A Statistical Learning Approach to Personalization in Revenue Management

Xi Chen[1], Zachary Owen[2], Clark Pixton[2], and David Simchi-Levi *[3]

[1]Leonard N. Stern School of Business, New York University
[2]Operations Research Center, Massachusetts Institute of Technology
[3]Engineering Systems Division, Department of Civil and Environmental Engineering and Operations Research Center, Massachusetts Institute of Technology

Abstract

We develop a general framework for modeling decision problems in which actions can be personalized by taking into account the information available to the decision maker. We demonstrate the application of our method to customized pricing and personalized assortment optimization. We show that learning under our model takes place reliably by establishing finite-sample convergence guarantees for model parameters which hold regardless of the number of customer types, which can be potentially uncountable. The parameter convergence guarantees are then extended to out-of-sample performance guarantees in decision problems. In particular, we provide 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 demand distribution. We conduct simulated experiments to demonstrate the performance of our method. We further test our method on real transaction data for airline seating reservations and show that our method is competitive with more sophisticated and computationally intensive methods while enjoying theoretical backing that these methods do not.

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 appropriately. This benefit is two-fold: 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 to 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.

*Corresponding author, please contact at dslevi@mit.edu

To implement a personalization strategy it has been widely proposed to divide the customer base into distinct segments and to tailor the service to each type appropriately as in Gurvich et al. [2009] and Bernstein et al. [2011], for example. This segmentation of customers is often accomplished in practice by dividing them into archetypical categories based on broad, easily observable characteristics such as business versus leisure travelers in the case of the airline industry as discussed in Talluri and van Ryzin [2004b]. Such methods have the potential to increase revenue over systems in which customers are assumed to be homogeneous; however, in practice, customer characteristics often vary smoothly over the population, which makes it difficult to cleanly separate customers along the axis of interest.

In this paper, we consider statistical models that use observed contextual information, rather than previously defined customer segments, to inform decisions. This approach represents a shift from thinking about personalization in terms of customer types towards personalized decisions as a function of the unique relevant information available at the time of each decision. In the language of machine learning, we build our personalized decision making framework based on a model learned via supervised learning techniques. In particular, past transactional data is used to discover the underlying relationship between customer behavior and the action of the decision maker as well as the contextual information, and predictions are made based on this relationship. This approach takes advantage of available data in at least two ways. First, it allows the seller to consider more complex relationships between context and customer behavior than those that can be captured with just a few customer segments. Second, our approach also allows previous learning to be generalized easily, even to previously unseen types, allowing for customization in the case of novel customer data.

As a demonstration of the advantages of this approach, consider again the example of business and leisure travelers in the airline industry. Within the business customer segment, there are large businesses with many assets and low price sensitivity, but also small firms that may not yet be financially established. Similarly, leisure travelers are often more price sensitive, but there are some wealthy travelers who may behave more like business customers. Transaction-specific contextual information is often rich enough to capture these "second-order" trends within the broad categories, which can then be leveraged to drive incremental revenue or to improve customer experience. With the granularity of information available today, it is likely that each new customer represents a unique pattern of information, and our personalization strategy should be flexible enough to both learn from and optimize for the full diversity of the customer population.

To implement this type of strategy, we model demand or customer choice with a binomial or multinomial logit function, where the arguments to the logit function are dependent on both observed customer attributes and seller decisions. This approach leads to a practical, data-driven algorithm with good theoretical guarantees. Our main contributions are summarized as follows:

1. *General Framework for Personalized Revenue Management:* We propose a general framework capturing a range of revenue management problems in which personalized contextual information is available. Instead of requiring prior knowledge of customer segments or using clustering procedures, our framework assumes a parametric form for the customer choice model which incorporates each customer attribute. This form permits information on the effect of attributes to be pooled across the sample, eliminating the need to estimate separately in every segment which significantly reduces the sample size. In the proposed framework, we allow for the estimation of customer choice across arbitrary types, even for customers with attribute values not before observed as well as the case in which the number of attributes greatly ex-

ceeds the size of previously observed transaction data (the so-called high-dimensional setting). In contrast to popular statistical machine learning models in which data of interest consist of contextual information (or attributes) and outcomes, *in revenue management problems, there is a third element that must be accounted for — the action of the seller.* This action may be a price set by a firm or an assortment offered to a customer, and it has an important effect on the observed outcome of the transaction. We provide ways to incorporate these actions into learning algorithms for customer behavior models.

2. *Statistical Bound on the Revenue Gap:* Given the model, we estimate the parameters by regularized maximum likelihood estimation. In contrast to statistical machine learning problems in which one cares about the generalization error, our goal is to provide an upper bound on the gap between the expected revenue of the proposed method and the *oracle revenue*, which holds *with high probability.* The oracle revenue is achieved by making the optimal decision with full knowledge of the true parameters specifying consumer behavior. In contrast to standard asymptotical results, our finite-sample bound gives the explicit number of transactional records needed to achieve any given level of revenue loss as compared to the oracle decision maker. Such a bound provides practical insights since it characterizes the trade-off between the cost of information, given by the cost of minimum number of transactional records to be collected, and the potential revenue loss. Further, our high-probability bound has an advantage over the usual expected loss bound since it is more robust against unexpected realizations of customers' choice. Technically speaking, we derive the desired revenue gap in two steps: (1) we first provide the bound on the distance between the estimated model parameters and true parameters holding with high probability; (2) we generalize the bound on estimation error into a bound on the revenue gap by exploiting properties of the revenue function. For the first step, although classical M-estimation theory has established asymptotic normality of the estimated parameter (see for example Theorem 3.2.16 in [van der Vaart and Wellner, 2000]), it fails to provide the finite-sample bound of the estimation error. We establish such a bound in this paper, which is used to derive the rate at which the revenue gap shrinks to zero as the number of transactional records grows. We further extend our theoretical results to two more practical and more challenging cases:

   (a) The misspecified model case, in which customer behavior does not in fact follow the assumed model. These results demonstrate the robustness of our method in the case when reality deviates from our models.

   (b) The high-dimensional case, in which the number of customer attributes is substantially larger than the number of collected transactional records. This case will be useful in the "big-data" scenario, where all the historical transactional information of a customer has been recorded, which can be represented by a high-dimensional attribute vector.

To illustrate the proposed framework for personalized revenue management, we consider two fundamental applications throughout the paper: customized pricing and personalized assortment optimization. In customized pricing a seller aims to offer her product to each customer at a price that maximizes expected profit, given information she has obtained about the transaction and the customer. Customized pricing, or price discrimination, is common in business to business transactions. Despite its limited application to general settings due to customer satisfaction and legal issues, it is predicted that the practice will become more widely accepted as more data becomes available (see Golrezaei et al. [2014]). In personalized assortment optimization, a seller aims to show to each customer the assortment of products which maximizes her expected profit for each individual customer arrival, given the contextual information. In the traditional brick-and-mortar context, implementation of a personalized assortment strategy is infeasible due to prohibitive setup costs. However, for online retailers personalized assortments are possible and even natural. To facilitate the illustration of our framework without unnecessary complication, we only consider the most classical setups for both problems, i.e., single-stage pricing and multinomial logit model (MNL) based assortment optimization (see Phillips [2005] or Talluri and van Ryzin [2004b]).

We demonstrate the performance of the proposed method for both customized pricing and personalized assortment optimization problems on simulated data. Further, we test the performance of our algorithm for customized pricing on transaction data from a European airline, and show a significant increase in revenue over the best single-price policy. However, our style of analysis is not limited to these specific instances, and can be applied in many other contexts such as online advertisement allocation, crowdsourcing task assignment, personalized medicine, and other problems in which personalization can aid in optimal decision-making.

2 Literature Review

Our work combines two recent themes within operations management: learning model parameters from historical data and making personalized decisions using contextual information. The primary applications of personalization that we consider here, customized pricing and personalized assortment optimization, have been extensively studied over the past decade and there is a vast body of related work that is out of scope for this summary. Therefore, we provide a brief review of papers in each application area, i.e., customized pricing and assortment optimization, which focus on learning and personalization.

Dynamic pricing with demand learning has been a popular theme in the revenue management literature; a brief survey of some early papers in this field can be found in Aviv et al. [2012]. In these settings it is often assumed that the demand function of the entire population is unknown, but it is possible to obtain information about the structure of demand through price experimentation. A common modeling assumption, (as used in Broder and Rusmevichientong [2012] and Keskin and Zeevi [2014], for example), is that the true market demand function is specified by a parametric choice model. Using price experimentation, they develop pricing policies that work to minimize regret in comparison to a clairvoyant who knows the full demand model. Another approach to learning and dynamic pricing was recently explored in Bertsimas and Vayanos [2014]. The authors formulated a robust optimization problem that captures the exploration-exploitation trade-off in dynamic pricing with unknown demand, and provide a tractable approximation. While their approach studied a more complicated setup of multi-stage dynamic pricing, they do not have performance guarantees and do not consider personalization using contextual information. In many practical scenarios, implementing frequent price changes may be infeasible due to business constraints. To address this Cheung et al. [2014] propose an algorithm which achieves strong performance with a limited number of such price adjustments.

At least in the field of operations research and management science, literature on personalized pricing is sparse. Carvalho and Puterman [2005] studied a multi-stage pricing problem and assumed a logit model for demand as a function of the offered price, and they suggested that their model could be extended to include customer-specific attributes. Aydin and Ziya [2009] considered 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 considered 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.

In many cases, models of price discrimination are actually cast as multi-product models, where the different price levels come with different qualifications and extras as in the airline industry. See Talluri and van Ryzin [2004a] and Belobaba [1989] for examples of this type. While in this paper we focus on customized pricing for a single product based on customer attributes, which would be appropriate in the insurance industry and in business to business transactions, our conception of contextual information is general enough to encompass other formulations. For example our model can also be applied in the case of dynamic pricing of differing products based on the individual attributes of each product, the effects of which may be learned over time.

Assortment optimization in the static case 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. [2008]. Some setups such as those of Talluri and van Ryzin [2004a] and Golrezaei et al. [2014] allow for a general model of customer choice, but others study structural properties specific to certain choice models. The multinomial logit model (MNL), the model that we use in this paper, is among the most commonly studied models of customer choice for assortment optimization. It is desirable for its simplicity and its flexibility in incorporating different effects on customer choice (see Du et al. [2015] for a recent example). There are several extensions of the MNL model, e.g., the mixed MNL model [McFadden and Train, 2000], the capacitated MNL [Rusmevichientong et al., 2010], the MNL with random choice parameters [Rusmevichientong et al., 2014], and the multi-level nested logit model [Li et al., 2015]. Our framework for personalized assortment optimization based on MNL can further be extended to these variations of the MNL model.

Recently, researchers have investigated the problem of learning choice models in assortment optimization. For example, Ulu et al. [2012] used Bayesian updates in a dynamic programming framework to balance learning customer choice probabilities against earning short-term revenue in offering assortments to customers. Vulcano et al. [2008] proposed an algorithm for estimating true demand from censored transaction data and proved convergence of their algorithm. Also, in Rusmevichientong et al. [2010], the authors considered a constrained version of assortment optimization and proposed a dynamic algorithm that incorporates learning the parameters of the MNL model. However, these papers have not incorporated the attributes of different customers into the modeling.

There are several recent papers that have studied personalized assortment optimization by assuming multiple customer types. For example, Bernstein et al. [2011] approached personalized assortment optimization by establishing one choice model for each customer type and showed properties of an optimal assortment policy for two products in the presence of inventory considerations. Golrezaei et al. [2014] also considered multiple customer types, providing a practical algorithm for personalization under inventory constraints and proving a strong worst-case performance bound. They discussed the issue of estimating model parameters, and showed that a version of their worst-case competitive ratio bound holds for the case when parameters are estimated from the data.

While not exactly personalization, we note a commonly-made assumption in the literature: 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. Our model generalizes this notion by specifying mean utility as linear in attributes of each *customer*. This setup allows for product effects to be modeled as attributes common to all customers but differing between products.

In addition to the literature of revenue management, we would like to discuss two recent related papers. Bertsimas and Kallus [2014] developed 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, while in our framework, the action is treated as both historical contextual information *and* a decision variable. Independently, Rudin and Vahn [2014] studied the newsvendor problem with feature information. However, as compared to our problem, the newsvendor problem has a simpler structure since observed demands are independent of order quantity decisions. Also, the obtained theoretical results in Rudin and Vahn [2014] involve generalization error, which is not an appropriate measure of performance in our case as we will explain at the beginning of Section 5.

The rest of the paper will proceed as follows. We give our approach and models in Section 3. In Section 4 we present our algorithm and describe its application to customized pricing and assortment optimization. Sections 5 and 6 are devoted to proving revenue bounds for the two problems under various assumptions, including a high-dimensional result. In Section 7 we show the results of experiments on both simulated and real data.

3 A General Model

In this section we present a general modeling framework for data-driven revenue management problems that include decision-specific context information, or customer features (we will use the terms 'feature' and 'attribute' interchangeably). We consider two applications in detail, namely customized pricing and assortment optimization, but we note that our approach is certainly not limited to these domains.

A decision maker observes a vector of features $z \in \mathcal{Z} \subseteq \mathbb{R}^m$ that encodes information about the context of the specific decision at hand. Taking into account z, she chooses an action a from a problem-specific action space $\mathcal{A}$. After the decision has been made, she observes an outcome y from a finite set $\mathcal{Y}$ and gains a random reward from a finite set $\mathcal{R}$. The probability of outcome y depends on both the feature vector z and the decision a. The reward $r_a(y)$ for outcome y may also depend on the decision a. She would like to make the decision that maximizes her expected reward given the context z.

The key to our modeling framework is a method for capturing the interaction between features, decision, and outcome using a logit model. This method gives conditional outcome probabilities $\mathbb{P}_z(y; a)$. We will derive the specific form for outcome probabilities as a function of features and decision for both customized pricing and assortment optimization later in the section.

Given the outcome probabilities from the problem-specific logit model, we can write the expected reward/revenue:

$$f_z(a) = \sum_{y \in \mathcal{Y}} r_a(y) \mathbb{P}_z(y; a). \tag{1}$$

The algorithm we present estimates the expected reward and then maximizes over all possible decisions. Before stating the algorithm in detail, we give specific models for customized pricing and personalized assortment optimization.

3.1 Customized Pricing Model

The first application of our model is in the case where a seller has a single product without inventory constraints and wishes to offer a price that will maximize her revenue. In basic form, the single-product pricing problem consists of a set $\mathcal{A} = \{p^1, \ldots, p^K\}$ of K distinct prices and a probability of purchase $\mathbb{P}(y = 1; p^k)$, $k = 1, \ldots, K$. Here the outcome y is a binary decision and is equal to one if the customer purchases the product and zero otherwise. Thus, the expected revenue function is $f(p) = p\mathbb{P}(y = 1; p)$. Without any other information, the seller would maximize $f(\cdot)$ over $p \in \mathcal{A}$.

In the customized pricing problem, we assume that the seller has the ability to offer a different price $p \in \mathcal{A}$ to each customer after observing the customer's associated feature vector $z \in \mathcal{Z} \subseteq \mathbb{R}^m$, which supplies the context for the current pricing decision. We assume that z contains a unit intercept term common to all customers that allows us to learn universal effects.

This structure allows us to define a *personalized* demand function $\mathbb{P}_z(y = 1; p)$, which is the probability that a customer with the feature vector z purchases the product at price p. We model purchase probabilities $\mathbb{P}_z(y = 1; p)$ using a logit model with true parameters $\beta^* \in \mathbb{R}^{K-1}$ and $\gamma^* \in \mathbb{R}^m$:

$$\log \frac{\mathbb{P}_z(y = 1; p, \beta^*, \gamma^*)}{1 - \mathbb{P}_z(y = 1; p, \beta^*, \gamma^*)} = \sum_{k=2}^{K} \beta_k^* \mathbb{I}(p = p^k) + \sum_{j=1}^{m} \gamma_j^* z_j, \tag{2}$$

where $\mathbb{I}(A)$ is the indicator function which takes the value 1 when the event A is true and 0 otherwise. We also note that since the price p is in a discrete set $\mathcal{A}$, we represent the price factor p as a group of dummy variables $(\mathbb{I}(p = p^2), \ldots, \mathbb{I}(p = p^K))$. As is common in linear models, to improve computational properties the effect of the highest price, p_1, is incorporated into the intercept instead of defining a separate effect β_1^*. For simplicity of notation, let $x = (\mathbb{I}(p = p^2), \ldots, \mathbb{I}(p = p^K), z) \in \mathbb{R}^{m+K-1}$ be the entire feature vector that contains both the seller's decision and customer features, let $\theta^* = (\beta^*, \gamma^*) \in \mathbb{R}^{K+m-1}$ be the entire true parameter vector, and define the total dimensionality $d = K + m - 1$. The logit model in (2) gives a specific form for the personalized demand function:

$$\mathbb{P}_z(y = 1; p, \theta^*) = \frac{1}{1 + \exp\left(-\left(\sum_{k=2}^{K} \beta_k^* \mathbb{I}(p = p^k) + \sum_{j=1}^{m} \gamma_j^* z_j\right)\right)} = \frac{1}{1 + \exp\left(-\langle x, \theta^* \rangle\right)}, \tag{3}$$

where $\mathbb{P}_z(y = 1; p, \theta^*)$ accomplishes the goal of linking customer attributes, seller decisions, and customer choice without defining any customer segmentation. To complete the model for customized pricing, the expected reward is given by

$$f_z(p, \theta^*) := p \; \mathbb{P}_z(y = 1; p, \theta^*), \tag{4}$$

We pause to make explicit the connection between customized pricing notation and the general notation in (1). For customized pricing, each action a is a price $p \in \mathcal{A}$. As mentioned above, outcomes $y \in \{0, 1\}$ correspond to purchase decisions, and reward $r_a(y) = r_p(y) = p$ if $y = 1$ and zero if $y = 0$. Using these mappings one can identify $\mathbb{P}_z(y = 1; p, \theta^*)$ and $f_z(p, \theta^*)$ as the problem-specific versions of $\mathbb{P}_z(y; a)$ and $f_z(a)$.

Remark 3.1. *Our model of the price effect via the parameters $\beta^* \in \mathbb{R}^{K-1}$ is simply for clarity of exposition and to highlight the model of the seller's decision (i.e., price) as a feature. In practice, this model directly extends to modeling interaction effects between offered prices and other features (e.g,. via introducing extra features pz_j). 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 our real transaction data. Further, our representation of the price in a finite set $\mathcal{A}$ as a group of dummy variables is a common way of modeling categorical feature in multifactor analysis-of-variance [Rao et al., 2008]. It can also be easily adapted to a continuous price set by using a single price parameter β.*

3.2 Personalized Assortment Optimization Model

Our framework also applies to the problem of personalized assortment optimization. For background, we introduce the standard random utility assortment optimization problem following the notation from [Rusmevichientong et al., 2014]. In this problem, the decision maker has J products $[J] \triangleq \{1, 2, \ldots, J\}$, with another "no-purchase" option. For each product j, let r_j be its associated revenue, fixed a priori and indexed such that $r_J \geq r_{J-1} \geq \ldots \geq r_1 \geq r_0 = 0$, where $r_0 = 0$ is the revenue for the no-purchase option. Whereas in the previous application the decision variable was price, here the decision maker must choose an *assortment* $S \in \mathcal{A}$ to show to the customer, where $\mathcal{A}$ is the collection of all feasible assortments.

In the general form of the problem, we assume that customers choose among the products according to some probability $\mathbb{P}(j; S)$, which denotes the probability that a customer chooses product j given that she was shown assortment S. The expected revenue for a given assortment S (without any contextual information) can then be written as

$$f(S) = \sum_{j \in S} r_j \mathbb{P}(j; S), \tag{5}$$

and the decision maker maximizes $f(S)$ over $S \in \mathcal{A}$.

A popular way to model the customer choice probability $\mathbb{P}(j; S)$ is to utilize the random utility model. That is, one assumes that a customer has the utility $U_j = V_j + \epsilon_j$ for each product j. Here, ϵ_j is a standard Gumbel random variable with mean zero, and V_j can be interpreted as the mean utility of product j. Without loss of generality, the utility V_0 of the no-purchase option is set to be zero. Given that the decision maker offers the assortment $S \in \mathcal{A}$ of products, the customer will choose the product in S with the highest U_j if this number is positive, but if all utilities are negative she leaves without purchasing any product. It is a well-known result in discrete choice theory [Train, 2009] that in this setting, a customer chooses product $j \in S$ with probability

$$\mathbb{P}(j; S) = \frac{e^{V_j}}{1 + \sum_{l \in S} e^{V_l}}. \tag{6}$$

The choice model in (6) is known as the multinomial logit (MNL) model, and is widely used to model discrete choice [Train, 2009]. It is an extension of the binary logit model to cases with more than two alternatives. Inserting the MNL choice probabilities (6) into (5) gives an expected revenue function in terms of the parameters $\{V_j\}_{j \in [J]}$ which can be estimated from data.

For personalized assortment optimization, suppose that before choosing an assortment S, the decision maker observes customer feature vector $z \in \mathcal{Z} \subseteq \mathbb{R}^d$. We can modify the assumptions on the choice probabilities to take into account this new customer information. In particular, we assume that each feature vector z corresponds to a different *personalized utility* for each product j, $U_j^z = V_j^z + \epsilon_j$, where ϵ_j is again a standard Gumbel random variable and we view V_j^z as the mean personalized utility of product j *for the customer with feature vector* z. We assume this mean utility is given by a linear model $V_j^z = \langle \theta_j^*, z \rangle$, where $\theta_j^* \in \mathbb{R}^d$ for $1 \le j \le J$. This parallels our assumption in Eq. (3), and reduces the parameter space from an arbitrary set to a dJ-dimensional linear space, thus allowing us to generalize learning from previous transactions to future customers with new feature vectors. With this structural assumption, we can modify (6) for the personalized case:

$$\mathbb{P}_z(j; S, \theta^*) = \frac{e^{V_j^z}}{1 + \sum_{l \in S} e^{V_l^z}} = \frac{\exp\{\langle \theta_j^*, z \rangle\}}{1 + \sum_{l \in S} \exp\{\langle \theta_l^*, z \rangle\}}. \tag{7}$$

Let $\theta^* = (\theta_1^*, \ldots, \theta_J^*) \in \mathbb{R}^{dJ}$ be the parameters for all J products. Offering the assortment S to a customer with feature vector z gives the following expected revenue:

$$f_z(S, \theta^*) = \sum_{j \in S} r_j \mathbb{P}_z(j; S, \theta^*). \tag{8}$$

As in the previous subsection, we note the mapping between the general framework notation and the notation specific to the personalized assortment optimization problem. The action a in this case is the offered assortment S, with action space $\mathcal{A}$ being the set of feasible assortments. The possible outcomes y are products with $y \in \{0, \ldots, J\}$, and thus $r_a(y) = r_S(j)$ is equal to r_j if $j \in S$ and zero otherwise. Using these mappings one can identify $\mathbb{P}_z(j; S, \theta^*)$ and $f_z(S, \theta^*)$ as the problem-specific versions of $\mathbb{P}_z(y; a)$ and $f_z(a)$.

Remark 3.2. *Although we assume a logit model for purchasing probability in (2) and linear model for the mean personalized utility $V_j^z = \langle \theta_j^*, z \rangle$ for simplicity of exposition, these models can be easily extended to more complicated semi-parametric or non-parametric models. For example, we could assume an additive model on V_j^z, i.e., $V_j^z = \sum_{l=1}^d g_l(z_l)$, where g_l is a smooth function associated with the l-th feature of z, which can be further approximated by a series of truncated power basis of splines (see, e.g., Hastie and Tibshirani [1990]).*

Our general model captures the relationship between features, decisions, and outcomes. Because we assume that past feature-decision-outcome data are observed and recorded, these models give rise to practical algorithms via maximum likelihood estimation for learning the underlying parameters, which facilitates the personalized decision-making for new customers, as in the next section.

4 Algorithm

With the model in place, we now present the general Personalized Revenue Maximization (PRM) algorithm in Algorithm 1. We assume the decision maker has access to a set $\mathcal{T} = \{(z_1, a_1, y_1), \ldots (z_n, a_n, y_n)\}$ of n samples of past features, decisions, and outcomes which serve as input to the algorithm. This is realistic, especially in the realm of online commerce. In preparation for maximum likelihood estimation (MLE), we calculate the negative log-likelihood $\ell_n(\mathcal{T}; \theta) = -\frac{1}{n} \sum_{i=1}^n \log\left(\mathbb{P}_{z_i}(y_i; a_i, \theta)\right)$, where the concrete form of $P_{z_i}(y_i; a_i, \theta)$ is given in (3) for customized pricing and (7) for personalized assortment optimization. We further pre-determine a positive number R such that $\|\theta^*\|_1 \le R$, where $\|v\|_1 = \sum_i |v_i|$ denotes the vector ℓ_1-norm, and adopt the regularized/constrained MLE with the constraint that $\|\theta\|_1 \le R$ (see (9)). 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 and to facilitate theoretical analysis, and it often leads to better empirical performance. In the

Algorithm 1 Personalized Revenue Maximization (PRM) Algorithm

Input: Data samples $\mathcal{T} = \{(z_1, a_1, y_1), \dots (z_n, a_n, y_n)\}$, regularization parameter R

Output: Decision policy $h : \mathcal{Z} \to \mathcal{A}$

1. Fit the regularized MLE on the observed data:

$$\widehat{\theta} = \arg\min_{\|\theta\|_1 \le R} \left[\ell_n(\mathcal{T}; \theta) = -\frac{1}{n} \sum_{i=1}^{n} \log\left(\mathbb{P}_{z_i}(y_i; a_i, \theta)\right) \right] \tag{9}$$

2. Obtain the estimate of outcome probabilities $\mathbb{P}_z(y; a, \widehat{\theta})$ for every $z \in \mathcal{Z}$ and $a \in \mathcal{A}$.

3. Construct the decision policy $h : \mathcal{Z} \to \mathcal{A}$ as $h(z) = \widehat{a}$ where

$$\widehat{a} = \arg\max_{a \in \mathcal{A}} f_z(a, \widehat{\theta}). \tag{10}$$

---

high-dimensional setting as discussed in Section 6.2, where the dimension of the attribute vector z is large compared to the number of observed transactions, ℓ_1-norm regularization is essential for both empirical performance and theoretical justification.

In step 1 of Algorithm 1, we minimize $\ell_n(\mathcal{T}; \theta)$ over θ under ℓ_1-regularization to get an estimate for θ. For both customized pricing and personalized assortment optimization, it is easy to see that $\ell_n(\mathcal{T}; \theta)$ is a convex function and ℓ_1-regularization is a convex constraint on θ. Therefore, any fast convex optimization procedure (e.g., accelerated projected gradient descent, alternating direction method of multipliers, etc) can be adopted to solve the optimization problem in (9). The reader might refer to F. Bach and Obozinski [2011] and S. Boyd and Eckstein [2010] for recent developments on ℓ_1-regularized convex optimization algorithms.

Given the estimated $\widehat{\theta}$, in step 2 we calculate the estimated outcome probabilities $\mathbb{P}_z(y; a, \widehat{\theta})$ for decision a given features z, which is used to approximate the expected reward $f_z(a, \theta^*)$ in (1). In step 3, we construct the decision policy to maximize the approximated expected reward. For customized pricing, the decision problem in step 3 can be simply solved via optimization over the finite price set $\mathcal{A}$. For assortment optimization, given the estimates of each parameter vector $\widehat{\theta}_j$, the estimated mean utility V_j^z is taken to be deterministic in the final step of policy construction. When $\mathcal{A}$ is taken to be all subsets of J products, the maximization over the power set of subsets of products can be reduced to maximization over *revenue-ordered assortments*, that is subsets of the form $S_k = \{1, \dots, k\}$ for some $k \in \{1, 2, \dots, J\}$, as demonstrated in Talluri and van Ryzin [2004a]. Further, since we assume that customer behavior is specified by the true parameter vectors θ_j^*, it follows that the optimal such set is also a revenue-ordered assortment. Thus this property ensures that our maximization scales linearly with number of products J and that both our estimated optimal assortment and the true optimal assortment belong to the same limited subset of $\mathcal{A}$. This reduction of the search space greatly reduces the computational cost.

Remark 4.1. *The algorithm is presented here in its full generality, however it can be easily adapted to incorporate problem-specific information through the use of additional constraints. For instance in the case of customized pricing, it is natural in many cases to assume that a higher price implies lower demand given the same context z. Let the prices in the candidate set $\mathcal{A}$ be ordered such* *that $p^K < p^{K-1} < \cdots < p^1$. To impose such an assumption, one can simply add an additional constraint to the regularized maximum likelihood estimation procedure in (9), i.e., β lies in the isotonic cone $\{\beta \in \mathbb{R}^{K-1} | 0 \leq \beta_2 \leq \ldots \leq \beta_K\}$. As we demonstrate in our numerical experiments, such extra information can be practically useful in aiding the learning algorithm, especially when the number of observations is small. In the case of assortment optimization, other papers such as Rusmevichientong et al. [2010] consider the maximization with constraints on $\mathcal{A}$ and give tractable methods of optimization. Our personalization scheme can be directly extended to these more general MNL-based assortment planning models through the use of additional constraints.*

5 Theory: Well-specified Model Setting

We now proceed to derive high-probability guarantees on the performance of the Algorithm 1 (PRM) in practical revenue management applications. In this section, we assume that the logit model is well-specified, that is there exist parameters θ^* such that for all observed data $i \in \{1, \ldots, n\}$, $\mathbb{P}(y_i) = \mathbb{P}_{z_i}(y_i; a_i, \theta^*)$, which means the logit model is the correct underlying model of outcome probabilities. In Section 6 we will discuss how our results can be adapted to cases in which transaction data does *not* follow such a logit model.

We use as our benchmark the oracle policy that knows this true θ^* and so is able to select the action a^* that maximizes expected revenue. Let $\widehat{a}$ be the action recommended by Algorithm 1 in (10). Using properties of the maximum likelihood estimates, for any feature vector z, our goal is to find a bound on the optimality gap $f_z(a^*, \theta^*) - f_z(\widehat{a}, \theta^*)$ which holds with high probability.

It is interesting to make a contrast to the *generalization error* bound which is important in the statistical machine learning literature (see, e.g., Bousquet et al. [2004]). In particular, a generalization error bound is an upper bound on the difference between the expected reward and the empirical reward on the historical data that holds for any action policy $\widehat{a} : \mathcal{Z} \to \mathcal{A}$, i.e., $|f_z(\widehat{a}) - \frac{1}{n}\sum_{i=1}^n r_{\widehat{a}}(y_i)|$ or $|\mathbb{E}_z(f_z(\widehat{a})) - \frac{1}{n}\sum_{i=1}^n r_{\widehat{a}}(y_i)|$ when z is random. Unfortunately, such a bound is not useful in our case since it requires that out-of-sample decision and all the past decisions are based on the same policy. However, for each historical data record i, the decision-maker might not follow the optimized action policy $\widehat{a}$ to offer the price. For example, a randomized price experiment might be adopted in order to better learn the decision function. Also, the generalization error bound is conservative in the sense that it requires the upper bound holding for *any policy*. Finally, it fails to provide an estimate on the potential revenue loss as compared to the oracle decision for a new customer with feature z, which is of interest in revenue management problems.

Before we prove our bound on the optimality gap in terms of the number of samples n, we first provide some necessary notation and preliminaries.

5.1 Notation and Preliminaries

We use $\|v\|_p$ to denote the p-norm of the vector v, given by $\|v\|_p = (\sum_i |v_i|^p)^{\frac{1}{p}}$ for $p > 0$ and $\|v\|_\infty = \max_i |v_i|$. Further, let $\lambda_{\max}(\Sigma)$ and $\lambda_{\min}(\Sigma)$ denote the largest and smallest eigenvalues of the matrix Σ, respectively, and let $\|\Sigma\|_{op}$ be the operator norm of Σ. For a symmetric positive semi-definite matrix Σ, we have $\|\Sigma\|_{\text{op}} = \lambda_{\max}(\Sigma)$. As our analysis focuses on rates of convergence, we will use c, C, and C' to denote universal constants. These constants may have different meanings in different sections, and so for our main results we will attach the subscripts *cp* and *ao* to the constants to differentiate between customized pricing and assortment optimization. Also, for conciseness we use $[n]$ to denote the set $\{1, \ldots, n\}$.

In our theoretical analysis, we assume that either the customers' feature vectors are fixed, in the deterministic design setting, or follow a sub-Gaussian distribution in the randomized design setting. The sub-Gaussian assumption on the feature vectors is a common and natural assumption for regression analysis since it captures a wide range of multivariate distributions. Examples include the multivariate Gaussian distribution, the multivariate Bernoulli distribution, the spherical distribution (for modelling normalized unit-norm feature vectors), and a uniform distribution on a convex set among many others. We briefly introduce sub-Gaussian random variables and vectors here and readers may refer to Vershynin [2012] for more details.

Formally, a sub-Gaussian random variable X is a random variable with moments that satisfy $(\mathbb{E}\,|X|^p)^{\frac{1}{p}} \leq K\sqrt{p}$ for all $p \geq 1$ for some $K > 0$. The corresponding sub-Gaussian norm $\psi_X = \|X\|_{\psi_2}$ is the smallest K for which the moment condition holds, i.e., $\|X\|_{\psi_2} = \sup_{p\geq 1} p^{-1/2}(\mathbb{E}\,|X|^p)^{\frac{1}{p}}$. It can be proven that such a moment condition is equivalent to a more natural tail condition of X that is similar to the super-exponential tail bound of a Gaussian random variable, i.e., $\mathbb{P}(|X| \geq t) \leq \exp\left(1 - ct^2/\|X\|_{\psi}^2\right)$ for some constant c (see Lemma 5.5. in Vershynin [2012]). Gaussian random variables and any bounded random variable (e.g., Bernoulli) are special cases of sub-Gaussian random variables. Given the definition of sub-Gaussian random variable, a random *vector* $x \in \mathbb{R}^d$ is sub-Gaussian if its one-dimensional marginals $\langle x, w\rangle$ are sub-Gaussian random variables for all $w \in \mathbb{R}^d$. The corresponding sub-Gaussian norm is defined by $\psi_x = \sup_{\|w\|_2\leq 1} \|\langle x, w\rangle\|_{\psi_2}$.

5.2 Theoretical Results for Customized Pricing

In this section, we will give the detailed revenue bound for the case of the customized pricing problem. We emphasize that for now we are considering the fixed-dimensional regime, where the dimensionality d remains fixed and $d < n$. The high-dimensional case will be studied in Section 6.2.

We consider both deterministic (or fixed) design, where the input feature vectors z_i are viewed as fixed quantities, and random design, where inputs z_i are randomly drawn from some distribution. Both of these design assumptions are popular in regression analysis (see Rao et al. [2008]). We formally state our assumptions before moving to the results.

Assumption 1. *For both deterministic and random design settings, we assume*

1. *Conditional independence: the observed outcomes $\{y_i\}_{i=1}^n$ are independent given each x_i (see definition of x_i in Section 3.1).*

2. *Bounded feature vector: there exists a universal constant $B' > 0$ such that for any customer feature vector z, $|z_j| \leq B'$ for all $j \in [m]$. This further implies that $|x_j| \leq \max(B', 1) \triangleq B$.*

This assumption is always satisfied in practice and guarantees that the transaction data used will not contain arbitrarily large elements which could have an outsize effect on our learning procedure. The remainder of our assumptions differ between the fixed and random design settings.

Assumption 2 (Deterministic Design)**.** *There exists a constant ρ such that $\lambda_{\min}(\Sigma_n) \geq \frac{\rho}{2} > 0$, where $\Sigma_n = \frac{1}{n}\sum_{i=1}^n x_i x_i^T$ is the sample Gram matrix.*

Assumption 3 (Random Design)**.** *1. The vectors $\{x_i\}_{i=1}^n$ are independent and identically distributed, following a sub-Gaussian distribution with the sub-Gaussian norm ψ.*

2. There exists a universal constant ρ such that $\lambda_{\min}(\Sigma) > \rho > 0$, where $\Sigma = \mathbb{E}(xx^T)$.

We note that the positive lower bound on either the sample Gram matrix Σ_n or the population second moment matrix Σ can easily be satisfied when the sample size $n > d$. The sub-Gaussian assumption for random design in Assumption 3 is satisfied as long as customer feature vectors z are sub-Gaussian with sub-Gaussian norm $\psi_z = \psi - 1$. We provide the justification of this simple claim in the online supplement to this paper. Finally, companies sometimes use periods of price experimentation, in which prices are offered at random, for the purpose of learning demand. In this case, the i.i.d. assumption on the vectors $\{x_i\}_{i=1}^n$ is satisfied at least for the price information that the x_i contain.

5.2.1 Estimation Error Bound

To establish the revenue gap, we first establish the rate of convergence of the parameter estimates $\widehat{\theta}$ provided by Algorithm 1 to the true parameters θ^*. Specifically, under Assumptions 1, 2, and 3 in the low dimensional setting, we will prove that $\|\widehat{\theta} - \theta^*\|_2 \leq C\sqrt{\frac{\log n}{n}}$ 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 $\frac{1}{\sqrt{n}}$. Using smoothness of our revenue functions, we will show that this rate of convergence of parameters can be translated into the revenue space. Below is the formal theorem statement.

Theorem 5.1 (Parameter Convergence Rate)**.** *In the deterministic design setting under Assumptions 1 and 2, we have with probability at least $1 - \frac{1}{n}$,*

$$\|\widehat{\theta} - \theta^*\|_2 \leq c\frac{B}{\rho}\frac{(1 + \exp(RB))^2}{\exp(RB)}\sqrt{\frac{d\log(nd)}{n}}.$$

In the randomized design setting under Assumptions 1 and 3, as long as $n \geq \frac{4C_{cp}(\psi)\log(n)d}{\min(\rho,1)^2}$ for some constant $C_{cp}(\psi)$ only depending on ψ, w.p. at least $1 - \frac{1}{n} - 2(\frac{1}{n})^d$,

$$\|\widehat{\theta} - \theta^*\|_2 \leq c\frac{\psi}{\rho}\frac{(1 + \exp(RB))^2}{\exp(RB)}\sqrt{\frac{d\log(nd)}{n}},$$

where c is a universal constant.

We note that from now on, we suppress the data argument $\mathcal{T}$ in the function ℓ_n for convenience. To prove Theorem 5.1, we first establish the strong convexity of the loss ℓ_n with strong convexity parameter $\eta > 0$. Let $\widehat{\Delta} = \widehat{\theta} - \theta^*$ denote the error in our parameter estimate with respect to θ^*, and recall that the goal of Theorem 5.1 is to provide a finite-sample upper bound on $\|\widehat{\Delta}\|_2$. The strong convexity of ℓ_n implies that

$$\frac{\eta}{2}\|\widehat{\Delta}\|_2^2 \leq \ell_n(\theta^* + \widehat{\Delta}) - \ell_n(\theta^*) - \langle \nabla \ell_n(\theta^*), \widehat{\Delta}\rangle, \tag{11}$$

where η is known as the strong convexity parameter. Since $\widehat{\theta}$ is the minimizer of the ℓ_n, we have $\ell_n(\widehat{\theta}) - \ell_n(\theta^*) = \ell_n(\theta^* + \widehat{\Delta}) - \ell_n(\theta^*) \leq 0$. Together with (11) and using Hölder's inequality, this implies that

$$\frac{\eta}{2}\|\widehat{\Delta}\|_2^2 \leq -\langle \nabla \ell_n(\theta^*), \widehat{\Delta}\rangle \leq \|\nabla \ell_n(\theta^*)\|_\infty \|\widehat{\Delta}\|_1 \leq \sqrt{d}\|\nabla \ell_n(\theta^*)\|_\infty \|\widehat{\Delta}\|_2,$$

which further implies that

$$\|\widehat{\Delta}\|_2 \leq \frac{2\sqrt{d}}{\eta}\|\nabla \ell_n(\theta^*)\|_\infty. \tag{12}$$

Intuitively, (12) 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 (12), we only need to (1) establish an upper bound on $\|\nabla \ell_n(\theta^*)\|_\infty$; (2) identify the strong-convexity parameter η. These steps are accomplished in the following lemmas, with proofs given in the online supplement. We begin by showing that $\|\nabla \ell_n(\theta^*)\|_\infty$ can be upper bounded with high probability in both deterministic and random design cases.

Lemma 5.2 (Gradient Bound)**.** *In the deterministic design setting under Assumptions 1 and 2, we have with probability at least $1 - \frac{1}{n}$, $\|\nabla \ell_n(\theta^*)\|_\infty \leq cB\sqrt{\frac{\log(nd)}{n}}$. In the randomized design setting under Assumptions 1 and 3, we have with probability at least $1 - \frac{1}{n}$, $\|\nabla \ell_n(\theta^*)\|_\infty \leq c\psi\sqrt{\frac{\log(nd)}{n}}$.*

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

Lemma 5.3 (Strong Convexity)**.** *In the deterministic design setting under Assumptions 1 and 2, we have that ℓ_n is strongly convex at the true parameter θ^* with*

$$\eta = \frac{\exp(RB)}{4(1 + \exp(RB))^2} \cdot \rho. \tag{13}$$

In the randomized design setting under Assumptions 1 and 3, as long as $n \geq \frac{4C_{cp}(\psi)\log(n)d}{\min(\rho,1)^2}$ for some constant $C_{cp}(\psi)$ only depending on ψ, ℓ_n is strongly convex at θ^* with strong convexity parameter given in (13), with probability at least $1 - 2(\frac{1}{n})^d$.

Proof of Theorem 5.1. By plugging both the upper bound on $\|\nabla \ell_n(\theta^*)\|_\infty$ in Lemma 5.2 and the strong convexity parameter η in (13) into (12), we obtain the result of Theorem 5.1 in both the fixed and the random design settings, which completes the proof of Theorem 5.1. □

5.2.2 Revenue Bound

Theorem 5.1 gives us a finite-sample (non-asymptotic) estimation bound that holds with high probability, i.e. the theorem tells us that the Algorithm 1 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. Recall the estimated best price $\hat{p}$ as defined in (10) and the oracle price,

$$\widehat{p} := \max_{p \in \mathcal{A}} f_z(p, \widehat{\theta}) \quad \text{and} \quad p^* := \max_{p \in \mathcal{A}} f_z(p, \theta^*).$$

We are interested in the revenue gap between the revenue generated by the oracle price p^* (so-called oracle revenue) and that generated by the offered price $\widehat{p}$ via Algorithm 1 when the customer's behavior is specified by the true parameters θ^*, i.e., $f_z(p^*, \theta^*) - f_z(\widehat{p}, \theta^*)$. 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*, since such a bound holds for any new customer with feature vector z.

Theorem 5.4 (Revenue Convergence Rate)**.** *Under Assumptions 1, 2 and 3, we have that with high probability, as long as $n \geq \frac{4C_{cp}(\psi)\log(n)d}{\min(\rho,1)^2}$, for any feature vector z, the expected revenue gap can be bounded by:*

$$f_z(p^*, \theta^*) - f_z(\widehat{p}, \theta^*) \leq \left(\max_{p \in \mathcal{A}} p\right) \frac{C'_{cp}(R, B, \psi)}{\rho} \sqrt{\frac{d^2 \log(nd)}{n}},$$

where $C'_{cp}(R, B, \psi)$ is a constant only depending on R, B and ψ.

To translate the parameter estimation error in Theorem 5.1 to the revenue gap, the key is to prove that $f_z(p, \theta)$ for any given z and p is Lipschitz continuous in θ with Lipschitz constant $(\max_{p \in \mathcal{A}} p)\frac{\sqrt{mB^2+1}}{4}$, where m is the dimension of z and B is given in Assumption 1. The details are in the online supplement.

As desired, we have bounded the optimality gap of the reward generated by the recommended price $\widehat{p}$ as compared to the oracle price p^*. This bound decreases as $O\left(\frac{1}{\sqrt{n}}\right)$, up to logarithmic terms, and thus quantifies the trade-off between the value of data and potential lost revenue.

5.3 Theoretical Results for Personalized Assortment Optimization

In this section, we provide revenue gap bounds for personalized assortment optimization problems, which depend on a MNL model rather than a binomial model of customer choice and thus feature a richer decision set. We state our assumptions and results, and the relevant proofs are provided in the online supplement.

In personalized assortment optimization, since there are J products (in contrast to the single-product setting of customized pricing), we must also assume that the seller has *sufficiently explored* each product. This assumption is natural since it is impossible to estimate a customer's reaction to a product if this product has never been offered in previous assortments. Formally, given the historical transaction data $\mathcal{T} = \{(z_i, S_i, y_i)\}_{i=1}^n$, where each observed outcome $y_i \in \{0, 1, \ldots, J\}$, let $\mathcal{I}_j = \{i : j \in S_i\}$, $j \in [J]$, be the indices of transactions that offer product j in their assortment. Thus, the number of transactions including product j is $n_j = |\mathcal{I}_j|$.

Assumption 4. *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.*

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 is 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. We will see this more clearly in the theoretical results in Theorem 5.5 and 5.6.

The following three assumptions exactly parallel those of the customized pricing model.

Assumption 5. *For both deterministic and random design settings, we assume*

1. *Conditional independence: the observed outcomes $\{y_i\}_{i=1}^n$ are independent given each z_i.*

2. *Bounded feature vector: there exists a universal constant $B > 0$ such that for any custom feature z, we have $|z_j| \leq B$ for all $j \in [m]$.*

Assumption 6 (Deterministic Design)**.** *There exists a constant ρ such that $\lambda_{\min}(\Sigma_n) \geq \frac{\rho}{2} > 0$, where $\Sigma_n = \frac{1}{n}\sum_{i=1}^n z_i z_i^T$ is the sample Gram matrix.*

Assumption 7 (Random Design)**.** 1. *The vectors $\{z_i\}_{i=1}^n$ are independent and identically distributed sub-Gaussian random vectors with sub-Gaussian norm ψ_z (see Section 5.1).*

2. *There exists a universal constant ρ such that $\lambda_{\min}(\Sigma) > \rho > 0$, where $\Sigma = \mathbb{E}(zz^T)$.*

Using these assumptions and a proof similar to that used in the case of customized pricing, we derive the following results:

Theorem 5.5. *Under Assumptions 4, 5, 6, and 7, we have the following: for deterministic design, with probability at least $1 - J/n$,*

$$\|\widehat{\theta}_j - \theta_j^*\|_2 \leq c\frac{B}{\rho}\frac{[1+\exp(RB)]^2}{\exp(-RB)}J^2\sqrt{\frac{d\log(nd)}{n\nu}}$$

simultaneously for all products $j = 1, \ldots, J$.

For the randomized design, as long as $n \geq \frac{4C_{ao}(\psi_z)\log(n)d}{\nu\min(\rho,1)^2}$ for some constant $C_{ao}(\psi_z)$ only depending on ψ_z, with probability at least $1 - J/n - 2J(1/n)^d$,

$$\|\widehat{\theta}_j - \theta_j^*\|_2 \leq c\frac{\psi_z}{\rho}\frac{[1+\exp(RB)]^2}{\exp(-RB)}J^2\sqrt{\frac{d\log(nd)}{n\nu}}$$

simultaneously for all products $j = 1, \ldots, J$.

Theorem 5.6. *Under Assumptions 4, 5, 6, and 7 and with high probability, the optimality gap of the algorithm as specified above can be bounded as follows:*

$$f_z(S^*, \theta^*) - f_z(\widehat{S}, \theta^*) \leq \frac{2C'_{ao}(R, B, \psi_z)}{\rho} r_1 J^4 \sqrt{\frac{d^2\log(nd)}{n\nu}},$$

provided that $n \geq \frac{4C_{ao}(\psi_z)\log(n)d}{\nu\min(\rho,1)^2}$.

Note that the form of this bound is very similar to that of the bound in Theorem 5.4. In particular, it decreases at a rate of $O\left(\frac{1}{\sqrt{n}}\right)$ with the number of samples n and holds for any (even previously unobserved) vector z. The added factor of J^4 reflects the fact that there are up to J alternatives in every transaction and also $J \times d$ parameters to be learned in the model. Although this factor weakens the bound in comparison with the pricing case, J is constant in practice and does not affect asymptotic performance.

The other main difference in Theorem 5.6 as compared to Theorem 5.4 is the appearance of the constant ν, which is inversely proportional to the revenue gap. This reflects intuition: for small ν, at least one of the products has been offered only rarely in previous transactions, increasing the difficulty of estimating parameters. Thus, we would expect overall prediction accuracy to improve and expected revenue gap to decrease for a higher value of ν (i.e., each product has been offered almost the same number of times).

6 Extensions of Theory

The theory presented so far has assumed that our model was well-specified and that the feature vectors were low-dimensional. We now demonstrate that it is possible for both of these assumptions to be relaxed. For the ease of presentation, we only focus on customized pricing in this section but all the results can also be shown for personalized assortment optimization.

6.1 Misspecified Model Setting

In practice it is quite possible that consumer behavior cannot be specified using a logit model, in which case the underlying demand structure is typically unknown. If we still proceed to use the logit loss function ℓ_n, then a natural benchmark is the performance of the logit model that most closely reflects actual consumer behavior, which we term the *oracle estimator*. The oracle estimator is obtained by minimizing $\mathbb{E}(\ell_n(\theta))$, where the expectation is taken with respect to the true underlying model. This corresponds to the maximum likelihood estimated parameters of a logit model in the limit as the sample size grows to infinity. As we only have a finite number of samples in practice, we are interested in the revenue gap between the action chosen based on the estimated parameters from our finite sample and that based on the oracle estimator.

In order for the notion of an infinitely large training set sampled from an underlying distribution to be well-defined, we consider random rather than fixed design. Thus, we suppose that the data $\mathcal{T} = \{(z_1, a_1, y_1), \ldots (z_n, a_n, y_n)\}$ is generated from some underlying random process. In accordance with our discussion above, we redefine

$$\theta^* = \arg\min_{\theta} \mathbb{E}(\ell_n(\theta)). \tag{14}$$

Thus in this section, the oracle estimator θ^* represents the estimated parameters under the *expected* negative log-likelihood (or the negative log-likelihood function based on an infinite number of samples from the underlying true distribution). This is in contrast to the interpretation of θ^* in Section 5 as the true underlying parameters, which are unavailable here because we do not make any modeling assumptions. We also note that since we consider the expectation of ℓ_n here, there is no need for a regularization constraint on θ in (14). This is because the main purpose of regularization is to penalize the model complexity when the number of samples is limited.

Under this setup, we can still prove meaningful revenue bounds. In fact, the same results from Theorems 5.1, 5.4, 5.5, and 5.6 still hold with the new definition of θ^*. As an example, we will prove here an analogous result to Theorem 5.4, the customized pricing revenue bound:

Theorem 6.1. *In the misspecified setting with θ^* defined as in (14) and under Assumptions 1 and 3, we have that with high probability, as long as $n \geq \frac{4C_{cp}(\psi)\log(n)d}{\min(\rho,1)^2}$, the gap between oracle estimator revenue and the revenue achieved with Algorithm 1 can be bounded as follows:*

$$f_z(p^*, \theta^*) - f_z(\hat{p}, \theta^*) \leq \left(\max_{p\in\mathcal{A}} p\right) \frac{C'_{cp}(R, B, \psi)}{\rho} \sqrt{\frac{d^2 \log(nd)}{n}},$$

for all feature vectors z.

The proof of Theorem 6.1 is rather straightforward since the form of the likelihood function ℓ_n has not changed from the well-specified to the misspecified setting. In addition, Assumptions 1 and 3 also remain unchanged and so Lemma 5.3 still holds. However, the form of the underlying randomness in the data and the meaning of θ^* has changed, so Lemma 5.2 can not be directly applied here. Examining the proof of Lemma 5.2, one can see that the proof (in particular, the concentration result) still holds so long as θ^* satisfies

$$\mathbb{E}\nabla\ell_n(\theta^*) = 0. \tag{15}$$

In the misspecified setting, because the oracle estimator θ^* is defined as an unconstrained minimum, we have $\nabla\mathbb{E}(\ell_n(\theta^*)) = 0$. Under certain mild regularity conditions on ℓ_n that hold for all our applications, the differentiation and integration can be interchanged, which means that (15) holds. Given that both lemmas hold, we have that Theorem 5.1 must hold for the oracle estimator, which further implies that the revenue gap bound in Theorem 5.4 holds since the revenue function f_z remains the same.

In the cases of both Theorem 5.4 and Theorem 6.1, we use an estimation error bound to quantify the rate at which the revenue gap shrinks to zero as the number of samples grows — a rate of $\frac{1}{\sqrt{n}}$. The difference between the two settings is in the interpretation of the oracle benchmark. In Theorem 5.4, the oracle policy knows the parameters of the true underlying distribution, whereas in Theorem 6.1 the oracle only knows the parameters of the logit model which best reflects customer choice.

6.2 High-Dimensional Setting

Here we will return to the well-specified setting and explore another extension. In previous sections, our bounds are increasing functions of the number of features d. This is reasonable when d remains fixed and n grows large. However, as companies continue to collect more granular data concerning their customers, there are applications where the number of customer features matches or even exceeds the number of data points. Examples could include highly granular GPS information, text information with words as features (e.g., using customers' online posts on Twitter), or detailed lists of item-to-item preferences that are not easily specified without introducing a combinatorial number of parameters.

In such a high-dimensional setting where d is comparable to or larger than n, a bound of the form presented above is of no use. Since standard statistical estimation methods such as regression are underdetermined in a high-dimensional setting, we are forced to make some assumptions on the structure of the data in order to derive performance bounds. One common assumption is the sparsity of the true parameters [Bühlmann and van de Geer, 2011].

Assumption 8 (Sparsity)**.** *We assume that there exists some index set $\mathcal{M} \subset \{1, \ldots, d\}$ with $|\mathcal{M}| = s \ll d$ such that $\theta^* \in \Theta(\mathcal{M}) := \{\theta \in \mathbb{R}^d : \theta_j = 0,\ \forall j \notin \mathcal{M}\}$.*

It is easy to see from Assumption 8 that if $j \notin \mathcal{M}$, then $\theta^*_j = 0$, which means that the j-th feature has no effect on the outcome. This implies that only a small portion of the available features are relevant for the prediction of customer purchase behavior. Such an assumption is often reasonable in practice. For example, if we take a customer's online posts (e.g., Facebook, Twitter) as contextual information with word counts as features (the so-called bag-of-words feature model), it is likely that most of the words are irrelevant to purchase behavior. In practice, the set of relevant features $\mathcal{M}$ is unknown to the decision-maker. Therefore, we would like to determine the subset of features that are most determinant of consumer purchase behavior algorithmically. A natural way to obtain an estimator of a sparse vector is to use ℓ_1-norm regularization as presented by Tibshirani [1996], which is well known to force degenerate parameters to zero. We point the reader to Bühlmann and van de Geer [2011] for a review of this and other techniques to deal with high-dimensional data.

The sparsity assumption allows us to prove theoretical revenue bounds analogous to those in the low dimensional setting. Using a Lagrangian form of (9), we obtain an estimator given by

$$\widehat{\theta}_{\lambda_n} \in \arg\min_{\theta\in\mathbb{R}^d} \{\ell_n(\mathcal{T},\theta) + \lambda_n\|\theta\|_1\} . \tag{16}$$

By standard duality theory (see Chapter 6 in Bertsekas et al. [2003] for example), (9) and (16) are equivalent in the sense that there is a one-to-one correspondence between λ_n and R. The change in form here is made only for convenience in proving the high-dimensional results. Although all of our results can be extended to the high-dimensional regime, the proofs become highly technical. For this reason we include here only the high-dimensional version of our revenue bound for the well-specified, random design, customized pricing problem.

Theorem 6.2. *Under Assumptions 1, 3, and 8, as long as $n = \Omega(s\log d)$, the objective function error associated with $(\widehat{\beta},\widehat{\gamma})$, the solution to (16) with regularization parameter $\lambda_n = c(\psi)\sqrt{\frac{\log p}{2n}}$ for suitable constant $c(\psi)$, can be bounded as follows:*

$$f_z(p^*,\theta^*) - f_z(\widehat{p},\theta^*) \leq \left(\max_{p\in\mathcal{A}} p\right)\frac{B\ C_{HD}(\psi)}{4\rho}s\sqrt{\frac{\log d}{n}},$$

for all bounded feature vectors z with probability at least $1-\frac{c}{n}$. Here c is a positive constant depending only on B, ψ, ρ, and Σ.

Notice that although we cannot apply Theorem 5.4 in the high-dimensional setting since d grows faster than n (and thus the bounds on the estimation error and revenue loss go to infinity as n grows), Theorem 6.2 does apply. In particular, the dimension parameter d only appears logarithmically. Since s remains small as d grows by the sparsity assumption, the bound shrinks to zero even when d is exponential in n^c with $c < 0.5$ (e.g., $\log(d) = o(\sqrt{n})$). We also note that in practice, we cannot directly rely on this guarantee because the regularization parameter λ_n is dependent on the unknown quantity ψ. However, one can tune λ_n by cross validation to achieve good performance.

7 Experiments

The theory we have presented so far suggests that our method provides an effective technique for estimating and optimizing data-driven decisions. In this section, we show here that our algorithm performs quite well in practice on both simulated and real data.

7.1 Simulated Experiments

We performed simulations for both customized pricing and personalized assortment optimization. For customized pricing we defined a problem class by specifying a number of prices $K \in \{2,4,10\}$ and a number of features $m \in \{5,10,15\}$, 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 $p^K \leq p^{K-1} \leq \cdots \leq p^1$. We also generated a d-dimensional true parameter vector γ^*, where each dimension was chosen i.i.d. from a normal distribution with mean zero and standard deviation 1.5, and a K-dimensional parameter vector β^* as the order statistics of K i.i.d. normal random variables with mean zero and standard deviation 3 sorting from lowest to highest. Here, the first term β_1 is incorporated into the intercept effect as in Section 3.1. 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 m, 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 Algorithm (PRM) in Algorithm 1,
2. The I-PRM Algorithm (PRM with the isotonic constraint on β as in Remark 4.1),
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, $\mathbb{R}^m$) 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 cross-validation in terms of the leaf size parameter.

We generated a test set of 1000 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 which knows the true parameters. This policy's performance was then used as a normalization to compare the four methods of interest.

Figure 1 shows the performance of all four methods under four representative problem classes, averaged across the 100 trials. The isotonic-constrained version of PRM (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 Algorithm 1 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 $\mathbb{P}_z(y = 1; p, \beta^*, \gamma^*, \xi^*) = \left(1 + \exp\left(-\left(\sum_{k=2}^{K} \beta_k^* \mathbb{I}(p = p^k) + \sum_{j=1}^{m} \gamma_j^* z_j + \sum_{j=1}^{m} \xi_j^* (z_j)^2\right)\right)\right)^{-1}$. 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

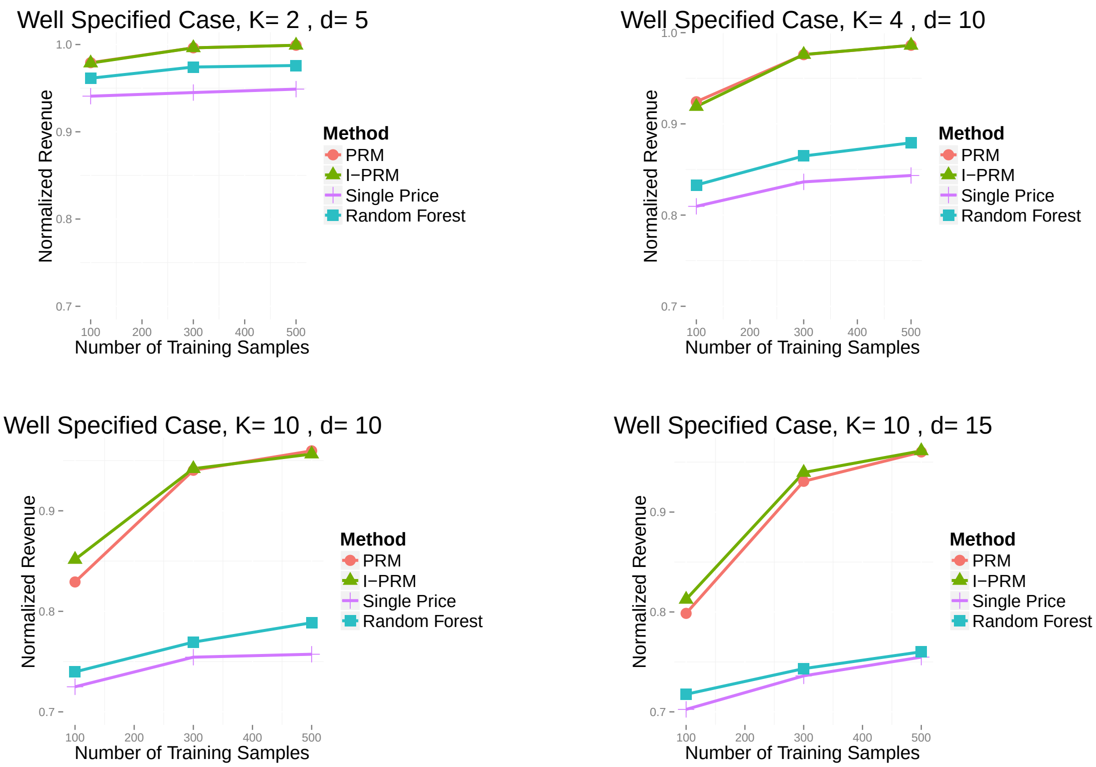

Figure 1: Performance of PRM, I-PRM, single price, and random forest algorithms in the well-specified setting for various problem classes

of PRM and I-PRM which knew the form of the underlying model. In other words, PRM and I-PRM learned only $\widehat{\beta}$ and $\widehat{\gamma}$ parameters, while 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 1000 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 to the oracle as given in (14), which was simulated by training a logistic regression model on $10,000$ training data points.

Figure 2 shows the results of these methods over $n = \{100, 300, 500, 1000\}$. 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 that 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 to 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 Forest method also struggled, and was out-performed by the PRM and I-PRM methods.

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

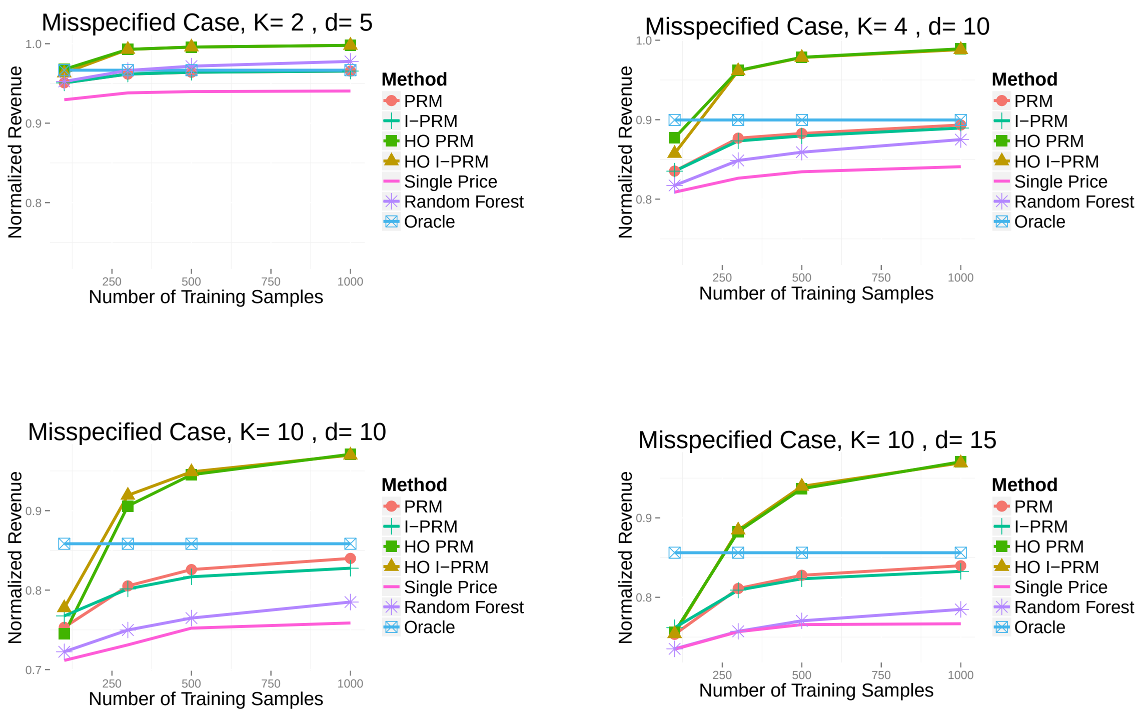

Figure 2: Performance of methods in the misspecified setting for various problem classes

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 3. 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 Personalized Revenue Maximization Algorithm (PRM)
2. A "Mean-Effect" Revenue Maximization Algorithm (MERM) that does not use any feature information, described below.

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,

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

where j_i indicates the item purchased for the i-th transaction. MERM then used these mean estimates $\widehat{V}^{\text{MERM}}$ in a 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 [2004a], the PRM, featureless, and full-knowledge policies all chose from a small subset of possible assortments, namely 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 of problem classes. As expected, PRM consistently and significantly outperforms the featureless approach, especially as the number of products grows large.

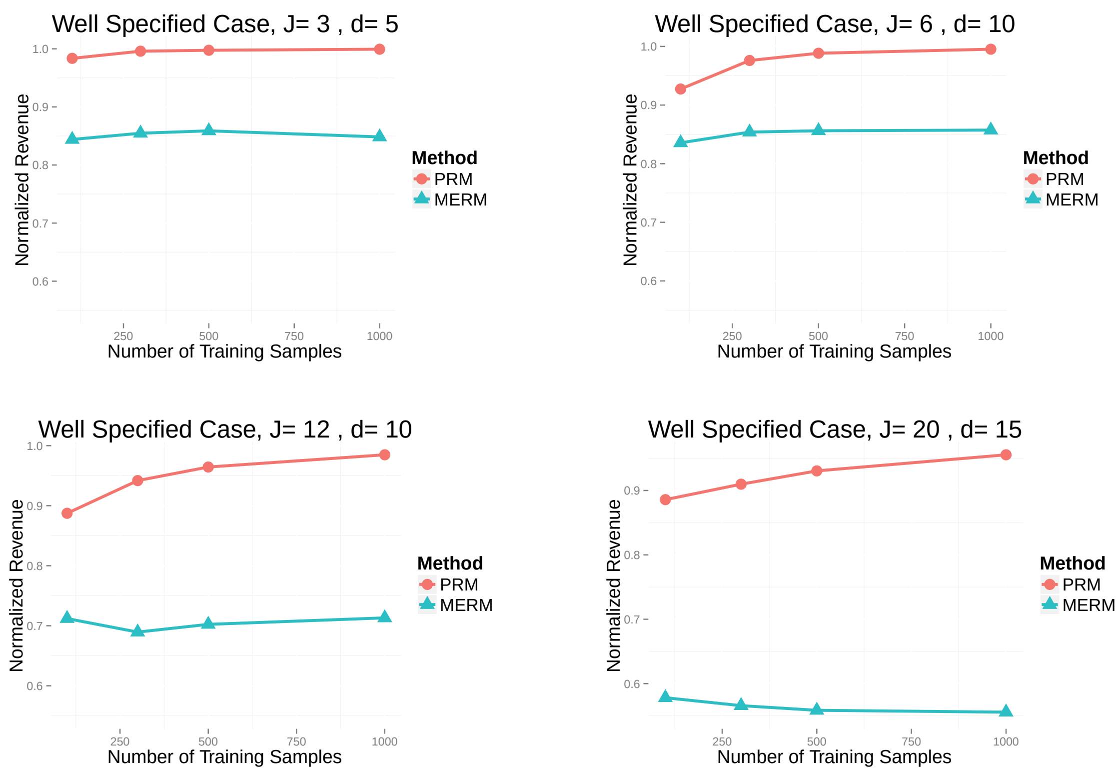

Figure 3: Performance of methods in the well-specified assortment optimization setting for various problem classes

7.2 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 of this carrier may pre-select a seat on their airplane that will be reserved for them on the day of their flight. Over a one month period, some 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 transaction records with the treatment price offered, data concerning attributes of the flight, data concerning attributes of the transaction, and the resulting purchase decision. More details about the data set including a list of features and an example record are presented in the online supplement.

The performance metric of interest in this case is the expected revenue from reserved seating per passenger. To allow us to fairly evaluate the success of our method, we began by splitting our data into a training and testing set chronologically, with the earliest 60% of transaction records used for training and the latest 40% used for testing. This ensures that any method selected based on the training set could have been implemented by the company during the subsequent testing period. Due to the relatively limited number of purchases, the non-purchase records were down-sampled in the training set. This down-sampling prevents machine learning techniques from over-generalizing due to the prevelance of non-purchases. Our primary benchmark is the performance of the best single price that could have been selected from the training data on the testing set. Since treatment prices were offered independently at random, we can estimate the single price performance unbiasedly, by restricting our attention to customers who were offered each specific price during the testing period. The highest single-price expected revenue per passenger on the training set is achieved using the highest price, and we will report the performance of our method normalized by the performance of this price.

Since each customer was only offered a single price, we are unable to perform a true counterfactual analysis to evaluate algorithm performance when the chosen price to offer a customer is different from the price in the data set. However, as treatment prices were offered to customers independently at random, we are able to estimate expected revenue per passenger by restricting our attention to the subset of records for which the best price selected by the algorithm matches the price offered in the data set. Although this limits the size of the test set, it allows us to obtain an unbiased estimate of expected revenue per passenger without the need for potentially erroneous counterfactual assumptions.

We compared the performance of our method, PRM, with that of two other widely used statistical techniques, regression trees and random forest, as the number of samples in the training set is increased. The random forest technique was described in Section 7.1. Regression trees are a standard recursive technique in statistical learning which in each iteration splits the data set into two pieces and computes an estimate in each resulting compartment. The recursive procedure terminates when all possible splits of the training data are unable to increase in-sample R^2 by a variable prescribed amount (see Chapter 9 of Hastie et al. [2009] for more details). In our case, for each offered price, we train a regression tree which attempts to learn purchase probability by splitting on the feature data.

To select our PRM model, in accordance with Algorithm 1, we fit a logit model to predict purchase probability using the price offered as a feature as well as the flight and transaction data discussed previously. We also included interaction effects, as discussed in Remark 3.1, between the offered price and other variables in the data set. To train both the regression tree and the random forest, we split the data into four sets corresponding to the offered prices and trained separate models for each as described in Section 7.1. In the case of PRM and regression trees, features were selected using a forward model selection procedure with 4-fold cross-validation, adding the feature to each model which yielded the highest incremental expected revenue per passenger on the held-out data sets. To avoid local optima common in such forward selection methods we began the feature selection process from a number of fixed initial features. This procedure was too

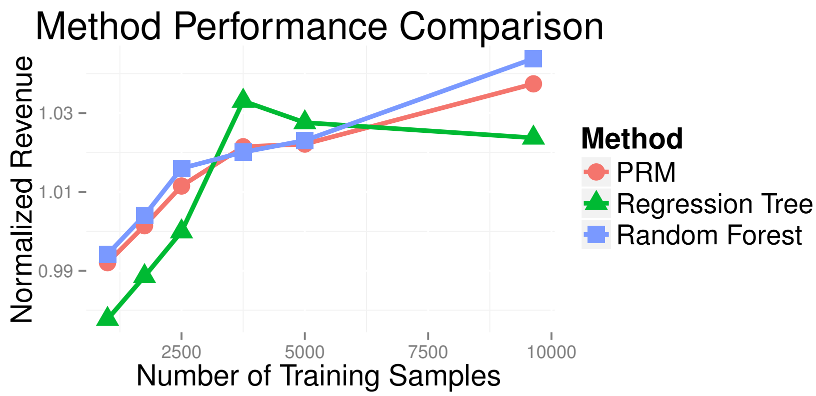

Figure 4: Performance of PRM, regression trees, and random forest on the testing set as the size of training set is increased. Performance is averaged over 45 randomly generated subsets of each size in the case of PRM and regression trees. In the case of random forest, the same 45 such sets were generated and the results for each are taken as the average of 5 separate random forests.

computationally intensive to be practical for random forests so we instead performed model selection using minimum node size for such models. For all strategies, in each fold of the cross-validation, we trained the models as described and cross-validation revenue was computed using the resulting pricing recommendations on the holdout set. Finally, to assess final algorithm performance, each model with its set of selected features was trained on the entire training set and the resulting model was applied to the testing set. This procedure was used separately for each training set size $n \in \{1000, 1750, 2500, 3750, 5000, 9635\}$ reflecting approximately 10%, 17.5%, 25%, 37.5%, 50%, and 100% of the available training data respectively.

Figure 4 displays the performance of each method on the test set normalized by the single-price revenue. We observe that although the PRM method performs competitively with both methods for smaller sample sizes and as the data size grows larger it continues to learn consistently achieving a 3.5% increase in revenue over the best single price. As sophisticated machine learning technique, random forest performs comparatively well on all training set sizes. This is due to its ability to construct non-linear functions of feature data including applicable interaction effects without the need for such functions to be explicitly pre-computed and provided. However, random forest models suffer from a lack of interpretability and it is often difficult to discern the important interaction effects, for example the inputs which affect price sensitivity, under such models.

8 Extensions and Future Work

We have developed a framework for modeling decision problems in which actions can be personalized by taking into account the information available to the decision maker. If we assume a logit model to describe outcome probabilities in terms of the features and the seller's decision, we demonstrate that learning takes place reliably by establishing finite-sample high probability convergence guarantees for model parameters which hold regardless of the number of customer types, which can be potentially uncountable. These bounds apply between the fitted model and the minimizer of expected loss whether or not the model has been well-specified. The parameter convergence guarantees can then be extended to performance bounds in operational problems as we show for the case of customized pricing and personalized assortment optimization. We have also shown how these bounds can be extended for the case in which a logit model does not specify underlying behavior and for the case of high-dimensional data.

Our results are presented here for single-period problems. However, some recent works have focused on multiple period settings in which the seller must balance learning demand with maximizing revenue. Our framework can be extended to a multi-stage problem to capture this exploiting-exploration tradeoff. For example, one simple strategy is as follows. In the first stage, referred to as an exploration stage, the seller conducts randomized price experimentation to learn the parameters of the demand function. Then in the exploitation stage, the seller chooses for each customer the action that maximizes personalized revenue based on the current estimate of the demand parameters and then updates the estimates after observing the customer's choice. Of course, one might design more complicated exploitation-exploration protocols which integrate exploitation and exploration phases together with a shrinking interval length (see Wang et al. [2014] for an example). An alternative approach, proposed by Johnson et al. [2015] incorporates a Thompson sampling-based approached to balance the trade-off between exploration and exploitation in an online setting with inventory constraints, and they extended their ideas to the personalized setting with contextual information.

Our framework gives a method for formulating personalized versions of classic operations models, and so it is natural to ask what insights from the classical models carry over to the personalized setting and what new insights we can gain. For example, the results of Talluri and van Ryzin [2004a] concerning revenue-ordered assortments have been successfully carried over to the personalized assortment optimization setting, in which the optimal assortment for each customer is also revenue-ordered. In the case of multi-product pricing it is known that under certain assumptions, *cost-plus pricing*, a pricing policy in which a single markup is added to the cost of all products, can be shown to be optimal [Li and Huh, 2011]. Interestingly, a *personalized* cost-plus policy, in which each individual is given their own markup, is optimal under similar assumptions for the personalized version of the problem. Another important class of classical models considers inventory constraints, which often strongly influence the structure of optimal policies. It will be important future work to explore generalizations of our approach to this capacitated setting.

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.

Acknowledgement

We would like to thank Ilan Lobel for the discussion of personalized assortment optimization and Jason Lee for the discussion of bounding the parameter estimation error.

References

X. Amatriain. Big & personal: data and models behind netflix recommendations. In *Proceedings of the 2nd International Workshop on Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications*, pages 1–6. ACM, 2013.

Y. Aviv, G. Vulcano, Ö. Özer, and R. Phillips. Dynamic list pricing. *The Oxford Handbook of Pricing Management. Oxford University Press, Oxford*, pages 522–584, 2012.

G. Aydin and S. Ziya. Technical note-personalized dynamic pricing of limited inventories. *Operations Research*, 57(6):1523–1531, 2009.

Goker Aydin and Serhan Ziya. Pricing promotional products under upselling. *Manufacturing & Service Operations Management*, 10(3):360–376, 2008.

P. Belobaba. Application of a probabilistic decision model to airline seat inventory control. *Operations Research*, 37:183–197, 1989.

F. Bernstein, A. G. Kök, and L. Xie. Dynamic assortment customization with limited inventories. Working Paper, 2011.

D. Bertsekas, A. Nedić, and A. Ozdaglar. *Convex Analysis and Optimization*. Athena Scientific, 2003.

D. Bertsimas and N. Kallus. From predictive to prescriptive analytics. *arXiv preprint arXiv:1402.5481*, 2014.

D. Bertsimas and P. Vayanos. Data-driven learning in dynamic pricing using adaptive optimization. Submitted to Operations Research, 2014.

O. Bousquet, S. Boucheron, and G. Lugosi. Introduction to statistical learning theory. In *Advanced Lectures on Machine Learning*. Springer, 2004.

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

J. Broder and P. Rusmevichientong. Dynamic pricing under a general parametric choice model. *Operations Research*, 60(4):965–980, 2012.

P. Bühlmann and S. van de Geer. *Statistics for High-Dimensional Data: Methods, Theory, and Applications*. Springer-Verlag, 2011.

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

Wang Chi Cheung, David Simchi-Levi, and He Wang. Dynamic pricing and demand learning with limited price experimentation. *Available at SSRN 2457296*, 2014.

C. Du, W. Cooper, and Z. Wang. Optimal pricing for a multinomial logit choice model with network effects. Working Paper, 2015.

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

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

I. Gurvich, M. Armony, and C. Maglaras. Cross-selling in a call center with a heterogeneous customer population. *Operations research*, 57(2):299–313, 2009.

T. Hastie and R. Tibshirani. *Generalized Additive Models*. Chapman and Hall, 1990.

Trevor Hastie, Robert Tibshirani, Jerome Friedman, T Hastie, J Friedman, and R Tibshirani. *The elements of statistical learning*, volume 2. Springer, 2009.

R. Horn and C. Johnson. *Matrix Analysis*. Cambridge University Press, 2nd edition, 2012.

Kris Johnson, David Simchi-Levi, and He Wang. Online network revenue management using thompson sampling. *Available at SSRN*, 2015.

N. B. Keskin and A. Zeevi. Dynamic pricing with an unknown demand model: Asymptotically optimal semi-myopic policies. *Operations Research*, 62(5):1142–1167, 2014.

A.G. Kök, M. Fisher, and R. Vaidyanathan. *Retail Supply Chain Management*, chapter Assortment Planning: Review of Literature and Industry Practice, pages 1–55. 2008.

G. Li, P. Rusmevichientong, and H. Topaloglu. The d-level nested logit model: Assortment and price optimization problems. *Operations Research (to appear)*, 2015.

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

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

D. McFadden and K. Train. Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15(5):447–470, 2000.

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

S. Negahban, P. Ravikumar, M. Wainwright, and B. Yu. A unified framework for high-dimensional analysis of m-estimators with decomposable regularizers. *arXiv pre-print*, 2010. arXiv:1010.2731v1.

S. Negahban, P. Ravikumar, M. Wainwright, and B. Yu. A unified framework for high-dimensional analysis of m-estimators with decomposable regularizers. *Statistical Science*, 27(4):538–557, 2012.

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

R. Phillips. *Pricing and Revenue Optimization*. Stanford Business Books, 2005.

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

C. Rudin and G. Vahn. The big data newsvendor: Practical insights from machine learning. Working Paper, 2014.

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

P. Rusmevichientong, D. Shmoys, C. Tong, and H. Topaloglu. Assortment optimization under the multinomial logit model with random choice parameters. *Production and Operations Management*, 23(11):2023–2039, 2014.

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

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

K. Talluri and G. van Ryzin. *The Theory and Practice of Revenue Management*. Kluwer Academic Publishers, 2004b.

R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.

K. Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, 2nd edition, 2009.

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

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

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

R. Vershynin. *Compressed Sensing, Theory and Application*, chapter 5, pages 210–268. Cambridge University Press, 2012.

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

Z. Wang, S. Deng, and Y. Ye. Close the gaps: A learning-while-doing algorithm for single-product revenue management problems. *Operations Research*, 62(2):318–331, 2014.

A One Explanation on Assumption 3

For Assumption 3, it is easy to show that when customer feature z is a sub-Gaussian vector with sub-Gaussian norm ψ_z, the feature vector x is also sub-Gaussian with sub-Gaussian norm $\psi = \psi_z + 1$. To see this, recall that $x = (v, z)$ where v is the vector of dummy variables $v = \big(\mathbb{I}(p_i = p^2), \ldots, \mathbb{I}(p_i = p^K)\big)$. Then, by the definition of sub-Gaussian norm of x in Section 5.1, we have

$$\begin{aligned}\psi &= \sup_{\|(w_v,w_z)\|_2\le 1} \|\langle v, w_v\rangle + \langle z, w_z\rangle\|_{\psi_2}\\ &\le \sup_{\|w_v\|\le 1}\|\langle v, w_v\rangle\|_{\psi_2} + \sup_{\|w_z\|_2\le 1}\|\langle z, w_z\rangle\|_{\psi_2}\\ &\le \psi_z + 1.\end{aligned}$$

B Proofs of Lemmas Used to Prove Theorem 5.1

We prove here the gradient norm bound and strong convexity lemmas, both restated below.

Lemma B.1 (Lemma 5.2)**.** *In the deterministic design setting under Assumptions 1 and 2, we have with probability at least* $1 - \frac{1}{n}$, $\|\nabla \ell_n(\theta^*)\|_\infty \le cB\sqrt{\frac{\log(nd)}{n}}$*. In the randomized design setting under Assumptions 1 and 3, we have with probability at least* $1 - \frac{1}{n}$, $\|\nabla \ell_n(\theta^*)\|_\infty \le c\psi\sqrt{\frac{\log(nd)}{n}}$.

Proof of Lemma 5.2. We note the j-th component of $\nabla \ell_n(\theta^*)$ takes the form $[\nabla \ell_n(\theta^*)]_j = \frac{1}{n}\sum_{i=1}^n W_{ij}$, where $W_{ij} = \left(\frac{e^{\langle x_i,\theta^*\rangle}}{1+e^{\langle x_i,\theta^*\rangle}} - y_i\right) x_{ij}$. Conditioned on x_i, W_{ij} is a zero-mean bounded random variable with $|W_{ij}| \le |x_{ij}| \le B$. For the fixed design setting, applying Hoeffding's inequality, we have $\Pr\left(|[\nabla \ell_n(\theta^*)]_j| \ge t\right) \le 2\exp\left(-\frac{nt^2}{2B^2}\right)$. By a union bound, we have

$$\Pr\left(\|\nabla \ell_n(\theta^*)\|_\infty \ge t\right) \le 2\exp\left(\log(d) - \frac{nt^2}{2B^2}\right). \tag{17}$$

By setting $t = B\sqrt{\frac{2\log(2nd)}{n}}$, we make the R.H.S. of (17) equal to $\frac{1}{n}$, which gives the desired result.

In the randomized design setting, W_{ij} is a centered sub-Gaussian random variable with the norm bounded above by ψ. To see this, observe that

$$\mathbb{E}\exp(tW_{ij}) \le \mathbb{E}\exp(|tW_{ij}|) \le \mathbb{E}\exp(|tx_{ij}|) \le \mathbb{E}\exp(|t|\langle \operatorname{sign}(x_{ij})e_j, x_i\rangle) \le \exp\left(ct^2\psi^2\right)$$

for some constant c. Applying Hoeffding's inequality this implies that, $\Pr\left(|[\nabla \ell_n(\theta^*)]_j| \ge t\right) \le 2\exp\left(-\frac{c_1 nt^2}{\psi^2}\right)$. By a union bound, we then have that

$$\Pr\left(\|\nabla \ell_n(\theta^*)\|_\infty \ge t\right) \le 2\exp\left(\log(d) - \frac{c_1 nt^2}{\psi^2}\right). \tag{18}$$

By setting $t = \psi\sqrt{\frac{\log(2nd)}{c_1 n}}$, we make the R.H.S. of (18) equal to $1 - \frac{1}{n}$, which gives the desired result. □

Lemma B.2 (Lemma 5.3)**.** *In the deterministic design setting under Assumptions 1 and 2, we have that ℓ_n is strongly convex at the true parameter θ^* with*

$$\eta = \frac{\exp(RB)}{4(1+\exp(RB))^2} \cdot \rho. \tag{19}$$

In the randomized design setting under Assumptions 1 and 3, as long as $n \geq \frac{4C_{cp}(\psi)\log(n)d}{\min(\rho,1)^2}$ for some constant $C_{cp}(\psi)$ only depending on ψ, ℓ_n is strongly convex at θ^* with strong convexity parameter given in (19), with probability at least $1 - 2(\frac{1}{n})^d$.

Proof. The Taylor expansion of ℓ_n implies that for some $\alpha \in (0,1)$ we have

$$\ell_n(\theta^* + \widehat{\Delta}) - \ell_n(\theta^*) - \langle \nabla \ell_n(\theta^*), \widehat{\Delta} \rangle = \frac{1}{2n} \sum_{i=1}^n \frac{\exp(x_i^T\theta^* + \alpha x_i^T \widehat{\Delta})}{(1+\exp(x_i^T\theta^* + \alpha x_i^T\widehat{\Delta}))^2} \widehat{\Delta}^T \left(x_i x_i^T\right) \widehat{\Delta}. \tag{20}$$

Further,

$$|x_i^T\theta^* + \alpha x_i^T\widehat{\Delta}| \leq \|x_i\|_\infty \|\theta^* + \alpha\widehat{\Delta}\|_1 \leq \|x_i\|_\infty \left(\alpha\|\theta^*\|_1 + (1-\alpha)\|\widehat{\theta}\|_1\right) \leq RB.$$

Since the function $\frac{\exp(a)}{(1+\exp(a))^2} \leq \frac{1}{4}$ is an even function and monotonically decreasing as $|a|$ increases, we have

$$\frac{\exp(x_i^T\theta^* + \alpha x_i^T \widehat{\Delta})}{(1+\exp(x_i^T\theta^* + \alpha x_i^T\widehat{\Delta}))^2} \geq \frac{\exp(RB)}{(1+\exp(RB))^2}, \tag{21}$$

which further implies that,

$$\frac{1}{2n} \sum_{i=1}^n \frac{\exp(x_i^T\theta^* + \alpha x_i^T \widehat{\Delta})}{(1+\exp(x_i^T\theta^* + \alpha x_i^T\widehat{\Delta}))^2} \widehat{\Delta}^T \left(x_i x_i^T\right) \widehat{\Delta} \geq \frac{1}{2}\frac{\exp(RB)}{(1+\exp(RB))^2} \widehat{\Delta}^T \left(\frac{1}{n}\sum_{i=1}^n x_i x_i^T\right) \widehat{\Delta}$$
$$\geq \frac{1}{2}\frac{\exp(RB)}{(1+\exp(RB))^2} \lambda_{\min}(\Sigma_n) \|\widehat{\Delta}\|_2^2.$$

In the deterministic setting, by the assumption that $\lambda_{\min}(\Sigma_n) \geq \frac{\rho}{2}$, we obtain the desired result in (19).

In the randomized setting, by Corollary 5.50 from [Vershynin, 2012], for any $\epsilon \in (0,1)$ and $t \geq 1$, when $n \geq C_{cp}(\psi)\left(\frac{t}{\epsilon}\right)^2 d$ (where $C_{cp}(\psi)$ is a constant which depends only on ψ), we have that with probability at least $1 - 2\exp(-t^2 d)$, $\|\Sigma_n - \Sigma\|_{\text{op}} \leq \epsilon$. By setting $\epsilon = \frac{1}{2}\min(\rho, 1)$ and $t = \sqrt{\log(n)}$, we have with probability at least $1 - 2(\frac{1}{n})^d$, $\|\Sigma_n - \Sigma\|_{\text{op}} \leq \frac{1}{2}\rho$, provided that $n \geq \frac{4C_{cp}(\psi)\log(n)d}{\min(\rho,1)^2}$.

By Weyl's theorem (see Theorem 4.3.1. in Horn and Johnson [2012], for example), we have $|\lambda_{\min}(\Sigma_n) - \lambda_{\min}(\Sigma)| \leq \|\Sigma_n - \Sigma\|_{\text{op}} \leq \frac{1}{2}\rho$, which further implies that $\lambda_{\min}(\Sigma_n) \geq \lambda_{\min}(\Sigma) - \frac{1}{2}\rho \geq \frac{1}{2}\rho$. This completes the proof. □

C Proof of Theorem 5.4

Theorem C.1 (Theorem 5.4)**.** *Under Assumptions 1, 2 and 3, we have that with high probability, as long as $n \geq \frac{4C_{cp}(\psi_z)\log(n)d}{\min(\rho,1)^2}$, for any feature vector z, the expected revenue gap can be bounded by:*

$$f_z(p^*, \theta^*) - f_z(\widehat{p}, \theta^*) \leq \left(\max_{p \in \mathcal{A}} p\right) \frac{C'_{cp}(R, B, \psi_z)}{\rho} \sqrt{\frac{d^2 \log(nd)}{n}},$$

where $C'_{cp}(R, B, \psi_z)$ is a constant only depending on R, B and ψ_z.

Proof. In order to translate the parameter bound into a bound on revenue, we exploit some of the structural properties of f_z as defined in (4). Here, we view f_z as a function of its parameters $\theta = (\beta, \gamma)$. For any given price $p \in \mathcal{A}$ and feature vector z, we have

$$\begin{aligned} \left| f_z(p, \theta^*) - f_z(p, \widehat{\theta}) \right| &= \left| \frac{p}{1 + \exp\left(-\left(\beta_k^* + \sum_j \gamma_j^* z_j\right)\right)} - \frac{p}{1 + \exp\left(-\left(\widehat{\beta}_k + \sum_j \widehat{\gamma}_j z_j\right)\right)} \right| \\ &\leq \frac{p}{4} \left| \langle (1, z), (\beta_k^* - \widehat{\beta}_k, \gamma^* - \widehat{\gamma}) \rangle \right| \qquad (22) \\ &\leq \frac{p}{4} \|(1, z)\|_2 \left\| (\beta_k^* - \widehat{\beta}_k, \gamma^* - \widehat{\gamma}) \right\|_2 \\ &\leq \left(\max_{p \in \mathcal{A}} p \right) \frac{\sqrt{mB^2 + 1}}{4} \left\| \theta^* - \widehat{\theta} \right\|_2, \end{aligned}$$

where k is the index of the corresponding price in $\mathcal{A}$ and we define β_k^* and $\widehat{\beta}_k$ to be equal to zero when $k = 1$ and to be equal to the corresponding element of β^* or $\widehat{\beta}$ respectively otherwise. We note that (22) follows from the fact that the derivative of the function $(1 + \exp(-a))^{-1}$ is bounded by $\frac{1}{4}$ for any a. Thus, using the fact that $f_z(\widehat{p}, \widehat{\theta}) \geq f_z(p^*, \widehat{\theta})$,

$$\begin{aligned} f_z(p^*, \theta^*) - f_z(\widehat{p}, \theta^*) &= f_z(p^*, \theta^*) - f_z(\widehat{p}, \widehat{\theta}) + f_z(\widehat{p}, \widehat{\theta}) - f_z(\widehat{p}, \theta^*) \\ &\leq f_z(p^*, \theta^*) - f_z(p^*, \widehat{\theta}) + f_z(\widehat{p}, \widehat{\theta}) - f_z(\widehat{p}, \theta^*) \\ &\leq \left| f_z(p^*, \theta^*) - f_z(p^*, \widehat{\theta}) \right| + \left| f_z(\widehat{p}, \widehat{\theta}) - f_z(\widehat{p}, \theta^*) \right| \\ &\leq \left(\max_{p \in \mathcal{A}} p \right) \frac{\sqrt{mB^2 + 1}}{2} \left\| \theta^* - \widehat{\theta} \right\|_2. \qquad (23) \end{aligned}$$

Having bounded the revenue gap by the parameter gap, we can apply Theorem 5.1 to get the result in the theorem statement. □

D Proofs of Assortment Optimization Results

Here we prove the assortment results from Section 5.3. The proofs parallel those in the customized pricing case.

D.1 Proof of Theorem 5.5

Theorem D.1 (Theorem 5.5)**.** *Under Assumptions 4, 5, 6, and 7, we have the following: for deterministic design, with probability at least $1 - J/n$,*

$$\|\widehat{\theta}_j - \theta_j^*\|_2 \leq c \frac{B}{\rho} \frac{[1 + \exp(RB)]^2}{\exp(-RB)} J^2 \sqrt{\frac{d \log(nd)}{n\nu}}$$

simultaneously for all products $j = 1, \ldots, J$.

For the randomized design, as long as $n \geq \frac{4C_{ao}(\psi_z)\log(n)d}{\nu \min(\rho,1)^2}$ *for some constant* $C_{ao}(\psi_z)$ *only depending on* ψ_z*, with probability at least* $1 - J/n - 2J(1/n)^d$*,*

$$\|\widehat{\theta}_j - \theta_j^*\|_2 \leq c\frac{\psi_z}{\rho}\frac{[1+\exp(RB)]^2}{\exp(-RB)}J^2\sqrt{\frac{d\log(nd)}{n\nu}}$$

simultaneously for all products $j = 1, \ldots, J$*.*

Viewing $\ell_n(\theta)$ as a function only of θ_j, with all other parameters vectors θ_ℓ fixed to $\widehat{\theta}_\ell$, we establish its strong convexity with parameter η. Letting $\widehat{\Delta}_j = \widehat{\theta}_j - \theta_j^*$, we have

$$\frac{\eta}{2}\left\|\widehat{\Delta}_j\right\|_2^2 \leq \ell_n(\theta_j^* + \widehat{\Delta}) - \ell_n(\theta_j^*) - \langle \nabla_j \ell_n(\theta^*), \widehat{\Delta}_j\rangle$$

We can use the fact that $\ell_n(\theta_j^* + \widehat{\Delta}_j) - \ell_n(\theta_j^*) \leq 0$ to arrive at

$$\frac{\eta}{2}\left\|\widehat{\Delta}_j\right\|_2^2 \leq -\langle \nabla_j \ell_n(\theta_j^*), \widehat{\Delta}_j\rangle \leq \left\|\nabla_j \ell_n(\theta_j^*)\right\|_\infty \left\|\widehat{\Delta}_j\right\|_1 \leq \sqrt{d}\left\|\nabla_j \ell_n(\theta_j^*)\right\|_\infty \left\|\widehat{\Delta}_j\right\|_2.$$

This further implies that

$$\left\|\widehat{\Delta}_j\right\|_2 \leq \frac{2\sqrt{d}}{\eta}\left\|\nabla_j \ell_n(\theta_j^*)\right\|_\infty.$$

Lemma D.2. *Under the assumptions of Theorem 5.5, for the fixed design, with probability* $1 - \frac{1}{n}$*,*

$$\left\|\nabla_j \ell_n(\theta_j^*)\right\|_\infty \leq cB\sqrt{\frac{\log(nd)}{n\nu}}.$$

For the random design, with probability $1 - \frac{1}{n}$*,*

$$\left\|\nabla_j \ell_n(\theta_j^*)\right\|_\infty \leq c\psi_z\sqrt{\frac{\log(nd)}{n\nu}}.$$

Proof. We note the k-th component of $\nabla \ell_n(\theta_j^*)$ takes the following form,

$$[\nabla \ell_n(\theta_j^*)]_k = \frac{1}{n_j}\sum_{i \in \mathcal{I}_j} W_{ijk}, \tag{24}$$

where $W_{ijk} = \left(\frac{\exp(\langle z_i, \theta_j^*\rangle)}{1+\sum_{\ell \in S_i}\exp(\langle z_i, \theta_\ell^*\rangle)} - \mathbb{I}[j_i = j]\right) z_{ik}$. Conditioned on z_i, W_{ijk} is a zero-mean bounded random variable with $|W_{ij}| \leq |z_{ik}| \leq B$. For the fixed design setting, applying Hoeffding's inequality,

$$\begin{aligned}\Pr\left(|[\nabla \ell_n(\theta_j^*)]_k| \geq t\right) &\leq 2\exp\left(-\frac{n_j t^2}{2B^2}\right)\\ &\leq 2\exp\left(-\frac{n\nu t^2}{2B^2}\right).\end{aligned}$$

By a union bound, we have

$$\Pr\left(\|\nabla \ell_n(\theta_j^*)\|_\infty \geq t\right) \leq 2\exp\left(\log(d) - \frac{n\nu t^2}{2B^2}\right). \tag{25}$$

By setting $t = B\sqrt{\frac{2\log(2nd)}{n\nu}}$, we make the R.H.S. of (25) equal to $1/n$, which gives the result.

For randomized design, W_{ijk} is a centered sub-Gaussian random variable with the norm bounded above by ψ_z. To see this,

$$\mathbb{E}\exp\left(tW_{ijk}\right) \leq \mathbb{E}\exp\left(|tW_{ijk}|\right) \leq \mathbb{E}\exp\left(|tz_{ik}|\right) \leq \mathbb{E}\exp\left(|t|\left(\mathrm{sign}(z_{ik})e_k\right)^T z_i\right) \leq \exp\left(Ct^2\psi_z^2\right).$$

Applying Hoeffding's inequality,

$$\Pr\left(|[\nabla \ell_n(\theta_j^*)]_k| \geq t\right) \leq 2\exp\left(-\frac{c_1 n_j t^2}{\psi_z^2}\right) \leq 2\exp\left(-\frac{c_1 n\nu t^2}{\psi_z^2}\right). \tag{26}$$

By a union bound, we have

$$\Pr\left(\|\nabla \ell_n(\theta_j^*)\|_\infty \geq t\right) \leq 2\exp\left(\log(d) - \frac{c_1 n\nu t^2}{\psi_z^2}\right). \tag{27}$$

By setting $t = \psi_z\sqrt{\frac{\log(2nd)}{c_1 n\nu}}$, we make the R.H.S. of (27) to be $1/n$, which gives the result.

□

Thus, in order to prove Theorem 5.5, it suffices to show the following lemma:

Lemma D.3. *Under the assumptions of Theorem 5.5, for deterministic design, we have strong convexity parameter*

$$\eta = \frac{\exp(-RB)}{4[1+\exp(-RB)+(J-1)\exp(RB)]^2}\rho.$$

For randomized design, as long as $n \geq \frac{4C_{\psi_z}\log(n)d}{\nu\min(\rho,1)^2}$ *for some constant* C_{ψ_z} *only depending on* ψ_z*, with probability at least* $1-2(1/n)^d$,

$$\eta = \frac{\exp(-RB)}{4[1+\exp(-RB)+(J-1)\exp(RB)]^2}\rho.$$

Proof. Fixing $j \in [J]$, we view ℓ_n as a function of θ_j, with all other parameters vectors θ_ℓ fixed to $\widehat{\theta}_\ell$. We have by Taylor expansion

$$\begin{aligned}&\ell_n(\theta_j^* + \widehat{\Delta}_j) - \ell_n(\theta_j^*) - \langle \nabla_j \ell_n(\theta_j^*), \widehat{\Delta}_j\rangle \\ &\qquad = \frac{1}{2n_j}\sum_{i\in\mathcal{I}_j}\frac{\exp(-\langle\theta_j^* + \alpha\widehat{\Delta}_j, z_i\rangle)}{[1+\exp(-\langle\theta_j^*+\alpha\widehat{\Delta}_j, z_i\rangle)+\sum_{\substack{\ell\in S_i\\ \ell\neq j}}\exp(\langle\widehat{\theta}_\ell, z_i\rangle)]^2}\widehat{\Delta}_j^T(z_i z_i^T)\widehat{\Delta}_j.\end{aligned} \tag{28}$$

Using the fact that the function $\frac{\exp(-x)}{1+\exp(-x)+c_1}$ dominates the function $\frac{\exp(-x)}{1+\exp(-x)+c_2}$ whenever $c_1 \leq c_2$ and a similar argument as the one directly preceding (21), we have

$$\frac{\exp(-\langle \theta_j^* + \alpha\widehat{\Delta}_j, z_i\rangle)}{[1+\exp(-\langle \theta_j^* + \alpha\widehat{\Delta}_j, z_i\rangle) + \sum_{\substack{\ell\in S_i\\ \ell\neq j}} \exp(\langle \widehat{\theta}_\ell, z_i\rangle)]^2} \geq \frac{\exp(-\langle \theta_j^* + \alpha\widehat{\Delta}_j, z_i\rangle)}{[1+\exp(-\langle \theta_j^* + \alpha\widehat{\Delta}_j, z_i\rangle) + (J-1)\exp(RB)]^2}$$

$$\geq \frac{\exp(-RB)}{[1+\exp(-RB)+(J-1)\exp(RB)]^2}. \tag{29}$$

We apply (29) to the right-hand side of (28) to get

$$\frac{1}{2n_j}\sum_{i\in\mathcal{I}_j} \frac{\exp(-\langle \theta_j^* + \alpha\widehat{\Delta}_j, z_i\rangle)}{[1+\exp(-\langle \theta_j^* + \alpha\widehat{\Delta}_j, z_i\rangle) + \sum_{\substack{\ell\in S_i\\ \ell\neq j}} \exp(\langle \widehat{\theta}_\ell, z_i\rangle)]^2}\widehat{\Delta}_j^T(z_iz_i^T)\widehat{\Delta}_j$$

$$\geq \frac{1}{2}\frac{\exp(-RB)}{[1+\exp(-RB)+(J-1)\exp(RB)]^2}\widehat{\Delta}_j^T(\frac{1}{n_j}\sum_{i\in\mathcal{I}_j} z_iz_i^T)\widehat{\Delta}_j$$

$$\geq \frac{\exp(-RB)}{2[1+\exp(-RB)+(J-1)\exp(RB)]^2}\lambda_{\min}(\Sigma_{n_j})\left\|\widehat{\Delta}_j\right\|_2^2.$$

In the deterministic inputs setting, we can apply the assumption that $\lambda_{\min}(\Sigma_{n_j}) \leq \frac{\rho}{2}$ to arrive at the result.

For the randomized setting, by Corollary 5.50 from [Vershynin, 2012], for any $\epsilon \in (0,1)$ and $t \geq 1$, when $n_j \geq C_{\psi_z}\left(\frac{t}{\epsilon}\right)^2 d$, where C_{ψ_z} is a constant depends only on ψ_z, with probability at least $1-2\exp(-t^2 d)$,

$$\|\Sigma_{n_j} - \Sigma\|_{\text{op}} \leq \epsilon.$$

By setting $\epsilon = \frac{1}{2}\min(\rho, 1)$ and $t = \sqrt{\log(n)}$, we have with probability at least $1-2(1/n)^d$,

$$\|\Sigma_{n_j} - \Sigma\|_{\text{op}} \leq \frac{1}{2}\rho,$$

provided that $n_j \geq \frac{4C_{\psi_z}\log(n)d}{\min(\rho,1)^2}$. By Weyl's theorem, we have

$$|\lambda_{\min}(\Sigma_{n_j}) - \lambda_{\min}(\Sigma)| \leq \|\Sigma_{n_j} - \Sigma\|_{\text{op}} \leq \frac{1}{2}\rho,$$

which further implies that,

$$\lambda_{\min}(\Sigma_{n_j}) \geq \lambda_{\min}(\Sigma) - \frac{1}{2}\rho \geq \frac{1}{2}\rho.$$

This completes the proof. □

D.2 Proof of Theorem 5.6

Theorem D.4 (Theorem 5.6)**.** *Under Assumptions 4, 5, 6, and 7 and with high probability, the optimality gap of the algorithm as specified above can be bounded as follows:*

$$f_z(S^*, \theta^*) - f_z(\widehat{S}, \theta^*) \leq \frac{2C'_{ao}(R, B, \psi_z)}{\rho} r_1 J^4 \sqrt{\frac{d^2\log(nd)}{n\nu}},$$

provided that $n \geq \frac{4C_{ao}(\psi_z)\log(n)d}{\nu\min(\rho,1)^2}$.

The proof is a simple consequence of the following proposition.

Proposition D.5. *For any offered assortment $S \subseteq \{1, \dots, J\}$, with high probability, as long as $n \geq \frac{4C_{ao}(\psi_z)\log(n)d}{\nu \min(\rho,1)^2}$, the error in our revenue forecast can be bounded as follows:*

$$\left| f_z(S, \theta^*) - f_z(S, \hat{\theta}) \right| \leq \frac{C'_{ao}(R, B, \psi_z)}{\rho} r_1 J^4 \sqrt{\frac{d^2 \log(nd)}{n\nu}}$$

for all bounded feature vectors z where $C'(R, B, \psi_z)$ is a constant depending only on R, B, and ψ_z.

Proof. Fix a customer feature vector $z \in \mathcal{Z}$ and a subset of products $S \subseteq \{1, \dots, J\}$. For a fixed subset of products the difference in expected revenue under the true model defined by θ^* and our estimated model specifed 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 $\frac{\delta}{\delta\theta_{jk}}\pi_z(j, S, \theta) \leq \frac{1}{4}\|z\|_\infty \leq \frac{1}{4}B$ for all $k \in S$ and for any value of θ and $z \in \mathcal{Z}$. Therefore we have the global bound on the gradient of $\mathbb{P}_z(j, S; \theta)$ with respect to θ of $\|\nabla \mathbb{P}_z(j, S, \theta)\|_\infty \leq \frac{1}{4}B$. Using this we can proceed to bound the forecast error as claimed using this Lipschitz constant:

$$\begin{aligned}
|f_z(S, \theta^*) - f_z(S, \hat{\theta})| &\leq \sum_{j \in S} r_j |\mathbb{P}_z(j, S; \theta^*) - \mathbb{P}_z(j, S; \hat{\theta})| \\
&\leq \sum_{j \in S} r_j \|\nabla \mathbb{P}_z(j, S, \theta)\|_\infty \|\theta^* - \widehat{\theta}\|_1 \\
&\leq \frac{B}{4} \sum_{j \in S} r_j \|\theta^* - \widehat{\theta}\|_1 \\
&\leq \frac{B\sqrt{Jd}}{4} \sum_{j \in S} r_j \|\theta^* - \widehat{\theta}\|_2 \\
&\leq \frac{J^2 r_1 B \sqrt{d}}{4} \max_{j=1,\dots,J} \|\theta^*_j - \widehat{\theta}_j\|_2.
\end{aligned}$$

Substituting in the result of Theorem 5.5 yields the desired result with high probability. ☐

To finish the proof of Theorem 5.6, we have

$$\begin{aligned}
f_z(S^*, \theta^*) - f_z(\widehat{S}_z, \theta^*) &= \left(f_z(S^*_z, \theta^*) - f_z(\widehat{S}_z, \widehat{\theta}) \right) + \left(f_z(\widehat{S}_z, \widehat{\theta}) - f_z(\widehat{S}_z, \theta^*) \right) \\
&\leq \left(f_z(S^*, \theta^*) - f_z(S^*, \widehat{\theta}) \right) + \left(f_z(\widehat{S}_z, \widehat{\theta}) - f_z(\widehat{S}_z, \theta^*) \right) \\
&\leq \frac{2C'(R, B, \psi_z)}{\rho} r_1 J^4 \sqrt{\frac{d^2 \log(nd)}{n\nu}},
\end{aligned}$$

where in the final step we have applied the result of Proposition D.5 twice.

E Proof of Theorem 6.2

Theorem 6.2 is the high-dimensional bound, restated here.

Theorem E.1 (Theorem 6.2)**.** *Under Assumptions 1, 3, and 8, as long as* $n = \Omega(s \log d)$*, the objective function error associated with* $(\widehat{\beta}, \widehat{\gamma})$*, the solution to (16) with regularization parameter* $\lambda_n = c(\psi)\sqrt{\frac{\log p}{2n}}$ *for suitable constant* $c(\psi)$*, can be bounded as follows:*

$$f_z(p^*, \theta^*) - f_z(\widehat{p}, \theta^*) \le \left(\max_{p \in \mathcal{A}} p\right) \frac{B\ C_{HD}(\psi)}{4\rho} s\sqrt{\frac{\log d}{n}},$$

for all bounded feature vectors z *with probability at least* $1 - \frac{c}{n}$*. Here* c *is a positive constant depending only on* B, ψ, ρ*, and* Σ*.*

Proof. Recall that for the high-dimensional proofs, we are concerned with the following optimization:

$$\widehat{\theta}_{\lambda_n} \in \arg\min_{\theta \in \mathbb{R}^d} \left\{ \ell_n(\mathcal{T}, \theta) + \lambda_n \|\theta\|_1 \right\}, \tag{30}$$

where for customized pricing $\widehat{\theta}_{\lambda_n} = (\widehat{\beta}_{\lambda_n}, \widehat{\gamma}_{\lambda_n})$. As before, our revenue guarantee depends on a high-probability bound for the distance between the estimated and optimal parameters. For the high dimensional case, such a bound is provided in Negahban et al. [2012] for general M-estimators, and the details for generalized linear models (which include logistic regression) are given in Negahban et al. [2012] and Negahban et al. [2010]. From their results (specifically Theorem 1 and Corollary 1 from Negahban et al. [2012]) we obtain

$$\|\widehat{\theta}_{\lambda_n} - \theta^*\|_1 \le \frac{C_{HD}(\psi)}{\rho} s\sqrt{\frac{\log(d)}{n}}, \tag{31}$$

for some constant $C_{HD}(\psi)$ depending on ψ but not on n, d, or s.

From the argument in Theorem 5.4, we have

$$\begin{aligned} f_z(\theta^*) - f_z(\widehat{\theta}_{\lambda_n}) &\le \frac{p^{k^*}}{4} \langle (1, z), ((\widehat{\beta}_{\lambda_n})_{k^*} - \beta^*_{k^*}, \widehat{\gamma}_{\lambda_n} - \gamma^*) \rangle \\ &\le \left(\max_{p \in \mathcal{A}} p\right) \frac{B}{4} \left\| \widehat{\theta}_{\lambda_n} - \theta^* \right\|_1. \end{aligned}$$

By substituting in the bound (31) we obtain the statement of the theorem, bounding objective function error with high probability. □

F Customized Pricing for Airline Priority Seating Details

Here we provide further details on the airline seating reservation data set and the resulting model derived in the style of Algorithm 1. Because of the potential legal issues with offering prices to customers based on their personal attributes, the data consists solely of information concerning the flight such as date or destination and transaction information such as date and time of website access. The data points we considered are presented in Table 1.

Due to issues often encountered when working with real transaction data we were forced to make a number of assumptions and simplifications to the data set in order to conduct our analysis. First, in the data collection process customers were actually offered two levels of product: one for reserved

Date Created	*Booking DW*	*Flight Date*	*Flight Hour*	*Flight DW*
12/01/2014 12:16	Monday	12/08/2014	8	Monday

Days until Flight	*Return Flight Date*	*Return Flight Hour*	*Return Flight DW*
5	12/11/2014	15	Thursday

Trip Length	*Qty. Sold to Date*	*Qty. Remaining*	*Departure City*	*Arrival City*
3	17	38	City1	City2

Dest. Type	*# Passengers*	*Avg. Fare*	*Offered Price Level*	*Purchase Decision*
Vacation	3	100	3	No Purchase

Table 1: A fabricated example record from our seating reservation transaction dataset.

seating and an alternative premium option that includes a reservation and additional benefits. These products were each offered at four different price levels but the price level of both products shown to each customer we aligned so that a customer seeing the lowest price for the standard reservation also saw the lowest price for the premium option. To extract the most information from the data set we treated both types of purchases as purchases of the lower level product at the lower level price. In effect, this modification makes the assumption that the seating reservation of the premium option can be decoupled from the other benefits and that purchasers of the premium option value the seating reservation at its marginal price. Second, each record in the data set contains the number of passengers in this particular flight booking. Some seating reservations were for large quantities and to avoid these records having an outsize effect in our testing scheme we considered each correct price offering to be worth the revenue of a single reservation.