Model-Free Assortment Pricing with Transaction Data

Ningyuan Chen
Rotman School of Management, University of Toronto, Toronto, ON, Canada, ningyuan.chen@utoronto.ca

Andre A. Cire
Rotman School of Management, University of Toronto, Toronto, ON, Canada, andre.cire@rotman.utoronto.ca

Ming Hu
Rotman School of Management, University of Toronto, Toronto, ON, Canada, ming.hu@rotman.utoronto.ca

Saman Lagzi
Rotman School of Management, University of Toronto, Toronto, ON, Canada, saman.lagzi@rotman.utoronto.ca

We study a problem in which a firm sets prices for products based on the transaction data, i.e., which product past customers chose from an assortment and what were the historical prices that they observed. Our approach does not impose a model on the distribution of the customers' valuations and only assumes, instead, that purchase choices satisfy incentive-compatible constraints. The individual valuation of each past customer can then be encoded as a polyhedral set, and our approach maximizes the worst-case revenue assuming that new customers' valuations are drawn from the empirical distribution implied by the collection of such polyhedra. We show that the optimal prices in this setting can be approximated at any arbitrary precision by solving a compact mixed-integer linear program. Moreover, we study special practical cases where the program can be solved efficiently, and design three approximation strategies that are of low computational complexity and interpretable. Comprehensive numerical studies based on synthetic and real data suggest that our pricing approach is uniquely beneficial when the historical data has a limited size or is susceptible to model misspecification.

Key words: data-driven; incentive-compatible; robust optimization; product pricing; small data

1. Introduction

Online retailing has seen steady growth in the last decade. According to the survey by Gaubys (2020), the global market share of e-commerce is expected to surpass 20% in 2022, and the trend is only accelerating. Retailers, however, face various challenges when transitioning to an online business model. First, the granularity of the data gathered from past customers, where firms observe what customers bought, the products they viewed, and their prices at the time of the purchase, far exceeds that of the offline setting. Second, the firm is able to roll out products rapidly and many new products (or new configurations of old products) can be displayed each day (Caro et al. 2020). Seemingly contradictory to the first point, there may be little purchase information for a firm to take timely actions on these new products, such as price correction and adjustments.

In an attempt to address this challenge, we study the pricing decision of an assortment of products for a firm based on its historical transaction data. The data records the assortment of products

viewed by past customers and their prices, which may vary across customers due to promotions. The data also records the decision made by the customers: the purchase of one of the products or leaving with no purchase. We are interested in the regime that the number of customers is not large relative to the number of available products, in order to accommodate the realistic settings mentioned above. The unique challenge arising from this regime, as noted by Abdallah and Vulcano (2020), is that the estimation of any discrete choice model from the "small" data is going to be noisy, even if it is correctly specified.

The common approach to handle this situation is what we denote by "model-estimate-optimize." The firm first builds a discrete choice model to characterize how customers form their utilities and make choices. For example, in the multinomial logit (MNL) model, the probability of a customer choosing product j from a set of n products priced at $(p_1, \ldots, p_n)$ can be expressed as

$$\frac{\exp(\alpha_j - \beta p_j)}{1 + \sum_{k=1}^{n} \exp(\alpha_k - \beta p_k)}$$

for some parameters $\{\alpha_k\}_{k=1}^n$ and β representing the average attractiveness of the products and the price sensitivity, respectively. It is equivalent to a random utility model in which a random customer has utility $\alpha_k - \beta p_k + \epsilon_k$ for product k, where ϵ_k follows an independent and identically distributed (i.i.d.) Gumbel distribution and customers choose the product, or no purchase, with the highest utility. The second step is to "estimate." Based on the historical data, the firm can estimate the parameters (see Train 2009 for the detailed steps). Finally, the firm "optimizes" its pricing decisions by setting the optimal prices for future customers using the estimated model. This approach has enjoyed wide popularity because of its simplicity and computational efficiency.

However, in the setting we consider, the model-estimate-optimize approach falls short in two aspects. First, because of the limited data for certain products, the estimation related to their parameters (such as α_j in the MNL model) may be noisy and unstable. It is not clear whether the firm can rely on the estimation to make adequate pricing decisions. Second, the discrete choice model may not be able to capture the behavioral pattern in the data, which is referred to as model misspecification. In this case, the firm can choose a more complex model, potentially with more parameters, to minimize the misspecification error. As pointed out by Abdallah and Vulcano (2020), this will inevitably exacerbates the first issue.

In this paper, we propose a data-driven approach to the assortment pricing problem. As opposed to the model-estimate-optimize approach, we consider no prior models on how customers form their utilities for the products. Instead, when a customer is observed to choose a product from a set of products with given prices, we assume that a (model-free) incentive-compatibility condition specifies her potential valuations for the products. For example, when a product is purchased by

a customer, the utility of that product must be necessarily at least as high as any of the non-purchased products. Leveraging this condition, the data for a transaction defines a polyhedral set containing the vector of valuations $(v_1, \ldots, v_n)$ of the particular customer, which is referred to as the IC polyhedron. When a new customer arrives, without knowing anything about her preferences, we uniformly sample from the IC polyhedra from customers in the historical data, and set prices to maximize the revenue for valuations in the sampled polyhedron. More precisely, we set prices such that the revenue derived from the arriving customer is maximized, when her valuations, drawn from the sampled IC polyhedron, lead to the least possible revenue for the firm.

The contribution of this work is threefold. First, to the best of our knowledge, this is the first study that leverages incentive-compatibility constraints to define a valuation polyhedron for the purposes of pricing. With this novel approach, we circumvent the need for pre-specified customer models as is customary in the random utility literature, investigating the optimal pricing directly based on the data. We believe this approach sheds new insights into the formulation of other related data-driven problems.

Second, by exploiting the structure of the polyhedra, we present a disjunctive model of the pricing problem and several structural results associated with the optimal prices. We show that the disjunctive model can be approximated at any desired precision by a compact bilinear program, itself solvable by a compact mixed-integer integer programming model after an appropriate reformulation. For scalability purposes, we also present low-complexity, interpretable approximation algorithms that are shown to achieve strong theoretical and numerical performance. We also study special cases of practical interest where the optimal prices can be obtained efficiently.

Third, we conduct a comprehensive numerical study based on synthetic and real data. In particular, we use the well-studied IRI dataset (Bronnenberg et al. 2008) and fit different choice models including the linear, MNL and mixed logit models. We then generate a small number of customer purchases based on the models. After applying our approach to the data, the generated revenues significantly outperform the incumbent prices in the data, for all the models, demonstrating the benefit of the data-driven model-free approach: the smaller sensitivity to model misspecification and promising performance when the sample size is limited.

Our findings in the numerical study suggest that (1) our best approximation algorithm consistently recovers at least 96% of the optimal revenue in various settings and scales linearly with the number of samples. (2) When the historical data is generated from commonly used discrete choice models such as the MNL model, our data-driven approach performs well, compared to the optimal prices under the correctly specified estimated model, especially for a limited data size. On the other hand, when the discrete choice model is misspecified, our approach is more robust and outperforms the misspecified optimal prices. (3) Applying our data-driven approach to real datasets leads to increase in revenues over the incumbent prices in the dataset.

2. Related Work

Our study is broadly related to three streams of literature. The first steam is the papers studying the estimation and optimization of the discrete choice models with prices, which provide the basis for the model-estimate-optimize approach. For the estimation of popular discrete choice models including the MNL model, Train (2009) provides an excellent review. Recently, two new choice models have drawn attention of scholars in the Operations Research community, the rank-based model (Farias et al. 2013) and the Markov chain choice model (Blanchet et al. 2016). Due to their flexibility, the estimation is not as straightforward as others such as the MNL model. Several studies propose various algorithms to address the issue (van Ryzin and Vulcano 2015, 2017, Şimşek and Topaloglu 2018). As for optimal pricing, the MNL model is studied in, e.g., Hopp and Xu (2005) and Dong et al. (2009). For the nested logit model, which is a generalization of the MNL model, Li and Huh (2011), Gallego and Wang (2014), Li et al. (2015) investigate its optimal pricing problem. Zhang et al. (2018) study the optimal pricing problem of the generalized extreme value random utility models with the same price sensitivity. In contrast, Mai and Jaillet (2019) use a robust framework to tackle the same problem for extreme value utilities.

Although we also adopt a robust approach, we do not rely on a particular distribution and the decision is fully guided by the data. Similarly, Rusmevichientong and Topaloglu (2012) and Jin et al. (2020) study the robust optimization problem of pricing or assortment planning in the MNL model. Several papers incorporate pricing into the rank-based model (Rusmevichientong et al. 2006, Jagabathula and Rusmevichientong 2017) and the Markov chain model (Dong et al. 2019) and study the optimal pricing problem. More recently, Yan et al. (2020) use the transaction data to fit a general discrete choice model of a representative consumer and solve the optimal pricing based on the fitted model by mixed-integer linear programming. While the problem they study is essentially similar to ours, their approach relies on the model-estimate-optimize paradigm and hence is conceptually distinct from the proposed approach here.

In the context of revenue management, the definition of the term "data-driven" typically depends on the problem context. When the agent makes decisions in the process of data collection, the data-driven approach is usually associated with a framework that integrates such a process with decision making, so that the agent is learning the unknown parameters or environment while maximizing revenues[1]. The previous works by Bertsimas and Vayanos (2017), Zhang et al. (2020), Cohen et al. (2018, 2020), Ettl et al. (2020), Ban and Keskin (2020) fall into this category. It is connected to a large body of literature on demand learning and dynamic pricing. We refer to den Boer (2015) for a comprehensive review of earlier papers in this setting.

[1] This is referred to as the online problem, which should not be confused with the notion of online retailing.

In contrast, our study essentially handles an offline setting, in which the data has been collected and given. Two recent papers also provide alternatives to the model-estimate-optimization approach. Bertsimas and Kallus (2020) leverage statistical methods such as k-nearest neighbors and kernel smoothing to integrate past observations into the current decision making problem given a covariate. It circumvents the step of estimating a statistical model. Elmachtoub and Grigas (2020) achieve a similar goal by skipping the minimization of the estimation error and directly focusing on the decision error. They design a new loss function that combines the errors in both stages: estimation and optimization. Both papers study a setting where the optimization is conditioned on a covariate. Our problem is less general than their formulation and does not have a covariate, which allows us to utilize the special structure of the multi-product pricing problem (i.e., the incentive-compatibility of customers) that does not apply to their general approach. Ban and Rudin (2019) propose an algorithm integrating historical demand data and newsvendor optimal order quantity, without estimating the demand model separately.

There are a few papers using model-free approaches in revenue management. Allouah and Besbes (2019) study the problem that the seller observes a sample from the buyer's evaluation but is agnostic to the underlying distribution. The optimal price is solved for a family of possible distributions in the maximin sense. Chen et al. (2019) adopt the estimate-then-model approach, which is model-free, to estimate choice models using random forests. They do not consider the optimization stage. Our problem is similar to the work by Ferreira et al. (2016), who study the demand forecasting and price optimization of a new product by a retailer. However, it focuses on a single product and it is unclear how to adapt it to multiple products. A few studies apply analytics to promotion planning (Cohen et al. 2017, Cohen and Perakis 2020). They consider more practice-based assumptions than ours and use various approximations for the demand model.

This paper uses the incentive-compatibility of past customers as a building block, and hence is related to the literature on auction designs, especially those papers using a data-driven or robust formulation. Bandi and Bertsimas (2014) study the multi-item auction design with budget constrained buyers and use a robust formulation for the set of valuations. The uncertainty set is constructed using the historical information, such as means and covariance matrices. Because of the polyhedral structure of the uncertainty sets, the robust optimization problem is tractable. Our formulation investigates the worst-case revenue in the IC polyhedron (see Section 3), which is similar to the idea of robust optimization for valuations drawn from uncertainty sets. However, there are three distinctions. First, in Bandi and Bertsimas (2014), the uncertainty set is constructed on the valuations of all bidders for a single product, while in our problem the valuations of a single customer for all the products fall into a polyhedron. This is because of the different applications. Second, our data-driven approach averages over the empirical distribution of historical customers,

while the historical information is used to construct the uncertainty set in Bandi and Bertsimas (2014). Third, the optimization problem in our problem cannot be solved efficiently and we resort to approximation algorithms. Derakhshan et al. (2019) consider a data-driven optimization framework to find the optimal personalized reservation price of buyers, when past bids are input to the algorithm. They do not impose assumptions on the valuation distributions and maximize the revenue when the valuation of the future customer is drawn from the empirical distribution of the historical data. This is similar to the motivation of this paper. However, because of the different contexts, the reduction and approximations have little in common with ours. For example, the LP relaxation has a 0.684-approximation ratio in Derakhshan et al. (2019) but performs poorly in our setting (see Section 5). Allouah and Besbes (2020) study the single-item auction for a group of buyers when the seller does not have access the valuation distribution of the buyers. Instead, the auction is designed for a general class of distributions, which is conceptually similar to our model-free approach.

3. Problem Description

We consider a firm that has observed the transaction data of m customers $\mathscr{C} = \{1, \ldots, m\}$ with respect to an assortment of n products $\mathscr{P} = \{1, \ldots, n\}$. Specifically, the firm's historical data includes the product prices $\mathbf{P}_i = (P_{i1}, \ldots, P_{in}) > \mathbf{0}$ viewed by each customer $i \in \mathscr{C}$, as well as the product $c_i \in \mathscr{P}$ that the customer chose from that assortment. We consider that customers viewed all products $\mathscr{P}$ and purchased one product from that assortment; both assumptions are made without loss of generality (Remarks 1 and 2 below). That is, we can incorporate the situation where historical customers may have not observed some of the products in the assortment during their transactions, and removing those customers who made no purchase during their shopping session does not change our results. The goal of the firm is to set the product prices $\mathbf{p} = (p_1, \ldots, p_n)$ for newly arriving customers that leverages this offline historical data.

In this study, we investigate a pricing approach that operates on the full set of possible customer utilities under incentive-compatibility constraints. More precisely, let $v_{ij} \geq 0$ be the *unknown* valuation that customer $i \in \mathscr{C}$ assigns to each product $j \in \mathscr{P}$. We assume that the utility of purchasing j is given by $v_{ij} - P_{ij}$ (and hence quasilinear in v_{ij}), and that such utilities must be compatible with observations from the data; i.e., the set of all possible valuations of customer i is

$$\begin{aligned} \mathscr{V}_i \equiv \big\{ (v_{i1}, \ldots, v_{in}) \in \mathbb{R}^n_+ : \; & v_{ic_i} - P_{ic_i} \geq 0, \\ & v_{ic_i} - P_{ic_i} \geq v_{ij'} - P_{ij'}, \; \forall j' \in \mathscr{P} \setminus \{c_i\} \big\}. \end{aligned} \tag{1}$$

The first inequality in (1) indicates that the utility of purchasing product c_i is non-negative. The second inequality specifies that valuations are incentive-compatible, i.e., that the utility of

purchasing product c_i must be either the same or larger than the utility for the remaining products $\mathscr{P} \setminus \{c_i\}$. We note that $\mathscr{V}_i$ is defined by a finite set of closed halfspaces and hence is a polyhedral set. We refer to $\mathscr{V}_i$ as the *incentive-compatible* (*IC*) *polyhedron* of customer i.

For a newly arriving customer, however, her valuation is not known to the firm. Based on historical data, it is reasonable to use the empirical distribution to form such an estimate. That is, we assume that her valuation is equally likely to fall into one of the IC polyhedra $\mathscr{V}_1, \mathscr{V}_2, \ldots, \mathscr{V}_m$. Given the new prices $\mathbf{p}$ and considering that the new customer's valuation falls into $\mathscr{V}_i$ for some i, we apply a robust approach where the arriving customer picks product $j \in \mathscr{P}$ that yields the lowest possible revenue and such a choice is consistent with $\mathscr{V}_i$ under prices $\mathbf{p}$. Such a revenue is described by the program

$$f_i(\mathbf{p}) \equiv \min_{\mathbf{v}_i \in \mathscr{V}_i, r \geq 0} \quad r \tag{DP}$$

$$\text{s.t.} \quad \bigvee_{j \in \mathscr{P}} \begin{pmatrix} r = p_j \\ v_{ij} - p_j \geq 0 \\ v_{ij} - p_j \geq v_{ij'} - p_{j'}, \ \forall j' \in \mathscr{P} \end{pmatrix} \vee \begin{pmatrix} r = 0 \\ v_{ij} - p_j \leq 0, \ \forall j \in \mathscr{P} \end{pmatrix}. \tag{2}$$

Model (DP) is a classical disjunctive program (Balas 1998), i.e., the set of feasible solutions is defined by a disjunction of polyhedra representing the feasible valuations for each product. Specifically, with the objective $\min r$ we select the product with the lowest revenue for which there exists a feasible valuation $\mathbf{v}_i \in \mathscr{V}_i$ that is also incentive-compatible under the new prices $\mathbf{p}$. This is modeled by the left-hand side disjunction of (2), where some product j (and hence price p_j) is selected only if its net utility (i.e., $v_{ij} - p_j$) is non-negative and at least as large as that of choosing other products. The right-hand side disjunction of (2) formulates the case where the customer chooses no product, only possible if the utility of choosing any product under $\mathbf{p}$ is non-positive.

Remark 1. As noted earlier, we assume that each historical customer $i \in \mathscr{C}$ has purchased one product from the assortment. In particular, if some customer i has not purchased any products, we can assume that i belongs to the polyhedron with zero revenue in model (DP), i.e., the right-hand side term in disjunction (2). Thus, revenue driven by customer i is always zero for any prices, and therefore she can be removed from the model. □

Based on (DP), we assign a weight $1/m$ and sum over $i \in \mathscr{C}$ to take the expectation with respect to the empirical distribution. To maximize over $\mathbf{p}$, the objective function is expressed as

$$\tau^* \equiv \sup_{\mathbf{p} \geq 0} \frac{1}{m} \sum_{i \in \mathscr{C}} f_i(\mathbf{p}). \tag{OP}$$

The problem (OP) highlights important distinctions from existing pricing approaches. First, it does not rely on a parametric discrete-choice model that explicitly specifies the distribution of customer utilities (see, e.g., Train 2009). Instead, we adopt a model-free approach and consider

the worst-case valuation for a given $\mathbf{p}$ under quasilinearity and incentive-compatible customer preferences, which are arguably weaker and more justifiable than existing parametric models to the best of our knowledge. Second, we do not attempt to estimate a non-parametric model (e.g., Jagabathula and Rusmevichientong 2017). Instead, the goal is to investigate the structure of (OP), trade-offs, and scenarios where prices derived from (OP) are beneficial in comparison to existing models, given its data-driven nature and emphasis on the historical data.

Remark 2. The choice model (DP) assumes that each customer $i \in \mathscr{C}$ has seen the same complete assortment $\mathscr{P}$ and their historical prices $\mathbf{P}_i$. While in practice customers may have seen different assortments, which can be subsets of $\mathscr{P}$, this assumption can be relaxed by setting a sufficiently large P_{ij} if customer i has not been offered product j (e.g., the sum of all historical prices). In that case, the structural results below can be rewritten accordingly without loss of generality. □

Remark 3. We note that the maximum of (OP) may not be attainable because of the discontinuity of the customer choices in (2), hence the use of supremum in the objective. For example, consider an instance with one customer ($m = 1$) and one product ($n = 1$), where $P_{11} = P^*$ for some $P^* > 0$ and $c_1 = 1$. Thus, from (1), the customer valuation satisfies $v_{11} \geq P^*$. Suppose now we assign a price $p_1 \geq P^*$ to the product. It follows from (2) that the optimal revenue is zero, since any non-negative valuation $P^* \leq v_{11} \leq p_1$ is incentive-compatible with the no-purchase option. However, for any $p_1 = P^* - \epsilon$ with $\epsilon > 0$, we have $\tau^* \to P^*$ as $\epsilon \to 0$. Thus, $f_i(\mathbf{p})$ is discontinuous in $\mathbf{p}$. □

4. Reformulations of the Pricing Model (OP)

In this section, we investigate linear reformulations of (OP) that serve as the basis of our exact and tractable approximation strategies. We start in §4.1 with an analysis of the optimal policy structure for the problem (DP), presenting a small linear programming model to compute $f_i(\mathbf{p})$ that is compact with respect to the number of products and customers. We leverage this model in §4.2 to derive an alternative, compact mixed-integer linear program that approximates τ^* at any desired absolute error with respect to the optimal solution of (OP). Finally, we use this reformulation in §4.3 to show the optimal price structure of two commonly seen special cases.

4.1. Optimal Revenue from an IC Polyhedron

We first characterize the IC polyhedron $\mathscr{V}_i$ and $f_i(\mathbf{p})$, for given $i \in \mathscr{C}$ and $\mathbf{p}$. It can be interpreted as how customer i behaves under the new price $\mathbf{p}$ based on her historical choice. Proposition 1 uses classical polyhedral results to transform (DP) into an equivalent linear program.

PROPOSITION 1. *The formulation (DP) is equivalent to the linear program*

$$f_i(\mathbf{p}) = \min_{\mathbf{v}_i^1,\ldots,\mathbf{v}_i^n,\mathbf{v}_i^\varnothing,\mathbf{x}\geq 0} \quad \sum_{j\in\mathscr{P}} p_j x_j \tag{DP-LP}$$

$$\text{s.t.} \quad v_{ij}^j - p_j x_j \geq 0, \qquad \forall j \in \mathscr{P}, \tag{3}$$

$$v_{ij}^j - p_j x_j \geq v_{ij'}^j - p_{j'} x_j, \qquad \forall j, j' \in \mathscr{P}, \tag{4}$$

$$v_{ij}^\varnothing \leq p_j x_\varnothing, \qquad \forall j \in \mathscr{P}, \tag{5}$$

$$\sum_{j\in\mathscr{P}} x_j + x_\varnothing = 1, \tag{6}$$

$$v_{ic_i}^j - P_{ic_i} x_j \geq 0, \qquad \forall j \in \mathscr{P}\cup\{\varnothing\}, \tag{7}$$

$$v_{ic_i}^j - P_{ic_i} x_j \geq v_{ij'}^j - P_{ij'} x_j, \qquad \forall j, j' \in \mathscr{P}\cup\{\varnothing\}, \tag{8}$$

where $\mathbf{v}_i^1,\ldots,\mathbf{v}_i^n,\mathbf{v}_i^\varnothing$ *and* $\mathbf{x} = (x_1,\ldots,x_n,x_\varnothing)$ *are* n*-dimensional and* $(n+1)$*-dimensional variable tuples, respectively.*

The model (DP-LP) is a linear program with $\mathcal{O}(n^2)$ variables and constraints. In particular, the variables $x_1,\ldots,x_n,x_\varnothing$ represent which product is picked by an arriving customer who draws the IC polyhedron $\mathscr{V}_i$, where $x_\varnothing$ encodes the no-purchase option. Constraint (6) ensures that the customer selects either one product or the no-purchase option. Constraints (3)-(4) imply that $\mathbf{v}_i^j$ is incentive-compatible with the choice x_j. They are equivalent to the j-th disjunctive term of (2), while the other inequalities $j' \neq j$ of the same family become redundant whenever $x_j = 0$. This same reason applies analogously to the no-purchase option and inequality (5). Constraints (7)-(8) ensure that $\mathbf{v}_i^j \in \mathscr{V}_i$, i.e., the valuation must be incentive-compatible with the historically chosen product c_i. The generated revenue $\sum_{j\in\mathscr{P}} p_j x_j$ is minimized over $\mathbf{v}$ and $\mathbf{x}$.

We now analyze the structure of (DP-LP) to draw insights into the optimal customer choices and reduce the size of the formulation. To this end, consider the set of valuations from (DP-LP) that are incentive-compatible with product $j \in \mathscr{P}$, i.e.,

$$\mathscr{W}_i^j(\mathbf{p}) \equiv \left\{\mathbf{v}_i^j \in \mathscr{V}_i \colon v_{ij}^j - p_j \geq 0,\ v_{ij}^j - p_j \geq v_{ij'}^j - p_{j'},\ \forall j' \in \mathscr{P}\right\} \tag{9}$$

and those incentive-compatible with the no-purchase option:

$$\mathscr{W}_i^\varnothing(\mathbf{p}) \equiv \left\{\mathbf{v}_i^\varnothing \in \mathscr{V}_i \colon v_{ij}^\varnothing \leq p_j,\ \forall j \in \mathscr{P}\right\}. \tag{10}$$

We now characterize in Lemma 1 when such valuation sets have at least one feasible point, that is, there exists a valuation vector for the products that is incentive-compatible with both the historical choice c_i under the historical price $\mathbf{P}_i$ and product j being chosen under the new price $\mathbf{p}$.

LEMMA 1. *For any price* $\mathbf{p} \geq 0$*, the following statements (a)-(c) hold:*

(a) *The no-purchase option is feasible to the i-th customer class ($\mathscr{W}_i^{\varnothing}(\mathbf{p}) \neq \emptyset$) if and only if $p_{c_i} \geq P_{ic_i}$, i.e., the new price of the historically chosen product c_i remains the same or increases.*

(b) *The purchase of the historically chosen product c_i by the i-th customer class is always feasible ($\mathscr{W}_i^{c_i}(\mathbf{p}) \neq \emptyset$) for all $\mathbf{p} \geq 0$.*

(c) *The purchase of $j \in \mathscr{P} \setminus \{c_i\}$ by the i-th customer class is feasible ($\mathscr{W}_i^{j}(\mathbf{p}) \neq \emptyset$) if and only if $p_j - p_{c_i} \leq P_{ij} - P_{ic_i}$, i.e., the price difference of j with respect to c_i remains the same or decreases.*

Lemma 1 provides easy-to-check conditions for whether the choice of a specific product or none is feasible. It also leads to a more compact formulation of (DP-LP). Intuitively, one can simply screen all products $j \in \mathscr{P}$ and the no-purchase option according to Lemma 1 for feasible options, and choose the one with the lowest revenue. Formally, let $\mathbb{I}(C)$ be the indicator function of the logical condition C, i.e., it is equal to 1 if C is true and 0 otherwise. Proposition 2 applies Lemma 1 to a reformulation of (DP-LP) via a projective argument.

Proposition 2. *The formulation (DP-LP) is equivalent to the linear program*

$$f_i(\mathbf{p}) = \min_{\mathbf{x} \geq 0} \quad \sum_{j \in \mathscr{P}} p_j x_j \tag{DP-C}$$

$$\text{s.t.} \quad \sum_{j \in \mathscr{P}} x_j = \mathbb{I}(p_{c_i} < P_{ic_i}), \tag{11}$$

$$x_j \leq \mathbb{I}(p_j - p_{c_i} \leq P_{ij} - P_{ic_i}), \qquad \forall j \in \mathscr{P}, \tag{12}$$

where $\mathbf{x} = (x_1, \ldots, x_n)$ is n-dimensional.

The formulation (DP-C) reveals the combinatorial structure of the problem for prices $\mathbf{p}$ and the IC polyhedron of customer i. Specifically, it is the minimum price p_j among the feasible products according to Lemma 1, with priority to the no-purchase option if available. Notice that (DP-C) is always feasible for any $\mathbf{p}$ since either $\sum_{j \in \mathscr{P}} x_j = 0$ or $x_{c_i} = 1$ is always a viable purchase option according to Lemma 1-(a) and (b). Furthermore, of particular importance to our methodology is the dual of (DP-C):

$$f_i(\mathbf{p}) = \max_{\mu_i \geq 0, \tau_i} \quad \mathbb{I}(p_{c_i} < P_{ic_i})\tau_i - \sum_{j \in \mathscr{P}} \mathbb{I}(p_j - p_{c_i} \leq P_{ij} - P_{ic_i})\mu_{ij} \tag{DP-C-Dual}$$

$$\text{s.t.} \quad \tau_i - \mu_{ij} \leq p_j, \qquad \forall j \in \mathscr{P}. \tag{13}$$

To draw insights on the above dual problem, suppose prices are ordered as $p_1 \leq p_2 \leq \cdots \leq p_n$. If $p_{c_i} < P_{ic_i}$, then the no-purchase option is not feasible by Lemma 1-(a) and the customer necessarily purchases one product from $\mathscr{P}$. In this case, the worst-case revenue is p_{j^*}, where

$$j^* \equiv \min_{j \in \mathscr{P}} \{ j : p_j - p_{c_i} \leq P_{ij} - P_{ic_i} \}.$$

In an optimal solution (τ_i^*, μ_i^*), variable $\tau_i^* = p_{j^*}$ yields the revenue obtained when prices are set to $\mathbf{p}$. The solution μ_{ij}^* for each $j \in \mathscr{P}$ captures the lost objective value if the product is feasible, i.e.,

$$\mu_{ij}^* = \begin{cases} p_{j^*} - p_j, & \text{if } j < j^*, \\ 0, & \text{otherwise.} \end{cases}$$

We note that these solutions are optimal since all are non-negative (due to the ascending order of prices), feasible to (13), and equal to the same solution value of (DP-C), as only terms indexed by $j \geq j^*$ have non-zero objective coefficients in the dual problem. The structure of the optimal duals also implies two immediate properties that we will leverage in our reformulations.

Proposition 3. *For any $\mathbf{p} \geq 0$, the following statements hold:*

(a) $f_i(\mathbf{p}) \leq p_{c_i} \leq P_{ic_i}$, i.e., the optimal revenue is bounded by the new price of the historically chosen product c_i.

(b) Let $P^{\max} \equiv \max_{i \in \mathscr{C}} P_{ic_i}$ be the maximum historical price paid by any customer. There exists some $\mathbf{p}' \geq 0$ such that $f_i(\mathbf{p}') \geq f_i(\mathbf{p})$ and $p'_j < P^{\max}$ for all $j \in \mathscr{P}$, i.e., pricing any product higher that $P^{\max}$ does not lead to any increase in revenue.

Given the dual interpretation above and Proposition 3, we obtain an equivalent version of the dual that is written in terms of the single variable τ_i:

$$f_i(\mathbf{p}) = \max_{\tau_i \geq 0} \quad \mathbb{I}(p_{c_i} < P_{ic_i})\tau_i \tag{DP-D}$$

$$\text{s.t.} \quad \tau_i \leq p_j + \mathbb{I}(p_j - p_{c_i} > P_{ij} - P_{ic_i})P_{ic_i}, \qquad \forall j \in \mathscr{P}. \tag{14}$$

If $p_j - p_{c_i} \leq P_{ij} - P_{ic_i}$ for some $j \neq c_i$, then the j-th product is feasible to purchase and inequality (14) reduces to $\tau_i \leq p_j$, i.e., the maximum revenue is bounded by p_j. Otherwise, the inequality becomes redundant given Proposition 3-(a). We also note that $\tau_i \leq p_{c_i}$ for $j = c_i$ in (14), i.e., c_i is always feasible and the model is consistent with Proposition 3-(a).

4.2. Robust Pricing Reformulation

We now develop reformulations to address our original pricing problem (OP). We consider the inner optimization of (OP) obtained by dropping the constant term $1/m$ and replacing $f_i(\cdot)$ by (DP-D), which has the same objective sense as the outer problem (hence, moving from a *sup min* problem to a *sup max* problem, which is simply *sup*):

$$\sup_{\mathbf{p}, \tau \geq 0} \quad \sum_{i \in \mathscr{C}} \mathbb{I}(p_{c_i} < P_{ic_i})\tau_i \tag{OP-C}$$

$$\text{s.t.} \quad \tau_i \leq p_j + \mathbb{I}(p_j - p_{c_i} > P_{ij} - P_{ic_i})P_{ic_i}, \qquad \forall i \in \mathscr{C}, j \in \mathscr{P}. \tag{15}$$

Multiplying the optimal value of (OP-C) by $1/m$ yields the optimal revenue τ^*. The difficulty in the above problem is its non-concave and discontinuous objective function, which is defined in

terms of indicator functions on both open and closed half-spaces. We will, however, exploit the constraint structure of (OP-C) to price products to any absolute error of τ^* by a two-step process, the first of which involves a significantly more computationally tractable model.

In particular, the indicator terms of (OP-C) can be reformulated in several ways (see, e.g., Belotti et al. 2016). For our purposes, we study the following equivalent bilinear mixed-integer model:

$$\sup_{\mathbf{p},\tau\geq 0,\mathbf{y}} \quad \sum_{i\in\mathscr{C}} y_{ic_i}\tau_i \tag{OP-B}$$

$$\text{s.t.} \quad \tau_i \leq p_j + (1-y_{ij})P_{ic_i}, \qquad \forall i\in\mathscr{C}, j\in\mathscr{P}, \tag{16}$$

$$p_{c_i} < P_{ic_i} + (P^{\max} - P_{ic_i})(1-y_{ic_i}), \qquad \forall i\in\mathscr{C}, \tag{17}$$

$$p_j - p_{c_i} > P_{ij} - P_{ic_i} - (P^{\max} + P_{ij} - P_{ic_i})y_{ij}, \qquad \forall i\in\mathscr{C}, j\in\mathscr{P}\setminus\{c_i\}, \tag{18}$$

$$\mathbf{y}\in\{0,1\}^{m\times n}. \tag{19}$$

The binary variables $\mathbf{y}$ encode the indicator functions in (OP-C). In particular, the assignment $y_{ij}=1$ for $j\in\mathscr{P}\setminus\{c_i\}$ is interpreted as product j being feasible to customer i under the new price, as per Lemma 1-(c), and $y_{ij}=0$ otherwise. The assignment $y_{ic_i}=1$ indicates that no-purchase option is unavailable for customer i (or, equivalently, that product c_i is purchasable). In such cases, note that the objective and constraints for each customer i reduce to problem (DP-D). Further, we also remark that inequalities (17)-(18) are big-M constraints that rely on $P^{\max}$ defined in Proposition 3. For completeness, we formalize the validity of model (OP-B).

Proposition 4. *At optimality, the solution values of* (OP-C) *and* (OP-B) *match.*

The set of feasible solutions to (OP-B), however, is not polyhedral because the linear inequalities (17)-(18) are defined by open halfspaces. We propose to solve a parameterized version of (OP-B) by setting a precision parameter ϵ on the constraint violation of (OP-B), which allow us to replace the supremum by a maximum:

$$g(\epsilon) \equiv \max_{\mathbf{p},\tau\geq 0,\mathbf{y}} \quad \sum_{i\in\mathscr{C}} y_{ic_i}\tau_i \tag{OP-ϵ}$$

$$\text{s.t.} \quad \tau_i \leq p_j + (1-y_{ij})P_{ic_i}, \qquad \forall i\in\mathscr{C}, j\in\mathscr{P}, \tag{20}$$

$$p_{c_i} \leq P_{ic_i} + (P^{\max} - P_{ic_i})(1-y_{ic_i}) - \epsilon, \qquad \forall i\in\mathscr{C}, \tag{21}$$

$$p_j - p_{c_i} \geq P_{ij} - P_{ic_i} - (P^{\max} + P_{ij} - P_{ic_i})y_{ij} + \epsilon, \qquad \forall i\in\mathscr{C}, j\in\mathscr{P}\setminus\{c_i\}, \tag{22}$$

$$\mathbf{y}\in\{0,1\}^{m\times n}. \tag{23}$$

Because (OP-ϵ) restricts the feasible space of (OP-B) for $\epsilon>0$, the optimal value must be no greater than the one from the original problem, i.e., $g(\epsilon)\leq m\tau^*$ for $\epsilon>0$. Conversely, when $\epsilon=0$, (OP-ϵ) serves as a relaxation and therefore $g(0)\geq m\tau^*$.

The challenge in solving $g(\epsilon)$ is to choose an appropriate $\epsilon > 0$ that leads to a sufficiently close approximation to the real supremum $m\tau^*$. Large values of ϵ may lead to poor approximations due to the discontinuity of (OP), while small values are not computationally tractable due to numerical limitations of solvers. Theorem 1, however, shows that it suffices to solve $g(0)$ to obtain a sufficiently close value to $m\tau^*$ at any desired (absolute) error. It also prescribes a set of final prices with a formal guarantee that can be used by the firm.

THEOREM 1. *Consider the formulation* (OP-ϵ) *with* $\epsilon = 0$. *The following statements hold.*

(a) The model has a finite optimal value attainable by some $\mathbf{p}^0 \geq 0$.

(b) If a product $j \in \mathscr{P}$ *is priced at zero by the optimal vector* $\mathbf{p}^0$, *we can reprice it at* $p_j^0 = \min_{i \in \mathscr{C}, j \in \mathscr{P}} P_{ij} > 0$ *without losing optimality.*

(c) Let $\mathbf{p}^0 > 0$ *be an optimal vector of positive prices after ordering, i.e.,* $0 < p_1^0 \leq p_2^0 \leq \cdots \leq p_n^0$. *For any desired error* $\delta > 0$, *let* $\mathbf{p}'$ *be a vector of prices such that*

$$\mathbf{p}' = \left(p_1^0 - \frac{\delta'}{mn}, p_2^0 - 2\frac{\delta'}{mn}, p_3^0 - 3\frac{\delta'}{mn}, \ldots, p_n^0 - n\frac{\delta'}{mn}\right)$$

for any sufficiently small $0 < \delta' \leq \delta$ *so that* $\mathbf{p}' > 0$. *Then,*

$$0 \leq g(0) - m\tau^* \leq g(0) - \sum_{i \in \mathscr{C}} f_i(\mathbf{p}') \leq \delta,$$

where τ^* *is the optimal solution of the original pricing problem* (OP).

By Theorem 1, we are able to solve (DP) for any desired error δ using the reformulation (OP-ϵ) with $\epsilon = 0$. More precisely, the set of prices $\mathbf{p}^*$ can be obtained by

1. Solving (OP-ϵ) for $\mathbf{p}$ with $\epsilon = 0$, which is guaranteed to exist due to Theorem 1-(a);
2. If any price is zero, increasing it to a positive value according to Theorem 1-(b); and
3. Applying the transformation from Theorem 1-(c) to any desired error δ.

While modern commercial solvers can address nonlinear models of the form (OP-ϵ), we also present an equivalent mixed-integer linear program by a standard big-M reformulation of the quadratic constraints. The program will be the key to the analysis of our approximation algorithms.

$$g(0) = \max_{\mathbf{p}, \tau, \bar{\tau} \geq 0, \mathbf{y}} \quad \sum_{i \in \mathscr{C}} \bar{\tau}_i \tag{OP-MIP}$$

$$\text{s.t.} \quad (20), (21), (22), (23) \text{ with } \epsilon = 0 \tag{24}$$

$$\bar{\tau}_i \leq y_{ic_i} P_{ic_i}, \qquad \forall i \in \mathscr{C}, \tag{25}$$

$$\bar{\tau}_i \leq \tau_i, \qquad \forall i \in \mathscr{C}, \tag{26}$$

$$\bar{\tau}_i \geq \tau_i - (1 - y_{ic_i}) P_{ic_i}, \qquad \forall i \in \mathscr{C}. \tag{27}$$

4.3. Special Cases

We now discuss two cases of practical interest where (OP-MIP) can be solved analytically. First, we consider the scenario where each individual customer is offered the same price for all products in the assortment. The prices, however, can be different per customer. This occurs when products are similar in nature and customers are offered personalized promotions over time. Proposition 5 states the structure of the optimal solution for this case.

PROPOSITION 5. *Suppose that, for each customer $i \in \mathscr{C}$, all products $j \in \mathscr{P}$ have the same historical price $P_{ij} = P_i$. Furthermore, without loss of generality, assume prices are ordered, i.e., $P_1 \leq P_2 \leq \cdots \leq P_m$. The price vector $\mathbf{p}^*$ defined by $p_j^* = P_{i^*}$ for all $j \in \mathscr{P}$, where*

$$i^* = \arg\max_{i \in \mathscr{C}} \{(m - i + 1)P_i\},$$

is optimal to (OP-MIP).

Proposition 5 reveals a connection with the classical pricing literature. Specifically, if we perceive the set $\{P_1, \ldots, P_m\}$ as the empirical distribution of the customer valuations, then our problem, when all products are offered at the same price, reduces to the well-studied revenue maximization problem $\max_p\{p \cdot d(p)\}$, where $d(p)$ is the demand function of price p.

As our second practical case, we consider a setting where the price for any product is fixed over time. More precisely, prices may differ per product but not per customer. Proposition 6 also indicates that the optimal prices have a simpler structure, i.e., it suffices to set them to their historical prices.

PROPOSITION 6. *Suppose that, for each product $j \in \mathscr{P}$, all customers observe the same historical price P_j. The price vector $\mathbf{p}^*$ such that $p_j^* = P_j$ for all $j \in \mathscr{P}$ is optimal to* (OP-MIP).

5. Approximate Pricing Strategies and Analysis

The formulation (OP-MIP), albeit amenable to state-of-the-art commercial solvers, may still be challenging to solve due to its difficult constraint structure (e.g., the presence of big-M constraints). Moreover, Abdallah and Vulcano (2020) point out that in retail, the number of SKUs in a family of products could be on the order of several hundred. Thus, even with a few data points per product, the size of (OP-MIP) could be significantly large. While recent works have focused on choice model estimation in such high-dimensional settings (Jiang et al. 2020), to the best of our knowledge, model-free pricing approaches are yet to be developed for such settings.

In this section, we analyze three interpretable and intuitive approximation algorithms that are of low-polynomial time complexity in the input size of the problem, and hence scalable to large problem sizes. We discuss their benefits and worst-case revenue performance in comparison to the

optimal solution of (OP-MIP). Specifically, we evaluate in §5.1 and §5.2 two standard heuristics based on historical prices and linear programming (LP) relaxations, respectively. In §5.3, we propose an approximation based on a simplified version of (OP-MIP), which provides the strongest approximation factor among the three policies, is efficient to compute, and also interpretable.

We note that in practice the firm might want to use simpler pricing algorithms, such as offering the products at their average historical prices or at the prices observed by a random historical customer. However, it can be shown that such pricing algorithms can lead to an arbitrarily poor revenue performance in comparison to the optimal solution of (OP-MIP). See Propositions 9 and 10 in the appendix. Thus, we do not discuss them in details in this section.

5.1. Conservative Pricing

A simple approximation a conservative firm may consider is to price all the products at their historically lowest purchase price. That is,

$$p_j = \min_{i \in \mathscr{C}:\, c_i = j} P_{ij}, \qquad \forall j \in \mathscr{P}. \tag{28}$$

By inequality (20) for $j = c_i$, the above prices guarantee that each customer i would purchase at least one product (i.e., the non-purchase option will not be chosen). Furthermore, it also follows that the final total revenue is at least $m\underline{P}$, where $\underline{P} \equiv \min_{i \in \mathscr{C}} P_{ic_i}$.

Such a pricing policy maximizes the demand at the cost of a lower profit margin, and is thus referred to as "conservative pricing." We show below that it has a strainghtforward worst-case performance bound, which is also tight.

Proposition 7. *Let $\underline{P} \equiv \min_{i \in \mathscr{C}} P_{ic_i}$ and $\bar{P} \equiv \max_{i \in \mathscr{C}} P_{ic_i}$ be the minimum and maximum historical purchase prices, respectively. The revenue from the conservative pricing* (28) *is at least $\underline{P}/\bar{P}$ of the optimal value of* (OP-MIP)*. Futhermore, this ratio is asymptotically tight.*

As one may expect, when products are not similar in nature and their prices vary in a wide range, this approximation is too conservative and would not perform well. However, the ratio $\underline{P}/\bar{P}$ provides a useful benchmark which can be used to gauge the performance of other approximations.

5.2. LP Relaxation Pricing

It is a natural practice to consider the LP relaxation of the program (OP-MIP), i.e., to replace the integrality constraint (23) by the continuous domain $\mathbf{y} \in [0,1]^{m \times n}$. The resulting LP model can be solved in (weakly) polynomial time, from which we can extract a candidate price vector $\mathbf{p}^{\mathrm{LP}}$. We wish to evaluate the quality of $\mathbf{p}^{\mathrm{LP}}$ with respect to the optimal prices of (OP-MIP). Notice that the remaining variables of (OP-MIP) can be easily determined when $\mathbf{p}$ is fixed to $\mathbf{p}^{\mathrm{LP}}$.

We next show in the following example that the resulting LP prices, unfortunately, have a worse worst-case performance than the conservative pricing approach.

Example 1. Consider an instance with $m = 3$ customers and $n = 2$ products. The historical prices P_{ij} and the customer choices (in bold) are listed in the table below:

Price	Product 1	Product 2
Customer 1	**1**	2
Customer 2	2	**3**
Customer 3	**1**	3

It can be shown that the LP relaxation yields the price vector $\mathbf{p}^{\mathrm{LP}} = (1.2, 2.3)$ and an upper bound of 4.8 to the optimal value of (OP-MIP). However, when evaluating (OP-MIP) with variables $\mathbf{p}$ fixed to $\mathbf{p}^{\mathrm{LP}}$, both customers 1 and 3 do not purchase any products, while customer 2 purchases product 1 in the worst-case. The following set of valuations, v_{ij}, that are drawn from the IC polyhedra of customers 1, 2 and 3 conform to these customer choices that lead to a worst-case revenue for the firm.

Valuation	Product 1	Product 2
Customer 1	1	0
Customer 2	2	3
Customer 3	1	0

Thus, the total revenue generated from $\mathbf{p}^{\mathrm{LP}}$ is \$1.2. The optimal solution of this instance is \$4.0 certified by prices $\mathbf{p}^* = (1.0, 2.0)$, which leads to an LP solution ratio of 30%, worse than the conservative price ratio of $\underline{P}/\bar{P} = 1/3 \approx 33\%$.

The optimal LP solution associated with variables $\mathbf{y}$ provides insights into the poor performance of the heuristic. In particular, consider inequality (20) for customer 2 ($i = 2$) and product 1 ($j = 1$):

$$\tau_2 \leq p_1 + P_{22}(1 - y_{21}).$$

At the LP optimality, $y_{21} = 0.4$ and the above right-hand side is equal to 3. The inequality is also tight, leading to a (relaxed) revenue of $\tau_2 = 3.0$. However, with integrality constraints, $y_{21} = 1.0$ and the constraint is again tight with $\tau_2 \leq p_1 = 1.2$, which is a significant decrease in revenue. The same issue is identified for the other customers.

More generally, the big-M structure of inequalities (20) tends to result in the optimal LP solution pricing p_j slightly higher than the historical prices for most of the customer i with $c_i = j$, in spite of the fact that the indicator $\mathbb{I}(p_j > P_{ic_j})$ has been encoded in (OP-MIP) to represent the no-purchase option. More precisely, y_{ij} can be set to $1 - \epsilon$ for any sufficiently small ϵ to overcome that condition, and the objective value of the LP is not impacted significantly since the customer is still assumed to purchase the product. However, when the integrality constraint is imposed, either the inequality becomes binding with respect to some price or the non-purchase option is chosen, leading to a smaller expected revenue. □

While the worst-case revenue is not superior relative to conservative pricing, we show in §6 that the LP relaxation pricing usually outperforms conservative pricing numerically in most cases. Intuitively, conservative pricing is more concerned with the worst-case, while LP relaxation can be close to maximizing the expected performance, specifically since we have tuned the big-M constrained to be very tight with respect to the input parameters.

5.3. Cut-off Pricing

In this policy, we relax customers' IC constraints to obtain a heuristic. In particular, suppose the IC polyhedron of $i \in \mathscr{C}$ is drawn for the new customer. When deciding which product to buy under $\mathbf{p}$, the new customer does not fully follow the IC constraints; rather, as long as the price of c_i is lower than the price of the historically chosen product by customer i, i.e., $p_{c_i} \leq P_{ic_i}$, all products are eligible for purchase by the new customer. Technically, we only keep part (a) of Lemma 1 and neglect the other two parts. Therefore, if the customer decides to purchase a product (i.e., $p_{c_i} \leq P_{ic_i}$), then she chooses the product with the lowest price $\arg\min_{j \in \mathscr{P}} p_j$ in the worst-case.

We formulate this setting in the following model:

$$\max_{\mathbf{p}, \tau \geq 0} \left\{ \sum_{i \in \mathscr{C}} \mathbb{I}(p_{c_i} \leq P_{ic_i}) \tau_i \,:\, \tau_i \leq p_j, \ \forall j \in \mathscr{P} \right\}. \tag{OP-CP}$$

Under the assumed purchase rule, it follows that regardless of the sampled IC polyhedron, as long as the new customer decides to purchase, she necessarily picks the same product. Thus, the key is to determine the minimum price p^*. Then setting all products at this price ($p_i \equiv p^*$) maximizes the objective in (OP-CP). As we vary the value of p^*, the indicator functions only change values when $p^* = P_{ic_i}$. Letting $p^* = P_{ic_i}$ for some $i \in \mathscr{C}$, formulation (OP-CP) simplifies to

$$\max_{i \in \mathscr{C}} \sum_{i' \in \mathscr{C}} \mathbb{I}(P_{i'c_{i'}} \geq P_{ic_i}) \, P_{ic_i}, \tag{29}$$

which can be solved in $\mathcal{O}(m)$ time complexity by inspecting one customer at a time. We denote the optimal solution to (OP-CP) by $p^* = P_{ic_i}$ for some customer i. In particular, p^* can be perceived as a cut-off price: if the new customer samples the IC polyhedron of customer i', then she does not purchase any product if and only if the historical price paid by customer i', $P_{i'c_{i'}}$, is below p^*. We leverage this to propose the cut-off pricing approximation $\mathbf{p}^{\text{CP}}$ as follows. For every $j \in \mathscr{P}$,

$$p_j^{\text{CP}} = \begin{cases} \min_{i \in \mathscr{C}} \{P_{ic_i} : c_i = j, \, P_{ic_i} \geq p^*\}, & \text{if } \{i \in \mathscr{C} : c_i = j, P_{ic_i} \geq p^*\} \neq \emptyset, \\ \bar{P}, & \text{otherwise}, \end{cases} \tag{30}$$

where each product j is priced at its lowest historical purchase price that was equal to or above the cut-off price p^*. We note that not all products are priced at the cut-off price. Rather, they are typically priced slightly higher than p^*. Such modifications do not change the values of the

indicator functions and thus the optimal value of (OP-CP) but lead to better and less conservative empirical performances.

We next show its performance in Proposition 8, recalling that $\underline{P} \equiv \min_{i \in \mathscr{C}} P_{ic_i}$ and $\bar{P} \equiv \max_{i \in \mathscr{C}} P_{ic_i}$ are the minimum and maximum historical purchase prices, respectively.

PROPOSITION 8. *The cut-off pricing* (30) *generates a revenue that is at least* $\frac{1}{1+\log(\bar{P}/\underline{P})}$ *of the optimal value of* (OP-MIP). *Furthermore, this ratio is asymptotically tight.*

Compared to Proposition 7, the cut-off pricing dramatically improves upon the worst-case scenario of the conservative pricing, especially when $\bar{P} \gg \underline{P}$. We also observe in §6 that this pricing policy is superior numerically to the the earlier proposed heuristics.

6. Numerical Analysis

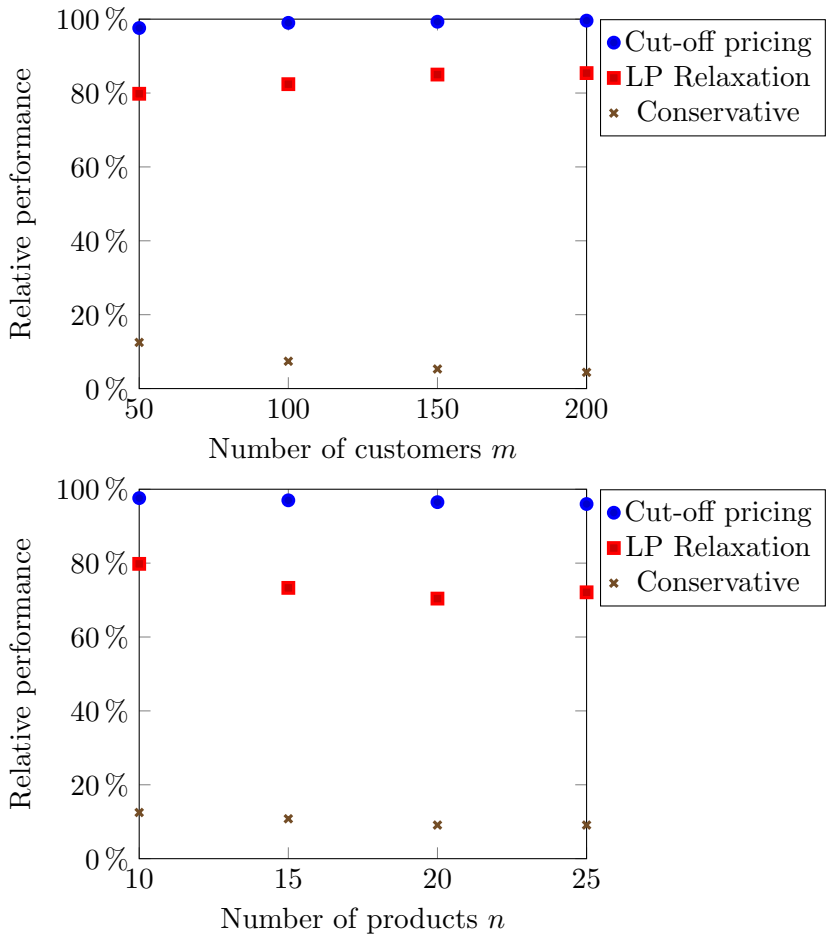
We now present a numerical study of the proposed approaches on both synthetic and real datasets, evaluating our methodologies with respect to classical model-based methodologies in scenarios of practical interest. We start in §6.1 with an analysis of the empirical performance of the approximation pricing strategies from §5. In §6.2, we evaluate our data-driven pricing approach on "small-data" regimes, which are typically challenging to classical model-based methods. Next, we consider a scenario in §6.3 where the firm misspecifies the pricing model. Finally, in §6.4 we compare all approaches on a large-scale dataset from the U.S. retail industry.

6.1. Approximation Performance

In this section, we use synthetic data to investigate the performance of the approximation algorithms developed in Section 5. We generate instances with $m \in \{50, 100, 150, 200\}$, $n \in \{10, 15, 20, 25\}$, and historical prices P_{ij} drawn uniformly at random from the interval $(0, 10)$. Each customer chooses a product in the assortment with equal probability $1/n$. Based on the synthetic data, we compute the optimal value $g(0)$ from (OP-MIP), and the objective of the three approximations in Section 5. We report the objective values of the approximations relative to the optimal value $g(0)$ for 200 independent instances.

Figure 1 depicts the relative performance ratio in percentage (i.e., $100\times$optimal revenue/heuristic revenue) of the conservative, LP relaxation, and cut-off pricing when the number of customers m and the number of products n vary in the historical data. The figures suggest that conservative pricing performs poorly (achieving less than 15% of the optimal revenues on average) and cut-off pricing does the best among the three, obtaining at least 96% of the optimal value in all the cases. Increasing the number of customers or decreasing the number of products improves the performance of LP relaxation pricing and cut-off pricing. Moreover, the numerical results suggest that the cut-off pricing significantly outperforms the conservative pricing, as expected from its theoretical

Figure 1 **The performance of the three approximation algorithms.**



performance guarantee (Proposition 8), as cut-off pricing generally has an average performance that is an order of magnitude higher than that of conservative pricing. Finally, we observe that the numerical cut-off price performance is far above that of its theoretical performance; we include additional tables in Appendix A with the expected theoretical performance of the tested instances.

Table 1 shows the solution time of the conservative, LP relaxation, cut-off pricing, and optimal problem when the number of customers m and the number of products n vary in the historical data. The results suggest that while the optimal solution time for mid-size problems is not too egregious, however it does not scale well when the problem size grows, as is expected. However, cut-off pricing pricing solves all problem sizes in under one second.

6.2. "Small" Data

When the data is limited, model-based methods risk misspecifying the model or incorporating noisy estimations. In this section, we evaluate revenues obtained from the proposed pricing approaches in such cases, comparing with a classical multinomial logit (MNL) model (Train 2009).

Table 1 The average solution time of the approximation strategies and optimal solution. Standard errors are reported in parentheses.

	Solution time in seconds			
(m,n)	Conservative	LP Relaxation	Cut-off	Optimal
$(50,10)$	0.041 (0.001)	0.453 (0.001)	0.040 (0.001)	1.477 (0.078)
$(50,15)$	0.065 (0.001)	1.049 (0.005)	0.062 (0.001)	4.250 (0.195)
$(50,20)$	0.080 (0.001)	1.762 (0.006)	0.083 (0.001)	5.691 (0.254)
$(50,25)$	0.103 (0.001)	2.767 (0.007)	0.109 (0.001)	8.003 (0.313)
$(100,10)$	0.081 (0.001)	0.929 (0.002)	0.078 (0.001)	17.260 (0.800)
$(150,10)$	0.117 (0.001)	1.419 (0.003)	0.132 (0.002)	89.050 (5.400)
$(200,10)$	0.165 (0.002)	2.018 (0.006)	0.181 (0.003)	525.800 (42.930)

Table 2 Instance parameters for study in §6.2.

Experiment	α	P_{ij}
Low utility	$[-2,0]$	$[2.5,4.5]$
High utility	$[1,3]$	$[5.5,8.5]$

We generate synthetic instances with $n=10$ products and varying number of customers $m \in \{20,30,\ldots,110\}$. The probability of customer i choosing product j from the assortment is given by

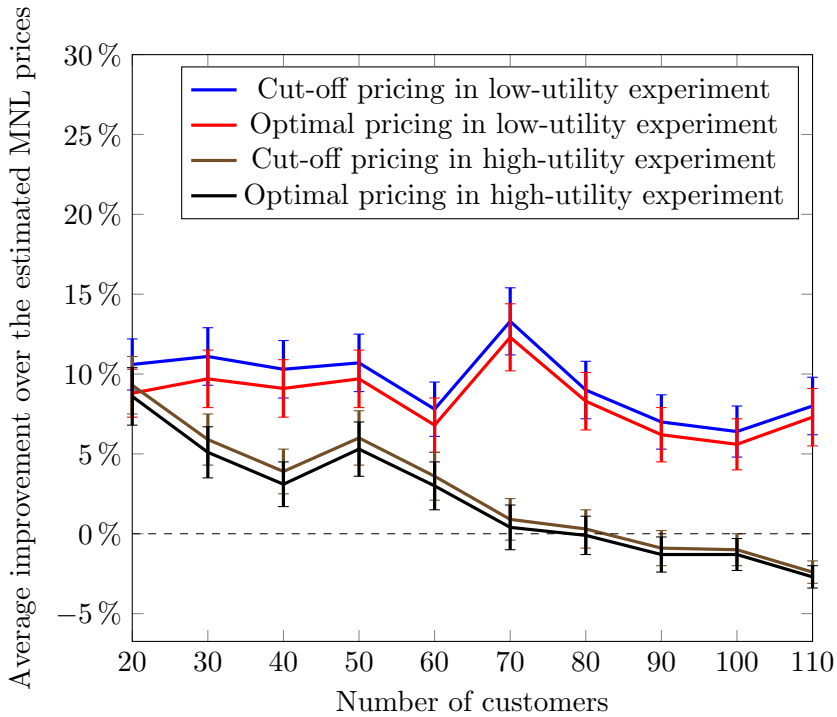
$$\frac{\exp(\alpha_j - \beta P_{ij})}{1+\sum_{k=1}^{n}\exp(\alpha_k - \beta P_{ik})},$$

where $\beta = 0.5$. We consider two sets of experiments for the remaining parameters α and the historical prices, representing low and high customer utility. Specifically, for the low-utility experiment, historical prices P_{ij} are drawn uniformly at random from the interval $[5.5,8.5]$, and $\{\alpha_j\}_{j=1}^{10}$ are independent and drawn uniformly at random from the interval $[1,3]$. In the high-utility experiment, historical prices are drawn uniformly at random from the interval $[2.5,4.5]$ and $\{\alpha_j\}_{j=1}^{10}$ are independent and drawn uniformly at random from the interval $[-2,0]$. Table 2 summarizes the setup of the two experiments. Note that, in both cases, this choice of parameters guarantees that the optimal price of the MNL model is within the range of the historical prices.

We simulate 200 independent instances for each m. In each instance, we calculate the optimal solution to (OP-MIP) and cut-off pricing. Moreover, we use the BIOGEME package developed by Bierlaire (2003) to estimate the parameters of the MNL model with the historical data, and then calculate the optimal price based on the fitted model. Note that they are the prices obtained from the model-estimate-optimize approach under the correct model specification. We then evaluate the three sets of prices with respect to the ground-truth model and compare their expected revenues.

In Figure 2, we illustrate the average difference of the revenues of cut-off (data-driven optimal) pricing and the estimated MNL prices, relative to the optimal MNL prices when the parameters

Figure 2 The performance of the data-driven optimal pricing (OP-MIP) and cut-off pricing relative to the optimal MNL prices estimated from the data. The difference in the revenues is converted to percentage by dividing it by the optimal revenue of the model and then averaged.



are known. Note that if the quantity is positive, then it implies that our approach outperforms the estimated MNL prices. From the figure, when the number of customers is less than 70, both data-driven pricing schemes outperform the estimated MNL prices, in both experiments. Note that the MNL prices are estimated based on the correct specification of the model. This experiment further suggests that data-driven approaches may be beneficial with respect to model-based approaches when the data size is small. It is also expected that as the data size grows, the model-estimate-optimize approach eventually converges to the optimal prices of the true model, given that the model is correctly specified. In this regime, data-driven approaches may not be beneficial.

6.3. Model Misspecification

One potential benefit of the proposed approach is that it is agnostic to the underlying model, and thus less sensitive to model misspecification. We assess this scenario in the next experiment with synthetic instances.

We consider $m = 50$ customers and $n = 10$ products in two experiment sets, also representing low and high utilities. The historical prices P_{ij} are drawn uniformly at random from $[2.5, 4.5]$ and $[5.5, 8.5]$ in the two experiments. For each customer, we generate their choices using a mixed

Table 3 Instance parameters for study in §6.3.

Experiment	α	P_{ij}
Low utility	$[-2,0]$	$[2.5,4.5]$
High utility	$[1,3]$	$[5.5,8.5]$

Table 4 Revenue between data-driven pricing and misspecified MNL model.

	Expected revenue in the mixed logit model	
Pricing method	Low-utility experiment	High-utility experiment
Data-driven optimal	0.725 (0.005)	2.393 (0.007)
Cut-off pricing	0.734 (0.005)	2.415 (0.007)
MNL optimal pricing	0.635 (0.014)	2.113 (0.035)

logit model with two classes (Train 2009). More precisely, given $(P_{i1},\ldots,P_{in})$, the probability of customer i choosing product j is

$$\frac{1}{2}\cdot\frac{\exp(\alpha_{1j}-\beta_1 P_{ij})}{1+\sum_{k=1}^{n}\exp(\alpha_{1k}-\beta_1 P_{ik})}+\frac{1}{2}\cdot\frac{\exp(\alpha_{2j}-\beta_2 P_{ij})}{1+\sum_{k=1}^{n}\exp(\alpha_{2k}-\beta_2 P_{ik})}.$$

Here we set $\beta_1=0.5$ and $\beta_2=2$. In each of the 200 instances, we randomly draw $(\alpha_{1j},\alpha_{2j})$ independently from $[-2,0]$ (for $P_{ij}\in[2.5,4.5]$) and $[1,3]$ (for $P_{ij}\in[5.5,8.5]$). Again, such choices of the ranges of P and α guarantees that the historical price ranges cover the optimal MNL with $\beta=0.5$ prices (See Table 3).

We investigate the case that the model is misspecified. In particular, we fit the MNL model instead of the mixed logit model using BIOGEME to the historical data. We calculate the optimal prices for the fitted MNL model and compare its expected revenue to our data-driven approaches (optimal solution and cut-off pricing) under the mixed logit model.

Table 4 suggests a better performance of data-driven assortment pricing as the optimal prices of the MNL model are designed for a misspecified model. Note that in this case, the misspecification error between the MNL model and the mixed logit model with two classes is arguably mild. We may expect the benefit of data-driven assortment pricing to be more substantial when there are strong irregular patterns in the data that cannot be captured by the assumed model. Combined with the results in §6.2, this further suggests that our proposed pricing models could be beneficial when the data size is small or the firm has little confidence in the modeling of the demand.

6.4. Real Datasets

In this section, we apply data-driven assortment pricing to the IRI Academic Dataset (Bronnenberg et al. 2008). The IRI data collects weekly transaction data from 47 U.S. markets from 2001 to 2012,

covering more than 30 product categories. Each transaction includes the week and store of the purchase, the universal product code (UPC) of the purchased item, the number of units purchased and the total paid dollars. We investigate the product category of razors and the transactions from the first two weeks in 2001. To construct the assortments, we focus on the top ten purchased products from all stores during the two weeks, that is, $n = 10$. The purchases of all other products are treated as "no-purchase." An assortment is thus defined as the products of the same store when the customer visits. Moreover, we follow the procedures in van Ryzin and Vulcano (2015) and Şimşek and Topaloglu (2018): for each purchase record, four no-purchase records of the same assortment are added to the dataset. The goal is to create a reasonable fraction of customers who do not buy any products.

To evaluate the performance of a pricing scheme, we resort to a model (estimated from the data) that describes how consumers choose products and calculate the expected revenue of the scheme under the model. We fit three models to the data to highlight the model-free or model-insensitive nature of our approach. More precisely, in the first MNL model we estimate $\{\alpha_j\}_{j=1}^{10}$ and β in the choice probability

$$\frac{\exp(\alpha_j - \beta P_{ij})}{1 + \sum_{k=1}^{10} \exp(\alpha_k - \beta P_{ik})}.$$

In the second mixed logit model, we estimate $\{w_l, \beta_l, \alpha_{l1}, \ldots, \alpha_{l,10}\}_{l=1}^{2}$ in the choice probability

$$w_1 \frac{\exp(\alpha_{1j} - \beta_1 P_{ij})}{1 + \sum_{k=1}^{10} \exp(\alpha_{1k} - \beta_1 P_{ik})} + w_2 \frac{\exp(\alpha_{2j} - \beta_2 P_{ij})}{1 + \sum_{k=1}^{10} \exp(\alpha_{2k} - \beta_2 P_{ik})}.$$

Both models are estimated using BIOGEME. We also fit a linear demand model with parameters $\{\alpha_j, \beta_j\}_{j=1}^{10}$ and $\{\beta_{jk}, \gamma_{jk}\}_{j \neq k}$. The choice probability of product j in the linear model is

$$\alpha_j - \beta_j P_{ij} + \sum_{k \neq j} (\beta_{jk} I_{ik} P_{ik} + \gamma_{jk}(1 - I_{ik})),$$

fitted using the ordinary least squares. Here $I_{ik} \in \{0, 1\}$ is the indicator for whether product k is included in the assortment seen by customer i. Note that we fit the choice probability separately for product $j = 1, \ldots, 10$ and the no-purchase probability is one minus their sum.

With the three estimated models, we generate three datasets. We set $m = 50$ and $n = 10$. The historical prices P_{ij} are drawn from $\{0.9\bar{P}_j, 0.95\bar{P}_j, \bar{P}_j, 1.05\bar{P}_j, 1.1\bar{P}_j\}$, where $\bar{P}_j$ is the average price of product j. The choices of customers are then generated using one of the three models. We then calculate the prices using the optimal solution to (OP-MIP) and cut-off pricing, and plug them into the three models to evaluate their expected revenues. We also compute the expected revenue of the incumbent prices, which is the average product prices in the IRI data, under the three models.

Note that we apply our proposed pricing techniques to the datasets generated from the models and then evaluate the expected revenues of the calculated prices, the optimal solution to (OP-MIP)

Table 5 Expected revenues of data-driven assortment pricing and the incumbent prices under the three models fitted using the IRI data.

	Fitted demand model		
Pricing method	MNL	Mixed logit	Linear Demand
Data-driven optimal	1.507 (0.007)	1.513 (0.008)	1.197 (0.018)
Data-driven cut-off pricing	1.730 (0.007)	1.731 (0.007)	1.477 (0.019)
Incumbent prices	1.464 (-)	1.468 (-)	0.944 (-)

or cut-off pricing, under the corresponding models. A seemingly more straightforward approach is to directly apply data-driven assortment pricing to the original IRI dataset, and then obtain the prices to evaluate in the three models. However, the major concern is the conflation of the model misspecification error and the efficacy of the pricing schemes. Suppose the data-driven prices calculated using the IRI data perform poorly under, say, the MNL model. There might be two reasons: (1) the estimated MNL model accurately captures the pattern in the IRI dataset but the data-driven approach fails to approximate the optimal prices of the MNL model, or (2) the MNL model does not fit the data and the data-driven approach, which is solely based on the data, cannot possibly approximate the optimal prices of the MNL model. We lack a reliable way to disentangle the two factors, the goodness of fit versus the performance of our approach. By the simulated datasets using the three models, we can control the goodness of fit and isolate the performance of the data-driven pricing schemes. Nevertheless, since the three models are fitted using the IRI data, they are expected to capture the choice patterns in the real world to a large degree.

Table 5 shows our results for this experiment with 200 instances, for each of the three estimated models. The linear model generates significantly lower expected revenues than the other two, possibly because the estimated demand is lower in the region around the incumbent prices in the linear model. Compared with the incumbent prices, data-driven assortment pricing significantly improves the expected revenue under all three demand models. It suggests that our approach offers a robust improvement in this setting.

7. Practical Considerations and Concluding Remarks

We point out a few practical considerations when our approach is applied to real-world problems. The first issue is censoring, i.e., when a customer walks away without purchasing and thus cannot be observed in the data. Many pricing approaches struggle to handle censoring and completely ignoring data censoring may result in price distortion. Fortunately, our approach can incorporated data censoring. Indeed, as mentioned in Remark 1, customers who do not buy anything can be removed from the dataset without affecting the resulting prices. As a result, the data-drive prices do not depend on the censored customers.

One implicit assumption we made is that consumers have price sensitivity identical to one, reflected in the quasilinear utilities. In fact, this assumption can be easily relaxed. The IC polyhedron can be constructed in the same fashion, as long as an individual customer has the same price sensitivity for all products. In this case, dividing (1) by the same factor results in the same polyhedron and the theoretical results still hold.

Another assumption we make is that the firm does not observe the new incoming customer's information and hence assumes she behaves similar to one of the previously seen customers with equal probability. In practice, the firm may potentially know that the customer is a returning customer. Then, one can analytically solve model (OP-MIP) for that customer. If the firm has seen the customer beforehand, it is optimal under the worst-case valuations of the customer to set the prices equal to the ones the customer observed in her historical purchase. Moreover, using consumer features to predict their shopping behavior is a popular practice in modern retailing. For example, an arriving consumer may have a similar background to a segment of past customers. To incorporate consumer features, we may put different weights (as opposed to equal weights) on the IC polyhedra $\mathscr{V}_1, \mathscr{V}_2, \ldots, \mathscr{V}_m$ in the formulation, based on how similar the arriving consumer is to a past one. Computationally, our approach can still accommodate this case. It remains an exciting research direction to capture the consumer features and properly reflect them in the weights.

While we are mainly interested in a situation where the number of customers is low relative to number of products, scalability is still important from a practical point of view. We have demonstrated in Section 6 that our mixed-integer programming reformulations can handle mid-scale problems with hundreds of samples. For large-scale data, our best approximation (cut-off pricing) is of low complexity and scales linearly with the number of samples. While it provides interpretable and intuitive prices, without the need for a commercial solver, such an approximation also has robust theoretical guarantees and good empirical performance as suggested by our numerical study.

References

Abdallah T, Vulcano G (2020) Demand estimation under the multinomial logit model from sales transaction data. *Manufacturing Service Oper. Management* Forthcoming.

Allouah A, Besbes O (2019) Sample-based optimal pricing. *Working Paper* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3334650.

Allouah A, Besbes O (2020) Prior-independent optimal auctions. *Management Sci.* Forthcoming.

Balas E (1998) Disjunctive programming: Properties of the convex hull of feasible points. *Discrete Appl. Math.* 89(1-3):3–44.

Ban GY, Keskin NB (2020) Personalized dynamic pricing with machine learning: High dimensional features and heterogeneous elasticity. *Management Sci.* Forthcoming.

Ban GY, Rudin C (2019) The big data newsvendor: Practical insights from machine learning. *Oper. Res.* 67(1):90–108.

Bandi C, Bertsimas D (2014) Optimal design for multi-item auctions: a robust optimization approach. *Math. Oper. Res.* 39(4):1012–1038.

Belotti P, Bonami P, Fischetti M, Lodi A, Monaci M, Nogales-Gómez A, Salvagnin D (2016) On handling indicator constraints in mixed integer programming. *Comput. Optim. Appl.* 65(3):545–566.

Bertsimas D, Kallus N (2020) From predictive to prescriptive analytics. *Management Sci.* 66(3):1025–1044.

Bertsimas D, Vayanos P (2017) Data-driven learning in dynamic pricing using adaptive optimization. *Working Paper* http://web.mit.edu/~dbertsim/www/papers/Robust%20Optimization/Data-driven%20learning%20in%20dynamic%20pricing%20using%20adaptive%20optimization.pdf.

Bierlaire M (2003) Biogeme: A free package for the estimation of discrete choice models. *Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland.*, www.strc.ch.

Blanchet J, Gallego G, Goyal V (2016) A Markov chain approximation to choice modeling. *Oper. Res.* 64(4):886–905.

Bronnenberg BJ, Kruger MW, Mela CF (2008) Database paper—the IRI marketing data set. *Marketing Sci.* 27(4):745–748.

Caro F, Kök AG, Martínez-de Albéniz V (2020) The future of retail operations. *Manufacturing Service Oper. Management* 22(1):47–58.

Chen N, Gallego G, Tang Z (2019) The use of binary choice forests to model and estimate discrete choice models. *Working Paper* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3430886.

Cohen MC, Leung NHZ, Panchamgam K, Perakis G, Smith A (2017) The impact of linear optimization on promotion planning. *Oper. Res.* 65(2):446–468.

Cohen MC, Lobel I, Paes Leme R (2020) Feature-based dynamic pricing. *Management Sci.* Forthcoming.

Cohen MC, Lobel R, Perakis G (2018) Dynamic pricing through data sampling. *Prod. Oper. Management* 27(6):1074–1088.

Cohen MC, Perakis G (2020) Optimizing promotions for multiple items in supermarkets. *Channel Strategies and Marketing Mix in a Connected World*, 71–97 (Springer).

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

Derakhshan M, Golrezaei N, Paes Leme R (2019) LP-based approximation for personalized reserve prices. *Proceedings of the 2019 ACM Conference on Economics and Computation*, 589–589.

Dong J, Simsek AS, Topaloglu H (2019) Pricing problems under the markov chain choice model. *Prod. Oper. Management* 28(1):157–175.

Dong L, Kouvelis P, Tian Z (2009) Dynamic pricing and inventory control of substitute products. *Manufacturing Service Oper. Management* 11(2):317–339.

Elmachtoub AN, Grigas P (2020) Smart "predict, then optimize". *Management Sci.* Forthcoming.

Elmachtoub AN, Gupta V, Hamilton M (2020) The value of personalized pricing. *Management Sci.* Forthcoming.

Ettl M, Harsha P, Papush A, Perakis G (2020) A data-driven approach to personalized bundle pricing and recommendation. *Manufacturing Service Oper. Management* 22(3):461–480.

Farias VF, Jagabathula S, Shah D (2013) A nonparametric approach to modeling choice with limited data. *Management Sci.* 59(2):305–322.

Ferreira KJ, Lee BHA, Simchi-Levi D (2016) Analytics for an online retailer: Demand forecasting and price optimization. *Manufacturing Service Oper. Management* 18(1):69–88.

Gallego G, Wang R (2014) Multiproduct price optimization and competition under the nested logit model with product-differentiated price sensitivities. *Oper. Res.* 62(2):450–461.

Gaubys J (2020) Ecommerce market share from 2017 to 2023. URL https://www.oberlo.ca/statistics/ecommerce-share-of-retail-sales.

Hopp WJ, Xu X (2005) Product line selection and pricing with modularity in design. *Manufacturing Service Oper. Management* 7(3):172–187.

Jagabathula S, Rusmevichientong P (2017) A nonparametric joint assortment and price choice model. *Management Sci.* 63(9):3128–3145.

Jiang ZZ, Li J, Zhang D (2020) A high-dimensional choice model for online retailing. *Jun and Zhang, Dennis, A High-Dimensional Choice Model for Online Retailing* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3687727.

Jin Q, Lin JY, Zhou SX (2020) Price discounts and personalized product assortments under multinomial logit choice model: A robust approach. *IISE Trans.* Forthcoming.

Li G, Rusmevichientong P, Topaloglu H (2015) The d-level nested logit model: Assortment and price optimization problems. *Oper. Res.* 63(2):325–342.

Li H, Huh WT (2011) Pricing multiple products with the multinomial logit and nested logit models: Concavity and implications. *Manufacturing Service Oper. Management* 13(4):549–563.

Mai T, Jaillet P (2019) Robust multi-product pricing under general extreme value models. *Working Paper* https://arxiv.org/abs/1912.09552.

Rusmevichientong P, Topaloglu H (2012) Robust assortment optimization in revenue management under the multinomial logit choice model. *Oper. Res.* 60(4):865–882.

Rusmevichientong P, Van Roy B, Glynn PW (2006) A nonparametric approach to multiproduct pricing. *Oper. Res.* 54(1):82–98.

Şimşek AS, Topaloglu H (2018) An expectation-maximization algorithm to estimate the parameters of the Markov chain choice model. *Oper. Res.* 66(3):748–760.

Train KE (2009) *Discrete choice methods with simulation* (Cambridge university press, Cambridge, England).

van Ryzin G, Vulcano G (2015) A market discovery algorithm to estimate a general class of nonparametric choice models. *Management Sci.* 61(2):281–300.

van Ryzin G, Vulcano G (2017) An expectation-maximization method to estimate a rank-based choice model of demand. *Oper. Res.* 65(2):396–407.

Yan Z, Cheng C, Natarajan K, Teo CP (2020) A representative consumer model in data-driven multi-product pricing optimization. *Working Paper* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2832385.

Zhang H, Rusmevichientong P, Topaloglu H (2018) Multiproduct pricing under the generalized extreme value models with homogeneous price sensitivity parameters. *Oper. Res.* 66(6):1559–1570.

Zhang M, Ahn HS, Uichanco J (2020) Data-driven pricing for a new product. *Working Paper* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3545574.

Table 6 The relative performance of the approximation strategies. Standard errors are reported in parentheses.

	Performance relative to the optimal solution		
(m,n)	Conservative	LP Relaxation	Cut-off
$(50,10)$	12.5% (0.6%)	79.8% (0.6%)	97.6% (0.1%)
$(50,15)$	10.8% (0.4%)	73.3% (0.6%)	97.0% (0.1%)
$(50,20)$	9.1% (0.4%)	70.4% (0.5%)	96.5% (0.1%)
$(50,25)$	9.1% (0.5%)	72.1% (0.5%)	96.0% (0.1%)
$(100,10)$	7.4% (0.3%)	82.4% (0.4%)	99.0% (0.1%)
$(150,10)$	5.3% (0.2%)	85.0% (0.3%)	99.3% (0.1%)
$(200,10)$	4.4% (0.2%)	85.4% (0.3%)	99.6% (0.1%)

Table 7 The performance guarantee of approximation strategies. Standard errors are reported in parentheses.

	Theoretical performance guarantee	
(m,n)	Conservative	Cut-off
$(50,10)$	2.2% (0.2%)	19.4% (0.3%)
$(50,15)$	1.9% (0.1%)	19.0% (0.3%)
$(50,20)$	1.7% (0.1%)	18.6% (0.3%)
$(50,25)$	1.9% (0.2%)	18.9% (0.3%)
$(100,10)$	1.0% (0.1%)	16.8% (0.2%)
$(150,10)$	0.6% (0.0%)	15.4% (0.2%)
$(200,10)$	0.5% (0.0%)	14.8% (0.2%)

Appendix A: Additional Tables and Figures

Table 6 shows the percentage performance ratio of the conservative, LP relaxation, and cut-off pricing as explained in Section 6.1. All the instances were solved to optimality using Gurobi 9.0.0 in Python with a desktop computer (Intel Core i7-8700, 3.2 GHz). Similarly, Table 7 portrays the average performance guarantee of conservative and cut-off pricing approximation algorithms.

Appendix B: Proofs and Supplementary Results

Proof of Proposition 1. The linear program that results from removing the disjunctive constraint (2) from (DP), i.e.,

$$\min_{r,\mathbf{v}_i}\{r\colon r \geq 0, \mathbf{v}_i \in \mathscr{V}_i\},$$

has the trivial finite optimum $r=0$. Thus, the model (DP-LP) is obtained by applying directly Corollary 2.1.2 by Balas (1998). □

Proof of Lemma 1. We begin with statement (a). If $p_{c_i} \geq P_{ic_i}$, then the valuation $\mathbf{v}_i^{\varnothing}$ defined by $v_{ic_i}^{\varnothing} = P_{ic_i}$ and $v_{ij}^{\varnothing} = 0$ for all $j \in \mathscr{P} \setminus \{c_i\}$ belongs to $\mathscr{W}_i^{\varnothing}(\mathbf{p})$. Conversely, if $\mathbf{v}_i^{\varnothing} \in \mathscr{W}_i^{\varnothing}(\mathbf{p})$, the inequalities $v_{ij}^{\varnothing} \leq p_j$ from (10) and $v_{ij}^{\varnothing} \geq P_{ic_i}$ from (1) together imply $p_{c_i} \geq P_{ic_i}$.

For statement (b), the valuation $\mathbf{v}_i^{c_i}$ defined by $v_{ic_i}^{c_i} = \max\{P_{ic_i}, p_{c_i}\}$ and $v_{ij}^{c_i} = 0$ for all $j \in \mathscr{P} \setminus \{c_i\}$ belongs to $\mathscr{W}_i^{c_i}(\mathbf{p})$.

Finally, for statement (c), consider the set of constraints that are satisfied by points in $\mathscr{W}_i^j(\mathbf{p})$ after rearranging the constant terms to the right-hand side of the inequalities:

$$v_{ij}^j \geq p_j, \tag{31}$$

$$v_{ij}^j - v_{ij'}^j \geq p_j - p_{j'}, \qquad \forall j' \in \mathscr{P}, \tag{32}$$

$$v_{ic_i}^j \geq P_{ic_i}, \tag{33}$$

$$v_{ic_i}^j - v_{ij'}^j \geq P_{ic_i} - P_{ij'}, \qquad \forall j' \in \mathscr{P} \setminus \{c_i\}. \tag{34}$$

If $p_j - p_{c_i} \leq P_{ij} - P_{ic_i}$, we construct a valuation $\mathbf{v}_i^j \in \mathscr{W}_i^j(\mathbf{p})$ where $v_{ij}^j = \max\{P_{ij}, p_j\}$, $v_{ic_i}^j = \max\{P_{ic_i}, P_{ic_i} - P_{ij} + p_j\}$, and $v_{ij'}^j = 0$ for all $j' \in \mathscr{P} \setminus \{c_i, j\}$. In particular, (31) and (33) are satisfied by construction. Assume now $p_j > P_{ij}$. For (32) and (34) with $j' = c_i$, we have

$$v_{ij}^j - v_{ic_i}^j = p_j - P_{ic_i} + P_{ij} - p_j = P_{ij} - P_{ic_i} \geq p_j - p_{c_i},$$

where the last inequality follows from the statement hypothesis. For $j' \neq c_i$, note that $v_{ij}^j - p_j$ is zero while $v_{ij'}^j - p_{j'}$ is non-positive in (32), and analogously $v_{ic_i}^j - P_{ic_i}$ is positive while $v_{ij'}^j - P_{ij'} = -P_{ij'}$ is non-positive. If $p_j \leq P_{ij}$, note that $v_{ij}^j = P_{ij}$ and $v_{ic_i}^j = P_{ic_i}$, and the same derivations above apply.

Finally, the sufficient conditions of (c) follow directly from (32) and (34) with $j' = c_i$ in (32) and $j' = j$ in (34). □

Proof of Proposition 2. Let $j \in \mathscr{P}$ and denote by G the set of feasible solutions to (DP-LP). The projection of G onto variable x_j is

$$\begin{aligned}
\text{Proj}_{x_j} G &= \{x_j \in [0,1] \colon \exists\ ((\mathbf{v}_i^1, \ldots, \mathbf{v}_i^n, \mathbf{v}_i^\varnothing), (x_1, \ldots, x_j, \ldots, x_n, x_\varnothing)) \in G\} \\
&= \{x_j \in [0,1] \colon\ \nexists \mathbf{v}_i^j \in \mathscr{W}_i^j(\mathbf{p}) \Rightarrow x_j = 0\} \\
&= \{x_j \in [0,1] \colon (p_j - p_{c_i} > P_{ij} - P_{ic_i}) \Rightarrow x_j = 0\} \\
&= \{x_j \in [0,1] \colon x_j \leq \mathbb{I}(p_j - p_{c_i} \leq P_{ij} - P_{ic_i})\},
\end{aligned}$$

where the second-to-last equality follows from Lemma 1. The same arguments follow for $x_\varnothing$. Since the objective is defined only in terms of $\mathbf{x}$, we can replace the inequalities of G by the projections depicted above, which results in the equivalent formulation

$$f_i(\mathbf{p}) = \min_{\mathbf{x} \geq 0} \quad \sum_{j \in \mathscr{P}} p_j x_j \tag{35}$$

$$\text{s.t.} \quad \sum_{j \in \mathscr{P}} x_j + x_\varnothing = 1, \tag{36}$$

$$x_\varnothing \leq \mathbb{I}(p_{c_i} \geq P_{ic_i}), \tag{37}$$

$$x_j \leq \mathbb{I}(p_j - p_{c_i} \leq P_{ij} - P_{ic_i}), \qquad \forall j \in \mathscr{P}. \tag{38}$$

Finally, note from (36) and (37) that

$$\sum_{j \in \mathscr{P}} x_j = 1 - x_\varnothing \geq 1 - \mathbb{I}(p_{c_i} \geq P_{ic_i}) = \mathbb{I}(p_{c_i} < P_{ic_i})$$

must hold at any feasible solution, and particularly tight at optimality since it is the only constraint that bounds $\mathbf{x}$ from below besides the non-negativity conditions. □

Proof of Proposition 3. Statement (a) follows from the fact that product c_i is always feasible (Lemma 1-(b)) and that the inequaliy (13) for $j = c_i$ holds with $\mu^*_{ic_i} = 0$ at optimality. For statement (b), if $f_i(\mathbf{p}) = 0$, the condition trivially holds. Assume $f_i(\mathbf{p}) > 0$. Let $P' = \max_{\{j \in \mathscr{P}: p_j < P^{\max}\}} p_j$. It can be easily shown that P' always exists. Moreover, let $j' \in \mathscr{P}$ be a product such that $p_{j'} \geq P^{\max}$. If $j' = c_i$, then $f_i(\mathbf{p}) = 0$ which is a contradiction. Thus, $p_{c_i} < P^{\max}$. If $j' \neq c_i$, $f_i(\mathbf{p}) < p_{j'}$ because of (a) and $p_{c_i} < p_{j'}$. Reducing the price of j' to P' therefore does not change the optimal value of (DP-C-Dual), and the same argument can be repeated for other products. □

Proof of Proposition 4. It suffices to show that both optimal solution values match when conditioned to a fixed $\mathbf{p} \geq 0$. First, by Proposition 3-(b), we can restrict our analysis to $p_j < P^{\max}$ for all $j \in \mathscr{P}$ without loss of generality.

Consider any customer $i \in \mathscr{C}$. If $p_{c_i} \geq P_{ic_i}$, then we must have $y_{ic_i} = 0$ in (OP-B); otherwise, we can assume $y_{ic_i} = 1$ since that can only be benefitial to the objective. Thus, the objective functions of both models match.

Suppose now that $p_j - p_{c_i} \leq P_{ij} - P_{ic_i}$ for $j \neq c_i$, i.e., product j is feasible to purchase by customer i. We necessarily must have $y_{ij} = 1$ because of (18) and, thus, (15) and (16) match. If otherwise $p_j - p_{c_i} > P_{ij} - P_{ic_i}$, then we may have either $y_{ij} = 0$ or $y_{ij} = 1$. Since the objective of (OP-B) maximizes τ_i, we can assume $y_{ij} = 0$, which can only relax the bound on τ_i in (16). Thus, (15) and (16) also match in this case, i.e., at optimality, the values of τ_i (and hence the optimal values of both models) are the same. □

Proof of Theorem 1. We first show (a). First, (OP-ϵ) with $\epsilon = 0$ is always feasible. That is, for any $\mathbf{p} \geq 0$ and customer $i \in \mathscr{C}$, we set $y_{ic_i} = 1$ if and only if $p_{c_i} \leq P_{ic_i}$, for all $j \in \mathscr{P} \setminus \{c_i\}$, $y_{ij} = 0$ if and only if $p_j - p_{c_i} \geq P_{ij} - P_{ij}$, and τ_i appropriately to satisfy (20). Thus, it remains to show that $g(0)$ is bounded from above. This follows from noting that, for all $i \in \mathscr{C}$, p_{c_i} is bounded by $P_{ic_i} + P^{\max}$ in inequality (21) and that τ_i is bounded by p_{c_i} in inequality (20) with $j = c_i$.

For (b), let $(\mathbf{p}^0, \tau^0, \mathbf{y}^0)$ be an optimal solution tuple and consider any $j' \in \mathscr{P}$ such that $p^0_{j'} = 0$. We show an alternative feasible solution with the same (optimal) value after increasing $p^0_{j'}$ as in the statement. If $y^0_{ic_i} = 0$ for some customer i, then increasing $p^0_{j'}$ does not affect τ_i nor the final solution value, given that y^0_{ij} is adjusted appropriately for $j \neq c_i$ to ensure feasibility. If otherwise $y^0_{ic_i} = 1$ for some customer i, we have two cases:

1. Case 1, $\tau^0_i = 0$. In such a scenario, we can equivalently set $y^0_{ic_i} = 0$ and apply the same adjustments to y^0_{ij} for all $j \neq c_i$ as above, preserving the solution value.
2. Case 2, $\tau^0_i > 0$. Due to inequality (20), we must have $j' \neq c_i$ and $y^0_{ij'} = 0$. Thus, $p^0_{j'} - p^0_{c_i} \geq P_{ij'} - P_{ic_i}$ from inequality (22). Increasing $p^0_{j'}$ thefore just increases the left-hand side of such inequality, and hence does not impact feasibility nor the solution value.

We now show (c). Since the feasible set of (OP-ϵ) relaxes and restricts that of (OP-B) for $\epsilon = 0$ and $\epsilon > 0$, respectively, the inequality

$$0 \leq g(0) - m\tau^* \leq g(0) - \sum_{i \in \mathscr{C}} f_i(\mathbf{p}')$$

follows directly. We will now show that

$$g(0) - \sum_{i \in \mathscr{C}} f_i(\mathbf{p}') \leq \delta' \leq \delta.$$

Consider the ordered vector of prices $\mathbf{p}^0 > 0$ in the statement and the associated tuple $(\mathbf{p}^0, \tau^0, \mathbf{y}^0)$. Let $i \in \mathscr{C}$ be a customer such that $y^0_{ic_i}\tau^0_i > 0$. The inequality (20) is tight for some $j' \leq c_i$, i.e., we can have $y^0_{ij} = 0$ for all $j < j'$. This implies that $p^0_j - p^0_{c_i} \geq P_{ij} - P_{ic_i}$ for those indices due to (22).

Next, for ease of notation, let $\sigma = \delta'/(mn)$ so that $\mathbf{p}' = (p^0_1 - \sigma, p^0_2 - 2\sigma, \ldots, p^0_n - n\sigma)$. For $j < j'$,

$$p'_j - p'_{c_i} = p^0_j - p^0_{c_i} + (c_i - j)\sigma > P_{ij} - P_{ic_i},$$

since $j < c_i$. Thus, when evaluating $f_i(\mathbf{p}')$, the constraints (14) in (DP-D) for $j < j'$ remains non-binding, i.e.,

$$\tau_i \leq p'_j + \mathbb{I}(p'_j - p'_{c_i} > P_{ij} - P_{ic_i})P^{\max} = p'_j + P^{\max}$$

for $j < j'$. Thus, since the price of any product $j \geq j'$ is decreased by $j\sigma \leq n\sigma$, evaluating the new price vector $\mathbf{p}'$ in (DP-D) yields: for all $i \in \mathscr{C}$,

$$f_i(\mathbf{p}') \geq \mathbb{I}(p'_{c_i} < P_{ic_i})(\tau^0_i - n\sigma) = \mathbb{I}(p^0_{c_i} - c_i\sigma < P_{ic_i})(\tau^0_i - n\sigma) \geq y^0_{ic_i}\tau^0_i - n\sigma,$$

where the last inequality follows from the fact that inequality (21) implies $p^0_{c_i} \leq P_{ic_i}$, which implies $p^0_{c_i} - c_i\sigma < P_{ic_i}$. Finally, summing the above inequality over all i, we have:

$$\sum_{i \in \mathscr{C}} f_i(\mathbf{p}') \geq \sum_{i \in \mathscr{C}} (y^0_{ic_i}\tau^0_i - n\sigma) = g(0) - \delta',$$

concluding the proof. □

Proof of Proposition 5. We first show by contradiction that all prices are the same at optimality. To this end, note from inequality (22) with $\epsilon = 0$ that, for any $i \in \mathscr{C}$, if a product $j \neq c_i$ is *not* eligible for purchase (i.e., $y_{ij} = 0$) then $p_j \geq p_{c_i}$, since all historical prices for i are the same. Consider now an optimal solution $(\mathbf{p}^*, \tau^*, \mathbf{y}^*)$ and let $p^{\min} \equiv \min_{j \in \mathscr{P}} p^*_j$ be the minimum optimal price. Suppose there exists some customer $i \in \mathscr{C}$ who selects a product j such that $\tau^*_i = p^*_j > p^{\min}$. But this implies that $p^{\min} < p^*_j \leq p^*_{c_i}$ (since c_i is always feasible), i.e., we must have $y^*_{ij} = 1$ which by inequality (20) implies that $\tau^*_i \leq p^{\min}$, a contradiction.

Finally, let p^* be the optimal (scalar) price of all products, and suppose $i' \in \mathscr{C}$ is the smallest customer index such that $p^* \leq P_{i'}$. It follows that any customer $i < i'$ does not purchase any product (since $p^* = p^*_{c_i} > P_i$), while all customers $i \geq i'$ yield a revenue of p^* (since $p^* = p^*_{c_i} \leq P_i$). Thus, we must have $p^* = P_{i'}$ at optimality, and the total revenue is $(m - i' + 1)P_{i'}$. The element i^* in the proposition statement is the index that maximizes this revenue. □

Proof of Proposition 6. The solution value of the proposed solution (i.e., a lower bound to the problem) is $\sum_{i \in \mathscr{C}} P_{c_i}$, since all customers would purchase their choice c_i. Due to inequality (20), this is also an upper bound, and hence is optimal. □

Proof of Proposition 7. By definition (28), the total revenue is at least $m\underline{P}$. Conversely, by inequalities (20) and (21), the total revenue is at most $\sum_{i\in\mathscr{C}} P_{ic_i} \le m\bar{P}$. The ratio hence follows from dividing the lower bound by the upper bound.

We now construct an instance where this ratio is asymptotically tight. Consider $n=1$ product, $m>1$ customers, and fix any P_1, P_2 such that $P_1 < P_2$. Customer 1 purchases the product at price P_1, while the remaining customers $2,\dots,m$ purchase it at P_2. The conservative pricing (28) will set the product's price at P_1, yielding a total revenue of mP_1. For any sufficiently large m, the optimal pricing strategy sets P_2 as the optimal price, yielding a total revenue of $(m-1)P_2$ (since customer 1 will not purchase any product). The performance ratio is therefore $mP_1/(m-1)P_2$. Taking the limit $m\to+\infty$ with respect to this ratio completes the proof. □

Proof of Proposition 8. Without loss of generality, suppose the customer index set $\mathscr{C}$ is ordered according to historical purchase prices, i.e., $0<\underline{P}=P_{1c_1}\le P_{2c_2}\le\cdots\le P_{mc_m}=\bar{P}$. Let τ^{OPT} and τ^{CP} be the optimal solution value of (OP-MIP) and the total revenue obtained by the cut-off price (30), respectively. It follows from (20) for $j=c_i$ that $\tau^{\mathrm{OPT}}\le\sum_{i\in\mathscr{C}} P_{ic_i}$. Furthermore, given the price ordering, notice that (29) evaluates to $\max_{i\in\mathscr{C}}(m-i+1)P_{ic_i}$. By the cut-off pricing definition (30), we therefore have $\tau^{\mathrm{CP}}\ge\max_{i\in\mathscr{C}}(m-i+1)P_{ic_i}$. This is because due to the incentive-compatibility constraint (22) some customers might not purchase the products priced at the cut-off price p^* and opt for a product that is priced higher than p^*. Thus,

$$\frac{\tau^{\mathrm{CP}}}{\tau^{\mathrm{OPT}}}\ge\frac{\max_{i\in\mathscr{C}}(m-i+1)P_{ic_i}}{\sum_{i\in\mathscr{C}} P_{ic_i}}.$$

Next, if we divide both the numerator and denominator of the right-hand side ratio above by the number of customers $m\ge1$, we obtain

$$\frac{\max_{i\in\mathscr{C}}(m-i+1)P_{ic_i}}{\sum_{i\in\mathscr{C}} P_{ic_i}}=\frac{\dfrac{\max_{i\in\mathscr{C}}(m-i+1)P_{ic_i}}{m}}{\dfrac{\sum_{i\in\mathscr{C}} P_{ic_i}}{m}}=\frac{\max_{i\in\mathscr{C}}[1-F_{\mathcal{X}}(P_{ic_i})]P_{ic_i}}{\mathbb{E}[\mathcal{X}]},$$

where $\mathcal{X}$ is a non-negative discrete random variable uniformily distributed on the set $\{P_{1c_1},\dots,P_{mc_m}\}$, and $F_{\mathcal{X}}(\cdot)$ and $\mathbb{E}[\mathcal{X}]$ denote the left continuous c.d.f. (i.e.,$F_{\mathcal{X}}(x)\equiv P(\mathcal{X}<x)$) and the expectation of $\mathcal{X}$, respectively. This problem now bears resemblance to the personalized pricing problem studied in Elmachtoub et al. (2020).

Let $R\equiv\max_{i\in\mathscr{C}}[1-F_{\mathcal{X}}(P_{ic_i})]P_{ic_i}$ be the numerator of the ratio above. We have $\underline{P}\le R\le\bar{P}$; in particular, the left-hand side inequality holds since $[1-F_{\mathcal{X}}(\underline{P})]\underline{P}=\underline{P}$. We can hence rewrite the expectation term as

$$\begin{aligned}\mathbb{E}[\mathcal{X}]=\sum_{i\in\mathscr{C}}\frac{1}{m}P_{ic_i}&=\sum_{i\in\mathscr{C}}\frac{1}{m}\left[\int_0^{P_{ic_i}}1\,dx\right]\\&=\sum_{i\in\mathscr{C}}\frac{1}{m}\left[\int_0^{P_{ic_i}}1\,dx+\int_{P_{ic_i}}^{+\infty}0\,dx\right]\\&=\sum_{i\in\mathscr{C}}\frac{1}{m}\left[\int_0^{+\infty}\mathbb{I}(x\le P_{ic_i})\,dx\right]\\&=\int_0^{+\infty}\sum_{i\in\mathscr{C}}\left[\frac{1}{m}\mathbb{I}(x\le P_{ic_i})\right]dx\end{aligned}$$

$$
\begin{aligned}
&= \int_0^{+\infty} [1 - F_{\mathcal{X}}(x)]\, dx \\
&= \int_0^{\bar{P}} [1 - F_{\mathcal{X}}(x)]\, dx \\
&= \int_0^{R} [1 - F_{\mathcal{X}}(x)]\, dx + \int_R^{\bar{P}} [1 - F_{\mathcal{X}}(x)]\, dx \\
&\le R + \int_R^{\bar{P}} [1 - F_{\mathcal{X}}(x)]\, dx \\
&\le R + \int_R^{\bar{P}} \frac{R}{x}\, dx \\
&= R + R \log\left(\frac{\bar{P}}{R}\right) \\
&\le R + R \log\left(\frac{\bar{P}}{\underline{P}}\right),
\end{aligned}
$$

where the previous-to-the-last inequality follows because $1 - F_{\mathcal{X}}(x) \le R/x$ for any $x \in [\underline{P}, \bar{P}]$ by definition. Finally, from the inequality above, we obtain

$$
\frac{\mathbb{E}[\mathcal{X}]}{R} \le 1 + \log\left(\frac{\bar{P}}{\underline{P}}\right) \Leftrightarrow \frac{\tau^{\mathrm{CP}}}{\tau^{\mathrm{OPT}}} \ge \frac{R}{\mathbb{E}[\mathcal{X}]} \ge \frac{1}{1 + \log\left(\frac{\bar{P}}{\underline{P}}\right)},
$$

proving the performance ratio.

It remains to show that the ratio is asymptotically tight. Consider an instance with any number $m > 0$ of customers and $m - k + 1$ products, where $k \in \mathbb{N}$ is any positive integer such that $k \le m$. The historical observed prices $\mathbf{P}_i$ by customer i, in turn, are defined by the following vectors, where δ is any scalar such that $0 < \delta < 1/(m-1)$:

$$
\begin{aligned}
\mathbf{P}_1 &= \left(\frac{m}{m}, \frac{m}{m-1}, \frac{m}{m-2}, \ldots, \frac{m}{k+1}, \frac{m}{k}\right), \\
\mathbf{P}_2 &= \left(\frac{m}{m} + \delta, \frac{m}{m-1}, \frac{m}{m-2}, \ldots, \frac{m}{k+1}, \frac{m}{k}\right), \\
\mathbf{P}_3 &= \left(\frac{m}{m} + \delta, \frac{m}{m-1} + \delta, \frac{m}{m-2}, \ldots, \frac{m}{k+1}, \frac{m}{k}\right), \\
&\ldots \\
\mathbf{P}_{m-k} &= \left(\frac{m}{m} + \delta, \frac{m}{m-1} + \delta, \frac{m}{m-2} + \delta, \ldots, \frac{m}{k+1}, \frac{m}{k}\right), \\
\mathbf{P}_{m-k+1} &= \left(\frac{m}{m} + \delta, \frac{m}{m-1} + \delta, \frac{m}{m-2} + \delta, \ldots, \frac{m}{k+1} + \delta, \frac{m}{k}\right), \\
\mathbf{P}_{m-k+2} &= \left(\frac{m}{m} + \delta, \frac{m}{m-1} + \delta, \frac{m}{m-2} + \delta, \ldots, \frac{m}{k+1} + \delta, \frac{m}{k}\right), \\
&\ldots \\
\mathbf{P}_m &= \left(\frac{m}{m} + \delta, \frac{m}{m-1} + \delta, \frac{m}{m-2} + \delta, \ldots, \frac{m}{k+1} + \delta, \frac{m}{k}\right).
\end{aligned}
$$

For the historical purchase choices, each customer $i \in \{1, \ldots, m-k\}$ purchased product $c_i = i$, while all remaining customers $i' \in \{m-k+1, \ldots, m\}$ pick the same product $c_{i'} = m - k + 1$.

An upper bound on the optimal revenue τ^{OPT} for this instance is:

$$
\tau^{\mathrm{OPT}} \le \sum_{i \in \mathscr{C}} P_{ic_i} = \sum_{i=1}^{m-k+1} P_{ii} + \sum_{i=m-k+2}^{m} P_{i(m-k+1)} = \sum_{i=0}^{m-k} \frac{m}{m-i} + \sum_{i=m-k+1}^{m-1} \frac{m}{k} = \sum_{i=0}^{m-k} \frac{m}{m-i} + m\frac{k-1}{k}.
$$

We now define prices whose resulting total revenue is arbitrarily close to τ^{OPT}. Specifically, consider the price vector

$$\mathbf{p} = \left(\frac{m}{m}, \frac{m}{m-1} - \delta, \frac{m}{m-2} - 2\delta, \ldots, \frac{m}{k} - (m-k)\delta\right).$$

From the definition above, since $p_{c_i} \le P_{ic_i}$ for all $i \in \mathscr{C}$, by inequality (21) every customer purchases a product. We next show that customer i purchases product c_i. For any $j < c_i$,

$$p_j - p_{c_i} = \frac{m}{m-j+1} - \frac{m}{m-c_i+1} - j\delta + c_i\delta \ge P_{ij} - P_{ic_i},$$

i.e., such products j violates incentive-compatibility constraints and will not be purchasable by customer i. Moreover, from our choice of $\delta < 1/(m-1)$, it can be easily verified that $p_1 < p_2 < \cdots < p_{m-k+1}$ and therefore $p_j > p_{c_i}$ for all i and $j > c_i$. Product p_{c_i} must be necessarily chosen by customer i under her worst-case valuation ($v_{ic_i} = P_{ic_i}$ and $v_{ik} = 0$ for all $k \in \mathscr{P} \setminus \{c_i\}$) and the revenue from this pricing is hence

$$\sum_{i \in \mathscr{C}} p_{c_i} = \sum_{i=1}^{m-k+1} p_i + (k-1)p_{m-k+1} = \sum_{i=0}^{m-k}\left(\frac{m}{m-i} - i\delta\right) + m\frac{k-1}{k} - (k-1)(m-k)\delta,$$

thus, as $\delta \to 0$, the total revenue obtained from $\mathbf{p}$ approximates that of τ^{OPT}.

We now show that the revenue obtained from the cut-off pricing (30) is m. Suppose that the customer index that solves (29) is i', and hence $p^* = P_{i'c_{i'}}$. If $c_{i'} < m-k+1$, then the cut-off prices are $p_j^{CP} = \frac{m}{k}$ for all $j < c_{i'}$, $p_j^{CP} = \frac{m}{m-j+1}$ for all $c_{i'} \le j < m-k+1$ and $p_{m-k+1}^{CP} = \frac{m}{k}$. However, if $c_{i'} = m-k+1$, the cut-off prices are $p_j^{CP} = \frac{m}{k}$ for all $j \in \mathscr{P}$. Suppose that $c_{i'} < m-k+1$. By the construction of historical prices,

$$P_{i'c_{i'}} = \frac{m}{m-c_{i'}+1},$$

and therefore $m - c_{i'} + 1$ customers would purchase a product since $P_{ic_i} \ge P_{i'c_{i'}}$ for all $i \ge i'$. Analogously, since by construction for all customers $i < i'$, $P_{ic_i} < P_{i'c_{i'}}$, none of the worst-case historical customers $i < i'$ would purchase any product. Moreover, for any of the worst-case customer $i \ge i'$, the cut-off price (30) of its chosen product c_i is

$$p_{c_i}^{\mathrm{CP}} = \frac{m}{m-c_i+1}.$$

This implies that, for any $i > i'$,

$$p_{c_{i'}}^{\mathrm{CP}} - p_{c_i}^{\mathrm{CP}} = \frac{m}{m-c_{i'}+1} - \frac{m}{m-c_i+1} < \left(\frac{m}{m-c_{i'}+1} + \delta\right) - \frac{m}{m-c_i+1} = P_{i'c_{i'}} - P_{ic_i},$$

i.e., product $c_{i'}$ is incentive-compatible with all customers $i > i'$. Because $p_{c_{i'}}^{\mathrm{CP}}$ is the lowest price across all the products, the total revenue of the cut-off solution

$$\tau^{\mathrm{CP}} = (m-c_{i'}+1)P_{i'c_{i'}} = (m-c_{i'}+1)\frac{m}{m-c_{i'}+1} = m$$

If $c_{i'} = m-k+1$, as mentioned before the cut-off prices are $p_j^{CP} = \frac{m}{k}$ for all $j \in \mathscr{P}$ and it can also be easily confirmed that $\tau^{\mathrm{CP}} = m$ as only the k historical customers who historically purchased product $m-k+1$ will make a purchase, under price $\frac{m}{k}$, in the worst-case.

Finally, as $\delta \to 0$,

$$\frac{\tau^{\mathrm{CP}}}{\tau^{\mathrm{OPT}}} \to \frac{m}{\sum_{i=0}^{m-k} \frac{m}{m-i} + m\frac{k-1}{k}} = \frac{m}{m\left(\sum_{i=0}^{m-k} \frac{1}{m-i} + \frac{k-1}{k}\right)}$$
$$= \frac{1}{\frac{1}{m} + \cdots + \frac{1}{k} - \frac{1}{k} + \frac{k}{k}}$$
$$= \frac{1}{\sum_{i=k+1}^{m} \frac{1}{i} + \frac{k}{k}}.$$

For a sufficiently large m and k (e.g., by multiplying both by the same constant), the ratio above can be made sufficiently close to

$$\frac{1}{\log m - \log k + \frac{1}{2m} - \frac{1}{2k} + 1} = \frac{1}{\log\left(\frac{m}{k}\right) + \frac{1}{2m} - \frac{1}{2k} + 1},$$

where the last equality follows from the fact that for large enough n, $\sum_{i=1}^{n} \frac{1}{i} = \log(n) + \gamma + \frac{1}{2n}$ where γ is the Euler–Mascheroni constant. The ratio above can approximate $1/(1+\log(\bar{P}/\underline{P}))$ at any desired precision since m/k can be made sufficiently close to $\bar{P}/\underline{P}$, while both numbers are also sufficiently large. Thus, the ratio is asymptotically tight for the constructed instance. □

Proposition 9. *If prices are set to their average historical values, the fraction between the revenue obtained with this approach and the optimal value w.r.t.* (OP-MIP) *is asymptotically zero as the number of historial customers grow.*

Proof of Proposition 9. Assume we have $n = 1$ product and $m > 1$ historical customers where the purchase price of customers $1, \cdots, (m-1)$, was 1, and the purchase price of customer m was 2. Then, the average historic price of the product is $\frac{m+1}{m}$. At this price, customers $1, \cdots, (m-1)$ will not make a purchase under their worst-case valuations while customer m will purchase the product at price $\frac{m+1}{m}$. However, setting the price of the product at 1 results in a revenue of m, since all customers will purchase the product. Thus, the ratio of the revenue from the average price to the optimal value of (OP-MIP) could be less than or equal to $\frac{m+1}{m^2}$. Taking the limit $m \to +\infty$ with respect to this ratio completes the proof. □

Proposition 10. *If prices are set uniformly at random based on the empirical distribution defined by historical prices, the fraction between the revenue obtained with this approach and the optimal value w.r.t.* (OP-MIP) *is asymptotically zero as the number of historial customers grow.*

Proof of Proposition 10. Assume we have n products and n historical customers. The historical prices $\mathbf{P}_i$ observed by customer i, are defined by the following vectors:

$$\mathbf{P}_1 = (1, 2, 2, \ldots, 2),$$
$$\mathbf{P}_2 = (2, 1, 2, \ldots, 2),$$
$$\mathbf{P}_3 = (2, 2, 1, 2, \ldots, 2),$$
$$\ldots$$
$$\mathbf{P}_n = (2, 2, \ldots, 2, 1).$$

For the historical purchase choices, each customer $i \in \{1, \ldots, n\}$ purchased product i.

Assume the vector of prices p is equal to one of $\mathbf{P}_1, \ldots, \mathbf{P}_n$ at random (with equal probability). If $p = \mathbf{P}_1$, then for $i = 1$, we have $P_{1c_1} \leq p_{c_1}$ and hence the revenue from customer 1 is 1. For any customer $i > 1$, we have $P_{ic_i} < p_{c_i}$, hence the revenue form any customer $i > 1$ is zero. Similarly we can show that if $p = \mathbf{P}_i$ for any $i > 1$, the total revenue will be 1, under the worst-case customer valuations.

However, for $p = (1, 1, \ldots, 1)$, the revenue will be n as every customer will make a purchase at price 1. Therefore, the ratio of the expected revenue from the price observed by a random historical customer to the optimal value of (OP-MIP) could be less than or equal to $\frac{1}{n}$. Taking the limit $n \to +\infty$ with respect to this ratio completes the proof. $\square$