$$\mathbb{E}[(\widehat{\lambda}\widehat{q}_{i,n} - \lambda_n q_i)^2] \sim \begin{cases} \sigma_{1,n} \sim q_i(1-q_i)\lambda_n\tau_n^{-1}\kappa_n & \text{if } q_i < 1 \\ \sigma_{2,n} \sim \lambda_n\tau_n^{-1} & \text{if } q_i = 1. \end{cases} \tag{22}$$

In the n-th problem, the sale count S_i in (15) is Poisson$(q_i\lambda_n\tau_n/\kappa_n)$, and thus $\mathbb{E}[(\widehat{\lambda q}_{i,n} - \lambda_n q_i)^2] = q_i\lambda_n\tau_n^{-1}\kappa_n$; combining this with (22), we obtain (18).

Proof of (19). This follows easily from the asymptotics in part (a).

Proof of (20). This is a consequence of the asymptotics in part (a), combined with the fact that the inter-price distance $\ell_n = (\overline{p} - \underline{p})/\kappa_n$ goes to zero. In more detail,

$$\begin{aligned}
&\mathbb{E}[\sum_{i=1}^{\kappa_n}(\widehat{\lambda}\widehat{q}_{i,n} - \lambda_n q_i)^2] \\
&\overset{(a)}{\sim} \ell_n^{-1}\left(\lambda_n\tau_n^{-1}\kappa_n \sum_{i:q_i<1} \ell_n q(p_i)[1-q(p_i)] + \lambda_n\tau_n^{-1}\sum_{i:q_i=1}\ell_n\right) \\
&\overset{(b)}{\sim} \ell_n^{-1}\left(\lambda_n\tau_n^{-1}\kappa_n \int_{[\underline{p},\overline{p}]\cap\{p:q(p)<1\}} q(p)[1-q(p)]\mathrm{d}p + \lambda_n\tau_n^{-1}\int_{[\underline{p},\overline{p}]\cap\{p:q(p)=1\}}\mathrm{d}p\right) \\
&\overset{(c)}{\sim} \ell_n^{-1}\lambda_n\tau_n^{-1}\kappa_n \int_{[\underline{p},\overline{p}]\cap\{p:q(p)<1\}} q(p)[1-q(p)]\mathrm{d}p
\end{aligned}$$

where step (a) uses (22); in step (b), it follows from $\ell_n \to 0$ that the sums are Riemann sums and converge to the indicated (Lebesgue) integrals; and step (c) uses that the second term in the sum to the left vanishes relative to the first term (due to $\kappa_n \to \infty$ and Assumption 4(c)). A similar argument gives $\mathbb{E}[\sum_{i=1}^{\kappa_n}(\widehat{\lambda q}_{i,n} - \lambda_n q_i)^2] \sim \ell_n^{-1}\lambda_n\tau_n^{-1}\kappa_n \int_{[\underline{p},\overline{p}]} q(p)\mathrm{d}p$. From the asymptotics given above, result (20) follows immediately. □

Remark 1. The asymptotic efficiencies (the right side of (18), (19), and (20)) are always no smaller than unity, and depend on $q()$ but not the arrival rate. They prove the theoretical superiority of the MLE and quantify the efficiency gain relative to using only sale data.

Example 1. Consider the (linear) case $q(p) = a - bp$ for $p \in [\underline{p}, \overline{p}]$, where $a > 0$, $b \geq 0$. The requirement $q(p) \in [0,1]$ for all $p \in [\underline{p}, \overline{p}]$ is equivalent to constraints $q(\underline{p}) = a - b\underline{p} \leq 1$ and $q(\overline{p}) = a - b\overline{p} \geq 0$. To give a simple formula for the asymptotic (global) efficiency, we take $[\underline{p}, \overline{p}] = [0,1]$ with constraints $q(0) = a \leq 1$ and $q(1) = a - b \geq 0$ and find $\mu_\infty = \mu_\infty(a,b) = 1/[1 - c - b^2/(12c)] > 1$, where $c = a - b/2 > 0$. Indicatively (for any $\underline{p} < \overline{p}$): $q(\underline{p}) = 1$ and $q(\overline{p}) = 0$ gives $\mu_\infty = 3$; $q(\underline{p}) = 1/2$ and $q(\overline{p}) = 0$ gives $\mu_\infty = 3/2$; and $q(\underline{p}) = 1$ 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 $\mathcal{D}$)

Following the ideas in Besbes and Zeevi (2009, Section 4.3), we consider a regime in which both the size of the initial inventory as well as potential demand grow proportionally large. In particular, for a market of size n, where n is a positive integer, the initial inventory and the arrival rate are now assumed to be given by

$$x_n = nx, \quad \lambda_n = n\lambda, \tag{23}$$

while the purchase-probability $q()$ is fixed for all n. Thus, the index n determines the order of magnitude of both inventory and rate of demand.

We elaborate the demand estimator (6) for the n-th system, including the connection with the following events during learning, which we show to be rare: (i) stock-out (inventory is depleted before selling ends); or (ii) no arrivals occur during a test period. For any test price p_i ($i \in \{1, \ldots, \kappa_n\}$) we write $q_i = q(p_i)$. Price p_i is applied from time $(i-1)\Delta_n$ to time $i\Delta_n$, where $\Delta_n = \tau_n/\kappa_n$; this is the *i-th test period.* The mean arrival count during any such test period is $\lambda n \Delta_n =: \lambda r_n$, where $r_n = n\Delta_n = n\tau_n/\kappa_n$. Let $N(\cdot)$ be a unit-rate Poisson process, and put $N_n = N(\lambda n \tau_n)$ and $N_{i,n} = N(\lambda i r_n) - N(\lambda(i-1)r_n)$. Moreover, for each $i = 1, \ldots, \kappa_n$, put $S_{i,n} = \sum_{j=1}^{N_{i,n}} I_j$, where conditionally on $N_{i,n}$, the sequence $I_1, I_2, \ldots$ consists of independent Bernoulli(q_i) variables that are independent of $N_{i,n}$. In particular, we have the probability laws $N_{i,n} \sim \text{Poisson}(\lambda r_n)$, $N_n \sim \text{Poisson}(\lambda n \tau_n)$, and $(S_{i,n}|N_{i,n)} \sim \text{Binomial}(N_{i,n}, q_i)$.

The estimator (6) has the representation $\widehat{\lambda}_n = \min(N_n, nx)/(n\tau_n)$ (estimated arrival rate); and $\widehat{q}_{i,n} = \min(S_{i,n}, nx) \min(N_{i,n}, nx)^{-1} \mathbb{1}_{[N_{i,n}>0]}$ (estimated purchase probability) for $i = 1, \ldots, \kappa_n$. The condition $N_n < nx$ will occur with high probability by enforcing $\tau_n \to 0$. This condition implies that no stock-out occurs during the learning phase $[0, \tau_n]$; and the latter implies the following: $\widehat{\lambda}_n = N_n/(n\tau_n)$, where N_n is observed as the total arrival count during the test period $[0, \tau_n]$; and $\widehat{q}_{i,n} = S_{i,n} N_{i,n}^{-1} \mathbb{1}_{[N_{i,n}>0]}$, where $N_{i,n}$ and $S_{i,n}$ are observed as the arrival count and sale count, respectively, during the i-th test period. Moreover, the condition $N_{i,n} > 0$ will occur with high probability by enforcing $r_n \to \infty$.

Consider the scale-n system associated to a fixed demand model $(\lambda, q) \in \mathcal{D}$. Let $J_n^\pi = J_n^\pi(x, T|\lambda, q)$ denote the expected revenue under policy π (Algorithm AS). The optimal value to the scale-n version of the full-information relaxation (3) is easily seen to be nJ^D, where J^D is that of the unscaled problem and is shown following (4) (Besbes and Zeevi 2009).

Our main result is as follows.

Theorem 1 *Consider the class $\mathcal{D}$ defined following Assumption 2. Let $\pi_n := \pi(\tau_n, \kappa_n)$ be given by Algorithm AS. If*

$$\tau_n \asymp \left(\frac{\log n}{n}\right)^{1/4}, \quad \kappa_n \asymp \left(\frac{n}{\log n}\right)^{1/4}, \tag{24}$$

then there exists a constant K_0 and a finite $\underline{n}$ such that for all $n \geq \underline{n}$,

$$\inf_{(\lambda,q)\in\mathcal{D}} \frac{J_n^{\pi_n}}{nJ^D} \geq 1 - K_0 \left(\frac{\log n}{n}\right)^{1/4}. \tag{25}$$

The main result supporting the proof (other than ideas from Besbes and Zeevi (2009)) is Proposition 5, which is presented and proven before proving Theorem 1. In this result, errors $\widehat{\lambda}_n - \lambda$ are handled via a large-deviation bound for i.i.d. (independent and identically distributed) Poisson summands (Lemma 1). Moreover, each error of the form $\widehat{q}_{i,n} - q_i$ is handled via a large-deviation bound for i.i.d. Bernoulli summands, known as Hoeffding's inequality (Lemma 2); in applying it, we need that the number of summands, which is the arrival count $N_{i,n}$, be large enough; we ensure that this condition holds with high probability by requiring that n be sufficiently large; this contrasts to Besbes and Zeevi (2009), whose result applies to all $n = 1, 2, \ldots$.

Proposition 5 will require Condition 1 below. This condition does not restrict τ_n and κ_n as Theorem 1 does. This makes the result more broadly applicable, and the proof more transparent, in our view. However, there is a small "cost": the construction of (formulas for) a finite $\underline{n}$ such that the result holds for all $n \geq \underline{n}$ is less specific than is possible. Formulas for $\underline{n}$ specific to Theorem 1 are given in Remark 2 following the proof of the theorem.

Condition 1 *Put $r_n := n\tau_n/\kappa_n$ and suppose the following hold:*

(a) $\kappa_n \to \infty$ as $n \to \infty$.
(b) For all $n \geq 1$, we have $\underline{c}_\tau n^{\psi_1} \leq n\tau_n \leq \overline{c}_\tau n^{\psi_2}$ for some positive $\underline{c}_\tau, \overline{c}_\tau$ and $0 < \psi_1 \leq \psi_2 < 1$.
(c) For all $n \geq 1$, we have $r_n \geq \underline{c}_r n^\beta$ for some $\underline{c}_r > 0$ and $\beta > 0$.

Proposition 5 *(Bounds on worst-case estimation error.) Fix $\eta \geq 2$. Let Condition 1 hold. Let $\delta_n = \delta_n(\lambda) = [4\eta\lambda \log n/(n\tau_n)]^{1/2}$; and $\ell_n = \ell_n(\lambda) = r_n(\lambda - \zeta_n)$, where $\zeta_n = \zeta_n(\lambda) = [2\eta\lambda\log(n)/r_n]^{1/2}$. Fix any $i \in \{1, \ldots, \kappa_n\}$.*

(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 $\epsilon_n = \epsilon_n(\lambda) = [\eta \log n/(2\ell_n)]^{1/2}$, we have

$$\mathbb{P}\{|\widehat{\lambda}_n\widehat{q}_{i,n} - \lambda q_i| \leq \delta_n + \lambda\epsilon_n\} \geq \mathbb{P}\{|\widehat{\lambda}_n - \lambda| \leq \delta_n,\ |\widehat{q}_{i,n} - q_i| \leq \epsilon_n\} \geq 1 - C_1 n^{-\eta} \tag{26}$$

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 $n \geq \underline{n}$,

$$\sup_{(\lambda,q)\in\mathcal{D}} \mathbb{P}\left\{|\widehat{\lambda}_n\widehat{q}_{i,n} - \lambda q_i| > (1+\alpha)\left(\frac{\overline{\lambda}\eta}{2}\right)^{1/2}\left(\frac{\log n}{r_n}\right)^{1/2}\right\} \leq C_1 n^{-\eta} \tag{27}$$

and

$$\sup_{(\lambda,q)\in\mathcal{D}} \mathbb{P}\left\{|\widehat{q}_{i,n} - q_i| > (1+\alpha)\left(\frac{\eta}{2\underline{\lambda}}\right)^{1/2}\left(\frac{\log n}{r_n}\right)^{1/2}\right\} \leq C_1 n^{-\eta}. \tag{28}$$

Proof of Proposition 5. All limits below are with respect to $n \to \infty$.

Note that $\ell_n = \lambda r_n\{1 - [(2\eta/\lambda)\log(n)/r_n]^{1/2}\}$ with $r_n \geq \underline{c}_r n^\beta$. We use (here and elsewhere) two properties of $\log(x)/x^\beta$ for real $x \geq 1$ and $\beta > 0$: (i) $\log(x)/x^\beta$ is maximized at $x_* = e^{1/\beta}$; moreover $\sup_{x\geq 1}\log(x)/x^\beta = \log(x_*)/x_*^\beta = 1/(\beta e)$; and (ii) for $x \in [e^{1/\beta}, \infty)$, $\log(x)/x^\beta$ decreases to zero.

We construct a finite $n_1(\lambda)$ such that

$$\ell_n(\lambda) > 0 \quad \text{for all } n \geq n_1(\lambda). \tag{29}$$

The condition $\ell_n(\lambda) > 0$ is equivalent to $\log(n)/r_n < \lambda/(2\eta)$; and since $r_n \geq \underline{c}_r n^\beta$, a sufficient condition for the latter is $\log(n)/n^\beta < \underline{c}_r^{-1}\lambda/(2\eta)$. Since $\log(n)/n^\beta$ is maximized at $x^* = e^{1/\beta}$ and decreases to zero for $x \in [x^*, \infty)$, it suffices to set $n_1(\lambda) := \min\{n : n \geq e^{1/\beta},\ \log(n)/n^\beta < \underline{c}_r^{-1}\lambda/(2\eta)\} < \infty$. Observe that $n_1(\lambda)$ is non-increasing in λ.

Next, we construct a finite $n_2(\lambda)$ such that

$$n\tau_n(\lambda + \delta_n) < nx \quad \text{for all } n \geq n_2(\lambda). \tag{30}$$

Since $n\tau_n \leq \overline{c}_\tau n^{\psi_2}$, a sufficient condition is $\overline{c}_\tau n^{\psi_2}(\lambda + \sup_{n\geq 1}\delta_n) < nx$. Note that $\sup_n \delta_n \leq \sup_n[4\eta\lambda\log n/(\underline{c}_\tau n^{\psi_1})]^{1/2} = [4\eta\lambda/(e\psi_1\underline{c}_\tau)]^{1/2}$. Thus it suffices to set

$$n_2(\lambda) = \lceil\{\overline{c}_\tau(\lambda + [4\eta\lambda/(e\psi_1\underline{c}_\tau)]^{1/2})/x\}^{1/(1-\psi_2)}\rceil. \tag{31}$$

Observe that $n_2(\lambda)$ is increasing in λ.

Put $n_0(\lambda) := \max(n_1(\lambda), n_2(\lambda))$; this number is finite and such that the conditions in (29) and (30) hold for all $n \geq n_0(\lambda)$.

Proof of part (a). Define the events

$$\begin{aligned} U_n &= \{|N_n - \lambda n\tau_n| \leq n\tau_n\delta_n\}, \\ L_n &= \{N_{i,n} \geq \ell_n\}, \\ D_n &= \{|S_{i,n} - N_{i,n}q_i| \leq N_{i,n}\epsilon_n\}. \end{aligned}$$

An essential idea behind the proof is: for $n \geq n_0(\lambda)$, we have

$$G_n := U_n \cap L_n \cap D_n \overset{(a)}{\subseteq} \{|\widehat{\lambda}_n - \lambda| \leq \delta_n, |\widehat{q}_{i,n} - q_i| \leq \epsilon_n\} \overset{(b)}{\subseteq} \{|\widehat{\lambda}_n\widehat{q}_{i,n} - \lambda q_i| \leq \delta_n + \lambda\epsilon_n\}, \tag{32}$$

which we now justify. To justify step (a), first note that on the set U_n a stock-out is excluded and thus $\widehat{\lambda}_n = N_n/(n\tau_n)$, since $N_n \leq n\tau_n(\lambda + \delta_n) < nx$, where the latter inequality holds because $n \geq n_2(\lambda)$. Further, note that on the set L_n a no-arrival event (that is, a zero arrival count during the i-th test period) is excluded since $\ell_n > 0$ by the fact that $n \geq n_1(\lambda)$. The set $U_n \cap L_n$ excludes both a stock-out and a no-arrival, and thus on this set we have $\widehat{q}_{i,n} = S_{i,n}/N_{i,n}$ with $N_{i,n} > 0$. Now $|\widehat{\lambda}_n - \lambda| = |N_n/(n\tau_n) - \lambda| \leq \delta_n$, where the inequality holds by definition of U_n. Moreover, $|\widehat{q}_{i,n} - q_i| = |S_{i,n}N_{i,n}^{-1} - q_i| \leq \epsilon_n$, where the inequality holds by definition of D_n. This completes the proof of step (a). Step (b) follows from the triangle inequality $|\widehat{\lambda}_n\widehat{q}_{i,n} - \lambda q_i| = |\widehat{q}_{i,n}(\widehat{\lambda}_n - \lambda) +$

$\lambda(\widehat{q}_{i,n} - q_i)| \leq |\widehat{\lambda}_n - \lambda| + \lambda|\widehat{q}_{i,n} - q_i|$. The other essential idea is a union bound on the complement of G_n, namely

$$\mathbb{P}({G_n}^c) \leq \mathbb{P}({U_n}^c) + \mathbb{P}({L_n}^c) + \mathbb{P}(L_n \cap {D_n}^c), \tag{33}$$

together with the fact that each probability on the right is of order $n^{-\eta}$, which we now show. By Lemma 1, result (44), we have for all $n \geq 1$,

$$\mathbb{P}({U_n}^c) = \mathbb{P}\{|N_n - \lambda n\tau_n| > n\tau_n\delta_n\} \leq 2C_0(\lambda)n^{-\eta}. \tag{34}$$

Moreover, Lemma 1, result (45), shows that for all $n \geq 1$,

$$\mathbb{P}({L_n}^c) = \mathbb{P}\{N_{i,n} < \ell_n\} \leq n^{-\eta}. \tag{35}$$

We claim that for all $n \geq n_1(\lambda)$

$$\begin{aligned} \mathbb{P}(L_n \cap {D_n}^c) &= \mathbb{P}\{N_{i,n} \geq \ell_n,\ |S_{i,n} - N_{i,n}q_i| > N_{i,n}\epsilon_n\} \\ &\overset{(a)}{\leq} 2\exp(-2\ell_n\epsilon_n^2) \\ &\overset{(b)}{=} 2n^{-\eta}, \end{aligned} \tag{36}$$

which we now explain. Since $S_{i,n}$ is the sum of $N_{i,n}$ independent Bernoulli(q_i) variables and since $N_{i,n}$ is nonnegative-integer-valued, the condition $N_{i,n} \geq \ell_n > 0$ (note that $\ell_n > 0$ holds because $n \geq n_1(\lambda)$) implies that $N_{i,n} \geq \max(\ell_n, 1)$. Now an application of Lemma 2 (for $N_{i,n}$ independent Bernoulli summands) justifies step (a). Finally, step (b) is a direct consequence of the choice of ϵ_n. Inserting the bounds in (34), (35), and (36) into (33), we obtain $\mathbb{P}({G_n}^c) \leq (2C_0(\lambda) + 3)n^{-\eta}$, and result (26) now follows from (32).

Proof of part (b). Recall that $n_1(\lambda)$ is non-increasing and $n_2(\lambda)$ is non-decreasing in λ. Put $m = \max\{n_1, n_2\}$, where $n_1 := \sup_{\lambda \in \mathcal{D}} n_1(\lambda) = n_1(\underline{\lambda})$ and $n_2 := \sup_{\lambda \in \mathcal{D}} n_2(\lambda) = n_2(\overline{\lambda})$. For all $n \geq m$, the bound (26) holds for all $(\lambda, q) \in \mathcal{D}$. Thus

$$\mathbb{P}\left\{|\widehat{\lambda}_n\widehat{q}_{i,n} - \lambda q_i| > \sup_{(\lambda,q)\in\mathcal{D}} [\delta_n(\lambda) + \lambda\epsilon_n(\lambda)]\right\} \leq C_1(\lambda)n^{-\eta} \quad \text{for all } (\lambda, q) \in \mathcal{D} \text{ and } n \geq m.$$

Thus, for all $n \geq m$,

$$\sup_{(\lambda,q)\in\mathcal{D}} \mathbb{P}\left\{|\widehat{\lambda}_n\widehat{q}_{i,n} - \lambda q_i| > \sup_{(\lambda,q)\in\mathcal{D}} [\delta_n(\lambda) + \lambda\epsilon_n(\lambda)]\right\} \leq \sup_{(\lambda,q)\in\mathcal{D}} C_1(\lambda)n^{-\eta} = C_1(\underline{\lambda})n^{-\eta}.$$

Result (27) now follows by noting that

$$\begin{aligned} &\sup_{(\lambda,q)\in\mathcal{D}} [\delta_n(\lambda) + \lambda\epsilon_n(\lambda)] \\ &\overset{(a)}{=} \sup_{(\lambda,q)\in\mathcal{D}} \left[2(\eta\lambda)^{1/2}\left(\frac{\log n}{n\tau_n}\right)^{1/2} + \lambda\left(\frac{\eta}{\lambda}\right)^{1/2}\left(\frac{\log n}{2r_n}\right)^{1/2}\left(1 - \sqrt{\frac{2\eta\log n}{\lambda r_n}}\right)^{-1/2}\right] \\ &\overset{(b)}{\sim} \left(\frac{\overline{\lambda}\eta\log n}{2r_n}\right)^{1/2} \end{aligned}$$

where step (a) simply inserts the expressions for δ_n and ϵ_n (via ℓ_n); step (b) uses that the first term vanishes relative to the second one, since $n\tau_n/r_n = k_n \to \infty$; and, on the rightmost term, we noted that $(1-\sqrt{(2\eta \log n)/(\lambda r_n)})^{-1/2} \sim 1$. This completes the proof of (27). Result (28) follows by an entirely analogous argument; here, the supremum across λ results in a $\underline{\lambda}^{-1/2}$ factor because in the pre-supremum, λ enters as $\lambda^{-1/2}$. □

Proof of Theorem 1. The proof adapts and refines the argument in the proof of Besbes and Zeevi (2009, Proposition 1). Our basic tool is Proposition 5; its role is somewhat analogous to that of Besbes and Zeevi (2009, Online Companion, Lemma 2) in their proof.

As an auxiliary step, we verify that Proposition 5 is in force. The assumption $\tau_n \asymp (\log n/n)^{1/4}$ means that for all $n \geq 1$ we have $\underline{c}_\tau(\log n/n)^{1/4} \leq \tau_n \leq \overline{c}_\tau(\log n/n)^{1/4}$ for some positive $\underline{c}_\tau, \overline{c}_\tau$. Similarly, the assumption $\kappa_n \asymp (n/\log n)^{1/4}$ means that for all $n \geq 1$ we have $\underline{c}_\kappa(n/\log n)^{1/4} \leq \kappa_n \leq \underline{c}_\kappa(n/\log n)^{1/4}$ for some positive $\underline{c}_\kappa, \underline{c}_\kappa$. We now verify that Condition 1 (stated immediately before Proposition 5) is in force. Part (a) is clearly in force. Since $n\tau_n \asymp n^{3/4}(\log n)^{1/4}$, part (b) is satisfied: take $\psi_1 = 3/4$ and any $\psi_2 > 3/4$. For part (c), we have $r_n := n\tau_n/\kappa_n \asymp (n \log n)^{1/2}$; thus part (c) is in force with $\beta = 1/2$. This completes the verification. Let $\underline{n} = \underline{n}(\alpha)$ be constructed as in Proposition 5 part (b), where the underlying $n_1(\lambda)$ and $n_2(\lambda)$ are tailored to $\{\tau_n\}$ and $\{\kappa_n\}$ as specified here; details on such $n_1(\lambda)$ and $n_2(\lambda)$ are given in Remark 2 following the proof.

Set

$$u_n = \max\left\{\tau_n, \kappa_n^{-1}, \left(\frac{\log n}{r_n}\right)^{1/2}\right\}. \tag{37}$$

The various errors involved in bounding the regret will all be shown to be of order u_n. Note that $u_n \asymp (\log n/n)^{1/4}$.

Onwards, the notation and proof structure are similar to that in Besbes and Zeevi (2009, Proposition 1). Let $X_n^{(L)} = \lambda \sum_{i=1}^{\kappa_n} q(p_i) n\Delta_n$, $X_n^{(P)} = \lambda q(\widehat{p})n(T-\tau_n)$, and put $Y_n^{(L)} = N(X_n^{(L)})$, $Y_n^{(P)} = N(X_n^{(P)})$, and $Y_n = N(X_n^{(L)} + X_n^{(P)})$.

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

$$J_n^\pi \geq \mathbb{E}[\widehat{p} \min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\}]. \tag{38}$$

Step 2. In what follows, we separate two cases: $\Lambda(\overline{p}) \leq x/T$ and $\Lambda(\overline{p}) > x/T$.

Case 1. Suppose that $\Lambda(\overline{p}) \leq x/T$. We start with auxiliary facts about the full-information relaxation (3). It is easy to see that $J^D = p^D \Lambda(p^D)T$, where $\Lambda(p) = \lambda q(p)$. The trivial case $J^D = 0$ is excluded by a simple argument paralleling Besbes and Zeevi (2009, Online Companion, Lemma 3); specifically, by Assumption 1(iii), there exists a price, p_0 say, such that $p_0 q(p_0) = m_a > 0$. Let $t_0 = \min\{T,\ x/\Lambda(\underline{p})\} = \min\{T, x/\underline{\lambda}\} > 0$. The solution that prices at p_0 up to time t_0 and at p_∞ afterwards is feasible for problem (3), since $\Lambda(p_0)t_0 \leq$

$\Lambda(\underline{p})t_0 \leq x$. This solution yields revenue $p_0\Lambda(p_0)t_0 \geq \underline{\lambda}m_a t_0 := m^D$. It follows that $J^D \geq m^D > 0$.

We will now show that for all $n \geq \underline{n}$, $J_n^{\pi_n}$ is at least nJ^D minus a loss of order $O(u_n)$. Note that $\min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\} \geq Y_n^{(P)} - (Y_n - nx)^+$. From (38), we have

$$J_n^\pi \geq \mathbb{E}[\widehat{p}Y_n^{(P)}] - \overline{p}\mathbb{E}[(Y_n - nx)^+] = \mathbb{E}[r(\widehat{p})]\lambda n(T - \tau_n) - \overline{p}\mathbb{E}[(Y_n - nx)^+] \quad (39)$$

by noting $\widehat{p} \leq \overline{p}$ 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. In Lemma 6, taking steps broadly similar to those in Besbes and Zeevi (2009, Lemma 4), and using the results in Proposition 5, we show that, for all $n \geq \underline{n}$, $\mathbb{E}[r(\widehat{p})] \geq r(p^D) - Ru_n - R_2/n^{\eta-1}$ for positive constants R and R_2.

Further, in Lemma 8, using steps similar to those in Besbes and Zeevi (2009, Lemma 5) and using the results in Proposition 5, we show that, for all $n \geq \underline{n}$, $\mathbb{E}[(Y_n - nx)^+] \leq K_E n u_n$ for some constant $K_E > 0$. We conclude that for all $n \geq \underline{n}$,

$$\begin{aligned} J_n^\pi &\geq \left[\lambda r(p^D) - Ru_n - \frac{R_2}{n^{\eta-1}}\right] n(T - \tau_n) - \overline{p}K_E n u_n \\ &\geq n\lambda r(p^D)T - K_1 n(u_n + \tau_n), \\ &= nJ^D - K_1 n(u_n + \tau_n), \end{aligned} \quad (40)$$

where K_1 is a suitable constant. We conclude that

$$\frac{J_n^\pi}{J_n^D} = \frac{J_n^\pi}{nJ^D} \geq 1 - \frac{K_1}{J^D}(u_n + \tau_n) \geq 1 - \frac{K_1}{m^D}(u_n + \tau_n). \quad (41)$$

Case 2. Suppose that $\Lambda(\overline{p}) > x/T$. Here, $p^D = \overline{p}$ and $J_n^D = n\overline{p}x$. Recall (38), where $\min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\}$ is a lower bound on the quantity sold during the pricing phase. In Lemma 9 we show that this bound is close to nx and moreover the price $\widehat{p}$ is close to p^D, with probability that is high for all $n \geq \underline{n}$. Specifically, the key event is $\mathcal{A} := \{\omega : \min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\} \geq nx - \tilde{K}_Y n u_n, |\widehat{p} - p^D| \leq K_c u_n\}$, where $\tilde{K}_Y$ and K_c are constants defined in Lemma 9. The (expected) revenue generated by π_n can be bounded as follows, for all $n \geq \underline{n}$:

$$\begin{aligned} J_n^\pi &\geq \mathbb{E}[\widehat{p}\min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\}] \\ &\overset{(a)}{\geq} \mathbb{E}[(p^D - K_c u_n)\min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\}|\mathcal{A}]\mathbb{P}(\mathcal{A}) \\ &\overset{(b)}{\geq} (p^D - K_c u_n)(nx - \tilde{K}_Y n u_n)\left(1 - \frac{C_2}{n^{\eta-1}}\right) \\ &\geq p^D nx - K_2 n u_n, \end{aligned} \quad (42)$$

where both (a) and (b) follow from the definition of $\mathcal{A}$ and Lemma 9, and where C_2 and K_2 are suitable constants. We conclude that for all $n \geq \underline{n}$,

$$\frac{J_n^\pi}{J_n^D} = \frac{J_n^\pi}{nJ^D} \geq 1 - \frac{K_2}{\overline{p}x}u_n. \quad (43)$$

Step 3. Combining the results (41) and (43), we set $K_0 = \max\{2K_1/m^D, K_2/(\overline{p}x)\}$ and have for all $n \geq \underline{n}$,

$$\inf_{(\lambda,q)\in\mathcal{D}} \frac{J_n^{\pi_n}}{nJ^D} \geq 1 - K_0 \left(\frac{\log n}{n}\right)^{1/4}.$$

This completes the proof. The full details of the proof are seen in Lemmata 1, 4, 5, 6, 7, 8, and 9. □

Remark 2. To specify more concretely the notion of a sufficiently large n, we now assume that $\tau_n = c_\tau(\log n/n)^{1/4}$ and $\kappa_n = c_\kappa(n/\log n)^{1/4}$ for some positive c_τ, c_κ. Such sequences would typically be used, with κ_n rounded to the closest integer; in what follows, we ignore this rounding to simplify the analysis. Then, we have: (i) $n\tau_n = c_\tau n^{3/4}(\log n)^{1/4}$; (ii) $\delta_n = [4\eta\lambda c_\tau^{-1}(\log n/n)^{3/4}]^{1/2}$; and (iii) $r_n = n\tau_n/\kappa_n = (c_\tau/c_\kappa)(n\log n)^{1/2}$ and thus $\log n/r_n = (c_\kappa/c_\tau)(\log n/n)^{1/2}$. $n_1(\lambda) := \min\{n : n \geq e,\ \log n/n < [(c_\kappa/c_\tau)^{-1}\lambda/(2\eta)]^2\}$ is sufficient for the requirement (29). Moreover, from (i)-(ii) it follows that $n_2(\lambda) := \min\{n : n \geq e, (\log n/n) < x/[c_\tau(\lambda+\delta_n)]^4\}$ is sufficient for the requirement (30) (since δ_n is monotone decreasing for $n \geq e$). 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

In this section, we briefly study the setting in Besbes and Zeevi (2009). We recall that their model postulates a demand function $\Lambda(p)$ directly, without reference to arrivals, and that their model class $\mathcal{D}_{BZ}$ is larger than ours (see below Assumption 2). Theorem 2 below shows that the algorithm in Besbes and Zeevi (2009, Section 4.1) can be adapted in order to improve the convergence rate. This is a sharpening of Besbes and Zeevi (2009, Proposition 3) and is achieved via deviation thresholds of a smaller order; specifically, in Besbes and Zeevi (2009, Proposition 3) these thresholds are (each) a constant times

$$u_n^{(BZ)} := (\log n)^{1/2} \max\left\{\frac{1}{\kappa_n}, \frac{1}{(n\Delta_n)^{1/2}}\right\} \asymp \frac{(\log n)^{1/2}}{n^{1/4}},$$

where $n\Delta_n := n\tau_n/\kappa_n \asymp n^{1/2}$. In Theorem 2 these thresholds are (each) a constant times the u_n defined in (37), for which $u_n \asymp (\log n/n)^{1/4}$. We see that $u_n/u_n^{(BZ)} \asymp (\log n)^{-1/4}$.

Theorem 2 *Let $\mathcal{D}_{BZ}$ be the class of demand models defined in Besbes and Zeevi (2009, Section 4.2). Let $\pi_n := \pi(\tau_n, \kappa_n)$ be given by Algorithm $\pi(\tau,\kappa)$ defined in Section 4.1 there. If we set*

$$\tau_n \asymp \left(\frac{\log n}{n}\right)^{1/4}, \quad \kappa_n \asymp \left(\frac{n}{\log n}\right)^{1/4}.,$$

then there exists a constant K_0' such that for all $n \geq 1$,

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

Proof of Theorem 2. The proof is a straightforward adaptation of the proof of Theorem 1; we give an overview and omit many details. In full analogy to Theorem 1, probability bounds are obtained for events defined via deviations proportional to u_n as in (37), for which we have $u_n \asymp (\log n/n)^{1/4}$. For example, to see how we 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 $K_c' = 4\underline{M}'^{-1}_\Lambda \max\{c_1, c_2\}$, where: $c_1 := \overline{M}'_\Lambda(\overline{p} - \underline{p})/2$ and $c_2 := 2(\eta M)^{1/2}$. The proof proceeds as in Lemma 5; in step (e) the Poisson large-deviation bound (Lemma 1) replaces (28); we obtain, for all $n \geq 1$, $\mathbb{P}\{|\widehat{p}^c - p^c| > K_c' u_n\} \leq C_0(M)n^{-\eta+1}$. (These deviations are of a smaller order than those in Besbes and Zeevi (2009, Online Companion, Lemma 4, Step 1), while the probability bound is the same.) In analogy to the above, we proceed as follows: the constants multiplying u_n to form the deviation events in Lemmata 4, 5, 6, 7, 8, and 9 are increased, when necessary, by a factor no larger than $2\sqrt{2}$; these deviations are no less than a positive constant times $(\log n/r_n)^{1/2}$ (since $(\log n/r_n)^{1/2} \asymp (\log n/n)^{1/4}$); it follows that the Poisson large-deviation bound (Lemma 1) applies, and it follows that, for all $n \geq 1$, the probabilities of these deviation events are no larger than a constant times $n^{-\eta+1}$. Onwards, the proof is the same as that of Theorem 1 and Besbes and Zeevi (2009, Proposition 3). □

We note that Theorem 2 holds for all $n \geq 1$, in contrast to Theorem 1, which requires that n be sufficiently large.

Remark 3. Theorem 2, in parallel with Theorem 1, prescribes that the parameters τ_n and κ_n grow at rates slightly different from those in Besbes and Zeevi (2009, Proposition 3). Specifically, the time spent learning, τ_n, is larger by a factor $(\log n)^{1/4}$; and the number of test prices, κ_n, is smaller by a factor $(\log n)^{-1/4}$. The regret upper bound is improved by the factor $(\log n)^{-1/4}$, and this results from a refinement of their proof technique.

6 Numerical Results

We compare five policies, which we index as follows: (1) our policy (Algorithm AS) (AS); (2) the policy in Besbes and Zeevi (2009, Section 4.1) (BZ); and two variants of the policy in Wang et al. (2014) (Section 7.1): (3) the starting interval in their step 3 is the last (best) price interval from their step 2 (W0); (4) the starting interval in their step 3 (which learns p^c when $p^c > p^u$) is $[\underline{p}, \overline{p}]$ (W1); and (5) the modification of policy BZ, as in Theorem 2 (BZ-M). The reason for considering policy W1 is that policy W0 (which they recommend over W1) does not always demonstrate the expected convergence rate. Whenever a policy prescribes that some price interval is tested (via a number of test prices), we use the midpoints of the relevant sub-intervals, merely because midpoints appear to work slightly better then left- or right-endpoints (for each policy).

The test problems here replicate fully those in Table 1 of Wang et al. (2014) (so that regret numbers are comparable) and additionally allow the presence of an arrival rate. Specifically, we fix the initial inventory $x = 20$; selling horizon $T = 1$; and feasible price set $[\underline{p}, \overline{p}] = [0.1, 10]$; and the demand model is drawn randomly from one of two families of demand functions:

- Linear: the demand function is $\Lambda(p) = \lambda - \alpha p$ with support $\lambda \in [20, 30]$ and $\alpha \in [2, 10]$. Our class $\mathcal{D}$ represents this as (λ, q), where $q(p) = 1 - (\alpha/\lambda)p$.
- Exponential: the demand function is $\Lambda(p) = \lambda \exp(-\beta p)$ with support $\lambda \in [40, 80]$ and $\beta \in [1/3, 1]$. Our class $\mathcal{D}$ represents this as (λ, q), where $q(p) = \exp(-\beta p)$.

The probability law is uniform for each parameter (to match the law used in their study, it is α that is uniformly sampled, not α/λ).

The theory behind each policy prescribes only an optimal asymptotic growth rate; for example, for policy AS, the optimal learning time is $\tau_n \asymp f(n)$ where $f(n) = (\log n/n)^{1/4}$; but n is a theoretical construct that does not exist in the real world; it may be replaced by cn, where $c > 0$ is a *scaling constant* (*scaling*, in short) that is fixed for all n. Put differently, since $f(cn) \sim c^{-1/4} f(n)$ and since a similar equivalence applies to all the relevant f, it follows that the regret convergence rate is unchanged when n is replaced by cn for any $c > 0$ that is fixed for all n. For this reason, in the key literature (Besbes and Zeevi 2009; Wang et al. 2014), sensitivity to c is not of interest, and $c = 1$ is usually chosen. However, we find that the optimal c differ drastically across policies. A fair comparison necessitates using for each policy the c that is nearly-optimal; forcing the same c across policies would tend to favor one against the others.

We examine the mean-regret sensitivity to scaling, for each policy i, when each parameter (duration of a certain phase, or number of test prices) is set as $f(c_i C_n)$, where $c_i > 0$ is a free variable; $C_n = 20n$ is the n-th capacity; and f is the optimal growth rate for each parameter specified in our theorems and in Besbes and Zeevi (2009); Wang et al. (2014). Figure 1 summarizes the sensitivity, showing the mean regret as a function of the ratio c_i/c_i^*, where c_i^* is the value of c_i that roughly minimizes the estimated mean regret for $n = 10^6$, and where: $c_1^* = c_5^* = 100$; $c_2^* = 5$; $c_3^* = c_4^* = 1$ for the linear family; and $c_3^* = c_4^* = 1/4$ for the exponential family; these values are not strict optimizers; they serve as reference points for measuring the sensitivity, and moreover they are not far from optimal in terms of regret, as the figure shows. The main finding is that policies W1 and W0 are the more sensitive: suboptimal scaling in either direction affects them negatively (increases the regret) more than the others.

Consider the ratio of learning-phase durations of policies AS to BZ, which is $(c_1/c_2)^{-1/4}(\log n)^{1/4}$. The value $(c_1^*/c_2^*)^{-1/4} \approx 0.47$ represents the reduction factor resulting from near-optimal scaling ($c_i = c_i^*$) relative to the arbitrary choice $c_1 = c_2$. This finding is consistent with the higher efficiency of the AS estimator (Proposition 4).

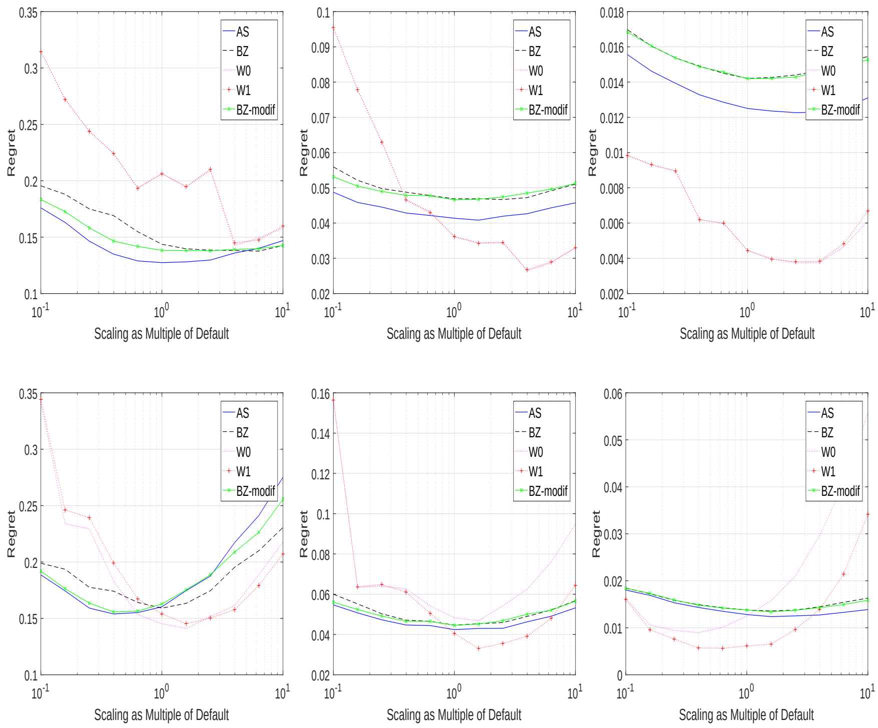

Fig. 1: Sensitivity of the mean regret to the scaling constant c_i for each policy i, for the linear family (top row) and for the exponential family (bottom row), and for varying system size $n \in \{10^2, 10^4, 10^6\}$ (from left to right). For each policy i, the x-axis is the ratio c_i/c_i^* and spans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

The estimated near-optimal performance of each policy (that is, with $c_i = c_i^*$) appears in Table 1; the standard error of estimates is $< 2\%$ for W0, W1 and $n = 10^6$ and $< 1\%$ otherwise.

Based on Figure 1 and Table 1, we conclude:

- The regret converges in line with theory for each policy, except possibly for W0.
- With all policies scaled near-optimally (as in Figure 1), and for sizes n smaller than some threshold, policy AS is superior; and above the threshold, policy W1 is superior. The threshold is between 10^3 and 10^4 for both families. This conclusion remains intact, even when the performance reported in Wang et al. (2014) replaces ours.

Table 1: Comparison of mean regret of the policies with near-optimal scaling.

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

– Policies BZ and BZ-M perform nearly identically, despite the latter's slightly better theoretical convergence.

In experiments with a purchase probability $q()$ that is Lipschitz but not differentiable, the comparison was not notably different; we therefore omit such results.

7 Conclusion

This paper studies a problem of pricing a finite and perishable inventory and a related estimation problem. In the pricing problem, the demand process is defined through an arrival rate and a purchase probability function, both of which are unknown to the seller prior to the beginning of sales.

The results on estimation (Section 4) apply more generally than the inventory-pricing problem, as we now summarize. We start with a Poisson process of unknown rate λ and an unknown thinning-probability function $q(p)$, where p is a control parameter whose domain may be finite or a bounded continuum. The *controlled process* is the Poisson process resulting by thinning the events of the original process with probability $q(p)$ whenever the control is set at p; this process has rate $\lambda q(p)$. Both the original process and the controlled process are observed. Section 4 primarily studies the estimation of the rate function $\lambda q(p)$. Proposition 4 characterizes the (mean-square) asymptotic efficiency of the estimator defined as the product of the empirical rate of the original process times the empirical thinning probabilities at certain points p_i, relative to the estimator defined as the empirical rate of the controlled process at these p_i; it proves and quantifies a gain in efficiency.

The regret of the two policies we studied converges at a rate slightly faster than that in Besbes and Zeevi (2009, Proposition 3). This theoretical improvement is a result of bounding methods that refine theirs, and is unrelated to the better efficiency of the arrivals-and-sales estimator, which converges to a finite positive constant (Proposition 4). The improved efficiency is, however, beneficial; the numerical study via near-optimal scaling constants revealed that our arrivals-and-sales-based policy (AS) spends less time learning relative to policy BZ, and its regret is modestly smaller.

A direction for future research would be to revisit existing pricing methods (Besbes and Zeevi 2009, 2012; Wang et al. 2014; Lei et al. 2014), replacing their sales-only estimator by the arrivals-and-sales one. Numerical results with the demand families in §6 suggest that this replacement, when applied to the algorithm in Wang et al. (2014), improves (reduces) the mean regret (but not the convergence rate). Another direction for future research would be to extend our approach to time-varying demand (Harrison et al. 2012; Besbes and Saure 2014; Keskin and Zeevi 2016), in particular with a time-varying arrival rate.

8 Proofs

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

$$A_n := \mathbb{P}\big(N(\lambda r_n) - \lambda r_n \geq r_n \epsilon_n\big) \leq \begin{cases} n^{-\eta} & \text{if } \epsilon_n < M \\ C_P n^{-\eta} & \text{if } \epsilon_n \geq M \end{cases} \tag{44}$$

Thus $\mathbb{P}(A_n) \leq C_0 n^{-\eta}$ holds for all $n \geq 1$, where $C_0 := C_0(M; \eta, \beta) = \max\{1, C_P\}$. Moreover, for all $n \geq 1$,

$$B_n := \mathbb{P}\big(N(\lambda r_n) - \lambda r_n < -r_n \tilde{\epsilon}_n\big) \leq n^{-\eta}. \tag{45}$$

Remark 4. Our proof of Lemma 1 uses the idea in Besbes and Zeevi (2009, Online Companion, Lemma 2) and corrects a notable error in their calculation of the constant leading the $n^{-\eta}$ term. Since this result is frequently used in the literature (Besbes and Zeevi 2009, 2012; Wang et al. 2014), we think it is important to bring this correction to light. Their constant is incorrect as a result of an incorrectly calculated Taylor expansion. Following the proof, we quantify the resulting error in Remark 5.

Proof of Lemma 1. We claim that for *any nonnegative* sequence $\{\epsilon_n\}$,

$$\mathbb{P}\{N(\lambda r_n) - \lambda r_n \geq r_n \epsilon_n\} \overset{(a)}{\leq} \exp\{-r_n f_*(\lambda + \epsilon_n; \lambda)\} \overset{(b)}{\leq} \exp\{-r_n f_*(M + \epsilon_n; M)\} \tag{46}$$

and

$$\mathbb{P}\{N(\lambda r_n) - \lambda r_n < -r_n \epsilon_n\} \overset{(a)}{\leq} \exp\{-r_n f_*(\lambda - \epsilon_n; \lambda)\} \overset{(b)}{\leq} \exp\{-r_n f_*(M - \epsilon_n; M)\}, \tag{47}$$

where $f_*(x; \lambda) := x \log(x/\lambda) + \lambda - x$, with $x \geq 0$, is the Fenchel-Legendre transform of the cumulant generating function (the logarithm of the moment-generating function) of the Poisson(λ) law (Dembo and Zeitouni 1998, Exercise 2.2.23). In both displays above, step (a) is the special case of more general theory seen in Dembo and Zeitouni (1998, Section 2.2). Specifically, Dembo and Zeitouni (1998, Equations (2.2.12) and (2.2.13)) apply to a sum of r_n independent Poisson(λ) variables, giving (46) and (47), respectively. Step (b) follows from the fact that the derivative of the exponent with respect to λ, which is $-r_n[\log(1+x/\lambda)-x/\lambda]$ with $x=\epsilon_n$ in (46) and with $x=-\epsilon_n$ in (47), is non-negative (an order-two Taylor expansion of $f(x)=\log(1+x)$ around the point zero implies that for any $x>-1$ there exists y with $y\le|x|$ such that $\log(1+x)=x-y^2/2$, so $\log(1+x)-x=-y^2/2\le 0$).

We make an order-two Taylor expansion of $f_*(x;M)$ (as a function of x), noting the first two derivatives of f_* are $f_*'(x;M):=\mathrm{d}f_*(x;M)/\mathrm{d}x=\log(x/M)$ and $f_*''(x;M):=\mathrm{d}^2f_*(x;M)/\mathrm{d}x^2=1/x$; and $f_*(M;M)=f_*'(M;M)=0$. Thus, there exists a $\xi=\xi_n$ in $[0,\epsilon_n]$ such that

$$f_*(M+\epsilon_n;M)=\frac{1}{2}f_*''(M+\xi;M)\epsilon_n^2=\frac{1}{2(M+\xi)}\epsilon_n^2. \tag{48}$$

Proof of (44). The upper bound in (46) will be bounded from above by bounding from below the term $[2(M+\xi)]^{-1}\epsilon_n^2$ in (48); we have two cases.

Case 1: $\epsilon_n<M$. Since $\xi\in[0,\epsilon_n]$, we have $\xi<M$ and $[2(M+\xi)]^{-1}>1/(4M)$; now (46) implies $A_n\le\exp\{-r_n\epsilon_n^2/(4M)\}$; 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 completes the proof of (44) for the case $\epsilon_n<M$.

Case 2: $\epsilon_n\ge M$. We have $[2(M+\xi)]^{-1}\epsilon_n^2\overset{(a)}{\ge}[2(M+\epsilon_n)]^{-1}\epsilon_n^2\overset{(b)}{\ge}M/4$, where step (a) follows from $\xi\le\epsilon_n$ and step (b) follows from the fact that the left side is an increasing function of ϵ_n in the set $[M,\infty)$, and hence its minimum occurs at $\epsilon_n=M$ and equals $M/4$. Thus, (46) gives $A_n\le\exp\{-r_nM/4\}\le\exp\{-n^\beta M/4\}$. For constant c to satisfy $\exp\{-n^\beta M/4\}\le cn^{-\eta}$ for all $n\ge1$, it suffices to set $\log(c)=\sup_{n\ge1}g(n)$, where $g(n)=\eta\log(n)-Mn^\beta/4$. By an elementary calculation, the maximum of g occurs at $n=(4\eta/M\beta)^{1/\beta}$ and equals $(\eta/\beta)\{\log[4\eta/(M\beta)]-1\}$; this gives $c=[4\eta/(M\beta e)]^{\eta/\beta}$. This completes the proof of (44) for the case $\epsilon_n\ge M$. The proof of (44) is now complete.

Proof of (45). If $\tilde\epsilon_n\ge M$, then the result holds trivially (since $\lambda-\tilde\epsilon_n\le M-\tilde\epsilon_n\le0$). Now assume $\tilde\epsilon_n<M$. An expansion of f_* to the left of M shows that there exists a $\xi=\xi_n$ in $[0,\tilde\epsilon_n]$ such that

$$f_*(M-\tilde\epsilon_n;M)=\frac{1}{2}f_*''(M-\xi;M)\tilde\epsilon_n^2=\frac{1}{2(M-\xi)}\tilde\epsilon_n^2. \tag{49}$$

Since $\xi\in[0,\tilde\epsilon_n]$, we have $0<M-\xi\le M$, and thus $[2(M-\xi)]^{-1}\tilde\epsilon_n^2\ge[2M]^{-1}\tilde\epsilon_n^2$. Now (47) implies $B_n\le\exp\{-r_n\tilde\epsilon_n^2/(2M)\}$, and (45) follows. □

Remark 5. Besbes and Zeevi (2009) (Online Companion, Lemma 2) work with the same Taylor expansion as ours, and claim (in their unnumbered display immediately following their (C-4)):

$$-\log\left(\frac{M+\epsilon_n}{M}\right)(M+\epsilon_n)+\epsilon_n=-\frac{1}{2(1+\xi)M}\epsilon_n^2\quad\text{for some }\xi\in[0,\epsilon_n]. \tag{50}$$

We claim that the correct expansion of $f_*(M+\epsilon_n; M)$ is (48) and not (50). The expansion (50) leads them to the result that for all $n \geq 1$, $A_n \leq C_{BZ} n^{-\eta}$, where $C_{BZ} = \max\{1,\ [4\eta M/(\beta e)]^{\eta/\beta}\}$ (after we correct a different minor error in C_{BZ} according to their argument). Comparing C_{BZ} to our constant $C_0 = \max\{1, C_P\}$, where $C_P = [4\eta/(M\beta e)]^{\eta/\beta}$, we see that the difference is an inversion of the factor $M^{\eta/\beta}$. As M moves away from unity, C_{BZ} gives increasingly misleading bounds: for $\eta = 2$ and $\beta = 1/2$, C_{BZ} is incorrectly inflated by a factor M^8 when $M > 1$ is large enough; and it is incorrectly deflated by the same factor when $M < 1$ is small enough. Besbes and Zeevi (2012) (Online Companion, Lemma 1) appears to suffer from the same error in the Taylor expansion.

Lemma 2 *Let $\{I_n\}$ be a sequence of independent Bernoulli(q) random variables with $q \in (0,1)$, and let $S_n := \sum_{i=1}^{n} I_i$ for all $n \geq 1$. Then for any nonnegative sequence $\{\epsilon_n\}$ and for all $n \geq 1$,*

$$A_n := \mathbb{P}\big(S_n - nq \geq n\epsilon_n\big) \leq \exp\big(-2n\epsilon_n^2\big) \tag{51}$$

and

$$B_n := \mathbb{P}\big(S_n - nq \leq -n\epsilon_n\big) \leq \exp\big(-2n\epsilon_n^2\big). \tag{52}$$

Proof of Lemma 2. Following Dembo and Zeitouni (1998, Section 2.2), we claim that for any nonnegative sequence $\{\epsilon_n\}$,

$$\mathbb{P}\big(S_n - nq \geq n\epsilon_n\big) \leq \exp\big(-nf_*(q+\epsilon_n; q)\big) \tag{53}$$

and

$$\mathbb{P}\big(S_n - nq \leq -n\epsilon_n\big) \leq \exp\big(-nf_*(q-\epsilon_n; q)\big) \tag{54}$$

where $f_*(x;q) := x\log(x/q) + (1-x)\log[(1-x)/(1-q)]$ for $x \in [0,1]$, and $f_*(x;q) = \infty$ otherwise, is the Fenchel-Legendre transform of the cumulant generating function of the Bernoulli(q) law (Dembo and Zeitouni 1998, Exercise 2.2.23). Specifically, Dembo and Zeitouni (1998, Equations (2.2.12) and (2.2.13)) for the case of a sum of n independent Bernoulli(q) variables imply (53) and (54), respectively.

We make two order-two Taylor expansions of $f_*(x;q)$ (as a function of x), one to the right of q and another to the left of q, noting the first two derivatives of f_* on $[0,1]$ are $f_*'(x;q) := \mathrm{d}f_*(x;q)/\mathrm{d}x = \log(x/q) - \log[(1-x)/(1-q)]$ and $f_*''(x;q) := \mathrm{d}^2 f_*(x;q)/\mathrm{d}x^2 = 1/[x(1-x)]$; and $f_*(q;q) = f_*'(q;q) = 0$. Thus, there exists $\xi = \xi_n$ with $\xi \in [0, \epsilon_n]$ such that

$$f_*(q \pm \epsilon_n; q) = \frac{1}{2} f_*''(q \pm \xi; q)\epsilon_n^2 = \frac{1}{2(q\pm\xi)[1-(q\pm\xi)]}\epsilon_n^2 \overset{(a)}{\geq} 2\epsilon_n^2, \tag{55}$$

where cases "+" and "−" result from expansion to the right and left, respectively (the ξ need not be the same); and step (a) results from maximizing the denominator: $\sup_{0\leq x\leq 1} x(1-x) = 1/4$. Now (51) follows from (53), and (52) follows from (54). □

8.1 Auxiliary Results and their Proofs

8.1.1 Results Supporting Section 4

The following result is used in Proposition 3.

Lemma 3 *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 of Lemma 3. Limits are meant as $n \to \infty$, unless otherwise indicated. We write $a_n \sim b_n$ to indicate $\lim_{n\to\infty} a_n/b_n = 1$. We claim that

$$Z_n \Rightarrow 1 \quad \text{as } n \to \infty \tag{56}$$

where "⇒" denotes convergence in distribution, and "1" denotes the degenerate distribution whose mass is concentrated at 1. We also claim that the sequence $Z_1, Z_2, \ldots$ is uniformly integrable, i.e.,

$$\lim_{\alpha\to\infty} \sup_n \mathbb{E}[Z_n \mathbb{1}_{[|Z_n|\geq\alpha]}] = 0. \tag{57}$$

The desired conclusion follows from these two results and by "passing the integration to the limit", e.g., Theorem 25.12 of Billingsley (1986).

Result (56) is a consequence of two facts: (a) $X_n/n \Rightarrow 1$; and (b) $Z_n = f(X_n/n)$, where the function $f(x) := x\mathbb{1}_{[x>0]}$ is continuous. Condition (a) above holds because X_n/n equals in distribution the average of n independent mean-one Poisson random variables, which converges to one almost surely and thus converges to one in distribution. Then, the Continuous Mapping Theorem (Billingsley 1986, Theorem 29.2) implies that $Z_n = f(X_n/n) \Rightarrow f(1) = 1$, and the proof of (56) is complete.

To verify the condition (57), we will bound from above the left side for $\alpha > 1$. We have

$$\mathbb{E}[Z_n \mathbb{1}_{[|Z_n|\geq\alpha]}] \stackrel{(a)}{=} \mathbb{E}[nX_n^{-1}\mathbb{1}_{[X_n\leq n/\alpha]}] \stackrel{(b)}{=} n \sum_{k=1}^{\lfloor n/\alpha\rfloor} p(k;n)\frac{1}{k}, \tag{58}$$

where step (a) uses the event equality $\{|Z_n| \geq \alpha\} = \{X_n \leq n/\alpha\}$, and step (b) uses that X_n is Poisson(n) so that $p(k;n) := \mathbb{P}(X_n = k) = e^{-n}\frac{n^k}{k!}$. The number of summands in the right side of (58) is no more than n/α; and each summand is at most

$$\max_{1\leq k\leq n/\alpha} p(k,n) \leq \frac{e^{-n}n^{n/\alpha}}{\lfloor n/\alpha\rfloor!} \sim c_\alpha^n \left(\frac{\alpha}{\sqrt{2\pi n}}\right)^{1/2} =: u_n(\alpha) \tag{59}$$

where $c_\alpha := \left(\alpha^{1/\alpha}e^{-1+1/\alpha}\right)$, and where the "∼" relation results by approximating $\lfloor n/\alpha\rfloor!$ by Stirling's approximation, $n! \sim \sqrt{2\pi n}(n/e)^n$. Thus

$$0 \leq \lim_{\alpha\to\infty} \sup_n \mathbb{E}[Z_n \mathbb{1}_{[Z_n\geq\alpha]}] \leq \lim_{\alpha\to\infty} \sup_n n\frac{n}{\alpha}u_n(\alpha) = 0,$$

where the last step follows from $\lim_{\alpha\to\infty} c_\alpha = e^{-1} < 1$. This proves (57) and completes the proof. □

8.1.2 Results Supporting Section 5.1

In this section, the conditions of Theorem 1 are in force; the sequence $\{u_n\}$ is the one that is motivated in (37), and thus $u_n \asymp (\log n/n)^{1/4}$.

To bound errors related to the revenue rate per arrival, $r(p) := pq(p)$, we first record a Lipschitz property of $r(\cdot)$. For any $p_1,\ p_2$, we have $|p_1 q(p_1) - p_2 q(p_2)| = |p_1(q(p_1) - q(p_2)) + q(p_2)(p_1 - p_2)|$. The latter difference is at most $\overline{p}\overline{M}|p_1 - p_2| + \overline{q}|p_1 - p_2|$, 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 4 *(Revenue Rate at Estimated Unconstrained Price.) Let $\eta \geq 2$. Let α be any small positive number. Let $R_u = R_u(\alpha) = 4\max\{c_1, c_2\}$, where $c_1 := \overline{M}_r(\overline{p} - \underline{p})/2$ and $c_2 := c_2(\alpha) = \overline{p}\eta^{1/2}(2\underline{\lambda})^{-1/2}(1+\alpha)$. Let C_1 and $\underline{n} = \underline{n}(\alpha)$ be as in Proposition 5 (b). For all $n \geq \underline{n}$,*

$$\mathbb{P}\{r(p^u) - r(\widehat{p}^u) \geq R_u u_n\} \leq \frac{C_1}{n^{\eta-1}}. \tag{60}$$

Proof of Lemma 4. For each n, put $\widehat{q}(p_i) = \widehat{q}_{i,n}$ and $\widehat{r}(p_i) := p_i \widehat{q}_{i,n}$ for $i = 1, \ldots, \kappa_n$, and let j be the interval $(p_{j-1}, p_j]$ that contains p^u; here we drop the dependence on n to avoid cluttering the notation. We have

$$\begin{aligned}
&r(p^u) - r(\widehat{p}^u) \\
&= [r(p^u) - r(p_j)] + [r(p_j) - \widehat{r}(p_j)] + [\widehat{r}(p_j) - \widehat{r}(\widehat{p}^u)] + [\widehat{r}(\widehat{p}^u) - r(\widehat{p}^u)] \\
&\leq \overline{M}_r(\overline{p} - \underline{p})\kappa_n^{-1} + 2\max_{1\leq i\leq \kappa_n} |r(p_i) - \widehat{r}(p_i)|,
\end{aligned} \tag{61}$$

where the inequality bounds each of the four terms in square brackets as follows: the first term is at most $\overline{M}_r(\overline{p} - \underline{p})\kappa_n^{-1}$ because $r(\cdot)$ is $\overline{M}_r$-Lipschitz and $|p^u - p_j| \leq (\overline{p} - \underline{p})\kappa_n^{-1}$ by definition of j; the third term is non-positive by the definition of $\widehat{p}^u = \arg\max_{1\leq j\leq \kappa_n} p_i \widehat{q}(p_i)$; each of the second and fourth term is at most $\max_{1\leq i\leq \kappa_n} |r(p_i) - \widehat{r}(p_i)|$. The proof of (61) is complete. By the definition of R_u, we have for all $n \geq 1$, $R_u u_n/4 \geq c_1 u_n \geq c_1 \kappa_n^{-1}$ and $R_u u_n/4 \geq c_2 u_n \geq c_2 (\log n/r_n)^{1/2}$, hence

$$\frac{R_u u_n}{2} - c_1 \kappa_n^{-1} \geq c_2 \left(\frac{\log n}{r_n}\right)^{1/2}. \tag{62}$$

Now, for all $n \geq \underline{n}$,

$$\begin{aligned}
&\mathbb{P}\{r(p^u) - r(\widehat{p}^u) > R_u u_n\} \\
&\overset{(a)}{\leq} \mathbb{P}\left\{\max_{1\leq i\leq \kappa_n} |r(p_i) - \widehat{r}(p_i)| > \frac{R_u u_n}{2} - c_1\kappa_n^{-1}\right\} \\
&\overset{(b)}{\leq} \mathbb{P}\left\{\overline{p}\max_{1\leq i\leq \kappa_n} |q(p_i) - \widehat{q}(p_i)| > \frac{R_u u_n}{2} - c_1\kappa_n^{-1}\right\} \\
&\overset{(c)}{\leq} \sum_{i=1}^{\kappa_n} \mathbb{P}\left\{|q(p_i) - \widehat{q}(p_i)| > \frac{1}{\overline{p}}\left(\frac{R_u u_n}{2} - c_1\kappa_n^{-1}\right)\right\} \\
&\overset{(d)}{\leq} \sum_{i=1}^{\kappa_n} \mathbb{P}\left\{|q(p_i) - \widehat{q}(p_i)| > \eta^{1/2}(2\underline{\lambda})^{-1/2}(1+\alpha)\left(\frac{\log n}{r_n}\right)^{1/2}\right\} \\
&\overset{(e)}{\leq} \kappa_n \frac{C_1}{n^\eta} \\
&\leq \frac{C_1}{n^{\eta-1}} \qquad (63)
\end{aligned}$$

where (a) follows from (61); in step (b) we observed that $\max_{1\leq i\leq\kappa_n} |r(p_i) - \widehat{r}(p_i)| = \max_{1\leq i\leq\kappa_n} p_i|q(p_i) - \widehat{q}(p_i)| \leq \overline{p}\max_{1\leq i\leq\kappa_n} |q(p_i) - \widehat{q}(p_i)|$; step (c) uses a union bound; step (d) uses $\overline{p}^{-1}(R_u u_n/2 - c_1\kappa_n^{-1}) \geq \overline{p}^{-1}c_2(\log n/r_n)^{1/2} = \eta^{1/2}(2\underline{\lambda})^{-1/2}(1+\alpha)(\log n/r_n)^{1/2}$, where the inequality is shown in (62); and step (e) follows from (28); it is for this step that we required $n \geq \underline{n}$. □

Next, we will bound errors related to p^c, the optimal constrained price. To this end, we need Lipschitz constants for the demand function $\Lambda(p) := \lambda q(p)$ and its "inverse" defined below. By Assumption 2,

$$|\Lambda(p_1;\lambda) - \Lambda(p_2;\lambda)| = \lambda|q(p_1) - q(p_2)| \leq \overline{\lambda M}|p_1 - p_2| \quad \text{for any } p_1,\ p_2, \lambda. \quad (64)$$

Viewing $\Lambda(p;\lambda)$ as a function of p, where $\lambda \in \mathcal{D}$, we call it $\overline{M}_\Lambda$-Lipschitz, where $\overline{M}_\Lambda := \overline{\lambda M}$ is the supremum across $\lambda \in \mathcal{D}$ of the Lipschitz constants of the functions $\{\lambda q(p) : \lambda \in \mathcal{D}\}$; this simply means that (64) holds. Moreover, by Assumption 2, we have $\inf_{\lambda,p_1,p_2\in\mathcal{D}} |\Lambda(p_1) - \Lambda(p_2)| \geq \underline{\lambda M}|p_1 - p_2|$; putting $\underline{M}_\Lambda := \underline{\lambda M}$, the latter is the same as

$$|p_1 - p_2| \leq \frac{|\Lambda(p_1) - \Lambda(p_2)|}{\underline{M}_\Lambda} \quad \text{for any } p_1,\ p_2, \lambda. \quad (65)$$

Noting that an arrival rate λ and a demand rate z determine the price uniquely as $q^{-1}(z/\lambda)$, we can view the set of mappings $\{z \mapsto q^{-1}(z/\lambda) : \lambda \in \mathcal{D}\}$, as being $\underline{M}_\Lambda^{-1}$-Lipschitz; this simply means that (65) holds.

Lemma 5 *(Estimated Constrained Price and Associated Revenue Rate per Arrival.) Let $\eta \geq 2$. Let α be any small positive number. Let $K_c = K_c(\alpha) = 4\underline{M}_\Lambda^{-1}\max\{c_1, c_2(\alpha)\}$, where $c_1 := \overline{M}_\Lambda(\overline{p} - \underline{p})/2$ and $c_2 := c_2(\alpha) = (\eta\overline{\lambda}/2)^{1/2}(1+$*

α). Let $\overline{M}_r := \overline{q} + \overline{M}\overline{p}$. Let C_1 and $\underline{n} = \underline{n}(\alpha)$ be as in Proposition 5 (b). For all $n \geq \underline{n}$,

$$\mathbb{P}\{|\widehat{p}^c - p^c| > K_c u_n\} \leq \frac{C_1}{n^{\eta-1}}, \tag{66}$$

$$\mathbb{P}\{|r(\widehat{p}^c) - r(p^c)| > \overline{M}_r K_c u_n\} \leq \frac{C_1}{n^{\eta-1}}. \tag{67}$$

Proof of Lemma 5. For each n, put $\widehat{\Lambda}(p_i) := \widehat{\lambda}_n \widehat{q}_{i,n}$ for $i = 1, \ldots, \kappa_n$, and let j be the interval $(p_{j-1}, p_j]$ that contains p^c; here we drop the dependence on n to avoid cluttering the notation. We have

$$|\Lambda(\widehat{p}^c) - \Lambda(p^c)| \leq |\Lambda(\widehat{p}^c) - \widehat{\Lambda}(\widehat{p}^c)| + |\widehat{\Lambda}(\widehat{p}^c) - \Lambda(p^c)|, \tag{68}$$

and observe that the first term on the right is at most $\max_{1\leq i\leq\kappa_n} |\widehat{\Lambda}(p_i) - \Lambda(p_i)|$. For the second term we have

$$\begin{aligned} |\widehat{\Lambda}(\widehat{p}^c) - \Lambda(p^c)| &\overset{(a)}{\leq} |\widehat{\Lambda}(p_j) - \Lambda(p^c)| \\ &\overset{(b)}{\leq} |\widehat{\Lambda}(p_j) - \Lambda(p_j)| + |\Lambda(p_j) - \Lambda(p^c)| \\ &\overset{(c)}{\leq} \max_{1\leq i\leq\kappa_n} |\widehat{\Lambda}(p_i) - \Lambda(p_i)| + \overline{M}_\Lambda(\overline{p} - \underline{p})\kappa_n^{-1}, \end{aligned} \tag{69}$$

where (a) follows from the definition of $\widehat{p}^c = \arg\min_{1\leq i\leq\kappa_n} |\widehat{\Lambda}(p_i) - x/T| = \arg\min_{1\leq i\leq\kappa_n} |\widehat{\Lambda}(p_i) - \Lambda(p^c)|$; (b) follows from the triangle inequality; and (c) holds because $\Lambda(\cdot)$ is $\overline{M}_\Lambda$-Lipschitz (seen in (64)) and $|p_j - p^c| \leq (\overline{p} - \underline{p})\kappa_n^{-1}$. Combining (68) and (69) we have

$$|\Lambda(\widehat{p}^c) - \Lambda(p^c)| \leq 2 \max_{1\leq i\leq\kappa_n} |\widehat{\Lambda}(p_i) - \Lambda(p_i)| + \frac{\overline{M}_\Lambda(\overline{p} - \underline{p})}{\kappa_n}. \tag{70}$$

For all $n \geq n_c$,

$$\begin{aligned} \mathbb{P}\{|\widehat{p}^c - p^c| > K_c u_n\} &\overset{(a)}{\leq} \mathbb{P}\{|\Lambda(\widehat{p}^c) - \Lambda(p^c)| > \underline{M}_\Lambda K_c u_n\} \\ &\overset{(b)}{\leq} \mathbb{P}\left\{\max_{1\leq i\leq\kappa_n} |\widehat{\Lambda}(p_i) - \Lambda(p_i)| > \frac{\underline{M}_\Lambda K_c u_n}{2} - c_1\kappa_n^{-1}\right\} \\ &\overset{(c)}{\leq} \mathbb{P}\left\{\max_{1\leq i\leq\kappa_n} |\widehat{\Lambda}(p_i) - \Lambda(p_i)| > (\eta\overline{\lambda}/2)^{1/2}(1+\alpha)\left(\frac{\log n}{r_n}\right)^{1/2}\right\} \\ &\overset{(d)}{\leq} \sum_{i=1}^{\kappa_n} \mathbb{P}\left\{|\widehat{\Lambda}(p_i) - \Lambda(p_i)| > (\eta\overline{\lambda}/2)^{1/2}(1+\alpha)\left(\frac{\log n}{r_n}\right)^{1/2}\right\} \\ &\overset{(e)}{\leq} \kappa_n \frac{C_1}{n^\eta} \\ &\leq \frac{C_1}{n^{\eta-1}}, \end{aligned} \tag{71}$$

where (a) follows from (65) (Lipschitz property of the class of mappings from demand rate to price); (b) follows from (70); step (c) is valid because $K_c u_n/2 - c_1\kappa_n^{-1} \geq (\eta\overline{\lambda}/2)^{1/2}(1+\alpha)(\log n/r_n)^{1/2}$ (the construction of K_c ensures the analog of (62) holds, with K_c replacing R_u, and c_1, c_2 defined here); step (d) uses a union bound; and step (e) follows from (27) in Proposition 5; it is for this step that we required that $n \geq \underline{n}$. This completes the proof of (66). Since $r(\cdot)$ is $\overline{M}_r$-Lipschitz, (67) follows from (66). □

Lemma 6 *(Revenue Rate at the Estimated Price.) Let α be any small positive number. Put $R = \max\{2R_u, \overline{M}_r^{-1}K_c, 2\overline{M}_r K_c\}$, where $K_c = K_c(\alpha)$ is defined in Lemma 5. Let C_1 and $\underline{n} = \underline{n}(\alpha)$ be as in Proposition 5 (b). For all $n \geq \underline{n}$,*

$$\mathbb{E}[r(\widehat{p})] \geq r(p^D) - Ru_n - r(p^D)\frac{2C_1}{n^{\eta-1}}. \tag{72}$$

Proof of Lemma 6. The proof is given in abbreviated form, and parallels Besbes and Zeevi (2009, Electronic Companion, Lemma 4, Step 3).

Case 1. Here we assume that $p^u \geq p^c$, hence $p^D = p^u$. We argue as Besbes and Zeevi (2009, Electronic Companion, Lemma 4, Step 3, Case 1). The only difference in the argument is that u_n in (24) and revenue rate per arrival $r(\cdot)$ replace their u_n and revenue rate per time $r(\cdot)$, respectively. By their argument and the conditions $R \geq R_u$ and $R \geq \overline{M}_r^{-1}K_c$, we obtain

$$\mathbb{P}\{r(p^D) - r(\widehat{p}) > Ru_n\} \leq \frac{C_1}{n^{\eta-1}}.$$

Case 2. Here we assume that $p^u > p^c$, hence $p^D = p^c$. We obtain

$$\begin{aligned}
&\mathbb{P}\{r(p^D) - r(\widehat{p}) > Ru_n\} \\
&\overset{(a)}{\leq} \mathbb{P}\{r(p^u) - r(\widehat{p}^u) > R_u u_n/2\} + \mathbb{P}\{r(p^c) - r(\widehat{p}^c) > \overline{M}_r K_c u_n/2\} \\
&\overset{(b)}{\leq} \frac{2C_1}{n^{\eta-1}}
\end{aligned}$$

where step (a) is argued exactly as Besbes and Zeevi (2009, Electronic Companion, Lemma 4, Step 3, Case 2), using the fact that $R/2 \geq \max\{R_u, \overline{M}_r K_c\}$; step (b) follows from (60) and (67).

Let $\mathcal{G} := \{\omega : r(p^D) - r(\widehat{p}) \leq Ru_n\}$. Putting together the results for cases 1 and 2, we have shown that

$$\mathbb{P}(\mathcal{G}^c) \leq \frac{2C_1}{n^{\eta-1}}. \tag{73}$$

Thus

$$\begin{aligned}
\mathbb{E}[r(\widehat{p})] &= r(p^D) - \mathbb{E}[r(p^D) - r(\widehat{p})|\mathcal{G}]\mathbb{P}(\mathcal{G}) - \mathbb{E}[r(p^D) - r(\widehat{p})|\mathcal{G}^c]\mathbb{P}(\mathcal{G}^c) \\
&\overset{(a)}{\geq} r(p^D) - Ru_n - r(p^D)\frac{2C_1}{n^{\eta-1}}
\end{aligned} \tag{74}$$

where step (a) bounds the second conditional term via (73) and the fact that $r(p^D) - r(\widehat{p}) \leq r(p^D) < \infty$. □

Lemma 7 *(Bound on the Sales During Learning.) Let $\eta \geq 2$. Let $M := \overline{\lambda}\overline{q}$ and $K_L = M + 2\eta^{1/2}M^{1/2}$. Take $C_0 = C_0(M)$ as in Lemma 1. For all $n \geq 1$,*

$$\mathbb{P}(Y_n^{(L)} > K_L n u_n) \leq \frac{C_0}{n^{\eta-1}}. \tag{75}$$

Proof of Lemma 7. Put $r_n := n\tau_n/\kappa_n$. We have $\mathbb{P}(Y_n^{(L)} > K_L n u_n) = \mathbb{P}\big(\sum_{i=1}^{\kappa_n} N(\lambda q_i r_n) > K_L n u_n\big) \leq \sum_{i=1}^{\kappa_n} \mathbb{P}\big(N(\lambda q_i r_n) > K_L n u_n/\kappa_n\big)$, which implies that

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

In the above, step (a) uses that $\lambda q_i \leq M$ for all i; in step (b) we use that

$$-M r_n + K_L \frac{n u_n}{\kappa_n} = -M\frac{n\tau_n}{\kappa_n} + K_L \frac{n u_n}{\kappa_n} \geq \frac{n}{\kappa_n}(K_L - M)u_n \geq 2\eta^{1/2}M^{1/2}\left(\frac{\log n}{r_n}\right)^{1/2},$$

where the first inequality follows from $u_n \geq \tau_n$ and the second inequality follows from $u_n \geq (\log n/r_n)^{1/2}$, $n/\kappa_n \geq 1$, and the definition of K_L; step (c) follows directly from Lemma 1; and step (d) uses that $\kappa_n = o(n)$. □

Lemma 8 *(Bound on the Expected Overshoot Above Capacity when FIR does not Force a Run-Out.) Put $\Lambda(\overline{p}) = \lambda q(\overline{p})$ and suppose that $\Lambda(\overline{p}) \leq x/T$. Let $u_n \asymp (\log n/n)^{1/4}$ and $u_n \geq c_u(\log n/n)^{1/4}$ for all n and for some $c_u > 0$. Take $\underline{n} = \underline{n}(\alpha)$ as in Proposition 5. For some positive constant K_E and for all $n \geq \underline{n}$,*

$$\mathbb{E}[(Y_n - nx)^+] \leq K_E n u_n. \tag{76}$$

Proof of Lemma 8. Let $K_\Lambda = \overline{M}_\Lambda K_c$, where K_c is defined in Lemma 5. Put $K_m = x + K_\Lambda T \sup_{n\geq 1} u_n$, which is finite since $u_n \to 0$. Let $K_Y = 2\max\{K_L, K_\Lambda T + 2c_u^{-1}\eta^{1/2}K_m^{1/2}e^{-1/4}\}$, where K_L is defined in Lemma 7.

Since $Y_n = Y_n^{(L)} + Y_n^{(P)}$, we have, for all $n \geq 1$,

$$\begin{aligned}
\mathbb{P}(Y_n - nx > K_Y n u_n) &\leq \mathbb{P}(Y_n^{(L)} > K_Y n u_n/2) + \mathbb{P}(Y_n^{(P)} > nx + K_Y n u_n/2) \\
&\overset{(a)}{\leq} \frac{C_0}{n^{\eta-1}} + \mathbb{P}(Y_n^{(P)} > nx + K_Y n u_n/2),
\end{aligned} \tag{77}$$

where step (a) follows from Lemma 7 and the fact that $K_Y/2 \geq K_L$, and where C_0 is defined in Lemma 1 (to which Lemma 7 refers).

To bound the last term in (77), we first note that $p^c - \widehat{p} = p^c - \max(\widehat{p}^u, \widehat{p}^c) \leq p^c - \widehat{p}^c$. By Lemma 5, for the constant C_1 defined there, we have for all $n \geq \underline{n}$, $\mathbb{P}(p^c - \widehat{p} > K_c u_n) \leq \mathbb{P}(p^c - \widehat{p}^c > K_c u_n) \leq C_1 n^{-\eta+1}$. From this and the fact that $\Lambda()$ is monotone decreasing and $\overline{M}_\Lambda$-Lipschitz, we obtain, for all $n \geq \underline{n}$,

$$\mathbb{P}\big\{\Lambda(\widehat{p}) > \Lambda(p^c) + K_\Lambda u_n\big\} \leq \mathbb{P}\{p^c - \widehat{p} > K_c u_n\} \leq \frac{C_1}{n^{\eta-1}}. \tag{78}$$

The above says that with probability at least $1 - C_1 n^{-\eta+1}$, the mean of $Y_n^{(P)}$, that is, $\Lambda(\widehat{p})n(T - \tau_n)$, is no more than nv_n, where

$$v_n := (\Lambda(p^c) + K_\Lambda u_n)(T - \tau_n) \leq x + K_\Lambda T u_n, \tag{79}$$

where the inequality uses that $\Lambda(p^c) \leq x/T$, which follows from the assumption $\Lambda(\overline{p}) \leq x/T$ (in more detail, since $\Lambda()$ is continuous and decreasing, either $\Lambda(\underline{p}) \geq x/T$, in which case $\Lambda(\overline{p}) = x/T$, or $\Lambda(\underline{p}) < x/T$, in which case $\Lambda(\overline{p}) < x/T$). Now, for all $n \geq \underline{n}$,

$$\begin{aligned}
&\mathbb{P}\{Y_n^{(P)} > nx + K_Y n u_n/2\} \\
&\leq \mathbb{P}\big\{N\big(\Lambda(\widehat{p})n(T-\tau_n)\big) > nx + K_Y n u_n/2,\ \Lambda(\widehat{p}) \leq \Lambda(p^c) + K_\Lambda u_n\big\} \\
&\quad + \mathbb{P}\big\{\Lambda(\widehat{p}) > \Lambda(p^c) + K_\Lambda u_n\big\} \\
&\overset{(a)}{\leq} \mathbb{P}\big\{N\big((\Lambda(p^c) + K_\Lambda u_n)n(T-\tau_n)\big) > nx + K_Y n u_n/2\big\} + \frac{C_1}{n^{\eta-1}} \\
&= \mathbb{P}\big\{N\big(nv_n\big) - nv_n > nx + K_Y n u_n/2 - nv_n\big\} + \frac{C_1}{n^{\eta-1}} \\
&\overset{(b)}{\leq} \frac{C_0}{n^{\eta-1}} + \frac{C_1}{n^{\eta-1}},
\end{aligned} \tag{80}$$

where step (a) uses (78); and step (b) follows by applying Lemma 1 with $r_n = n$ and $M = K_m$ (since $\sup_n v_n \leq x + K_\Lambda T \sup_n u_n = K_m$); this application is valid because the deviation threshold satisfies

$$\begin{aligned}
nx + K_Y n u_n/2 - nv_n &\overset{(b1)}{\geq} nx + K_Y n u_n/2 - n(x + K_\Lambda T u_n) \\
&\overset{(b2)}{\geq} (K_Y/2 - K_\Lambda T) n c_u \left(\frac{\log n}{n}\right)^{1/4} \overset{(b3)}{\geq} 2\eta^{1/2} K_m^{1/2} (n \log n)^{1/2},
\end{aligned}$$

where (b1) follows from (79); (b2) follows from $u_n \geq c_u(\log n/n)^{1/4}$; and (b3) uses that $K_Y/2 - K_\Lambda T \geq 2c_u^{-1}\eta^{1/2}K_m^{1/2}\sup_{n\geq 1}(\log n/n)^{1/4}$ and that $\sup_{n\geq 1}(\log n/n) = e^{-1}$. This completes the proof of (80). Combining (77) and (80), we have, for all $n \geq \underline{n}$,

$$\mathbb{P}(Y_n - nx > K_Y n u_n) \leq \frac{2C_0 + C_1}{n^{\eta-1}}. \tag{81}$$

Now, for all $n \geq \underline{n}$,

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

where K_E is a suitable finite constant. In the above, step (a) uses the fact 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). □

Lemma 9 *(Simultaneous Bounds on the Sale Quantity and Price when FIR Forces a Run-Out). Suppose that $\Lambda(\overline{p}) > x/T$. Let $u_n \asymp (\log n/n)^{1/4}$ and $u_n \geq c_u(\log n/n)^{1/4}$ for all n and for some $c_u > 0$. Let $\tilde{K}_Y = \max\{K_L, \Lambda(\overline{p}) + 2c_u^{-1}\eta^{1/2}[\Lambda(\overline{p})T]^{1/2}e^{-1/4}\}$, where K_L is defined in Lemma 7. Define $\mathcal{A} = \{\omega : \min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\} \geq nx - \tilde{K}_Y n u_n, |\widehat{p} - p^D| \leq K_c n u_n\}$, where K_c is defined in Lemma 5. Take $\underline{n} = \underline{n}(\alpha)$ as in Proposition 5. Then for all $n \geq \underline{n}$,*

$$\mathbb{P}(\mathcal{A}) \geq 1 - \frac{2C_0 + C_1}{n^{\eta-1}}, \tag{82}$$

where C_0 is defined in Lemma 1 and C_1 is defined in Proposition 5 (b).

Proof of Lemma 9. Note that $\mathcal{A} = \{Y_n^{(P)} \geq nx - \tilde{K}_Y n u_n,\ Y_n^{(L)} \leq \tilde{K}_Y n u_n,\ |\widehat{p} - p^D| \leq K_c n u_n\}$. For all $n \geq \underline{n}$,

$$\begin{aligned}\mathbb{P}(\mathcal{A}^c) &\overset{(a)}{\leq} \mathbb{P}(Y_n^{(P)} < nx - \tilde{K}_Y n u_n) + \mathbb{P}(Y_n^{(L)} > \tilde{K}_Y n u_n) + \mathbb{P}(|\widehat{p} - p^D| \leq K_c n u_n) \\ &\overset{(b)}{\leq} \mathbb{P}(Y_n^{(P)} < nx - \tilde{K}_Y n u_n) + \frac{C_0}{n^{\eta-1}} + \mathbb{P}(|\widehat{p}^c - p^c| \leq K_c n u_n) \\ &\overset{(c)}{\leq} \mathbb{P}(Y_n^{(P)} < nx - \tilde{K}_Y n u_n) + \frac{C_0}{n^{\eta-1}} + \frac{C_1}{n^{\eta-1}},\end{aligned} \tag{83}$$

where (a) follows from a union bound; in step (b), the second term was bounded by using Lemma 7 and the third term was bounded by using that $p^D = p^c = \overline{p}$ and therefore $|\widehat{p} - p^D| = |\max(\widehat{p}^c, \widehat{p}^u) - p^c| \leq |\widehat{p}^c - p^c|$; and in step (c) we used Lemma 5 to bound the third term. For the remaining term we have, for all $n \geq 1$,

$$\begin{aligned}&\mathbb{P}\{Y_n^{(P)} < nx - \tilde{K}_Y n u_n\} \\ &= \mathbb{P}\big\{N\big(\Lambda(\widehat{p})n(T - \tau_n)\big) < nx - \tilde{K}_Y n u_n\big\} \\ &\overset{(a)}{\leq} \mathbb{P}\big\{N\big((\Lambda(\overline{p})n(T - \tau_n)\big) < n\Lambda(\overline{p})T - \tilde{K}_Y n u_n\big\} \\ &= \mathbb{P}\big\{N\big((\Lambda(\overline{p})n(T - \tau_n)\big) - n\Lambda(\overline{p})(T - \tau_n) < -n(\tilde{K}_Y u_n - \Lambda(\overline{p})\tau_n)\big\} \\ &\overset{(b)}{\leq} \frac{C_0}{n^{\eta-1}},\end{aligned} \tag{84}$$

where step (a) uses the fact that $\Lambda(\widehat{p}) \geq \Lambda(\overline{p})$ and $x \leq \Lambda(\overline{p})T$; and step (b) follows by applying Lemma 1 with $r_n = n$ and $M = \Lambda(\overline{p})T$; this application is valid because the deviation threshold satisfies

$$n(\tilde{K}_Y u_n - \Lambda(\overline{p})\tau_n) \overset{(b1)}{\geq} (\tilde{K}_Y - \Lambda(\overline{p}))nu_n \overset{(b2)}{\geq} (\tilde{K}_Y - \Lambda(\overline{p}))nc_u \left(\frac{\log n}{n}\right)^{1/4} \overset{(b3)}{\geq} 2\eta^{1/2}[\Lambda(\overline{p})T]^{1/2}(n\log n)^{1/2},$$

where (b1) follows from $u_n \geq \tau_n$; (b2) follows from $u_n \geq c_u(\log n/n)^{1/4}$; and (b3) uses $\tilde{K}_Y - \Lambda(\overline{p}) \geq 2c_u^{-1}\eta^{1/2}[\Lambda(\overline{p})T]^{1/2}\sup_{n\geq 1}(\log n/n)^{1/4}$ and that $\sup_{n\geq 1}(\log n/n) = e^{-1}$. This completes the proof of (84); inserting the latter into (83) completes the proof. □

References

Araman, Victor F., René Caldentey. 2009. Dynamic pricing for nonperishable products with demand learning. *Oper. Res.* **57** 1169–1188. doi:10.1287/opre.1090.0725. URL `http://dx.doi.org/10.1287/opre.1090.0725`.

Auer, P., N. Cesa-Bianchi, P. Fischer. 2002. Finite-time analysis of the multiarmed bandit problem. *Machine Learning* **47** 235–256.

Aviv, Y., A. Pazgal. 2005. Pricing of short life-cycle products through active learning. Available at `http://www.olin.wustl.edu/faculty/aviv/papers/Pricing_MS_Dec_2005.pdf`.

Ball, M. O., M. Queyranne. 2009. Toward robust revenue management: Competitive analysis of online booking. *Operations Research* **57** 950–963.

Besbes, O., D. Saure. 2014. Dynamic pricing strategies in the presence of demand shifts. *Manufacturing and Service Operations Management* **16** 513–528.

Besbes, Omar, Assaf Zeevi. 2009. Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Operations Research* **57** 1407–1420.

Besbes, Omar, Assaf Zeevi. 2012. Blind Network Revenue Management. *Operations Research* **60** 1537–1550.

Besbes, Omar, Assaf Zeevi. 2015. On the (Surprising) Sufficiency of Linear Models for Dynamic Pricing with Demand Learning. *Management Science* **61** 723–739.

Bickel, P. J., K. A. Doksum. 1977. *Mathematical Statistics: Basic Ideas and Selected Topics*. Holden-Day Company, Oakland, California.

Billingsley, P. 1986. *Probability and Measure*. 2nd ed. Wiley, New York.

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

Broder, J., P. Rusmevichientong. 2012. Dynamic pricing under a general parametric choice model. *Operations Research* **60** 965–980.

Dembo, A., O. Zeitouni. 1998. *Large Deviation Techniques and Applications*. Springer.

den Boer, A. V. 2015. Dynamic pricing and learning: historical origins, current research, and new directions. *Surveys in Operations Research and Management Science* **20** 1–18.

den Boer, A. V., N. B. Keskin. 2017. Discontinuous demand functions: estimation and pricing. `https://ssrn.com/abstract=3003984`.

den Boer, A. V., B. Zwart. 2014. Simultaneously learning and optimizing using controlled variance pricing. *Management Science* **60** 770–783.

den Boer, A. V., B. Zwart. 2015. Dynamic pricing and learning with finite inventories. *Operations Research* **63** 965–978.

Elmaghraby, W., P. Keskinocak. 2003. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management Science* **49** 1287–1309.

Farias, Vivek F., Benjamin Van Roy. 2010. Dynamic pricing with a prior on market response. *Oper. Res.* **58** 16–29. doi:10.1287/opre.1090.0729. URL `http://dx.doi.org/10.1287/opre.1090.0729`.

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

Harrison, J. M., N. B. Keskin, A. Zeevi. 2012. Bayesian dynamic pricing policies: learning and earning under a binary prior distribution. *Management Science* **58** 570–586.

Keskin, N. B., A. Zeevi. 2014. Dynamic pricing with an unknown demand model: Asymptotically optimal semi-myopic policies. *Operations Research* **62** 1142–1167.

Keskin, N. B., A. Zeevi. 2016. Chasing demand: Learning and earning in a changing environment. *Mathematics of Operations Research* **42** 277–307.

Kleinberg, Robert, Tom Leighton. 2003. The value of knowing a demand curve: Bounds on regret for online posted-price auctions. *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*. FOCS '03, IEEE Computer Society, Washington, DC, USA, 594–605. URL `http://dl.acm.org/citation.cfm?id=946243.946352`.

Lei, Y., S. Jasin, A. Sinha. 2014. Near-optimal bisection search for nonparametric pricing with inventory constraint. `https://deepblue.lib.umich.edu/handle/2027.42/108717`.

Lim, A., J. Shanthikumar. 2007. Relative entropy, exponential utility, and robust dynamic pricing. *Operations Research* **55** 198–214.

Lin, K. Y. 2006. Dynamic pricing with real-time demand learning. *European Journal of Operational Research* **174** 522–538.

Talluri, K., G. van Ryzin. 2005. *Theory and Practice of Revenue Management*. Springer.

Wang, Z., S. Deng, Y. Ye. 2014. Close the gaps: a learning-while-doing algorithm for a class of single-product revenue management problems. *Operations Research* **62** 318–331.