Loss-Leader Pricing Strategies for Personalized Bundles Under Customer Choice

Zhengliang Xue
IBM T.J Watson Research Center
Yorktown Heights, NY 10598
zxue@us.ibm.com

Shivaram Subramanian
IBM T.J Watson Research Center
Yorktown Heights, NY 10598
subshiva@us.ibm.com

Markus Ettl
IBM T.J Watson Research Center
Yorktown Heights, NY 10598
msettl@us.ibm.com

Abstract

This paper considers the pricing of multi-product request-for-quotes (RFQs) that are configured by a buyer based on a large number of products or services offered in a seller's product catalog. The buyer submits an RFQ for a desired bundle of line items in a bid configuration to a seller. The seller reviews the configuration and offers an approved price for each line item in the bundle. The buyer can selectively purchase any combination of products or services in the bid configuration at the seller's approved prices. In addition to the line item pricing approach, we propose a novel loss-leader model that uses machine learning to calibrate the buyer's preferences among correlated line items, and dynamically optimizes the prices of any configuration to maximize the seller's expected profit. The pricing strategies were implemented in a business-to-business (B2B) sales environment with a multinational technology company. Counterfactual analysis shows that loss-leader pricing can generate more than ten percent lift in gross profit over existing pricing practices.

Keywords: Bundling, loss-leader pricing, counterfactual analysis, data sparsity, win probability

1. Introduction

We analyze pricing strategies for request-for-quotes (RFQs) for multi-product bid configurations, where a buyer submits an RFQ for a desired personalized bundle of products or services in the bid configuration to a seller. The seller reviews the bid configuration and offers a system-approved price for each item in the bid configuration to the buyer. The buyer then decides whether to purchase one or more line items in the configuration at the price offered by the seller. The process of determining an optimized price based on estimated valuations and win probability assessments is shown in Figure 1.

Artificial intelligence (AI) based pricing in an environment of personalized bundles that are dynamically configured from a large variety of products or services presents challenges not found in single commodity purchasing environments. This is not only because the correlations of products within a bundle are highly complex, but also because the configurations do not have enough transaction repeats.

In an analysis conducted with a large technology corporation, which provides customers with the option to tailor a comprehensive IT solution from a vast array of tens of thousands of products and services, it was observed that the overwhelming majority of chosen bundle configurations were distinct and one-of-a-kind. Analysis of historical sales data also showed the many bid configurations have a core product that plays a leading role in a bid configuration. Such products can be viewed as "loss-leaders" where a discounted price can significantly boost the buyer's overall propensity to purchase. As opposed to complete bid execution that requires the buyer to purchase all line items in a bid configuration ("all-or-nothing"), we analyze partial, or unconstrained, bid execution that allows the buyer to selectively purchase any number of line items from the bid configuration at the approved price. The additional flexibility offered with partial bid execution allows customers to selectively purchase line items in the bid configuration that are offered at the most attractive price, a behavior called "cherry-picking", that may lead to an actual decline in the seller's revenue or profit. We show that the loss-leader pricing strategy proposed in this paper increases the seller's expected revenue/profit,

URI: https://hdl.handle.net/10125/106569
978-0-9981331-7-1

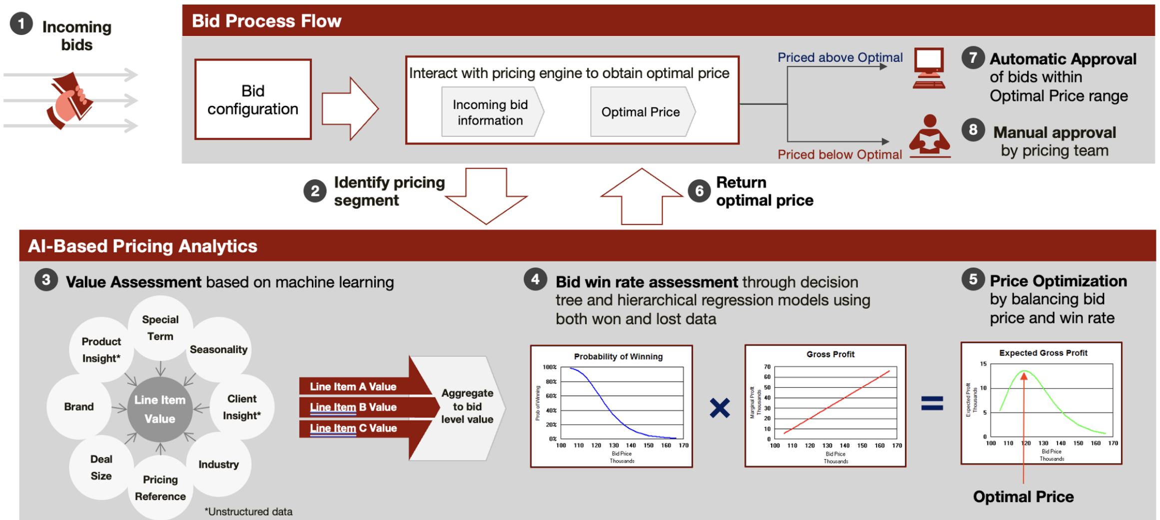

Figure 1. Process for determining the optimal price in response to a Request-For-Quote (RFQ).

all while anticipating the customer's selective choices.

In this paper, we propose a general method to identify loss-leader products in the aforementioned RFQ setting that drive the customers' purchase decisions, estimate the self-price and cross-price elasticity through machine learning, and effectively optimize the seller's pricing strategy that maximizes expected profit or revenue. We develop win probability estimation models that statistically compute the probability of winning a bid configuration at a given set of line item prices, and profit optimization models that compute the optimal pricing strategy by balancing the likelihood of winning each configuration at the prices of line items with the profitability of selling each product at the given price.

The remainder of the paper is organized as follows: Section 2 provides a review of the prior literature on bundle pricing and delineates our contribution to the existing literature. Section 3 analyzes the buyers' purchase behavior of multi-product RFQs, and presents a line-item pricing model to study their response to a seller's approved bid prices. More sophisticated strategies that consider product correlations are studied in Section 4. Given the tremendous computational challenges of analyzing all possible demand correlations among a large number of line items, we propose a loss-leader model which is computationally efficient and highly scaleable. Section 5 presents a case study based on actual sales data collected from a large technology corporation. Counterfactual comparisons between loss-leader pricing and incumbent pricing practices shows a significant profit increase from the loss leader pricing strategy. Our findings are summarized in Section 6.

2. Literature Review

Bundling, a strategy of pricing and selling product combinations, has long been a topic of research in the literature of economics, marketing, and operations management. Typically, there are three types of bundling strategies that offer different degrees of purchase flexibility: seller-configured bundles, customer-configured bundles (or personalized bundles) with complete execution, and customer-configured bundles with unconstrained execution.

A focus of previous research on seller-configured bundling strategies was to identify consumer purchasing behavior and cost structures under which bundling is profitable. It was shown the driving factors for bundling include complementarity and substitutability of products, heterogeneity of consumers and marginal cost. Prominent works in this literature include Adams and Yellen (1976), McAfee et al. (1989), Hanson and Martin (1999), Salinger (1995), Bakos and Brynjolfsson (1999), Bakos and Brynjolfsson (2000), Venkatesh and Kamakura (2003) and Stremersch and Tellis (2002). Early case studies and implementations of bundling were discussed in Eisenhardt (1989), Garfinkel et al. (2006), Schoenherr and Mabert (2006), Ozkul et al. (2012) and Li et al. (2015).

Computational complexity is a major concern while applying these traditional bundling models to price a large number of products for bundling in near real-time. The computational time increases exponentially with the number of combinations, particularly when the objective function of profit or revenue maximization is not jointly concave in their prices. Several papers have analyzed the performance of simple bundle policies to address this challenge. Chu et al. (2011) showed that setting prices that depend only on the size of bundle purchased tends to be more profitable than offering the individual products priced separately and tends to closely approximate the profits from mixed bundling. Ma and Simchi-Levi (2016) studied how to make bundling a more attractive strategy than selling individual items when the items have high production costs in a static model. They proposed a new mechanism that sells all items in a single bundle, but allows the customer to return any subset of items for a refund equal to their total production cost. In contrast, our model offers complete flexibility to customers in the final execution by allowing a customer to selectively purchase any subset of items from the original configuration request. Abdallah et al. (2017) considered a simple policy that prices a bundle based on the size rather than the different possible combinations of bundles. Abdallah (2018) analyzed situations where a simple pure bundling mechanism in the presence of non-negative marginal costs and correlated valuations is preferable to more complicated mixed bundling approached. Song and Xue (2021) analyzed the bundling strategy for vertically differentiated bundles and showed that the product inventory status and supply chain agility also drive the seller's bundling strategy.

The literature on personalized bundle pricing is relatively scarce. The challenge with personalization is that historical data for any given configuration is sparse which limits the effectiveness of machine learning in directly predicting customer preference or market responses for a given bid configuration. Hitt and Chen (1999) proposed a customized bundle pricing scheme where customers can select a fixed number of goods for a certain fixed price and compare it to other traditional bundling approaches. Their focus is on bundle design as their pricing is not really personalized to each bundle configuration. Several subsequent works take a similar approach, e.g., Wu et al. (2008), Jiang et al. (2011), Basu and Vitharana (2011). In more recent practice, configuration recommendation systems are designed based on machine learning to meet customer's demand for personalization (Chowdhary et al., 2021).

Lastly, Xue et al. (2016) and Chu et al. (2021) used a top-down and bottom-up method to analyze the customer response to the bundle price of personalized configurations based on sales data including over a thousand products. The top-down method decomposes the configurations into products and estimates the market value of each item. These market values are aggregated bottom-up to characterize the features of each configuration. The procedure allows to cluster the configurations into different segments and estimate the customers' purchase probability in each segment based on the bundle price and configuration features, and the optimal bundle price is calculated to maximize the seller's expect profit or revenue. However, this method requires complete execution of the entire bundle.

To the best of our knowledge, this paper is the first to propose a real-time strategy for loss-leader pricing, allowing customers the flexibility to selectively purchase products or services from their 'wish list' at seller's approved prices. The proposed approach optimizes pricing for individual line items by considering the relationships between the loss-leading items and the supplementary products or services within the bid configuration. Simultaneously, it anticipates the customer's selective preferences and choices.

3. Customer Choice and Bid Execution

Consider an RFQ for a bid configuration consisting of a core product or service and a set of secondary products or services under an unconstrained execution policy. We define a purchase option as a subset of line items that can be purchased by the buyer from a given configuration denoted by $\mathcal{C}$.

We use discrete choice models to estimate the win probability of a product, which is also the likelihood of the buyer purchasing the product. We develop a set of extended logit models based on the buyer utilities for various purchase options, including the no-purchase option. We model the buyer's utility for a purchase option, denoted by $\mathcal{S} \subseteq \mathcal{C}$, as a function of the following attributes: approved price (p_j), list price ($\bar{p}_j$) and manufacturing or service cost (c_j) for each line item j in the configuration $\mathcal{C}$. The 'competitive advantage' of a line item can be defined as $a_j = \bar{p}_j / c_j$. Another useful derived attribute of a line item is its contribution ratio, defined as $w_j = c_j / C_{\mathcal{C}}$, or the weight of cost c_j relative to the total configuration cost $C_{\mathcal{C}} = \sum_{j \in \mathcal{C}} c_j$.

Next, we provide the notations for the consumer choice model. The seller offers J line items to customers, indexed by $j \in \mathcal{J} = \{1, ...J\}$. Here all vectors, expressed by boldface letters, are column vectors with the same dimension as the corresponding product set, for instance $\mathbf{v}_{\mathcal{J}} = \{v_j\}_{j \in \mathcal{J}}$. We use $\mathbf{v}_{\mathcal{S}}$ to denote a subvector of $\mathbf{v}_{\mathcal{J}}$ whose component indices are restricted to the subset $\mathcal{S}$, for any $\mathcal{S} \subseteq \mathcal{J}$. The seller offers the price $\mathbf{p}_{\mathcal{C}}$ for a configuration $\mathcal{C} \subseteq \mathcal{J}$. Let $q_j(\mathbf{p}_{\mathcal{C}})$ be the seller's win probability with respect to product j in configuration $\mathcal{C}$, which is the likelihood of a customer purchasing product j at the price of $\mathbf{p}_{\mathcal{C}}$, for any product $j \in \mathcal{C}$. These win probabilities are estimated based on p_j, $\bar{p}_j$, c_j, w_j, a_j, etc., for any $j \in \mathcal{C}$.

In the following, we describe three different choice models that characterize the customer's purchase behavior. We first describe a Line Item Choice (LIC) model that ignores cross-item price correlations and calibrates the win probability of each line item based on its own-price effect. Next, we present a Whole Bundle Choice (WBC) model and Loss Leader Choice (LLC) model. Both models take into account correlations between line items, and demonstrate how to estimate their cross-price effects in a bid configuration. To learn the consumer choice behavior, we formulate a buyer utility function and calibrate the purchase probability based on the utility. We use logit models as an example to demonstrate the win probability estimation and price optimization. The methods can be extended to other discrete choice models.

3.1. Line Item Choice Model (LIC)

Here we assume that the execution of each line item is independent of other line items in the bid configuration with utility function $U_j = \Theta_j - \gamma_{j1} p_j + \varepsilon_j$, for any customer buying product j. Here, Θ_j summarizes the utility from all non-pricing related factors, for example, $\Theta_j = \gamma_{j2} c_j + \gamma_{j3}(\bar{p}_j - c_j)$. Here, we assume that the product value is positively correlated to the manufacturing cost c_j and list price $\bar{p}_j$. The own-price effect is measured by γ_{j1}. Customer heterogeneity is characterized by a random factor ε_j, representing a buyer's deviation from the average perceived value.

In practice, it is difficult to estimate $(\gamma_{j1}, \gamma_{j2}, ...)$, if product j has sparse sales history. We therefore group line items with similar characteristics to obtain sufficient training data by clustering them based on their features (e.g., server, memory, processor, hard drive, on-site support, warranty etc.), and competitive advantage (ratio of list price to cost a_j), among others. With respect to the customer response, further segmentation is applied based on factors such as RFQ size (total configuration revenue), customer incumbency (e.g., acquisition, development, or retention account), and business sector (financial services, automotove, government, etc.). We assume that the coefficients $(\gamma_1, \gamma_2, ...)$ are the same over all products within a given segment.

To consistently measure the utility of different products in a segment, we normalize the utility function U_j by the cost c_j, and have

$$u_j = \theta_j - \beta_1 z_j + \epsilon_j = u_j + \epsilon_j. \quad (1)$$

Here, $z_j = (p_j - c_j)/c_j$ is the profit margin based on cost. $\theta_j = \beta_2 + \beta_3 a_j$ is the normalized value of the product, measuring the value created from each dollar input to production. Intuitively, we can define high-end products to be those with the highest θ_j (or a_j). Such normalizations allow us to group sales data for similar line items into clusters, because it measures their attributes based on a percentage value instead of an absolute value. This significantly mitigates the challenges associated with data sparsity in machine learning. Normalization can also be applied to other attributes such as list price $\bar{p}_j$ when the variable cost c_j is equal to zero or is not well defined as is often the case in the services sector.

Compared to other non-linear classifying machine learning models such as the probit model that assumes normal distributed errors, the logit model shows minor difference in the probability estimation but significant advantage in the price optimization, e.g., Agrawal and Ferguson (2007). Hence, we also use the logit model to study the customer choice based on the normalized utility function as in (1).

Assuming ϵ_j satisfies the i.i.d. doubly exponential distribution in each segment, a customer will purchase product j with probability $q_j = \frac{\exp(u_j)}{\exp(0)+\exp(u_j)}$, where the no-purchase option has zero utility. The coefficients $(\beta_1, \beta_2, \beta_3)$ corresponding to (z_j, w_j, a_j) can be obtained as outputs of the logistic regression. Now, the coefficients β can also be calibrated for a group of products compared to the γ_j's defined for each line item. Therefore, it enables the regression model to use a hierarchical classification tree to splits the data set into clusters at different tiers (e.g., from product category to product family to line item) consequently alleviating the issue of data sparsity.

3.2. Whole Bundle Choice Model (WBC)

Next, we analyze the choice models for unconstrained execution in the presence of cross-product price correlations. Consider a buyer's utility for a configuration $\mathcal{C}$, at the seller's approved prices $\mathbf{p}_{\mathcal{C}}$. If the whole configuration $\mathcal{C}$ has been executed, it provides a net utility $U_{\mathcal{C}} = \Theta_{\mathcal{C}} - \gamma_{\mathcal{C}1} P_{\mathcal{C}} + \varepsilon_{\mathcal{C}}$, with $\Theta_{\mathcal{C}} = \sum_{j \in \mathcal{C}} \Theta_j$.

Under complete execution, Xue et al. (2016) proposed an approach to analyze the win probability of such personalized bundles at a total price $P_{\mathcal{C}} = \sum_{j \in \mathcal{C}} p_j$. Under unconstrained execution, however, the buyer might purchase any subset of the whole configuration, $\mathcal{S} \subseteq \mathcal{C}$, at the approved line item prices ($\mathbf{p}_{\mathcal{C}}$) rather than a single price ($P_{\mathcal{C}}$) of the entire bundle. Ideally, we can evaluate the utility of each subset, and estimate the corresponding coefficients such as $\Theta_{\mathcal{S}}$ and $\gamma_{\mathcal{S}1}$. These coefficients depend on the correlation between line items, for instance $\Theta_{\mathcal{S}} = g\big(\{\Theta_j\}_{j\in\mathcal{C}}\big)$.

With respect to the configurations that include a core product, we impose the following assumption on the correlation between the core product and secondary products or services:

Assumption 1 *For any configuration that consists of a core product $k \in \mathcal{K}$ and secondary items $\mathcal{L}_k$, consider any combination $\mathcal{S} \subseteq \mathcal{C}$. We have $\Theta_{\mathcal{S}} = -\infty$, if k is not in $\mathcal{S}$, and $\Theta_{\mathcal{S}} = \sum_{j\in\mathcal{S}} \Theta_j$, if $k \in \mathcal{S}$.*

I.e., we assume a strongly negative utility for purchasing any subset that does not include a core product.

To enable the analysis on a variety of product combinations, we normalize the utility function by the total cost $C_{\mathcal{C}} = \sum_{j\in\mathcal{C}} c_j$. For any subset $\mathcal{S}$ including the core product k, we have the following utility function:

$$u_{\mathcal{S}} = \theta_{\mathcal{S}} - \beta_1 \sum_{j\in\mathcal{S}} (p_j - c_j)/C_{\mathcal{C}} + \epsilon_{\mathcal{S}}, \text{ for } k \in \mathcal{S}. \quad (2)$$

Here, $\theta_{\mathcal{S}} = \beta_2 w_{\mathcal{S}} + \beta_3 w_{\mathcal{S}} a_{\mathcal{S}}$, with $C_{\mathcal{S}} = \sum_{j\in\mathcal{S}} c_j$, $w_{\mathcal{S}} = C_{\mathcal{S}}/C_{\mathcal{C}}$ and $a_{\mathcal{S}} = \bar{P}_{\mathcal{S}}/c_{\mathcal{S}}$ respectively.

Compared to the LIC model where the win/loss of a line item is explicitly labeled, the win/loss of a configuration cannot be simply defined in a binary manner for RFQs where line products were selectively purchased. There are two possible ways to label the win/loss for a partially executed configuration.

Firstly, we can consider labeling a configuration $\mathcal{C}$ as a win if any subset $\mathcal{S} \subseteq \mathcal{C}$ of the original configuration was eventually purchased by the customer. Accordingly, label a configuration as a loss if no line item has been purchased. Then we can use the method proposed by Xue et al. (2016) and Chu et al. (2021) to estimate the seller's win probability. However, we may overestimate the chance to win by ignoring the non-purchased items $j \in \mathcal{C}/\mathcal{S}$ in any partially executed RFQ. Moreover, this approach can only optimize the total price of entire configuration and not the price distribution for each line item.

Secondly, we can consider all possible subsets of each configuration, in which each meaningful combination of products is seen as a 'purchase option'. In this case, each purchase option has plenty of sales data. However, such method will encounter the curse of dimensionality when the number of line products increases. Let $\mathcal{F}$ be the set of all the possible product combinations corresponding to the configuration $\mathcal{C}$. Let $\mathcal{F}_{\mathcal{S}}$ be the set of all the possible product combinations that contain $\mathcal{S}$, for any subset $\mathcal{S} \subseteq \mathcal{C}$. Particularly, let $\mathcal{F}_j = \mathcal{F}_{\{j\}}$ be the set of all the possible product combinations that contain product j. Then we have $\mathcal{F}_{\{k,l\}} = \mathcal{F}_k \cap \mathcal{F}_l$. For example, consider the configuration $\mathcal{C} = \{1, 2, 3\}$. We have $\mathcal{F} = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$, $\mathcal{F}_1 = \{\{1\}, \{1, 2\}, \{1, 3\}, \{1, 2, 3\}\}$, and $\mathcal{F}_{\{1,2\}} = \{\{1, 2\}, \{1, 2, 3\}\}$.

Now we use these notations to analyze a configuration that contains a core product indexed by k and multiple accessories indexed by $l \in \mathcal{L}$, and have the win probability of product k:

$$q_k = Pr(k \in \mathcal{S}) = \frac{\sum_{\mathcal{S}\in\mathcal{F}_k} \exp(u_{\mathcal{S}})}{\exp(0) + \sum_{\mathcal{S}\in\mathcal{F}_k} \exp(u_{\mathcal{S}})}.$$

Also, we have the joint probability of purchasing core product k together with some secondary product l:

$$q_{kl} = Pr(k, l \in \mathcal{S}) = \frac{\sum_{\mathcal{S}\in\mathcal{F}_k\cap\mathcal{F}_l} \exp(u_{\mathcal{S}})}{\exp(0) + \sum_{\mathcal{S}\in\mathcal{F}_k} \exp(u_{\mathcal{S}})}.$$

Similarly, we can define the win probability for all product combinations.

A logit model works well when customers are fully aware of each combination in their choice set. Also, it is well-known that the IIA (independence of irrelevant alternatives) assumption of the logit model is difficult to satisfy in practice. Alternatives like the nested-logit or mixed logit model can help in this regard but unfortunately tend to increase the computational complexity of both machine learning and price optimization for large bids. Considering this, the logit model was an acceptable tradeoff in the RFQ pricing context.

Next, we propose a new method to overcome the data sparsity in the analysis of product correlations, especially for configurations that consist of one or more core products and many secondary products or services.

3.3. Loss Leader Choice Model (LLC)

An LIC model ignores the price correlation among products, whereas a WBC model exhaustively considers all the possible correlations of product prices in each configuration. The former has the least computational challenges, but may yield sub-optimal solutions. On the other hand, the latter would theoretically provide optimal solutions, but any real-time implementation is strongly impacted by data availability and computational complexity. In this section, we propose a new loss-leader model that detects the most important pairwise correlations between a core product and accessories.

Consider a basic configuration composed of a core product, e.g., server, indexed by k and multiple secondary products, or accessories, indexed by $l \in \mathcal{L}$.

In this setting it is reasonable to assume that the buyer would not purchase the accessories in the absence of the server, because their add-on value is completely tied to the server purchase. Corresponding to Assumption 1, we have

Assumption 2 *Consumers will not buy any accessory unless they buy the server in the configuration, i.e.,* $Prob(l \in \mathcal{S}|k \notin \mathcal{S}) = 0$.

Any subset $\mathcal{S} \neq \emptyset$ of the initial configuration is purchased if and only if the server is included, $k \in \mathcal{S}$. The win probability $Pr(\mathcal{S} \neq \emptyset) = q_k$. Let $q_{j|k} = Pr(l \in \mathcal{S}|k \in \mathcal{S})$ be the conditional probability of purchasing accessory j given that the customer purchased the server. We use a set of extended logit models to calibrate the win probabilities. Again, using the normalized price, or profit margin, z_j as the explanatory variable we have

$$\begin{aligned} q_k(z_k, \mathbf{z}_{\mathcal{L}}) &= \frac{\exp(\theta_k - \beta_k z_k - \sum_{j\in\mathcal{L}} \beta_{kj} z_j)}{1+\exp(\theta_k - \beta_k z_k - \sum_{j\in\mathcal{L}} \beta_{kj} z_j)}, \\ q_{lk}(z_k, \mathbf{z}_{\mathcal{L}}) &= \frac{\exp(\theta_l - \beta_{lk} z_k - \sum_{j\in\mathcal{L}} \beta_{lj} z_j)}{1+\exp(\theta_l - \beta_{lk} z_k - \sum_{j\in\mathcal{L}} \beta_{lj} z_j)}. \end{aligned}$$

The computational complexity heavily depends on the number of cross-price effects characterized in q_k and q_{lk}. The following assumption effectively simplifies the win-probability computations:

Assumption 3 *Assume*

H_1 *: q_k depends on its own profit margin z_k, but not on the profit margin of its accessory, $\mathbf{z}_{\mathcal{L}}$.*

H_2 *: q_{lk} depends on its own profit margin z_l and the server's margin z_k, but not on the profit margin of the other accessory $\mathbf{z}_{\mathcal{L}\setminus\{l\}}$.*

If H_1 does not hold, the win probability of the server depends on $(z_k, \mathbf{z}_{\mathcal{J}})$. If H_2 does not hold, the win probability of the accessory depends on the profit margin of the other accessories. In the following, we explore pricing optimization under hypotheses H_1 and H_2.

4. Optimal Pricing Strategy

First, we consider a line item pricing strategy (LIP) focused on individual products where cross-price effects in the customer's choice are ignored. Then, a whole bundle pricing (WBP) strategy is proposed to include all possible product combinations and price correlations. However, the computational time rapidly increases for large-scale bid configurations. Thus, a loss leader pricing (LLP) strategy, with favorable computational efficiency, is applied to configurations that consist of a core product and several accessories.

To develop an explicit formulation, we use the logistic model to illustrate the pricing optimization. For ease of notation, we drop the subscript "$\mathcal{C}$" in this section, since all the variables and coefficients are defined in a configuration for the pricing optimization of all products. Let G be the expected profit and R be the expected revenue of a configuration $\mathcal{C}$. We have

$$\begin{aligned} &\max_{\mathbf{p}} G(\mathbf{p}) = \sum_{j\in\mathcal{C}} (p_j - c_j) q_j(\mathbf{p}) \qquad (3) \\ \text{s.t.} \quad & R(\mathbf{p}) = \mathbf{p}\cdot\mathbf{q}(\mathbf{p}) \geq \underline{R}, \\ & \bar{\mathbf{p}} \geq \mathbf{p} \geq \underline{\mathbf{p}} \end{aligned}$$

Here $\underline{R} \leq \max_{\mathbf{p}} R(\mathbf{p})$ must be a meaningful lower bound on the expected revenue target. Moreover, there can be some bounds on the pricing decisions. Typically, we have the list price as the upper-bound and the cost as the lower-bound.

4.1. Line Item Pricing (LIP)

Obtaining the optimal price based on the line item choice model is a straightforward solution to any configuration where the cross-price effects have been ignored among all products. Thus, we have $q_j(\mathbf{p}) = q_j(p_j)$ as in (3), which means $G(\mathbf{p}) = \sum_{j\in\mathcal{C}}(p_j - c_j)q_j(p_j) = \sum_{j\in\mathcal{C}} G_j(p_j)$. To this end, maximizing the total expected profit of a configuration is equivalent to maximizing the expected profit of each line item respectively, based on its self-price effect estimated in the LIC model.

To solve the constrained optimization problem, we employ a variable transformation, using the profit margin z_j as the explanatory variable in the win probability instead of the price p_j and using the purchase probability q_j as the decision variable instead of the price p_j. Since $q_j(z_j)$ is an decreasing function, it admits a monotone inverse function $z_j(q_j)$ that we use to revise the constrained maximization problem. Thus, we have $G_j(p_j) = G_j(z_j) = c_j z_j q_j(z_j) = c_j z_j(q_j) q_j = G_j(q_j)$, throughout the transformation. Such transformation is particularly useful when the win probability q_j is characterized by a logit model. We have

$$\begin{aligned} \max_{q_j} \quad & G_j(q_j) = c_j z_j(q_j) q_j && (4) \\ s.t. \quad & R_j(q_j) = c_j \cdot (z_j(q_j)+1)\cdot q_j \geq \underline{R}_j && (5) \\ & \bar{z}_j \geq z_j(q_j) \geq \underline{z}_j. \end{aligned}$$

Let $q_j^\circ = \arg\max_{q_j} G_j(q_j)$ and $\bar{q}_j^\circ = \arg\max_{q_j} R_j(q_j)$ be the unconstrained profit maximizer and revenue maximizer respectively. Note, we have $q_j^\circ \leq \bar{q}_j^\circ$, and $R_j(q_j)$ is increasing in $0 \leq q_j \leq \bar{q}_j^\circ$, which leads to an inverse function $Q_j(R_j)$.

Proposition 1 *The following properties hold for the revised optimization problem as in (4).*

(1) $G_j(q_j)$ is strictly concave in q_j. It admits a unique optimum solution $q_j^\circ = \arg\max_{q_j} G_j(q_j)$;

(2) In the presence of the revenue constraint, the optimal solution $q_j^* = \max(q_j^\circ, Q_j(\underline{R}_j))$;

(3) In the presence of the price bounds, the optimal solution $q_j^* = \min(q_j(\underline{z}_j), \max(q_j^\circ, q_j(\bar{z}_j))$.

Proposition 1 shows that the logit model has attractive properties that support real-world industrial implementation. A similar optimization is applicable to the probit model and other generalized choice models, although a closed-form solution may not exist.

Accordingly, the computational complexity of pricing optimization is linearly increasing in the number of line items, denoted by $O(n)$. To this end, the LIP strategy can be used as a benchmark as well as a starting point for finding more sophisticated pricing solutions.

4.2. Whole Bundle Pricing (WBP)

A whole bundle pricing model requires a large volume of data that contains sufficient data samples for each possible combination of products, which is typically not satisfied in practice for a large product assortment. Furthermore, it needs to estimate and optimize for all possible demand interactions among n, resulting in an exponential complexity of computation, $O(e^n)$. Given these limitations, WBP is ill-suited for our RFQ application that requires near real-time pricing for an incoming RFQ.

4.3. Loss Leader Pricing (LLP)

Unlike WBP, the loss leader model does not significantly increase the computational time and has a polynomial complexity, $O(n^c)$ to price n line items by focusing on the most important product correlations in the configuration. Thus it can price large configurations selected from thousands of product catalog offerings.

The business application shown in Figure 1 requires an RFQ to be priced on-the-fly, i.e., within 200-300ms. To satisfy this practical requirement, we consider the win probability expressed in the form of $q_k(z_k)$ and $q_{lk}(z_k, z_l)$ under H_1 and H_2 and optimize the product prices of each configuration as follows:

$$G^*(z_k, \mathbf{z}_\mathcal{L}) = \max_{z_k} \left(c_k z_k q_k + \max_{\mathbf{z}_\mathcal{L}} \sum_{l \in \mathcal{L}} (c_l z_l q_{lk}) \right) \quad (6)$$

$$q_k(z_k) = \frac{\exp(\theta_k - \beta_k z_k)}{1 + \exp(\theta_k - \beta_k z_k)},$$

$$q_{lk}(z_k, z_l) = \frac{\exp(\theta_l - \beta_{lk} z_k - \beta_{ll} z_l)}{1 + \exp(\theta_l - \beta_{lk} z_k - \beta_{ll} z_l)}, \forall l \in \mathcal{L}.$$

With regard to q_{lk}, it is common to assume the self-price effect dominates the cross-price effect.

Assumption 4 *We have $|\beta_{kk}| \geq |\beta_{lk}|$.*

Under this assumption, the following properties hold for the total expected profit of each configuration.

Proposition 2 *Under Assumption 4, we have:*

(1) The win probability functions are convertible to inverse functions, where we have a unique profit margin function, $z_k(q_k)$ and $z_l(q_k, q_{lk})$, for $l \in \mathcal{L}$.

(2) The total expected profit is jointly strictly concave in q_k and q_{lk}, for $\forall l \in \mathcal{L}$.

Accordingly, the profit margins can be expressed as the function of win probabilities.

$$z_k = \left(\theta_k - \ln(\frac{q_k}{1 - q_k}) \right) / \beta_{kk},$$

$$\mathbf{z}_\mathcal{L} = \mathrm{B}^{-1} \left[\theta_\mathcal{L} - \beta_{\mathcal{L},k} z_k - \ln(\frac{q_{lk}}{1 - q_{lk}})_{l \in \mathcal{L}} \right].$$

As long as Assumption 4 holds, we can use $(q_k, \mathbf{q}_\mathcal{L})$ as the decision variables instead of profit margins $(z_k, \mathbf{z}_\mathcal{L})$. Here B^{-1} is the inverse matrix of correlation coefficients of all the accessories in each configuration. Although we are able to explicitly characterize the objective function, it still takes a considerable time to compute the optimal prices for thousands of line items. The pricing optimization can be further simplified under additional assumptions.

Assume the conditional win probability $q_{l|k}(z_l)$ of accessory j given the configuration win only depends on its own price, then we have $q_{lk}(z_k, z_l) = q_k(z_k) \times q_{l|k}(z_l)$ and the total expected profit function (6) can be further simplified as follows:

$$G^* = \max_{z_k} \left(c_k z_k + \sum_{l \in \mathcal{L}} \max_{z_l} (c_l z_l q_{l|k}) \right) q_k.$$

Also we assume the cross-price effect of the server on the accessory l is much stronger than that of any of the other accessories, that is $\beta_{kl} \gg \beta_{jl}$, and $\beta_{ll} \gg \beta_{jl}$, for any $j \in \mathcal{C} \setminus \{k\}$. Accordingly, we formulate the profit margin as the function of win probabilities.

$$z_k(q_k) = \left(\theta_k - \ln(\frac{q_k}{1-q_k})\right)/\beta_{kk}, \tag{7}$$

$$z_l(q_{lk}, q_k) = \left[\theta_l - \beta_{lk} z_k - \ln(\frac{q_{lk}}{1-q_{lk}})\right]/\beta_{ll}. \tag{8}$$

The server's profit margin has a one-to-one mapping to its win probability as in (7), but it does not depend on any accessories. Given the purchase probability of a server q_k, the optimal price margin $z_l^*(q_k)$ can be calculated for each accessory l by optimizing q_{lk}. The corresponding win probability is characterized as $q_{lk}^*(q_k)$. Meanwhile, an accessory's profit margin (8) can be characterized by the server's profit margin (z_k) and a function of its own purchase probability (q_{lk}). The former is concave in q_k whereas the latter is concave in q_{lk}. We have the optimal $q_{lk}^*(q_k)$ monotone in q_k, then we have $z_l^*(q_k)$ concave in q_k, which is the win probability of a server regardless of any purchase of accessories. Therefore, the profit maximization can be solved efficiently under Assumptions H_1 and H_2. If these assumptions are relaxed, the computational complexity will increase accordingly.

5. Counterfactual Analysis

The pricing models presented in the previous sections were implemented for a large technology corporation and deployed for general business and large enterprise customers across several countries. The data generation process was fully automated and all the required information for pricing including win/loss outcomes, product, customer, and bid attributes were recorded for each RFQ. Additional features derived from the raw data were generated and added to the data. To illustrate the value that the business derived from the new pricing models, we present a case study based on 1,133 RFQs that were configured from more than 200 product catalog offerings. RFQ data was collected over a three-month period from direct (15%) and indirect sales channels (85%). The accessories included in the study consisted of memory, processors and hard drives.

An initial data exploration showed that 343 out of the 1,133 RFQs included in the data set involved the purchase of a core item (server). In instances where a server was purchased, the win rate of accessories reached 72.5%. In the 790 remaining cases where a server was not bought, the success rate for accessories was notably lower, standing at 8.1%. In other words, customers that purchased a server that was quoted in a bid configuration were 9 times more likely to also purchase accessories. Table 1 displays the win rates under specific conditions for memory, processors, and hard disk drives. It distinguishes between RFQs that included the acquisition of a server and those that did not.

Product Category	Server win	Server loss
Memory	73.0%	7.1%
Processor	62.0%	3.1%
Hard disk drive	77.7%	11.2%

Table 1. Conditional win probabilities of accessories.

Next, we compare the optimal prices recommended by the line-item pricing (LIP) and loss-leader pricing (LIP) models by defining a ratio p_j^{LIP}/p_j^{LLP} for line item j. A ratio greater than 1 denotes cases where the LIP-recommended price is higher than the LLP-recommended price.

Figure 2 depicts a histogram representing the price ratio for all RFQs in the data set, categorized by commodity. Under the LLP strategy, the loss-leader items (i.e., server in this instance) were priced 0.5% lower on average compared to the LIP strategy, in particular in configurations with a high nominal bid value. On the other hand, accessories were priced between .7% and 4% higher under the LLP strategy.

To evaluate the influence of the proposed pricing strategies on the achievable gross profit, we conduct a counterfactual analysis following the approach described in Xue et al. (2016) and Ye et al. (2018):

- Scenario 1: if the optimal price p_k^* is greater than or equal to the actual price p_k, and p_k resulted in a win, the optimal price would have led to a win with probability $\frac{q(p_k^*)-c_k)}{q_k(p_k)}$. This conditional probability measures the odds of winning at the optimal price p_k^*, given that the approved price resulted in a win. The incremental profit pertaining to this scenario is $\Delta G_1 = (p_k^* - c_k)\frac{q(p_k^*)}{q(p_k)} - (p_k - c_k)$.

- Scenario 2: if the optimal price p_k^* is less than p_k, and p_k led to a win, the optimal price is certainly accepted by the buyer. However, the incremental profit pertaining to this scenario is a net loss of $\Delta G_2 = p_k^* - p_k < 0$.

- Scenario 3: if the optimal price p_k^* is greater or equal p_k, and p_k yielded a loss, the higher optimal price would also have resulted in a loss and the

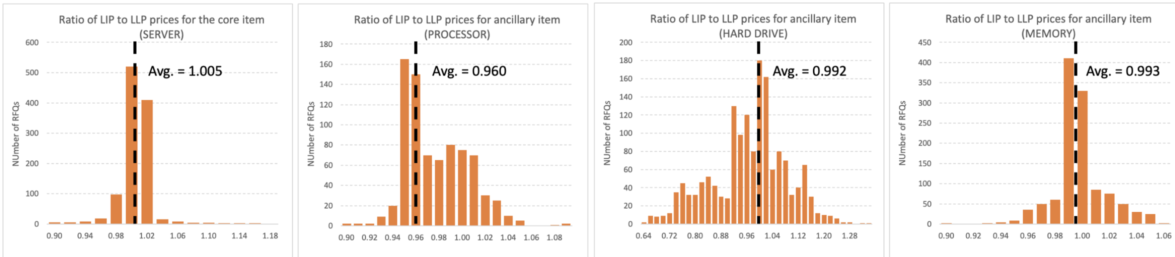

Figure 2. Frequency diagram of the ratio of LIP to LLP recommended prices for all commodities.

net profit gain pertaining to this scenario is zero, $\Delta G_3 = 0$.

- Scenario 4: if the optimal price p_k^* is less than p_k, and p_k resulted in a loss, the lower optimal price would have converted the loss to a win with conditional probability $1 - \frac{1-q_k(p_k^*)}{1-q_k(p_k)}$. The incremental profit pertaining to this scenario is $\Delta G_4 = \left(1 - \frac{1-q_k(p_k^*)}{1-q_k(p_k)}\right)(p_k^* - c_k)$.

With this, the total incremental gross profit improvement can be estimated as

$$\begin{aligned}\Delta G &= \left\{(p_k^* - c_k)\left[\frac{q_k(p_k^*)}{q_k(p_k)} - 1\right]^- + (p_k^* - p_k)\right\}\mathbf{1}_w \\ &+ \left\{\left[1 - \frac{1-q_k(p_k^*)}{1-q_k(p_k)}\right]^+ (p_k^* - c_k)\right\}(1 - \mathbf{1}_w).\end{aligned}$$

where $[x]^+ = \max(x, 0)$ and $[x]^- = -\min(x, 0)$. The indicator function $\mathbf{1}_w$ is the indicator of a win at the approved price. The findings of the counterfactual analysis are summarized in Figure 3.

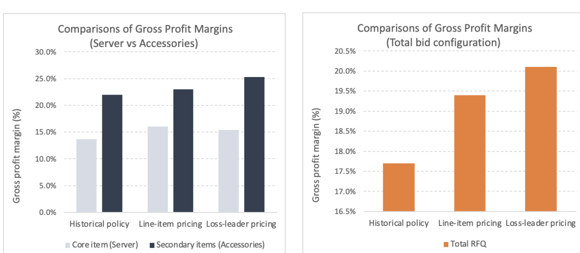

Figure 3. Counterfactual comparisons of gross profit margins.

For comparison, we report the historical gross profit that was measured from business-as-usual pricing practice as a benchmark. Upon consolidating the total gross profit gains across all product categories, we observe that the LIP strategy increases the gross profit margin from 17.7% as achieved by the incumbent pricing strategy to 19.4% for a relative lift of 9.5 percent. Under the LLP strategy, the gross profit margin increases even further to 20.1% for a relative lift of 13.6 percent over the benchmark.

6. Concluding Remarks

In this paper we described the challenges of the pricing problem for multi-product request-for-quotes (RFQs) and presented a novel loss-leader pricing model that was deployed in production at a multi-national technology corporation. The loss-leader model proves especially valuable in scenarios where customized bundles are tailored to individual buyer preferences, with a focus on a core product. Offline counterfactual evaluations showed that the proposed loss-leader pricing model performs significantly better than the incumbent pricing strategy at the company, and that it also outperforms a direct line-item pricing strategy.

We illustrated the use of advanced bundle pricing strategies in the context of manufacturing, but the concepts and benefits apply equally to the services sector. For instance, a bundle pricing strategy in consumer-focused industries that involves offering multiple products or services together as a package is widely used to attract and retain customers, increase sales, and maximize revenue. Customers with a perceived value proposition can purchase a combination of items at a lower price than if they were to buy each item individually. This encourages consumers to make larger purchases. Many businesses use bundle pricing during holidays or special events to encourage spending. Allowing customers to customize their bundles by choosing from a selection of products or services can further increase customer satisfaction and engagement. Pricing bundles strategically can also play on consumers' psychological tendencies. For example, offering three items in a bundle can make it seem like a better deal than two. A data-driven pricing approach as outlined in this paper allows companies to fine-tune their bundle offerings for maximum effectiveness.

References

Abdallah, T. (2018). On the benefit (or cost) of large-scale bundling. *Production Oper. Management*, (28(4)), 955–969.

Abdallah, T., Asadpour, A., & Reed, J. (2017). Large-scale bundle size pricing: A theoretical analysis. *Working paper, New York University, New York, NY*.

Adams, W., & Yellen, J. (1976). Commodity bundling and the burden of monopoly. *The Quarterly Journal of Economics*, (90(3)), 475–498.

Agrawal, V., & Ferguson, M. (2007). Bid-response models for customised pricing. *Journal of Revenue and Pricing Management*, (6), 212–228.

Bakos, J., & Brynjolfsson, E. (1999). Bundling information goods: Pricing, profit and efficiency. *Management Sci.*, (45(12)), 1613–1630.

Bakos, J., & Brynjolfsson, E. (2000). Bundling and competition on the internet. *Marketing Sci.*, (19(1)), 63–82.

Basu, A., & Vitharana, P. (2011). Impact of customer knowledge heterogeneity on bundling strategy. *Marketing Sci.*, (28(4)), 792–801.

Chowdhary, P., Ettl, M., Mukherjee, S., & Xue, Z. (2021). Product configuration recommendation and optimization. *Patent: US11100570B2*.

Chu, S., Jiang, H., Xue, Z., & Deng, X. (2021). Adaptive convex clustering of generalized linear models with application in purchase likelihood prediction. *Technometrics*, (63(2)), 171–183.

Chu, S., Leslie, P., & Sorensen, A. (2011). Bundle-size pricing as an approximation to mixed bundling. *Amer. Econom. Rev.*, (101(2)), 263–303.

Eisenhardt, K. (1989). Building theories from case study research. *The Academy of Management Review*, (14(4)), 532–550.

Garfinkel, R., Gopal, R., Tripathi, A., & Yin, F. (2006). Design of a shopbot and recommender system for bundle purchases. *Decision Support Systems*, (42(3)), 1974–1986.

Hanson, W., & Martin, R. (1999). Optimal bundling pricing. *Management Sci.*, (36(2)), 155–174.

Hitt, L., & Chen, P. (1999). Bundling with customer self-selection: A simple approach to bundling low-marginal-cost goods. *Management Sci.*, (51(10)), 1481–1493.

Jiang, Y., Shang, J., Kemerer, C., & Liu, Y. (2011). Optimizing e-tailer profits and customer savings: Pricing multistage customized online bundles. *Marketing Sci.*, (30(4)), 735–752.

Li, S., Sun, H., Yan, J., & Yu, J. (2015). Bundling decisions in procurement auctions with sequential tasks. *Journal of Public Economics*, (128), 96–106.

Ma, W., & Simchi-Levi, D. (2016). Reaping the benefits of bundling under high production costs. *Working paper, MIT, Cambridge, MA*.

McAfee, R., McMillan, J., & Whinston, M. (1989). Multiproduct monopoly, commodity bundling, and correlation of values. *The Quarterly Journal of Economics*, (104(2)), 371–383.

Ozkul, A., Yilmaz, H., & Barut, M. (2012). Commodity bundling in the supply chain: A simulation study. *International Journal of Integrated Supply Chain Management*, (7), 114–137.

Salinger, M. (1995). A graphical analysis of bundling. *The Journal of Business*, (68(1)), 85–98.

Schoenherr, T., & Mabert, V. (2006). Bundling for b2b procurement auctions: Current state and best practices. *International Journal of Integrated Supply Management*, (2(3)), 189–213.

Song, J.-S., & Xue, Z. (2021). Demand shaping through bundling and product configuration: A dynamic multiproduct inventory-pricing model. *Oper. Res.*, (69(2)), 525–544.

Stremersch, S., & Tellis, G. (2002). Strategic bundling of products and prices: A new synthesis for marketing. *Journal of Marketing*, (66(1)), 55–72.

Venkatesh, R., & Kamakura, W. (2003). Optimal bundling and pricing under a monopoly: Contrasting complements and substitutes from independently valued products. *Journal of Business*, (76(2)), 211–231.

Wu, S., Hitt, L., Chen, P., & Anandalingam, G. (2008). Customized bundle pricing for information goods: A nonlinear mixed-integer programming approach. *Management Sci.*, (54(3)), 608–622.

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

Ye, P., Qian, J., Chen, J., Wu, C.-h., Zhou, Y., De Mars, S., Yang, F., & Zhang, L. (2018). Customized regression model for airbnb dynamic pricing. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 932–940. https://doi.org/10.1145/3219819.3219830