

The Blessing of Strategic Customers in Personalized Pricing

Zhi Chen

Department of Decisions, Operations and Technology, The Chinese University of Hong Kong, zhi.chen@cuhk.edu.hk

Bradley Sturt

Department of Information and Decision Sciences, University of Illinois Chicago, bsturt@uic.edu

Weijun Xie

H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, wxie@gatech.edu

We consider a feature-based personalized pricing problem in which the buyer is *strategic*: given the seller’s pricing policy, the buyer can augment the features that they reveal to the seller to obtain a low price for the product. We model the seller’s pricing problem as a stochastic program over an infinite-dimensional space of pricing policies where the radii by which the buyer can perturb the features are strictly positive. We establish that the sample average approximation of this problem is *asymptotically consistent*; that is, we prove that the objective value of the sample average approximation converges almost surely to the objective value of the stochastic problem as the number of samples tends to infinity under mild technical assumptions. This consistency guarantee thus shows that incorporating strategic consumer behavior into a data-driven pricing problem can, in addition to making the pricing problem more realistic, also help prevent overfitting.

Key words: sample average approximation, stochastic programming, data-driven pricing.

History: August 16, 2024

1. Introduction

Personalized pricing is a problem in revenue management in which a seller is tasked with selecting a price for a single product to offer to a buyer. The buyer’s valuation of the product is not directly observed by the seller, but the seller observes features of the buyer (e.g., the buyer’s age, credit score, purchase timing) that can be used to predict the buyer’s valuation. The goal of the seller is to find a pricing policy that specifies what price to offer to the buyer as a function of the buyer’s features to maximize the seller’s expected revenue. We model the personalized pricing problem as a stochastic program of the form

$$\underset{\pi: \mathcal{X} \rightarrow \mathcal{P}}{\text{maximize}} \quad \mathbb{E}[\pi(\mathbf{X}) \cdot \mathbb{I}\{\pi(\mathbf{X}) \leq V\}], \quad (1)$$

where $\pi(\cdot)$ is the pricing policy, $\mathcal{X} \subseteq \mathbb{R}^D$ is the feasible set for the buyer’s D features, $\mathcal{P} \subseteq \mathbb{R}$ is the set of feasible prices that can be offered by the seller, and the buyer is represented by the random vector (\mathbf{X}, V) consisting of the features $\mathbf{X} \in \mathcal{X}$ and product valuation $V \in \mathcal{V} \subseteq \mathbb{R}$. For an introduction to the class of stochastic programs (1), see Elmachtoub et al. (2021) and references therein.

A classical technique in the field of operations research for estimating the objective value of a stochastic program is the *sample average approximation* (Shapiro et al. 2009). This technique consists of two steps: first, we use historical data or simulation to obtain independent samples of the random variables; second, we approximate the stochastic program by replacing the true probability distribution with the empirical probability distribution constructed from the samples. The sample average approximation in application to the personalized pricing problem (1) is denoted by

$$\underset{\pi: \mathcal{X} \rightarrow \mathcal{P}}{\text{maximize}} \quad \frac{1}{N} \sum_{i=1}^N \pi(\mathbf{X}^i) \cdot \mathbb{I} \{ \pi(\mathbf{X}^i) \leq V^i \}, \quad (2)$$

where $(\mathbf{X}^1, V^1), \dots, (\mathbf{X}^N, V^N)$ are independent samples drawn from the same distribution as (\mathbf{X}, V) . The objective value of (2) is the sample average approximation's estimate of the objective value of (1).

Unfortunately, if the buyer's features are drawn from a continuous probability distribution, then the objective value of (2) can yield a highly inaccurate estimate of the objective value of (1), even in the asymptotic regime in which the number of samples tends to infinity (see, e.g., Xie et al. 2024). We illustrate this by the following example:

EXAMPLE 1. Suppose that $D = 1$, $\mathcal{P} = \mathcal{X} = \mathcal{V} = \mathbb{R}$, and that $X, V \in \mathbb{R}$ are independent random variables that are each uniformly distributed over $[0, 1]$. The fact that X and V are independent implies that there exists an optimal pricing policy for (1) that offers a constant price for all features; hence, (1) = $\max_p \mathbb{E}[p \mathbb{I} \{p \leq V\}] = \max_p \int_p^1 p dv = \max_p \{p - p^2\} = 1/4$. However, it follows from the fact that $X^i \neq X^j$ for all $1 \leq i < j \leq N$, almost surely, that there exists an optimal pricing policy for (2) that satisfies $\pi(X^i) = V^i$ for all $1 \leq i \leq N$; hence, (2) = $\frac{1}{N} \sum_{i=1}^N V^i \xrightarrow{N \rightarrow \infty} 1/2$ almost surely.

As illustrated in Example 1, the key issue with the sample average approximation is that (2) is optimizing over an infinite-dimensional space of pricing policies, which leads (2) to output a pricing policy that overfits the samples.

In this note, we study a variant of (1) in which the buyer is *strategic*: given the seller's pricing policy, the buyer can augment the features that are revealed to the seller to obtain a low price for the product. The radii by which the buyer can perturb their features are positive random variables that are jointly distributed with the buyer's features and valuation. Such incorporation of strategic behavior with multi-dimensional features has received significant attention in recent years as an accurate approximation of how customers make decisions (Tsirtsis et al. 2024), and our model for incorporating strategic behavior into (1) can be viewed as a multi-dimensional generalization of intertemporal price discrimination in which customers are strategic in timing their purchases (Besbes and Lobel 2015).

Our main contribution of this note is establishing a *phase transition* for the sample average approximation between the non-strategic and strategic versions of personalized pricing. Specifically, for the strategic version of (1), we prove under mild assumptions that the objective value of the corresponding version of (2) is *asymptotically consistent*; that is, the objective value of the sample average approximation converges almost surely to the objective value of the stochastic problem as the number of samples tends to infinity. The asymptotic consistency guarantee holds when optimizing over the infinite-dimensional space of all pricing policies. Our result thus shows that accounting for strategic consumer behavior not only makes the personalized pricing problem (1) more realistic: it also helps prevent (2) from outputting pricing policies that overfit the samples. More generally, our result shows the potential for using strategic consumer behavior as a tool for augmenting the sample average approximation in order to obtain asymptotic consistency guarantees in stochastic programs with infinite-dimensional policy spaces.

Organization. In §2, we formalize the strategic version of personalized pricing and state our main result, Theorem 1. In §3, we present the proof. In §4, we illustrate the phase transition implied by Theorem 1 through a numerical experiment. In §5, we present technical details that were omitted from §3. Throughout the paper, random variables are represented by uppercase letters (e.g., \mathbf{L}) or by lowercase letters with a hat (e.g., $\hat{\nu}_N$). Uppercase letters can also represent a constant or function. Curly upper case letters (e.g., $\mathcal{F}, \mathfrak{A}$) represent sets. Lowercase letters represent a realization of a random variable or a number. A variable is bold if and only if it is a vector.

Related Literature. Our consistency result contributes to the literature on data-driven stochastic programming over infinite-dimensional policy spaces. For the non-strategic version of (1) in which the buyer has a one-dimensional feature (i.e, the setting where $D = 1$), Xie et al. (2024) propose nonparametric ways of imposing restrictions on the infinite-dimensional space of policies in (2) in order to obtain asymptotic consistency guarantees. In recent years, other papers have provided nonparametric techniques for preventing overfitting in stochastic programs with infinite-dimensional policy spaces based on incorporating adversarial noise into the sample average approximation (Xu et al. 2012, Shafieezadeh-Abadeh et al. 2019, Sturt 2023) and applying kernel smoothing to the samples in the sample average approximation (Pflug and Pichler 2016). We contribute to the aforementioned literature by showing that incorporating strategic consumer behavior into the sample average approximation can prevent overfitting in stochastic programs with infinite-dimensional policy spaces.

2. Problem Setting and Main Result

2.1. Problem Setting

We consider a strategic variant of the personalized pricing problem defined by a set of allowable prices, $\mathcal{P} \subseteq \mathbb{R}$, and a joint probability distribution for the buyer. The buyer drawn from that joint

probability distribution is represented by a random vector $(\mathbf{L}, \mathbf{U}, V) \in \mathcal{X} \times \mathcal{X} \times \mathcal{V}$ whose support is contained in $\mathcal{F} \times \mathcal{V}$ where $\mathcal{F} \triangleq \{(\boldsymbol{\ell}, \mathbf{u}) \in \mathcal{X} \times \mathcal{X} : \boldsymbol{\ell} \leq \mathbf{u}\}$. The random variable $V \in \mathcal{V}$ denotes the buyer's valuation, and the random vectors $\mathbf{L} \equiv (L_1, \dots, L_D) \in \mathcal{X}$ and $\mathbf{U} \equiv (U_1, \dots, U_D) \in \mathcal{X}$ define the random hyperrectangle $[\mathbf{L}, \mathbf{U}] \equiv \times_{d=1}^D [L_d, U_d]$ from which the buyer chooses the feature vector $\mathbf{x} \in [\mathbf{L}, \mathbf{U}]$ that is revealed to the seller.¹

A pricing policy refers to a measurable function $\pi : \mathcal{X} \rightarrow \mathcal{P}$ that specifies the price to offer to the buyer based on their revealed feature vector. In other words, if the buyer $(\mathbf{L}, \mathbf{U}, V)$ reveals $\mathbf{x} \in [\mathbf{L}, \mathbf{U}]$ as their feature vector, then π will offer the buyer a price of $\pi(\mathbf{x}) \in \mathcal{P}$. Given a pricing policy π and a buyer realization $(\boldsymbol{\ell}, \mathbf{u}, v) \in \mathcal{F} \times \mathcal{V}$, the revenue received by the seller is

$$R^\pi(\boldsymbol{\ell}, \mathbf{u}, v) \triangleq \mathbb{I} \left\{ \inf_{\mathbf{x} \in [\boldsymbol{\ell}, \mathbf{u}]} \pi(\mathbf{x}) \leq v \right\} \cdot \inf_{\mathbf{x} \in [\boldsymbol{\ell}, \mathbf{u}]} \pi(\mathbf{x}).$$

The zero-one indicator $\mathbb{I} \{ \inf_{\mathbf{x} \in [\boldsymbol{\ell}, \mathbf{u}]} \pi(\mathbf{x}) \leq v \}$ implies that the buyer realization $(\boldsymbol{\ell}, \mathbf{u}, v)$ will purchase the product if and only if there exists a feature vector $\mathbf{x} \in [\boldsymbol{\ell}, \mathbf{u}]$ that the buyer can reveal to the seller that generates a price $\pi(\mathbf{x})$ that is less than or equal to the buyer's valuation v . If the buyer purchases the product, then the buyer will maximize their utility by revealing the feature vector $\mathbf{x} \in [\boldsymbol{\ell}, \mathbf{u}]$ that leads to the minimum price. The goal of the seller is to choose a pricing policy that maximizes the seller's expected revenue, which is captured by the following stochastic program:

$$\nu^* \triangleq \sup_{\pi: \mathcal{X} \rightarrow \mathcal{P}} \left\{ J^*(\pi) \triangleq \mathbb{E} [R^\pi(\mathbf{L}, \mathbf{U}, V)] \right\}. \quad (\text{OPT})$$

2.2. Main Result

Our main result establishes the asymptotic consistency of the objective value of the sample average approximation with respect to the objective value ν^* of (OPT). We make the following assumptions:

ASSUMPTION 1. $(\mathbf{L}, \mathbf{U}, V), (\mathbf{L}^1, \mathbf{U}^1, V^1), (\mathbf{L}^2, \mathbf{U}^2, V^2), \dots$ is a sequence of independent and identically distributed random vectors with support contained in $\mathcal{F} \times \mathcal{V}$.

ASSUMPTION 2. $\mathcal{X} = [0, 1]^D$, $\mathcal{V} = [0, 1]$, and $\mathcal{P} \subseteq [0, 1]$ is closed and nonempty.

ASSUMPTION 3. $(\mathbf{L}, \mathbf{U}) \in \mathbb{R}^{2D}$ is a continuous random vector with probability density function $g : \mathbb{R}^{2D} \rightarrow [0, \infty)$.

Assumption 1 is a standard assumption in the stochastic programming literature and allows the random variables $\mathbf{L}, \mathbf{U}, V$ to be correlated with one another. Assumption 2 is a boundedness assumption on the feature vector and product valuations, which is imposed to simplify our proofs,

¹The random hyperrectangle $[\mathbf{L}, \mathbf{U}] \subseteq \mathbb{R}^D$ can be equivalently interpreted as the hyperrectangle $[\mathbf{X} - \mathbf{E}, \mathbf{X} + \mathbf{E}]$ where $\mathbf{X} \triangleq \frac{1}{2}(\mathbf{L} + \mathbf{U})$ is the buyer's random feature vector and $\mathbf{E} \triangleq \frac{1}{2}(\mathbf{U} - \mathbf{L})$ is the random radii by which the buyer can perturb each of their features.

and a regularity assumption on the set of prices \mathcal{P} . Similarly as Xie et al. (2024), our main result can be extended to the case of features and valuations with unbounded supports under mild probabilistic assumptions using standard truncation arguments. We also note that the regularity assumption on \mathcal{P} is satisfied if there exist prices $p_1 < \dots < p_K$ such that $\mathcal{P} = \{p_1, \dots, p_K\}$. Assumption 3 says that (\mathbf{L}, \mathbf{U}) is drawn from a continuous probability distribution with respect to \mathbb{R}^{2D} , and thus implies that $\mathbb{P}((\mathbf{L}, \mathbf{U}) \in \mathcal{Z}) = 0$ for any measurable set $\mathcal{Z} \subseteq \mathbb{R}^{2D}$ with zero Lebesgue measure. We note that Assumption 3 is satisfied if $\mathbf{L} = \mathbf{X} - \boldsymbol{\epsilon}$ and $\mathbf{U} = \mathbf{X} + \boldsymbol{\epsilon}$ for fixed radii $\boldsymbol{\epsilon} > \mathbf{0}$ and a random vector $\mathbf{X} \in \mathbb{R}^D$ that satisfies $\mathbb{P}(\mathbf{X} \in [\boldsymbol{\epsilon}, \mathbf{1} - \boldsymbol{\epsilon}]) = 1$ and is absolutely continuous with respect to the Lebesgue measure on \mathbb{R}^D . We show in Appendix A that Assumption 3 is a necessary condition for asymptotic consistency, in the sense that relaxing this assumption can invalidate our main result.

Our main result is the following:

THEOREM 1. *Suppose Assumptions 1, 2, and 3 hold, and let*

$$\widehat{\nu}_N \triangleq \sup_{\pi: \mathcal{X} \rightarrow \mathcal{P}} \left\{ \widehat{J}_N(\pi) \triangleq \frac{1}{N} \sum_{i=1}^N R^\pi(\mathbf{L}^i, \mathbf{U}^i, V^i) \right\} \quad (\text{SAA})$$

for each integer $N > 0$. Then $\lim_{N \rightarrow \infty} \widehat{\nu}_N = \nu^*$ almost surely.

3. Proof of Theorem 1

We assume throughout this section that Assumptions 1–3 hold. We additionally assume that:

ASSUMPTION 4. *There exist $0 < p_1 < \dots < p_K \leq 1$ such that $\mathcal{P} = \{p_1, \dots, p_K\}$.*

This assumption is unnecessary for establishing Theorem 1 but simplifies the presentation of its proof. In §5.1, we prove that Theorem 1 holds in the absence of Assumption 4.

It follows from Shapiro et al. (2009, Equation (5.6)) that (SAA) is an asymptotically optimistic estimate of (OPT), in the sense that $\liminf_{N \rightarrow \infty} \widehat{\nu}_N \geq \nu^*$ almost surely. It remains to prove the other direction of that inequality. To this end, we introduce a grid-based partitioning of the feature space. Specifically, for each integer $S > 0$, we consider partitioning $[0, 1]^D$ into hypercubes of width $1/S$, where the hypercubes are indexed by vectors $\boldsymbol{\sigma} \equiv (\sigma_1, \dots, \sigma_D) \in \{1, \dots, S\}^D$. The set of feature vectors in the hypercube corresponding to $\boldsymbol{\sigma}$ is denoted by $\mathcal{H}_S(\boldsymbol{\sigma}) \triangleq \times_{d=1}^D \mathcal{H}_{S,d}(\sigma_d)$ where

$$\mathcal{H}_{S,d}(\sigma_d) \triangleq \begin{cases} \left[\frac{\sigma_d - 1}{S}, \frac{\sigma_d}{S} \right) & \text{if } d \in \{1, \dots, D-1\} \\ \left[\frac{\sigma_d - 1}{S}, \frac{\sigma_d}{S} \right] & \text{if } d = D. \end{cases}$$

We readily observe that $\{\mathcal{H}_S(\boldsymbol{\sigma}) : \boldsymbol{\sigma} \in \{1, \dots, S\}^D\}$ is a collection of disjoint sets whose union is equal to $[0, 1]^D$. This implies that every feature vector $\mathbf{x} \in [0, 1]^D$ is contained in exactly one hypercube. For notational convenience, we denote the indices of the hypercube that contains feature

vector $\mathbf{x} \in [0, 1]^D$ by $\boldsymbol{\sigma}_S(\mathbf{x}) \equiv (\sigma_{S,1}(x_1), \dots, \sigma_{S,D}(x_D)) \in \{1, \dots, S\}^D$; that is, $\boldsymbol{\sigma}_S(\mathbf{x})$ is defined as the unique vector of indices that satisfies $\mathbf{x} \in \mathcal{H}_S(\boldsymbol{\sigma}_S(\mathbf{x}))$.

Let the set of all pricing policies that are piecewise constant over the grid-based partitioning of $[0, 1]^D$ into hypercubes of width $1/S$ be denoted by

$$\Pi_S \triangleq \left\{ \pi \in \Pi \mid \text{if } \boldsymbol{\sigma} \in \{1, \dots, S\}^D \text{ and } \mathbf{x}, \mathbf{x}' \in \mathcal{H}_S(\boldsymbol{\sigma}), \text{ then } \pi(\mathbf{x}) = \pi(\mathbf{x}') \right\}.$$

It follows from the above definitions that the restricted space of pricing policies Π_S has a finite number of elements. More specifically, we observe from Assumption 4 that for each integer $S > 0$, the restricted space of pricing policies has a cardinality of $|\Pi_S| = K^{S^D}$. It follows from this observation that the objective value of the sample average approximation is a consistent estimator of the objective value of the stochastic problem when optimizing over the restricted space Π_S , in the sense that the following line holds for every integer $S > 0$:²

$$\lim_{N \rightarrow \infty} \sup_{\pi \in \Pi_S} \widehat{J}_N(\pi) = \sup_{\pi \in \Pi_S} J^*(\pi) \quad \text{almost surely.} \quad (3)$$

The remainder of the proof of Theorem 1 consists of showing for every sufficiently large integer $S > 0$ and for every pricing policy $\pi : [0, 1]^D \rightarrow \mathcal{P}$ that there exists a pricing policy in the restricted space Π_S yielding approximately the same objective value with respect to the sample average approximation. To show this, let $\Pi \triangleq \{\pi : [0, 1]^D \rightarrow \mathcal{P}\}$ denote the space of all measurable pricing policies, and let $T_S : \Pi \rightarrow \Pi_S$ be an operator defined as

$$(T_S \circ \pi)(\mathbf{x}) \triangleq \min_{\mathbf{x}' \in \mathcal{H}_S(\boldsymbol{\sigma}_S(\mathbf{x}))} \pi(\mathbf{x}') \quad \forall \pi \in \Pi, \mathbf{x} \in [0, 1]^D.$$

The pricing policy $T_S \circ \pi$ can be interpreted as a ‘rounded’ version of pricing policy $\pi \in \Pi$ that, for each $\boldsymbol{\sigma} \in \{1, \dots, S\}^D$, offers the constant price $\min_{\mathbf{x}' \in \mathcal{H}_S(\boldsymbol{\sigma})} \pi(\mathbf{x}')$ to all feature vectors in the hypercube $\mathcal{H}_S(\boldsymbol{\sigma})$. The following relates the asymptotic performance of $T_S \circ \pi$ relative to π .

PROPOSITION 1. $\limsup_{S \rightarrow \infty} \limsup_{N \rightarrow \infty} \sup_{\pi \in \Pi} \left\{ \widehat{J}_N(\pi) - \widehat{J}_N(T_S \circ \pi) \right\} = 0$ almost surely.

The proof of Proposition 1 is found in §3.1. Combining Proposition 1 and (3), we have

$$\begin{aligned} \limsup_{N \rightarrow \infty} \widehat{v}_N &= \limsup_{N \rightarrow \infty} \sup_{\pi \in \Pi} \widehat{J}_N(\pi) = \limsup_{S \rightarrow \infty} \limsup_{N \rightarrow \infty} \sup_{\pi \in \Pi} \left\{ \widehat{J}_N(\pi) - \widehat{J}_N(T_S \circ \pi) + \widehat{J}_N(T_S \circ \pi) \right\} \\ &\leq \limsup_{S \rightarrow \infty} \limsup_{N \rightarrow \infty} \sup_{\pi \in \Pi} \left\{ \widehat{J}_N(\pi) - \widehat{J}_N(T_S \circ \pi) \right\} + \limsup_{S \rightarrow \infty} \limsup_{N \rightarrow \infty} \sup_{\pi \in \Pi_S} \widehat{J}_N(\pi) \\ &= \limsup_{S \rightarrow \infty} \lim_{N \rightarrow \infty} \sup_{\pi \in \Pi_S} \widehat{J}_N(\pi) \quad \text{almost surely} \end{aligned}$$

² More precisely, (3) follows from Assumption 1, from the fact that $|\Pi_S| < \infty$, and from the fact that each $\pi \in \Pi_S$ is measurable with respect to the Lebesgue measure over $[0, 1]^D$, which implies the function $(\boldsymbol{\ell}, \mathbf{u}, v) \in \mathcal{F} \times [0, 1] \mapsto R^\pi(\boldsymbol{\ell}, \mathbf{u}, v) \in [0, 1]$ is integrable (Shapiro et al. 2009, Proposition 5.2 and Theorem 7.48).

$$\begin{aligned}
&= \limsup_{S \rightarrow \infty} \sup_{\pi \in \Pi_S} J^*(\pi) \quad \text{almost surely} \\
&\leq \nu^*.
\end{aligned}$$

The first and second equalities and the first inequality follow from algebra. The third equality follows from Proposition 1. The fourth equality follows from (3). The final inequality follows from the fact that $\Pi_S \subseteq \Pi$ for all $S > 0$. Combining the above with the fact that $\liminf_{N \rightarrow \infty} \widehat{\nu}_N \geq \nu^*$ almost surely completes the proof of Theorem 1.

3.1. Proof of Proposition 1

Given any integer $S > 0$, we define a bucket as any tuple of the form $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_S$ where

$$\mathcal{B}_S \triangleq \{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \{1, \dots, S\}^D \times \{1, \dots, S\}^D : \boldsymbol{\lambda} \leq \boldsymbol{\mu}\}.$$

We say henceforth that a buyer realization $(\boldsymbol{\ell}, \mathbf{u}, v) \in \mathcal{F} \times [0, 1]$ is mapped to bucket $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_S$ if and only if $\boldsymbol{\lambda} = \boldsymbol{\sigma}_S(\boldsymbol{\ell})$ and $\boldsymbol{\mu} = \boldsymbol{\sigma}_S(\mathbf{u})$. That is, buyer realization $(\boldsymbol{\ell}, \mathbf{u}, v)$ is mapped to bucket $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ if and only if $\boldsymbol{\ell}$ is an element of the hypercube $\mathcal{H}_S(\boldsymbol{\lambda})$ and \mathbf{u} is an element of the hypercube $\mathcal{H}_S(\boldsymbol{\mu})$. It follows that every buyer realization is mapped to exactly one bucket.

Let the buyer realizations that are mapped to bucket $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ be denoted by

$$\mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu}) \triangleq \{(\boldsymbol{\ell}, \mathbf{u}, v) \in \mathcal{F} \times [0, 1] : \boldsymbol{\lambda} = \boldsymbol{\sigma}_S(\boldsymbol{\ell}), \boldsymbol{\mu} = \boldsymbol{\sigma}_S(\mathbf{u})\}.$$

For all integers $S, N > 0$ and buckets $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_S$, let the subset of the N random buyers that are mapped to bucket $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ be denoted by

$$\widehat{\mathcal{N}}_{S,N}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \triangleq \{i \in \{1, \dots, N\} : (\mathbf{L}^i, \mathbf{U}^i, V^i) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})\}.$$

For each pricing policy $\pi \in \Pi$, let the objective value of the sample average approximation corresponding to π for the random buyers that are mapped to bucket $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ be denoted by

$$\widehat{J}_{S,N}(\pi; \boldsymbol{\lambda}, \boldsymbol{\mu}) \triangleq \frac{1}{|\widehat{\mathcal{N}}_{S,N}(\boldsymbol{\lambda}, \boldsymbol{\mu})|} \sum_{i \in \widehat{\mathcal{N}}_{S,N}(\boldsymbol{\lambda}, \boldsymbol{\mu})} R^\pi(\mathbf{L}^i, \mathbf{U}^i, V^i),$$

where we use the convention that $1/0 \times 0 = 0$. Finally, let the probability that a random buyer is mapped to bucket $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ be denoted by $p_S(\boldsymbol{\lambda}, \boldsymbol{\mu}) \triangleq \mathbb{P}(\boldsymbol{\sigma}_S(\mathbf{L}) = \boldsymbol{\lambda}, \boldsymbol{\sigma}_S(\mathbf{U}) = \boldsymbol{\mu})$.

It follows from the above notations that

$$\begin{aligned}
&\limsup_{S \rightarrow \infty} \limsup_{N \rightarrow \infty} \sup_{\pi \in \Pi} \left\{ \widehat{J}_N(\pi) - \widehat{J}_N(T_S \circ \pi) \right\} \\
&= \limsup_{S \rightarrow \infty} \limsup_{N \rightarrow \infty} \sup_{\pi \in \Pi} \left\{ \sum_{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_S} \frac{|\widehat{\mathcal{N}}_{S,N}(\boldsymbol{\lambda}, \boldsymbol{\mu})|}{N} \left(\widehat{J}_{S,N}(\pi; \boldsymbol{\lambda}, \boldsymbol{\mu}) - \widehat{J}_{S,N}(T_S \circ \pi; \boldsymbol{\lambda}, \boldsymbol{\mu}) \right) \right\}
\end{aligned}$$

$$\begin{aligned}
&\leq \limsup_{S \rightarrow \infty} \limsup_{N \rightarrow \infty} \sup_{\pi \in \Pi} \left\{ \sum_{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_S} \frac{|\widehat{\mathcal{N}}_{S,N}(\boldsymbol{\lambda}, \boldsymbol{\mu})|}{N} \sup_{(\boldsymbol{\ell}, \mathbf{u}, v) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})} \{R^\pi(\boldsymbol{\ell}, \mathbf{u}, v) - R^{T_{S \circ \pi}}(\boldsymbol{\ell}, \mathbf{u}, v)\} \right\} \\
&= \limsup_{S \rightarrow \infty} \sup_{\pi \in \Pi} \left\{ \sum_{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_S} p_S(\boldsymbol{\lambda}, \boldsymbol{\mu}) \sup_{(\boldsymbol{\ell}, \mathbf{u}, v) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})} \{R^\pi(\boldsymbol{\ell}, \mathbf{u}, v) - R^{T_{S \circ \pi}}(\boldsymbol{\ell}, \mathbf{u}, v)\} \right\} \text{ almost surely, } (4)
\end{aligned}$$

where the first equality and the first inequality follow from algebra and the second equality follows from the uniform law of large numbers.³

The remainder of the proof of Proposition 1 consists of showing, roughly speaking, that the inner-most supremum in line (4) is nonpositive for *most* buckets. More precisely, for each integer $S > 0$ and pricing policy $\pi \in \Pi$, let the buckets for which the inner-most supremum is positive be denoted by

$$\mathcal{A}_S^\pi \triangleq \left\{ (\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_S : \sup_{(\boldsymbol{\ell}, \mathbf{u}, v) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})} \{R^\pi(\boldsymbol{\ell}, \mathbf{u}, v) - R^{T_{S \circ \pi}}(\boldsymbol{\ell}, \mathbf{u}, v)\} > 0 \right\}.$$

The following proposition shows that the probability of a random buyer mapping to a bucket in \mathcal{A}_S^π converges to zero as S tends to infinity, uniformly over all pricing policies.

PROPOSITION 2. $\limsup_{S \rightarrow \infty} \sup_{\pi \in \Pi} \sum_{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{A}_S^\pi} p_S(\boldsymbol{\lambda}, \boldsymbol{\mu}) = 0$.

The proof of Proposition 2 is found in §3.2. Equipped with the above proposition, we observe that

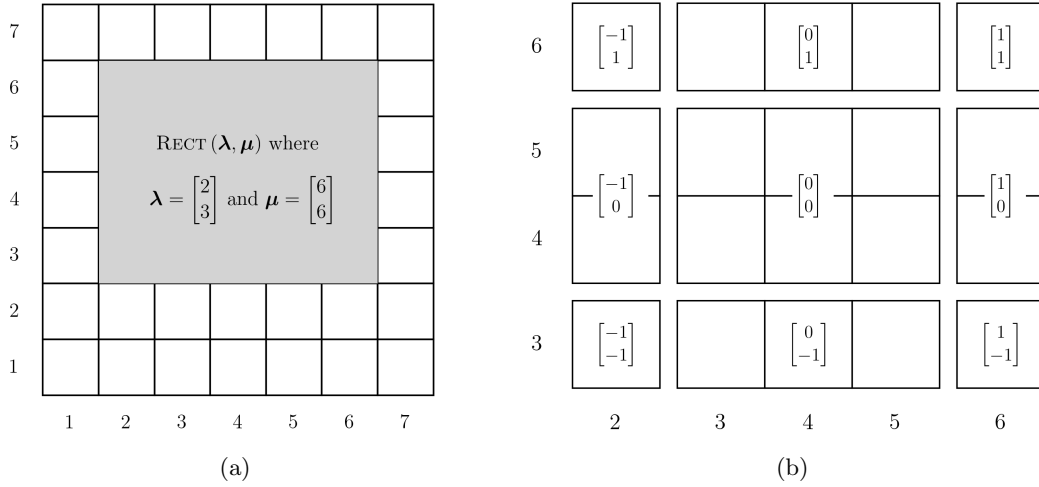
$$\begin{aligned}
(4) &\leq \limsup_{S \rightarrow \infty} \sup_{\pi \in \Pi} \left\{ \sum_{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_S \setminus \mathcal{A}_S^\pi} p_S(\boldsymbol{\lambda}, \boldsymbol{\mu}) \sup_{(\boldsymbol{\ell}, \mathbf{u}, v) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})} \{R^\pi(\boldsymbol{\ell}, \mathbf{u}, v) - R^{T_{S \circ \pi}}(\boldsymbol{\ell}, \mathbf{u}, v)\} \right\} \\
&\quad + \limsup_{S \rightarrow \infty} \sup_{\pi \in \Pi} \left\{ \sum_{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{A}_S^\pi} p_S(\boldsymbol{\lambda}, \boldsymbol{\mu}) \sup_{(\boldsymbol{\ell}, \mathbf{u}, v) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})} \{R^\pi(\boldsymbol{\ell}, \mathbf{u}, v) - R^{T_{S \circ \pi}}(\boldsymbol{\ell}, \mathbf{u}, v)\} \right\} \\
&\leq 0 + \limsup_{S \rightarrow \infty} \sup_{\pi \in \Pi} \left\{ \sum_{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{A}_S^\pi} p_S(\boldsymbol{\lambda}, \boldsymbol{\mu}) \cdot 1 \right\} = 0,
\end{aligned}$$

where the first inequality follows from algebra, the second inequality follows from the fact that $\sup_{(\boldsymbol{\ell}, \mathbf{u}, v) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})} \{R^\pi(\boldsymbol{\ell}, \mathbf{u}, v) - R^{T_{S \circ \pi}}(\boldsymbol{\ell}, \mathbf{u}, v)\} \leq 0$ for all $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_S \setminus \mathcal{A}_S^\pi$ and from the fact that $R^\pi(\boldsymbol{\ell}, \mathbf{u}, v) - R^{T_{S \circ \pi}}(\boldsymbol{\ell}, \mathbf{u}, v) \leq 1$ for all $(\boldsymbol{\ell}, \mathbf{u}, v) \in \mathcal{F} \times [0, 1]$, and the equality follows from Proposition 2. That concludes the proof of Proposition 1.

3.2. Proof of Proposition 2

Consider any arbitrary $S > 1$ and pricing policy $\pi \in \Pi$.

³The uniform law of large numbers (Shapiro et al. 2009, Theorem 7.48) says for every $S > 0$ and $\delta > 0$, there exists a random variable \bar{N} satisfying $\bar{N} < \infty$ almost surely such that for all $N \geq \bar{N}$ and all $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_S$, we have $p_S(\boldsymbol{\lambda}, \boldsymbol{\mu}) - \delta \leq |\widehat{\mathcal{N}}_{S,N}(\boldsymbol{\lambda}, \boldsymbol{\mu})|/N \leq p_S(\boldsymbol{\lambda}, \boldsymbol{\mu}) + \delta$. Hence, line (4) follows from choosing $\delta > 0$ to be arbitrarily small.

Figure 1 Visualizations from §3.2

Note. (a) Visualization of grid-based partitioning of the feature space $[0, 1]^2$ into hypercubes of width $\frac{1}{7}$. (b) Visualization of $\text{FACE}((2, 3), (6, 6), \delta)$ for each $\delta \in \{-1, 0, 1\}^2$. Note that the union of the δ -faces in (b) is equal to the hyperrectangle $\text{RECT}((2, 3), (6, 6))$ which is shown shaded in (a).

We begin by introducing a necessary condition (Lemma 1) that must be satisfied by every bucket that is contained in \mathcal{A}_S^π . To state this necessary condition, we use the following terminology. For each bucket $(\lambda, \mu) \in \mathcal{B}_S$, we let the hyperrectangle formed by the set of hypercubes $\sigma \in \{1, \dots, S\}^D$ that satisfy $\lambda \leq \sigma \leq \mu$ be denoted by $\text{RECT}(\lambda, \mu) \triangleq \times_{d=1}^D \{\lambda_d, \dots, \mu_d\}$. Moreover, for each bucket $(\lambda, \mu) \in \mathcal{F}_S$ and each vector $\delta \in \{-1, 0, 1\}^D$, we define the δ -face of $\text{RECT}(\lambda, \mu)$ as

$$\text{FACE}(\lambda, \mu, \delta) \triangleq \left\{ \begin{array}{ll} \sigma_d = \lambda_d & \text{for all } d \in \{1, \dots, D\} \text{ such that } \delta_d = -1 \\ \sigma : \sigma_d = \mu_d & \text{for all } d \in \{1, \dots, D\} \text{ such that } \delta_d = 1 \\ \sigma_d \in \{\lambda_d + 1, \dots, \mu_d - 1\} & \text{for all } d \in \{1, \dots, D\} \text{ such that } \delta_d = 0 \end{array} \right\}.$$

It is easy to verify for each bucket $(\lambda, \mu) \in \mathcal{B}_S$ that $\{\text{FACE}(\lambda, \mu, \delta) : \delta \in \{-1, 0, 1\}^D\}$ is a collection of sets whose union is equal to $\text{RECT}(\lambda, \mu)$. In Figures 1a and 1b, we visualize hyperrectangles and δ -faces. We now present our necessary condition for a bucket to be an element of \mathcal{A}_S^π .

LEMMA 1. *If $(\lambda, \mu) \in \mathcal{A}_S^\pi$, then $\min_{\sigma \in \text{FACE}(\lambda, \mu, \mathbf{0})} \min_{\mathbf{x} \in \mathcal{H}_S(\sigma)} \pi(\mathbf{x}) > \min_{\delta \in \{-1, 0, 1\}^D : \delta \neq \mathbf{0}} \min_{\sigma \in \text{FACE}(\lambda, \mu, \delta)} \min_{\mathbf{x} \in \mathcal{H}_S(\sigma)} \pi(\mathbf{x})$.*

The proof of the above lemma (as well as the proof of each subsequent lemma in §3) can be found in §5.2. In words, the above lemma says that if a bucket satisfies $(\lambda, \mu) \in \mathcal{A}_S^\pi$, then there must exist a feature vector in one of the hypercubes in the boundary of the hyperrectangle corresponding to the bucket, $\mathbf{x} \in \cup_{\delta \in \{-1, 0, 1\}^D : \delta \neq \mathbf{0}} \cup_{\sigma \in \text{FACE}(\lambda, \mu, \delta)} \mathcal{H}_S(\sigma)$, for which the pricing policy offers a price, $\pi(\mathbf{x})$, that is strictly less than the price offered by the pricing policy for all feature vectors in the

hypercubes in the interior of the hyperrectangle corresponding to the bucket. For every integer $M \in \{1, \dots, S-1\}$, it follows from Lemma 1 that

$$\begin{aligned} \mathcal{A}_S^\pi &\subseteq \bigcup_{\delta \in \{-1, 0, 1\}^D: \delta \neq \mathbf{0}} \left\{ (\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_S : \min_{\boldsymbol{\sigma} \in \text{FACE}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{0})} \min_{\mathbf{x} \in \mathcal{H}_S(\boldsymbol{\sigma})} \pi(\mathbf{x}) > \min_{\boldsymbol{\sigma} \in \text{FACE}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\delta})} \min_{\mathbf{x} \in \mathcal{H}_S(\boldsymbol{\sigma})} \pi(\mathbf{x}) \right\} \\ &\subseteq \mathcal{B}_{S,M}^< \cup \underbrace{\bigcup_{\delta \in \{-1, 0, 1\}^D: \delta \neq \mathbf{0}} \left\{ (\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_{S,M}^> : \min_{\boldsymbol{\sigma} \in \text{FACE}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{0})} \min_{\mathbf{x} \in \mathcal{H}_S(\boldsymbol{\sigma})} \pi(\mathbf{x}) > \min_{\boldsymbol{\sigma} \in \text{FACE}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\delta})} \min_{\mathbf{x} \in \mathcal{H}_S(\boldsymbol{\sigma})} \pi(\mathbf{x}) \right\}}_{\triangleq \mathcal{A}_{S,M,\delta}^\pi} \end{aligned} \quad (5)$$

where the sets $\mathcal{B}_{S,M}^<$ and $\mathcal{B}_{S,M}^>$ in (5) are defined below as

$$\mathcal{B}_{S,M}^< \triangleq \left\{ (\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_S : \min_{d \in \{1, \dots, D\}} \{\mu_d - \lambda_d\} < M \right\} \text{ and } \mathcal{B}_{S,M}^> \triangleq \left\{ (\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_S : \min_{d \in \{1, \dots, D\}} \{\mu_d - \lambda_d\} \geq M \right\}.$$

For each integer $M \in \{1, \dots, S-1\}$, it follows from algebra that

$$|\mathcal{B}_{S,M}^<| \leq \sum_{m=0}^{M-1} \sum_{d=1}^D |\{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_S : \mu_d - \lambda_d = m\}| \leq MDS^{2D-1}. \quad (6)$$

Moreover, the following proposition establishes a conservative upper bound on the cardinality of $\mathcal{A}_{S,M,\delta}^\pi$ for every integer $M \in \{1, \dots, S-1\}$ and every vector $\boldsymbol{\delta} \in \{-1, 0, 1\}^D$ that satisfies $\boldsymbol{\delta} \neq \mathbf{0}$:

PROPOSITION 3. $|\mathcal{A}_{S,M,\delta}^\pi| \leq \frac{2DKS^{2D}}{M}$.

Since $\pi \in \Pi$ was chosen arbitrarily, we observe from the above reasoning that the following holds for every arbitrary $S > 2$:

$$\begin{aligned} \sup_{\pi \in \Pi} \frac{|\mathcal{A}_S^\pi|}{S^{2D}} &\leq \sup_{\pi \in \Pi} \left\{ \frac{1}{S^{2D}} \cdot \min_{M \in \{1, \dots, S-1\}} \left\{ |\mathcal{B}_{S,M}^<| + \sum_{\delta \in \{-1, 0, 1\}^D: \delta \neq \mathbf{0}} |\mathcal{A}_{S,M,\delta}^\pi| \right\} \right\} \\ &\leq \frac{1}{S^{2D}} \cdot \min_{M \in \{1, \dots, S-1\}} \left\{ MDS^{2D-1} + \sum_{\delta \in \{-1, 0, 1\}^D: \delta \neq \mathbf{0}} \frac{2DKS^{2D}}{M} \right\} \\ &\leq \underbrace{\frac{D}{S} \left(\lceil \sqrt{S} \rceil + \frac{(3^D - 1)2KS}{\lceil \sqrt{S} \rceil} \right)}_{\triangleq \beta_S}. \end{aligned} \quad (7)$$

Indeed, the first inequality follows from (5), the second inequality follows from (6) and Proposition 3, the third inequality follows from setting $M = \lceil \sqrt{S} \rceil$. Since S was chosen arbitrarily, we have

$$\limsup_{S \rightarrow \infty} \sup_{\pi \in \Pi} \frac{|\mathcal{A}_S^\pi|}{S^{2D}} \leq \lim_{S \rightarrow \infty} \beta_S = 0, \quad (8)$$

where the inequality follows from (7) and the equality follows from the definition of β_S .

To interpret the upper bound from line (8), we recall that the number of buckets satisfies $|\mathcal{B}_S| = S^{2D}$. Line (8) can thus be interpreted as showing that the proportion $\frac{|\mathcal{A}_S^\pi|}{|\mathcal{B}_S|}$ of buckets that are in

the set \mathcal{A}_S^π converges to zero as S tends to infinity, uniformly over all $\pi \in \Pi$. To conclude the proof of Proposition 2, we use Assumption 3 to show that the probability of a random buyer being contained in any vanishingly small proportion of buckets converges to zero as S tends to infinity. This result is shown in the following lemma.

LEMMA 2. *Given any sequence $\alpha_1, \alpha_2, \dots \in (0, 1)$ that satisfies $\lim_{S \rightarrow \infty} \alpha_S = 0$, we have*

$$\lim_{S \rightarrow \infty} \sup_{\mathcal{A}_S \subseteq \mathcal{B}_S: \frac{|\mathcal{A}_S|}{|\mathcal{B}_S|} \leq \alpha_S} \sum_{(\lambda, \mu) \in \mathcal{A}_S} p_S(\lambda, \mu) = 0.$$

We conclude that

$$\limsup_{S \rightarrow \infty} \sup_{\pi \in \Pi} \sum_{(\lambda, \mu) \in \mathcal{A}_S^\pi} p_S(\lambda, \mu) \leq \limsup_{S \rightarrow \infty} \sup_{\mathcal{A}_S \subseteq \mathcal{B}_S: \frac{|\mathcal{A}_S|}{|\mathcal{B}_S|} \leq \beta_S} \sum_{(\lambda, \mu) \in \mathcal{A}_S} p_S(\lambda, \mu) = 0,$$

where the first inequality follows from line (7) and the equality follows from Lemma 2 and line (8). This completes our proof of Proposition 2.

3.3. Proof of Proposition 3

Consider any arbitrary $S > 1$, pricing policy $\pi \in \Pi$, vector $\delta \in \{-1, 0, 1\}^D$ that satisfies $\delta \neq \mathbf{0}$, and integer $M \in \{1, \dots, S-1\}$. We say henceforth that a bucket $(\lambda, \mu) \in \mathcal{B}_{S,M}^\geq$ is a δ -anchor if and only if there exists $d \in \{1, \dots, D\}$ such that $\delta_d \neq 0$, $\lambda_d = 1$ if $\delta_d = 1$, and $\mu_d = S$ if $\delta_d = -1$. For convenience, we denote the set of all δ -anchors by

$$\mathfrak{A}_{S,M}(\delta) \triangleq \bigcup_{d \in \{1, \dots, D\}: \delta_d = 1} \{(\lambda, \mu) \in \mathcal{B}_{S,M}^\geq : \lambda_d = 1\} \cup \bigcup_{d \in \{1, \dots, D\}: \delta_d = -1} \{(\lambda, \mu) \in \mathcal{B}_{S,M}^\geq : \mu_d = S\}.$$

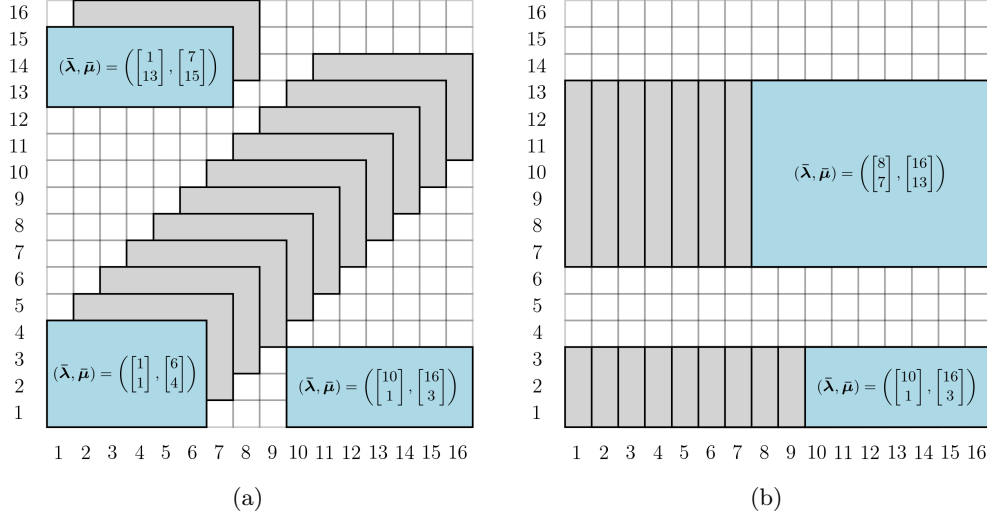
Furthermore, we say henceforth that $(\lambda, \mu) \in \mathcal{B}_{S,M}^\geq$ is contained in the δ -line emanating from $(\bar{\lambda}, \bar{\mu}) \in \mathfrak{A}_{S,M}(\delta)$ if and only if there exists an integer $\iota \geq 0$ such that $(\lambda, \mu) = (\bar{\lambda}, \bar{\mu}) + \iota(\delta, \delta)$. For convenience, we denote the δ -line emanating from $(\bar{\lambda}, \bar{\mu}) \in \mathfrak{A}_{S,M}(\delta)$ by

$$\mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta) \triangleq \{(\lambda, \mu) \in \mathcal{B}_{S,M}^\geq : \exists \text{ integer } \iota \geq 0 \text{ such that } (\lambda, \mu) = (\bar{\lambda}, \bar{\mu}) + \iota(\delta, \delta)\}.$$

A visualization of δ -anchors and δ -lines is shown in Figure 2. In the following two intermediary lemmas, we give a conservative upper bound on the number of δ -anchors, and we show that the δ -lines emanating from the δ -anchors form a cover of $\mathcal{B}_{S,M}^\geq$.

LEMMA 3. $|\mathfrak{A}_{S,M}(\delta)| \leq DS^{2D-1}$.

LEMMA 4. $\{\mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta) : (\bar{\lambda}, \bar{\mu}) \in \mathfrak{A}_{S,M}(\delta)\}$ is a collection of sets whose union is equal to $\mathcal{B}_{S,M}^\geq$.

Figure 2 Visualizations of δ -anchors and δ -lines

Note. Visualization of a subset of δ -anchors and the δ -lines emanating from those δ -anchors for the case of $S = 16$. Specifically, (a) shows three examples of δ -anchors for the case of $\delta = (1,1)$, and (b) shows two examples of δ -anchors for the case of $\delta = (-1,0)$. In both (a) and (b), the δ -anchors are shown in blue. For each δ -anchor $(\bar{\lambda}, \bar{\mu})$, the blue rectangle together with the corresponding grey rectangles constitute the δ -line emanating from $(\bar{\lambda}, \bar{\mu})$. For example, in (a), the δ -line emanating from the δ -anchor $(\bar{\lambda}, \bar{\mu}) = ((1,1), (6,4))$ is equal to the set of buckets $\{((1,1) + \iota(1,1), (6,4) + \iota(1,1)) : \iota \in \{0, \dots, 9\}\} = \{((1,1), (6,4)), \dots, ((11,16), (11,14))\}$.

It follows from Lemmas 3 and 4 that

$$|\mathcal{A}_{S,M,\delta}^\pi| \leq DS^{2D-1} \max_{(\bar{\lambda}, \bar{\mu}) \in \mathfrak{A}_{S,M}(\delta)} |\mathcal{A}_{S,M,\delta}^\pi \cap \mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta)|. \quad (9)$$

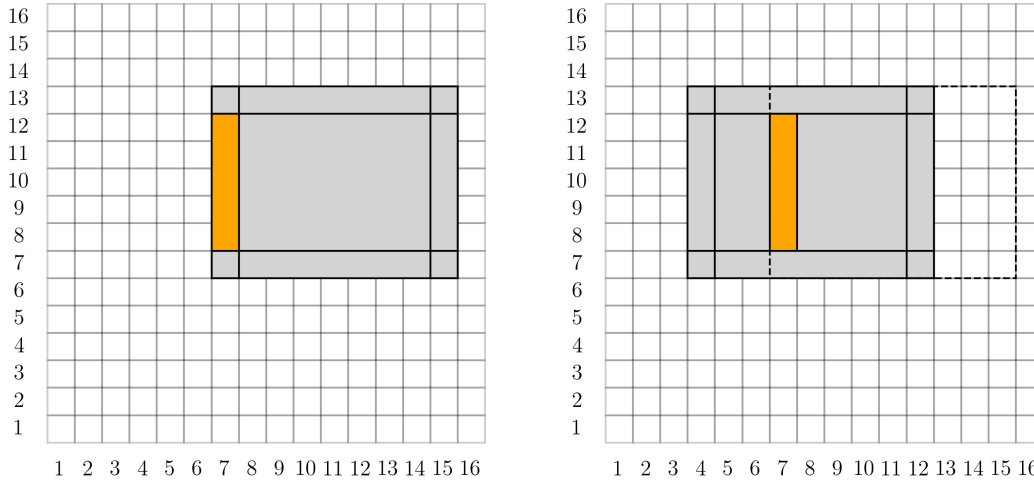
Now consider any arbitrary δ -anchor $(\bar{\lambda}, \bar{\mu}) \in \mathfrak{A}_{S,M}(\delta)$. To develop an upper bound on $|\mathcal{A}_{S,M,\delta}^\pi \cap \mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta)|$, we begin by decomposing $\mathcal{A}_{S,M,\delta}^\pi$ into K subsets of buckets,

$$\mathcal{A}_{S,M,\delta}^\pi = \bigcup_{k=1}^K \underbrace{\left\{ (\lambda, \mu) \in \mathcal{B}_{S,M}^\geq : p_k = \min_{\sigma \in \text{FACE}(\lambda, \mu, \mathbf{0})} \min_{\mathbf{x} \in \mathcal{H}_S(\sigma)} \pi(\mathbf{x}) \text{ and } p_k > \min_{\sigma \in \text{FACE}(\lambda, \mu, \delta)} \min_{\mathbf{x} \in \mathcal{H}_S(\sigma)} \pi(\mathbf{x}) \right\}}_{\triangleq \mathcal{A}_{S,M,\delta,k}^\pi},$$

where the equality follows from the definition of $\mathcal{A}_{S,M,\delta}$ and from Assumption 4, which implies that π is a function of the form $\pi : [0, 1]^D \rightarrow \{p_1, \dots, p_K\}$. Therefore, we observe that

$$|\mathcal{A}_{S,M,\delta}^\pi \cap \mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta)| \leq K \max_{k \in \{1, \dots, K\}} |\mathcal{A}_{S,M,\delta,k}^\pi \cap \mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta)|. \quad (10)$$

Now consider any arbitrary $k \in \{1, \dots, K\}$. To develop an upper bound on $|\mathcal{A}_{S,M,\delta,k}^\pi \cap \mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta)|$, our high-level strategy is to show that any two buckets $(\lambda, \mu), (\lambda', \mu')$ that are contained in $\mathcal{A}_{S,M,\delta,k}^\pi \cap \mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta)$ must be sufficiently far apart from one another. We show this using an intermediary lemma denoted below by Lemma 5. The lemma shows that if two buckets on the δ -line emanating from the δ -anchor $(\bar{\lambda}, \bar{\mu})$ are sufficiently close to one another, then the δ -face of the

Figure 3 Visualization of Lemma 5

Note. Visualization of Lemma 5 for the δ -line emanating from the δ -anchor $(\bar{\lambda}, \bar{\mu}) = ((8, 7), (16, 13))$ for $\delta = (-1, 0)$ (see Figure 2b). The gray rectangles in the left and right figures correspond to $\text{RECT}(\bar{\lambda} + \iota\delta, \bar{\mu} + \iota\delta)$ for the case of $\iota = 1$ and $\iota = 4$, respectively. The orange rectangle in both figures corresponds to $\text{FACE}(\bar{\lambda} + 1\delta, \bar{\mu} + 1\delta, \delta)$. We observe from the right figure that $\text{FACE}(\bar{\lambda} + 1\delta, \bar{\mu} + 1\delta, \delta) \subseteq \text{FACE}(\bar{\lambda} + 4\delta, \bar{\mu} + 4\delta, \mathbf{0})$.

hyperrectangle corresponding to one bucket must be contained in the interior of the hyperrectangle corresponding to the other bucket.

LEMMA 5. *Let $(\lambda, \mu), (\lambda', \mu') \in \mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta)$. If there exists an integer $\iota \in \{1, \dots, M-1\}$ such that $(\lambda', \mu') = (\lambda, \mu) + \iota(\delta, \delta)$, then $\text{FACE}(\lambda, \mu, \delta) \subseteq \text{FACE}(\lambda', \mu', \mathbf{0})$.*

To provide further intuition for Lemma 5, we recall that if two buckets satisfy $(\lambda, \mu), (\lambda', \mu') \in \mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta)$, there must exist nonnegative integers ι, ι' such that $(\lambda, \mu) = (\bar{\lambda}, \bar{\mu}) + \iota(\delta, \delta)$ and $(\lambda', \mu') = (\bar{\lambda}, \bar{\mu}) + \iota'(\delta, \delta)$. If $\iota' - \iota \in \{1, \dots, M-1\}$, then Lemma 5 establishes that the δ -face of the hyperrectangle corresponding to (λ, μ) must be contained in the interior of the hyperrectangle corresponding to (λ', μ') . A visualization of Lemma 5 is found in Figure 3.

Equipped with Lemma 5, we proceed to establish our upper bound on $|\mathcal{A}_{S,M,\delta,k}^\pi \cap \mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta)|$. Indeed, let $\bar{\iota} \triangleq |\mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta)|$ denote the maximal integer such that $(\bar{\lambda}, \bar{\mu}) + \bar{\iota}(\delta, \delta) \in \mathcal{B}_{S,M}^\geq$. Note that it follows from the fact that $\delta \neq \mathbf{0}$ and from the fact that $(\bar{\lambda}, \bar{\mu}) \in \mathcal{B}_{S,M}^\geq$ that the integer $\bar{\iota}$ satisfies $\bar{\iota} \leq S-1$. We observe that

$$\begin{aligned} \mathcal{A}_{S,M,\delta,k}^\pi \cap \mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta) &= \mathcal{A}_{S,M,\delta,k}^\pi \cap \{(\bar{\lambda} + \iota\delta, \bar{\mu} + \iota\delta) : \iota \in \{0, \dots, \bar{\iota}\}\} \\ &= \left\{ (\bar{\lambda} + \iota\delta, \bar{\mu} + \iota\delta) : \begin{array}{l} \iota \in \{0, \dots, \bar{\iota}\} \\ p_k = \min_{\sigma \in \text{FACE}(\bar{\lambda} + \iota\delta, \bar{\mu} + \iota\delta, \mathbf{0})} \min_{\mathbf{x} \in \mathcal{H}_S(\sigma)} \pi(\mathbf{x}) \\ p_k > \min_{\sigma \in \text{FACE}(\bar{\lambda} + \iota\delta, \bar{\mu} + \iota\delta, \delta)} \min_{\mathbf{x} \in \mathcal{H}_S(\sigma)} \pi(\mathbf{x}) \end{array} \right\}, \quad (11) \end{aligned}$$

where the first equality follows from the definition of $\mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta)$ and the second equality follows from the definition of $\mathcal{A}_{S,M,\delta,k}^\pi$. Consider any arbitrary $\iota \in \{0, \dots, \bar{\iota}\}$ that satisfies $(\bar{\lambda} + \iota\delta, \bar{\mu} + \iota\delta) \in \mathcal{A}_{S,M,\delta,k}^\pi \cap \mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta)$ and any arbitrary $\iota' \in \{\iota + 1, \dots, \min\{\iota + M - 1, \bar{\iota}\}\}$. It follows from line (11) that $p_k > \min_{\sigma \in \text{FACE}(\bar{\lambda} + \iota\delta, \bar{\mu} + \iota\delta, \delta)} \min_{\mathbf{x} \in \mathcal{H}_S(\sigma)} \pi(\mathbf{x})$, and it follows from Lemma 5 that $\text{FACE}(\bar{\lambda} + \iota\delta, \bar{\mu} + \iota\delta, \delta) \subseteq \text{FACE}(\bar{\lambda} + \iota'\delta, \bar{\mu} + \iota'\delta, \mathbf{0})$, which together imply that $p_k > \min_{\sigma \in \text{FACE}(\bar{\lambda} + \iota'\delta, \bar{\mu} + \iota'\delta, \mathbf{0})} \min_{\mathbf{x} \in \mathcal{H}_S(\sigma)} \pi(\mathbf{x})$. We thus conclude from line (11) that $(\bar{\lambda} + \iota'\delta, \bar{\mu} + \iota'\delta) \notin \mathcal{A}_{S,M,\delta,k}^\pi \cap \mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta)$. More generally, we have shown that if two integers $\hat{\iota}, \hat{\iota}' \in \{0, \dots, \bar{\iota}\}$ satisfy $\hat{\iota} < \hat{\iota}'$ and $(\bar{\lambda} + \hat{\iota}\delta, \bar{\mu} + \hat{\iota}\delta), (\bar{\lambda} + \hat{\iota}'\delta, \bar{\mu} + \hat{\iota}'\delta) \in \mathcal{A}_{S,M,\delta,k}^\pi \cap \mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta)$, then it must be the case that $\hat{\iota}' - \hat{\iota} \geq M$. Therefore, we conclude that

$$|\mathcal{A}_{S,M,\delta,k}^\pi \cap \mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta)| \leq \left\lceil \frac{\bar{\iota} + 1}{M} \right\rceil \leq \left\lceil \frac{S}{M} \right\rceil, \quad (12)$$

where the first inequality follows from our earlier reasoning that $\hat{\iota}' - \hat{\iota} \geq M$ for all $\hat{\iota}, \hat{\iota}' \in \{0, \dots, \bar{\iota}\}$ that satisfy $\hat{\iota} < \hat{\iota}'$ and $(\bar{\lambda} + \hat{\iota}\delta, \bar{\mu} + \hat{\iota}\delta), (\bar{\lambda} + \hat{\iota}'\delta, \bar{\mu} + \hat{\iota}'\delta) \in \mathcal{A}_{S,M,\delta,k}^\pi \cap \mathcal{L}_{S,M}(\bar{\lambda}, \bar{\mu}, \delta)$, and the second inequality follows from our earlier observation that $\bar{\iota} \leq S - 1$.

Because the vector $\delta \in \{-1, 0, 1\}^D$ that satisfies $\delta \neq \mathbf{0}$, the δ -anchor $(\bar{\lambda}, \bar{\mu}) \in \mathfrak{A}_{S,M}(\delta)$, and the integer $k \in \{1, \dots, K\}$ were chosen arbitrarily, we conclude from lines (9), (10), and (12) that

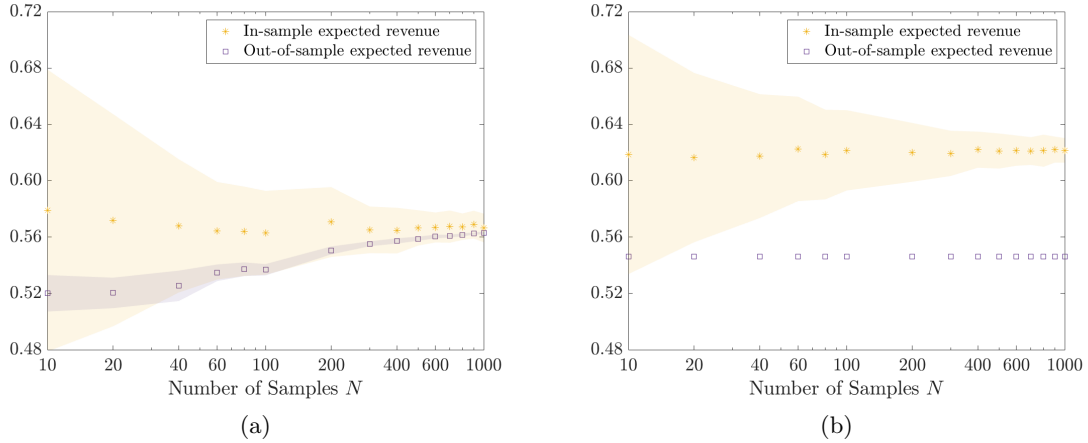
$$|\mathcal{A}_{S,M,\delta}^\pi| \leq DS^{2D-1}K \left\lceil \frac{S}{M} \right\rceil \leq \frac{2DS^{2D}K}{M},$$

which completes the proof of Proposition 3.

4. Numerical Experiment

To illustrate the phase transition established by Theorem 1, we conduct numerical experiments that compare the non-strategic and strategic versions of a simple personalized pricing problem with two features, *i.e.*, $D = 2$. The probability distribution for the buyer in our experiments is given by $(\mathbf{L}, \mathbf{U}, V) = (\mathbf{X} - \epsilon, \mathbf{X} + \epsilon, V)$, where \mathbf{X} is uniformly distributed over $[0.1, 0.9]^2$, the radii ϵ are chosen as either $\epsilon = (0.09, 0.09)$ or $\epsilon = (0, 0)$, and the buyer's valuation of the product is a deterministic function of \mathbf{L}, \mathbf{U} of the form $V = L_1(1 - L_1) + L_2(1 - L_2) + U_1(1 - U_1) + U_2(1 - U_2)$. Our probability distribution is chosen such that the buyer has a higher (resp., lower) valuation if \mathbf{X} is closer to the center (resp., the boundary) of $[0, 1]^2$. The set of feasible prices for the seller is given by $\mathcal{P} = \{p_1, p_2\}$ where $p_1 = 0.65$ and $p_2 = 0.83$. It can be readily verified that the above problem satisfies Assumptions 1-3 if $\epsilon = (0.09, 0.09)$ and violates Assumption 3 if $\epsilon = (0, 0)$.

In Figure 4, we visualize the phase transition between the strategic and non-strategic versions of (SAA). Specifically, let $\hat{\pi}_N$ denote an optimal solution for (SAA). In Figure 4a, we observe for the case of $\epsilon = (0.09, 0.09)$ that the in-sample expected revenue $\hat{v}_N \equiv \hat{J}_N(\hat{\pi}_N)$ and the out-of-sample expected revenue $J^*(\hat{\pi}_N)$ of the pricing policy $\hat{\pi}_N$ converge to one another as the number of samples

Figure 4 Visualization of in-sample and out-of-sample performances.

Note. Each plot shows the (yellow) in-sample expected revenue $\widehat{J}_N(\widehat{\pi}_N)$ and the (purple) out-of-sample expected revenue $J^*(\widehat{\pi}_N)$, where $\widehat{\pi}_N$ is an optimal solution for (SAA) with (a) $\epsilon = (0.09, 0.09)$ and (b) $\epsilon = (0, 0)$. The solid lines describe the mean statistics over 50 random instances, whereas the shaded regions cover mean \pm standard deviation. The experiment details for solving (SAA) can be found in Appendix B.

N increases. That is, Figure 4a illustrates Theorem 1 by showing that $\widehat{\nu}_N$ converges to ν^* . In contrast, Figure 4b shows for the case of $\epsilon = (0, 0)$ that the in-sample objective value of (SAA) always overfits, even as the number of samples N increases.

5. Omitted Proofs from §3

5.1. Relaxation of Assumption 4

To show that Assumption 4 is unnecessary for the proof of Theorem 1, we prove the following:

PROPOSITION 4. *Let Assumptions 1, 2 hold. For each $K > 0$, there exist $p_1, \dots, p_K \in \mathcal{P}$ such that*

$$0 \leq \sup_{\pi: [0,1]^D \rightarrow \mathcal{P}} \frac{1}{N} \sum_{i=1}^N R^\pi(\mathbf{L}^i, \mathbf{U}^i, V^i) - \sup_{\pi: [0,1]^D \rightarrow \{p_1, \dots, p_K\}} \frac{1}{N} \sum_{i=1}^N R^\pi(\mathbf{L}^i, \mathbf{U}^i, V^i) \leq \frac{1}{K}$$

for every integer $N > 0$, almost surely, and

$$0 \leq \sup_{\pi: [0,1]^D \rightarrow \mathcal{P}} \mathbb{E}[R^\pi(\mathbf{L}, \mathbf{U}, V)] - \sup_{\pi: [0,1]^D \rightarrow \{p_1, \dots, p_K\}} \mathbb{E}[R^\pi(\mathbf{L}, \mathbf{U}, V)] \leq \frac{1}{K}.$$

The above proposition implies that Assumption 4 is unnecessary for the proof of Theorem 1 because the proposition shows, for every arbitrary $K > 0$, that there exist prices $p_1, \dots, p_K \in \mathcal{P}$ such that (OPT) and (SAA) can be approximated within a $\frac{1}{K}$ -additive error by restricting to pricing policies of the form $\pi: [0,1]^D \rightarrow \{p_1, \dots, p_K\}$. To prove Proposition 4, we use the following lemmas.

LEMMA 6. *Let Assumption 2 hold. For every integer $K > 0$, there exist prices $p_1, \dots, p_K \in \mathcal{P}$ such that for every $p \in \mathcal{P}$, there exists $k \in \{1, \dots, K\}$ such that $0 \leq p - p_k \leq \frac{1}{K}$.*

Proof of Lemma 6. Consider any integer $K > 0$. The Weierstrass extreme value theorem, in combination with the fact that the set $\mathcal{P} \neq \emptyset$ is closed and bounded (Assumption 2), implies that any continuous real-valued function over \mathcal{P} must attain a maximum and minimum. Hence, we have that the quantities $\underline{p}, \bar{p}, p_1, \dots, p_K$ defined below exist and are elements of \mathcal{P} :

$$\underline{p} \triangleq \inf_{p \in \mathcal{P}} p; \quad \bar{p} \triangleq \sup_{p \in \mathcal{P}} p; \quad p_k \triangleq \inf \left\{ p \in \mathcal{P} : p \geq \underline{p} + (\bar{p} - \underline{p}) \frac{k-1}{K} \right\} \quad \forall k \in \{1, \dots, K\}.$$

Now consider any arbitrary $p \in \mathcal{P}$. It follows from the fact that $p \in \mathcal{P} \subseteq [\underline{p}, \bar{p}]$ that there must exist some $k \in \{1, \dots, K\}$ such that $\underline{p} + (\bar{p} - \underline{p}) \frac{k-1}{K} \leq p \leq \underline{p} + (\bar{p} - \underline{p}) \frac{k}{K}$. For that k , we observe that $p_k = \inf \{ p' \in \mathcal{P} : \underline{p} + (\bar{p} - \underline{p}) \frac{k-1}{K} \leq p' \} \leq p$, where the inequality follows from the fact that $\underline{p} + (\bar{p} - \underline{p}) \frac{k-1}{K} \leq p$ and from the fact that $p \in \mathcal{P}$. We have thus shown that $p_k \leq p \leq \underline{p} + (\bar{p} - \underline{p}) \frac{k}{K}$, which implies that

$$\begin{aligned} 0 \leq p - p_k &\leq \underline{p} + (\bar{p} - \underline{p}) \frac{k}{K} - p_k \\ &\leq \underline{p} + (\bar{p} - \underline{p}) \frac{k}{K} - \left(\underline{p} + (\bar{p} - \underline{p}) \frac{k-1}{K} \right) \end{aligned} \quad (13)$$

$$\begin{aligned} &= (\bar{p} - \underline{p}) \frac{1}{K} \\ &\leq \frac{1}{K}, \end{aligned} \quad (14)$$

where line (13) follows from the fact that $p_k \geq \underline{p} + (\bar{p} - \underline{p}) \frac{k-1}{K}$ and line (14) follows from the fact that $\mathcal{P} \subseteq [0, 1]$. Since $p \in \mathcal{P}$ was chosen arbitrarily, our proof of Lemma 6 is complete. \square

LEMMA 7. *Let Assumption 2 hold. For every integer $K > 0$, there exist prices $p_1, \dots, p_K \in \mathcal{P}$ such that for every pricing policy $\pi : [0, 1]^D \rightarrow \mathcal{P}$, there exists a pricing policy $\pi' : [0, 1]^D \rightarrow \{p_1, \dots, p_K\}$ such that for all $(\ell, \mathbf{u}, v) \in \mathcal{F} \times [0, 1]$ we have $R^{\pi'}(\ell, \mathbf{u}, v) \geq R^\pi(\ell, \mathbf{u}, v) - \frac{1}{K}$.*

Proof of Lemma 7. Consider any integer $K > 0$, and let $p_1, \dots, p_K \in \mathcal{P}$ be the prices obtained from Lemma 6. Consider any pricing policy $\pi : [0, 1]^D \rightarrow \mathcal{P}$. It follows from Lemma 6 and from the fact that $\pi : [0, 1]^D \rightarrow \mathcal{P}$ that there exists a pricing policy $\pi' : [0, 1]^D \rightarrow \{p_1, \dots, p_K\}$ satisfying $0 \leq \pi(\mathbf{x}) - \pi'(\mathbf{x}) \leq \frac{1}{K}$ for all $\mathbf{x} \in [0, 1]^D$. Therefore, for all $(\ell, \mathbf{u}, v) \in \mathcal{F} \times [0, 1]$ we have

$$\begin{aligned} R^\pi(\ell, \mathbf{u}, v) &= \mathbb{I} \left\{ \inf_{\mathbf{x} \in [\ell, \mathbf{u}]} \pi(\mathbf{x}) \leq v \right\} \cdot \inf_{\mathbf{x} \in [\ell, \mathbf{u}]} \pi(\mathbf{x}) \leq \mathbb{I} \left\{ \inf_{\mathbf{x} \in [\ell, \mathbf{u}]} \pi'(\mathbf{x}) \leq v \right\} \cdot \inf_{\mathbf{x} \in [\ell, \mathbf{u}]} \pi(\mathbf{x}) \\ &\leq \mathbb{I} \left\{ \inf_{\mathbf{x} \in [\ell, \mathbf{u}]} \pi'(\mathbf{x}) \leq v \right\} \cdot \inf_{\mathbf{x} \in [\ell, \mathbf{u}]} \left\{ \pi'(\mathbf{x}) + \frac{1}{K} \right\} \leq \frac{1}{K} + \mathbb{I} \left\{ \inf_{\mathbf{x} \in [\ell, \mathbf{u}]} \pi'(\mathbf{x}) \leq v \right\} \cdot \inf_{\mathbf{x} \in [\ell, \mathbf{u}]} \pi'(\mathbf{x}) \\ &= \frac{1}{K} + R^{\pi'}(\ell, \mathbf{u}, v). \end{aligned}$$

The first equality is the definition of $R^\pi(\ell, \mathbf{u}, v)$. The first inequality holds because $0 \leq \pi(\mathbf{x}) - \pi'(\mathbf{x})$. The second inequality holds because $\pi(\mathbf{x}) - \pi'(\mathbf{x}) \leq \frac{1}{K}$. The third inequality follows from algebra, and the second equality follows from the definition of $R^{\pi'}(\ell, \mathbf{u}, v)$. \square

We conclude §5.1 by using Lemma 7 to prove Proposition 4.

Proof of Proposition 4. Consider any arbitrary integer $K > 0$, and let $p_1, \dots, p_K \in \mathcal{P}$ be the prices obtained from Lemma 7. For each arbitrary integer $N > 0$, it follows from algebra that

$$0 \leq \sup_{\pi: [0,1]^D \rightarrow \mathcal{P}} \frac{1}{N} \sum_{i=1}^N R^\pi(\mathbf{L}^i, \mathbf{U}^i, V^i) - \sup_{\pi: [0,1]^D \rightarrow \{p_1, \dots, p_K\}} \frac{1}{N} \sum_{i=1}^N R^\pi(\mathbf{L}^i, \mathbf{U}^i, V^i).$$

Moreover, choose any arbitrary $\eta > 0$, and let $\hat{\pi}_{N,\eta}: [0,1]^D \rightarrow \mathcal{P}$ denote a pricing policy that satisfies

$$\frac{1}{N} \sum_{i=1}^N R^{\hat{\pi}_{N,\eta}}(\mathbf{L}^i, \mathbf{U}^i, V^i) \geq \sup_{\pi: [0,1]^D \rightarrow \mathcal{P}} \frac{1}{N} \sum_{i=1}^N R^\pi(\mathbf{L}^i, \mathbf{U}^i, V^i) - \eta.$$

It follows from Lemma 7 that there exists a pricing policy $\hat{\pi}'_{N,\eta}: [0,1]^D \rightarrow \{p_1, \dots, p_K\}$ that satisfies $R^{\hat{\pi}'_{N,\eta}}(\ell, \mathbf{u}, v) \geq R^{\hat{\pi}_{N,\eta}}(\ell, \mathbf{u}, v) - \frac{1}{K}$ for all $(\ell, \mathbf{u}, v) \in \mathcal{F} \times [0,1]$. Therefore,

$$\begin{aligned} \sup_{\pi: [0,1]^D \rightarrow \mathcal{P}} \frac{1}{N} \sum_{i=1}^N R^\pi(\mathbf{L}^i, \mathbf{U}^i, V^i) &\leq \eta + \frac{1}{N} \sum_{i=1}^N R^{\hat{\pi}_{N,\eta}}(\mathbf{L}^i, \mathbf{U}^i, V^i) \\ &\leq \eta + \frac{1}{K} + \frac{1}{N} \sum_{i=1}^N R^{\hat{\pi}'_{N,\eta}}(\mathbf{L}^i, \mathbf{U}^i, V^i) \text{ almost surely} \\ &\leq \eta + \frac{1}{K} + \sup_{\pi: [0,1]^D \rightarrow \{p_1, \dots, p_K\}} \frac{1}{N} \sum_{i=1}^N R^\pi(\mathbf{L}^i, \mathbf{U}^i, V^i). \end{aligned}$$

The first inequality follows from the definition of $\hat{\pi}_{N,\eta}$. The second inequality follows from the fact that $(\mathbf{L}^1, \mathbf{U}^1, V^1), \dots, (\mathbf{L}^N, \mathbf{U}^N, V^N) \in \mathcal{F} \times [0,1]$ almost surely (Assumption 1). The third inequality follows from the fact that $\hat{\pi}'_{N,\eta}$ is a feasible but possibly suboptimal solution for the optimization problem $\sup_{\pi: [0,1]^D \rightarrow \{p_1, \dots, p_K\}} \frac{1}{N} \sum_{i=1}^N R^\pi(\mathbf{L}^i, \mathbf{U}^i, V^i)$. Since $N > 0$ and $\eta > 0$ were chosen arbitrarily, we conclude that the following inequality holds almost surely for every integer $N > 0$:

$$\sup_{\pi: [0,1]^D \rightarrow \mathcal{P}} \frac{1}{N} \sum_{i=1}^N R^\pi(\mathbf{L}^i, \mathbf{U}^i, V^i) - \sup_{\pi: [0,1]^D \rightarrow \{p_1, \dots, p_K\}} \frac{1}{N} \sum_{i=1}^N R^\pi(\mathbf{L}^i, \mathbf{U}^i, V^i) \leq \frac{1}{K}.$$

An identical reasoning is used to show the bound for the expectations. \square

5.2. Proofs of Intermediary Lemmas

Proof of Lemma 1. Consider any $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{A}_S^\pi$. We use the following two intermediary claims.

CLAIM 1. *If $(\ell, \mathbf{u}, v) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})$, then $\cup_{\sigma \in \text{FACE}(\boldsymbol{\lambda}, \boldsymbol{\mu}, 0)} \mathcal{H}_S(\sigma) \subseteq [\ell, \mathbf{u}]$.*

Proof of Claim 1. It follows from the definition of $\mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})$ that the following holds for all buyer realizations $(\ell, \mathbf{u}, v) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})$ and every $d \in \{1, \dots, D\}$:

$$\ell_d \in \mathcal{H}_{S,d}(\lambda_d) \subseteq \left[\frac{\lambda_d - 1}{S}, \frac{\lambda_d}{S} \right] \text{ and } u_d \in \mathcal{H}_{S,d}(\mu_d) \subseteq \left[\frac{\mu_d - 1}{S}, \frac{\mu_d}{S} \right]. \quad (15)$$

Thus, for every $(\ell, \mathbf{u}, v) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})$, $\boldsymbol{\sigma} \in \text{FACE}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{0})$, and $d \in \{1, \dots, D\}$, we have

$$\ell_d \leq \frac{\lambda_d}{S} \leq \frac{\sigma_d - 1}{S} \text{ and } u_d \geq \frac{\mu_d - 1}{S} \geq \frac{\sigma_d}{S}, \quad (16)$$

where the first inequalities for ℓ_d and u_d follow from (15), and the second inequalities follow from the fact that $\boldsymbol{\sigma} \in \text{FACE}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{0})$ (which implies that $\sigma_d \in \{\lambda_d + 1, \dots, \mu_d - 1\}$). We conclude from (16) that for every buyer realization $(\ell, \mathbf{u}, v) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})$, $\boldsymbol{\sigma} \in \text{FACE}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{0})$, and $d \in \{1, \dots, D\}$, it must be the case that $\mathcal{H}_{S,d}(\sigma_d) \subseteq [\frac{\sigma_d - 1}{S}, \frac{\sigma_d}{S}] \subseteq [\ell_d, u_d]$. We have therefore shown for every buyer realization $(\ell, \mathbf{u}, v) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})$ and $\boldsymbol{\sigma} \in \text{FACE}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{0})$ that $\mathcal{H}_S(\boldsymbol{\sigma}) \subseteq [\ell, \mathbf{u}]$. Since this reasoning holds for every $\boldsymbol{\sigma} \in \text{FACE}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{0})$, our proof of Claim 1 is complete. \square

CLAIM 2. *If $(\ell, \mathbf{u}, v) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})$, then $[\ell, \mathbf{u}] \subseteq \cup_{\boldsymbol{\sigma} \in \text{RECT}(\boldsymbol{\lambda}, \boldsymbol{\mu})} \mathcal{H}_S(\boldsymbol{\sigma})$.*

Proof of Claim 2. It follows from the definition of $\mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})$ that the following holds for all buyer realizations $(\ell, \mathbf{u}, v) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})$ and every $d \in \{1, \dots, D\}$:

$$\frac{\lambda_d - 1}{S} \leq \ell_d \text{ and } \begin{cases} u_d < \frac{\mu_d}{S} & \text{if } \mu_d < S \\ u_d \leq \frac{\mu_d}{S} & \text{if } \mu_d = S. \end{cases} \quad (17)$$

Thus, it follows from the fact that $\ell_d \leq u_d$ that

$$\begin{aligned} [\ell_d, u_d] &\subseteq \begin{cases} [\frac{\lambda_d - 1}{S}, \frac{\mu_d}{S}] & \text{if } \mu_d < S \\ [\frac{\lambda_d - 1}{S}, \frac{\mu_d}{S}] & \text{if } \mu_d = S \end{cases} \\ &= \mathcal{H}_{S,d}(\lambda_d) \cup \dots \cup \mathcal{H}_{S,d}(\mu_d). \end{aligned}$$

Hence, we conclude for every buyer realization $(\ell, \mathbf{u}, v) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})$ that

$$[\ell, \mathbf{u}] = \bigtimes_{d=1}^D [\ell_d, u_d] \subseteq \bigtimes_{d=1}^D \bigcup_{\sigma_d \in \{\lambda_d, \dots, \mu_d\}} \mathcal{H}_{S,d}(\sigma_d) = \bigcup_{\boldsymbol{\sigma} \in \text{RECT}(\boldsymbol{\lambda}, \boldsymbol{\mu})} \mathcal{H}_S(\boldsymbol{\sigma}),$$

which completes the proof of Claim 2. \square

Equipped with Claims 1 and 2, we proceed to complete the proof of Lemma 1. Indeed, it follows from the fact that $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{A}_S^\pi$ that there exist a buyer realization $(\ell, \mathbf{u}, v) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})$ and $k \in \{1, \dots, K\}$ satisfying $R^{T_S \circ \pi}(\ell, \mathbf{u}, v) < R^\pi(\ell, \mathbf{u}, v) = p_k$. Consider any such buyer realization and k . We observe that

$$\min_{\boldsymbol{\sigma} \in \text{FACE}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{0})} \min_{\mathbf{x} \in \mathcal{H}_S(\boldsymbol{\sigma})} \pi(\mathbf{x}) = \min_{\mathbf{x} \in \cup_{\boldsymbol{\sigma} \in \text{FACE}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{0})} \mathcal{H}_S(\boldsymbol{\sigma})} \pi(\mathbf{x}) \geq \min_{\mathbf{x} \in [\ell, \mathbf{u}]} \pi(\mathbf{x}) = R^\pi(\ell, \mathbf{u}, v) = p_k. \quad (18)$$

Indeed, the first equality follows from algebra. The inequality follows from Claim 1. The second equality follows from Assumption 4 (which implies that $p_1, \dots, p_K > 0$) and from the fact that $R^\pi(\ell, \mathbf{u}, v) = p_k$, which together imply that $\mathbb{I}\{\min_{\mathbf{x} \in [\ell, \mathbf{u}]} \pi(\mathbf{x}) \leq v\} = 1$. The third equality follows from our choice of the buyer realization $(\ell, \mathbf{u}, v) \in \mathcal{C}_S(\boldsymbol{\lambda}, \boldsymbol{\mu})$. Moreover, we observe that

$$p_k > R^{T_S \circ \pi}(\ell, \mathbf{u}, v) = \mathbb{I}\left\{ \min_{\mathbf{x} \in [\ell, \mathbf{u}]} (T_S \circ \pi)(\mathbf{x}) \leq v \right\} \cdot \min_{\mathbf{x} \in [\ell, \mathbf{u}]} (T_S \circ \pi)(\mathbf{x})$$

$$\begin{aligned}
&\geq \mathbb{I} \left\{ \min_{\mathbf{x} \in [\ell, \mathbf{u}]} \pi(\mathbf{x}) \leq v \right\} \cdot \min_{\mathbf{x} \in [\ell, \mathbf{u}]} (T_S \circ \pi)(\mathbf{x}) = \min_{\mathbf{x} \in [\ell, \mathbf{u}]} (T_S \circ \pi)(\mathbf{x}) \\
&\geq \min_{\mathbf{x} \in \cup_{\sigma \in \text{RECT}(\lambda, \mu)} \mathcal{H}_S(\sigma)} (T_S \circ \pi)(\mathbf{x}) = \min_{\sigma \in \text{RECT}(\lambda, \mu)} \min_{\mathbf{x} \in \mathcal{H}_S(\sigma)} (T_S \circ \pi)(\mathbf{x}) \\
&= \min_{\sigma \in \text{RECT}(\lambda, \mu)} \min_{\mathbf{x} \in \mathcal{H}_S(\sigma)} \min_{\mathbf{x}' \in \mathcal{H}_S(\sigma_S(\mathbf{x}))} \pi(\mathbf{x}') = \min_{\sigma \in \text{RECT}(\lambda, \mu)} \min_{\mathbf{x} \in \mathcal{H}_S(\sigma)} \pi(\mathbf{x}) \\
&= \min_{\delta \in \{-1, 0, 1\}^D} \min_{\sigma \in \text{FACE}(\lambda, \mu, \delta)} \min_{\mathbf{x} \in \mathcal{H}_S(\sigma)} \pi(\mathbf{x}) = \min_{\delta \in \{-1, 0, 1\}^D: \delta \neq \mathbf{0}} \min_{\sigma \in \text{FACE}(\lambda, \mu, \delta)} \min_{\mathbf{x} \in \mathcal{H}_S(\sigma)} \pi(\mathbf{x}). \quad (19)
\end{aligned}$$

Indeed, the first inequality follows from our choice of the buyer realization $(\ell, \mathbf{u}, v) \in \mathcal{C}_S(\lambda, \mu)$. The first equality is the definition of $R^{T_S \circ \pi}(\ell, \mathbf{u}, v)$. The second inequality follows from the definition of $T_S \circ \pi$, which implies that $(T_S \circ \pi)(\mathbf{x}) \leq \pi(\mathbf{x})$ for all $\mathbf{x} \in [0, 1]^D$. The second equality follows from the fact that $R^\pi(\ell, \mathbf{u}, v) > 0$, which implies that $\mathbb{I} \{ \min_{\mathbf{x} \in [\ell, \mathbf{u}]} \pi(\mathbf{x}) \leq v \} = 1$. The third inequality follows from Claim 2. The third equality follows from algebra. The fourth equality is the definition of $T_S \circ \pi$. The fifth equality follows from the fact that $\sigma_S(\mathbf{x}) = \sigma$ for all $\mathbf{x} \in \mathcal{H}_S(\sigma)$. The sixth equality follows from the fact that $\text{RECT}(\lambda, \mu) = \cup_{\sigma \in \{-1, 0, 1\}^D} \text{FACE}(\lambda, \mu, \delta)$. The seventh equality follows from the fact that $\min_{\sigma \in \text{FACE}(\lambda, \mu, \mathbf{0})} \min_{\mathbf{x} \in \mathcal{H}_S(\sigma)} \pi(\mathbf{x}) \geq p_k$, which is established in line (18). Combining lines (18) and (19), our proof of Lemma 1 is complete. \square

Proof of Lemma 2. For notational convenience, let $\mathbf{Z} \triangleq (\mathbf{L}, \mathbf{U}) \in \mathcal{F}$. For each Borel set $\mathcal{Z} \in \mathcal{B}(\mathbb{R}^{2D})$, recall that $\int_{\mathcal{Z}} d\mathbf{z}$ denotes the volume of \mathcal{Z} with respect to the Lebesgue measure. Our proof of Lemma 2 will use the following intermediary claim.

CLAIM 3. $\lim_{\rho \downarrow 0} \sup_{\mathcal{Z} \in \mathcal{B}(\mathbb{R}^{2D}): \int_{\mathcal{Z}} d\mathbf{z} \leq \rho} \mathbb{P}(\mathbf{Z} \in \mathcal{Z}) = 0$.

Proof of Claim 3. We observe that the function $h(\rho) \triangleq \sup_{\mathcal{Z} \in \mathcal{B}(\mathbb{R}^{2D}): \int_{\mathcal{Z}} d\mathbf{z} \leq \rho} \mathbb{P}(\mathbf{Z} \in \mathcal{Z})$ is bounded and monotonically increasing for $\rho > 0$. Therefore, it follows from the fact that $h(\rho) \in [0, 1]$ that the limit $\lim_{\rho \downarrow 0} h(\rho)$ exists. Hence,

$$\lim_{\rho \downarrow 0} \sup_{\mathcal{Z} \in \mathcal{B}(\mathbb{R}^{2D}): \int_{\mathcal{Z}} d\mathbf{z} \leq \rho} \mathbb{P}(\mathbf{Z} \in \mathcal{Z}) = \lim_{n \rightarrow \infty} \sup_{\mathcal{Z} \in \mathcal{B}(\mathbb{R}^{2D}): \int_{\mathcal{Z}} d\mathbf{z} \leq \frac{1}{n^2}} \mathbb{P}(\mathbf{Z} \in \mathcal{Z}). \quad (20)$$

Choose any arbitrary constant $\eta > 0$. For each integer $n \in \mathbb{N}$, let $\mathcal{Z}_n^\eta \in \mathcal{B}(\mathbb{R}^{2D})$ denote a Borel set that satisfies both the inequality $\int_{\mathcal{Z}_n^\eta} d\mathbf{z} \leq \frac{1}{n^2}$ and the inequality $\sup_{\mathcal{Z} \in \mathcal{B}(\mathbb{R}^{2D}): \int_{\mathcal{Z}} d\mathbf{z} \leq \frac{1}{n^2}} \mathbb{P}(\mathbf{Z} \in \mathcal{Z}) \leq \mathbb{P}(\mathbf{Z} \in \mathcal{Z}_n^\eta) + \eta$. Therefore, we observe that

$$\lim_{n \rightarrow \infty} \sup_{\mathcal{Z} \in \mathcal{B}(\mathbb{R}^{2D}): \int_{\mathcal{Z}} d\mathbf{z} \leq \frac{1}{n^2}} \mathbb{P}(\mathbf{Z} \in \mathcal{Z}) \leq \limsup_{n \rightarrow \infty} \mathbb{P}(\mathbf{Z} \in \mathcal{Z}_n^\eta) + \eta \leq \mathbb{P} \left(\mathbf{Z} \in \limsup_{n \rightarrow \infty} \mathcal{Z}_n^\eta \right) + \eta, \quad (21)$$

where the first inequality is the definition of \mathcal{Z}_n^η and the second inequality follows from Fatou's lemma. Finally, we observe that $\sum_{n=1}^{\infty} \int_{\mathcal{Z}_n^\eta} d\mathbf{z} \leq \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$, and so it follows from the Borel-Cantelli lemma that the equality $\int_{\limsup_{n \rightarrow \infty} \mathcal{Z}_n^\eta} d\mathbf{z} = 0$ holds. Combining this equality with lines (20)-(21) and the fact that \mathbf{Z} is absolutely continuous with respect to the Lebesgue measure on \mathbb{R}^{2D} (Assumption 3), we have shown that $\lim_{\rho \downarrow 0} \sup_{\mathcal{Z} \in \mathcal{B}(\mathbb{R}^{2D}): \int_{\mathcal{Z}} d\mathbf{z} \leq \rho} \mathbb{P}(\mathbf{Z} \in \mathcal{Z}) \leq \eta$. Since $\eta > 0$ was chosen arbitrarily, our proof of Claim 3 is complete. \square

Equipped with Claim 3, we are ready to prove Lemma 2. We observe that

$$\begin{aligned} \lim_{S \rightarrow \infty} \sup_{\mathcal{A}_S \subseteq \mathcal{B}_S: \frac{|\mathcal{A}_S|}{|\mathcal{B}_S|} \leq \alpha_S} \sum_{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{A}_S} p_S(\boldsymbol{\lambda}, \boldsymbol{\mu}) &= \lim_{S \rightarrow \infty} \sup_{\mathcal{A}_S \subseteq \mathcal{B}_S: |\mathcal{A}_S| \leq \alpha_S S^{2D}} \sum_{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{A}_S} p_S(\boldsymbol{\lambda}, \boldsymbol{\mu}) \\ &\leq \lim_{S \rightarrow \infty} \sup_{\mathcal{Z} \in \mathcal{B}(\mathbb{R}^{2D}): \int_{\mathcal{Z}} dz \leq \alpha_S} \mathbb{P}(\mathbf{Z} \in \mathcal{Z}) = 0. \end{aligned}$$

Indeed, the first equality follows from algebra and from the fact that $|\mathcal{B}_S| = S^{2D}$. The inequality follows from algebra and from the fact that $\int_{\{(\boldsymbol{\ell}, \mathbf{u}) \in \mathcal{F}: \boldsymbol{\lambda} = \boldsymbol{\sigma}_S(\boldsymbol{\ell}) \text{ and } \boldsymbol{\mu} = \boldsymbol{\sigma}_S(\mathbf{u})\}} dz = \frac{1}{S^{2D}}$ for all $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_S$. The second equality follows from Claim 3 and from the fact that $\lim_{S \rightarrow \infty} \alpha_S = 0$. \square

Proof of Lemma 3. Recall that $\boldsymbol{\delta} \neq \mathbf{0}$. Therefore,

$$\begin{aligned} |\mathfrak{A}_{S,M}(\boldsymbol{\delta})| &= \left| \bigcup_{d \in \{1, \dots, D\}: \delta_d = 1} \{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_{S,M}^{\geq} : \lambda_d = 1\} \cup \bigcup_{d \in \{1, \dots, D\}: \delta_d = -1} \{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_{S,M}^{\geq} : \mu_d = S\} \right| \\ &\leq \sum_{d \in \{1, \dots, D\}: \delta_d = 1} |\{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_{S,M}^{\geq} : \lambda_d = 1\}| + \sum_{d \in \{1, \dots, D\}: \delta_d = -1} |\{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_{S,M}^{\geq} : \mu_d = S\}| \\ &\leq \sum_{d \in \{1, \dots, D\}: \delta_d = 1} S^{2D-1} + \sum_{d \in \{1, \dots, D\}: \delta_d = -1} S^{2D-1} \leq DS^{2D-1}. \end{aligned}$$

The first equality is the definition of $\mathfrak{A}_{S,M}(\boldsymbol{\delta})$. The first inequality follows from algebra. The second inequality follows from the fact that $\mathcal{B}_{S,M}^{\geq} \subseteq \mathcal{B}_S$. The third inequality follows from algebra. \square

Proof of Lemma 4. Our proof consists of showing for all $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_{S,M}^{\geq}$ that there exists an anchor $(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}) \in \mathfrak{A}_{S,M}(\boldsymbol{\delta})$ that satisfies $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{L}_{S,M}(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}, \boldsymbol{\delta})$. Indeed, consider any arbitrary $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_{S,M}^{\geq}$, and define $\iota \triangleq \min \{\iota' \in \{0, 1, \dots, S\} : (\boldsymbol{\lambda} - \iota' \boldsymbol{\delta}, \boldsymbol{\mu} - \iota' \boldsymbol{\delta}) \in \mathcal{B}_{S,M}^{\geq}\}$ and $(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}) \triangleq (\boldsymbol{\lambda} - \iota \boldsymbol{\delta}, \boldsymbol{\mu} - \iota \boldsymbol{\delta})$. It follows immediately from these definitions that $(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}) \in \mathcal{B}_{S,M}^{\geq}$, that $(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}})$ is a $\boldsymbol{\delta}$ -anchor, and that $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ is contained in the line emanating from anchor $(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}) \in \mathfrak{A}_{S,M}(\boldsymbol{\delta})$. Since $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathcal{B}_{S,M}^{\geq}$ was chosen arbitrarily, our proof is complete. \square

Proof of Lemma 5. Let $(\boldsymbol{\lambda}, \boldsymbol{\mu}), (\boldsymbol{\lambda}', \boldsymbol{\mu}') \in \mathcal{L}_{S,M}(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}, \boldsymbol{\delta})$, and suppose there exists an integer $\iota \in \{1, \dots, M-1\}$ such that $(\boldsymbol{\lambda}', \boldsymbol{\mu}') = (\boldsymbol{\lambda}, \boldsymbol{\mu}) + \iota(\boldsymbol{\delta}, \boldsymbol{\delta})$. We observe that

$$\begin{aligned} \text{FACE}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\delta}) &= \left\{ \begin{array}{ll} \sigma_d = \lambda'_d - \iota \delta_d & \text{for all } d \in \{1, \dots, D\} \text{ such that } \delta_d = -1 \\ \sigma_d = \mu'_d - \iota \delta_d & \text{for all } d \in \{1, \dots, D\} \text{ such that } \delta_d = 1 \\ \sigma_d \in \{\lambda'_d - \iota \delta_d + 1, \dots, \mu'_d - \iota \delta_d - 1\} & \text{for all } d \in \{1, \dots, D\} \text{ such that } \delta_d = 0 \end{array} \right\} \\ &= \left\{ \begin{array}{ll} \sigma_d = \lambda'_d + \iota & \text{for all } d \in \{1, \dots, D\} \text{ such that } \delta_d = -1 \\ \sigma_d = \mu'_d - \iota & \text{for all } d \in \{1, \dots, D\} \text{ such that } \delta_d = 1 \\ \sigma_d \in \{\lambda'_d + 1, \dots, \mu'_d - 1\} & \text{for all } d \in \{1, \dots, D\} \text{ such that } \delta_d = 0 \end{array} \right\} \\ &\subseteq \{\boldsymbol{\sigma} : \sigma_d \in \{\lambda'_d + 1, \dots, \mu'_d - 1\} \text{ for all } d \in \{1, \dots, D\}\} = \text{FACE}(\boldsymbol{\lambda}', \boldsymbol{\mu}', \mathbf{0}). \end{aligned}$$

The first equality follows from the definition of $\text{FACE}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\delta})$ and from the fact that $(\boldsymbol{\lambda}', \boldsymbol{\mu}') = (\boldsymbol{\lambda}, \boldsymbol{\mu}) + \iota(\boldsymbol{\delta}, \boldsymbol{\delta})$. The second equality follows from substituting δ_d . To see why the inclusion holds,

we observe that $(\boldsymbol{\lambda}', \boldsymbol{\mu}') \in \mathcal{B}_{S,M}^{\geq}$ implies that $\mu'_d - \lambda'_d \geq M$ for all $d \in \{1, \dots, D\}$. Therefore, for all $d \in \{1, \dots, D\}$, it follows from the fact that $\mu'_d - \lambda'_d \geq M$ and from the fact that $\iota \in \{1, \dots, M-1\}$ that $\lambda'_d + \iota \in \{\lambda'_d + 1, \dots, \lambda'_d + M - 1\} \subseteq \{\lambda'_d + 1, \dots, \mu'_d - 1\}$ and $\mu'_d - \iota \in \{\mu'_d - (M-1), \dots, \mu'_d - 1\} \subseteq \{\lambda'_d + 1, \dots, \mu'_d - 1\}$. The third equality is the definition of $\text{FACE}(\boldsymbol{\lambda}', \boldsymbol{\mu}', \boldsymbol{\delta})$. \square

References

- Besbes, Omar, Ilan Lobel. 2015. Intertemporal price discrimination: Structure and computation of optimal policies. *Management Science* **61**(1) 92–110.
- Elmachtoub, Adam, Vishal Gupta, Michael Hamilton. 2021. The value of personalized pricing. *Management Science* **67**(10) 6055–6070.
- Pflug, Georg, Alois Pichler. 2016. From empirical observations to tree models for stochastic optimization: convergence properties. *SIAM Journal on Optimization* **26**(3) 1715–1740.
- Shafieezadeh-Abadeh, Soroosh, Daniel Kuhn, Peyman Mohajerin Esfahani. 2019. Regularization via mass transportation. *Journal of Machine Learning Research* **20**(103) 1–68.
- Shapiro, Alexander, Darinka Dentcheva, Andrzej Ruszczyński. 2009. *Lectures on stochastic programming: modeling and theory*. SIAM.
- Sturt, Bradley. 2023. A nonparametric algorithm for optimal stopping based on robust optimization. *Operations Research* **71**(5) 1530–1557.
- Tsirsis, Stratis, Behzad Tabibian, Moein Khajehnejad, Adish Singla, Bernhard Schölkopf, Manuel Gomez-Rodriguez. 2024. Optimal decision making under strategic behavior. *Management Science*.
- Xie, Haitian, Ying Zhu, Denis Shishkin. 2024. On the limitations of data-based price discrimination. *Theoretical Economics* (forthcoming).
- Xu, Huan, Constantine Caramanis, Shie Mannor. 2012. A distributional interpretation of robust optimization. *Mathematics of Operations Research* **37**(1) 95–110.

Appendix A: Necessity of Assumption 3

We show in this section that Assumption 3 is necessary for asymptotic consistency, in the sense that relaxing the absolute continuity requirement can invalidate Theorem 1.

EXAMPLE EC.1. Let $D = 2$ and $\epsilon \in [0, 1/4]$. Let $\mathbf{X} \in [0, 1]^2$ be a random vector drawn uniformly from the perimeter of the circle centered at $(1/2, 1/2)$ with radius $1/4$. Let $\mathbf{L} = \mathbf{X} - \mathbf{1}\epsilon$ and $\mathbf{U} = \mathbf{X} + \mathbf{1}\epsilon$. If $\epsilon > 0$, then we observe that $\mathbb{P}(L_d < U_d) = 1$ for all $d \in \{1, 2\}$. However, we observe for all $\epsilon \in [0, 1/4]$ that (\mathbf{L}, \mathbf{U}) is not drawn from an probability distribution that is absolutely continuous with respect to the Lebesgue measure on \mathbb{R}^{2D} ; indeed, Figure EC.1a shows that $\mathbb{P}((\mathbf{L}, \mathbf{U}) \in \mathcal{Z}_L \times \mathcal{Z}_U) = 1$ for $\mathcal{Z}_L \triangleq \{(1/4 \cos(\theta) + 1/2 - \epsilon, 1/4 \sin(\theta) + 1/2 - \epsilon) : \theta \in [0, 2\pi)\}$ and $\mathcal{Z}_U \triangleq \{(1/4 \cos(\theta) + 1/2 + \epsilon, 1/4 \sin(\theta) + 1/2 + \epsilon) : \theta \in [0, 2\pi)\}$, and we observe from Figure EC.1a that the Lebesgue measure with respect to \mathbb{R}^{2D} of $\mathcal{Z}_L \times \mathcal{Z}_U \subset \mathbb{R}^{2D}$ is equal to zero. Hence, we observe for every $\epsilon \in [0, 1/4]$ that Assumption 3 is violated.

To show that the asymptotic consistency guarantee from Theorem 1 is invalidated for this example, let the set of feasible prices be $\mathcal{P} = \{1/3, 1/2\}$, and let V be a random variable that is independent of (\mathbf{L}, \mathbf{U}) and satisfies $\mathbb{P}(V = 1/3) = \mathbb{P}(V = 1/2) = 1/2$. Since V is independent of (\mathbf{L}, \mathbf{U}) , it is easy to see that there exists an optimal solution for (OPT) that satisfies $\pi^*(\mathbf{x}) = 1/3$ for all $\mathbf{x} \in [0, 1]^2$ and that the objective value of (OPT) is $\nu^* = 1/3$.⁴ However, Figure EC.1b illustrates that there almost surely exists a pricing policy $\pi : [0, 1]^2 \rightarrow \{1/3, 1/2\}$ that satisfies the equality $R^\pi(\mathbf{L}^i, \mathbf{U}^i, 1/3) = 1/3$ for all $i \in \{1, \dots, N\}$ such that $V^i = 1/3$ and satisfies the equality $R^\pi(\mathbf{L}^i, \mathbf{U}^i, 1/2) = 1/2$ for all $i \in \{1, \dots, N\}$ such that $V^i = 1/2$. Hence, we conclude from the strong law of large numbers that $\lim_{N \rightarrow \infty} \widehat{\nu}_N = (1/2)(1/3) + (1/2)(1/2) = 5/12 \neq \nu^*$ almost surely.

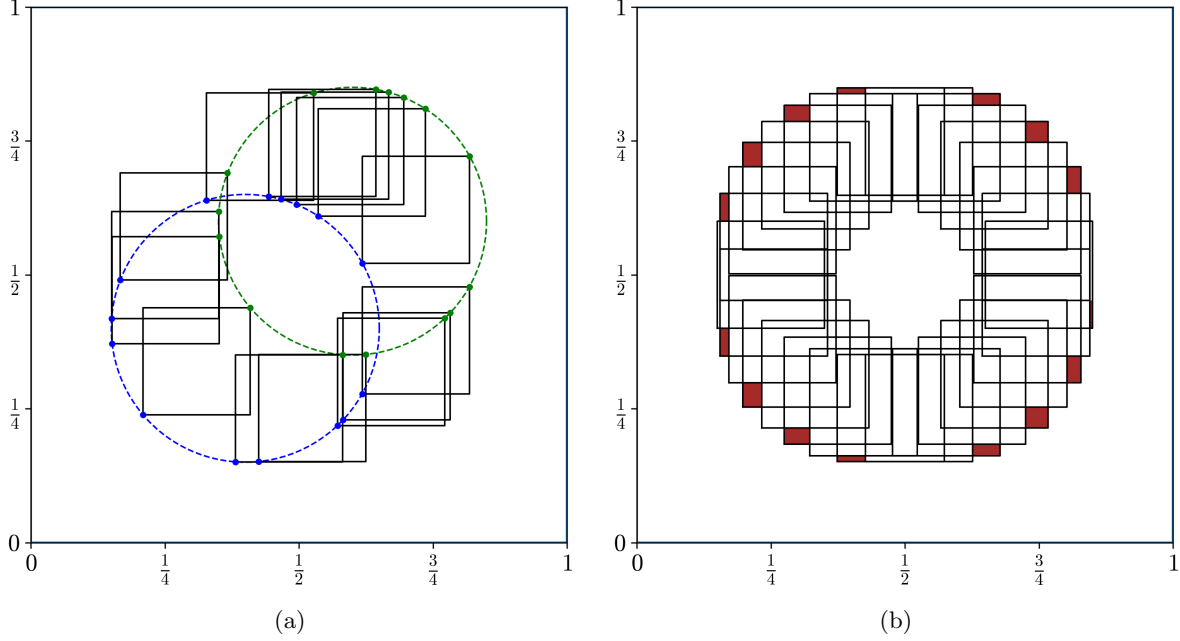
Appendix B: Experiment Details

Our numerical experiment in §4 involves solving (SAA). In this appendix, we provide the mixed-integer linear program reformulation of (SAA) under Assumption 4 that is used in the numerical experiment.

First, let $\mathcal{U}^i = [\mathbf{L}^i, \mathbf{U}^i]$ be the hyperrectangle of buyer i 's features. We then define the following collection of subsets of buyers:

$$\mathcal{I} \triangleq \left\{ \mathcal{I} \in 2^{\{1, \dots, N\}} : \bigcap_{i \in \mathcal{I}} \mathcal{U}^i \setminus \bigcup_{i \in \{1, \dots, N\} \setminus \mathcal{I}} \mathcal{U}^i \neq \emptyset \right\}.$$

⁴ The fact that V is independent of (\mathbf{L}, \mathbf{U}) implies that either $\pi_1(\cdot) = \frac{1}{3}$ or $\pi_2(\cdot) = \frac{1}{2}$ is an optimal solution for (OPT). We observe that $J^*(\pi_1) = \frac{1}{2} \mathbb{E} [R^{\pi_1}(\mathbf{L}, \mathbf{U}, \frac{1}{3})] + \frac{1}{2} \mathbb{E} [R^{\pi_1}(\mathbf{L}, \mathbf{U}, \frac{1}{2})] = \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{3}$ and $J^*(\pi_2) = \frac{1}{2} \mathbb{E} [R^{\pi_2}(\mathbf{L}, \mathbf{U}, \frac{1}{3})] + \frac{1}{2} \mathbb{E} [R^{\pi_2}(\mathbf{L}, \mathbf{U}, \frac{1}{2})] = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$. Thus, we conclude that an optimal solution for (OPT) is $\pi^* = \pi_1$ and the objective value of (OPT) is $\nu^* = \frac{1}{3}$.

Figure EC.1 Visualization of Example EC.1

Note. Both figures show case where $\epsilon = 0.1$. (a) $N = 15$. Blue dots show $\mathbf{L}^1, \dots, \mathbf{L}^{15}$ and green dots show $\mathbf{U}^1, \dots, \mathbf{U}^{15}$. Blue and green dashed lines show the sets \mathcal{Z}_L and \mathcal{Z}_U , respectively. (b) $N = 30$. The figure shows a pricing policy $\pi : [0, 1]^2 \rightarrow \{1/3, 1/2\}$ in which the low price $\pi(\mathbf{x}) = 1/3$ is offered if and only if $\mathbf{x} \in [0, 1]^2$ is in a red shaded region. We observe that this pricing policy offers the low price only in the outermost corners of alternating squares, and that the low price is not offered in any of the other squares. Following this strategy, we conclude for a given sample $(\mathbf{L}^1, \mathbf{U}^1, V^1), \dots, (\mathbf{L}^N, \mathbf{U}^N, V^N)$ that we construct a pricing policy that satisfies the equality $\min_{\mathbf{x} \in [\mathbf{L}^i, \mathbf{U}^i]} \pi(\mathbf{x}) = 1/3$ for all $i \in \{1, \dots, N\}$ such that $V^i = 1/3$ and satisfies the equality $\min_{\mathbf{x} \in [\mathbf{L}^i, \mathbf{U}^i]} \pi(\mathbf{x}) = 1/2$ for all $i \in \{1, \dots, N\}$ such that $V^i = 1/2$. Note for this figure that $\mathbf{X}^1, \dots, \mathbf{X}^{30}$ are equally spaced around the perimeter of the circle, rather than drawn uniformly over the perimeter, for the sake of visual clarity.

For each $\mathcal{I} \in \mathcal{J}$, let $\mathcal{W}_{\mathcal{I}} \triangleq \bigcap_{i \in \mathcal{I}} \mathcal{U}^i \setminus \bigcup_{i \in \{1, \dots, N\} \setminus \mathcal{I}} \mathcal{U}^i$ denote the subset of features corresponding to \mathcal{I} . Note that $\{\mathcal{W}_{\mathcal{I}} : \mathcal{I} \in \mathcal{J}\}$ is a collection of disjoint sets whose union is equal to $\bigcup_{i \in \{1, \dots, N\}} \mathcal{U}^i$.

Next, for each buyer i , let \mathcal{J}^i denote the subset of \mathcal{J} such that $\mathcal{U}^i = \bigcup_{\mathcal{I} \in \mathcal{J}^i} \mathcal{W}_{\mathcal{I}}$. We observe that

$$\begin{aligned}
\sup_{\pi: \mathcal{X} \rightarrow \mathcal{P}} \frac{1}{N} \sum_{i=1}^N \mathcal{R}^{\pi}(\mathbf{L}^i, \mathbf{U}^i, V^i) &= \sup_{\pi: \mathcal{X} \rightarrow \mathcal{P}} \frac{1}{N} \sum_{i=1}^N \mathbb{I} \left\{ \inf_{\mathbf{x} \in [\mathbf{L}^i, \mathbf{U}^i]} \pi(\mathbf{x}) \leq V^i \right\} \cdot \inf_{\mathbf{x} \in [\mathbf{L}^i, \mathbf{U}^i]} \pi(\mathbf{x}) \\
&= \sup_{\pi: \mathcal{X} \rightarrow \mathcal{P}} \frac{1}{N} \sum_{i=1}^N \mathbb{I} \left\{ \min_{\mathcal{I} \in \mathcal{J}^i} \inf_{\mathbf{x} \in \mathcal{W}_{\mathcal{I}}} \pi(\mathbf{x}) \leq V^i \right\} \cdot \min_{\mathcal{I} \in \mathcal{J}^i} \inf_{\mathbf{x} \in \mathcal{W}_{\mathcal{I}}} \pi(\mathbf{x}) \\
&= \sup_{\tilde{\pi}: \mathcal{J} \rightarrow \mathcal{P}} \frac{1}{N} \sum_{i=1}^N \mathbb{I} \left\{ \min_{\mathcal{I} \in \mathcal{J}^i} \tilde{\pi}(\mathcal{I}) \leq V^i \right\} \cdot \min_{\mathcal{I} \in \mathcal{J}^i} \tilde{\pi}(\mathcal{I}),
\end{aligned}$$

where the first equality follows from the definition of revenue function, the second equality is due to $\mathcal{U}^i := [\mathbf{L}^i, \mathbf{U}^i] = \bigcup_{\mathcal{I} \in \mathcal{J}^i} \mathcal{W}_{\mathcal{I}}$, and the last equality is because for any pricing policy $\pi(\cdot)$, we can

define an equivalent policy based on set \mathfrak{I} as $\tilde{\pi}(\mathcal{I}) := \inf_{\mathbf{x} \in \mathcal{W}_{\mathcal{I}}} \pi(\mathbf{x})$. Hence, we have the following mixed-integer linear program reformulation:

$$\begin{aligned}
& \underset{\mathbf{b}, \boldsymbol{\lambda}}{\text{maximize}} && \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K p_k b_{ik} \\
& \text{subject to} && b_{ik} \leq \mathbb{I}\{p_k \leq V^i\} && \forall i = 1, \dots, N, k = 1, \dots, K \\
& && b_{ik} \leq \sum_{\kappa=1, \dots, K: \kappa \geq k} \lambda_{\mathcal{I}\kappa} && \forall i = 1, \dots, N, k = 1, \dots, K, \mathcal{I} \in \mathfrak{I}: i \in \mathfrak{N}_{\mathcal{I}} \\
& && b_{ik} \leq \sum_{\mathcal{I} \in \mathfrak{I}: i \in \mathfrak{N}_{\mathcal{I}}} \lambda_{\mathcal{I}k} && \forall i = 1, \dots, N, k = 1, \dots, K \\
& && \sum_{k=1}^K \lambda_{\mathcal{I}k} = 1 && \forall \mathcal{I} \in \mathfrak{I} \\
& && \sum_{k=1}^K b_{ik} \leq 1 && \forall i = 1, \dots, N \\
& && \mathbf{b} \in \{0, 1\}^{N \times K}, \boldsymbol{\lambda} \in \{0, 1\}^{|\mathfrak{I}| \times K}.
\end{aligned}$$

Here, for each $\mathcal{I} \in \mathfrak{I}$, $\mathfrak{N}_{\mathcal{I}} = \{i \in \{1, \dots, N\} : \mathcal{W}_{\mathcal{I}} \subseteq \mathcal{U}^i\}$ denotes the largest subset of buyers whose hyperrectangle of features contains $\mathcal{W}_{\mathcal{I}}$, $\lambda_{\mathcal{I}k} = 1$ if and only if a region is offering the price p_k , while $b_{ik} = 1$ if and only if

$$k \in \arg \max_{\kappa=1, \dots, K} \left\{ \kappa : \min_{\mathcal{I} \in \mathfrak{I}: i \in \mathfrak{N}_{\mathcal{I}}} p_{\kappa} \lambda_{\mathcal{I}\kappa} \leq V^i \right\} = \arg \max_{\kappa=1, \dots, K} \left\{ \kappa : p_{\kappa} \cdot \min_{\mathcal{I} \in \mathfrak{I}: i \in \mathfrak{N}_{\mathcal{I}}} \lambda_{\mathcal{I}\kappa} \leq V^i \right\}.$$