Multi-product Pricing with Transaction Data

Ton Meesena, Mike Qu, Jiaqi Shi

December 20, 2024

Abstract

We study a monopolist's pricing problem for multiple substitutable products, aiming to maximize expected revenue. We present different loss functions using transaction data to prescribe prices, including the MNL likelihood loss, combined-revenue-likelihood loss, generalized hinge loss from [1], and the direction loss.

We then establish a revenue lower bound for a pricing policy using the direction loss without relying on specific choice models. Numerical experiments show that our method outperforms the standard estimate-then-optimize approach using the MNL likelihood and the model-free method in [2] across various settings.

1 Introduction

In an increasingly data-driven economy, firms' ability to leverage transaction data and develop effective pricing models can offer significant profitability and competitive advantages. In many cases, firms may offer products that are substitutable, each catering to a specific segmented group of customers, making it necessary to incorporate cross-product interactions such as cannibalization and demand spillover. These interactions complicate the pricing decision as the optimal price for one product can affect the optimal price for another. By modeling for these factors, firms can price product that maximize revenue and better understand their targeted customers

One widely-used approach for contextual pricing is Estimate-then-Optimize. Within this paradigm, a discrete choice model is used to characterize customer valuation of a given product through measures such as utility estimation and price sensitivity. This is achieved by closely aligning the model's estimate the probability of a customer purchasing the product at a specific price, with observed patterns in historical transaction data. Subsequently, a pricing policy can be obtained by optimizing to maximize revenue under the predicted choice probabilities. However, there are several potential issues that can arise. Firstly, the selected choice model may fail to adequately capture the patterns present in historical transaction data, leading to poor predictive performance and by extension suboptimal pricing policies. This is commonly referred to as model misspecification. Secondly, an overly complex choice model may introduce non-convexities during the downstream revenue optimization problem, increasing the likelihood of the optimization process being trapped in local minima and failing to identify the true optimum.

Other approaches [1] [2] do not rely on parametric models that describe customer behaviors. Rather, they use observed transaction data to directly optimize a surrogate loss function to approximate the revenue-maximizing objective. Also known as "model-free" approaches, they are more robust to misspecification as they do not rely on assumptions about the structure of the demand or the customers' behavior.

In this project, we propose a combined Estimate-then-Optimize MNL Loss, hinge loss, and a direction Loss. Empirically, our methods manage to outperform existing baselines in both well-specified and misspecified cases. We also provide theoretical guarantees for the direction loss function.

2 Model

Consider a monopolist facing a pricing problem for multiple substitutable products with the objective of maximizing expected revenue. The seller offers n products, represented by the set $[n] = \{1, 2, \ldots, n\}$. We let 0 denote the "no-purchase" option. The seller needs to determine the price vector $p = (p_1, \ldots, p_n) \in \mathbb{R}_+^n$. Given the price vector p, customers choose product $j \in [n] \cup \{0\}$ with probability $q_j(p)$. However, the true choice probabilities are unknown to the seller. Instead, the seller has access to T historical transaction data points. Each transaction entry includes a historical price vector $p^t \in \mathbb{R}_+^n$ and the corresponding customer decision $y^t \in [n] \cup \{0\}$, indicating which product was purchased.

Our goal is to leverage the transaction data to prescribe an effective pricing policy that maximized the expected revenue. In the following subsections, we explore and propose various data-driven methods to design high-performance pricing policies.

2.1 Estimate-then-Optimize MNL Loss

A standard approach to solving the data-driven multi-product pricing problem is known as Estimate-Then-Optimize (ETO). In this method, we assume that the ground truth follows a specific choice model and use the Maximum Likelihood Estimation (MLE) method to estimate customer choice behavior. Then, we solve the corresponding price optimization problem based on the estimated choice model.

For instance, we can consider that customers' choice can be represented by a classical Multinomial Logit (MNL) model with linear utilities. Given price vector p, the choice probability of choosing product $i \in [n]$ is given by

$$q_i(p) = \frac{e^{\alpha_i - \beta p_i}}{1 + \sum_{j=1}^n e^{\alpha_j - \beta p_j}},$$

where β represents the price sensitivity and α_j takes into account all feature values excluding the price. An important property of the MNL model is that the optimal prices are uniform and can be captured using the Lambert W function.

Proposition 1. *For the MNL model, uniform pricing is optimal. Let p^* denote the optimal prices and $\mathcal{R}(p^*)$ denote the optimal revenue under pricing policy p^*. Then, it holds that*

$$\mathcal{R}(p^*) = \frac{W(\sum_{j=1}^n e^{\alpha_j - 1})}{\beta}, \quad p_i^* = \frac{1 + W(\sum_{j=1}^n e^{\alpha_j - 1})}{\beta},$$

where $W(\cdot)$ is the Lamert W function.

Under the ETO framework, we begin by applying MLE to fit the transaction data. This involves optimizing the following objective function to determine the best α and β that align with the historical data.

$$\max_{\alpha,\beta} \mathcal{L}_{MLE}(\alpha, \beta) = \frac{1}{T} \sum_{t=0}^{T} \Big(\mathbb{I}_{y^{(t)} \neq 0} \cdot \log \frac{e^{\alpha_{y^{(t)}} - \beta p_{t y^{(t)}}}}{1 + \sum_{j=1}^n e^{\alpha_j - \beta p_{tj}}} + \mathbb{I}_{y^{(t)} = 0} \cdot \log \frac{1}{1 + \sum_{j=1}^n e^{\alpha_j - \beta p_{tj}}} \Big).$$

Let $\hat{\alpha}$ and $\hat{\beta}$ denote the optimal solution of the above problem. Then, for $i \in [n]$, our proposed price is $p_i = \frac{1 + W(\sum_{j=1}^n e^{\hat{\alpha}_j - 1})}{\hat{\beta}}$.

2.2 Combined Estimation and Decision Loss

As an alternative to the ETO method, we may consider the revenue maximization objective directly. This is akin to minimizing the decision error induced by an estimation rather than the estimation error. Specifically, we consider the MNL Model again, whose closed-form expressions for optimal revenue $\mathcal{R}(p^*)$ given utility and price sentivity parameters α and β may be used as the revenue-maximizing decision objective:

$$\max_{\alpha,\beta} \mathcal{R}(\alpha,\beta) := \frac{W(\sum_{j=1}^{n} e^{\alpha_j - 1})}{\beta}$$

However, optimizing for maximum revenue directly may lead to the selection of an arbitrarily large α or a very small β. Therefore, it is necessary to introduce additional penalty terms on α and β. Furthermore, revenue optimization does not take into account customer decisions from historical data, which are important in establishing the relative utilities of different products. Therefore, we formulate a combined estimation and decision loss that encourage both alignment with transaction data and revenue maximization, along with penalization terms for the parameters.

$$\max_{\alpha,\beta} \mathcal{L}_{combined}(\alpha,\beta) := c \cdot \mathcal{R}(\alpha,\beta) + (1-c)\mathcal{L}_{MLE} - \lambda_\alpha \Big(\sum_{j=1}^{n} \alpha_j^2 \Big) + \lambda_\beta \log \beta$$

Note that c is a tunable parameter that determines the weight of revenue and MLE in the loss function. Let $\hat{\alpha}$ and $\hat{\beta}$ denote the optimal solution of the above problem. Then, for $i \in [n]$, our proposed price is $p_i = \frac{1+W(\sum_{j=1}^{n} e^{\hat{\alpha}_j - 1})}{\hat{\beta}}$.

2.3 Hinge Loss Function

[1] introduce a hinge loss function that does not rely on specific assumptions about the customer's underlying purchase model. The paper shows that minimizing the hinge loss function can lead to an effective pricing policy for selling a single product. Inspired by this idea, we generalize this loss function to the multi-product case. Let $\pi \in \mathbb{R}_+^n$ denote the suggested price vector. For historical price vector P and customer decision Y, our proposed hinge loss function is defined as:

$$L_c^h(\pi, Y, P) = \frac{1}{\phi(P)} \cdot \begin{cases} \sum_{i=1}^{n} (\pi_i - P_i)^+, & Y = 0 \\ c(P_Y - \pi_Y)^+ + (1-c) \cdot \sum_{i=1}^{n} (\pi_i - P_i)^+, & Y \in N, \end{cases}$$

where $\phi(P)$ denotes the density function of the historical pricing policy, and c is a parameter chosen by the seller. The intuition behind this loss function is straightforward: (i) When a customer selects the no-purchase option, the loss function encourages the reduction of prices for all products as the offered prices likely exceed the customer's valuations. (ii) When item i is sold, the loss function encourages lowering the prices of unsold items and increasing the price of the sold item. Given the proposed hinge loss function, the prescribed price vector is obtained by solving the optimization problem below:

$$\min_{\pi} \frac{1}{T} \sum_{t=1}^{T} \mathcal{L}_c^h(\pi, Y_t, P_t).$$

2.4 Direction Loss Function

The direction loss function, inspired by [3], reformulates the multi-product pricing problem as a one-dimensional search problem. This policy begins with a fixed vector of prices $f \in \mathbb{R}_+^n$, referred to as the "direction." These prices are then rescaled by a one-dimensional decision variable π^f to determine an optimal scaling factor. The loss function, defined for arbitrary historical transaction data, takes a form similar to the hinge loss:

$$L_c^d(\pi^f, Y, P) = \frac{1}{\phi(P)} \cdot \begin{cases} \sum_{i=n}^{n} (\pi^f \cdot f_i - P_i)^+, & Y = 0 \\ c(P_Y - \pi^f \cdot f_i)^+ + (1-c) \cdot \sum_{i=1}^{n} (\pi^f \cdot f_i - P_i)^+, & Y \in N, \end{cases}$$

where $\phi(P)$ represents the density function of the historical pricing policy, and c is a parameter chosen by the seller. In essence, instead of optimizing over an n-dimensional price vector, the optimization occurs over a scalar π^f. Similar to the hinge loss function, it penalizes prices that are below the listed prices when one item was sold and penalizes prices that are above the listed prices when no items are sold. Based on the direction loss function, the prescribed price vector is obtained by solving the following optimization problem:

$$\min_{\pi^f} \frac{1}{T} \sum_{t=1}^{T} \mathcal{L}_c^d(\pi^f, Y_t, P_t). \tag{1}$$

An important remaining question is how to select appropriate directions. In the next section, we show that when the direction is set to the unit vector, i.e., $f = (1, \ldots, 1)$, the prices derived from the above minimization problem can provide a meaningful revenue guarantee under mild assumptions. In addition, in the numerical experiments, we consider two directions: the unit vector and the average historical price vector.

3 Performance Guarantee for Direction Loss Function

In this section, we establish the performance guarantee for our proposed direction loss function. We consider that customer has independent random valuations V_i for $i \in [n]$ and selects the one with the highest non-negative utility, defined as $V_i - p_i$.

Our approach leverages results from [1] and [3] to establish a reasonable lower bound on revenue without imposing model-specific assumptions on the choice model. Using the analysis in [3], we bound the revenue gap between the optimal non-uniform pricing policy and the optimal uniform pricing policy. Building on this, we apply the results from [1] to demonstrate that transaction data can reveal useful information about the aggregate valuation function. This insight enables us to set a price that guarantees the uniform pricing policy's performance from the direction loss is lower bounded relative to the optimal uniform pricing policy. We first impose several standard assumptions in [1] and [3].

Assumption 1 (Weak-Rationality)**.** *Choice probabilities $q_i(p)$, $i \in [n]$ satisfy the substitutable property:*

(a) For all $k \neq i$, $q_i(p)$ is increasing in p_k for $k \neq i$.

(b) $\sum_{i=1}^{n} q_i(p)$ is decreasing in p_k for all $k \in [n]$.

Assumption 2 (Positive and Finite Optimal Price)**.** *We assume that $\infty > p_i^* > 0$ for all $i \in [n]$.*

Assumption 3 (Log-Concavity)**.** *The complementary CDF of the valuation $\bar{F}_{V_i}(v) = \mathbb{P}(V_i > v)$ is log-concave, i.e., for all $x, y \in dom\, F_{V_i}$ and $0 < \theta < 1$, it satisfies*

$$\bar{F}_{V_i}(\theta x + (1-\theta) y) \geq \bar{F}_{V_i}(x)^{\theta} \bar{F}_{V_i}(y)^{(1-\theta)}, \quad \forall i \in [n].$$

Let R^* denote the optimal revenue. We also denote p_h as the minimizer of (1) when the direction f is the unit vector. Then, the corresponding revenue is denoted as $\mathcal{R}^f(p_h)$. Under the above assumptions, we have the performance guarantee for our direction loss function. The proof can be found in the Appendix.

Theorem 1. *Suppose assumptions 1, 2, and 3 hold. Moreover, suppose that the optimal price is bounded by $p_{min} \leq p_i^* \leq p_{max}$. Then, we can choose $c^* = 0.8234$ to obtain a price p_h derived from the direction loss, and the revenue is guaranteed, relative to the non-uniform optimal pricing policy:*

$$\frac{\mathcal{R}^f(p_h)}{\mathcal{R}^*} \geq \frac{0.7715}{(1 + \ln(p_{max}/p_{min}))}.$$

4 Numerical Experiments

In this section, we conduct numerical experiments to evaluate the performance of different pricing policies proposed in section 2. We consider two synthetic settings where the underlying choice models are governed by the MNL model (Section 4.1) and the Markov Chain choice model (Section 4.2). The second setting is used to evaluate the performance of our MNL-based approaches (Sections 2.1 and 2.2) when the true underlying structure is misspecified.

4.1 Numerical performance on MNL choice model

We start with the case where the ground truth model follows the MNL model. The data generation process follows the approach outlined in [2], allowing us to compare the proposed pricing policies with the benchmark policy studied in that work. Specifically, we generate synthetic instances with $n = 10$ products and varying number of data points $T \in [20, 60, \ldots, 300]$.Under the MNL model, the probability of choosing product $j \in [n]$ when facing prices $(p_{t1}, \ldots, p_{tn})$ is given by

$$\frac{\exp(\alpha_j - \beta p_{tj})}{1 + \sum_{i=1}^{n} \exp(\alpha_i - \beta p_{ti})}.$$

We consider $\beta = 0.5$. Two parameter configurations are considered: (i) High-utility experiments: The parameters $\{\alpha_j\}_{j=1}^{1}0$ are independently drawn from the uniform distribution over $[1, 3]$ and the historical prices are sampled uniformly from $[5.5, 8.5]$. (ii) Low-utility experiments: The parameters $\{\alpha_j\}_{j=1}^{1}0$ are independently drawn from the uniform distribution over $[-2, 0]$ and the historical prices are sampled uniformly from $[2.5, 4.5]$. For each value of T, we simulate 60 independent instances and report the average revenue ratios of all methods compared to the optimal revenue, as shown in Figure 1 and 2.

From the figures, we observe that when the dataset size is small, the MLE method struggles because it is impossible to get a good estimation of 11 variables with only 20 data points. However, both the combined method and direction loss method achieve near-optimal revenues. This is reasonable because the combined method prescribes moderate prices, which are effective in this setting. Additionally, since uniform pricing is optimal for the MNL model, the direction-loss method performs near-optimally, as the unit vector is one of the suggested directions. Moreover, our proposed hinge loss function demonstrates better or equivalent performance compared to the model-free method studied in [2]. Finally, as the number of data points increases, the performance of the MLE method improves significantly.

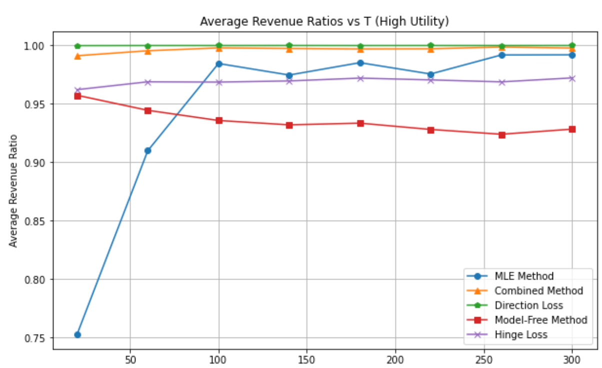

Figure 1: Model performance under high-utility MNL model.

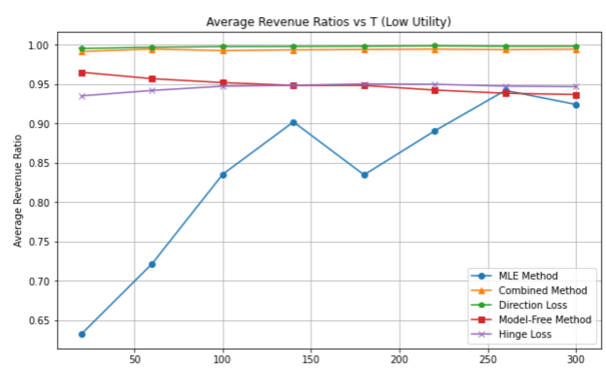

Figure 2: Model performance under low-utility MNL model.

4.2 Numerical performance on misspecified choice model (Markov Chain choice model)

We then consider the case where the ground truth model follows the Markov Chain choice model. Specifically, we generate synthetic instances with $n = 10$ products and varying number of data points $T \in [40, 80, \ldots, 320]$. We evaluate performance in two settings: (i) High-price experiments: The historical prices are samples uniformly from [6.5,10.5]. (ii) Low-price experiments: The historical prices are samples uniformly from [2,5]. For each value of T, we simulate 60 independent instances and report the average expected revenues of the proposed pricing policies, as shown in Figure 3 and 4.

For the high-price setting, both the combined method and the direction-loss method consistently achieve relatively high expected revenues across different dataset sizes. Additionally, the MLE method outperforms the hinge-loss method in this context. A plausible explanation for the strong performance of these three methods is their tendency to prescribe moderate prices, which are more effective in the high-price setting. In contrast, for the low-price setting, the MLE method shows the weakest performance among all methods, as expected, given that the ground truth model does not follow the MNL structure. Meanwhile, the combined method and the direction-loss method continue to exhibit strong performance. Finally, all proposed pricing policies consistently outperform the model-free method from [2], highlighting the robustness and effectiveness of our approaches even when the underlying model is misspecified.

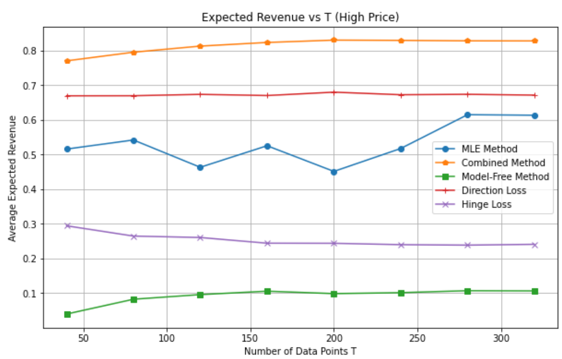

Figure 3: Model performance under high-price Markov Chain Choice Model.

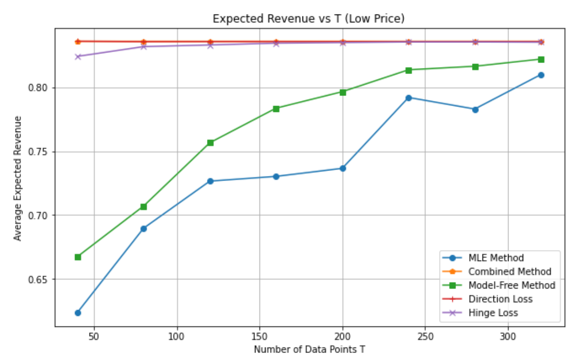

Figure 4: Model performance under low-price Markov Chain Choice Model.

5 Conclusion and Future Work

We consider various loss functions for multi-product pricing problems. Initially, we adopt a direct ETO approach by imposing a MNL structure on the transaction data. We estimate price sensitivity and subsequently determine the optimal price. Next, we explore the combine-revenue-likelihood loss, which accounts for the potential revenue from the proposed price. Intuitively, this approach not only seeks to maximize the likelihood of the parameters but also looks for a price vector that yields high revenue. We then extend the hinge loss proposed in [1] to multi-product setting. A simpler version of this is the direction loss, where we have a base price vector and rescale the prices of all products simultaneously. Under mild assumptions, we can guarantee the revenue from the direction-loss pricing policy. We perform numerical experiments to verify the near-optimal performance of our proposed policies.

For future work, we plan to explicitly find or bound $p_{\min}$ and $p_{\max}$ as shown in the theoretical results. We hypothesize that if the value distributions V_i and V_j are very close (with respect to some divergence measure) and are independent, the gap between $q_{\min}$ and $q_{\max}$ may be small.

References

[1] M. Biggs, "Convex surrogate loss functions for contextual pricing with transaction data," *arXiv preprint arXiv:2202.10944*, 2022.

[2] N. Chen, A. A. Cire, M. Hu, and S. Lagzi, "Model-free assortment pricing with transaction data," *Management Science*, vol. 69, no. 10, pp. 5830–5847, 2023.

[3] G. Gallego and G. Berbeglia, "Bounds and heuristics for multi-product personalized pricing," *arXiv preprint arXiv:2102.03038*, 2021.

Appendix

Proof of Proposition 1

The mean utility of a product is given by $u_i = \alpha_i - \beta p_i$. As stated previously, the expected revenue under the MNL choice model is defined as

$$\mathcal{R}(p) = \sum_{i=1}^{n} p_i q_i(p) = \sum_{i=1}^{n} p_i \frac{e^{\alpha_i - \beta p_i}}{1 + \sum_{j=1}^{n} e^{\alpha_j - \beta p_j}}$$

For clarity, Let $A_i = e^{\alpha_i - \beta p_i}$, and $B = 1 + \sum_{j=1}^{n} e^{\alpha_j - \beta p_j}$. Taking the derivative of the revenue with respect to prices, we obtain

$$\frac{\partial \mathcal{R}}{\partial p_i} = \frac{A_i}{B} + p_i\Big(- \frac{\beta A_i}{B} + \frac{\beta A_i^2}{B^2}\Big) + \sum_{j \neq i} p_j \frac{\beta A_j A_i}{B^2}$$

Setting derivative equal to zero, and assuming uniform pricing $p = p_i = p_j$, and noting that $\sum_j A_j = B - 1$, we have

$$\begin{aligned} BA_i + p\beta\big(- BA_i + A_i^2\big) + \sum_{j \neq i} p\beta A_j A_i = 0 \\ B - p\beta B + p\beta(B-1) = 0 \\ p\beta = B = 1 + \sum_{j=1}^{n} e^{\alpha_j - \beta p} \end{aligned}$$

This expression may be aranged as a Lambert W function with some more algebraic manipulations and the result follows:

$$\begin{aligned} p\beta - 1 &= e^{-(\beta p - 1)} \sum_{j=1}^{n} e^{\alpha_j - 1} \\ (p\beta - 1)e^{(p\beta - 1)} &= \sum_{j=1}^{n} e^{\alpha_j - 1} \\ p\beta - 1 &= W\big(\sum_{j=1}^{n} e^{\alpha_j - 1}\big) \\ p &= \frac{1 + W(\sum_{j=1}^{n} e^{\alpha_j - 1})}{\beta} \end{aligned}$$

Proof of Theorem 1

Proof. Under Assumptions 1 and 2, we obtain the following result, proven in [3]:

Proposition 2. *Suppose that assumption 1 and assumption 2 hold. We also assume that* $p_{\min} \leq p_i^* \leq p_{\max}$ *for all* $i \in [n]$*. Then, we can select a uniform price* $f = (1, 1, \ldots, 1)$ *and obtain the following guarantee:*

$$\mathcal{R}^* \leq (1 + \ln(p_{\max}/p_{\min}))\mathcal{R}^f(p_f^*)$$

where $\mathcal{R}^*$ *and* $\mathcal{R}^f(p_f^*)$ *are the optimal revenues using the non-uniform and uniform pricing, respectively.*

Next, we lower bound the ratio $\frac{\mathcal{R}^f(p_h)}{\mathcal{R}^f(p_f^*)}$ for uniform direction $f = (1, \ldots, 1)$. To simplify the analysis, we further assume that our historical transaction data is generated using a directional pricing policy. Then, the expected loss is given by:

$$\mathbb{E}[L_c^d(\pi^f, y, p)] = \int_p \frac{1}{\phi(p)} \left[(1 - c \Pr(y \neq 0|p)) \left(\pi^f - p\right)^+ + c \Pr(y \neq 0|p) \left(p - \pi^f\right)^+ \right] \phi(p) dp.$$

Following the argument in [1], we obtain that the minimizer p_h satisfy:

$$p_h = c \int_0^\infty \Pr(y \neq 0|p) dp = c \int_0^\infty \Pr\left(V_i \geq p f_i \text{ for some } i\right) dp = c \int_0^\infty \Pr(W \geq p) dp = c\mathbb{E}[W],$$

where $W = \max\{V_1, \ldots, V_n\}$.

In [1], the assumption of log-concavity on the valuation distribution is used to derive the revenue guarantee. We adopt this log-concavity assmuption and show that W, the maximum valuation among the products, also possesses the log-concavity property:

Lemma 1 (Log-Concavity of the Maximum Valuation)**.** *Assume that* $F_i(x)$*,* $f_i(x)$*, and* $f_i'(x)$ *exist for all* $x \in [0, \infty)$ *and* $i \in [n]$*. Suppose further that* $\bar{F}_i(x)$ *is log-concave, and the random valuations* V_i *are independent. Then, the complementary CDF of* $W = \max\{V_1, \ldots, V_n\}$ *is also log-concave.*

Proof. See the next section. □

With uniform pricing, we can treat W as a random valuation of a single product. Therefore, from [1], we can choose $c^* = 0.8234$ such that the revenue ratio is lower bounded when using the directional loss price p_h, i.e.,

$$\frac{\mathcal{R}^f(p_h)}{\mathcal{R}^f(p_f^*)} \geq 0.7715$$

where p_f^* is the optimal uniform price. Specifically, note that the revenue from using a p_h-uniform pricing policy is

$$\mathcal{R}^f(p_h) = p_h \mathbb{P}\left(\max\{V_1, V_2, \ldots, V_n\} \geq p_h\right) = p_h \mathbb{P}\left(W \geq p_h\right)$$

From [1], if W is a log-concave valuation distribution and $p_h = c\mathbb{E}\left[W\right]$, then

$$\frac{\mathcal{R}^f(p_h)}{\mathcal{R}^f(p_f^*)} = \frac{p_h \mathbb{P}\left(W \geq p_h\right)}{\sup_{p>0} p\mathbb{P}\left(W \geq p\right)} \geq \min\left\{ \min_{z \leq -2c} \frac{cze^{-z\left(\frac{1}{c}-1\right)-1}}{z+c}, \min_{0<z<1} \frac{c(z-1)e^{c(z-1)}}{z\ln(z)} \right\}$$

where the denominator term is the revenue from the optimal uniform pricing policy. We choose $c = 0.8234$ so that the ratio is lower bounded by 0.7715.

From the results above, the ratio between the uniform direction-loss price and the optimal uniform pricing policy is lower bounded by 0.7715. Moreover, the ratio between the optimal uniform pricing policy and the optimal non-uniform pricing policy is lower bounded by $1/\left(1 + \ln(p_{max}/p_{min})\right)$. Combining these two results yields the proposed bound. □

Proof of Lemma 1

Proof. For simplicity, we assume that $F_i(x) < 1$ for all $x \in \mathbb{R}$ and $i \in [n]$.

Since all the valuation are independent, we thus have that

$$G(x) = \mathbb{P}\left(W \leq x\right) = \mathbb{P}\left(\max V_1, \dots, V_n \leq x\right) = \mathbb{P}\left(V_1 \leq x\right) \dots, \mathbb{P}\left(V_n \leq x\right) = \prod_{i=1}^{n} F_i(x).$$

Since $\bar{F}_i$ is log-concave, we know that

$$\bar{F_i}'' = \frac{\partial^2}{\partial x^2} \log\left(1 - F_i(x)\right) = \frac{-f_i'(x)(1 - F_i(x)) + f_i^2(x)}{\left(1 - F_i(x)\right)^2} \leq 0.$$

Then, we consider each $x \in \mathbb{R}^+$ and omit x as an input of functions. That is, from the above log-concavity result, we can write it as

$$f_i'\left(1 - F_i\right) \geq f_i^2$$

for all $i \in [n]$.

We then note that

$$\begin{aligned}
\bar{G}'' &= \frac{\partial^2}{\partial x^2} \log\left(1 - \prod_{i=1}^{n} F_i\right) \\
&= \frac{-\left(\sum_{j=1}^{n} F_1 \dots F_{j-1} f_j' F_{j+1} \dots F_n + 2\left(\sum_{1 \leq j < k \leq n} F_1 \dots F_{j-1} f_j F_{j+1} \dots F_{k-1} f_k F_{K+1} \dots F_n\right)\right)\left(1 - \prod_{i=1}^{n} F_i\right)}{\left(1 - \prod_{i=1}^{n} F_i\right)^2} \\
&\quad + \frac{\left(\sum_{j=1}^{n} F_1 \dots F_{j-1} f_j F_{j+1} \dots F_n\right)^2}{\left(1 - \prod_{i=1}^{n} F_i\right)^2}.
\end{aligned}$$

We note that $F_i, f_i \geq 0$, and W is log-concave if

$$\left(\sum_{j=1}^{n} F_1 \dots F_{j-1} f_j' F_{j+1} \dots F_n\right)\left(1 - \prod_{i=1}^{n} F_i\right) \geq \left(\sum_{j=1}^{n} F_1 \dots F_{j-1} f_j F_{j+1} \dots F_n\right)^2$$

We can rewrite each term in the LHS as follows:

$$\begin{aligned}
&F_1 \dots F_{j-1} f_j' F_{j+1} \dots F_n \left(1 - \prod_{i=1}^{n} F_i\right) \\
&= F_1 \dots F_{j-1} f_j' F_{j+1} \dots F_n \left(1 - F_j + F_j\left(1 - F_1\right) + F_j F_1\left(1 - F_2\right) + \dots + F_1 \dots F_j\left(1 - F_{j+1}\right) + F_1 \dots F_{n-1}\left(1 - F_n\right)\right) \\
&\geq F_1 \dots F_{j-1} f_j^2 F_{j+1} \dots F_n + f_j^2 \sum_{k \neq j} \left(\prod_{i=1}^{n} F_i\right)\left(\prod_{i=1; i \neq j}^{k-1} F_i\right) \frac{1 - F_k}{1 - F_j}
\end{aligned}$$

Since $F_i < 1$, we thus have that

$$F_1 \dots F_{j-1} f_j' F_{j+1} \dots F_n \left(1 - \prod_{i=1}^{n} F_i\right) \geq F_1 \dots F_{j-1} f_j^2 F_{j+1} \dots F_n + f_j^2 \sum_{k \neq j} \left(\prod_{i=1}^{n} F_i\right)^2 \frac{1 - F_k}{\left(1 - F_j\right) F_j F_k}$$

For $j \neq k$, we know that

$$f_j^2 \left(\prod_{i=1}^{n} F_i\right)^2 \frac{1 - F_k}{\left(1 - F_j\right) F_j} + f_k^2 \left(\prod_{i=1}^{n} F_i F_k\right)^2 \frac{1 - F_j}{\left(1 - F_k\right) F_k F_j} \geq 2\left(\prod_{i=1}^{n} F_i\right)^2 f_j f_k / F_j F_k$$

by A.M.-G.M. Therefore,

$$\begin{aligned}
&\left(\sum_{j=1}^{n} F_1 \dots F_{j-1} f_j' F_{j+1} \dots F_n\right)\left(1 - \prod_{i=1}^{n} F_i\right) \\
&\geq \sum_{i=1}^{n} F_1 \dots F_{j-1} f_j^2 F_{j+1} \dots F_n + 2\sum_{j<k}\left(\prod_{i=1}^{n} F_i\right)^2 f_j f_k / F_j F_k \\
&= \left(\sum_{j=1}^{n} F_1 \dots F_{j-1} f_j F_{j+1} \dots F_n\right)^2 .
\end{aligned}$$

Therefore, W is log-concave.

$\square$