Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Statistical Learning Approach to Personalization in Revenue Management

Xi Chen, Zachary Owen,, Clark Pixton, David Simchi-Levi

To cite this article:
Xi Chen, Zachary Owen,, Clark Pixton, David Simchi-Levi (2021) A Statistical Learning Approach to Personalization in Revenue Management. Management Science

Published online in Articles in Advance 18 Jan 2021

. https://doi.org/10.1287/mnsc.2020.3772

http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–15
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

A Statistical Learning Approach to Personalization in Revenue Management

Xi Chen,[a] Zachary Owen, Clark Pixton,[b] David Simchi-Levi[c]

[a] Leonard N. Stern School of Business, New York University, New York, New York 10012-1126; [b] Marriott School of Business, Brigham Young University, Provo, Utah 84602; [c] Institute for Data, Systems, and Society, Department of Civil and Environmental Engineering and Operations Research Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142

Contact: xc13@stern.nyu.edu, https://orcid.org/0000-0002-9049-9452 (XC); zowen@alum.mit.edu, https://orcid.org/0000-0002-3962-2654 (ZO); cpixton@byu.edu, https://orcid.org/0000-0002-5501-0367 (CP); dslevi@mit.edu, https://orcid.org/0000-0002-4650-1519 (DS-L)

Received: May 11, 2019
Revised: December 26, 2019; May 1, 2020
Accepted: June 23, 2020
Published Online in Articles in Advance: January 18, 2021

https://doi.org/10.1287/mnsc.2020.3772

Abstract. We consider a logit model-based framework for modeling joint pricing and assortment decisions that take into account customer features. This model provides a significant advantage when one has insufficient data for any one customer and wishes to generalize learning about one customer's preferences to the population. Under this model, we study the statistical learning task of model fitting from a static store of precollected customer data. This setting, in contrast to the popular learning and earning paradigm, represents the situation many business teams encounter in which their data collection abilities have outstripped their data analysis capabilities. In this learning setting, we establish finite-sample convergence guarantees on the model parameters. The parameter convergence guarantees are then extended to out-of-sample performance guarantees in terms of revenue, in the form of a high-probability bound on the gap between the expected revenue of the best action taken under the estimated parameters and the revenue generated by a decision maker with full knowledge of the choice model. We further discuss practical implications of these bounds. We demonstrate the personalization approach using ticket purchase data from an airline carrier.

History: Accepted by J. George Shanthikumar, special issue on data-driven prescriptive analytics.
Funding: X. Chen is supported by the National Science Foundation [Grant IIS-1845444]. The work of Z. Owen, C. Pixton, and D. Simchi-Levi was partially sponsored by the Massachusetts Institute of Technology Data Science Laboratory, a laboratory focused on the development of analytic techniques and tools for improving decision making in environments that involve uncertainty and require statistical learning.
Supplemental Material: The online appendix is available at https://doi.org/10.1287/mnsc.2020.3772.

Keywords: assortment optimization • personalization • price discrimination • statistical learning

1. Introduction

The increasing prominence of electronic commerce has given businesses an unprecedented ability to understand their customers as individuals and to tailor their services for them accordingly. This benefit is twofold: customer profiles and data repositories often provide information that can be used to predict which products and services are most relevant to a customer, and the fluid nature of electronic services allows for this information to be used to optimize their experience in real time (see Murthi and Sarkar 2003). For instance, Linden et al. (2003) document how Amazon.com has used personalization techniques to optimize the selection of products it recommends to users for many years, dramatically increasing click-through and conversion rates as compared with static sites. Other companies, such as Netflix, have implemented personalization through recommender systems as described by Amatriain (2013) to drive revenue indirectly by improving customer experience.

In this paper, we study the application of personalization strategies to problems in revenue management, specifically assortment optimization and pricing. A key factor of a successful assortment or pricing strategy is the ability to understand and predict customer demand or preferences. In many online applications, a firm will have enough customer-specific data (i.e., purchase history) to model the preferences of each customer individually. However, the e-commerce business trend has also led to a faster-paced, more uncertain business environment, with accelerated product life cycles and higher rates of change in customer base. When dealing with new products or new customers, feature information is useful in leveraging knowledge of demand or preferences across customers. In such a setting, the goal is to mathematically understand the relationship between customer or situational *features* and the resulting purchase decisions.

The logit model is by far the most commonly used tool for understanding customer preferences/demand in

practice (Train 2009). It is ubiquitous in both the operations management (OM) and marketing literatures. In the OM context, this model is often used to understand the effects of *product* features on customer preference, but the framework has also been widely used to study the effects of *customer* or *situational* features on choice. For example, logit models have been used as a framework for personalization in various applications, including conjoint analysis (Arora and Huber 2001), targeted advertising (Luo et al. 2013), personalized promotions (Zhang and Krishnamurthi 2004), pricing (Zhang et al. 2014, Xue et al. 2016), and assortment optimization (Golrezaei et al. 2014). The logit model has the advantages of interpretability and simplicity, which are key in many applications. Thus, even with the proliferation of extremely sophisticated and complex statistical models, the logit framework remains a useful tool in practice.

In this work, we analyze a unified logit-based modeling framework that captures personalized decision making for both assortment and pricing decisions. Our analysis is highly relevant to practical settings because we explicitly consider the need to estimate the model from a corpus of past transaction and firm decision data. Based on the classical M-estimation theory, we provide sample complexity results for learning the multinomial logistic model, bounding the number of transactional data points needed to achieve any given level of accuracy. This bound can be applied when data samples represent transactions in which different customers are shown different assortments and in which prices are the result of past decisions. We also give finite-sample results in the setting where data points are dependent across time but satisfy a mixing condition. These results help to build the link between theory and practice, where the offered assortments are highly personalized and data are dependent.

To further tailor the estimation error results for revenue management applications, we provide a high-probability upper bound on the gap between the expected revenue of the proposed method and the *oracle revenue*. Such a bound provides practical insights because it characterizes the trade-off between the cost of information, given by the cost of the minimum number of transactional records to be collected, and the potential revenue loss.

Although we will focus our exposition on familiar problems in revenue management, the analysis technique is not limited to this domain and can be applied in many other contexts that incorporate estimation of personalized preferences, such as online advertisement allocation, crowdsourcing task assignment, and personalized medicine.

2. Literature Review

We focus on two revenue management applications: assortment optimization and pricing. Assortment optimization was brought to the attention of the revenue management community by van Ryzin and Mahajan (1999). Since that time, assortment optimization techniques and models have been heavily researched, with much past work well summarized in Kök et al. (2015). The multinomial logit model (MNL), our focus in this paper, is among the most commonly studied models of customer choice for assortment optimization (Talluri and van Ryzin 2004, Rusmevichientong et al. 2010, and Du et al. 2016 are some examples). Train (2009) gives a good practical and theoretical summary of the logit model in the context of choice modeling. A commonly made assumption in the literature on MNL models is that the utility of each product is linear in the attributes of the product. See Vulcano et al. (2008) and Rusmevichientong et al. (2010) for discussions of this assumption. However, incorporating the attributes of different *customers* is a more recent trend, which we discuss later.

Within the assortment optimization literature, two trends are especially relevant to our work: model estimation and personalization. Recently, researchers have investigated the problem of estimating choice models for assortment optimization. For example, Vulcano et al. (2008) proposed an algorithm for estimating true demand from censored transaction data and proved convergence of their algorithm. Rusmevichientong et al. (2010), Ulu et al. (2012), and Sauré and Zeevi (2013) propose dynamic policies that balance learning customer choice behavior against earning short-term revenue. Rusmevichientong et al. (2010) and Sauré and Zeevi (2013) use the multinomial logit model. None of these papers consider personalization.

Bernstein et al. (2015) approach personalized assortment optimization by establishing one choice model for each customer type and show properties of an optimal assortment policy for two products in the presence of inventory considerations. Golrezaei et al. (2014) also consider multiple customer types, providing a practical algorithm for personalization under inventory constraints and proving a worst-case performance bound (i.e., the competitive ratio) for arbitrary customer arriving sequence.

Turning to our other application area, there has been much literature on pricing recently because of the popularity of dynamic pricing, sometimes with demand learning. A brief survey of some early papers in this field can be found in Aviv and Vulcano (2012). Aydin and Ziya (2009) consider the

case of customized pricing in which customers belong to either a high- or low-reservation price group and provide a signal to the seller that gives some information as to how likely they are to belong to the higher-price group. Netessine et al. (2006) and Aydin and Ziya (2008) each consider a form of personalized dynamic pricing in their treatments of cross-selling, in which the offer to each customer is customized, based on the other items that they have purchased or are considering purchasing. Carvalho and Puterman (2005) study a multistage pricing problem and assume a logit model for demand as a function of the offered price, and they suggest that their model could be extended to include customer-specific attributes. In many cases, models of price discrimination are actually cast as multiproduct models, where the different price levels come with different qualifications and extras as in the airline industry. See Belobaba (1989) and Talluri and van Ryzin (2004) for examples of this type. Additionally, Javanmard and Nazerzadeh (2019) consider a high-dimensional dynamic pricing problem with multiple products, including feature-based customer choice and maximum likelihood estimation (MLE). Qiang and Bayati (2016) also consider dynamic pricing with covariates, showing the optimality of a greedy least-squares method in their setting.

Though in some cases personalized pricing is not recommended because of considerations of customer satisfaction and legality, there are some areas in which it is natural, such as in business to business transactions and even customer transactions where personalization happens on features such as location and time rather than on customer demographics and purchase history. Additionally, personalized pricing has recently found its way into more traditional industries such as retail via personalized discounts and web pricing (Clifford 2012) and airline tickets (made possible by IATA 2014). From a theoretical perspective, Li and Jain (2016) develop a model of behavior-based pricing that, contrary to previous literature, suggests that behavior-based pricing may increase revenues and allow firms to offer discounts to loyal customers, especially when fairness is important to consumers. On the other hand, in situations where personalized pricing is illegal or unwise, our approach is completely valid for pricing applications with nondemographic features, such as time until takeoff in the airline industry, that when taken into account may allow for a better pricing strategy.

In a related paper, Bertsimas and Kallus (2019) develop a method of extending machine learning predictions to actionable prescriptions that takes into account the uncertainty in the estimate of the joint distribution between outcomes and observable information. However, in their paper, the observed outcome depends only on contextual information and not on the decision, whereas in our framework, the action is treated as both historical contextual information and a decision variable. In another related paper, Cohen et al. (2016) consider pricing products that are characterized by a set of features. The value of the product is a linear function of the features, much like how we assume that the utility of a product is linear in the customer features. They analyze an online algorithm for this setting in the case of adversarially chosen customer features. By contrast, our setting is offline and uses previously collected transaction data, and we consider randomly chosen features rather than adversarial.

Other work has treated the problem of pricing an assortment of products subject to customer choice behavior. For example, Li and Huh (2011) demonstrate that under the nested logit model, the total profit function is concave in the vector of market shares that map one to one with pricing decisions, enabling optimal prices to be determined efficiently. Gallego and Wang (2014) generalize to the case of product-specific price sensitivities and demonstrate an optimal markup scheme. In Gallego and Topaloglu (2014), the authors offer a slightly different approach and demonstrate a polynomial time algorithm for selecting optimal prices from a discrete set of candidate prices under the nested logit model. In contrast to these approaches, Jagabathula and Rusmevichientong (2016) employ a nonparametric model and demonstrate a method for estimating the choice model from data and an approximation scheme to solve the resulting price and assortment optimization problem.

Our work is also related to some learning and earning literature with covariate information, where the focus of those papers is mainly on regret analysis. For dynamic pricing, with a parametric demand model, Cohen et al. (2016) establish the regret of $O(d^2 \log(T/d))$ for a binary search algorithm inspired by the ellipsoid method, where d is the dimension of a feature vector and T is the length of the time horizon. With an additional sparsity assumption on features (with the sparsity level s), Ban and Keskin (2020) propose a policy with expected regret $O(s\sqrt{T}(\log d + \log T))$, and Javanmard and Nazerzadeh (2019) develop an online policy with regret $O(s \log d \log T)$. Chen and Gallego (2018) consider a fully parametric demand model and provide an algorithm with the regret $O(T^{(2+d)/(4+d)} \log(T)^2)$. For the assortment optimization problem, Agrawal et al. (2017, 2019) and Chen and Wang (2018) study dynamic assortment optimization without any covariate information and establish the tight regret bound of $O(\sqrt{nT})$, where n is the total number of products. Cheung and Simchi-Levi (2017) and Chen et al. (2018) further incorporate feature information

into dynamic assortment optimization and establish an $\widetilde{O}(d\sqrt{T})$ regret. However, these dynamic assortment efforts do not take the pricing decision into consideration.

Our results build on some classical theorems from statistics, notably *M*-estimation theory. Traditionally, these results have focused on asymptotic results (see, for example, theorem 3.2.16 in van der Vaart and Wellner 2000), though finite-sample bounds are becoming more useful in modern data-driven applications. In the nonindependent feature setting, Banna et al. (2016) provide important results that undergird our analysis.

The rest of the paper proceeds as follows. In Section 3, we present our modeling approach and formalize the algorithm, detailing its application to customized pricing and assortment optimization. Section 4 is devoted to proving revenue bounds for the two problems under various assumptions. The experimental results are provided in Section 5, followed by the conclusion in Section 6.

3. A Unified Logit Model and Estimation Guarantees

In this section, we develop a unified logit modeling framework to integrate decision-specific context information into revenue management problems. We consider a decision maker performing joint pricing and assortment decisions.

In our setting, the decision maker has J potential products to offer (the set of products is denoted by $[J] := \{1, 2, \ldots, J\}$). We include another "no-purchase" option as is common in choice models. Prior to making pricing and assortment decisions for the products, the decision maker observes a customer feature vector $z \in \mathcal{Z}$, where $\mathcal{Z} \subseteq \mathbb{R}^d$ is the space of possible contexts. We assume the z vectors are scaled such that $|z|_\infty \le 1$ for all z. After observing z, the decision maker selects a price p_j for each product j, from the interval $[p_{\min}, p_{\max}]$ with $p_{\min} > 0$. By an appropriate scaling, we can take $p_{\max} = 1$. We use the notation $\boldsymbol{p} = (p_1, \ldots, p_J)$ to denote the vector of current pricing decisions, and we use $\mathcal{P} = [p_{\min}, p_{\max}]^J$ to denote the space of possible pricing decisions. Because the customer will see the result of the pricing decision, the price and the context vector combine to provide a signal of customer choice behavior. This is done via a personalized utility for each product j, $U_j^z = V_j^z + \epsilon_j$, where we view V_j^z as the base personalized utility of product j for the customer with feature vector z at the price level p_j. Here, V_j^z is specified by a linear model $V_j^z = \langle \gamma_j^*, z \rangle + \beta_j^* p_j$ for some $\gamma_j^* \in \mathbb{R}^m$ and $\beta_j^* \in \mathbb{R}$ and for $1 \le j \le J$. Without loss of generality, the mean utility V_0 of the no-purchase option is taken to be zero, for each feature vector z. If the current customer has feature vector z and the decision maker offers the assortment $S \in \mathcal{S}$ of products at prices specified by the vector $\boldsymbol{p} \in \mathbb{R}^J$, the customer will choose the product in S with the highest U_j^z. We assume that the ϵ_j values are Gumbel distributed and independent, in which case it is a well-known result from discrete choice theory (Train 2009) that in this setting, the customer chooses product $j \in S$ with probability

$$\mathbb{P}_z(j; S, \boldsymbol{p}, \gamma^*, \beta^*) = \frac{e^{V_j^z}}{1 + \sum_{l \in S} e^{V_l^z}} = \frac{\exp\left\{\langle \gamma_j^*, z \rangle + \beta_j^* p_j\right\}}{1 + \sum_{l \in S} \exp\left\{\langle \gamma_l^*, z \rangle + \beta_l^* p_l\right\}}, \quad (1)$$

for $j \ne 0$. In what follows, we use the variable y to denote the customer's decision between offered products with $y \in \{0, \ldots, J\}$, where product 0 is used to represent the no-purchase alternative. We also note that this model is different from some contextual assortment optimization models in the existing literature (Chen et al. 2018) because it involves the price information p_j.

Before discussing the estimation procedure, we note that our model of the price effect via the parameters $\beta_j^* \in \mathbb{R}$ is simply for clarity of exposition and to highlight the model of the seller's decision (i.e., price) as a feature. It is equally possible to model the effect of price using other, possibly nonlinear, functions. In particular, if the set of possible prices is discrete and finite, then using indicator random variables is a common way of modeling the nonlinear impact of categorical features in multifactor analysis of variance (Rao et al. 2008). We demonstrate the application of this methodology in our numerical experiments focused on customized pricing as detailed in Section 5.1. This model directly extends to modeling interaction effects between offered prices and other features (e.g., via introducing extra features). These interaction effects allow us to measure the change in price sensitivity given specific customer features, and we have found such effects to be especially useful in practice.

3.1. Maximum Likelihood Estimation Approach

No manager has access to a correct and fully specified model. Every model must be learned from data. We focus on issues of learning and statistics, which depend on the estimation procedure, described next.

We assume the decision maker has access to a set $\mathcal{T} = \{(z_1, S_1, \boldsymbol{p}_1, y_1), \ldots (z_n, S_n, \boldsymbol{p}_n, y_n)\}$ of n past samples, each consisting of the associated features, the offered pricing and assortment decisions, and the resulting outcome that serves as input to the algorithm. In this case $\boldsymbol{p}_i = (p_{i1}, \ldots, p_{iJ})$ is the vector of historical

pricing decisions for all products in period i. We stress that, in contrast to the literature on learning and earning, this data set is assumed to be analyzed all at once, rather than online. It is also worthwhile to note that this corpus of data is not the result of a personalized pricing strategy, or else, the prices would be a function of the observed features, destroying the ability to learn the price parameter of the model. One potential realistic source for these data could be company data from before any personalization was implemented, when prices were varied over time, but on a coarser scale without regard to the customer features. In this case, our setup would model a firm that is in the process of implementing personalized revenue management for the first time. Another setting is when a random pricing strategy is interlaced with a personalized pricing strategy, and only the data from the random prices are included in the training set. Despite the popularity of dynamic learning and earning in academic literature, we believe that the static setting considered in this paper is more prevalent in practice.

In preparation for MLE, we calculate the negative log-likelihood $\ell_n(\mathcal{T};\gamma,\beta) = -1/n\sum_{i=1}^{n}\log(\mathbb{P}_{z_i}(y_i;S_i,\boldsymbol{p}_i,\gamma,\beta))$, where the concrete form of $P_{z_i}(y_i;S_i,\boldsymbol{p}_i,\gamma,\beta)$ is given in (1). We further predetermine a positive number R such that $\|(\gamma^*,\beta^*)\|_1 \le R$, where $\|v\|_1 = \sum_i |v_i|$ denotes the vector ℓ_1-norm, and adopt the regularized/constrained MLE with the constraint that $\|(\gamma,\beta)\|_1 \le R$. In practice, one can either tune this R for better performance or simply fix a large-enough number R. This regularization is useful to control model complexity, to facilitate theoretical analysis, and often leads to better empirical performance. For our theoretical results, we assume $\|(\gamma^*,\beta^*)\|_1 \le R$.

To learn the parameters of the model, we minimize $\ell_n(\mathcal{T};\gamma,\beta)$ over (γ,β) under ℓ_1 regularization to get an estimate for γ and β. It is easy to see that $\ell_n(\mathcal{T};\gamma,\beta)$ is a convex function and ℓ_1 regularization is a convex constraint on (γ,β). Therefore, any fast convex optimization procedure (e.g., accelerated projected gradient descent, alternating direction method of multipliers, etc.) can be adopted to solve this optimization problem. The reader might refer to Boyd et al. (2010) and Bach et al. (2011) for recent developments on ℓ_1-regularized convex optimization algorithms.

3.2. Applications of the Model

Although we consider pricing and assortment decisions simultaneously, the model can capture equally well either customized pricing or personalized assortment optimization, with the other decision fixed. In the case of customized pricing, the seller has a single product ($J = 1$) without inventory constraints and wishes to offer a price to each customer that will maximize her revenue. Here, the outcome y is a binary decision and is equal to one if the customer purchases the product and zero otherwise. Given a pricing decision p, we can express the expected revenue in this scenario by

$$f_z(p;\gamma^*,\beta^*) := p\ \mathbb{P}_z(1;\gamma^*,\beta^*), \tag{2}$$

where $\mathbb{P}_z(1;\gamma^*,\beta^*) = \frac{\exp\{\langle\gamma^*,z\rangle+\beta^*p\}}{1+\exp\{\langle\gamma^*,z\rangle+\beta^*p\}}$. In the case of personalized assortment optimization, the prices of each product are taken as fixed, and instead, the seller has the choice of which assortment $S \subseteq \mathcal{S}$ to show each arriving customer. Because in this case, prices are fixed, the effect on purchase decisions can be incorporated into the product-specific intercept term, and we suppress dependence on β^*. Offering the assortment S to a customer with feature vector z gives the following expected revenue:

$$f_z(S;\gamma^*) = \sum_{j\in S} p_j \mathbb{P}_z(j;S,\gamma^*), \tag{3}$$

where $\mathbb{P}_z(j;S,\gamma^*) = \frac{\exp\{\langle\gamma_j^*,z\rangle\}}{1+\sum_{l\in S}\exp\{\langle\gamma_l^*,z\rangle\}}$. In both of these applications, the seller wishes to learn the personalized choice model from past transaction data. Critically, this model includes both the effect of an individual customer's features and the effect of the seller's pricing and assortment decisions. In the following section, we provide statistical bounds on the accuracy of the estimation procedure and show how these bounds can be extended to provide similar guarantees on the revenue of the seller's resulting decisions.

3.3. Discussions of Linearity Assumption

In this section, we comment on the linearity assumption of utility parameters and some natural extensions of the linear model.

First, the recent work (Besbes and Zeevi 2015) has examined a similar linear demand assumption in the context of demand learning through dynamic pricing. They found that even when true demand is nonlinear, the "price of misspecification" (i.e., the revenue loss from using a linear demand model) is much smaller than would be expected.

If one believes, however, that a linear demand assumption is too strong, we discuss two potential ways to weaken this assumption. The first is to include basis functions of features (e.g., B-spline basis) as part of the feature vector. This requires no change to the model assumptions, only an increase in the dimension of the feature vector. Such an approach is standard in the literature (see, for example, Ban and Rudin 2019). Another strategy is that one can replace the linearity assumption with a generalized additive model (GAM), in which the utility is taken to be a sum

of arbitrary one-dimensional functions of the features: that is,

$$V_j^z = \sum_{i=1}^{d} f_{ij}(z_i) + g_j(p_j),$$

where f_{ij} and g_j are univariate smooth functions. Such an approach remains very computationally tractable and is common in statistical analysis. We refer the interested reader to Hastie and Tibshirani (1990) for more information, including a backfitting algorithm for learning GAMs from data.

4. Theory

We now proceed to derive high-probability guarantees on the performance of the proposed regularized maximum likelihood estimate technique. Specifically, we will show that the parameters recovered by this estimation technique lie within an L_2 ball centered at the true parameters with high probability, and quantify the rate at which the radius of this ball converges to zero as the number of historical samples grows. Subsequently, we show how this result can be extended to bound the revenue lost in operational contexts in comparison with an operator with full knowledge of the system parameters. Throughout, we assume that the logit model is well specified (i.e., it is the correct underlying model of outcome probabilities). Let $\theta := (\gamma_1^T, \beta_1, \ldots, \gamma_J^T, \beta_J)^T$, with θ^* the underlying true parameters for generating the observed data $i \in \{1, \ldots, n\}$ (i.e., $\mathbb{P}_{z_i}(y_i; S_i, \boldsymbol{p}_i) = \mathbb{P}_{z_i}(y_i; S_i, \boldsymbol{p}_i, \theta^*)$).

We make the following assumptions.

Assumption 1.

a. *Conditional independence: the observed outcomes* $\{y_i\}_{i=1}^n$ *are independent given each* z_i, S_i, *and* $\boldsymbol{p}_i$.

b. *The vectors* $\{x_{ij} := (z_i, p_{ij})^T\}_{i=1}^n$ *for each* j *are independent across* i, *though not necessarily identically distributed. For each* i, *the distribution is sub-Gaussian with uniform sub-Gaussian norm* ψ.

Conditional independence between transactions is a standard assumption in the revenue management literature. By allowing for the possibility of nonidentically distributed features and prices, our model affords the flexibility to capture a number of scenarios of practical interest. For example, past prices could have been set by random experimentation or deterministically through a promotion schedule fixed at the beginning of the selling season. In addition, it is possible that the distribution of customer attributes was itself changing over the course of the historical period. We will only require that the features and pricing decisions are not too correlated over time in a sense that will be made precise shortly (see Assumption 2b).

We also assume that the customers' feature vectors obey a sub-Gaussian distribution. The assumption of sub-Gaussian feature vectors is a common and natural assumption for regression analysis because it captures a wide range of multivariate distributions. Examples include the multivariate Gaussian distribution, the multivariate Bernoulli distribution, the spherical distribution (for modeling normalized unit-norm feature vectors), and a uniform distribution on a convex set, among many others. We refer interested readers to the online appendix and to Vershynin (2012) for more details.

We note that the assumption of bounded features directly implies that the feature vectors are sub-Gaussian. However, the sub-Gaussian norm implied by the boundedness assumption for a feature vector of dimension d could be as large as $\sqrt{d}$ in the worst case. The assumed ψ could be much smaller than this naive bound in practice.

Assumption 2.

a. *Let* $\mathcal{I}_j = \{i : j \in S_i\}$, $j \in [J]$, *be the indices of transactions that offer product* j *in their assortment, with* $n_j = |\mathcal{I}_j|$. *We assume that for all* $j \in [J]$, $n_j \geq \nu n$, *where* $\nu \equiv \min_j \frac{n_j}{n} \in (0, \frac{1}{J}]$ *is a constant.*

b. *Define* $\Sigma_{\mathcal{I}_j} = \frac{1}{n_j} \sum_{i \in \mathcal{I}_j} \mathbb{E}(x_{ij} x_{ij}^T)$. *Let* $\lambda_{\min}(\cdot)$ *denote the minimum eigenvalue of the argument matrix. We assume there exists a constant* ρ *such that* $\lambda_{\min}(\Sigma_{\mathcal{I}_j}) \geq \rho > 0$ *for all* $j \in [n]$. *We also assume that* $\max_{i,j}\{\lambda_{\min}(\mathbb{E}(x_{ij} x_{ij}^T))\} > 0$.

The first part of this assumption requires that the seller has *sufficiently explored* each product. In the context of assortment optimization, there are J products, and it is impossible to estimate a customer's reaction to a product if this product has never been offered in previous assortments. Intuitively, when ν is small, some products are explored rarely in the data, which renders the estimation task difficult. On the other hand, when ν is close $\frac{1}{J}$, the collected data are more balanced in the sense that each product is explored roughly the same number of times. In such a case, the estimation task is simpler, and the revenue gap becomes smaller.

The second part of this assumption simply requires that the contextual variables and the pricing decisions are not collinear. From a theoretical perspective, such collinear features add no representational capacity to the model, and in practice, it is standard procedure in regression modeling to check variables for such collinearity and to adjust the model in response (see Bertsimas and King 2015 for detailed discussion). This assumption also highlights the need for managers to be aware of possible sources of price endogeneity in their data and to work to minimize these effects to ensure that the data will be useful for subsequent price optimization. As a practical example, a retailer

that only schedules pricing promotions on weekends is likely to introduce correlations between offered prices and features leading to incomplete learning and potentially suboptimal operational decisions.

4.1. Estimation Error Bound

To establish the revenue gap, we first establish the rate of convergence of the parameter estimates $\widehat{\theta}$ to the true parameters θ^*. Specifically, under Assumptions 1 and 2, we will prove that $\|\widehat{\theta} - \theta^*\|_2 \le C\sqrt{\frac{\log n}{n}}$ with high probability for some constant C. Intuitively, this says that the parameters of the choice model that are estimated from the data converge to the true parameters at a rate of $1/\sqrt{n}$ (up to a logarithmic factor). Using the smoothness of our revenue functions, we will show that this rate of convergence of parameters can be translated into the revenue space. The formal theorem statement is given here.

Theorem 1 (Parameter Convergence Rate). *Under Assumptions 1 and 2, we have the following: as long as* $n \ge \frac{4C_\psi \log(\nu n)}{\nu \min(\rho,1)^2}$ *for some constant* C_ψ *only depending on* ψ, *with probability at least* $1 - 1/n - 2J/(\nu n)$,

$$\|\widehat{\theta} - \theta^*\|_2 \le 2\,\frac{\left[1 + \exp(-R) + (J-1)\exp(R)\right]^2}{\exp(-R)\nu\rho} \times \sqrt{\frac{2J(d+1)\log(2nJ(d+1))}{n}}.$$

With respect to its dependence on n, this bound is rate optimal as it matches the Cramer–Rao lower bound up to logarithmic terms (see, for example, theorem 6.1 in Lehmann and Casella 1998, p. 124). Here, J, d, and R are viewed as constants. Providing a sharp lower bound with respect to the constants requires an explicit characterization of the inverse Fisher information matrix, which unfortunately, has no closed-form expression in this case.

To prove Theorem 1, we first establish the strong convexity of the loss $\ell_n(\theta)$ with strong convexity parameter $\eta > 0$. Let $\widehat{\Delta} = \widehat{\theta} - \theta^*$ denote the error in the parameter estimate with respect to θ^*, and recall that the goal of Theorem 1 is to provide a finite-sample upper bound on $\|\widehat{\Delta}\|_2$. The strong convexity of ℓ_n implies that

$$\frac{\eta}{2}\left\|\widehat{\Delta}\right\|_2^2 \le \ell_n\left(\theta^* + \widehat{\Delta}\right) - \ell_n(\theta^*) - \left\langle \nabla\ell_n(\theta^*), \widehat{\Delta}\right\rangle. \tag{4}$$

Because $\widehat{\theta}$ is the minimizer of ℓ_n, we have $\ell_n(\theta^* + \widehat{\Delta}) - \ell_n(\theta^*) = \ell_n(\widehat{\theta}) - \ell_n(\theta^*) \le 0$. Together with (4) and using Hölder's inequality, this implies that

$$\frac{\eta}{2}\left\|\widehat{\Delta}\right\|_2^2 \le -\left\langle \nabla\ell_n(\theta^*), \widehat{\Delta}\right\rangle \le \left\|\nabla\ell_n(\theta^*)\right\|_\infty \left\|\widehat{\Delta}\right\|_1 \le \sqrt{J(d+1)}\left\|\nabla\ell_n(\theta^*)\right\|_\infty \left\|\widehat{\Delta}\right\|_2.$$

This further implies that

$$\left\|\widehat{\Delta}\right\|_2 \le \frac{2\sqrt{J(d+1)}}{\eta}\left\|\nabla\ell_n(\theta^*)\right\|_\infty. \tag{5}$$

Intuitively, (5) tells us that when ℓ_n has sufficient curvature near θ^* (quantified by η), a small gradient must imply that the true parameter θ^* is near optimal for the empirical log-likelihood function ℓ_n.

Therefore, to establish an upper bound on $\|\widehat{\Delta}\|_2 = \|\widehat{\theta} - \theta^*\|_2$ using (5), we only need to (1) establish an upper bound on $\|\nabla\ell_n(\theta^*)\|_\infty$ and (2) identify the strong convexity parameter η. These steps are accomplished in the following lemmas, with proofs given in the online appendix. We begin by showing that $\|\nabla\ell_n(\theta^*)\|_\infty$ can be upper bounded with high probability.

Lemma 1 (Gradient Bound). *Under the assumptions of Theorem 1, with probability at least* $1 - \frac{1}{n}$,

$$\left\|\nabla\ell_n(\theta^*)\right\|_\infty \le \sqrt{\frac{2\log(2nJ(d+1))}{n}}.$$

Now we show that ℓ_n is a strongly convex function with strong convexity parameter η, which is independent of sample size n.

Lemma 2 (Strong Convexity). *Under the assumptions of Theorem 1, as long as* $n \ge \frac{4C_\psi \log(\nu n)}{\nu \min(\rho,1)^2}$ *for some constant* C_ψ *only depending on* ψ, *with probability at least* $1 - 2J/(\nu n)$, *we have strong convexity parameter*

$$\eta = \frac{\exp(-R)\nu\rho}{2\left[1 + \exp(-R) + (J-1)\exp(R)\right]^2}. \tag{6}$$

Proof of Theorem 1. By plugging both the upper bound on $\|\nabla\ell_n(\theta^*)\|_\infty$ in Lemma 1 and the strong convexity parameter η in (6) into (5), we obtain the result of Theorem 1, which completes the proof of Theorem 1. ∎

4.2. Revenue Bound

Theorem 1 gives us a finite-sample (nonasymptotic) estimation bound that holds with high probability (i.e., the theorem tells us that the parameter estimates $\widehat{\theta}$ converge to the true parameters θ^* at a rate of $\frac{1}{\sqrt{n}}$). We can now present the associated bound on expected revenue, which is much more important in the context of revenue management problems. In what follows, we fix any customer feature vector $z \in \mathcal{Z}$. Given this feature vector, we define the expected revenue function for the assortment and pricing decision specified by $S \in \mathcal{S}$ and $\boldsymbol{p} \in \mathcal{P}$ under parameters θ by

$$f_z(S, \boldsymbol{p}; \theta) = \sum_{j \in S} p_j \mathbb{P}_z(j; S, \boldsymbol{p}, \theta).$$

Let $(\widehat{S},\widehat{\boldsymbol{p}})$ and $(S^*,\boldsymbol{p}^*)$ denote the personalized assortment and pricing decisions determined under the estimated parameters $\widehat{\theta}$ and the true parameters θ^*, respectively. Formally, we have

$$\left(\widehat{S},\widehat{\boldsymbol{p}}\right) := \arg\max_{S\in\mathcal{S},\boldsymbol{p}\in\mathcal{P}} f_z\left(S,\boldsymbol{p};\widehat{\theta}\right) \quad \text{and}$$

$$\left(S^*,\boldsymbol{p}^*\right) := \arg\max_{S\in\mathcal{S},\boldsymbol{p}\in\mathcal{P}} f_z\left(S,\boldsymbol{p};\theta^*\right).$$

Here, we assume access to an optimization oracle for solving these optimization problems. For example, Li and Huh (2011) show that the problem of finding optimal market shares, and thereby determining optimal prices, can be solved as a convex optimization problem under the nested logit model of which the multinomial logit is a special case. When prices are fixed, Talluri and van Ryzin (2004) demonstrate that the optimal assortment is revenue ordered in the unconstrained case, whereas Rusmevichientong et al. (2010) provide an efficient algorithm when feasible assortments must meet a cardinality constraint.

After estimating the parameters and determining the action $(\widehat{S},\widehat{\boldsymbol{p}})$, we are interested in the revenue gap between the revenue of this decision and that generated by the oracle decision $(S^*,\boldsymbol{p}^*)$ when the customer's behavior is specified by the true parameters θ^*. This gap is expressed mathematically as $f_z(S^*,\boldsymbol{p}^*;\theta^*) - f_z(\widehat{S},\widehat{\boldsymbol{p}};\theta^*)$. The next theorem demonstrates that this revenue gap decreases at a quantifiable rate as the sample size n is increased. We note that this revenue gap is an *out-of-sample guarantee* because such a bound holds for any new customer with feature vector z.

Theorem 2 (Revenue Convergence Rate)**.** *Under Assumptions 1 and 2, we have that with high probability, as long as* $n \geq \frac{4C_\psi \log(\nu n)}{\nu \min(\rho,1)^2}$*, for any feature vector* z*, the expected revenue gap as a fraction of the maximal price can be bounded by*

$$f_z\left(S^*,\boldsymbol{p}^*;\theta^*\right) - f_z\left(\widehat{S},\widehat{\boldsymbol{p}};\theta^*\right) \leq \frac{C(R,\psi)}{\nu\rho} J^4(d+1) \times \sqrt{\frac{\log(2nJ(d+1))}{n}},$$

where $C(R,\psi)$ *is a constant only depending on* R *and* ψ*.*

The proof is a simple consequence of the following proposition.

Proposition 1. *For any offered assortment* $S \subseteq \{1,\dots,J\}$ *and price vector* $\boldsymbol{p}\in\mathcal{P}$*, with high probability, as long as* $n \geq \frac{4C_\psi \log(n)}{\nu \min(\rho,1)^2}$*, the error in our revenue forecast as a fraction of the maximal price can be bounded as follows:*

$$\left|f_z(S,\boldsymbol{p};\theta^*) - f_z(S,\boldsymbol{p};\hat{\theta})\right| \leq \frac{C(R,\psi)}{2\nu\rho} J^4(d+1) \times \sqrt{\frac{\log(2nJ(d+1))}{n}}$$

for all bounded feature vectors z *where* $C(R,\psi)$ *is a constant depending only on* R *and* ψ*.*

Proof of Proposition 1. Fix a customer feature vector $z\in\mathcal{Z}$, a subset of products $S\subseteq\mathcal{S}$, and a price vector $\boldsymbol{p}\in\mathcal{P}$. For a fixed subset of products, the difference in expected revenue under the true model defined by θ^* and our estimated model specified by $\widehat{\theta}$ depends only on the difference in purchase probabilities they suggest. To bound the difference in these purchase probabilities for each item j, we observe that $\frac{\delta}{\delta\theta_{jk}}\mathbb{P}_z(j;S,\boldsymbol{p},\theta) \leq \frac{1}{4}\|z\|_\infty \leq \frac{1}{4}$ for all $k\in S$ for any value of θ, and $z\in\mathcal{Z}$. Therefore, we have the global bound on the gradient of $\mathbb{P}_z(j;S,\boldsymbol{p},\theta)$ with respect to θ of $\|\nabla\mathbb{P}_z(j;S,\boldsymbol{p},\theta)\|_\infty \leq \frac{1}{4}$. Using this, we can proceed to bound the forecast error as claimed using this Lipschitz constant:

$$\begin{aligned}
&\left|f_z(S,\boldsymbol{p};\theta^*) - f_z(S,\boldsymbol{p};\hat{\theta})\right| \\
&\quad\leq \sum_{j\in S} p_j \left|\mathbb{P}_z(j;S,\boldsymbol{p},\theta^*) - \mathbb{P}_z(j;S,\boldsymbol{p},\hat{\theta})\right| \\
&\quad\leq \sum_{j\in S} p_j \frac{1}{4}\|\theta^* - \widehat{\theta}\|_1 \\
&\quad\leq \frac{J\,p_{\max}}{4}\|\theta^* - \widehat{\theta}\|_1 \\
&\quad\leq \frac{J}{4}\sqrt{J(d+1)}\|\theta^* - \widehat{\theta}\|_2
\end{aligned}$$

Substituting in the result of Theorem 1 yields the desired result with high probability. ∎

Proof of Theorem 2. We have

$$\begin{aligned}
&f_z\left(S^*,\boldsymbol{p}^*;\theta^*\right) - f_z\left(\widehat{S}_z,\widehat{\boldsymbol{p}}_z;\theta^*\right) \\
&\quad= \left(f_z\left(S_z^*,\boldsymbol{p}^*;\theta^*\right) - f_z\left(\widehat{S}_z,\widehat{\boldsymbol{p}}_z;\widehat{\theta}\right)\right) \\
&\qquad+ \left(f_z\left(\widehat{S}_z,\widehat{\boldsymbol{p}}_z;\widehat{\theta}\right) - f_z\left(\widehat{S}_z,\widehat{\boldsymbol{p}}_z;\theta^*\right)\right) \\
&\quad\leq \left(f_z\left(S^*,\boldsymbol{p}^*;\theta^*\right) - f_z\left(S^*,\boldsymbol{p}^*;\widehat{\theta}\right)\right) \\
&\qquad+ \left(f_z\left(\widehat{S}_z,\widehat{\boldsymbol{p}}_z;\widehat{\theta}\right) - f_z\left(\widehat{S}_z,\widehat{\boldsymbol{p}}_z;\theta^*\right)\right) \\
&\quad\leq \frac{C(R,\psi)}{\nu\rho} J^4(d+1)\sqrt{\frac{\log(2nJ(d+1))}{n}},
\end{aligned}$$

where in the final step, we have applied the result of Proposition 1 twice. ∎

As desired, we have bounded the optimality gap of the reward generated by the recommended action $(\widehat{S},\widehat{\boldsymbol{p}})$ as compared with the oracle decision $(S^*,\boldsymbol{p}^*)$. This bound decreases as $O(\frac{1}{\sqrt{n}})$, up to logarithmic terms, and thus, quantifies the trade-off between the value of data and potential lost revenue. We can now specialize these results to the cases of customized pricing and assortment optimization. In the case of

customized pricing, there is only a single product ($J = 1$), which allows the bound to sharpen considerably.

Corollary 1 (Customized Pricing). *In the setting of customized pricing, under Assumptions 1 and 2, we have that with high probability, as long as* $n \geq \frac{4C_\psi \log(n)}{\min(\rho,1)^2}$, *for any feature vector* z, *the expected revenue gap as a fraction of the maximal price can be bounded by*

$$f_z(p^*;\theta^*) - f_z(\widehat{p};\theta^*) \leq \frac{C(R,\psi)}{\rho}(d+1) \times \sqrt{\frac{\log(2n(d+1))}{n}}.$$

Corollary 2 (Assortment Optimization). *In the setting of assortment optimization with fixed prices, under Assumptions 1 and 2, we have that with high probability, as long as* $n \geq \frac{4C_\psi \log(\nu n)}{\nu \min(\rho,1)^2}$, *for any feature vector* z, *the expected revenue gap can be bounded by*

$$f_z(S^*;\gamma^*) - f_z\left(\widehat{S};\gamma^*\right) \leq \frac{C(R,\psi)}{\nu\rho} J^4 d\sqrt{\frac{\log(2nJd)}{n}}.$$

4.3. Dependent Observations

In some cases, the assumption of independent feature vectors across customers may be too strong. It is very possible that customers who have similar features will arrive in the system close together. One reason for this correlation of features is that a customer is likely to communicate about a purchase and thus, influence similar customers' purchase decisions. The results of Sections 4.1 and 4.2 can be generalized to the case of dependent features by replacing the independence assumption with a *mixing* assumption. Intuitively, a random process is mixing if pairs of events tend toward independence as their intervening distance in the sequence grows.

There are many notions of mixing (Bradley 2005). We adopt geometric absolute regularity. Let $\mathcal{A}$ and $\mathcal{B}$ be σ fields on a set Ω. Define

$$b(\mathcal{A},\mathcal{B}) = \frac{1}{2}\sup\left\{\sum_{i\in I}\sum_{j\in J}|P(A_i \cap B_j) - P(A_i)P(B_j)|\right\},$$

where the supremum is taken over all finite partitions $(A_i)_{i\in I}$ and $(B_j)_{j\in J}$ of Ω where $A_i \in \mathcal{A}$ and $B_j \in \mathcal{B}$ for all i, j. We also define $b_0 := 1$ and for $m = 1, \ldots, n$,

$$b_m := \sup_{j\in[n-m]} b\left(\sigma(x_i, i \leq j), \sigma(x_i, i \geq j+m)\right),$$

where $\sigma(\cdot)$ denotes the σ algebra generated by a set of random variables.

A stochastic process is said to be geometrically absolutely regular if $b_m \leq \exp(-c(m-1))$ for some constant $c > 0$.

Consider this generalization of Assumption 1.

Assumption 3.

1. *Conditional independence: the observed outcomes* $\{y_i\}_{i=1}^n$ *are independent given each* z_i, S_i, *and* $\boldsymbol{p}_i$.
2. *For each* j, *the vectors* $\{x_{ij} := (z_i, p_{ij})^T\}_{i=1}^n$ *are geometrically absolutely regular with uniform constant* c. *For each* i, *the distribution is sub-Gaussian with uniform sub-Gaussian norm* ψ.

We can now prove the following result, which is a generalization of our Theorem 1.

Proposition 2. *Under Assumptions 2 and 3, we have*

$$\left\|\widehat{\theta} - \theta^*\right\|_2 \leq \widetilde{O}\left(\sqrt{\frac{d}{n}}\right),$$

where $\widetilde{O}(\cdot)$ *denotes the order of the expression up to logarithmic factors.*

This result includes finite, irreducible, aperiodic Markov chains as a special case (Samson 2000). The proof of Proposition 2 will be relegated to the online appendix. Proposition 2 can be extended to derive a revenue bound analogous to the bounds in Section 4.2.

5. Numerical Experiments

The theory we have presented so far suggests that our method provides an effective technique for estimating and optimizing data-driven decisions. We performed simulations for both customized pricing and personalized assortment optimization to demonstrate the effectiveness of our proposed techniques. Additionally, we corroborate our theoretical results using data from a European airline carrier by showing that the model performs well with a real data set, even with relatively small training sets.

5.1. Customized Pricing

For customized pricing, we defined a problem class by specifying a number of prices $K \in \{2, 4, 10\}$ and a number of features $d \in \{5, 10, 15\}$ and then performed 100 trials for each problem class.

We used a price set of size K by evenly spacing prices on the interval $[5, 20]$, ordered such that $p^K \leq p^{K-1} \leq \cdots \leq p^1$. For the K discrete prices, we adopt a dummy variable encoding that represents the price using $K-1$ binary features, which requires K coefficients, including the intercept. We generated the $K-1$-dimensional parameter vector β^* as the $K-1$ highest-order statistics of K independent and identically distributed (i.i.d.). normal random variables with mean zero and standard deviation three, sorting

from lowest to highest with the lowest incorporated into the intercept. This demonstrates an extension of our methodology to incorporate price-specific sensitivity parameters, which allow price effects to have a nonlinear impact on the log odds of the purchase probability. Here, the first term β_1 is incorporated into the intercept effect to remove a redundant parameter. We also generated a d-dimensional true parameter vector γ^*, where each dimension was chosen i.i.d. from a normal distribution with mean 0 and standard deviation 1.5. The difference in standard deviations for the two parameters allows the price effect to dominate the effects of the other features. The sorting of the β_k is motivated by the fact that in almost all cases, demand for a product is decreasing in its price.

We then generated a training set, where each data point consists of a feature vector z of size d, drawn i.i.d. from a multivariate normal distribution, a price drawn uniformly at random from the constructed set, and a purchase decision given according to the logistic regression model with the true parameters γ^* and β^*. Each method in each problem class was tested with $n = 100, 300$, and 500 training data points. The distribution of the vectors z had mean zero and a covariance structure such that $\text{var}(Z_i) = 1$ and $\text{cov}(Z_i, Z_j) = 0.3$, $i \neq j$. We trained the following using the training data.

1. The Personalized Revenue Maximization (PRM) algorithm, our maximum likelihood-based approach. We first estimate the parameters as described in Section 3.1 and then select prices greedily.
2. The I-PRM algorithm (PRM with the isotonic constraint that $\beta_k < \beta_l$ for all $k < l$).
3. A single-price policy.
4. A random forest-based (RF) classification algorithm.

The random forest algorithm (see Breiman 2001) splits the data into K subsets by offered price. For each subset k, RF then trains a forest of predictors, using customer features as splitting criteria. The output of forest k is a mapping from the space of feature vectors (in this case, $\mathbb{R}^d$) to a probability of purchase at price p^k. We then used these as inputs to the revenue optimization in order to choose a price. In training this model, we used the default splitting criteria (see Breiman 2001). We experimented with crossvalidation in terms of the leaf size parameter.

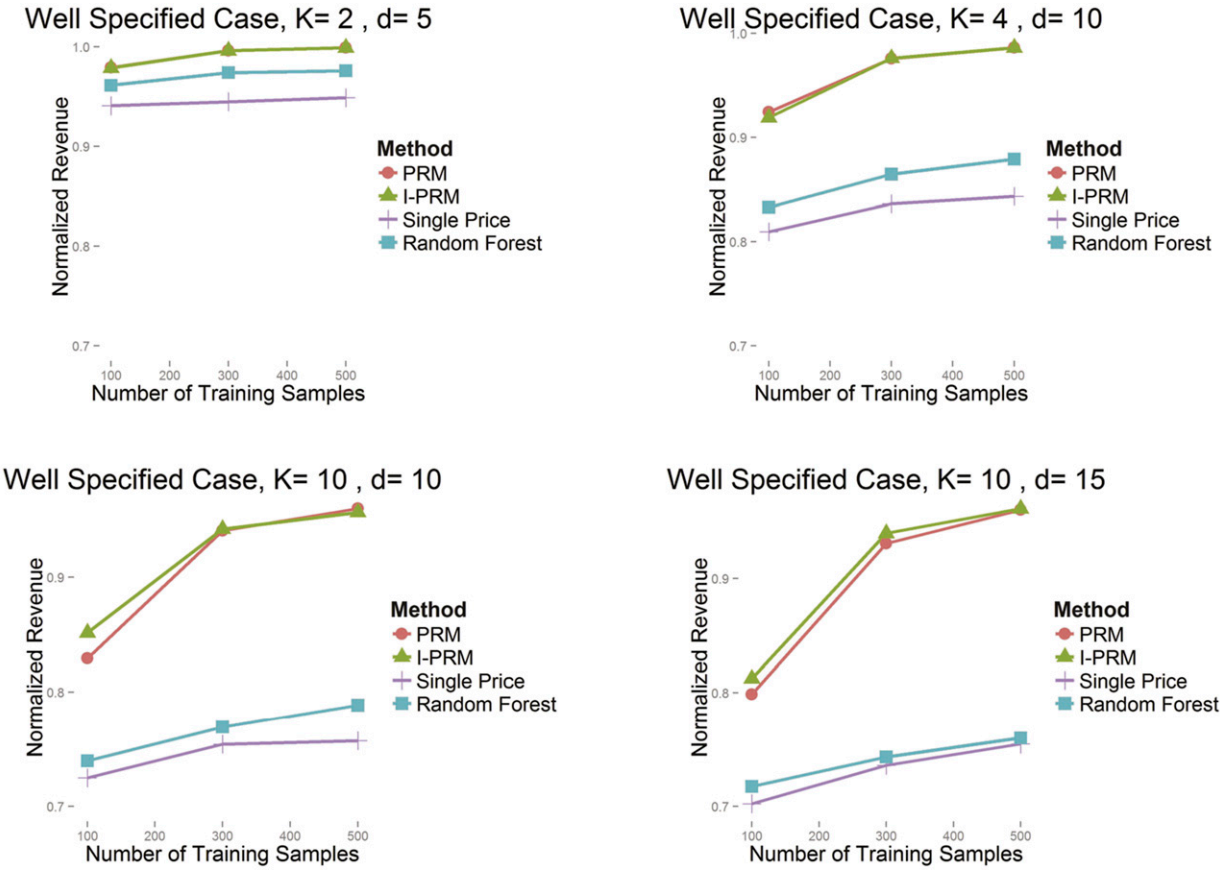
We generated a test set of 1,000 data points consisting of feature vectors and a reservation price generated according to the true model. We then used the test set to calculate empirical expected revenue for each method. We also calculated empirical expected revenue for a policy that knows the true parameters, selecting the price p^* that maximizes the contextual expected revenue $f_z(p; \theta^*) = p\mathbb{P}_z(p, \theta)$. We present the performance of each policy as the expected revenue of that policy divided by this optimal expected revenue.

Figure 1 shows the performance of all four methods under four representative problem classes, averaged across the 100 trials. The I-PRM slightly outperforms the regular version, and both are significantly better than a single-price strategy. Not very many samples are required for the algorithms to recover almost all of the full-knowledge revenue. The random forest also does better than single price but not by much. In this setting, we do not expect it to perform as well as PRM because PRM has extra information about the underlying generative model.

Recognizing that the generation of simulated data from an underlying logistic distribution favors the PRM algorithm over other algorithms, we also performed tests in the misspecified case. For these tests, we included second-order effects for each feature, resulting in underlying demand function $\mathbb{P}_z(y = 1; p, \beta^*, \gamma^*, \xi^*) = (1 + \exp(-(\sum_{k=2}^{K} \beta_k^* \mathbb{I}(p = p^k) + \sum_{j=1}^{d} \gamma_j^* z_j + \sum_{j=1}^{d} \xi_j^* (z_j)^2)))^{-1}$. We performed 100 trials for the same problem classes as in the well-specified case, generating the features, β^*, and γ^* in the same way. The ξ^* parameters were generated i.i.d. standard normal. In addition to the PRM, I-PRM, single-price, and random forest algorithms, we also trained versions of PRM and I-PRM that knew the form of the underlying model. In other words, PRM and I-PRM learned only $\widehat{\beta}$ and $\widehat{\gamma}$ parameters, whereas the new versions learned $\widehat{\beta}$, $\widehat{\gamma}$, and $\widehat{\xi}$ parameters. We dubbed these methods "HO PRM" and "HO I-PRM" for "higher-order" PRM and I-PRM. We tested all methods on a 1,000 data point test set generated with the new underlying demand model. We again normalized all empirical revenues by the empirical revenue of the full-knowledge method, which knows the true underlying distribution. Additionally, we compared results with the oracle model, the PRM model with the true maximum expected log likelihood using only the variables available to the PRM model. This model was simulated by training a logistic regression model on 10,000 training data points.

Figure 2 shows the results of these methods over $n = \{100, 300, 500, 1{,}000\}$. The PRM and I-PRM methods converge to the oracle revenue rather than the true optimal revenue in this case because of the misspecified model, but the oracle policy still collects a large fraction of the full-knowledge policy, more than 95% in smaller problem classes. The HO PRM and HO I-PRM converge, as expected, to the full-knowledge method, but for small amounts of data, their performance is comparable with PRM and I-PRM. Also, for $n < 500$ the HO methods struggled to converge because of the larger number of parameters to be estimated. With small amounts of data, the random

Figure 1. (Color online) Performance of PRM, I-PRM, Single-Price, and Random Forest Algorithms in the Well-specified Setting for Various Problem Classes



forest method also struggled and was outperformed by the PRM and I-PRM methods.

5.2. Assortment Optimization

For the assortment optimization experiments, the problem classes were given by specifying a number of products $J \in \{3, 6, 12, 20\}$ and a number of features $d \in \{5, 10, 15\}$. As in customized pricing, for each problem class we performed 100 trials. The revenues r_j, $j = 1, \ldots, J$ were evenly spaced on $[5, 25]$. We also generated a $d \times J$-dimensional true parameter matrix θ^*, where each matrix entry was chosen i.i.d from a normal distribution with mean zero and standard deviation three. To reflect the fact that features are often correlated with price of product, we sorted the first two rows of θ^*. We then generated a training set as in customized pricing, where each data point consists of a multivariate normal feature vector z of size d, an assortment drawn uniformly at random, and a purchase decision given according to the multinomial logit model with the true parameters γ^*. We trained the following using the training data:

1. the PRM Algorithm as applied to the case of assortment optimization,

2. a "Mean-Effect" Revenue Maximization (MERM) algorithm that does not use any feature information.

We did not train a tree-based or an empirical single-assortment method in the multinomial case because of the exponential number of possible assortments. The featureless algorithm MERM learned the mean effect V_j for each product by performing maximum likelihood estimation on the offered assortment and purchase decision data as follows:

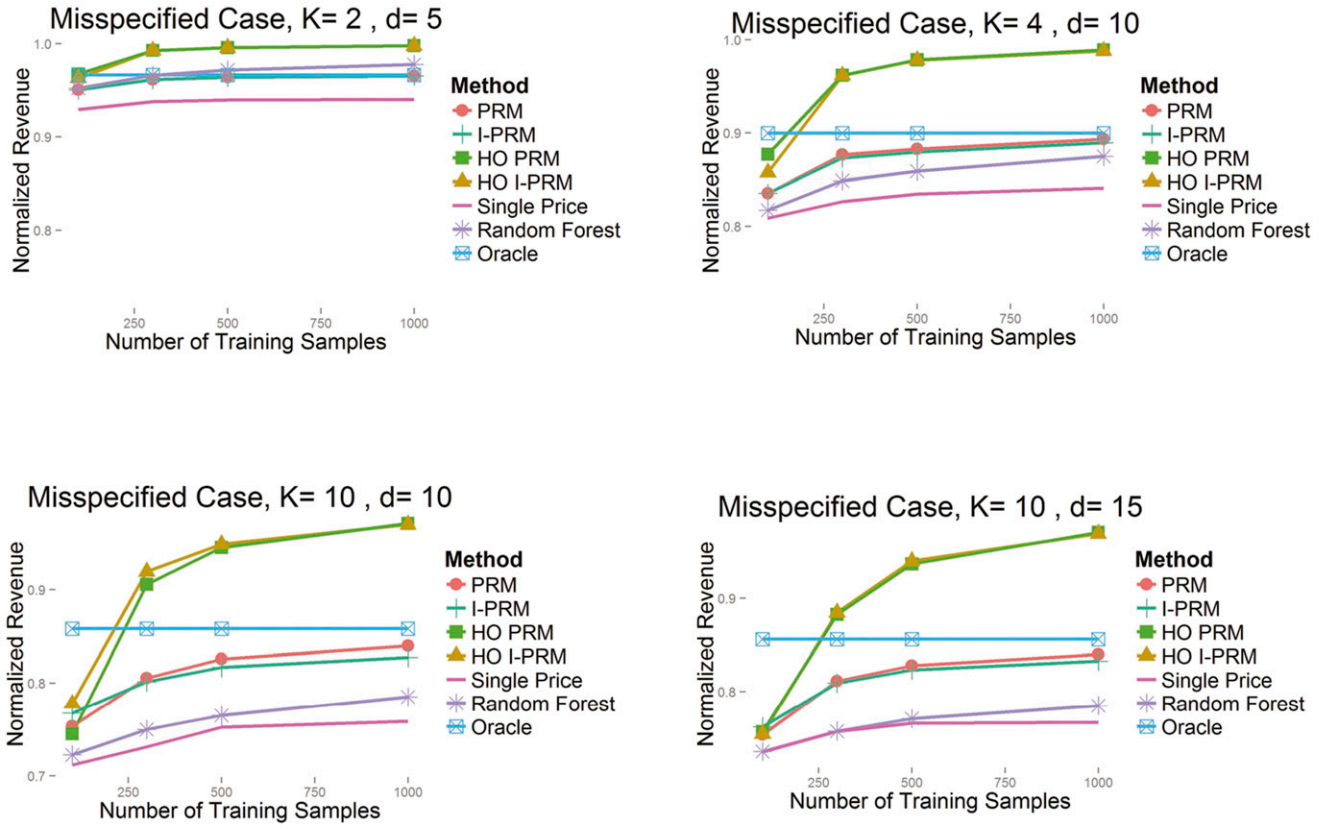
$$\widehat{V}_1^{\mathrm{MERM}}, \ldots, \widehat{V}_J^{\mathrm{MERM}} = \arg\min_{V_1,\ldots,V_J} \frac{1}{n} \sum_{i=1}^{n} \left[-V_{j_i} + \log\left(1 + \sum_{l \in S_i} \exp\{V_l\}\right)\right],$$

where j_i indicates the item purchased for the ith transaction. MERM then used these mean estimates $\widehat{V}^{\mathrm{MERM}}$ in an MNL model of choice to pick the best assortment.

Our test set consisted of feature vectors and a product ranking list generated according to the true model. In keeping with Talluri and van Ryzin (2004), all assortment policies were restricted to the revenue-ordered assortments $S_k = \{1, \ldots, k\}$, $k \in [J]$.

Figure 3 shows the performance of PRM and MERM algorithms (both normalized by the full-knowledge revenue) across a representative selection

Figure 2. (Color online) Performance of Methods in the Misspecified Setting for Various Problem Classes



of problem classes. As expected, PRM consistently and significantly outperforms the featureless approach, especially as the number of products grows large.

5.3. Customized Pricing for Airline Priority Seating

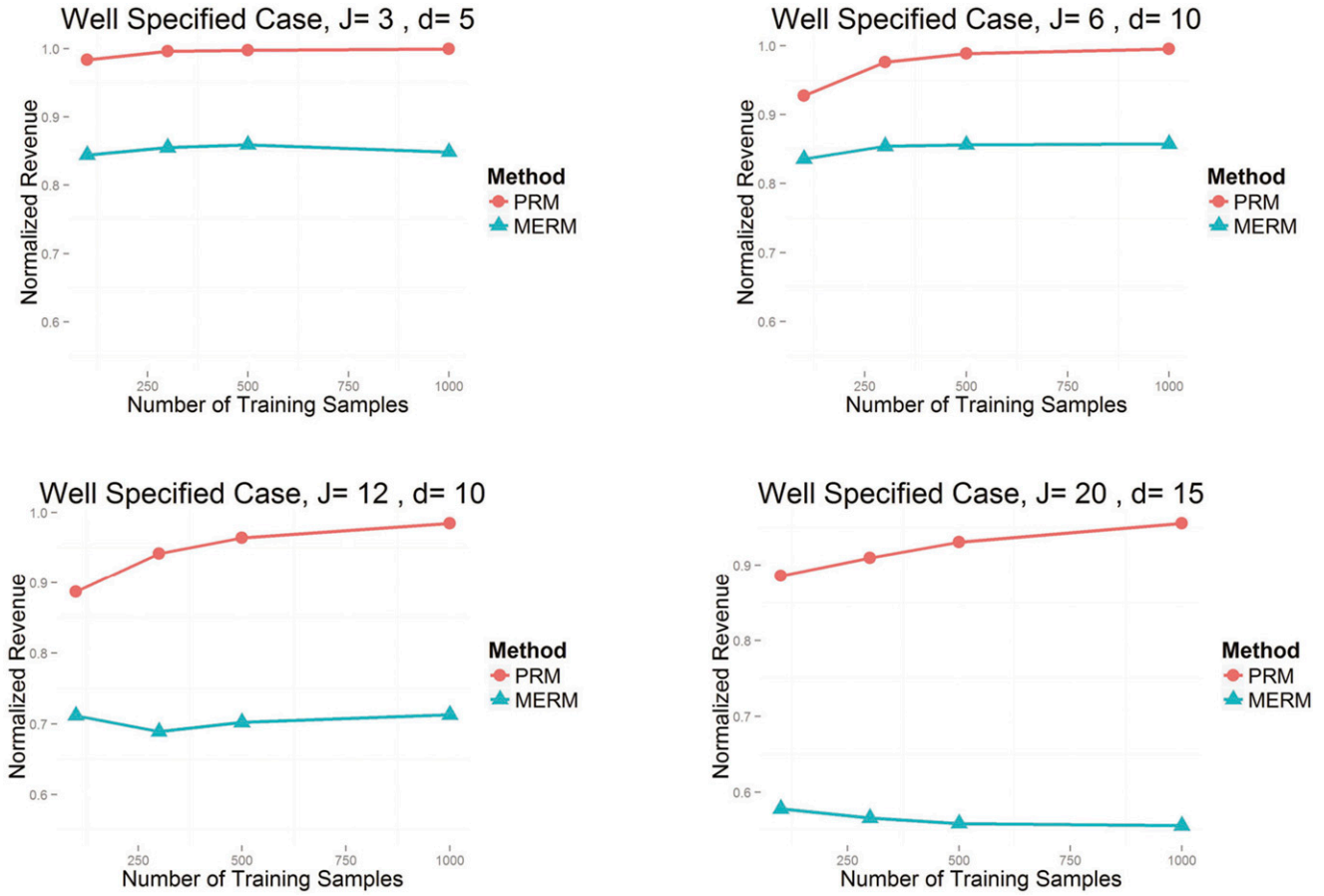
In addition to simulated experimental results, we tested our method for data-driven customized pricing with sales data from a European airline carrier. The data set concerns sales of passenger seating reservations, in which for an extra fee, passengers may select a seat at booking that will be reserved for them on the day of their flight. Over a one-month period, a fraction of customers who purchased domestic airline tickets was offered the opportunity to purchase a seating reservation at a treatment price randomly selected with equal probability from four candidate prices. The resulting data set consists of around 300,000 records with the treatment price offered, data concerning attributes of the flight, data concerning attributes of the transaction, and the resulting purchase decision. Other than the treatment price, the features we used are listed. We also considered interaction effects between each of these features and the treatment price.

- *Booking DW*: the day of the week of booking, as a categorical variable.
- *Flight DW*: the day of the week of the outbound flight, as a categorical variable.
- *Flight Hour*: the hour of the day of the outbound flight, as a numerical value.
- *Return Flight DW*: the day of the week of the return flight, as a categorical variable. This includes a value indicating that the outbound flight was one way.
- *Days Out*: the number of days between the date of the flight and the date of booking, as a numerical value.
- *Trip Days*: the length of the trip in days, as a numerical value.
- *Number of Passengers*: the number of passengers on the reservation, as a numerical value.
- *Trip Type*: a carrier-provided classification of the trip, as a categorical variable.
- *Fare Class*: the fare class of the base ticket, aggregated into three groups, as a categorical variable.

The numerical-valued features are taken to have linear impact on the log odds of purchase in our base model. Moreover, as we suggested in Section 3.3, we also model nonlinear effects using a GAM, described in more detail.

To allow us to fairly evaluate the extent to which our model can explain the observed data, we began by

Figure 3. (Color online) Performance of Methods in the Well-specified Assortment Optimization Setting for Various Problem Classes

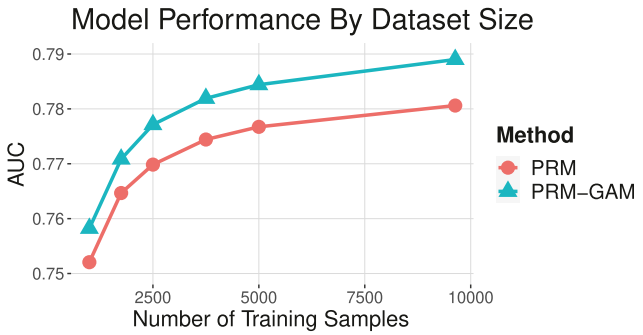


splitting our data into a training set and a testing set, with the earliest 60% of transaction records used for training and the latest 40% used for testing. Because of the relatively limited number of purchases, the nonpurchase records were down-sampled in the training set, resulting in 9,635 training examples. This down-sampling prevents machine learning techniques from overgeneralizing because of the prevalence of nonpurchases.

We fit and evaluated our model 100 times over each of the following training set sizes $n \in \{1{,}000, 1{,}750, 2{,}500, 3{,}750, 5{,}000, 9{,}635\}$ using the following steps. First, we randomly select n examples from the training pool and subsequently divide this set into four folds for cross-validation. We perform feature selection on each of these folds using a forward model selection procedure that at each step adds the feature that most improves the model, and we select the set of features with the highest average cross-validation performance. Finally, using these features, we refit the model on the entire training set and test its performance on the fully held-out test set. In addition to fitting the linear PRM model, we also fit a variation that utilizes the modeling flexibility of GAM on the features selected in this manner. The PRM-GAM learns a flexible nonlinear relationship between the numerical features (flight hour, days out, trip days, number of passengers) and the purchase decision. Interactions between the treatment price and these features are learned by fitting a separate nonlinear relationship for each such feature under each treatment price. Under this model, the effect of categorical features is learned in the same manner as for PRM.

In Figure 4, we report the average out-of-sample AUC (*Area Under* the Receiver Operating Characteristic *Curve*) obtained on the held-out test set over the 100 runs of this procedure. AUC is a statistical measure of model quality defined by the area under the curve given by the false positive rate and true positive rate of a classifier at each possible decision threshold (see, e.g., Fawcett 2006 for an introduction to this metric). If we compute the purchase probability using a given model for every testing set transaction record, the AUC is the fraction of purchase records that is given a higher probability than nonpurchase records. We observe that the AUC on the held-out sample grows from 0.75 to 0.78 as the training set size is increased from 1,000 to 9,635. Our results also demonstrate the

Figure 4. (Color online) Model Performance in AUC by Training Set Size



Note. Test set AUC grows from 0.752 to 0.781 for the linear model-based PRM and from 0.758 to 0.789 for the GAM-based PRM as the training sample size increases.

benefit of modeling numerical features in a flexible way as the nonlinear GAM outperforms the linear model-based PRM at every sample size, and this performance differential seems to grow as the sample size grows larger.

6. Extensions and Future Work

For a class of personalized revenue management problems, we demonstrate that learning takes place reliably by establishing finite-sample, high-probability convergence guarantees for model parameters. We also extend the parameter convergence guarantees to performance bounds for the joint pricing and assortment problem.

Beyond problems in revenue management, our approach is relevant in many other situations in which decisions resulting in discrete outcomes can benefit from taking into account explicit contextual information. One such example is in online advertisement allocation in which we would like to predict click-through rates and make the optimal advertisement selection, taking into account information we have about each viewer. Another example application is crowdsourcing in which we would like to specialize our work schedule based on information we have gathered concerning our workers, the available tasks, and the interaction between their attributes. Finally, beyond the specific domain of operations management, we envision applications in personalized medicine in which the likelihood of success of a treatment or the probability of disease could be predicted and decisions optimized by taking into account information concerning each patient. It is of great interest to explore such applications in the future.

Acknowledgments

The authors thank the department editor, the associated editor, and the anonymous referees for many useful suggestions and feedback, which greatly improved the paper.

References

Agrawal S, Avadhanula V, Goyal V, Zeevi A (2017) Thompson sampling for the MNL-bandit. Satyen K, Ohad S, eds. *Proc. Conf. Learning Theory (COLT)* 65:76–78.

Agrawal S, Avadhanula V, Goyal V, Zeevi A (2019) MNL-bandit: A dynamic learning approach to assortment selection. *Oper. Res.* 67(5):1453–1485.

Amatriain X (2013) Big & personal: Data and models behind Netflix recommendations. *Proc. 2nd Internat. Workshop Big Data, Streams Heterogeneous Source Mining: Algorithms, Systems, Programming Models Appl.* (ACM, New York), 1–6.

Arora N, Huber J (2001) Improving parameter estimates and model prediction by aggregate customization in choice experiments. *J. Consumer Res.* 28(2):273–283.

Aviv Y, Vulcano G (2012) Dynamic list pricing. Özer Ö, Phillips R, eds. *The Oxford Handbook of Pricing Management* (Oxford University Press, Oxford, United Kingdom), 522–584.

Aydin G, Ziya S (2008) Pricing promotional products under upselling. *Manufacturing Service Oper. Management* 10(3):360–376.

Aydin G, Ziya S (2009) Technical note—personalized dynamic pricing of limited inventories. *Oper. Res.* 57(6):1523–1531.

Bach F, Jenatton R, Mairal J, Obozinski G (2011) Optimization with sparsity-inducing penalties. *Foundations Trends Machine Learn.* 4(1):1–106.

Ban GY, Keskin NB (2020) Personalized dynamic pricing with machine learning: High dimensional features and heterogeneous elasticity. Preprint, submitted April 20, http://dx.doi.org/10.2139/ssrn.2972985.

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

Banna M, Merlevède F, Youssef P (2016) Bernstein type inequality for a class of dependent random matrices. *Random Matrices Theory Appl.* 5(2):1–28.

Belobaba P (1989) Application of a probabilistic decision model to airline seat inventory control. *Oper. Res.* 37(2):183–197.

Bernstein F, Kök AG, Xie L (2015) Dynamic assortment customization with limited inventories. *Manufacturing Service Oper. Management* 17(4):538–553.

Bertsimas D, Kallus N (2019) From predictive to prescriptive analytics. *Management Sci.* 66(3):1025–1044.

Bertsimas D, King A (2015) An algorithmic approach to linear regression. *Oper. Res.* 64(1):2–16.

Besbes O, Zeevi A (2015) On the (surprising) sufficiency of linear models for dynamic pricing with demand learning. *Management Sci.* 61(4):723–739.

Boyd S, Parikh N, Chu E, Peleato B, Eckstein J (2010) Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations Trends Machine Learn.* 3(1):1–122.

Bradley RC (2005) Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surveys* 2(2005): 107–144.

Breiman L (2001) Random forests. *Machine Learn.* 45(1):5–32.

Carvalho A, Puterman M (2005) Learning and pricing in an Internet environment with binomial demands. *J. Revenue Pricing Management* 3(4):320–336.

Chen N, Gallego G (2018) Nonparametric learning and optimization with covariates. Preprint, submitted May 3, https://arxiv.org/abs/1805.01136.

Chen X, Wang Y (2018) A note on tight lower bound for MNL-bandit assortment selection models. *Oper. Res. Lett.* 46(5):534–537.

Chen X, Wang Y, Zhou Y (2018) Dynamic assortment optimization with changing contextual information. Preprint, submitted October 31, https://arxiv.org/abs/1810.13069.

Cheung WC, Simchi-Levi D (2017) Thompson sampling for online personalized assortment optimization problems with

multinomial logit choice models. Preprint, submitted January 2017, https://www.researchgate.net/publication/321640292_Thompson_Sampling_for_Online_Personalized_Assortment_Optimization_Problems_with_Multinomial_Logit_Choice_Models.

Clifford S (2012) Shopper alert: Price may drop for you alone. *New York Times* (August 9), http://www.nytimes.com/2012/08/10/business/supermarkets-try-customizing-prices-for-shoppers.html.

Cohen MC, Lobel I, Paes Leme R (2016) Feature-based dynamic pricing. *Proc. ACM Conf. Econom. Comput.*(Association for Computing Machinery, New York), 817.

Du C, Cooper WL, Wang Z (2016) Optimal pricing for a multinomial logit choice model with network effects. *Oper. Res.* 64(2):441–455.

Fawcett T (2006) An introduction to roc analysis. *Pattern Recognition Lett.* 27(8):861–874.

Gallego G, Topaloglu H (2014) Constrained assortment optimization for the nested logit model. *Management Sci.* 60(10):2583–2601.

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.

Golrezaei N, Nazerzadeh H, Rusmevichientong P (2014) Real-time optimization of personalized assortments. *Management Sci.* 60(6):1532–1551.

Hastie T, Tibshirani R (1990) *Generalized Additive Models* (Chapman and Hall, London, United Kingdom).

IATA (2014) Resolution 787: Enhanced Airline Distribution (new). Accessed October 23, 2016, https://www.iata.org/contentassets/6de4dce5f38b45ce82b0db42acd23d1c/ndc-resolution-787.pdf.

Jagabathula S, Rusmevichientong P (2016) A nonparametric joint assortment and price choice model. *Management Sci.* 63(9):3128–3145.

Javanmard A, Nazerzadeh H (2019) Dynamic pricing in high-dimensions. *J. Machine Learn. Res.* 20(9):1–49.

Kök A, Fisher M, Vaidyanathan R (2015) Assortment planning: Review of literature and industry practice.Agrawal N, Smith S, eds. *Retail Supply Chain Management*, International Series in Operations Research & Management Science, vol. 223 (Springer, Boston), 175–236.

Lehmann EL, Casella G (1998) *Theory of Point Estimation*, Springer Texts in Statistics (Springer-Verlag, New York).

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

Li KJ, Jain S (2016) Behavior-based pricing: An analysis of the impact of peer-induced fairness. *Management Sci.* 62(9):2705–2721.

Linden G, Smith B, York J (2003) Amazon. com recommendations: Item-to-item collaborative filtering. *IEEE Internet Comput.* 7(1):76–80.

Luo X, Andrews M, Fang Z, Phang CW (2013) Mobile targeting. *Management Sci.* 60(7):1738–1756.

Murthi B, Sarkar S (2003) The role of the management sciences in research on personalization. *Management Sci.* 49(10):1344–1362.

Netessine S, Savin S, Xiao W (2006) Revenue management through dynamic cross selling in e-commerce retailing. *Oper. Res.* 54(5):893–913.

Qiang S, Bayati M (2016) Dynamic pricing with demand covariates. Preprint, submitted April 25, https://arxiv.org/abs/1604.07463.

Rao C, Toutenburg H, Shalabh, Heumann C (2008) *Linear Models and Generalizations: Least Squares and Alternatives* (Springer-Verlag, Berlin).

Rusmevichientong P, Shen Z, Shmoys D (2010) Dynamic assortment optimization with a multinomial logit choice model and a capacity constraint. *Oper. Res.* 58(6):1666–1680.

Samson PM (2000) Concentration of measure inequalities for Markov chains and φ-mixing processes. *Ann. Probab.* 28(1):416–461.

Sauré D, Zeevi A (2013) Optimal dynamic assortment planning with demand learning. *Manufacturing Service Oper. Management* 15(3):387–404.

Talluri K, van Ryzin G (2004) Revenue management under a general discrete choice model of consumer behavior. *Management Sci.* 50(1):15–33.

Train K (2009) *Discrete Choice Methods with Simulation*, 2nd ed. (Cambridge University Press, Cambridge, United Kingdom).

Ulu C, Honhon D, Alptekinoglu A (2012) Learning consumer tastes through dynamic assortments. *Oper. Res.* 60(4):833–849.

van der Vaart A, Wellner J (2000) *Weak Convergence and Empirical Processes: With Applications to Statistics* (Springer, New York).

van Ryzin G, Mahajan S (1999) On the relationship between inventory costs and variety benefits in retail assortments. *Management Sci.* 45(11):1496–1509.

Vershynin R (2012) Introduction to the non-asymptotic analysis of random matrices. Eldar Y, Kutyniok G, eds. *Compressed Sensing: Theory and Application* (Cambridge University Press, Cambridge, United Kingdom), 210–268.

Vulcano G, van Ryzin G, Ratliff R (2008) Estimating primary demand for substitutable products from sales transaction data. *Oper. Res.* 60(2):313–334.

Xue Z, Wang Z, Ettl M (2016) Pricing personalized bundles: A new approach and an empirical study. *Manufacturing Service Oper. Management* 18(1):51–68.

Zhang J, Krishnamurthi L (2004) Customizing promotions in online stores. *Marketing Sci.* 23(4):561–578.

Zhang JZ, Netzer O, Ansari A (2014) Dynamic targeted pricing in B2B relationships. *Marketing Sci.* 33(3):317–337.