

Bundle Pricing via Learning the Market from Customer Preferences

Ezgi Karabulut^{*1} and Burak Gökgür²

¹Faculty of Engineering and Natural Sciences, Sabancı University, Istanbul,
TURKEY

²Sabancı Business School, Sabancı University, Istanbul, TURKEY

Abstract

We study the problem of learning revenue-maximizing bundle prices when customer valuations and market composition are unknown and the firm observes only customer choices. This problem is challenging because the number of bundle-price decisions grows exponentially with the number of items offered. We develop BLMP (Bundle Pricing via Learning the Market from Customer Preferences), a learning-and-optimization approach that transforms observed bundle choices into item-level revealed-preference information, sequentially estimates valuation characteristics and segment fractions, and then solves a mixed-integer bundle-pricing model using the estimated parameters. The framework accommodates both deterministic and bounded probabilistic valuations and reduces the dimensionality of the valuation learning problem from $O(2^n)$ to $O(n)$. We benchmark BLMP against CMA-ES and GP-UCB across different problem dimensions and valuation distributions. BLMP outperforms CMA-ES and remains competitive with GP-UCB while avoiding the latter's computational difficulties as the number of observations increases. BLMP produces higher-quality final pricing solutions and moves closer to the full-information benchmark, whereas CMA-ES shows little improvement. The results reveal a trade-off in allocating customer interactions across BLMP's two stages: more interactions in the first stage generally improve final pricing quality, whereas allocating relatively more interactions to the second stage increases revenue collected during learning. These findings show how exploiting the structure embedded in customer choices can make bundle-price learning more tractable and shed light on the allocation of a finite budget of customer interactions.

Keywords— Mixed-Bundle Pricing, Preference Learning, Optimization, Revenue Management

^{*}Corresponding Author: ezgi.turkseven@sabanciuniv.edu

1 Introduction

Mixed bundling, the practice of offering two or more different products at a single discount in addition to their stand-alone buying opportunities (Schmalensee, 1984), is prevalent among retail firms that aim to enlarge their customer base and to enhance their bottom line. Such practices arise across a wide range of industries, from telecommunications and digital services to grocery and consumer packaged goods. In grocery retailing, for example, Tesco sells eligible mains, snacks, and drinks individually while simultaneously allowing customers to combine a selected main, snack, and drink at a fixed meal-deal price (Tesco, 2026). In storable consumer goods, for instance, Target offers a 20-oz Heinz Squeeze Tomato Ketchup and 16-oz/10-ct Oscar Mayer Original Wieners Hot Dogs jointly for \$6.63, compared with \$6.98 when the products are purchased separately (Target, 2026); analogous discounted bundle examples can be found across household-case and personal-care categories. These examples illustrate the economic rationale of mixed bundling: by keeping both stand-alone purchase and bundle purchase options, a retailer can accommodate heterogeneous purchasing needs while encouraging purchases that may not arise under stand-alone pricing alone.

Relying on this opportunity requires information that retailers generally cannot observe directly. A customer’s willingness to pay (or valuation) for each product is private, and customers are heterogeneous in the valuations they attach to products; willingness to pay is inherently distributional across customers rather than represented by a single market-wide value (He et al., 2024). Consequently, the firm observes customers’ responses to the purchase options and prices it offers rather than their underlying valuations. This informational limitation has motivated a growing literature on demand and preference learning, which shows how firms can use observations generated through market interactions to improve their knowledge of otherwise unknown demand and customer preferences (Huang et al., 2018). For a firm employing mixed bundling, the relevant unknowns may extend beyond aggregate demand to the underlying structure of the market, i.e., the product-level valuations comprising heterogeneous customer segments and the relative sizes of those segments.

Pricing interactions can provide a potential means of acquiring such information; prices can serve not only as revenue-generating lever but also instruments for learning an unknown market. For example, Cheung et al. (2017) study price experimentation as a demand-learning tool when the seller initially faces an unknown demand function and report a field implementation at Groupon with improvements on both revenue and bookings. Misra et al. (2019) similarly study online retailers that set prices for large number of products with incomplete demand information and use price experiments, relying on an extended version of multiarmed bandit algorithms, to learn product-level demand and profitable prices. In a large online-retailing field study, Fisher et al. (2018) use randomized prices to obtain unbiased estimates of price elasticities and consumer choice behavior; their live pricing experiment increases revenue by 11%. In a mixed-bundling environment, pricing interactions can be informative because customers choose among multiple stand-alone and bundle options. Such choices demonstrate comparisons among the

available alternatives and can contain information about the heterogeneous product valuations underlying the observed purchase preferences. The resulting problem is not simply to optimize stand-alone and bundle prices for a known market, but to determine how mixed-bundle pricing interactions can be used to learn an initially unknown heterogeneous market and support revenue enhancement.

The first stream of the literature related to this study is mixed-bundle pricing. Seminal papers establish bundling as a mechanism for exploiting heterogeneity in consumer valuations (Adams and Yellen, 1976; Schmalensee, 1984; McAfee et al., 1989). Subsequent papers examine the design and optimization of bundle pricing under various market structures and consumer preferences (Hanson and Martin, 1990; Stremersch and Tellis, 2002; Venkatesh and Kamakura, 2003). The scope of mixed-bundle pricing literature has been expanded with models that consider additional features such as consumer subscription (Cai et al., 2026), supply chain coordination (Gökgür and Kocabıykoğlu, 2026), consumer regret (G. Zhang et al., 2023), and personalized bundling (Ettl et al., 2020). A common feature of much of this literature is that the information describing the market structure (i.e., consumer valuation distributions and segment characteristics) is available to the firm. A notable exception is Fang et al. (2025), who study bundling when the proportions of consumers with low and high valuations are initially unknown in a single-segment market environment. They develop a Thompson-sampling-based learning mechanism through which the firm updates its beliefs about these proportions and dynamically chooses its selling strategy. Under their binary valuation structure, however, the mixed-bundling strategies considered are shown to be equivalent to corresponding pure-bundling strategies, reducing the firm’s dynamic decision essentially to a choice between pure component selling and pure bundling. Thus, while Fang et al. (2025) introduce valuation uncertainty and learning into the bundling problem, the role of mixed bundling as a pricing mechanism in learning and exploiting an unknown heterogeneous market remains unexplored.

A second stream studies pricing when the underlying demand environment is initially unknown. A central issue in this stream of the literature is the learning and earning trade-off: pricing decisions bring about immediate revenue while also producing information that can improve future decisions. Early studies develop dynamic pricing policies for learning unknown demand characteristics and examine the implications of learning and price experimentation (Şen and A. X. Zhang, 2009; Harrison et al., 2012; Keskin and Zeevi, 2014). This literature has been extended to several directions. Wang (2021), for example, characterizes optimal Bayesian demand-learning policies over finite selling horizons. Xiao and W. Yang (2024) develop a Bayesian framework for simultaneously learning multiple correlated parameters of an unknown demand function. Other studies incorporate demand learning into different operational and behavioral environments, including price experimentation (Cheung et al., 2017), reference-price effects (Cao et al., 2019; Boer and Keskin, 2022), strategic customer behavior (X. Chen et al., 2022), and joint capacity and pricing decisions (J. Chen et al., 2026). This literature demonstrates how pricing decisions can be adapted as information regarding an initially unknown demand environment accumulates.

Within this broader learning literature, another group of studies approaches pricing through multi-armed bandit and online learning formulations. Tehrani et al. (2012) exploit dependence among prices induced by a common underlying demand model. On the other hand, Misra et al. (2019) use revealed-preference restrictions to extract information about demand at prices other than the one currently offered. Bandit-based approaches have also been extended to multi-product revenue management environments; Ferreira et al. (2018), for instance, employ Thompson sampling to learn unknown demand parameters while managing prices and inventory. Recently, nonparametric approaches have been developed for multi-product pricing without requiring a known parametric demand function. X. Yang et al. (2024) use simultaneous price perturbations to learn and optimize a multi-product revenue function.

Bayesian optimization provides another model-free approach when the relationship between pricing decisions and their outcomes can only be evaluated through sequential observations. Gaussian-process-based methods construct a probabilistic surrogate for the unknown function and use predictions and uncertainty from this surrogate to guide subsequent experimentation (Shahriari et al., 2016). Recent applications employ Bayesian optimization for dynamic pricing under unknown demand (Anand et al., 2025), while Jain et al. (2023) exploit known compositional structure within Bayesian optimization and demonstrate its application to multi-product dynamic pricing. Related Gaussian-process approaches combined with bandit policies demonstrate the potential of nonparametric learning for pricing problems involving multiple products (Ringbeck and Huchzermeier, 2019).

As the number of items increases, the number of decisions to make, i.e., the bundle prices to set, increases exponentially. Given n items, black-box optimization methods applied to this problem face a search space of dimension $O(2^n)$. As n grows, this exponential search space makes it increasingly difficult for such methods to find high-quality solutions within a fixed number of iterations. Our motivation in this research is to translate customer preference feedback into item-specific information, enabling learning in an $O(n)$ -dimensional space rather than $O(2^n)$. The market structure learned in this lower-dimensional space can then be used to solve for the optimal bundle prices via the proposed optimization model.

In this paper, we study how a seller offering mixed bundles can learn the structure of a multi-segmented market (the item-level valuations of each segment and the relative sizes of these segments) from a sequence of pricing interactions with customers, and use this information to compute near-optimal bundle prices. We propose BLMP, an algorithm that transforms customer purchase data into a form that enables the estimation of segment valuations and probabilities, and then uses these market parameter estimates to approximate the optimal prices. Given the same number of customer interactions, we expect our algorithm to learn the underlying market structure more efficiently than generic black-box optimization methods, and to converge to higher-quality solutions.

This paper contributes to the literature by proposing this approach to learning and pricing in mixed-bundling markets. We evaluate the resulting algorithm, BLMP, against two representative

black-box optimization benchmarks, CMA-ES and GP-UCB, and find that it offers a more scalable and practical alternative for this problem.

The remainder of this paper is structured as follows. Section 2 defines the problem setting. Section 3 details our proposed algorithm. Experimental results are presented in Section 4, followed by concluding remarks in Section 5.

2 Problem Definition

We consider a single seller offering n items to a multi-segmented market. Each item i is associated with a low valuation L_i and a high valuation H_i . These are the maximum amounts a customer is willing to pay for item i . L_i and H_i can be deterministic, or come from a nontrivial probability distribution. We assume that the market consists of 2^n segments, each corresponding to a unique combination of L/H values across the n items, and let δ_j denote the probability (or population fraction) that a customer belongs to segment $j \in \{1, 2, \dots, 2^n\}$, where $\sum_{j \in J} \delta_j = 1$. Markets with this structure have been studied in the literature (Aydin and Ziya, 2008; Gökgür and Karabatı, 2019).

In each period a random customer from this multi-segmented market arrives, with valuations $v_i \in \{L_i, H_i\}$ for each item i . The seller offers all 2^n subsets of the n items as purchase options and must set a price for each, except the empty subset, which is fixed at zero to represent the no-purchase option. The price vector is denoted by p , and $p \in R_+^n$.

For all options $B \subseteq \{1, \dots, n\}$, the customer calculates the consumer surplus as the sum of the valuations of the items in B minus the price of B , and chooses the option with the greatest consumer surplus. Consistent with the relevant literature (Adams and Yellen, 1976; Venkatesh and Kamakura, 2003), we assume that each customer purchases at most one unit of each product.

$$B^* = \arg \max_{B \subseteq \{1, \dots, n\}} \sum_{i \in B} v_i - p_B \quad (1)$$

Although valuations may be stochastic, we assume that, conditional on their realization, the customer's choice is deterministic - always the option with the highest surplus. In the case of multiple maximizers of the consumer surplus, the customer chooses the option containing the largest number of items.

In each period, the seller is allowed to change the prices before the customer arrives. However, the seller never observes the customer's segment or valuations, and therefore cannot customize prices for individual customers. The seller's only feedback is the customer's preference, i.e. selected option B^* , and the revenue collected, p_{B^*} , and the learning data is created as $(p, (B^*, p_{B^*}))$, where p is the input data, and (B^*, p_{B^*}) is the label.

The goal of the seller is to calculate the revenue maximizing prices by interacting with T customers and observing their preferences, where T is finite and known in advance.

3 Methodology

Rather than feeding price–preference data directly into black-box optimization tools, our approach extracts the underlying market parameters from this data and uses them to solve for revenue-maximizing prices. This allows the learning problem to operate over an $O(n)$ -dimensional search space, instead of the $O(2^n)$ -dimensional space induced by the full set of bundle prices.

Our proposed algorithm consists of three components: (1) Decompose preference data by item, (2) Learn the market parameters, (3) Approximate the optimal solution. We describe each component in detail in the following subsections.

3.1 Decompose Preference Data

The price vector and the resulting preference feedback are jointly dependent across items, so the feedback we observe is associated with a 2^n -dimensional space. From this data, we aim to extract information pertaining only to the individual item valuations v_i . Specifically, for each item we construct a separate training dataset associated only with v_i , which we then use in the second component (Learn the Market Parameters). In this section, we describe this extraction procedure.

The fact that the customer’s chosen option is B^* , as determined by the decision mechanism described in (1), implies the following:

1. For any item $i \in B^*$, B^* has a higher consumer surplus than $B^* \setminus i$.

$$\begin{aligned} \sum_{j \in B^*} v_j - p_{B^*} &\geq \sum_{j \in B^* \setminus i} v_j - p_{B^* \setminus i} \\ \sum_{j \in B^* \setminus i} v_j + v_i - p_{B^*} &\geq \sum_{j \in B^* \setminus i} v_j - p_{B^* \setminus i} \\ v_i &\geq p_{B^*} - p_{B^* \setminus i} \end{aligned}$$

If the customer t is offered item i at a price of $p_{B^*} - p_{B^* \setminus i}$, since their valuation for item i is above that price, the customer would purchase the item. This information is transferred to the training data of item i as the point $x_i^t = p_{B^*} - p_{B^* \setminus i}$ with its label $y_i^t = 1$.

2. For any item $i \notin B^*$, B^* has a higher consumer surplus than $B^* \cup i$.

$$\begin{aligned} \sum_{j \in B^*} v_j - p_{B^*} &> \sum_{j \in B^* \cup i} v_j - p_{B^* \cup i} \\ \sum_{j \in B^*} v_j - p_{B^*} &> \sum_{j \in B^*} v_j + v_i - p_{B^* \cup i} \\ p_{B^* \cup i} - p_{B^*} &> v_i \end{aligned}$$

If the customer t is offered item i at a price of $p_{B^* \cup i} - p_{B^*}$, since their valuation for item i is below that price, the customer would not purchase the item. This information is transferred to the training data of item i as the point $x_i^t = p_{B^* \cup i} - p_{B^*}$ with its label $y_i^t = 0$.

Repeating this procedure over all T customer interactions, we obtain, for each item i , a training dataset $\{(x_i^t, y_i^t)\}$ consisting of price points and their associated purchase/non-purchase labels. This decomposition reduces the original 2^n -dimensional preference feedback into n independent, item-level datasets, one for each v_i .

3.2 Learn Market Parameters

There are two market parameters that determine the optimal bundle prices: the item valuations (either deterministic constants or the parameters of their underlying probability distributions) and the segment fractions. These two parameters are learned sequentially.

3.2.1 Learning Valuations

For each item i , the training data generated according to Section 3.1 has the structure shown in Figure 1. Each point (x_i^t, y_i^t) represents a customer interaction. Given the price x_i^t offered for item i , $y_i^t = 1$ if the customer would purchase item i at that price, and $y_i^t = 0$ otherwise.

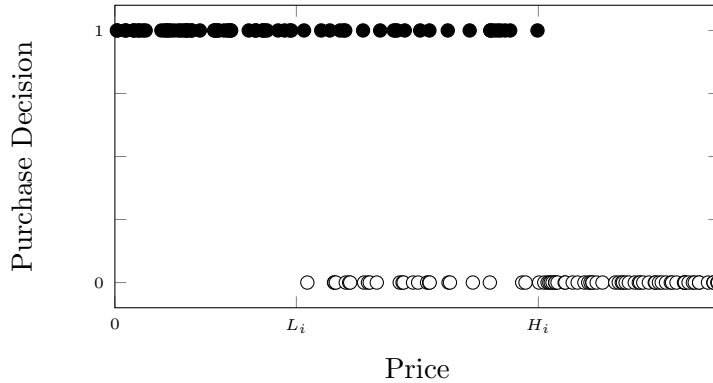


Figure 1: Sample training data for item i

Although the market consists of 2^n segments, from the perspective of item i there are only two segments: a segment with a low valuation L_i for item i , and a segment with a high valuation H_i for item i . In this reduced two-segmented view, if the offered price x_i^t is less than L_i , all customers purchase i ($y_i^t = 1$); if it is above H_i , no customer purchases i ($y_i^t = 0$); and for prices in between, only customers with high valuation make the purchase. In this structure, we search for breakpoints in the training data to estimate L_i and H_i .

We propose two methods: one for deterministic L_i and H_i , and one for bounded probability distributions for L_i and H_i , i.e. $L_i \in [L_{ia}, L_{ib}]$ and $H_i \in [H_{ia}, H_{ib}]$.

Deterministic Valuations

For each unknown valuation, we propose three estimates: lower, upper and middle, that are expected to converge to the same value with sufficient training data.

The smallest x_i^t value where we observe $y_i^t = 0$ is the upper estimate for L_i , L_i^u . This is the smallest price for which a no-purchase decision is observed. The largest x_i^t value that is smaller than L_i^u is the lower estimate for L_i , L_i^l . Their average is the middle estimate L_i^m .

$$L_i^u = \min_t \{x_i^t : y_i^t = 0\} \quad (2)$$

$$L_i^l = \max_t \{x_i^t : x_i^t < L_i^u\} \quad (3)$$

$$L_i^m = (L_i^u + L_i^l)/2 \quad (4)$$

Similarly, the largest x_i^t value where we observe $y_i^t = 1$ is the lower estimate for H_i , H_i^l . This is the highest price for which a purchase decision is observed. The smallest x_i^t value that is larger than H_i^l is the upper estimate for H_i , H_i^u . Their average is the middle estimate H_i^m .

$$H_i^l = \max_t \{x_i^t : y_i^t = 1\} \quad (5)$$

$$H_i^u = \min_t \{x_i^t : x_i^t > H_i^l\} \quad (6)$$

$$H_i^m = (H_i^u + H_i^l)/2 \quad (7)$$

Our preliminary experiments indicate that, across most of the settings we tested, the difference in estimation error between L_i^l and L_i^m was not statistically significant, and that both are smaller than the estimation error of L_i^u ; and the difference in estimation error between H_i^u and H_i^m was not statistically significant, and that both are smaller than the estimation error of H_i^l . However, we also observed that among the three estimates, the revenue collected under the optimal price computed from the estimated valuations is less sensitive to the magnitude of the estimation error, and more sensitive to whether the valuations are over- or underestimated.

As an example, suppose the true low valuation of item i is 30, and the low- and high-paying segments each occur with probability 0.5. Suppose we underestimate the low valuation as 29, and the optimal solution accordingly sets the price at 29, targeting the entire market. Since $29 \leq L_i$, the realized revenue matches this expectation: all customers purchase, yielding an expected revenue of 29. Now suppose instead we overestimate the low valuation as 30.5, and the optimal solution again sets the price at the (estimated) low valuation, 30.5. However, since $30.5 > L_i$, only the high-paying segment purchases at this price, and we lose the low-paying segment entirely; the realized expected revenue drops sharply to $30.5 \times 0.5 = 15.25$. A smaller estimation error leads to drastic deviation from the optimal revenue because of its direction.

For this reason, we recommend using L_i^l and H_i^l , defined in (3) and (5), respectively, the

estimates with the lowest probability of overestimation.

Bounded Probabilistic Valuations

When L_i and H_i come from bounded non-overlapping probability distributions with $L_i \in [L_{ia}, L_{ib}]$ and $H_i \in [H_{ia}, H_{ib}]$, the expected purchase probability function, i.e. $E[y_i^t | x_i^t]$ taken over the training data in Figure 1 has the structure of the function in Figure 2. For $x_i^t \leq L_{ia}$, $E[y_i^t] = 1$, for $x_i^t > H_{ib}$, $E[y_i^t] = 0$, for $x_i^t \in (L_{ib}, H_{ia}]$, $E[y_i^t] = \Delta_i$, where Δ_i is the sum of the probabilities of segments with high valuation for item i ; and for the regions $x_i^t \in (L_{ia}, L_{ib}]$ and $x_i^t \in (H_{ia}, H_{ib}]$, the function $E[y_i^t]$ is the complementary cumulative density function scaled to $[1, \Delta_i]$ and $[\Delta_i, 0]$ respectively.

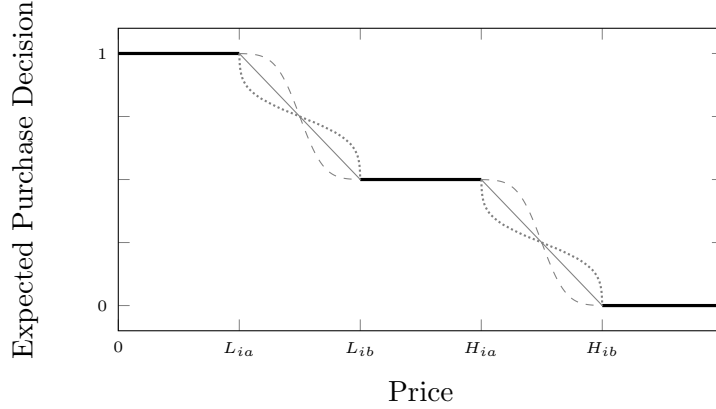


Figure 2: Expected purchase probability function structure for item i

In order to detect this structure, we must work with $E[y_i]$ rather than y_i . We apply k NN to the training data to obtain an approximation $\hat{y}_i$ of $E[y_i]$. The resulting approximation $\{(x_i^t, \hat{y}_i^t)\}$ is exemplified in Figure 3.

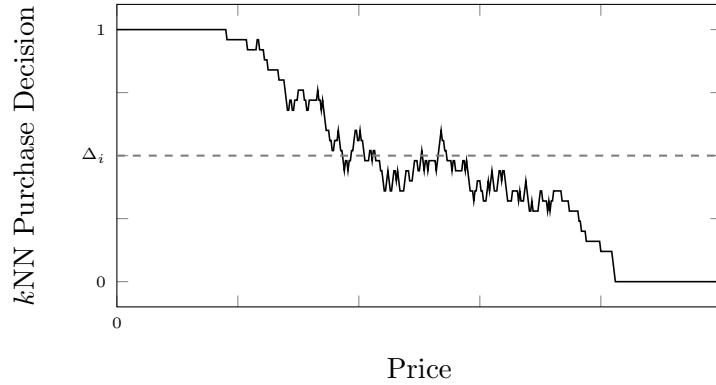


Figure 3: k NN applied to the sample training data of item i

From this data set, we estimate, in the following order: (1) L_{ia} and H_{ib} , (2) Δ_i , (3) L_{ib} and H_{ia} .

Step 1: L_{ia} is calculated according to (3) as the lower threshold for the no-purchase decision, and H_{ib} is calculated according to (6) as the upper threshold for the purchase decision.

Step 2: We filter the data set and let $\mathcal{Y}_i = \{\hat{y}_i^t : x_i^t \in [L_{ia}, H_{ib}]\}$. $\mathcal{Y}_i$ is expected to consist of three blocks: the upper tail $(1, \Delta_i)$, the main body concentrated around Δ_i , and the lower tail $(\Delta_i, 0)$. We apply k -Means clustering to $\mathcal{Y}_i$ with $k = 3$ and take the centroid of the second cluster as the estimate for Δ_i .

Step 3: We locate the points where the upper and lower tails cross the level $\hat{y}_i^t = \Delta_i$. While increasing the price x_i^t from L_{ia} , as $\hat{y}_i^t$ starts decreasing from 1, the first time it crosses Δ_i is the estimation of L_{ib} . While decreasing the price x_i^t from H_{ib} , as $\hat{y}_i^t$ starts increasing from 0, the first time it crosses Δ_i is the estimation of H_{ia} .

$$L_{ib} = \min_t \{x_i^t : \hat{y}_i^t \leq \Delta_i\} \quad (8)$$

$$H_{ia} = \max_t \{x_i^t : \hat{y}_i^t \geq \Delta_i\} \quad (9)$$

We observe that L_{ia} and H_{ib} converge to the correct region quite fast. L_{ib} and H_{ia} estimations require more data points to stabilize due to the fluctuations in the k NN approximation.

Notice that the convergence guarantee of this method is independent of the distribution itself. However, how prone the algorithm is to over or underestimate L_{ib} and H_{ia} depends on the shape of the cumulative density function.

3.2.2 Learning Segment Probabilities

The process of estimating segment probabilities is almost identical for deterministic and probabilistic valuations, and requires that the data be restricted to customers t such that $x_i^t \in (L_i, H_i)$ in the deterministic case, $x_i^t \in (L_{ib}, H_{ia})$ in the probabilistic case for all $i = 1, \dots, n$.

This can be achieved by filtering the existing data with this condition; however, the number of data points in the remaining set cannot be controlled this way. Alternatively, another set of customer interactions can be conducted, where the price of each item is restricted to the desired interval, and no bundle discounts are applied to the other sales options. This way, the marginal price $x_i^t = p_{B^* \cup i} - p_{B^*}$ or $x_i^t = p_{B^*} - p_{B^* \setminus i}$ reduces to the (undiscounted) price of item i itself, restricted to the desired bounds.

Let's illustrate the main idea of the algorithm with an example with $n = 3$, and for a customer t from a segment with valuation profile (L_1, H_2, L_3) . If the offered item prices are in the given intervals (L_i, H_i) , we expect the price of items 1 and 3 to exceed the customer's valuation and the price of item 2 to fall below it; and therefore have $y_1^t = y_3^t = 0$ and $y_2^t = 1$. To generalize this logic, we represent the purchase decision of segment j as a vector $\mathbf{Y}^j \in \{0, 1\}^n$, where $\mathbf{Y}_i^j = 0$ if the valuation of item i in segment j is L_i , and $\mathbf{Y}_i^j = 1$ if the valuation of item i in segment j is H_i .

For all 2^n segment profiles $\mathbf{Y}^j$, we count all customers t whose purchase decision matches the profile (i.e. $\mathbf{Y}_i^j = y_i^t$ for all $i = 1, \dots, n$). The frequency of these observations gives the estimates

for the segment probabilities δ_j .

3.3 Optimize

Hanson and Martin (1990) study a mixed-integer programming formulation for optimal bundle prices under deterministic valuations, in a more general setting than ours. In this paper, we make minor modifications to their formulation to fit our problem setting.

Given n items, the MIP model assumes a set of sales options, $L = \{1, \dots, 2^n\}$, and a set of customer segments $S = \{1, \dots, 2^n\}$. For each option i , we let $B(i) \subseteq \{1, \dots, n\}$ be the items included in option i . The valuation of option i for customers in segment k , R_{ki} , and the fraction of segment k in the market, N_k , are parameters of the formulation. The model solves for the price of option i (p_i), whether customers in segment k choose option i as their favorite (θ_{ki}), and the revenue generated by customers in segment k selecting option i (z_{ki}).

The MIP formulation, referred to as OPT-BP (Optimal Bundle Prices), is as follows:

$$\max \quad \sum_{i \in L} \sum_{k \in S} N_k z_{ki} - \epsilon \sum_i p_i \quad (10)$$

$$\text{s.t.} \quad \sum_{i \in L} \theta_{ki} = 1 \quad k \in S \quad (11)$$

$$M_1(1 - \theta_{ki}) + R_{ki} - p_i \geq R_{kj} - p_j \quad k \in S, i, j \in L \quad (12)$$

$$z_{ki} \leq M_2 \theta_{ki} \quad k \in S, i \in L \quad (13)$$

$$z_{ki} \leq p_i \quad k \in S, i \in L \quad (14)$$

$$p_1 = 0 \quad (15)$$

$$p_i \leq \sum_{j \in I} p_j \quad i \in L, \cup_{j \in I} B(j) = B(i) \quad (16)$$

$$p_i, z_{ki} \geq 0, \theta_{ki} \in \{0, 1\} \quad k \in S, i \in L \quad (17)$$

The objective function (10) maximizes the expected revenue and uses the minimum prices as a tie-breaker. (11) ensures that every segment has a selected option, and (12) forces the selected option to have the highest consumer surplus. The value $R_{max} = \max_{k,i} R_{k,i}$ can be used for both M_1 and M_2 , as it upper-bounds the largest possible surplus or revenue term in the respective constraints. Segment k 's revenue generated by option i is allowed to be nonzero only if option i is selected by segment k (13), and this revenue is upperbounded by the price of option i (14). The empty bundle must have a price of zero (15). As for the subadditivity of the prices, (16) ensures that the price of any option i cannot exceed the sum of the prices of any set of options $I \subseteq L$ whose bundles union to $B(i)$. Note that this set of constraints may be exponential in the number of options, and therefore doubly exponential in the number of items.

To the best of our knowledge, no well-defined optimization model exists in the literature for

settings with probabilistic valuations when $n > 2$. The special case of $n = 2$ products with uniform distributions has been studied by Bhargava (2013), and a nonlinear optimization model has been formulated to calculate the optimal bundle prices.

For probabilistic valuations with bounded distributions in general, we propose approximating the optimal solutions by solving OPT-BP using the estimated lower bounds of each item's valuation (L_{ia} and H_{ia}) to calculate the deterministic R_{ki} . We choose the lower bounds of the intervals to avoid the overestimation errors discussed in Section 3.2.1.

With the three components detailed in Sections 3.1, 3.2, 3.3, the proposed algorithm in this paper BLMP (Bundle Pricing via Learning the Market from Customer Preferences) is summarized in Algorithm 1 below.

Algorithm 1 BLMP: Bundle Pricing via Learning the Market from Customer Preferences

Require: Bounds (g, G) on the valuations

```

1: for  $t = 1, \dots, T_1$  customers do
2:   Generate item prices in  $U(g, G)$  and bundle discounts as desired
3:   Collect preference information from a random customer
4:   Decompose the preference data for each item  $\{(x_i^t, y_i^t)\}_{i=1}^n$ 
5: end for
6: for  $i = 1, \dots, n$  items do
7:    $L, H$  = Estimate valuations  $\triangleright L_a, L_b, H_a, H_b$  if probabilistic
8: end for
9: for  $t = 1, \dots, T_2$  customers do
10:  Generate item prices in  $U(L, H)$  and no bundle discounts  $\triangleright U(L_b, H_a)$  if probabilistic
11:  Collect preference information from a random customer
12:  Decompose the preference data for each item  $\{(x_i^t, y_i^t)\}_{i=1}^n$ 
13: end for
14:  $\delta$  = Estimate Segment Probabilities
15: Solve OPT-BP with  $L, H, \delta$   $\triangleright L_a, H_a, \delta$  if probabilistic

```

4 Experiments

4.1 Benchmark

To evaluate the performance of BLMP, we use the following two black-box optimization algorithms as benchmark:

1. **CMA-ES**: Covariance Matrix Adaptation Evolution Strategy (CMA-ES) is a derivative-free evolutionary method commonly used for blackbox optimization in continuous domains (Hansen and Ostermeier, 2001). The strength of this algorithm among various black-box optimization methods has also been shown in the literature (Pošík and Kubalík, 2012).

For n items, the search space of CMA-ES is $2^n - 1$ dimensional (as the price of the no-purchase option is zero, it is discarded as a variable). The first n elements correspond

to individual item prices scaled to $[0, 1]$, and the remaining elements associated with the bundles of these items correspond to the bundle discount rate, rather than bundle prices.

Beside the initialization points, we allow CMA-ES the same number of customer interactions as BLMP ($T_1 + T_2$); population size is set to 10 (10 interactions per iteration) and the maximum number of iterations is $(T_1 + T_2)/10$.

2. **GP-UCB**: Gaussian Process Upper Confidence Bound (GP-UCB) is a Bayesian optimization method that balances exploration and exploitation using a Gaussian process surrogate model (Srinivas et al., 2010).

Similarly, for n items, the search space of GP-UCB is $2^n - 1$ dimensional. The first n elements correspond to individual item prices, and the remaining elements associated with the bundles of these items correspond to the bundle discount rate.

To improve computational efficiency, we employ batch Bayesian optimization by selecting multiple candidate price vectors per iteration via the qUCB acquisition function, reducing the number of costly GP refitting operations while maintaining comparable solution quality. Beside the initialization points, we allow GP-UCB the same number of customer interactions as BLMP ($T_1 + T_2$); 5 candidates (interactions) per iteration and the maximum number of iterations is $(T_1 + T_2)/5$.

4.2 Experimental Setup

Our experiments consider $n = 2, 3, 4$ items. In each randomly generated market, for each item $i = 1, \dots, n$, L_i is a fixed value drawn from $U(20, 40)$, and H_i from $U(60, 80)$. We consider three settings: (1) each customer’s valuation is fixed at these values (denoted **D**, for Deterministic); (2) each customer’s valuation is drawn from $U(0.9L_i, 1.1L_i)$ or $U(0.9H_i, 1.1H_i)$, depending on the customer’s segment (denoted **U**, for Uniform); and (3) each customer’s valuation is drawn from $L_i(0.9 + 0.2\text{Beta}(2^{n-1}, 2^{n-1}))$ or $H_i(0.9 + 0.2\text{Beta}(2^{n-1}, 2^{n-1}))$, depending on the customer’s segment (denoted **B**, for Beta). For both **U** and **B**, we set $L_{ia} = 0.9L_i$, $L_{ib} = 1.1L_i$, $H_{ia} = 0.9H_i$, and $H_{ib} = 1.1H_i$. In all markets, segment probabilities are equal across segments, i.e., $\delta_j = 2^{-n}$ for all j . For each combination of dimension n and distribution setting, we generate 100 random markets.

We compare all three methods (BLMP, CMA-ES, GP-UCB) by giving each the same number of customer interactions T . During our experiments, we observed that the per-iteration runtime of GP-UCB, that is the time required to compute the next price to offer, increased substantially as the number of iterations grew. This is because each iteration requires recomputing the covariance matrix over all points observed so far, an operation whose cost scales as $O(m^3)$ in the number of observed points m . As a result, practical application of GP-UCB became infeasible for larger T . Therefore, while BLMP is tested with $T_1, T_2 \in \{1000, 2000, 3000, 4000, 5000\}$ and CMA-ES is tested with $T \in \{2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000\}$, GP-UCB is

tested only at $T = 2000$, and over a reduced set of 20 random markets (rather than the 100 used for the other methods).

As a benchmark for algorithm performance, we compute a *full-information benchmark*: the best price we can compute assuming knowledge of the true market parameters. This benchmark coincides with the true optimal solution whenever computing the optimum is tractable; when it is not, we instead compute the best price attainable using the same solution approach as our algorithm, given full information rather than estimated parameters.

For **D**, the optimal price is computed directly, since customer valuations are known exactly. For $n = 2$ in the **U** setting, the optimal price can be formulated as a nonlinear optimization problem following (Bhargava, 2013), which we solve using differential evolution. For $n = 2$ in the **B** setting, no closed-form formulation is available, so the true optimum cannot be computed; we find that using the $n = 2$ **U**-optimal price outperforms solving OPT-BP with lower-bound values, and therefore use the expected revenue of this price, evaluated under the **B** distribution, as the benchmark.

For $n > 2$, no closed-form solution is available in either the **U** or **B** setting, and the true optimum cannot be computed even under full information. In this case, analogous to the Optimize step of our algorithm (Section 3.3), we compute the benchmark by solving OPT-BP using the *true* values of L_{ia} , H_{ia} , and δ , rather than the estimates used by our algorithm. This isolates the effect of estimation error: the only difference between our algorithm’s final solution and this benchmark is the use of true versus estimated parameter values.

The algorithms are evaluated using the following two metrics:

1. Average revenue per period, collected throughout the algorithm’s execution.
2. Expected revenue under the final prices returned by the algorithm.

These two metrics capture, respectively, the quality of the learning process itself and the quality of its final output. For both metrics, we report the raw value as well as its exact and relative deviation from the corresponding benchmark value.

4.3 Results

We first discuss several implementation-level observations for BLMP, before comparing its performance to the baseline methods.

For deterministic customer valuations, it has been observed (Naderivandi et al., 2025) that the penalty of overestimating the valuations significantly outweighs underestimating them. This phenomenon was also discussed in Section 3.2.1. The results of our experiments also verify this. As a remedy, we propose subtracting a correction term α from the valuation estimates before calculating the optimal price in Line 15 of Algorithm 1. Of the two estimation methods, (5) is

guaranteed not to overestimate H_i , while (3) can potentially overestimate L_i ; the correction is therefore applied only to the low-valuation estimates. Table 1 below shows the impact of the value of α on the performance of the algorithm’s proposed price, as a function of T_1 and T_2 , for $n = 2$ and **D** setting.

Table 1: Mean relative deviation (%) from optimal, by T_1 , T_2 , and α for $n = 2$

T_1	T_2	$\alpha = 0.0$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.15$	$\alpha = 0.2$	$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 1.0$
1000	1000	-9.49	-8.14	-6.41	-5.30	-4.10	-2.94	-1.45	-1.09
1000	2000	-11.40	-9.65	-7.41	-5.90	-4.52	-3.81	-2.16	-1.01
1000	3000	-9.41	-7.44	-5.94	-4.41	-3.52	-2.91	-2.02	-1.02
1000	4000	-9.75	-8.61	-6.45	-4.90	-4.09	-3.26	-1.27	-1.28
1000	5000	-11.52	-8.59	-6.94	-6.49	-5.58	-4.03	-1.66	-1.11
2000	1000	-9.57	-5.06	-2.46	-1.56	-1.61	-1.20	-0.89	-1.01
2000	2000	-9.06	-5.27	-3.55	-1.96	-1.26	-0.96	-0.54	-0.96
2000	3000	-9.50	-6.37	-3.88	-3.24	-1.93	-1.76	-0.74	-0.93
2000	4000	-8.79	-5.40	-3.74	-2.97	-2.21	-1.41	-0.77	-0.95
2000	5000	-9.78	-5.03	-2.88	-2.29	-1.86	-0.86	-0.52	-0.94
3000	1000	-10.71	-5.94	-2.36	-1.05	-0.40	-0.31	-0.56	-0.96
3000	2000	-10.19	-5.15	-2.16	-0.97	-0.34	-0.29	-0.51	-0.94
3000	3000	-11.21	-4.05	-2.19	-1.16	-1.00	-0.79	-0.53	-0.93
3000	4000	-8.31	-3.99	-1.76	-0.64	-0.57	-0.28	-0.50	-0.92
3000	5000	-10.39	-6.90	-3.17	-1.45	-1.16	-0.84	-0.76	-1.18
4000	1000	-10.25	-2.46	-0.75	-0.27	-0.31	-0.36	-0.56	-0.98
4000	2000	-9.34	-3.38	-1.82	-0.81	-0.50	-0.29	-0.50	-0.92
4000	3000	-9.82	-3.15	-0.70	-0.17	-0.21	-0.26	-0.48	-0.92
4000	4000	-9.09	-4.00	-2.11	-0.72	-0.35	-0.27	-0.50	-0.90
4000	5000	-8.91	-3.17	-1.99	-1.02	-0.35	-0.27	-0.49	-0.91
5000	1000	-8.56	-2.77	-0.76	-0.29	-0.34	-0.28	-0.50	-0.94
5000	2000	-8.65	-1.73	-0.78	-0.68	-0.73	-0.53	-0.75	-0.90
5000	3000	-11.01	-2.70	-0.94	-0.29	-0.23	-0.27	-0.50	-0.91
5000	4000	-10.17	-2.59	-1.46	-0.88	-0.49	-0.53	-0.73	-1.15
5000	5000	-8.87	-3.34	-0.80	-0.49	-0.30	-0.25	-0.46	-0.88

Table 1 suggests two important conclusions. First, the impact of α depends on T_1 rather than T_2 . This is expected, since T_1 determines the length of the process used to estimate the valuations, whereas T_2 plays no role in valuation estimations. Second, as T_1 increases, valuation estimates improve, so a smaller correction term leads to higher expected revenues. The same trend is observed for $n = 3$ and $n = 4$.

The value of α can be fine-tuned depending on the problem setting and the number of customer interactions; however for simplicity, we fix $\alpha = 0.25$ for all results reported under the **D** setting from this point forward.

Given a fixed budget of customer interactions T , we analyze the impact of the split between T_1 and T_2 on algorithm performance. Figures 4, 5, and 6 show heatmaps of the average final

expected revenue of BLMP's returned prices, for each (T_1, T_2) combination, across all n values, for the **D**, **B**, and **U** settings respectively.

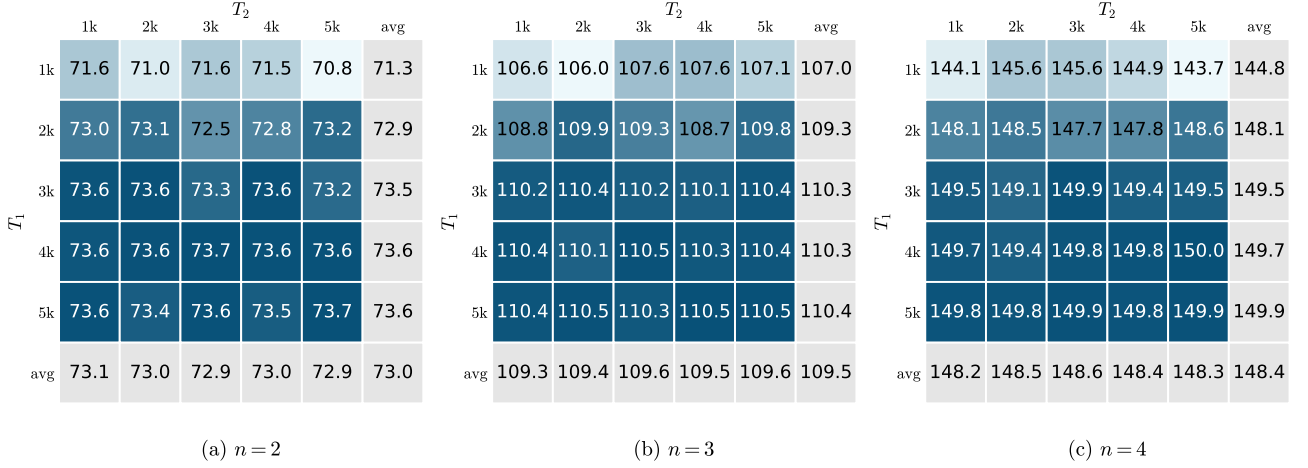


Figure 4: Mean expected revenue of the final solution of BLMP for the **D** distribution

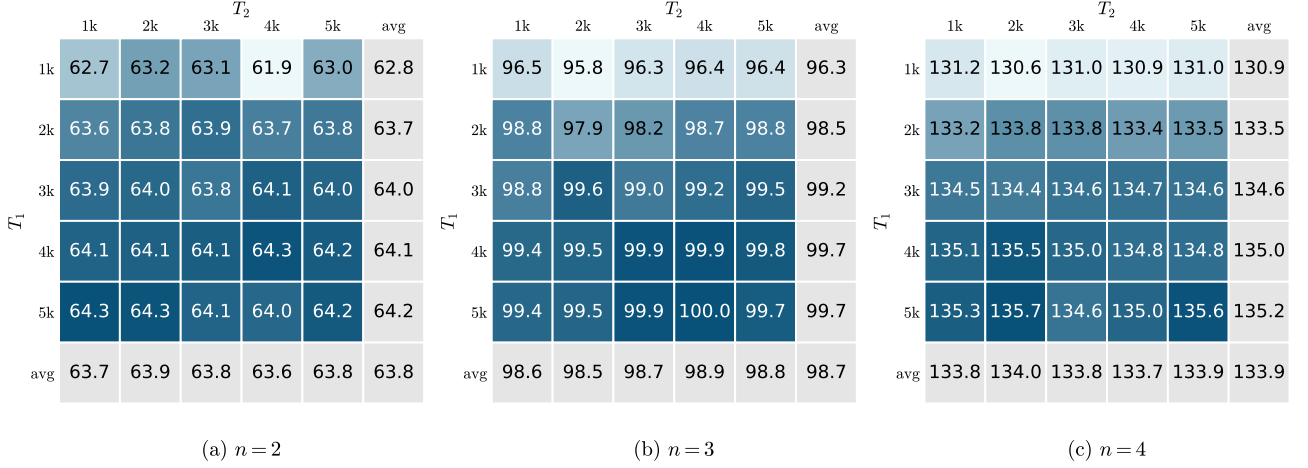


Figure 5: Mean expected revenue of the final solution of BLMP for the **B** distribution

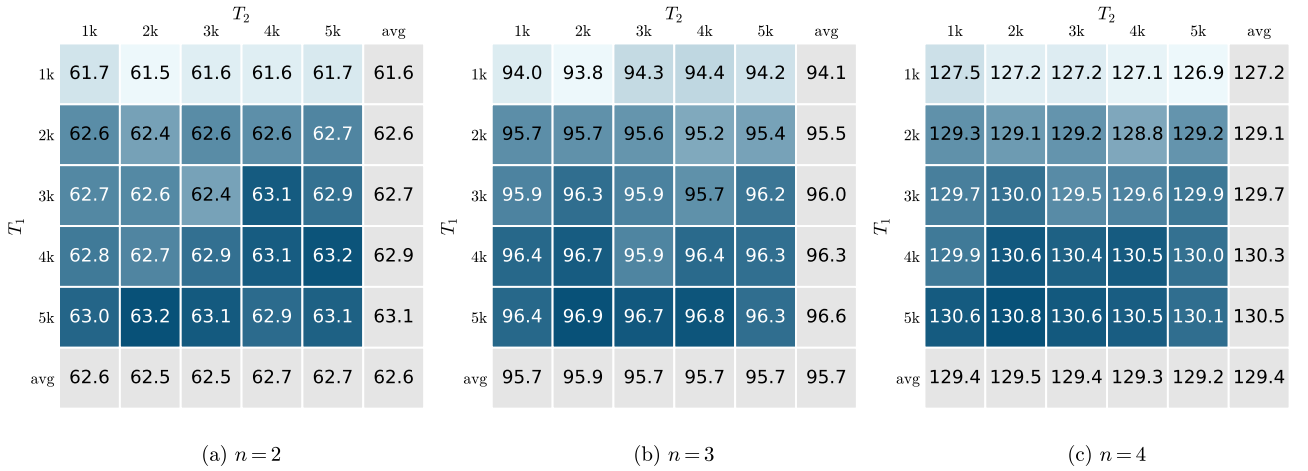


Figure 6: Mean expected revenue of the final solution of BLMP for the **U** distribution

A common pattern observed in all nine heatmaps is that $T_2 = 1000$ is sufficient to saturate the learning process, and increasing T_2 further has no significant impact on performance. In

contrast, increasing T_1 has a consistently positive impact on performance. This analysis of final expected revenue favors partitions with higher T_1 values.

As for the second performance metric, average collected revenue, recall that BLMP offers T_1 prices from the full search space (Algorithm 1, Line 2), and T_2 prices from a restricted search space lying between the high and low valuations (Algorithm 1, Line 10). The expected revenue of the second phase is naturally higher than that of the first. Therefore, we expect higher average collected revenue for settings with a higher T_2 relative to T_1 . Figure 7 illustrates this result, showing average collected revenue as a function of the T_1/T_2 ratio, separately for each dimension n . Across all dimensions and distributions, average collected revenue decreases substantially as the T_1/T_2 ratio increases. This result conflicts with the recommendation from the final expected revenue analysis, and leaves the user with a trade-off between the two performance metrics.

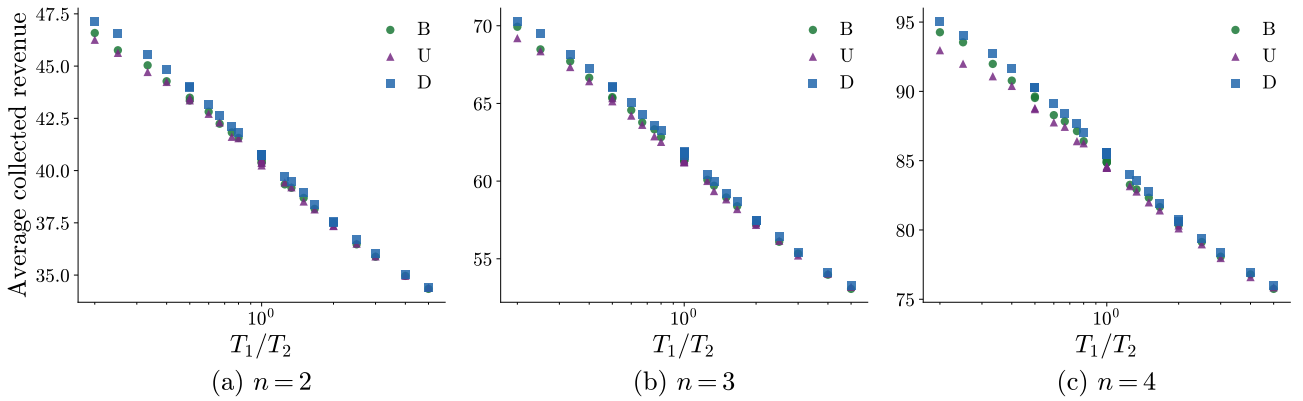
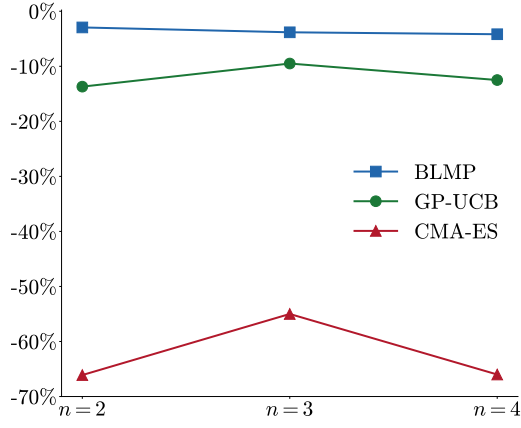


Figure 7: Average revenue collected by BLMP as a function of T_1/T_2 ratio

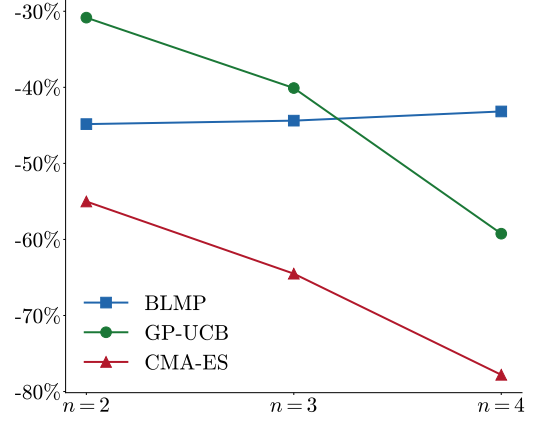
To compare BLMP with the benchmark methods CMA-ES and GP-UCB, we restrict the setting to $T_1 = T_2 = 1000$, as $T = 2000$ was the largest budget at which GP-UCB produced stable results, despite its impractical runtime (as discussed in Section 4.2). Figures 8 and 9 show, in panels (a) and (b) respectively, the percent deviation from the full-information benchmark of (a) the expected revenue of the algorithms' final solutions, and (b) the average revenue collected by the algorithms.

Figure 8, generated using $T = 2000$ and the **D** setting, examines the impact of dimension on algorithm performance. In Figure 8a, we observe that BLMP returns final prices that significantly outperform both benchmarks across all dimensions. As for the average revenue collected, shown in Figure 8b, both benchmarks show collected revenue decreasing with dimension, unlike BLMP, which shows a slightly positive trend. As a result, the dominant algorithm shifts from GP-UCB to BLMP as n increases.

Figure 9, generated using $T = 2000$ and $n = 2$, examines the impact of the valuation distribution on algorithm performance. In Figure 9a, we observe that GP-UCB slightly outperforms BLMP in terms of the quality of the final solution in probabilistic settings. As for the average revenue collected, shown in Figure 9b, the relative performance of all three algorithms is similar across distributions. One conclusion that can be drawn from Figures 8 and 9 is that GP-UCB

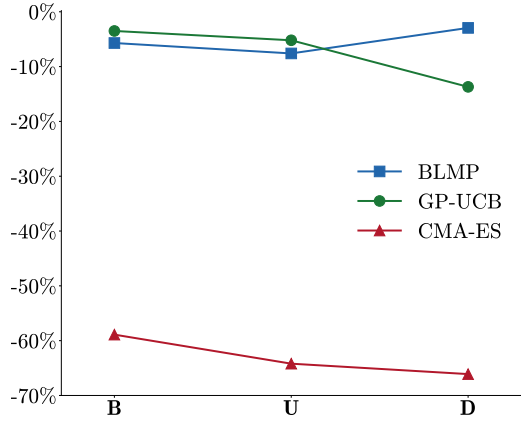


(a) Expected revenue of the final solutions

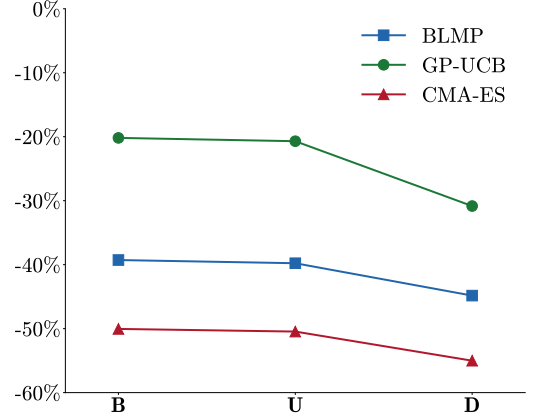


(b) Average revenue collected

Figure 8: Percent deviation of algorithm performance from the full-information benchmark, across dimensions, for the **D** setting.



(a) Expected revenue of the final solutions



(b) Average revenue collected

Figure 9: Percent deviation of algorithm performance from the full-information benchmark, across valuation distributions, for $n = 2$.

and BLMP each have advantages in different settings and metrics, while CMA-ES is consistently outperformed by both alternatives across all settings and metrics considered.

Finally, we study the impact of the number of customer interactions T on the performance of BLMP and CMA-ES. Figures 10a and 10b display the expected revenue of the algorithms' final solutions and the average revenue collected by the algorithms, respectively. The performance metrics are calculated as a function of T only; therefore, for each T value, Figure 10 shows results averaged across all distribution settings, all dimensions, and (for BLMP) all (T_1, T_2) combinations such that $T_1 + T_2 = T$. The averages of the full-information benchmarks are denoted by the red dashed line.

CMA-ES fails remarkably at this optimization problem, regardless of the amount of training data. Individual simulation results indicate that the generated populations often shift toward higher prices (ones that would yield high revenue if accepted, but have a low probability of being accepted) which degrades the overall effectiveness of the search process. BLMP, on the other

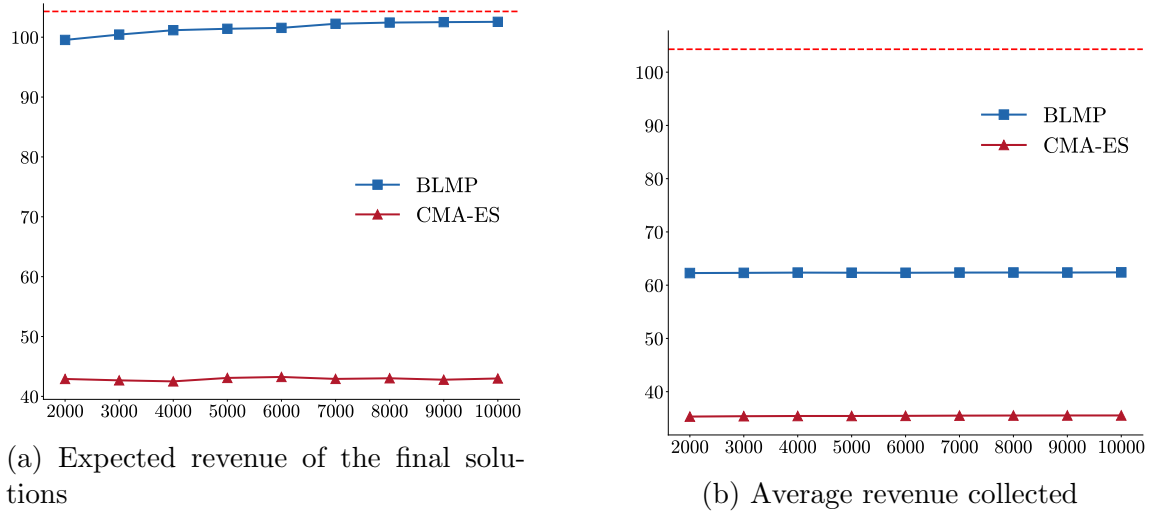


Figure 10: BLMP vs. CMA-ES performance as a function of T

hand, benefits from additional training data, producing higher-quality final solutions.

5 Conclusion

The goal of this study was to design an algorithm for a system in which bundle discounts can be applied to products, that learns the underlying market segment structure by transforming customer purchase preferences into item-specific data, and then uses this information to estimate optimal prices. We aimed to show that this algorithm outperforms conventional black-box optimization methods.

Two methods representing different approaches to black-box optimization, CMA-ES and GP-UCB, were used as benchmarks. CMA-ES was entirely unsuccessful in solving this problem. GP-UCB, while achieving results comparable to our method, encounters stability issues when applied to larger amounts of data, making it infeasible to implement in practice. By decomposing customer preference data into item-specific observations, our approach reduces the dimensionality of the learning problem from $O(2^n)$ to $O(n)$, allowing market parameters to be estimated efficiently even as the number of items grows. Across all tested settings, BLMP’s final solutions remained within 1% of the full-information benchmark under deterministic valuations, and within 5% under both probabilistic settings. Considering both performance and practical applicability, BLMP, the algorithm presented in this paper, can be said to be the strongest of the three.

From a managerial perspective, these results suggest that firms do not need to consider mixed-bundle pricing only as a direct search over a large number of bundle-price decisions. By using the item-level information obtained from observed customer preferences, firms can learn the underlying market structure and then use this information for their pricing decisions. This becomes more important as the number of items increases, since the number of possible bundle prices grows exponentially, while BLMP learns valuations at the item level.

BLMP also gives firms some flexibility in how they allocate customer interactions between the two learning stages. Our results show that using more interactions in the first stage generally improves the quality of the final pricing solution, whereas allocating relatively more interactions to the second stage increases the revenue collected during the learning process. Therefore, firms can adjust this allocation depending on whether they give more importance to improving future pricing decisions or to collecting more revenue during learning.

An important advantage of our algorithm lies in the modularity of its two main components, learning and optimization, each of which can be improved or replaced independently without altering the overall structure of the method. In this paper, for optimal bundle-pricing settings that lack a closed-form solution and cannot be expressed as a tractable optimization problem, we used an approximation based on the deterministic optimization model; other heuristics could equally be substituted in this step, and a similar performance analysis, comparing against the full-information setting, could then be carried out for any such heuristic. Likewise, the bound estimation procedures used for probabilistic valuations in Section 3.2.1 could be improved or replaced with alternative estimators, potentially tightening the gap to the full-information benchmark observed under the **B** and **U** settings. This modularity makes the algorithm robust to future improvements in either component, and offers a natural direction for future work.

Acknowledgement

This work was supported by TÜBİTAK (Scientific and Technological Research Council of Turkey), grant number 221M392. The authors would like to thank Ahmet Karakose for his help in the preliminary analysis phase.

Data Availability Statement

The code used to generate the results presented in this manuscript is available at

<https://github.com/EzgiKarabulutTurkseven/Bundle-Pricing-Problem>.

References

- Adams, William James and Janet L. Yellen (1976). “Commodity Bundling and the Burden of Monopoly”. In: *The Quarterly Journal of Economics* 90.3, pp. 475–498.
- Anand, Anush, Pranav Agrawal, and Tejas Bodas (2025). “Bayesian optimization for dynamic pricing and learning”. In: *Performance Evaluation* 170, p. 102519.
- Aydin, Goker and Serhan Ziya (2008). “Pricing promotional products under upselling”. In: *Manufacturing & Service Operations Management* 10.3, pp. 360–376.

- Bhargava, Hemant K. (2013). “Mixed Bundling of Two Independently Valued Goods”. In: *Management Science* 59.9, pp. 2170–2185.
- Boer, Arnoud V. den and N. Bora Keskin (2022). “Dynamic Pricing with Demand Learning and Reference Effects”. In: *Management Science* 68.10, pp. 7112–7130.
- Cai, Shoujie et al. (2026). “Product bundling strategy and pricing for co-livestreaming rooms considering consumer subscription”. In: *Omega* 138, p. 103379.
- Cao, Ping, Nenggui Zhao, and Jie Wu (2019). “Dynamic pricing with Bayesian demand learning and reference price effect”. In: *European Journal of Operational Research* 279.2, pp. 540–556.
- Chen, Jian et al. (2026). “Capacity and Pricing Management with Demand Learning”. In: *Management Science*. Article in Advance. DOI: 10.1287/mnsc.2023.03749. In press.
- Chen, Xi et al. (2022). “Bayesian dynamic learning and pricing with strategic customers”. In: *Production and Operations Management* 31.8, pp. 3125–3142.
- Cheung, Wang Chi, David Simchi-Levi, and He Wang (2017). “Technical Note—Dynamic Pricing and Demand Learning with Limited Price Experimentation”. In: *Operations Research* 65.6, pp. 1722–1731.
- Ettl, Markus et al. (2020). “A Data-Driven Approach to Personalized Bundle Pricing and Recommendation”. In: *Manufacturing & Service Operations Management* 22.3, pp. 461–480.
- Fang, Lang, Jiafu Tang, and Zhendong Pan (2025). “Learning-Based Bundling Strategy for Two Products Under Uncertain Consumer’s Valuations”. In: *IEEE Transactions on Engineering Management* 72, pp. 1970–1982.
- Ferreira, Kris Johnson, David Simchi-Levi, and He Wang (2018). “Online Network Revenue Management Using Thompson Sampling”. In: *Operations Research* 66.6, pp. 1586–1602.
- Fisher, Marshall, Santiago Gallino, and Jun Li (2018). “Competition-Based Dynamic Pricing in Online Retailing: A Methodology Validated with Field Experiments”. In: *Management Science* 64.6, pp. 2496–2514.
- Gökgür, Burak and Selçuk Karabatı (2019). “Dynamic and targeted bundle pricing of two independently valued products”. In: *European Journal of Operational Research* 279.1, pp. 184–198.
- Gökgür, Burak and Ayşe Kocabıyıkoglu (July 2026). “Bundle pricing in a two-stage supply chain with customer heterogeneity”. In: *IMA Journal of Management Mathematics*, dpag022. URL: <https://doi.org/10.1093/imaman/dpag022>.
- Hansen, Nikolaus and Andreas Ostermeier (2001). “Completely derandomized self-adaptation in evolution strategies”. In: *Evolutionary computation* 9.2, pp. 159–195.
- Hanson, Ward and R Kipp Martin (1990). “Optimal bundle pricing”. In: *Management Science* 36.2, pp. 155–174.
- Harrison, J. Michael, N. Bora Keskin, and Assaf Zeevi (2012). “Bayesian Dynamic Pricing Policies: Learning and Earning Under a Binary Prior Distribution”. In: *Management Science* 58.3, pp. 570–586.
- He, Sharlene, Eric T. Anderson, and Derek D. Rucker (2024). “Measuring Willingness to Pay: A Comparative Method of Valuation”. In: *Journal of Marketing* 88.3, pp. 50–68.

- Huang, Tingliang, Chao Liang, and Jingqi Wang (2018). “The Value of “Bespoke”: Demand Learning, Preference Learning, and Customer Behavior”. In: *Management Science* 64.7, pp. 3129–3145.
- Jain, Kunal, K. J. Prabuchandran, and Tejas Bodas (2023). “Bayesian Optimization for Function Compositions with Applications to Dynamic Pricing”. In: *Learning and Intelligent Optimization*. Ed. by Meinolf Sellmann and Kevin Tierney. Vol. 14286. Lecture Notes in Computer Science. Springer, pp. 62–77. DOI: 10.1007/978-3-031-44505-7_5.
- Keskin, N. Bora and Assaf Zeevi (2014). “Dynamic Pricing with an Unknown Demand Model: Asymptotically Optimal Semi-Myopic Policies”. In: *Operations Research* 62.5, pp. 1142–1167.
- McAfee, R. Preston, John McMillan, and Michael D. Whinston (1989). “Multiproduct Monopoly, Commodity Bundling, and Correlation of Values”. In: *The Quarterly Journal of Economics* 104.2, pp. 371–383.
- Misra, Kanishka, Eric M. Schwartz, and Jacob Abernethy (2019). “Dynamic Online Pricing with Incomplete Information Using Multiarmed Bandit Experiments”. In: *Marketing Science* 38.2, pp. 226–252.
- Naderivandi, Khosroparviz, Ezgi Karabulut, and Burak Gökgür (2025). “Optimizing Pricing Strategies through Learning the Market Structure”. In: Available at SSRN. DOI: 10.2139/ssrn.5910343. URL: <https://ssrn.com/abstract=5910343>.
- Pošík, Petr and Jiří Kubalík (2012). “Experimental comparison of six population-based algorithms for continuous black box optimization”. In: *Evolutionary computation* 20.4, pp. 483–508.
- Ringbeck, Daniel and Arnd Huchzermeier (2019). “Dynamic Pricing and Learning: An Application of Gaussian Process Regression”. In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.3406293.
- Schmalensee, Richard (1984). “Gaussian Demand and Commodity Bundling”. In: *The Journal of Business* 57.1, S211–S230.
- Şen, Alper and Alex X. Zhang (2009). “Style goods pricing with demand learning”. In: *European Journal of Operational Research* 196.3, pp. 1058–1075.
- Shahriari, Bobak et al. (2016). “Taking the Human Out of the Loop: A Review of Bayesian Optimization”. In: *Proceedings of the IEEE* 104.1, pp. 148–175.
- Srinivas, Niranjan et al. (2010). “Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design”. In: *Proceedings of the 27th International Conference on Machine Learning (ICML)*. Omnipress, pp. 1015–1022.
- Stremersch, Stefan and Gerard J. Tellis (2002). “Strategic Bundling of Products and Prices: A New Synthesis for Marketing”. In: *Journal of Marketing* 66.1, pp. 55–72.
- Target (2026). *Grocery Value Bundles: Condiments and Sauces*. <https://www.target.com/c/grocery-value-bundles/condiments-and-sauces/-/N-cqgcdZ2a03n>. Accessed August 25, 2026.
- Tehrani, Pouya, Yixuan Zhai, and Qing Zhao (2012). *Dynamic Pricing under Finite Space Demand Uncertainty: A Multi-Armed Bandit with Dependent Arms*. eprint: 1206.5345. URL: <https://arxiv.org/abs/1206.5345>.

- Tesco (2026). *Lunch Meal Deals*. <https://www.tesco.com/shop/en-GB/browse/fresh-food/chilled-soup-sandwiches-and-salad-pots/lunch-meal-deals>. Accessed August 25, 2026.
- Venkatesh, R. and Wagner Kamakura (2003). “Optimal Bundling and Pricing under a Monopoly: Contrasting Complements and Substitutes from Independently Valued Products”. In: *The Journal of Business* 76.2, pp. 211–231.
- Wang, Jue (2021). “Optimal Bayesian Demand Learning over Short Horizons”. In: *Production and Operations Management* 30.4, pp. 1154–1177.
- Xiao, Baichun and Wei Yang (2024). “A Bayesian learning and pricing model with multiple unknown demand parameters”. In: *Annals of Operations Research* 343.1, pp. 493–513.
- Yang, Xiangyu et al. (2024). “Nonparametric multi-product dynamic pricing with demand learning via simultaneous price perturbation”. In: *European Journal of Operational Research* 319.1, pp. 191–205.
- Zhang, Guoquan, Guohao Li, and Jennifer Shang (2023). “Optimizing mixed bundle pricing strategy: Advance selling and consumer regret”. In: *Omega* 115, p. 102782.