The Power of Opaque Products in Pricing

Adam N. Elmachtoub
Department of Industrial Engineering and Operations Research, Columbia University, New York, NY, 10027, adam@ieor.columbia.edu
Michael L. Hamilton
Katz Graduate School of Business, University of Pittsburgh, Pittsburgh, PA 15260 mhamilton@katz.pitt.edu

We study the power of selling opaque products, i.e., products where a feature (such as color) is hidden from the customer until after purchase. Opaque products, which are sold with a price discount, have emerged as a powerful vehicle to increase revenue for many online retailers and service providers that offer horizontally differentiated items. In the opaque selling models we consider, each of the items are sold at a single common price alongside opaque products which may correspond to various subsets of the items. We consider two types of customers, *risk-neutral* ones who assume they will receive a truly random item from the opaque product, and *pessimistic* ones who assume they will receive their least favorite item from the opaque product. We benchmark opaque selling against two common selling strategies: discriminatory pricing, where one explicitly charges different prices for each item, and single pricing, where a single price is charged for all the items.

We give a sharp characterization of when opaque selling outperforms discriminatory pricing. Namely, this result holds for situations where all customers are pessimistic, or the item valuations are supported on two points. In the latter case, we also show that opaque selling with just one opaque product guarantees at least 71.9% of the revenue from discriminatory pricing. We then provide upper bounds on the potential revenue increase from opaque selling strategies over single pricing, and describe cases where the increase can be significantly more than that of discriminatory pricing. Finally, we provide pricing algorithms and conduct an extensive numerical study to assess the power of opaque selling under for a variety valuation distributions and model extensions.

Key words: opaque products; price discrimination; probabilistic selling; random utility maximization

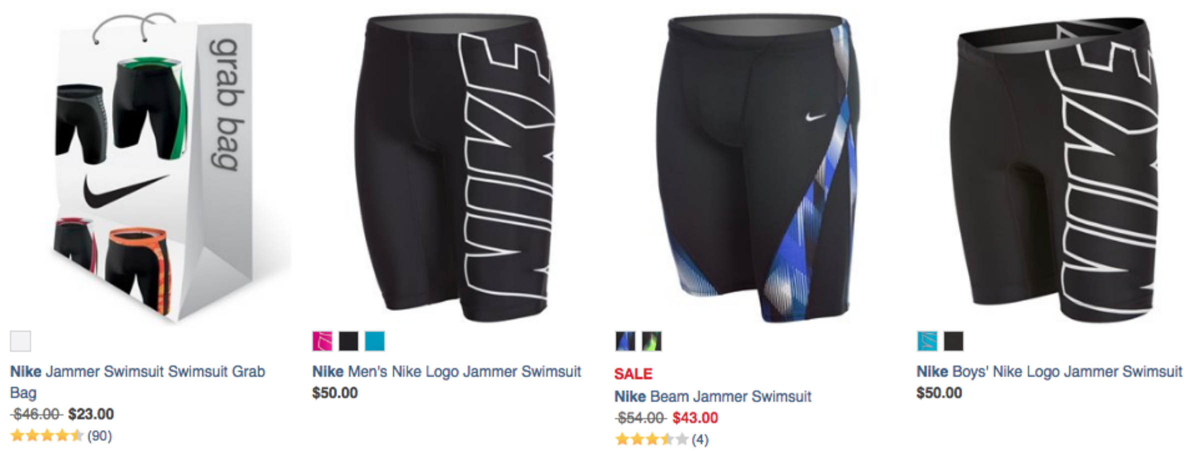
1. Introduction

An opaque product is a product where one or more features (such as color, brand, or location) are hidden from the customer until after the purchase is made. In recent years, several online retailers have begun selling opaque products. For example, Amazon.com offers various colors of Swingline staplers alongside a "colors may vary" option, which is an opaque product over the various colors (see Fig. 1). In another example, SwimOutlet.com offers various styles of Nike swimsuits, as well as a "Grab Bag" over all the different styles offered (see Fig. 2).

Figure 1 Opaque Selling on Amazon.com.

Note. Swingline offers their "SmartTouch" staplers on Amazon.com traditionally alongside a "colors may vary" option i.e., a single opaque product where the color of the stapler is hidden until after purchase. The opaque product is offered at the discounted price of $15.99, which is $1.63 less than the traditional price of $18.62.

Figure 2 Opaque Selling on SwimOutlet.com.



Note. SwimOutlet.com offers various styles of Nike brand swim trunks for prices between $43.00-$50.00 alongside an opaque "Grab Bag" for $23.00.

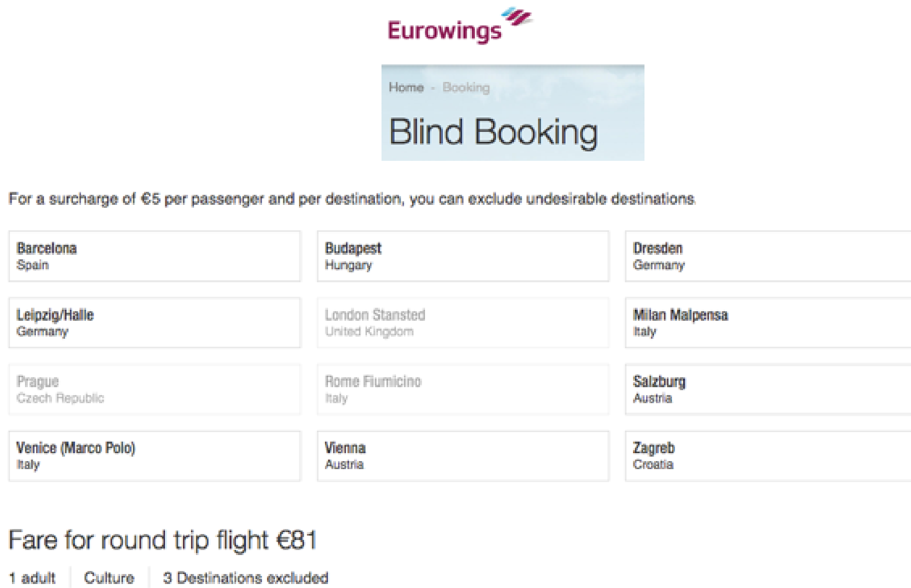
In both of these examples, customers who purchase the opaque product sacrifice exact knowledge of the item they will receive in exchange for a price discount. This allows the seller to price discriminate between customers with strong and weak preferences, and therefore earn more revenue. The goal of this work is to showcase the power of opaque selling compared to more traditional price discrimination tactics, and quantify the potential extra revenue that a seller can obtain.

In our framework, we consider a seller offering N items, each of which are similar but may differ in a secondary attribute such as color or style. Customers draw valuations for each of the items from a joint probability distribution that is known to the seller. We focus on the class of exchangeable distributions which naturally represent horizontally differentiated items (see Definition 1 for a formal definition). Exchangeable joint distributions have the properties that *(i)* the marginal distribution of each item is identical, capturing the notion of horizontal differentiation, and *(ii)* there can be (some form of) correlation between the items. Special cases of exchangeable distributions include i.i.d. valuations, the Hotelling model, and Salop's circle model (when the number of items is fewer than 3, see Section C for an in depth discussion). In the absence of opaque products, customers simply choose the item that maximizes their utility, i.e., their valuation for the item minus its price. No item is purchased in the case where the utility from all items is negative.

Interestingly, when the valuation distribution is exchangeable (even i.i.d.), the optimal pricing strategy in this model may use different prices for different items (Chawla et al. (2007)). This strategy, that we refer to as *discriminatory pricing*, is a natural benchmark for our opaque selling strategies. Due to symmetry, discriminatory pricing arbitrarily chooses some items to have high prices, e.g. in Fig. 12 customers who prefer blue are charged higher prices. This may cause customers to think the pricing strategy is unfair, and it may be particularly problematic when certain items (colors) that have a higher price are correlated with demographic information such as race or gender. In some settings, the items are often constrained to have the same price by the manufacturer or by regulatory bodies to ensure impartiality to customers. Thus another natural benchmark to consider in this context is the best *single price* strategy.

We now carefully describe our opaque selling strategy, where the seller offers opaque products in addition to offering the N items. Specifically, an *opaque product* is an explicit subset of items from which a customer will receive one item upon purchase. An opaque selling strategy can offer all possible $2^N - N - 1$ opaque products alongside the N original items. For practicality and tractability of the model, we assume that opaque products corresponding to subsets of the same size must have the same price. Moreover, we impose a restriction that the prices of the items must also be the same. This selling structure is employed by the company Eurowings which sells round trip flights to opaque destinations (see Fig. 3 and Post and Spann (2012)). In their setting, customers may narrow down the possible destinations of their trip in exchange for an increased price. As a result, an opaque selling strategy is parameterized by N prices, similar to a discriminatory pricing strategy. However, customers interested in an item or opaque products of the same size always pay the same price, which prevents the opaque strategy from arbitrarily discriminating against customers that prefer specific items. *In essence, an opaque selling strategy balances the impartiality of a single price strategy with the price discrimination capability of discriminatory item pricing.*

Figure 3 Opaque products for travel services.



Note. Eurowings.com uses an opaque selling strategy to offer airline tickets with a base price of €66.00. Here the destination of the flight is opaque, with $N = 12$ possible destinations. The site allows the customer to exclude as many destinations as they desire, where each exclusion adds an additional €5.00 to the price. In the figure above, three destinations are excluded and the price for the desired opaque product is €81.00.

In order to study opaque selling strategies, we must also specify how customers value an opaque product. In practice, the seller never reveals the probabilities of receiving individual items in an opaque product, leaving the customers to formulate their valuations based on their judgment. We consider two types of customers motivated by realistic interpretations of opaque products, which we call pessimistic and risk-neutral. A customer is said to be *pessimistic* if they value an opaque product as the *minimum* of their valuations among the corresponding subset of items. A pessimistic customer essentially wants an ex-post or worst-case guarantee that purchasing the opaque product yields positive utility. Such an assumption is particularly natural when some of the items are infeasible for the customer (e.g., unfavorable colors or incompatible destinations). Another reason a customer may be pessimistic is that they are fundamentally mistrustful of the seller's motives, and fears the seller will allocate the product they desire least.

A customer is said to be *risk-neutral* if they value an opaque product as the *average* of their valuations among the corresponding subset of items. A risk-neutral customer is essentially optimistic, and believes that the seller is impartial in the sense that the probability of receiving any item in an opaque product is uniformly distributed. Such an assumption is natural when the customer has to no reason to believe the seller knows their item preferences (for example a first time customer or anonymous online customer). Although uniform allocation and risk-neutral risk preferences may not always apply to all sellers and customers, the risk-neutral assumption has been the dominant assumption in the literature. In this work we study the power of opaque selling in markets consisting of mixtures of both customer types. We use α to denote the probability that a customer is pessimistic and call such markets α*-mixed*.

We next outline our contributions, which formally describe conditions under which opaque selling performs well with respect to discriminatory and single pricing strategies.

1. We give sufficient conditions for when opaque selling dominates the optimal discriminatory pricing strategy. In particular, we show opaque selling is *guaranteed* to provide at least as much revenue as discriminatory pricing when the valuations are drawn from an exchangeable distribution and either of the following conditions hold: *(i)* the market is homogeneously pessimistic or *(ii)* the valuations can only take two values (high or low), and supply counter-examples when neither condition is satisfied. Intuitively, under these conditions the customers naturally separate into sufficiently distinct groups that allows an opaque strategy to effectively price each group. On the other hand, discriminatory pricing can effectively price against only a subset of customers whose ranking of the items correspond to the ranking of the item prices. One surprising consequence of our result is that the seller may benefit when customers are pessimistic, thus sellers may be encouraged to not reveal the opaque product allocation probabilities in the hopes that customers adopt a worst-case valuation.
2. In the important special case when valuations are drawn from an exchangeable distribution and can only take two values, we show a single opaque product can always guarantee *at least 71.9%* of the revenue from discriminatory pricing. This result is independent of the fraction of the customers that are pessimistic.

3. We then show that in α-mixed markets, opaque selling can earn *up to and at most a factor of* αN more than the best single pricing, and there are instances where opaque selling earns $\frac{\alpha}{2}\frac{N}{1+\log(N)}$ more than discriminatory pricing. We complement this result by showing that in the restricted case of i.i.d. valuations, these gap falls to constant factors. We also show that offering just a single opaque product can increase revenue by *up to and at most a factor of* $2-(1-\alpha)\frac{N-1}{N}$ over the best single pricing.
4. We perform an extensive numerical study which bears out our results for several typical distributions. To conduct the study we first derive an *efficient* algorithm for finding the optimal prices when N is small. We empirically observe that the seller earns more revenue as α, the fraction of pessimistic customers, increases. We also see up to a 5% increase in revenue using opaque selling compared to discriminatory pricing. We next show that α can be effectively estimated from a modest amount sales data and induce opaque selling strategies that garner nearly all the available revenue. Further, even when the underlying assumptions of the model are violated, we show that using our opaque selling framework can still generate effective pricing strategies.

Taken together, our results provide strong support for opaque selling as a customer-friendly alternative to discriminatory pricing, often achieving comparable or higher revenues.

1.1. Literature Review

Our work connects into several streams of literature across operations, marketing, economics, and computer science. We first review literature on monopolistic sellers offering opaque products. The parallel works of Jiang (2007) and Fay and Xie (2008) both consider opaque selling frameworks when customers have valuations drawn from a Hotelling or Salop's circle choice model. Note that this assumes perfect correlation between the items, and may not necessarily represent customer behavior well although it does fall into the exchangeable distribution assumption. Both works provide conditions for when opaque selling can have strictly positive increase in profit over single pricing strategies. Fay and Xie (2008) also show that opaque products can be used to hedge against possibly incorrect demand estimation. Our work does not make any assumption about the valuation distribution, and benchmarks primarily against discriminatory pricing.

A separate stream of work has shown the power opaque products for managing capacity and inventory. Gallego and Phillips (2004) and Gallego et al. (2004) considers the notion of a flexible product in revenue management, where customers who buy the flexible option are allocated a product after the completion of the time horizon. Fay and Xie (2014) show how to use opaque products to protect inventory when one of the two items is strictly preferred over the other by all customers, and Xiao and Chen (2014) provide dynamic programming algorithms to decide when to use opaque products. Elmachtoub and Wei (2015), Elmachtoub et al. (2019) quantify the value of opaque products in real-time inventory management environments. In our work, we avoid any notion of cost and focus purely on the price discrimination effect offered by opaque products.

There are several works on opaque products when used among competitors (Shapiro and Shi (2008), Jerath et al. (2010)), in name-your-own-price channels (Chen et al. (2014), Huang et al. (2017)), in empirical analysis (Xie et al. (2016), Granados et al. (2018)), and in queueing systems (Xu et al. (2016), Geng (2016)). Post and Spann (2012) and Post (2010) consider settings where multiple opaque products are offered simultaneously. We also mention a stream of work in economics that considers optimal mechanism design with opaque products, along the lines of Balestrieri et al. (2015), and Balestrieri and Izmalkov (2016).

Our results related to purely risk-neutral markets ($\alpha = 0$) connects to a stream of literature on pricing with lotteries. A lottery, as described in the literature, is a probability distribution over the items that is sold by the seller and announced to the customer. Customers are risk-neutral and use the expected valuation of the lottery when deciding what to buy. If all customers are risk-neutral, then our opaque selling strategy can be thought of as a special case of lottery pricing where the items are allocated uniformly at random. Under arbitrary valuation distributions, Briest et al. (2015) and Hart and Nisan (2014) show that lottery pricing can earn infinitely more revenue than any discriminatory pricing when $N \geq 3$ and $N = 2$, respectively. When customers draw their valuations independently, Chawla et al. (2015) show the optimal lottery pricing earns at most four times discriminatory pricing.

Our work is also related to, and draws on, the algorithmic pricing literature which studies when pricing can compete with richer classes of selling mechanisms. When valuations are independent, Chawla et al. (2007, 2010), Alaei et al. (2019) show that an optimally chosen single price strategy is a constant factor approximation to the revenue of an optimal discriminatory pricing. Cai and Daskalakis (2011) then provide a polynomial-time approximation scheme for computing the optimal discriminatory pricing when valuations are drawn independently, while Chen et al. (2018) show the problem is NP-Hard even when valuations are drawn i.i.d. or drawn independently and supported on three points.

We note that our work resembles that of bundle pricing on the surface due to the nature in which items are aggregated into opaque products, although bundling results generally assume customers are interested in purchasing multiple items. The one exception is that of Briest and Roglin (2010) who frame opaque products as 'unit demand bundles' and provide hardness results. Finally, our work fits in parallel to recent work on simple mechanisms for problems in auctions (Celis et al. 2014, Hartline and Roughgarden 2009, Alaei et al. 2019, Jin et al. 2019), bundling (Ma and Simchi-Levi 2016, Abdallah et al. 2017), and first or third-degree price discrimination (Elmachtoub et al. 2018, Bergemann et al. 2020).

2. Selling Models

We now formally describe the selling models that we study throughout this work. We consider a seller who has $N \geq 2$ items for sale, described by the index set $[N] := \{1, 2, \ldots, N\}$. The seller may also offer one or more opaque products, each of which is described by a subset $S \in 2^{[N]}$ where $|S| \geq 2$. The seller simultaneously offers the items and potentially some number of opaque products to a utility-maximizing, unit-demand customer. (Note that this is equivalent to selling to many customers with no inventory constraints.) The customer has a non-negative random valuation for each item i denoted by V_i, and the joint valuation $V = (V_1, V_2, \ldots, V_N)$ is drawn from a known joint distribution F. For every selling model we consider, the customer maximizes their own utility, which is the valuation of the item or opaque product purchased minus its price. If no item or opaque product results in a non-negative utility, then the customer does not purchase anything. In the case where the customer has multiple options that maximize their utility, we assume without loss of generality, that the customer purchases the product with the highest price (see Chen et al. (2018) for detailed discussion of tie-breaking rules in this context). When there are multiple products with the same price providing maximum utility, we assume the customer breaks ties arbitrarily.

We note that the notion of valuation and utility described thus far does not extend in an obvious way to opaque products. That is, the way a customer values an opaque product depends on the customer's belief about the seller's allocation process and the customer's personal risk preferences. Next, we describe two natural frameworks for modeling valuations of opaque products.

2.1. Valuations for Opaque Products

We model the customer's valuation for an opaque product as a function over their valuations for the items the opaque product can return. We consider two natural assumptions for how to model customer behavior with respect to opaque products, which we call *pessimistic* and *risk-neutral.* For any subset of items $S \in 2^N$, we let V^S denote the random valuation of the opaque product corresponding to S. For pessimistic customers, V^S is the minimum over all the valuations in S, and for risk-neutral customers, V^S is the average over all the valuations in S, i.e.,

$$\text{(Pessimistic)} \quad V^S = \min_{i \in S}\{V_i\}, \qquad\qquad \text{(Risk-Neutral)} \quad V^S = \frac{\sum_{i \in S} V_i}{|S|}.$$

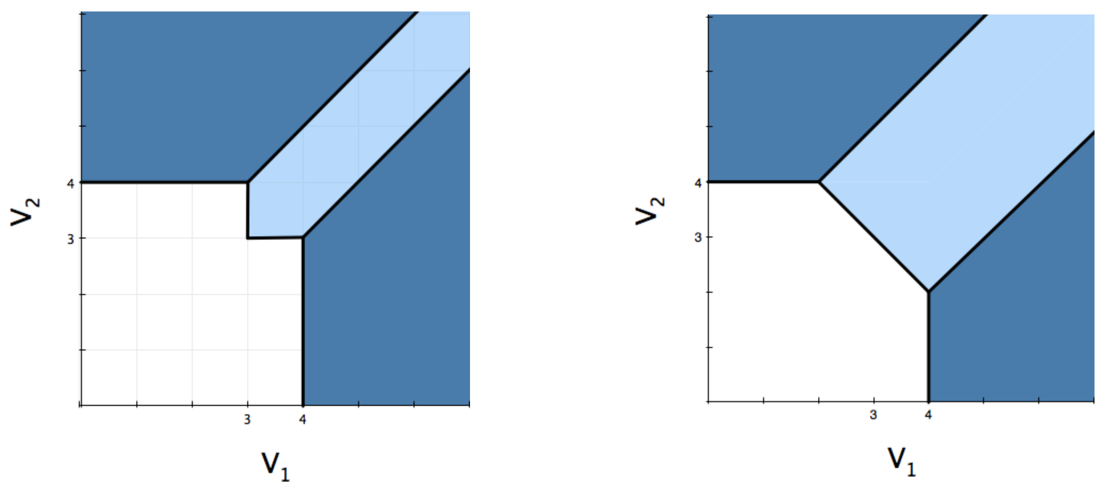
We assume that a customer is pessimistic with probability α and risk-neutral with probability $1 - \alpha$. We let X_α be the random variable corresponding to the customer type.

In practice, the seller never announces the allocation probabilities for an opaque product, forcing the customer to form their own beliefs. A pessimistic customer believes the seller will allocate the product that is desired least by the customer. Given that the allocation probabilities are entirely unknown, this corresponds to a customer placing a worst-case allocation distribution on the outcome of the opaque product. The pessimistic preference also captures another important and practical situation where even if the customers know the opaque allocation probabilities, they are extremely risk-averse. In other words, the customer wants their purchasing decision to be ex-post optimal, i.e., there is no regret even after the item in the opaque product is revealed. This particular situation can arise when customers know that certain items provide no value, which can happen when particular colors or flight destinations are completely undesirable (see Figures 2 and 3).

A risk-neutral customer believes that the seller will allocate the items in the opaque product uniformly at random, which is an optimistic belief. With respect to this fair allocation, they are also risk-neutral in their valuation of the opaque product. Thus, a risk-neutral customer simply averages their valuations across the product, even though the allocation is most likely not uniformly at random. With limited information, it is natural for some customers to form this valuation, in particular when the valuations of each item are reasonably close together (in which case the difference between risk-neutral and pessimistic is small). Furthermore, in some special cases it is ex-ante revenue optimal for the seller to design the allocation of opaque products in a balanced way, conforming to and lending additional motivation for the risk-neutral assumption. We explore this idea further in Appendix A. We also note that the risk-neutral assumption has been the primary focus in the literature (Gallego and Phillips (2004), Fay and Xie (2008), Jerath et al. (2010)), while the pessimistic case has not been studied to our knowledge.

Finally, we highlight that the delineation between pessimistic and risk-neutral customers is quite important from a technical perspective. Fig. 4 illustrates this distinction geometrically in the valuation space in the case where $N = 2$. Note the shape and size of the valuations regions where customers purchase the opaque product are quite different, which explains the dependence on α in our analysis.

Figure 4 Visualizing the valuation space.



(a) Valuation space for pessimistic customers

(b) Valuation space for risk-neutral customers

Note. Above are two valuation spaces for opaque selling strategies with prices $(p, p^2) = (4, 3)$. The darkened regions correspond to customer valuations that yield purchases of an item at a price of 4. The lighter regions correspond to purchases of the opaque product which has a price of 3. The unshaded regions correspond to valuations that result in no purchase.

2.2. Selling Strategies

We now describe four specific selling strategies that we use throughout the work. For notational convenience and improved exposition of this subsection, we assume that F is continuous to avoid tie-breaking scenarios (which would go to the highest price option w.l.o.g.). In the *single price selling model* (SP), the seller offers all N items all at the same price. In other words, the price of item i is the same for all $i \in [N]$. We refer to this single price as p. We denote $\mathcal{R}^F_{SP}(p)$ and $\mathcal{R}^F_{SP}$ as the expected revenue using single pricing with joint distribution F under price p and the optimal price, respectively. Formally,

$$\mathcal{R}^F_{SP}(p) = p\mathbb{P}(\max_{i\in[N]}\{V_i\} \geq p) \text{ and } \mathcal{R}^F_{SP} = \max_p \mathcal{R}^F_{SP}(p).$$

In the *discriminatory pricing model* (DP), the prices may differ between the items. Without loss of generality, we always relabel the indices so that $p_1 \geq p_2 \geq \ldots \geq p_N$. We denote the vector of prices as $\vec{p}$. We note that even when valuations are i.i.d., discriminatory item pricing may provide strictly more revenue than single pricing strategies (Chawla et al. (2007)). However, discriminatory pricing may be difficult for customers to accept (especially when valuations are i.i.d.), and for similar reasons may be infeasible for the seller due to business constraints. We denote the expected revenue under F using prices $\vec{p}$ by $\mathcal{R}^F_{DP}(\vec{p})$ and the optimal item pricing by $\mathcal{R}^F_{DP}$. Formally,

$$\mathcal{R}^F_{DP}(\vec{p}) = \sum_{i\in[N]} p_i \mathbb{P}\left(V_i - p_i \geq \max_{j\neq i}\{V_j - p_j, 0\}\right) \text{ and } \mathcal{R}^F_{DP} = \max_{\vec{p}} \mathcal{R}^F_{DP}(\vec{p}).$$

Although DP seems at first unnatural and counterintuitive when valuations are i.i.d., the revenue function creates a natural tension to segment the market and capture high valuation customers without sacrificing market size. Every time an item is priced high, selling another item at a low price becomes more valuable since its market share will increase. Although DP can provide more revenue in many i.i.d. settings, including two point (high-low) distributions, it may not be beneficial in other settings. For example, when i.i.d. valuations correspond to a multinomial logit (MNL) choice model, the 'constant markup property' (Anderson et al. (1992)) implies SP is optimal. We provide a longer primer delving further into DP in Appendix B.

In the *single opaque selling model* (1OPQ), the seller offers only one opaque product associated with the set $[N]$ at a price p^N, alongside the traditional items all at a fixed price p. This model is important and common in practice due to its simplicity, impartiality, and ease of implementation. We denote the expected revenue under F, in α-mixed markets, using prices (p, p^N) by $\mathcal{R}^{F,\alpha}_{1OPQ}(p, p^N)$ and the optimal pricing by $\mathcal{R}^{F,\alpha}_{1OPQ}$. Formally,

$$\begin{aligned}\mathcal{R}^{F,\alpha}_{1OPQ}(p, p^N) = {} & \alpha p \mathbb{P}\left(\max_i\{V_i - p\} \geq \max\{V^{[N]} - p^N, 0\} | \text{Pessimistic}\right) \\ & + \alpha p^N \mathbb{P}\left(V^{[N]} - p^N > \max_i\{V_i - p, 0\} \cap V^{[N]} - p^N \geq 0\} | \text{Pessimistic}\right)\end{aligned}$$

$$
\begin{aligned}
&+(1-\alpha)p\mathbb{P}\left(\max_i\{V_i-p\}\geq\max\{V^{[N]}-p^N,0\}|\text{Risk Neutral}\right)\\
&+(1-\alpha)p^N\mathbb{P}\left(V^{[N]}-p^N>\max_i\{V_i-p,0\}\cap V^{[N]}-p^N\geq 0\}|\text{Risk Neutral}\right)\\
&=E_{X_\alpha}\left[p\mathbb{I}\left(\max_i\{V_i-p\}\geq\max\{V^{[N]}-p^N,0\}|X_\alpha\right)\right.\\
&\left.+p^N\mathbb{I}\left(V^{[N]}-p^N>\max_i\{V_i-p,0\}\cap V^{[N]}-p^N\geq 0\}|X_\alpha\right)\right]
\end{aligned}
$$

and $\mathcal{R}^{F,\alpha}_{1OPQ}=\max_{p,p^N}\mathcal{R}^{F,\alpha}_{1OPQ}(p,p^N)$

The first equation is derived by considering four events that generate revenue. The first event is that a customer is pessimistic, which has probability α, and has a valuation for their favorite item, $\max_i\{V_i-p\}$, which is larger than the opaque product and 0, $\max\{V^{[N]}-p^N,0\}$. In this event, the customer buys an item and the seller earns p. The remaining events are similar and enumerate the other cases where an opaque product is purchased and/or the customer is risk-neutral.

In the general *opaque selling model* (OPQ), the seller offers all possible opaque products, alongside the items which are offered at a single price p. For simplicity, tractability, and impartiality, opaque products of the same cardinality are assigned the same price (see Figure 3 for an example of this exact scenario). That is for all $S,S'\in 2^{[N]}$ s.t. $|S|=|S'|\geq 2$, the opaque products corresponding to the subsets S and S' must have the same price. For subsets of size k, the corresponding price is p^k, and the vector of the $N-1$ opaque product prices is denoted by $\vec{p}$. We denote $\mathcal{R}^{F,\alpha}_{OPQ}(p,\vec{p})$ and $\mathcal{R}^{F,\alpha}_{OPQ}$ as the expected revenue under F, in α-mixed markets, using prices $(p,\vec{p})$ and the optimal pricing, respectively. More formally,

$$
\begin{aligned}
&\mathcal{R}^{F,\alpha}_{OPQ}(p,\vec{p})=E_{X_\alpha}\left[p\mathbb{I}\left(\max_{i\in[N]}\{V_i-p\}\geq\max_{S\in 2^{[N]},|S|\geq 2}\{V^S-p^{|S|},0\}|X_\alpha\right)\right.\\
&\left.+\sum_{S\in 2^{[N]},|S|\geq 2}p^{|S|}\mathbb{P}\left(V^S-p^{|S|}\geq\max_{S'\in 2^{[N]},|S'|\geq 2}\{V^{S'}-p^{|S'|},0\}\cap V^S-p^{|S|}\geq\max_{i\in[N]}\{V_i-p\}|X_\alpha\right)\right]
\end{aligned}
$$

and $\mathcal{R}^F_{OPQ}=\max_{p,\vec{p}}\mathcal{R}^F_{OPQ}(p,\vec{p})$.

In this expression, the expectation is taken with respect to the customer type, which affects the opaque product valuations. The first summand corresponds to the revenue in the case where an item is bought, and the remaining summands corresponds to the revenues of the opaque products. Note an opaque product is sold only if it provides nonnegative utility and has more utility than all of the items and other opaque products.

We note that implementing the OPQ strategy in practice is simple, despite the exponentially large number of products (see Figure 3). The seller simply displays the price ladder corresponding to the size of the opaque product purchased, and then users simply select (click) their top k products if they chose to purchase an opaque product of size k. We also note that OPQ and 1OPQ are equivalent when $N = 2$. For readability we often omit the superscripts F and α when they can be inferred from context. In general, subscripts always refer to item prices (p_i) and superscripts refer to opaque product prices ($p^{|S|}$).

2.3. Valuation Distributions

In this work we focus on the class of *exchangeable* valuation distributions, which is a generalization of i.i.d. valuations that allows for symmetric correlation between items.

Definition 1. We call the random valuation vector $V = (V_1, \ldots, V_N)$ *exchangeable* if every permutation of the item valuations results in the same joint distribution. The corresponding distribution is also said to be exchangeable in this case. □

Exchangeable valuation distributions are a natural model for horizontally differentiated items as they allow for individual preferences between items, but enforce a distributional symmetry as the items are all alike. One important example is the Hotelling model which has been the primary focus of previous works (Fay and Xie (2008), Jerath et al. (2010)) which focus on scenarios with two items. Salop's circle is a generalization of the Hotelling model for more than two items, and is a standard choice model for capturing horizontal differentiation (Salop (1979), Fay and Xie (2008)). When $N = 3$, Salop's circle model is an exchangeable distribution. For $N \geq 4$, a more general, but complex, notion of exchangeability is needed to capture Salop's circle. We provide this definition in Appendix C and note that many of our results extend to this more general definition. For ease of exposition, we shall focus on Definition 1 throughout the paper.

3. The Power of Opaque Products

In this section, we focus on the revenue from the general opaque strategy (OPQ) when item valuations are drawn from an exchangeable distribution. In Sections 3.1 and 3.2, we provide conditions for when the expected revenue of OPQ is guaranteed to exceed that of discriminatory pricing (DP). When neither of these conditions hold, there is no dominance in either direction, and we supply counterexamples where discriminatory pricing is better. In Section 3.3, we quantify how much more revenue OPQ selling strategies can potentially earn over single pricing (SP), and show that the extra revenue garnered by OPQ strategies can be on the order of αN more than DP. In the special case where item valuations are i.i.d., we show this gap collapses to a constant factor.

3.1. Benchmarking against Discriminatory Pricing

We now give sufficient conditions for when OPQ is guaranteed to garner more revenue than DP. In particular, we show that when all customers are pessimistic or when valuations can take only two values (high or low), opaque selling is guaranteed to earn more revenue than discriminatory pricing. In Example 1, we give a valuation distribution where neither condition holds and $\mathcal{R}_{DP} > \mathcal{R}_{OPQ}$. This counterexample assumes valuations can take three values, and assumes that $0 \leq \alpha \leq .846$. Next, we formally state our result in Theorem 1 and defer the proof to Section 3.2.

THEOREM 1 **(When OPQ dominates DP)**. *Suppose customers are α-mixed and draw their valuations from an exchangeable distribution. If (i) $\alpha = 1$ or (ii) the item valuations take only two values, then opaque selling dominates discriminatory pricing, i.e.,*

$$\mathcal{R}_{OPQ} \geq \mathcal{R}_{DP}.$$

Interpretation and Implications of Theorem 1: While restricted, both cases of Theorem 1 where the dominance result holds represent situations of significant interest. When α is near 1, most customers assume a worst-case behavior with respect to opaque product allocation. This situation may arise in markets where opaque products have been recently introduced and there is no information for customers to be had. For markets where customers tend to view a particular item (color or destination) as unacceptable, this pessimistic behavior may also be common. A trivial implication of Theorem 1(i) is that $\mathcal{R}_{OPQ} \geq \alpha\mathcal{R}_{DP}$, which follows from simply ignoring the revenue from all risk-neutral customers. Thus in highly pessimistic markets where α is close to 1, an opaque selling strategy is guaranteed to preserve almost all the gains from discriminatory pricing, and potentially earn even more.

Further, we show in Corollary 2 that Theorem 1(i) extends to another important class of distributions for horizontally differentiated items known as Salop's circle, often used as a standard tool in the literature. We provide a short primer on Salop's circle model along with the proof of Corollary 2 in Appendix C.

When valuations are supported on two points, the market is highly differentiated and has binary 'high/low' valuations for the items. This setting has been the subject of Fay and Xie (2008), Huang and Yu (2014) in the literature on opaque selling. Note that the case of binary valuations is exactly when discriminatory pricing is most profitable compared to single pricing: Dutting and Klimm (2016) shows that for every N, there exists an i.i.d. two point distribution such that $\mathcal{R}_{DP} = \left(2 - \frac{1}{N}\right)\mathcal{R}_{SP}$ and that this is the largest possible revenue gap. Thus in markets where retailers would be most inclined to consider discriminatory pricing strategies, an opaque selling strategy is even more profitable.

It is important to note that when the conditions of Theorem 1 do not hold, that either DP or OPQ may be preferred depending on the market assumptions. Thus it is worth noting that OPQ may have other advantages over DP. For example, discriminatory selling can be unnatural and undesirable for customers, particularly in the settings we consider where the items only differ superficially. Charging different prices for what are essentially equivalent products may increase revenue, however it may be perceived as unfair by customers (and cause strategic behavior) or even disallowed by manufacturers altogether. In contrast, the OPQ strategy is impartial and will never result in a customer paying more simply for preferring a particular item (color). Collectively, we believe these arguments show that opaque selling should always be considered as an alternative to discriminatory pricing, and in many cases may result in more profit.

When the Conditions in Theorem 1 Are Not Met. Both the exchangeability and sufficient conditions ($\alpha = 1$ or two point valuations) for Theorem 1 are critical for the result to hold. In Example 1, we construct a simple three-point distribution from which item valuations are drawn i.i.d. and $\mathcal{R}_{DP} > \mathcal{R}_{OPQ}$ for any $\alpha \leq .846$. Thus Example 1 precludes generalizing Theorem 1 to situations beyond two point valuations and purely pessimistic markets. Surprisingly, it also implies that when $\alpha = 1$, the revenue from OPQ may be higher than when $\alpha = 0$. In other words, the seller may actually *benefit* from customers adopting a pessimistic attitude towards opaque products, as this helps segment the market more favorably. In Example 2, we describe a valuation distribution that is not exchangeable and where $\mathcal{R}_{DP} > \mathcal{R}_{OPQ}$

Example 1 (When Assumptions (i) and (ii) Do Not Hold). For $N = 2$, we construct a three point distribution F where, when customers are risk-neutral, the optimal discriminatory selling strategy earns strictly more revenue than an opaque strategy. Let $\alpha = 0$ and suppose i.i.d valuations for two items drawn according to,

$$V_i = \begin{cases} 0 & : \text{w.p } 8/27 \\ .1 & : \text{w.p } 2/3 \\ .9 & : \text{w.p } 1/27 \end{cases}$$

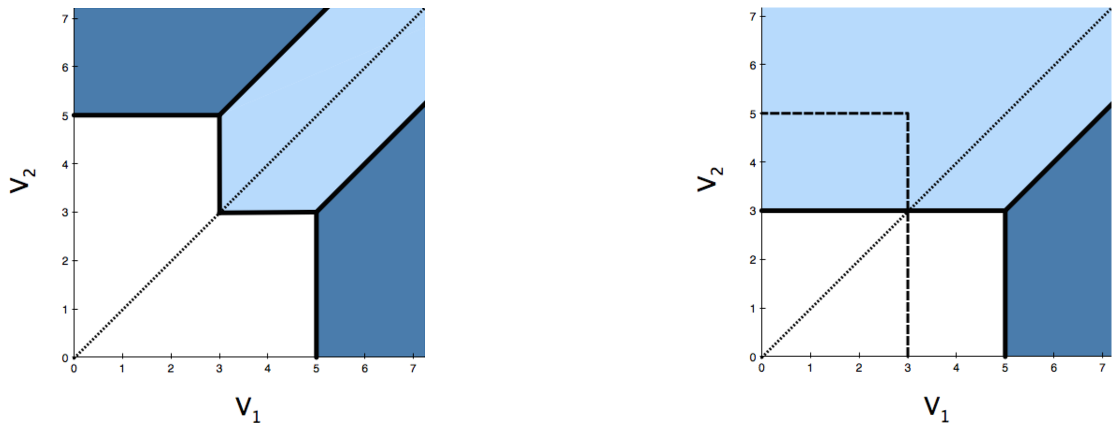
Then using the algorithm described in Theorem 7 we can compute $\mathcal{R}_{SP} = \mathcal{R}_{OPQ} = 0.091220... < 0.0913$ achieved by pricing both items at 0.1. However $\mathcal{R}_{DP}(0.9, 0.1) = 0.1$, 11% more revenue than the optimal opaque selling strategy.

Further, it can be computed that when $\alpha > .69231$, the optimal opaque pricing switches from (.1,.1) to a mixed pricing (.1,.9) earning revenue $\approx \alpha(1.08779) + (1-\alpha)(.0517146)$. The revenue from this optimal mixed opaque selling strategy overtakes the revenue from discriminatory pricing when $\alpha > .846$. Thus for $\alpha < .69231$, $\mathcal{R}_{SP} = \mathcal{R}_{OPQ} < \mathcal{R}_{DP}$. For $\alpha \in (.69231, .846)$, $\mathcal{R}_{SP} < \mathcal{R}_{OPQ} < \mathcal{R}_{DP}$. For $\alpha > .846$, $\mathcal{R}_{SP} < \mathcal{R}_{DP} < \mathcal{R}_{OPQ}$. □

Example 2 (When Exchangeability Does Not Hold). Consider a market where $N = 2$ and $\alpha = 1$, and where valuations for two items are drawn independently from V_1, which is two times a Bernoulli r.v. with probability 1/2, and V_2 which is distributed as a Bernoulli r.v. with probability 1/2. Since V_1 and V_2 are independent but not identical, the market is therefore not exchangeable. However, by a simple enumeration, one can see that $\mathcal{R}_{DP}(2, 1) = \frac{5}{4}$ whereas $\mathcal{R}_{1OPQ} \leq 1$. □

Geometric Proof of Theorem 1 when $N = 2$ and $\alpha = 1$. Before delving into the formal proof in Section 3.2, we provide some geometric intuition in the special case when $N = 2$ and $\alpha = 1$. Suppose the optimal discriminatory pricing uses prices (p_1, p_2) with $p_1 > p_2$. We show that an opaque selling strategy with prices $(p, p^2) = (p_1, p_2)$ exceeds the revenue of the optimal discriminatory pricing. Fig. 5a and Fig. 5b show the different purchase behaviors under OPQ and DP, respectively, where a darker color corresponds to a more expensive customer purchase. Due to exchangeability, it is then visually clear that the following are all equal: (i) the revenue of OPQ conditioned on $V_1 \geq V_2$, (ii) the revenue of OPQ conditioned on $V_2 \geq V_1$, and (iii) the revenue of DP conditioned on $V_1 \geq V_2$. To complete the proof, we claim that the revenue of DP conditioned on the event $V_1 \geq V_2$ is at least the revenue of DP conditioned on the event $V_2 \geq V_1$. If this were not the case, then reducing p_1 to p_2 would increase the revenue in the event that $V_1 \geq V_2$ without changing the revenue in the event $V_2 \geq V_1$, which would contradict the optimality of (p_1, p_2).

Figure 5 Geometric proof of Theorem 1(i)



(a) Valuation space for pessimistic customers facing a OPQ strategy with $(p, p^2) = (5, 3)$. Note the purchasing behavior is symmetric across the line $V_2 = V_1$.

(b) Valuation space for a customer facing a DP strategy with $(p_1, p_2) = (5, 3)$. Note that below the line $V_2 = V_1$, the purchasing behavior in (a) and (b) are identical.

Note. The valuation space and purchasing behaviors for a pessimistic customer facing OPQ and DP selling strategies respectively. Customers with valuations in the darkened regions buy at price 5 in both figures. Customers with valuations in the lightly shaded regions buy at price 3 (i.e., purchase the opaque product or item 2, respectively). Customers in the unshaded region do not purchase.

One interesting consequence of this geometric argument is that, when N=2 and $\alpha = 1$, $\mathcal{R}_{DP} \leq \frac{\mathcal{R}_{OPQ}+\mathcal{R}_{SP}}{2}$. Suppose that $\mathcal{R}_{DP} = (1+\gamma)\mathcal{R}_{SP}$, for some $\gamma > 0$. Then rearranging $\mathcal{R}_{DP} \leq \frac{\mathcal{R}_{OPQ}+\mathcal{R}_{SP}}{2}$ gives

$$\mathcal{R}_{OPQ} \geq \frac{1+2\gamma}{1+\gamma}\mathcal{R}_{DP}$$

which implies the inequality in Theorem 1 is strict whenever $\mathcal{R}_{DP} > \mathcal{R}_{SP}$. In Corollary 1 we expound on this observation to show more generally, whenever the conditions of Theorem 1(i) are met, and $\mathcal{R}_{DP} > \mathcal{R}_{SP}$, it follows that $\mathcal{R}_{OPQ} > \mathcal{R}_{DP}$.

When the conditions of Theorem 1(ii) hold, no such result can be shown. When α is small, there are cases when $\mathcal{R}_{DP} = \mathcal{R}_{OPQ}$ even if $\mathcal{R}_{DP} > \mathcal{R}_{SP}$, see Fig. 6c for an example. Instead we show an analogous result for when the conditions of Theorem 1(ii) are met, $\mathcal{R}_{DP} > \mathcal{R}_{SP}$, *and when α is sufficiently large*, it follows that $\mathcal{R}_{OPQ} > \mathcal{R}_{DP}$. The proof can be found in Appendix I.1.

COROLLARY 1. *Suppose customers are α-mixed and draw their valuations from an exchangeable distribution. Suppose DP earns more than SP and let $\gamma > 0$ denote the gap, i.e., $\mathcal{R}_{DP} = (1+\gamma)\mathcal{R}_{SP}$.*

(i) If $\alpha = 1$*, then*

$$\mathcal{R}_{OPQ} > \frac{1 + \frac{N}{N-1}\gamma}{1+\gamma}\mathcal{R}_{DP}.$$

(ii) If the valuations are supported on two points and $\alpha \geq 1 - \frac{\gamma}{\gamma N + N - 1}$*, then*

$$\mathcal{R}_{OPQ} > \alpha\left(1 + \frac{\gamma}{(N-1)(1+\gamma)}\right)\mathcal{R}_{DP}.$$

3.2. Proof of Theorem 1.

Proof of Theorem 1. We will consider the two cases separately.

Case (i): Let $\alpha = 1$, F be an exchangeable distribution over N items, and w.l.o.g. let $p_1 \geq p_2 \geq \ldots \geq p_N$ be the optimal prices corresponding to $\mathcal{R}_{DP}$. For ease of exposition we assume F is continuous and ignore ties, although the same argument follows when F is not continuous and one carefully considers the tie-breaking procedure. Let Σ be the set of permutations $\sigma : [N] \to [N]$, and let $\sigma(i)$ be the mapping of index i under σ. For every $\sigma \in \Sigma$, define the event $E_\sigma := \{V_{\sigma(1)} \geq V_{\sigma(2)} \geq \ldots \geq V_{\sigma(N)}\}$. Note that each $\{E_\sigma\}_{\sigma\in\Sigma}$ is equally likely by exchangeability. We define $q_{i|\sigma}$ to be the probability of a customer purchasing i under the DP strategy $(p_1, \ldots, p_N)$ conditioned on the event E_σ. We define $Rev(p_1, \ldots, p_N|\sigma)$ to be the expected revenue of the DP strategy conditioned on the event E_σ, i.e.,

$$Rev(p_1, \ldots, p_N|\sigma) = \sum_{i=1}^{N} p_i q_{i|\sigma}.$$

Define σ^* such that $Rev(p_1, \ldots, p_N|\sigma^*) \geq Rev(p_1, \ldots, p_N|\sigma)$ over all $\sigma \in \Sigma$, i.e., E_{σ^*} is the event that leads to the most revenue. This implies that $Rev(p_1, \ldots, p_N|\sigma^*) \geq \mathcal{R}_{DP}$.

Now consider an opaque selling strategy OPQ that uses prices $p^i = p_{\sigma^*(i)}$. (Note that p^1 is the price of the items.) We shall show that this opaque strategy has expected revenue of at least $Rev(\vec{p}|\sigma^*)$. Under our opaque strategy, we call the probability of a customer buying an opaque product of size i to be q^i and the probability of a customer buying an item to be q^1. We let $V^{(i)}$ be the i^{th} order statistic such that $V^{(1)} = \max_i\{V_i\}$ and $V^{(N)} = \min_i\{V_i\}$. We now show that $q^i = q_{\sigma^*(i)|\sigma^*}$ for all i by

$$\begin{aligned} q^i &= \mathbb{P}\left(\max_{S,|S|=i}\{V^S - p^i\} \geq \max_{j\neq i,|S'|=j}\{V^{S'} - p^j, 0\}\right) \\ &= \mathbb{P}\left(V^{(i)} - p^i \geq \max_{j\neq i}\{V^{(j)} - p^j, 0\}\right) \end{aligned}$$

$$= \mathbb{P}\left(V^{(i)} - p^i \geq \max_{j \neq i} \{V^{(j)} - p^j, 0\} | E_{\sigma^*} \right)$$
$$= \mathbb{P}\left(V_{\sigma^*(i)} - p_{\sigma^*(i)} \geq \max_{j \neq i} \{V_{\sigma^*(j)} - p_{\sigma^*(j)}, 0\} | E_{\sigma^*} \right)$$
$$= q_{\sigma^*(i)|\sigma^*}.$$

The first equality follows from the definition of OPQ strategies and q^i. The second equality follows from noting that a customer only needs to consider the best opaque product of each possible size $i = 2, \ldots, N$ and the best item. The best opaque product of size i has a valuation of the minimum of the top i valuations, which is the i'th order statistic. The third equality follows from the fact that the valuations are exchangeable, and thus an event on the order statistics is independent of E_σ for all $\sigma \in \Sigma$. The fourth equality follows from our pricing rule and the definition of σ^*. The last equality follows from the definition of $q_{i|\sigma^*}$. Combining our findings yields

$$\mathcal{R}_{OPQ} \geq \mathcal{R}_{OPQ}(p^1, \ldots, p^N) = \sum_{i=1}^{N} p^i q^i = \sum_{i=1}^{N} p_{\sigma^*(i)} q_{\sigma^*(i)|\sigma^*} = Rev(p_1, \ldots, p_N | \sigma^*) \geq \mathcal{R}_{DP}. \quad \square$$

Case (ii): Fix a distribution F supported on two points $\{a, b\}$ where $a < b$. Note for such distributions, the optimal discriminatory pricing uses prices $\vec{p} = (a, a, \ldots, a)$, $(b, b, \ldots, b)$ or a mixed pricing where exactly one price (since F is exchangeable it doesn't matter which price) is low $(a, b, b, \ldots, b)$. If either $(a, a, \ldots, a)$ or $(b, b, \ldots, b)$ are the the optimal discriminatory pricing given F, then $\mathcal{R}_{SP} = \mathcal{R}_{DP}$ and the claim follows automatically. Suppose $\mathcal{R}_{DP} > \mathcal{R}_{SP}$, then the optimal pricing is the mixed strategy and under a mixed pricing, a discriminatory selling strategy always sells an item. Further we will restrict ourselves to opaque pricings where $p^N = a$, and thus always sell the item. Since the item is always sold in both strategies, we may normalize the support of F to $\{1, 1 + \delta\}$ without changing the ratio $\frac{\mathcal{R}_{DP}}{\mathcal{R}_{OPQ}}$. Now let U be a random variable representing the number of valuations that are equal to $1 + \delta$. When $U = 0$, DP earns revenue of 1. When $U = i \geq 1$, DP earns revenue of $1 + \delta$ with probability $\frac{\binom{N-1}{i}}{\binom{N}{i}} = \frac{N-i}{N}$ and 1 otherwise. (The customer buys the cheap item when they value it at $1 + \delta$.) Then for $i \geq 1$,

$$E[\mathcal{R}_{DP} | U = i] = 1 + \frac{N-i}{N}\delta. \tag{1}$$

Consider the following opaque pricing where for $i \in [N]$ we let $p^i = 1 + \frac{N-i}{N}\delta$. When $U = 0$, the customer buys the opaque product of size N at price 1, paying the same in the corresponding case in DP. When $U = i \geq 1$, we claim that regardless of whether the customer is pessimistic or risk-neutral, they will purchase an opaque product of size i (or item if $i = 1$) earning revenue $1 + \frac{N-i}{N}\delta$, which is the same revenue in the corresponding case in DP and therefore would complete the proof. First suppose the customer is pessimistic, then when $U = i$ the customer values the size i product as $1 + \delta$ and garners utility $\frac{i}{N}\delta$. For $j < i$, the customer values the opaque product the same but has to pay a higher price, while for $j > i$ the customer values the opaque product at 1 and does not buy. Thus a pessimistic customer yields revenue $1 + \frac{N-i}{N}\delta$ when $U = i$.

When the customer is risk-neutral and $U = i$, they again value the size i product as $1+\delta$ and garners utility $\frac{i}{N}\delta$ for purchasing it. Products of size $j < i$ have the same valuation, but at a higher price, and thus offer less utility. For products of size $j > i$, the utility of the size j opaque product is

$$\frac{i(1+\delta) + (j-i)\cdot 1}{j} - (1 + \frac{N-j}{N}\delta) = \left(\frac{i}{j} - \frac{N-j}{N}\right)\delta$$

which is strictly less than $\frac{i}{N}\delta$. Finally, the above expression also shows that the utilities of the opaque products of size i and N are the same, in which case the customer buys i (since we have assumed w.l.o.g. that ties are broken in favor of the more expensive option). Thus both pessimistic and risk-neutral customers have the same purchase behavior under this opaque pricing, and yield the same expected revenue as $\mathcal{R}_{DP}$. $\square$

In the proof of Theorem 1(i), we essentially view opaque selling with pessimistic customers as discriminatory pricing where the ordering of the item valuations is known a priori to the seller. That is, the valuations for the best item and opaque products take on exactly the valuations of the original items - only the assignment of valuations to items/products differs between the two strategies. It is then tempting to assume that OPQ is trivially more profitable than DP, where the ordering of the item valuations is not known to the seller and hence "less information" is available. Unfortunately this line of argument would just as easily extend to other settings where our result does not hold, specifically settings where valuations are not exchangeable (see Example 2 for counterexample). The extra information offered by OPQ comes with an additional constraint: w.l.o.g. the highest valued item is sold at the highest price, the second highest valued item is sold at the second highest price, and so on, which need not be optimal.

3.3. Benchmarking against Single Pricing

In this section, we seek to quantify the potential gains that opaque selling offers over a simple single pricing strategy. This question has also been studied in the context of discriminatory pricing. For example, when valuations are drawn i.i.d., Chawla et al. (2007) shows that discriminatory pricing can earn at most $2 - \frac{1}{N}$ more than single price strategies, and Dutting and Klimm (2016) shows that this bound is tight. Interestingly, in the same setting of i.i.d. valuations, opaque selling can also earn up to a constant factor more than single pricing. In Theorem 2, we describe this upper bound as a function of α and N. A direct consequence of this theorem is that when valuations are i.i.d., OPQ and DP are always within a constant factor of each other. We defer the proof in Appendix I.2.

THEOREM 2 **(Revenue Upper Bound when Valuations are I.I.D.)**. *Suppose customers are α-mixed and their item valuations are i.i.d. Then,*

$$\mathcal{R}_{OPQ} \leq \left(3 + (1-\alpha)\left(1 - \frac{2}{N}\right)\right)\mathcal{R}_{SP}.$$

In the more general case of exchangeable distributions, no results comparing DP to SP are available to the best of our knowledge. In Theorem 3, we show that DP earns at most $1 + \log(N)$ more than SP, while OPQ can earn up to and at most N times more than SP. This implies that OPQ can earn up to $O\left(\frac{N}{\log(N)}\right)$ more revenue than DP, which we also show is indeed possible in Theorem 3 and Example 4. We defer the proof to Appendix I.5

THEOREM 3 **(Revenue Upper Bound when Distribution is Exchangeable)**. *Suppose customers are α-mixed and draw their valuations from an exchangeable distribution. Then, (i) $\mathcal{R}_{DP} \leq (1 + \log(N))\mathcal{R}_{SP}$, (ii) $\mathcal{R}_{OPQ} \leq N\mathcal{R}_{SP}$, and (iii) there exists a distribution F such that $\mathcal{R}_{OPQ} \geq \frac{\alpha}{2}\frac{N}{1+\log(N)}\mathcal{R}_{DP}$.*

4. The Power of One Opaque Product

In this section, we study the revenue gained by using a strategy with a single opaque product (1OPQ), where the seller offers all N items at a single price alongside a single opaque product corresponding to the set $[N]$. 1OPQ represents the easiest use-case for opaque selling, simply offering one opaque option made up of all N items. Fig. 1 shows an example of 1OPQ for staplers on Amazon.com.

We note that since the 1OPQ strategy only offers two prices, a comparison to discriminatory pricing which offers N prices becomes significantly more challenging. Nevertheless, we show in Section 4.1 that 1OPQ guarantees 71.9% of the revenue of DP in the special case of two-point distributions. In comparison to single pricing, we show that 1OPQ can earn at most a factor of $\left(2-(1-\alpha)\frac{1}{N}\right)$ more than SP in Section 4.2. When $N=2$, our upper bounds are tight and the revenue increase can be larger than that of DP.

4.1. Benchmarking against Discriminatory Pricing

In Theorem 4 we show that 1OPQ guarantees at least 71.9% of the revenue that DP provides when the distribution is exchangeable and valuations are supported on two points (low or high). As previously mentioned, such distributions are a natural and well-studied model of customers with binary preferences, and may be used to approximate bimodal distributions. Further, as seen in Example 5 and Chawla et al. (2007), two point distributions represent natural best cases for price discrimination for both 1OPQ and DP strategies. We emphasize that Theorem 4 is a strict improvement on the best approximation possible by $\mathcal{R}_{SP}$, which is $0.50\mathcal{R}_{DP}$ in this setting. Specifically, Chawla et al. (2007) give a two point distribution such that when scaling the number of items N, $\lim_{N\to\infty}\frac{\mathcal{R}_{SP}}{\mathcal{R}_{DP}}=.5$.

THEOREM 4 **(When 1OPQ Approximates DP)**. *Suppose customers are α-mixed and draw their valuations from an exchangeable distribution supported on two values, then*

$$\mathcal{R}_{1OPQ}\geq 0.719\mathcal{R}_{DP}.$$

Our proof follows from observing that when the probability of customers having high valuations is large, a single pricing strategy is a good approximation. Otherwise, if the probability of customers having high valuations is small, we show that augmenting single price strategies with a single opaque product is a good approximation of the optimal discriminatory pricing. We defer the complete proof to Appendix I.6.

4.2. Benchmarking against Single Pricing

In this section, we show that the addition of a single opaque product over the set $[N]$ can increase the revenue by at most $\left(2-(1-\alpha)\frac{1}{N}\right)$. Although our bound holds under all exchangeable distributions, Examples 5 and 6 shows that our analysis is tight in the special cases of $\alpha=0$ and $\alpha=1$, even when customers are restricted to have i.i.d. valuations. We defer the proof to Appendix I.7.

Theorem 5 **(Revenue Upper Bounds for 1OPQ)**. *Suppose customers are α-mixed and draw their valuations from any distribution. Then,*

$$\mathcal{R}_{1OPQ} \le \left(2-(1-\alpha)\frac{1}{N}\right)\mathcal{R}_{SP}$$

Further, when $\alpha = 0$ or $\alpha = 1$ there exists an i.i.d valuation distribution such that the bounds are tight.

Theorem 5 fully describes the possible revenue increase a seller could hope to garner using a single opaque product. It is of interest to note that when $N = 2$ and $\alpha = 1$, Theorem 5 implies the existence of a valuation distribution such that $\mathcal{R}_{1OPQ} = 2\mathcal{R}_{SP}$. However by Theorem 3(i), $\mathcal{R}_{DP} \le (1+\log(2))\mathcal{R}_{SP}$ for any distribution. Together these results show that 1OPQ can sometimes achieve higher revenue lifts than DP.

5. Numerical Experiments

In this section we conduct numerical experiments to explore the relationships between $\mathcal{R}_{SP}, \mathcal{R}_{1OPQ}, \mathcal{R}_{DP}$, and $\mathcal{R}_{OPQ}$ for various valuation distributions and under various market assumptions. To perform the experiments, we must solve for the optimal prices for any of these strategies. That is given the distribution F, and α, we must solve for the price vector that maximizes revenue. However, in general solving for the optimal prices in multi-item settings is quite difficult. Even in the special case when the valuations are i.i.d., solving for the optimal discriminatory pricing strategy is NP-Hard (Chen et al. (2018)).

Although not the focus of this work, as a prerequisite step in Appendix D we address this issue by developing a simple enumerative algorithm which is computationally efficient in the special case when the support of the valuations is discrete and the number of items is not large. Note that a simple brute force search over the support is not sufficient, as optimal prices do not necessarily lie on the support (Chawla et al. (2007), Chen et al. (2015, 2018)). Given any distribution F, our approach then will be to first discretize the distribution and then run Algorithm 1. We emphasize that carefully discretizing the support and then solving still yields near-optimal solutions for the true underlying distribution. Indeed, Hartline and Koltun (2005) show that when valuations are supported on $[l,h]$, restricting the price optimization to $\log_{1+\epsilon}\left(\frac{h}{l}\right)$ discrete price points results in prices that garner revenue within a factor of $1+\epsilon$ of the optimal revenue. Finally, when the number of items is large, we note that one can employ a standard MIP approach along the lines of (Hanson and Martin (1990), see Appendix E for a complete formulation of the MIP) however such an approach runs in exponential time in both m and n in the worst case.

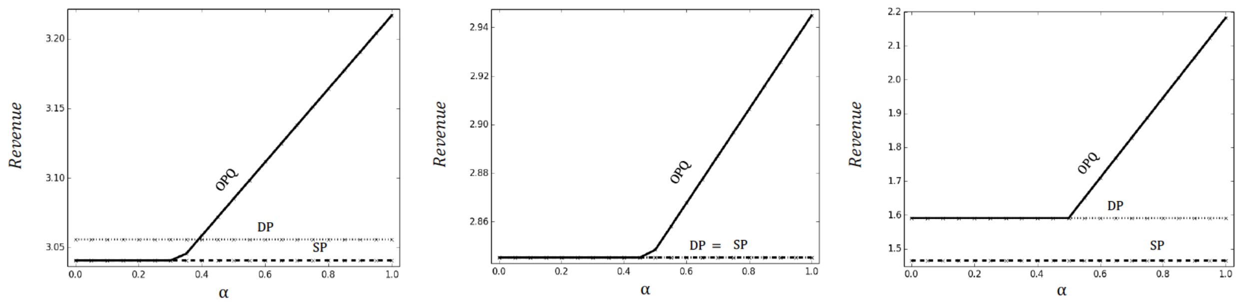
In what follows we conduct our numerical experiments. In Section 5.1, we introduce the valuation distributions we study in this section and perform computations to establish baseline comparisons between $\mathcal{R}_{SP}, \mathcal{R}_{1OPQ}, \mathcal{R}_{DP}$, and $\mathcal{R}_{OPQ}$. In Section 5.2 we show how opaque selling strategies perform when the model is misspecified. Specifically, we first consider the case when the percentage of pessimistic customers in the market (α) is unknown, and demonstrate that α can be effectively estimated from sales data such that the resulting opaque selling model garners almost all of the available revenue. Then we consider a case when customers are neither pessimistic nor risk-neutral, but something in between, and again show fitting our original model and optimizing results in an opaque selling strategy that extracts nearly all of the revenue. Finally, in Appendix G, we conduct two additional experiments to study the efficacy of opaque selling as the variance of the valuation distribution changes and when valuations are sampled from multiple exchangeable distributions.

5.1. Baseline Computational Results

In this section we set up a baseline for numerical experiments on three typical valuation distributions which bear out the relationships between $\mathcal{R}_{SP}, \mathcal{R}_{1OPQ}, \mathcal{R}_{DP}$, and $\mathcal{R}_{OPQ}$ that we have studied in the previous sections. We shall assume item valuations are drawn i.i.d. from the following distributions: *(i)* a triangular distribution supported on $[1,7]$ with mode 3, *(ii)* a normal distribution with mean 3 and standard deviation 2 truncated on $[1,7]$, and *(iii)* a Bernoulli distribution supported on $\{1,7\}$ with probability of a 7 being 1/9. These three valuations distributions are natural and widely studied, while also varied enough to showcase the differences between the pricing strategies. In order to apply Theorem 7, we discretize these distributions by rounding valuations to their nearest integer value. We compute the revenue of SP, 1OPQ, DP, and OPQ for every value of α from 0 to 1 in increments of 0.05. Note that the revenue of SP and DP do not change with α, which is a parameter for the mixture of pessimistic and risk-neutral customers with respect to opaque products. Fig. 6 displays our results when $N = 2$ (in which case OPQ is equivalent to 1OPQ) and Fig. 7 displays our results when $N = 3$.

Each of the three distributions we study result in fundamentally different behaviors. In Fig. 6(a) we note that $\mathcal{R}_{DP} > \mathcal{R}_{SP}$, and further when $\alpha < 0.3$, $\mathcal{R}_{DP} > \mathcal{R}_{OPQ}$. However as α increases towards one, the relationship between $\mathcal{R}_{DP}$ and $\mathcal{R}_{OPQ}$ reverses. When α is

Figure 6 Revenue in simulated markets when $N = 2$.



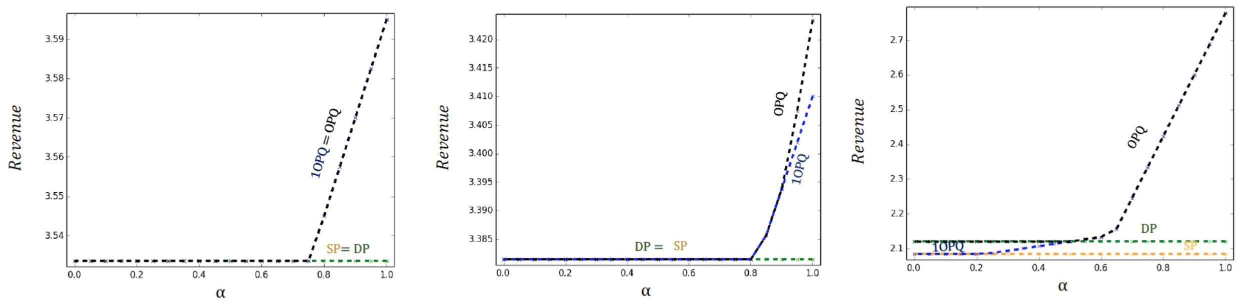
(a) Triangular distribution. (b) Truncated normal distribution. (c) Bernoulli distribution.

Note. Illustrates the relationship between $\mathcal{R}_{SP}$ (dashed line), $\mathcal{R}_{DP}$ (dotted line), and $\mathcal{R}_{OPQ}$ (solid line) as the proportion of pessimistic customers increases.

close to 1, $\mathcal{R}_{OPQ}$ significantly outperforms $\mathcal{R}_{DP}$, garnering up to approximately 5% more revenue. In Fig. 6(b) we note that $\mathcal{R}_{SP} = \mathcal{R}_{DP}$, meaning that discriminatory pricing alone does not add value over a single price. However, OPQ can earn strictly more revenue than either strategy when $\alpha > 0.5$. Finally in Fig. 6(c), $\mathcal{R}_{OPQ} \geq \mathcal{R}_{DP}$ for any value of α, which is known directly from Theorem 1(ii). The gap is positive and increasing when $\alpha > 0.5$, which is implied by Corollary 1. Interestingly, for all three distributions $\mathcal{R}_{OPQ}$ is a non-decreasing function in α. This is counter-intuitive: as the number of pessimistic customers increase, more customers have lower values for the opaque products but the overall revenue from OPQ increases. This suggests that the revenue non-monotonicity noted in Section 3.2 is quite pervasive. We believe that this occurs since pessimistic behavior results in the customers being much more separated naturally than risk-neutral customers. This helps ensure that the opaque products are not cannibalizing the full priced traditional products too much. One practical insight from this is that retailers should not reveal allocation probabilities to consumers, as knowledge of such allocation probabilities would more likely result in more risk-neutral customers and thus less revenue.

In Fig. 7(a), we see that that lifting the problem from $N = 2$ to $N = 3$ collapses the revenue gap between $\mathcal{R}_{SP}$ and $\mathcal{R}_{DP}$, but does not diminish the impact of opaque products. Further we note that a single opaque product performs just as well as the general opaque strategy does. In Fig. 7(b), we observe that when $\alpha > 0.8$, 1OPQ and OPQ can outearn DP. As α approaches 1, eventually there is a revenue gap between 1OPQ and OPQ. Finally in Fig. 7(c), we see that when $\alpha < 0.5$, DP and OPQ are equivalent and outperform

Figure 7 Revenue in simulated markets when $N = 3$.



(a) Triangular distribution. (b) Truncated normal distribution. (c) Bernoulli distribution.

Note. The relationship between $\mathcal{R}_{SP}, \mathcal{R}_{DP}, \mathcal{R}_{1OPQ}$ and $\mathcal{R}_{OPQ}$ as the proportion of pessimistic customers increases.

1OPQ. When $\alpha > 0.5$, 1OPQ and OPQ are equivalent and outperform DP. These these experiments demonstrate a wide range of behavior, but show, generally, that OPQ and 1OPQ strategies tend to outperform DP almost always, and to improve as α increases.

In Appendix G, we vary the variance of the distributions and find that a low to medium amount of variance results in the most benefit for DP and OPQ, otherwise SP becomes optimal as variance becomes large.

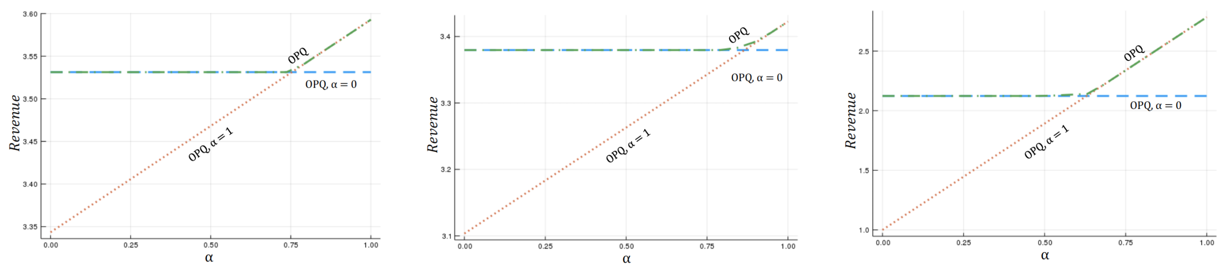
5.2. Robustness of the Model

In this subsection we conduct numerical experiments on the three valuation distributions described above to test the performance of opaque selling when the model is misspecified.

When α is misspecified. In Figs. 6 and 7 we compared the revenues of the opaque and discriminatory pricing strategies when the model is well-specified, i.e. the underlying valuation distribution is known and the mixture of pessimistic vs. risk-neutral customers, α, is explicitly known. In the next set of computations we compare the revenue of opaque strategies when α is misspecified.

Specifically, in Fig. 8 we plot the revenue of two opaque selling strategies trained under the assumption of $\alpha = 0$, and $\alpha = 1$, respectively, on markets where the true proportion of pessimistic customers α varies between zero and one. We find that prices trained under the assumption of risk-neutrality (OPQ, $\alpha = 0$) have stable performance regardless of α reflecting that fact that risk-neutral valuations are more tightly concentrated and thus lead to prices which are less finely tuned to exploiting differences in the valuations across the items. On the other hand, prices for opaque selling strategies trained under the pessimistic assumption (OPQ, $\alpha = 1$), are calibrated to capture differences in the valuations between

Figure 8 Revenue in simulated markets when $N = 3$ and α is misspecified.



(a) Triangular distribution. (b) Truncated normal distribution. (c) Bernoulli distribution.

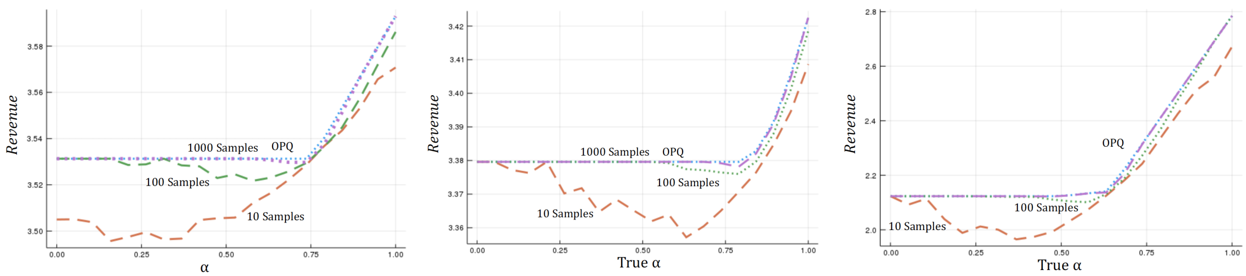
Note. Illustrates the relationship between the revenue of an opaque selling strategy fitted under the assumption that $\alpha = 0$ (dashed line) and $\alpha = 1$ (dotted line), as well as a correctly fitted opaque selling strategy (dashed-dotted line) as the true proportion of pessimistic customers increases.

items and only perform well when correctly specified. We observe that the performance of such strategies smoothly degrades as α tends to 0. Most importantly, we see that the best of the two strategies $\alpha = 0$ or $\alpha = 1$, nearly approximates the optimal opaque strategy. This suggests that α can be productively thought of as a binary parameter, and that effectively only two strategies need to be tested.

When α must be estimated from data. Given the sensitivity of opaque selling to the proportion of pessimistic customers in the market, a natural question then is whether α can be easily estimated from sample data. We assume the underlying valuation distribution F is known and focus on estimating α from purchase data. This situation arises for a seller who has been engaged in discriminatory pricing for a long time and thus has knowledge of F and the underlying choice model already. Knowing F, the seller switching to an opaque strategy and simply needs to estimate α.

In Fig. 9, we compare the revenues of opaque selling strategies when α is estimated from a fixed number of sample purchase histories given a fixed price $\vec{p}$. To estimate α we use a maximum likelihood estimate under a fixed test price, for details of the estimation procedure see Appendix F. As α ranges from 0 to 1, we compute opaque selling strategies by estimating α using a small number of randomly generated sample purchases (10), a medium number of sample purchases (100), and a large number (1000), and then we solve for the optimal opaque prices under the estimated α and plot its performance. We repeat the estimation of α and subsequent price computation 100 times for each true value of α. For plots of descriptions of the estimated α's themselves, see Appendix F.

Figure 9 Revenue in simulated markets when $N = 3$ and α is estimated.



(a) Triangular distribution. (b) Truncated normal distribution. (c) Bernoulli distribution.

Note. Illustrates the relationship between $\mathcal{R}_{OPQ}$ (dotted line) when the true value must be estimated from data. The dashed lines represent the average revenue obtained under an opaque selling strategy for α was estimated from ten samples (10 Samples), one hundred samples (100 Samples), and one thousand samples (1000 Samples).

We find that using only ten samples is enough to recover most of the performance of opaque selling in the cases that the true opaque selling revenue outperforms SP, but given the small number of samples, the opaque strategy trained on only ten samples may in fact lose revenue by overestimating the portion of risk-averse customers in the market. Increasing the number of samples to the still modest number of 100 eliminates most of this instability and recovers almost all the revenue of the optimal OPQ in each of the three markets. When the number of observed purchases histories is 1000, the performance of OPQ with an estimated α is nearly identical to the performance of the opaque selling strategy that knows the true value of α.

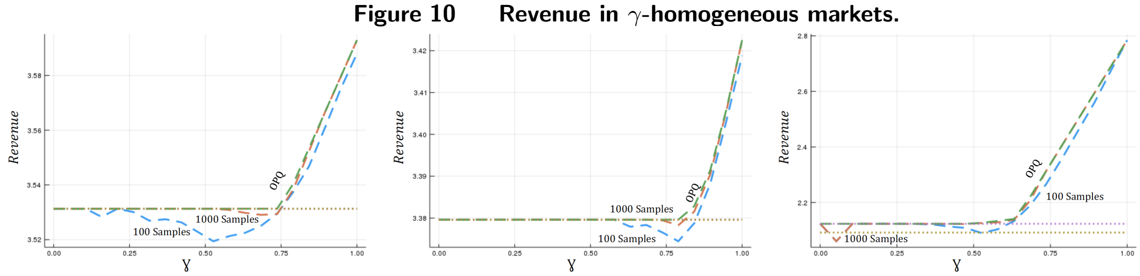
When customers are neither pessimistic or risk-neutral. Thus far, we assume customers value the opaque product either in a pessimistic or a risk-neutral fashion. One natural extension beyond our model is to imagine customers value the opaque product not merely as pessimistic or risk-neutral, but as some convex combination of those valuations representing the spectrum of ways individuals respond to risk. In the next set of experiments we consider the efficacy of decisions made under our model when in fact customers are neither pessimistic nor risk-neutral. Specifically, we call a market γ-*homogeneous* when all customers value an opaque product over the set of items S as a γ-convex combination of the minimum valued item and the average valuation for the items, i.e., they value an opaque product over items $S \subset [N]$ as

$$V^S(\gamma) = \gamma \min_{i \in S}\{V_i\} + (1-\gamma)\frac{\sum_{i \in S} V_i}{|S|}.$$

Note that a 1-mixed (0-mixed) and 1-homogeneous (0-homogeneous) market are equivalent.

To assess the efficacy of our model on γ-homogeneous markets we conduct the following experiment. As before, we assume knowledge of the underlying valuation distributions but not the proportion of pessimistic customers. Then we estimate α from sample purchases over the γ-homogeneous market using the same test price and procedure as above and in Appendix F, using either 100 or 1000 samples. Once α is estimated, we optimize our OPQ strategy with the misspecified α market to compute a price $\vec{p}$, and finally measure the performance of that price on the true γ-homogeneous market.

Figure 10 **Revenue in γ-homogeneous markets.**



(a) Triangular dist. (b) Normal dist. (c) Bernoulli dist.

Note. Illustrate the relationship between $\mathcal{R}_{OPQ}$ (OPQ) trained with full knowledge that the market is γ mixed, and three opaque strategies generated by fitting an α and computing a price under the assumption that the market is α mixed. The dashed lines represent the average revenue obtained on the γ mixed market, under an opaque selling strategy which estimated α from one hundred samples and one thousand samples.

In Fig. 10 we plot the results of our experiment and make the following findings. First, opaque selling (trained with full knowledge of customer preferences) continues to achieve significant gains over single price and discriminatory pricing strategies when the market is γ-homogeneous. Second fitting an opaque selling strategy according to the model studied in this paper, even if customers are γ-homogeneous and value the opaque goods in neither a pessimistic nor risk-neutral fashion, still achieves almost all of the gain of an optimal opaque strategy that knows the true model (c.f. Fig. 10).

6. Conclusion

In this paper, we studied opaque selling strategies in the context of selling horizontally differentiated items to unit-demand customers. We considered mixtures of two practical models of customer behavior corresponding to pessimistic and risk-neutral customers, motivated by the customer's lack of knowledge about how opaque products are allocated by the seller. When the valuation distribution is exchangeable and either customers are pessimistic or have binary preferences, we showed that opaque selling dominates discriminatory pricing. We also explicitly quantified the best possible revenue lift from using opaque products, which can be significantly higher than discriminatory pricing. Finally, we considered the practical case where only one opaque product is offered, and offered theoretical and numerical evidence of the strength of this simplified strategy.

We believe our results provide strong motivation for the use of opaque products as a vehicle for price discrimination, especially in online sales channels. Since our opaque model imposes a single price for opaque products of the same size, it is impartial to customers with particular preferences. It is also particularly advantageous in situations where discriminatory pricing could be effective, but disallowed due to business constraints and/or poor customer perception. It would interesting for future research to consider the impact of competition and finite inventory constraints on opaque selling, as well as behavioral studies for how consumers value opaque products in various markets.

References

Abdallah, Tarek, Arash Asadpour, Josh Reed. 2017. Large-scale bundle size pricing: A theoretical analysis .

Alaei, Saeed, Jason Hartline, Rad Niazadeh, Emmanouil Pountourakis, Yang Yuan. 2019. Optimal auctions vs. anonymous pricing. *Games and Economic Behavior* **118** 494 – 510.

Anderson, Simon P, Andre De Palma, Jacques François Thisse. 1992. *Discrete choice theory of product differentiation*. MIT press.

Balestrieri, F., S. Izmalkov. 2016. Informed seller in a hotelling market. *SSRN* URL `http://dx.doi.org/10.2139/ssrn.2398258`.

Balestrieri, Filippo, Sergei Izmalkov, Joao Leao. 2015. The market for surprises: selling substitute goods through lotteries .

Bergemann, Dirk, Francisco Castro, Gabriel Y Weintraub. 2020. Uniform pricing versus third-degree price discrimination .

Briest, P., S. Chawla, R. Kleinberg, S.M. Weinberg. 2015. Pricing lotteries. *Journal of Economic Theory*, vol. 156. 144–174.

Briest, P., H. Roglin. 2010. The power of uncertainty: Bundle-pricing for unit-demand customers. *Proceedings of the Workshop on Approximation and Online Algorithms (WAOA)*. 47–58.

Cai, Y., C. Daskalakis. 2011. Extreme value theorems for optimal multidimensional pricings. *Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. IEEE, 522531.

Celis, L Elisa, Gregory Lewis, Markus Mobius, Hamid Nazerzadeh. 2014. Buy-it-now or take-a-chance: Price discrimination through randomized auctions. *Management Science* **60**(12) 2927–2948.

Chawla, S., J. Hartline, R. Kleinberg. 2007. Algorithmic pricing via virtual valuations. *Proceedings of the 8th ACM Conference on Electronic Commerce (EC)*. 243–251.

Chawla, S., J. Hartline, D. Malec, B. Sivan. 2010. Mutli-parameter mechanism design and sequential posted pricing. *Proceedings of the Forty-Second ACM Symposium on Theory of Computing (STOC)*. 311–320.

Chawla, S., D. Malec, B. Sivan. 2015. The power of randomness in bayesian optimal mechanism design. *Games and Economic Behavior*, vol. 91. 297–317.

Chen, R., E. Gal-Or, P. Roma. 2014. Opaque distribution channels for competing service providers: Posted price vs. name-your-own-price mechanisms. *Operations research* **62**(4) 733–750.

Chen, X., I. Diakonikolas, A. Orfanou, D. Paparas, X. Sun, M. Yannakakis. 2015. On the complexity of optimal lottery pricing and randomized mechanisms. *the Proceedings of the 2015 IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*.

Chen, X., I. Diakonikolas, D. Paparas, X. Sun, M. Yannakakis. 2018. The complexity of optimal multidimensional pricing for a unit-demand buyer. *Games and Economic Behavior* **110** 139 – 164.

Dutting, P., M. Klimm. 2016. Revenue gaps for discriminatory and anonymous sequential posted pricing URL `arXiv:cs.GT/1607.07105`.

Elmachtoub, A., Y. Wei. 2015. Retailing with opaque products. *SSRN* URL `http://dx.doi.org/10.2139/ssrn.2659211`.

Elmachtoub, Adam N, Vishal Gupta, Michael Hamilton. 2018. The value of personalized pricing. *Available at SSRN 3127719* .

Elmachtoub, Adam N, David Yao, Yeqing Zhou. 2019. The value of flexibility from opaque selling. *Available at SSRN* .

Fay, S., J. Xie. 2008. Probabilistic goods: A creative way of selling products and services. *Marketing Science*, vol. 27. 674–690.

Fay, Scott, Jinhong Xie. 2014. Timing of product allocation: Using probabilistic selling to enhance inventory management. *Management Science* **61**(2) 474–484.

Gallego, G., G. Iyengar, R. Phillips, A. Dubey. 2004. Managing flexible products on a network .

Gallego, G., R. Phillips. 2004. Revenue management of flexible products. *Manufacturing & Service Operations Management*, vol. 6. 321–337.

Geng, Xin. 2016. Opaque selling in congested systems. *Operations Research Letters* **44**(6) 737–741.

Granados, Nelson, Kunsoo Han, Dan Zhang. 2018. Demand and revenue impacts of an opaque channel: Evidence from the airline industry. *Production and Operations Management* **27**(11) 2010–2024.

Hanson, Ward, R Kipp Martin. 1990. Optimal bundle pricing. *Management Science* **36**(2) 155–174.

Hart, S., N. Nisan. 2014. How good are simple mechanisms for selling multiple goods? *The Hebrew University of Jerusalem, Center for Rationality*. 893–906.

Hartline, Jason D., Vladlen Koltun. 2005. Near-optimal pricing in near-linear time. Frank Dehne, Alejandro López-Ortiz, Jörg-Rüdiger Sack, eds., *Algorithms and Data Structures*. Springer Berlin Heidelberg, Berlin, Heidelberg, 422–431.

Hartline, Jason D, Tim Roughgarden. 2009. Simple versus optimal mechanisms. *Proceedings of the 10th ACM conference on Electronic commerce*. 225–234.

Huang, T., Y. Yu. 2014. Sell probabilistic goods? a behavioral explanation for opaque selling. *Marketing Science* **33**(5) 743–759.

Huang, Xiao, Greys Sošić, Gregory Kersten. 2017. Selling through priceline? on the impact of name-your-own-price in competitive market. *IISE Transactions* **49**(3) 304–319.

Jerath, K., S. Netessine, S. Veeraraghavan. 2010. Revenue management with strategic customers: Last-minute selling and opaque selling. *Management Science*, vol. 56. 430–448.

Jiang, Y. 2007. Price discrimination with opaque products. *Journal of Revenue & Pricing Management*, vol. 6. 118–134.

Jin, Yaonan, Pinyan Lu, Qi Qi, Zhihao Gavin Tang, Tao Xiao. 2019. Tight approximation ratio of anonymous pricing. *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2019, Association for Computing Machinery, New York, NY, USA, 674685.

Ma, Will, David Simchi-Levi. 2016. Reaping the benefits of bundling under high production costs .

Post, D., M. Spann. 2012. Improving airline revenues with variable opaque products: "blind booking" at germanwings. *Interfaces*, vol. 42. 329–338.

Post, David. 2010. Variable opaque products in the airline industry: A tool to fill the gaps and increase revenues. *Journal of Revenue and Pricing Management* **9**(4) 292–299.

Salop, Steven C. 1979. Monopolistic competition with outside goods. *The Bell Journal of Economics* **10**(1) 141–156. URL `http://www.jstor.org/stable/3003323`.

Shapiro, Dmitry, Xianwen Shi. 2008. Market segmentation: The role of opaque travel agencies. *Journal of Economics & Management Strategy* **17**(4) 803–837.

Xiao, Yongbo, Jian Chen. 2014. Evaluating the potential effects from probabilistic selling of similar products. *Naval Research Logistics (NRL)* **61**(8) 604–620.

Xie, Xiaoqing Kristine, Rohit Verma, Chris K Anderson. 2016. Demand growth in services: A discrete choice analysis of customer preferences and online selling. *Decision Sciences* **47**(3) 473–491.

Xu, Xiaoya, Zhaotong Lian, Xin Li, Pengfei Guo. 2016. A hotelling queue model with probabilistic service. *Operations Research Letters* **44**(5) 592–597.

Appendix. E-companion.

A. Theoretical Motivation for the Risk-Neutral Assumption

In this work we consider customers two types of customers, pessimistic ones and risk-neutral ones. In the risk-neutral case, the customer's assumption is equivalent to assuming that upon purchase of an opaque product of size i, each item in the product is allocated by the seller with probability $\frac{1}{i}$. Such an assumption is simple and natural when the seller does not announce the allocation probabilities. However when forming beliefs about the seller's allocation, it is worthwhile to consider the situation from the seller's point of view. If allocation probabilities were announced, what would be the revenue maximizing way for the seller to allocate items? When goods are horizontally differentiated is it revenue maximizing of the seller to always allocate goods uniformly at random? In this section we consider this question and show that, in general, allocating goods uniformly at random from the opaque product is not revenue-optimal. However, in the special case when valuations are supported on two points, we show that uniformly random allocations are optimal. Along the way we highlight an interesting non-linear relationship between the allocation proportions and the revenue of an opaque strategy when valuations are supported on two points (c.f. Fig. 11).

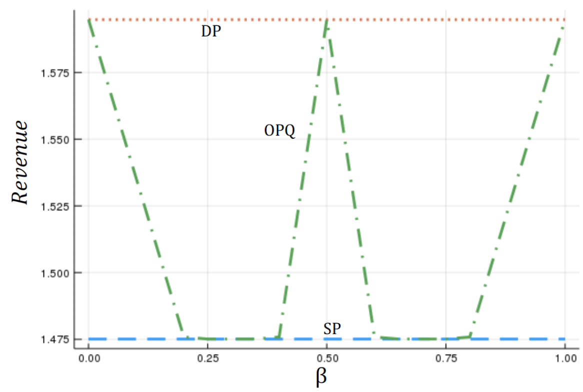
First, we show that uniformly random allocations is not always revenue-optimal. To this end, suppose $N = 2$ and consider an allocation rule where the opaque product always yields item 1 with probability $\beta > .5$ and item 2 with probability $1-\beta$. We denote the revenue of an opaque strategy which allocates in this way by $\mathcal{R}_{OPQ}(\vec{p}, \beta)$. In Fay and Xie (2008), they show that $\beta = 1/2$ is revenue optimal when the valuations follow a generalized Hotelling model, i.e., $1/2 = \arg\max_\beta \max_{\vec{p}} \mathcal{R}_{OPQ}(\vec{p}, \beta)$. Unfortunately, such a result does not hold for all exchangeable valuation distributions. Even when $N = 2$ and valuations are i.i.d., the following example implies there are cases when $\beta = 1/2$ is suboptimal.

Example 3 ($\mathcal{R}_{DP}(\vec{p}) = \mathcal{R}_{OPQ}(\vec{p}, 1)$ when $N = 2$.). Let $N = 2$ and F be an exchangeable valuation distribution. Suppose (p_1, p_2) are the prices used by an optimal DP strategy, and $p_1 \geq p_2$. When $\beta = 1$ we claim the optimal opaque selling strategy uses prices $(p^1, p^2) = (p_1, p_2)$, which implies $\mathcal{R}_{DP}(p_1, p_2) = \mathcal{R}_{OPQ}(p_1, p_2, 1)$. To see this, note that when $\beta = 1$ OPQ can never sell item 1 at the traditional price p^1 since $V_1 - p^1 < \beta V_1 + (1-\beta)V_2 - p^2 = V_1 - p^2$. Thus the opaque strategy only allocates item 1 via the opaque option, and only when $(V_1 - p^2) \geq (V_2 - p^1)^+$. Similarly, item 2 is never allocated via the opaque option, and is sold only when $V_2 - p^1 \geq (V_1 - p^2)^+$. It then follows that the opaque selling strategy with $\beta = 1$ induces the same sale probabilities as the discriminatory pricing with the same prices, and thus $\mathcal{R}_{DP}(p_1, p_2) = \mathcal{R}_{OPQ}(p_1, p_2, \beta = 1)$.

Combined with Example 1 which furnishes an example where $\mathcal{R}_{DP} > \mathcal{R}_{OPQ}$ when $\alpha = 0$, the equivalence between opaque selling to risk-neutral customers when $\beta = 1$ and discriminatory pricing implies that $\beta = 1/2$ is not revenue-optimal for all exchangeable distributions. □

One consequence of Example 3 is that $\beta = 1/2$ can only be optimal if in situations where OPQ selling is guaranteed to dominate the revenue of discriminatory pricing. As shown previously, this is not always true. However in Theorem 1(ii) we proved that when $\beta = 1/2$ and valuations are drawn from a two-point distribution, that OPQ dominates DP. In the remainder of this section we explore optimal allocation rules for opaque selling when valuations are supported on two points.

Figure 11 Revenue when $N = 2$, two point distributions, risk-neutral customers as β varies.



Note. Illustrates the relationship between β (the probability that item 1 is allocated by the opaque product) and the revenue of the optimal pricing strategies: $\mathcal{R}_{SP}, \mathcal{R}_{DP}$, and $\mathcal{R}_{OPQ}(\beta)$, as β varies between 0 and 1.

In Fig. 11, we plot the revenues of SP, DP, and OPQ when valuations are supported on two points and as β changes. Specifically, we suppose valuations are drawn i.i.d. according to the Bernoulli distribution described in Section 5.1. As one can see, the revenue is highly non-concave in β, but does peak when $\beta = 0.5$ and additionally when $\beta = 0$ or 1. This figure highlights another counter-intuitive aspect of multi-product pricing, and further motivates the need for simple assumptions about how customers behave. In Theorem 6 below we show that although there may be other optimal allocation rules, when valuations are supported on two points, $\beta = 1/2$ is always optimal.

THEOREM 6 **(When Equal Allocation is Optimal)**. *Suppose customers are risk-neutral and draw their valuations from an exchangeable distribution supported on two values. Then for any $\beta \in (.5, 1]$,*

$$\mathcal{R}_{OPQ}(\cdot, 1/2) \geq \mathcal{R}_{OPQ}(\cdot, \beta).$$

Proof. Fix $\beta > 1/2$. Following the proof of Theorem 1(ii) let the support of V be normalized to 1 and $1+\delta$, let U be the random variable counting the number of valuations that are equal to $1+\delta$, and let $q_i = \mathbb{P}(U = i)$. Since $N = 2$, there are four possible valuations for the opaque product: $\mathcal{S} = \{1, 1+(1-\beta)\delta, 1+\beta\delta, 1+\delta\}$ and further by Lemma 10, we know $p^2 \in \mathcal{S}$, and $p^1 = s_1 - (s_2 - p^2)^+$ for some $s_1, s_2 \in \mathcal{S}$. We will now enumerate the possible prices for the opaque strategy and check them individually, omitting prices such that $p^1 = p^2$ which can easily be seen to be less than or equal to $\mathcal{R}_{SP}$.

Suppose $p^2 = 1$. If the opaque strategy uses prices $(p^1, p^2) = (1+\delta, 1)$ one can easily check these price earn less than the revenue of an optimal discriminatory pricing and thus less than $\mathcal{R}_{OPQ}(\cdot, 1/2)$ by Theorem 1(ii). If $(p^1, p^2) \in \{(1+\beta\delta, 1), (1+(1-\beta)\delta, 1)\}$, then

$$\mathcal{R}_{OPQ}((1+\beta\delta, 1), \beta) = q_0 + (1+\beta\delta/2)q_1 + q_2,$$

$$\mathcal{R}_{OPQ}((1+(1-\beta)\delta, 1), \beta) = q_0 + (1+(1-\beta)\delta)q_1 + q_2,$$

which are both less than $q_0 + (1+\delta/2)q_1 + q_2 = \mathcal{R}_{OPQ}((1+\delta/2, 1), 1/2)$. Otherwise, if $p^2 > 1$ then $\mathcal{R}_{OPQ}((p_1, p_2), \beta) \leq (q_1 + q_2)p^1 = \mathcal{R}_{SP}(p^1)$. Thus, we have checked all cases and conclude $\arg\max_\beta \mathcal{R}_{OPQ}(\cdot, \beta) = 1/2$. $\square$

B. A Primer on Discriminatory Pricing

In this work, we consider pricing strategies for customers with valuations drawn from an exchangeable distribution. Due to the symmetry of exchangeable valuation distributions, it is natural to assume that the optimal pricing strategy would be to offer an identical price for each item. Surprisingly, this is not the case; discriminatory pricing where some items are priced higher than others can yield significantly more revenue. (Note any permutation of the optimal prices is also optimal.) Consider the following simple example where $N=2$, and customers draws their item valuations from an i.i.d. distributions V_1 and V_2:

$$V_1, V_2 \sim \begin{cases} 2 : \text{w.p. } 1/3 \\ 1 : \text{w.p. } 2/3 \end{cases}$$

The optimal single price strategy offers both items at price 2, and earns revenue of $\mathcal{R}_{SP} = 2 \cdot \mathbb{P}(\max\{V_1, V_2\} \geq 2) = \frac{10}{9}$. Now consider a discriminatory pricing where item 1 is sold at a price of 1 and item 2 is sold at a price of 2, then $\mathcal{R}_{DP}(1,2) = 1 \cdot (\mathbb{P}(V_1 = 1, V_2 = 1) + \mathbb{P}(V_1 = 2, V_2 = 1)) + \mathbb{P}(V_1 = 2, V_2 = 2)) + 2 \cdot \mathbb{P}(V_1 = 1, V_2 = 2) = \frac{11}{9}$ (note that the higher price item is purchased in the event of a tie w.l.o.g.). This extra $\frac{1}{9}$ is from two opposing forces at play. The low price of 1 allows the discriminatory strategy to extract revenue from customers with low valuations for both items, an additional expected revenue of $\frac{4}{9}$. The downside of the low price is that it cannibalizes sales from the high priced item when $V_1 = 2$, resulting in a loss of $\frac{1}{3}$. Overall, the upside outweighs the downside and increased revenue can be had by offering different prices for these i.i.d. valued items.

Given that the revenue from discriminatory pricing can exceed single pricing, a natural question is then how much more can revenue can discriminatory pricing earn. In i.i.d. settings, the question has been fully resolved.

LEMMA 1 **(Dutting and Klimm (2016) Theorems 3 and 4)**. *Let N be the number of items. When valuations are drawn i.i.d.,*

$$\mathcal{R}_{DP} \leq \left(2 - \frac{1}{N}\right) \mathcal{R}_{SP}.$$

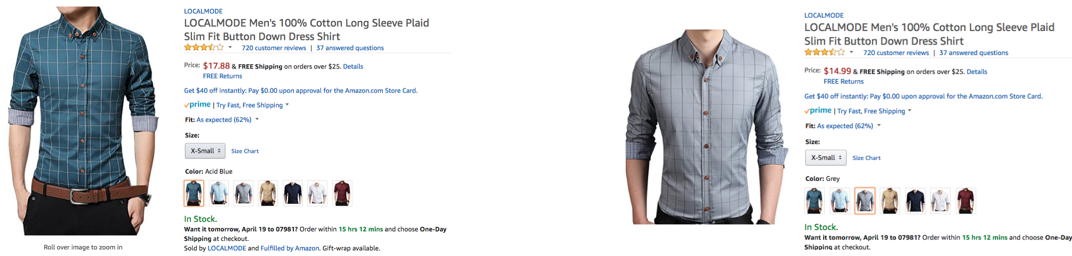
Further, this bound is tight for each N.

In addition to being theoretically interesting, discriminatory pricing strategies are common in online marketplaces, even for basic retail goods. As an example in Fig. 12, two nearly identical shirts are offered for different prices on Amazon.com.

C. An Extension of Exchangeability that Encompasses Salop's Circle Model

In this section we describe Salop's circle model, a core model for horizontally differentiated items, and discuss how Theorem 1(i) can be extended to this case. In Salop's circle model, each item is represented as a point on a circle. The N points are equidistant on the circumference of the circle, and denoted by $y_1, \ldots, y_N$. Each customer corresponds to a random point X on the circumference of the circle, distributed uniformly at random. The customer then values each item i according to $V_i = a - b\|X - y_i\|$, where the norm corresponds to the distance traveled on the circle and a, b are tunable parameters. Note that the underlying joint distribution for item valuations arising from Salop's circle is not independent, as closeness to one item on the circle necessarily implies the customer is farther from (lower valuations for) the other items.

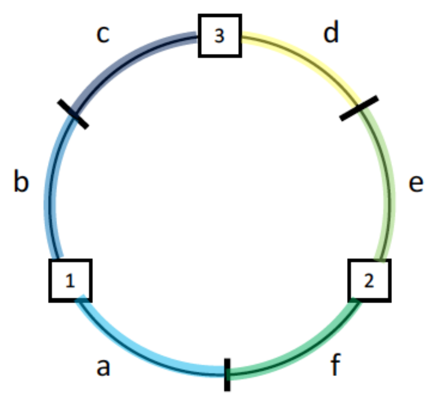
Figure 12 Discriminatory Pricing on Amazon.com.



(a) A dark blue men's shirt listed for \$17.88. (b) A light blue men's shirt listed for \$14.99.

When $N = 2$, Salop's circle reduces to the Hotelling model, which is well-known to be exchangeable. We now show that when $N = 3$, Salop's circle still gives rises to an exchangeable valuation distribution. First, observe from Fig. 13 that the six possible valuation orders of the 3 items are equally likely. For example, if a customer is in region c, they prefer item 3, then 1, then item 2. If a customer is in region d, they prefer item 3, then item 2, then item 1. All 6 orderings are possible when $N = 3$. Since X is drawn uniformly at random, then exchangeability follows immediately.

Figure 13 Salop's Circle when N=3.



Note. Salop's circle for N=3 products, divided into six equal regions.

Unfortunately this argument for exchangeability does not extend beyond $N = 3$. One easy way to see this is to note that, when $N = 4$, the customers are partitioned into 8 regions corresponding to 8 possible valuation orderings (analogous to Figure 13). However, there are 24 possible valuation orderings when $N = 4$, and thus it cannot be the case that every permutation yields identical marginal distributions. More generally, Salop's circle has $2N$ possible orderings arising from $2N$ customer regions, while exchangeability requires all $N!$ valuation orderings be equally likely.

To capture Salop's circle, we relax our definition of exchangeability from requiring every permutation to yield identical joint distributions to only a subset of permutations, S, to have identical joint distributions.

Definition 2. Let Σ be the set of all permutations on $[N]$. Let sort$(\cdot)$ be the function that sorts a vector in descending order. We call the random valuation vector $V = (V_1, \ldots, V_N)$ *S-exchangeable* if S is a non-empty subset of Σ such that for all $\sigma \in S$, $\sigma(\text{sort}(V))$ has the same joint distribution as V.

In essence, S-exchangeability limits the possible orderings of the valuations for items to the set S and enforces that inside S, each of those orderings is equally likely. Note that when $S = \Sigma$, then S-exchangeability is equivalent to our earlier notion of exchangeability. It is now easy to see that Salop's circle model is S-exchangeable, where S describes the orderings arising from the $2N$ regions. The joint distributions for each of these orderings are all the same since X is uniformly distributed on the circle.

Finally, we note that when customers are pessimistic and their valuations are S-exchangeable, opaque selling always dominates the revenue from discriminatory pricing. This implies Corollary 2, which is an extension of Theorem 1(i).

Corollary 2. *Assume $\alpha = 1$ and customers valuations are S-exchangeable, then*

$$\mathcal{R}_{OPQ} \geq \mathcal{R}_{DP}.$$

Proof. The proof is exactly the same as Theorem 1(i) where Σ, the set of all permutations, is replaced by S instead. □

D. An Enumerative Algorithm for Finding Optimal Pricings

In this section, we describe an algorithm for finding the optimal pricing in the special case when the support of the distribution is discrete. (As mentioned previously, we assume that we have approximated the original distribution by a discrete distribution.) When the number of items is small, this algorithm is relatively efficient. Specifically, we show that if N is assumed to be constant, then the optimal prices for any strategy (SP, DP, OPQ, or 1OPQ) can be found in time that is polynomial in the size of the support of the valuation distribution.

We let m denote the number of points (valuation vectors) in the support of F. Each support point j corresponds to a customer type with a valuation vector $\vec{v}_j = (v_{j,1}, v_{j,2}, \ldots, v_{j,N})$. When referring to DP, $v_{j,i}$ denotes type j's valuation for item i. When referring to OPQ, $v_{j,i}$ denotes type j's valuation for the opaque product of size i. Note that opaque valuation vectors can easily be generated given a discretized distribution by computing the opaque valuations for pessimistic and risk-neutral customers directly. The type vector can also be made to correspond to the type vector for SP by replacing each vector in the DP case with $(\max_i v_{j,i})$, and 1OPQ by replacing each vector in the OPQ case with $(v_{j,1}, v_{j,N})$. In the case of 1OPQ, this reduction ensures that regardless of how many items are being sold, since only two options are being priced the effective number of items is two.

In Theorem 7 we show that the optimal prices can be found in time $O(m^N)$ ($O(m^2)$ for 1OPQ). The idea of Algorithm 1 is to identify a set of $(m+1)^N$ candidate prices which is guaranteed to contain the optimal price. The algorithm then enumerates over the set of candidate prices and returns the price that yields the highest revenue. We summarize the result in the following theorem.

Theorem 7 **(Algorithm for Computing Optimal Prices)**. *Let F be an exchangeable distribution over m customer types. Then both the optimal opaque pricing and discriminatory pricing can be computed in $O(m^N)$ time by Algorithm 1.*

Proof. The proof will depend on the following structural lemma which asserts that the prices can be found by carefully combing through the support of the valuation distribution. Suppose $\vec{p}$ is the optimal price vector. By Lemma 10 in Appendix J, and for every $i \in [N]$, either there exists a type j such that $p_i = v_{j,i} - \max_{k>i}\{(v_{j,k} - p_k)^+\}$, or no customer type buys item i. Using this observation, we can inductively enumerate the prices starting from the lowest price and working upwards.

Algorithm 1: Enumerative Algorithm

Main `Enumerate Price Tree(`F`):`

Input: Distribution F, supported on m types $v_i \in \mathbb{R}^N$.

Initialize: $P^N = \cup_{j=1}^m v_{j,N}$

for ($i = N-1:1$) {

  for ($\tilde{p} \in P^{i+1}$) {

    $P^i = P^i \cup_{j=1}^m v_{j,i} - \max_{k>i} (v_{j,k} - \tilde{p}_k)^+$

  }

}

return $\arg\max_{\tilde{p} \in P^1} \mathcal{R}_{DP}^F(\vec{p})$

We focus on optimal pricing for DP, and the same analysis holds for OPQ. Consider the following algorithm that proceeds by guessing the prices in order from low to high. Fix some ordering on the prices $p_1 \geq p_2 \geq \ldots \geq p_N$, by exchangeability this is w.l.o.g. By Lemma 10, it must be the case that the lowest price $p_N \in \{v_{j,N}\}_{j=1}^m$ or else that item N is not purchased by any customer. If that item is supposed to not be bought, we can set the price to ∞ effectively discarding the item. Thus there are $m+1$ choices for p_N, one for each customer type and the ∞ no-purchase option. Under each of these choices, compute $\{\tilde{v}_{j,N-1}\}_{j=1}^m$ where $\tilde{v}_{j,i} = v_{j,i} - \max_{k>i} (v_{j,k} - p_k)^+$. Again, it must be the case that $p_{N-1} \in \{\tilde{v}_{j,N-1}\}_{j=1}^m$ by Lemma 10 or else it is not bought and we can set the price to ∞. Proceeding in this way we create a tree of size of depth N with $m+1$ branches, terminating in $(m+1)^N$ leaf nodes each corresponding to a potential optimal solution. For each leaf node, one can compute the revenue from the candidate price in linear time. Thus, the overall runtime is simply $O(m^N)$. $\square$

E. An Integer Programming Formulation of the Optimal Pricing Problem

In Section D we described a simple recursive algorithm for computing optimal prices. In this section we formulate the optimal pricing problem given sample valuations as a mixed integer linear program (MILP). We emphasize that the worst case run-time of this MILP is N^m which can be exponentially worse than the enumerative scheme in the common case when m, the size of the type space, is large but N, the number of items, is small. The details of this MILP are implicit in the work of Hanson and Martin (1990) who solve a broader version of this problem.

The input for the MILP is a distribution over m customer types $\{v_i\}_{i=1}^m \subset \mathbb{R}^N_{\geq 0}$. We will use q_i to be the probability of a customer of type v_i occurring. The decision variables for the MILP are:

1. $p_{i,j}$, the price customer of type i would pay for item j under the pricing.
2. $\theta_{i,j}$, boolean variables that equals 1 when a customer of type i buys item j, and 0 otherwise.
3. p_i, the price of item i.
4. $s_{i,j}$, the utility (or surplus) a customer of type i obtains from from a purchase of item j.

The optimal pricing is a solution to the following formulation.

$$\begin{aligned}
&\max \sum_i q_i \sum_j p_{i,j} && \\
&p_{i,j} \leq p_j, \quad p_{i,j} \geq p_j - \max_{i,j}\{v_{i,j}\}(1-\theta_{i,j}) && \forall (i,j) \\
&s_{i,j} = v_{i,j}\theta_{i,j} - p_{i,j} && \forall (i,j) \\
&s_i \geq v_{i,j} - p_j, \quad s_i = \sum_{j=1}^N s_{i,j} && \forall i \in [m] \\
&\sum_{i=1}^m \theta_{i,j} \leq 1 && \forall j \in [N] \\
&p_j, p_{i,j}, s_i, s_{i,j}, \geq 0, \quad \theta_{i,j} \in \{0,1\} &&
\end{aligned}$$

The objective function is the expected revenue earned by the pricing. The first set of constraints enforce that the prices charged to a customer of type i for item j are less than the price for item j and are consistent with $\theta_{i,j}$, the Boolean expressing whether or not a customer purchases. The second set of constraints enforce that the item a customer purchases maximizes their utility among the N items. The third set of constraints enforces that each customer only purchases at most one item.

F. Fitting Models

In this work we consider markets that are mixtures of pessimistic and risk-neutral customers, and use α to represent the proportion of pessimistic customers in the market. In this section we show how α can be estimated via a maximum likelihood procedure. Let F be the distribution of valuations assumed to be computed earlier and fully known to the seller, and suppose a seller experiments with opaque selling using the vector of distinct prices $\vec{p}$ such that $p^i > p^{i+1}$ for each i. Under this pricing, the seller observes sales data for m customers. Specifically, each arriving customers purchases either a traditional item (opaque product of size 1), an opaque product of size $2,\dots,N$, or no product at all. Partition the customers into sets $\{E_i\}_{i=0}^N$, where E_i is the set of customers who purchased the opaque product of size $i = 1,2,\dots,N$, and E_0 are the customers who did not purchase.

Since F and $\vec{p}$ are fixed, the expected purchase probabilities when $\alpha = 0$ and $\alpha = 1$ under $\vec{p}$ can be computed either explicitly or by simulation. We will denote these vectors of purchase probabilities as $\vec{q}^0 = (q_1^0, \dots, q_N^0)$

and $\vec{q}^1 = (q_1^1, \ldots, q_N^1)$, and more generally we will denote the vector of true purchase probabilities for a fixed α by $\vec{q}^\alpha = \alpha\vec{q}^0 + (1-\alpha)\vec{q}^1$.

Now the likelihood of observing $\{E_i\}_{i=0}^N$ under a hypothesis for the proportion of pessimistic customers $\tilde{\alpha}$ is simply $\prod_i(\tilde{\alpha}q_i^0 + (1-\tilde{\alpha})q_i^1)^{|E_i|}$ and the log-likelihood is

$$\mathcal{L}(\{E_i\}, \tilde{\alpha}) = \sum_i |E_i| \log\left(\tilde{\alpha}q_i^0 + (1-\tilde{\alpha})q_i^1\right).$$

Thus the maximum likelihood estimate for the true mixture proportion α is $\hat{\alpha} = \arg\max_{\tilde{\alpha}} \mathcal{L}(\{E_i\}, \tilde{\alpha})$ is the solution to a straightforward one dimensional concave maximization problem. To demonstrate the efficacy of our estimation procedure, in Fig. 14 we plot how well this MLE procedure fits the true α given 10, 100, and 1000 customer samples, respectively, for the three test valuation distributions described in Section 5.1. For each test distribution, a single test price vector is found via local search such that it induced non-zero purchase probabilities for each of the opaque goods. Specifically, prices are adjusted down one at a time, starting with the highest price for which the probability of purchase is zero, until all the products have nonzero probability of purchase. For the triangular distributions, the test price found is $\vec{p} = (5.6094, 4.399, 3.8357)$, for normal distributions the test price found is $\vec{p} = (5.6962, 3.732, 2.6008)$, and for Bernoulli distributions the test price found is $\vec{p} = (7, 6, 1)$. For each $\alpha \in \{.01, .02, \ldots, .99, .1\}$, we perform the estimation procedure 100 times. We will denote of the i^{th} estimation when the true value was α by $\hat{\alpha}_{i,\alpha}$. For each estimated $\hat{\alpha}_{i,\alpha}$ we plot a black dot in Fig. 14. The mean estimation for α over all 100 experiments, $\frac{1}{100}\sum_{i=1}^{100}\hat{\alpha}_{i,\alpha}$, is given by the central line along with lines representing the estimate one sample standard deviation above the mean and one standard deviation below the mean. We notice that in all cases, our estimation is centered around the true value and the mean of all estimates closely tracks the true value of α regardless of the number of samples.

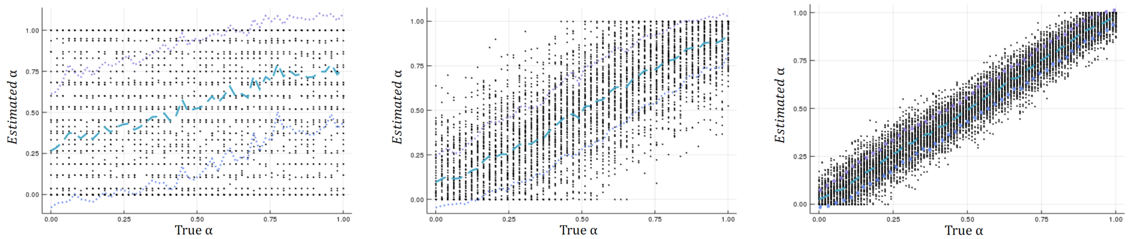
For triangular distributions using ten, one hundred, and one thousand samples, our MLE estimation achieves an average estimation error, $\frac{1}{100}\sum_{\alpha\in\{.01,.02,\ldots,.99,.1\}}\frac{1}{100}\sum_{i=1}^{100}|\alpha - \hat{\alpha}_{\alpha,i}|$, of 0.318, 0.150, and 0.051, respectively, and results in average percent revenue loss (from using $\hat{\alpha}$ rather than α) of 0.58%, 0.1394%, and 0.0123%. For normal distributions our MLE estimation achieves an average estimation error of 0.210, 0.073, and 0.024, respectively, and results in average percent revenue loss of 0.244%, 0.0435%, and 0.0052%. Finally, for Bernoulli distributions our MLE estimation achieves an average estimation error of 0.231, 0.082, and 0.027, respectively, and results in average percent revenue loss of 3.58%, 0.628%, and 0.052%. Thus we conclude that our estimation is effective, even with a modest number of samples, in approximating the true value of α, and very effective in inducing correct pricing decisions for opaque selling strategies.

G. Additional Experiments

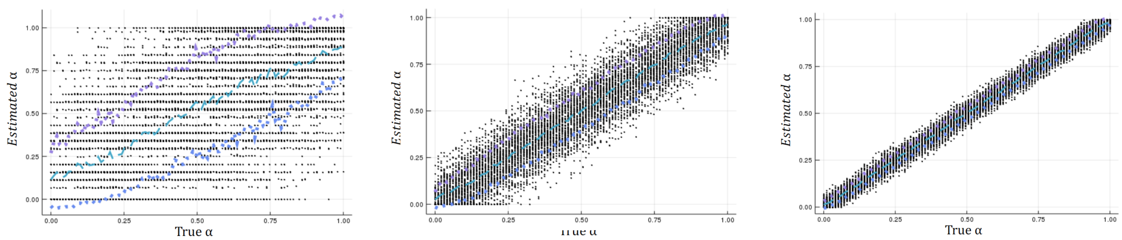
In this section we conduct two numerical experiments to further understand the performance of opaque selling, one as the variance of the valuation distribution increases, and one when valuations are not simply i.i.d. but drawn from a complicated market where some customers have strong preferences and other customers are indifferent.

One natural conjecture for explaining the performance of discriminatory pricing strategies and thus, by Theorem 1, the performance of opaque selling, is that the differentiation in prices captures variations in the

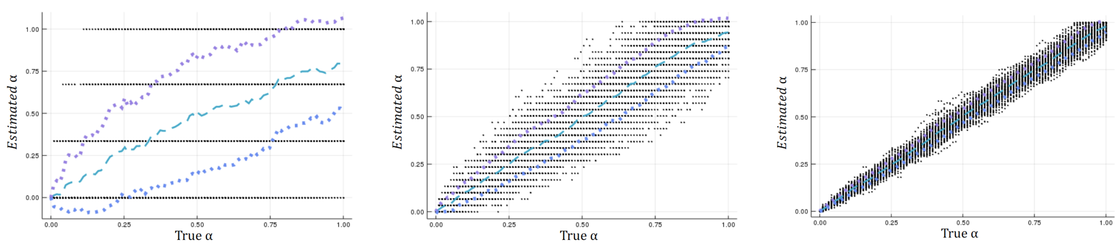
Figure 14 **Fitted α's in simulated markets when $N = 3$.**



(a) Triang. dist., 10 Samples. (b) Triang. dist., 100 Samples. (c) Triang. dist., 1000 Samples.



(d) Normal dist., 10 Samples. (e) Normal dist., 100 Samples. (f) Normal dist., 1000 Samples.

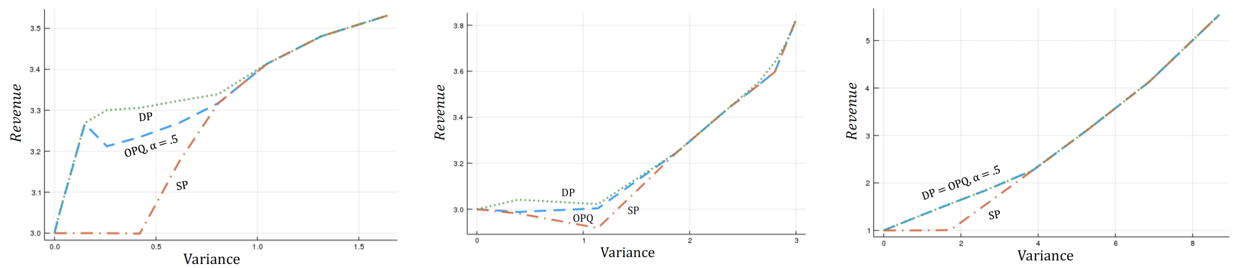


(g) Bernoulli dist., 10 Samples. (h) Bernoulli dist., 100 Samples. (i) Bernoulli dist., 1000 Samples.

Note. Plotted above are 5000 simulated estimations of α as α varies from 0 to 1 by .01, each represented by a single black dot. We repeat the experiment three times for each valuation distribution using 10, 100, and 1000 randomly generated sample purchases for each estimation, respectively. For each plot, the dashed line is the average estimate of α, and the dotted lines represent one sample standard deviation above the estimated α and below the estimated α, respectively.

customer's valuations. Thus as the variance of the underlying valuations increase so too should the benefit of price discrimination. In Fig. 15 we test this hypothesis by considering the revenue of SP, DP, and OPQ for i.i.d. valued items as the variance of the valuation distribution increases. In the OPQ strategy, we let α be .5 and note that OPQ selling revenues are higher for larger α and lower for smaller α. As expected, when the variance is close to zero SP, DP, and OPQ all perform similarly. As the variance begins to increase, both OPQ and DP begin to outperform SP by comparable amounts. However as the variance becomes quite large we observe that the gaps between the revenues of the strategies vanish, all three strategies perform nearly

Figure 15 Revenue in simulated markets when $N = 3$ as the variance increases.

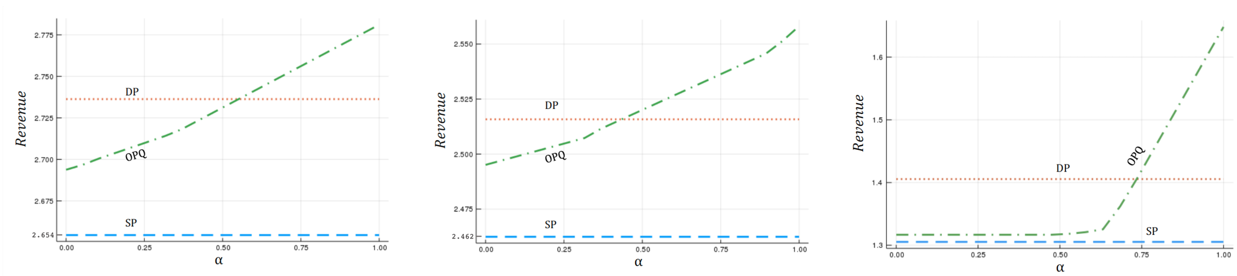


(a) Triangular distribution. (b) Truncated normal distribution. (c) Bernoulli distribution.

Note. The relationship between $\mathcal{R}_{SP}$ (dashed line), $\mathcal{R}_{DP}$ (dotted line), and $\mathcal{R}_{OPQ}$ (solid line) as the variance increases.

identically, and the optimal opaque and discriminatory prices become uniform. This experiment suggests that for highly variable distributions, it is more important to capture the highest valuation with a single optimally chosen price than it is to try and capture intra-item variation in valuations with differentiated prices due to the risk of cannibalization that can occur from offering a low price to a customer with high valuations.

Figure 16 Revenue in mixture of exchangeable markets when $N = 3$.



(a) Triangular distribution. (b) Truncated normal distribution. (c) Bernoulli distribution.

Note. Illustrates the relationship between $\mathcal{R}_{SP}$ (dashed line), $\mathcal{R}_{DP}$ (dotted line), and $\mathcal{R}_{OPQ}$ (solid line) as the proportion of pessimistic customers increases.

In our final set of numerical experiments we consider valuations that are exchangeable but not i.i.d. by incorporating two additional customer types: 1) *pointed* customers who's valuations are all zero except for a single item, and where the non-zero valued item is choose uniformly at random and is distributed according to one of the three distributions described in the beginning, and 2) *uniform* customers who's valuations are the same for all of the items, and that valuation is also distributed according to one of the three distributions described in the beginning. Specifically in Fig. 16 we consider a market where 40% of the market has valuations drawn i.i.d., 30% of the market is pointed, and the remaining 30% of the market is uniform. As in our previous experiments, both DP and OPQ greatly outperform SP and further, DP appears to outperform

OPQ when the customers are primarily risk-neutral and the relationship reverses as the market becomes more pessimistic.

Overall, our numerics support that idea that OPQ selling has comparable performance to DP in a wide variety of markets and settings, and that the parameter α can be easily and effectively estimated.

H. Missing Examples

Example 4 (OPQ Can Earn O(N) Times More Revenue than SP). We construct an exchangeable distribution F over N items such that $\mathcal{R}_{SP} \le 2$ and $\mathcal{R}_{OPQ} \ge \alpha N$, implying a gap between OPQ and SP on the order of αN. Note by Theorem 3(i) this implies $\mathcal{R}_{DP} \le 2(1+\log(N))$, and thus $\mathcal{R}_{OPQ} \ge \frac{\alpha}{2}\frac{N}{1+\log(N)}\mathcal{R}_{DP}$.

To construct $V \sim F$, we specify $N+1$ possible valuation vectors in $\mathbb{R}^N$ and assume that each permutation of the specified valuation vector is equally likely. The first vector has valuations where one item is valued at 2^N and all others are valued at 0. The second vector has valuations where two items are valued at 2^{N-1} and all other are zero and so on. All vectors have probabilities chosen so that $\mathcal{R}_{SP}(2^{N-i}) \le 2$. Formally,

$$V = \begin{cases} \text{Uniformly some permutation of } (2^N, 0, \ldots, 0) & : \text{w.p. } 2^{-N} \\ \text{Uniformly some permutation of } (2^{N-1}, 2^{N-1}, 0, ..., 0) & : \text{w.p. } 2^{-(N-1)} \\ \ldots & \\ \text{Uniformly some permutation of } (2, 2, \ldots, 2, 2) & : \text{w.p. } 2^{-1} \\ \text{Uniformly some permutation of } (1, 1 \ldots, 1, 1) & : \text{w.p. } 2^{-N} \end{cases}$$

Then $\mathcal{R}_{SP} = \max_i 2^{N-i}\mathbb{P}\left(V^{(1)} \ge 2^{N-i}\right) = \max_i 2^{N-i}\sum_{j=N-i}^{N} 2^{-j} \le \max_i 2^{N-i}2^{-N+i+1} = 2$. Now we show $\mathcal{R}_{OPQ} \ge \alpha N$. Let prices be $p^i = 2^{N-i+1}$. Now consider a pessimistic customer with $V^{(1)} = V^{(2)} = \ldots = V^{(i)} = 2^{N-i+1}$, which occurs w.p. $\alpha 2^{-(N-i+1)}$. By construction $V^{(i+1)} = 0$ and thus this customer purchases the size i opaque product at price 2^{N-i+1}. The total revenue is then

$$\mathcal{R}_{OPQ}(2^N, 2^{N-1}, \ldots, 2) \ge \alpha \sum_{i=1}^{N} 2^i 2^{-i} = \alpha N.$$

Thus OPQ earns at least αN from pessimistic customers under this pricing strategy and $\mathcal{R}_{OPQ} \ge \alpha N$. □

Example 5 (Tightness of Theorem 5 when $\alpha = 1$). We describe a distribution F such that $\mathcal{R}^F_{1OPQ}/\mathcal{R}^F_{SP} = 2$ when customers are pessimistic, demonstrating tightness of Theorem 5. Fix $z \in (0,1)$ and let $V_1, \ldots, V_N$ be i.i.d. where

$$V_i = \begin{cases} 1 & : \text{w.p. } z \\ 1-(1-z)^N & : \text{w.p. } 1-z \end{cases}$$

Then,

$$\mathcal{R}_{SP} = \mathcal{R}_{SP}(1) = \mathcal{R}_{SP}(1-(1-z)^N) = 1-(1-z)^N.$$

Similarly we compute the revenue of a 1OPQ strategy with prices $(1, 1-(1-z)^N)$. In this strategy, the opaque product is only purchased if the customer has a high valuation for all items or a low valuation for all items. Thus,

$$\mathcal{R}_{1OPQ}(1, 1-(1-z)^N) = 1(1-z^N-(1-z)^N) + (1-(1-z)^N)(z^N+(1-z)^N).$$

Then,

$$\begin{aligned} \frac{\mathcal{R}_{1OPQ}}{\mathcal{R}_{SP}} \ge \frac{\mathcal{R}_{1OPQ}(1, 1-(1-z)^N)}{\mathcal{R}_{SP}} &= \frac{(1-z^N-(1-z)^N) + (1-(1-z)^N)(z^N+(1-z)^N)}{1-(1-z)^N} \\ &= 1 + z^N + (1-z)^N - \frac{z^N}{1-(1-z)^N}. \end{aligned} \tag{2}$$

Then Eq. (2) can be arbitrarily close to 2 as z goes to zero. Note this example holds for any N. □

Example 6 (Tightness of Theorem 5 when $\alpha=0$). We describe a distribution F such that $\mathcal{R}^F_{1OPQ}/\mathcal{R}^F_{SP}=\frac{3}{2}$ when customers are risk-neutral, demonstrating tightness of Theorem 5 when $N=2$. Let valuations for the two items be drawn i.i.d. according to the CDF

$$\mathbb{P}(V_i \le x)=\sqrt{1-\frac{1}{x}}, \quad x\in[1,\infty).$$

Note that for any $x\in[1,\infty)$, we have that $\mathbb{P}(\max\{V_1,V_2\}\le x)=1-\frac{1}{x}$ and $\mathbb{P}(\max\{V_1,V_2\}\ge x)=\frac{1}{x}$. Thus $R_{SP}=1$ since $p\mathbb{P}(\max\{V_1,V_2\}\ge p)=1\,\forall p\in[1,\infty)$.

Now consider the 1OPQ strategy $(p,1)$ for $p>1$. To compute $\mathcal{R}_{1OPQ}(p,1)$, let $V^{(1)}=\max_i\{V_i\}$, $V^{(2)}=\min_i\{V_i\}$, $u=V^{(1)}-p$, and $u^2=\frac{V^{(1)}+V^{(2)}}{2}-1$. Note that u and u^2 are the utilities of buying the best item or the opaque product, respectively. By conditioning on the event $V^{(1)}\ge p$, we show that

$$\begin{aligned}
\frac{\mathcal{R}_{1OPQ}}{\mathcal{R}_{SP}} &\ge \mathcal{R}_{1OPQ}(p,1)\\
&=\mathbb{P}(V^{(1)}\ge p)\left(p\mathbb{P}(u\ge(u^2)^+|V^{(1)}\ge p)+\mathbb{P}(u^2>(u)^+|V^{(1)}\ge p)\right)\\
&+\mathbb{P}(V^{(1)}<p)\left(p\mathbb{P}(u\ge(u^2)^+|V^{(1)}<p)+\mathbb{P}(u^2>(u)^+|V^{(1)}<p)\right)\\
&=\mathbb{P}(V^{(1)}\ge p)(p\mathbb{P}(V^{(1)}-V^{(2)}\ge 2p-2|V^{(1)}\ge p)+\mathbb{P}(V^{(1)}-V^{(2)}<2p-2|V^{(1)}\ge p))\\
&+\mathbb{P}(V^{(1)}<p)(p\mathbb{P}(V^{(1)}-V^{(2)}\ge 2p-2|V^{(1)}<p)+\mathbb{P}(V^{(1)}-V^{(2)}<2p-2|V^{(1)}<p))\\
&=\mathbb{P}(V^{(1)}\ge p)(p\mathbb{P}(V^{(1)}-V^{(2)}\ge 2p-2|V^{(1)}\ge p)+\mathbb{P}(V^{(1)}-V^{(2)}<2p-2|V^{(1)}\ge p))\\
&+\mathbb{P}(V^{(1)}<p)\\
&=\frac{1}{p}(p\mathbb{P}(V^{(1)}-V^{(2)}\ge 2p-2|V^{(1)}\ge p)+\mathbb{P}(V^{(1)}-V^{(2)}\le 2p-2|V^{(1)}\ge p))+(1-\frac{1}{p})\\
&\ge \mathbb{P}(V^{(1)}\ge 2p-2|V^{(1)}\ge p)+\frac{1}{p}\mathbb{P}(V^{(1)}-V^{(2)}\le 2p-2|V^{(1)}\ge p)+1-\frac{1}{p}\\
&=\frac{\frac{1}{2p-2}}{\frac{1}{p}}+\frac{1}{p}\mathbb{P}(V^{(1)}-V^{(2)}\le 2p-2|V^{(1)}\ge p)+1-\frac{1}{p} \qquad (3)
\end{aligned}$$

The first inequality follows since $\mathcal{R}_{SP}=1$ and $(p,1)$ is feasible for 1OPQ. The first equality follows from the definition of $\mathcal{R}_{1OPQ}(p,1)$ conditioning on $V^{(1)}\ge p$. The second equality follows from the definitions of u and u^2. The third equality follows from the fact that $V^{(2)}\ge 1$ combined with the case where $V^{(1)}<p$. The fourth equality follows from the distribution F. The second inequality follows since $V^{(2)}\ge 0$. The last equality follows from Bayes rule. As p goes to ∞, the expression in Eq. (3) goes to $\frac{3}{2}$, matching the upper bound in Theorem 5. $\square$

I. Missing Proofs

I.1. Proof of Corollary 1.

Case (i): From the proof of Theorem 1(i), the revenue from OPQ dominates the revenue under SP conditional on every ordering of the valuations. When $V_N=V^{(1)}$, only the highest-valued item can be purchased under DP. Conditional on this event, the revenue from DP is clearly at most $\mathcal{R}_{SP}$. Thus,

$$\mathcal{R}_{DP}\le\frac{N-1}{N}\mathcal{R}_{OPQ}+\frac{1}{N}\mathcal{R}_{SP}=\frac{N-1}{N}\mathcal{R}_{OPQ}+\frac{1}{N+N\gamma}\mathcal{R}_{DP}.$$

Rearranging the inequality gives the result.

Case (ii): Recall from the proof Theorem 1(ii) that w.l.o.g. we may assume F is supported on $\{1, 1+\delta\}$. Suppose the OPQ prices are $p^i = 1+\delta$ for $i < N$ and $p^N = 1$. Recall that U is a r.v. denoting the number of items for which the customer has a valuation of $1+\delta$. A pessimistic customers buys an item at a price of $1+\delta$ if $1 \le U \le N-1$ and the opaque product of size N at price 1 otherwise. Letting $u_i = \mathbb{P}(U = i)$, then

$$\mathcal{R}_{OPQ} \ge 1 + \delta(1 - u_0 - u_N). \tag{4}$$

Similarly from Eq. (7) in Appendix I.6 we have that

$$\mathcal{R}_{DP} = 1 + \delta \sum_{qi=1}^{N-1} \frac{N-i}{N} u_i \le 1 + \delta \frac{N-1}{N}(1 - u_0 - u_N).$$

Now since $\mathcal{R}_{DP} = (1+\gamma)\mathcal{R}_{SP}$ and $\mathcal{R}_{SP} \ge \mathcal{R}_{SP}(1) = 1$, then

$$\delta \frac{N-1}{N}(1 - u_0 - u_N) \ge \gamma \mathcal{R}_{SP}.$$

Combining these inequalities with Eq. (4) we obtain

$$\begin{aligned}
\mathcal{R}_{OPQ} &\ge 1 + \delta(1 - u_0 - u_N) \\
&\ge \mathcal{R}_{DP} + \frac{1}{N}\delta(1 - u_0 - u_N) \\
&\ge \mathcal{R}_{DP} + \frac{\gamma}{N-1}\mathcal{R}_{SP} = \left(1 + \frac{\gamma}{(N-1)(1+\gamma)}\right)\mathcal{R}_{DP}
\end{aligned}$$

When customers are risk-neutral, they always purchase the opaque product of size N garnering revenue 1. Thus $\mathcal{R}_{OPQ} \ge \alpha \left(1 + \frac{\gamma}{(N-1)(1+\gamma)}\right)\mathcal{R}_{DP} + 1 - \alpha$. Thus when $\alpha \ge 1 - \frac{\gamma}{\gamma N + N - 1}$ we obtain $\mathcal{R}_{OPQ} > \mathcal{R}_{DP}$. □

I.2. Proof of Theorem 2

We divide the proof of Theorem 2 into two lemmas. Lemma 2 states that when customers are purely pessimistic $\mathcal{R}_{OPQ} \le 3\mathcal{R}_{SP}$. Lemma 4 states that when customers are purely risk-neutral $\mathcal{R}_{OPQ} \le \left(4 - \frac{2}{N}\right)\mathcal{R}_{SP}$. To obtain Theorem 2, we relax OPQ to observe X_α and price pessimistic and risk-neutral customers separately. Using Lemmas 2 and 4, we get that $\mathcal{R}_{OPQ} \le \left(\alpha \cdot 3 + (1-\alpha)\left(4 - \frac{2}{N}\right)\right)\mathcal{R}_{SP}$ which proves the desired result.

Lemma 2. *Assume all customers are pessimistic. Then when item valuations are i.i.d.,*

$$\mathcal{R}_{OPQ} \le 3\mathcal{R}_{SP}.$$

The proof of Lemma 2, found in Appendix I.3, relies on connecting the revenue generated by OPQ to a Myerson auction, and makes use of the following lemma.

Lemma 3 **(Chawla et al. (2010) Theorem 8)**. *Let $\mathcal{R}^{\mathcal{M}}$ be the expected revenue from the Myerson auction for one item, run on N bidders with i.i.d. valuations. Then*

$$\mathcal{R}^{\mathcal{M}} \le 2\mathcal{R}_{SP}.$$

We now consider the case of risk-neutral customers in Lemma 4.

Lemma 4. *Assume all customers are risk-neutral. When item valuations are i.i.d., then*

$$\mathcal{R}_{OPQ} \le \left(4 - \frac{2}{N}\right)\mathcal{R}_{SP}.$$

The proof of Lemma 2 bounds the revenue from exponentially many opaque products by the highest valuation any opaque product could receive from a pessimistic customer. We noted that the highest valuation for an opaque product is bounded by the expected value of a second price auction, which allowed us to apply Lemma 3. Such an argument fails for risk-neutral customers since valuations for opaque products can be higher than $V^{(2)}$, the second order statistic of V. To circumvent this difficulty, we recast opaque selling with risk-neutral customers in the language of lotteries.

DEFINITION 3. A *lottery* over N items denoted by $(p, \vec{q})$ consists of a price p and probabilities q_i for receiving each item i, s.t. $\sum_{i=1}^{N} q_i \leq 1$.

A customer with valuation vector $\vec{v}$ values a lottery $(p, \vec{q})$ as $\sum_{i=1}^{n} v_i q_i - p$. Note that selling lotteries can simulate deterministic item pricing by defining N lotteries where lottery $l_i = (p, e_i)$, where e_i is the i^{th} unit vector. An opaque product over a set S can be cast as a lottery with price $p^{|S|}$ and allocation probabilities $q_i = \frac{1}{|S|}$ for each $i \in S$ and $q_i = 0$ for each $i \notin S$. We call a collection of offered lotteries a *lottery pricing*, denoted by $\mathcal{L}$. Using this framework, we can prove that OPQ can obtain at most $4 - \frac{2}{N}$ times more revenue than SP. The proof can be found in Appendix I.4, draws on lottery pricing results of Chawla et al. (2015), who proved an upper bound of 4 in their setting.

I.3. Proof of Lemma 2

Proof. Let $(p, \vec{p})$ denote the prices of an optimal OPQ strategy under F, where p is the price of items and $\vec{p}$ are the prices of the opaque products. The proof follows by separately bounding revenue from items priced at p and the the revenue from opaque products. Let $V^{(k)}$ be the k^{th} order statistic (counting so that $V^{(1)} = \max_i\{V_i\}$), and note that the highest valuation a customer has for opaque products of size k is just $V^{(k)}$. Then,

$$\begin{aligned}
\mathcal{R}_{OPQ} &= p\mathbb{P}(V^{(1)} - p \geq \max_{k=2,\ldots,N}\{V^{(k)} - p^k, 0\}) + \sum_{k=2}^{N} p^k \mathbb{P}(\text{buys opaque product of size } k) \\
&\leq p\mathbb{P}(V^{(1)} - p \geq 0) + \sum_{k=2}^{N} p^k \mathbb{P}(\text{buys opaque product of size } k) \\
&\leq p\mathbb{P}(V^{(1)} - p \geq 0) + E[V^{(2)}] \\
&\leq \mathcal{R}_{SP} + E[V^{(2)}] \\
&\leq \mathcal{R}_{SP} + \mathcal{R}^{\mathcal{M}} \\
&\leq 3\mathcal{R}_{SP}.
\end{aligned}$$

The equality follows from the definitions of $\mathcal{R}_{OPQ}, p$, and $\vec{p}$. The first inequality follows from non-negativity of $\max_{k=2,\ldots,N}\{V^{(k)} - p^k, 0\}$. The second inequality follows from realizing that the highest valued opaque product is valued at $V^{(2)}$, and thus customers pay at most $V^{(2)}$ when buying an opaque product. The third inequality follows from the optimality of $\mathcal{R}_{SP}$. The fourth inequality follows from the fact that $E[V^{(2)}]$ is the revenue of a second price auction, which is at most the revenue of the optimal (Myerson) auction. The final inequality follows from Lemma 3. $\square$

I.4. Proof of Lemma 4

The proof is based on following lemma which makes a fundamental connection between lottery pricings and the Myerson auctions.

LEMMA 5 **(Lemmas 3 and 4 in Chawla et al. (2015))**. *Consider a customer with a valuation draw $\vec{v}$ and let $i^* = argmax_i v_i$. Let $\mathcal{L}$ be a lottery pricing such that the custo mer buys lottery $(p, q_1, \ldots, q_N) = l \in \mathcal{L}$. Let $\mathcal{M}$ be the Myerson auction for one item, run on N bidders with valuations drawn from F. Then*

$$\mathcal{R}_{\mathcal{L}}(\vec{v}) \leq \mathcal{R}^{\mathcal{M}}(\vec{v}) + \sum_{i \neq i^*} q_i v_i.$$

where $\mathcal{R}_{\mathcal{L}}(\vec{v})$ and $\mathcal{R}^{\mathcal{M}}(\vec{v})$ denote the revenue earned by the lottery pricing and Myerson auction when the valuation draw is $\vec{v}$.

Proof. Let $(p, \vec{p})$ denote the prices of an optimal OPQ strategy under F, where p is the price of items and $\vec{p}$ are the prices of the opaque products. Note that every opaque product S can be written as a lottery with the same price and a uniform allocation probability over S. Furthermore, we can describe the items as N individual lotteries, each priced at p with a deterministic allocation. Thus for risk-neutral customers, our opaque selling strategy can be recast as a lottery pricing which we call $\mathcal{L}_{\mathcal{OPQ}}$, i.e.,

$$\mathcal{R}_{OPQ} = \mathcal{R}_{\mathcal{L}_{\mathcal{OPQ}}}.$$

From Lemma 5, we have that

$$\mathcal{R}_{\mathcal{L}_{\mathcal{OPQ}}}(\vec{v}) \leq \mathcal{R}^{\mathcal{M}}(\vec{v}) + \sum_{i \neq i^*} q_i v_i. \tag{5}$$

We note that if a customer with valuation $\vec{v}$ and $i^* = \text{argmax}_i v_i$ purchases an item, then $\sum_{i \neq i^*} q_i v_i = 0$ since $q_{i^*} = 1$. Otherwise if an opaque product is purchased, $\sum_{i \neq i^*} q_i \leq \frac{N-1}{N}$ and $v_i \leq v^{(2)}$ for $i \neq i^*$. Combining these facts with Eq. (5) yields

$$\mathcal{R}_{\mathcal{L}_{\mathcal{OPQ}}}(\vec{v}) \leq \mathcal{R}^{\mathcal{M}}(\vec{v}) + \frac{N-1}{N} v^{(2)}. \tag{6}$$

Thus

$$\begin{aligned}
\mathcal{R}_{OPQ} &\leq \mathcal{R}^{\mathcal{M}} + \frac{N-1}{N} E[V^{(2)}] \\
&\leq \mathcal{R}^{\mathcal{M}} + \frac{N-1}{N} \mathcal{R}^{\mathcal{M}} \\
&\leq (4 - \frac{2}{N}) \mathcal{R}_{SP}
\end{aligned}$$

The first inequality follows from taking the expectation of Eq. (6) over $\vec{v}$. The second inequality follows from the fact that $E[V^{(2)}]$ is the revenue of a second price auction, which is dominated by the Myerson auction. The third inequality follows from Lemma 3. $\square$

I.5. Proof of Theorem 3

Proof. First we prove part (i). Fix an exchangeable distribution F over N items. For ease of exposition we assume F is continuous and ignore ties, although the same argument follows when F is not continuous and one carefully considers the tie-breaking procedure. Let $p_1 \geq p_2 \geq \ldots \geq p_N$ be the optimal discriminatory pricing and let $q_1, q_2, \ldots, q_N$ be the probability item $1, 2, \ldots, N$ is sold under this pricing. Define $Q_i = \{v \in \text{supp}(F) \,|\, v_i - p_i \geq (v_j - p_j)^+ \, \forall j\}$. Note that $q_i = \mathbb{P}(Q_i)$. Let $\sigma_{i,j} : \mathbb{R}^N \to \mathbb{R}^N$ be the map that interchanges v_i and v_j in a vector v.

We first observe that $\mathbb{P}(\sigma_{i,j}(Q_i)) = q_i$ by exchangeability. Further, for all $i < j$ and any $v \in Q_i$, notice that $\sigma_{i,j}(v) \notin Q_i$ since $p_i \geq p_j$. With that established note,

$$\mathbb{P}(V^{(1)} \geq p_i) \geq \mathbb{P}(\cup_{j \geq i} \sigma_{i,j}(Q_i)) = \cup_{j \geq i} \mathbb{P}(\sigma_{i,j}(Q_i)) = (N - i + 1) q_i.$$

In words, the above equation says the probability the highest valuation is greater than p_i is lower bounded by the probability of selling item i under a discriminatory pricing, union with disjoint permutations of the event. Using this observation we can bound the gap between $\mathcal{R}_{SP}$ and $\mathcal{R}_{DP}$ as

$$\frac{\mathcal{R}_{DP}}{\mathcal{R}_{SP}} \leq \frac{\sum_{i=1}^N q_i p_i}{\max_i \mathbb{P}(V^{(1)} \geq p_i) p_i} \leq \frac{\sum_{i=1}^N q_i p_i}{\max_i (N - i + 1) q_i p_i} \leq H_N.$$

The final inequality follows from Lemma 6 in Appendix J with $C_1 = 0, C_2 = 1$, $K = 1$, and H_N is the N^{th} harmonic number. Recalling the fact that $H_N \leq 1 + \log(N)$ yields the result.

Part (ii) follows from the observation that one of the N prices in an optimal opaque selling strategy garners the most revenue. Call that price p_i and let q_i be the probability an opaque product of size i is sold. Then we must have that $q_i \leq \mathbb{P}(V^{(1)} \geq p_i)$ and thus $p_i q_i \leq \mathcal{R}_{SP}(p_i) \leq \mathcal{R}_{SP}$. Therefore, $\mathcal{R}_{OPQ} \leq N p_i q_i \leq N \mathcal{R}_{SP}$.

Part (iii) follows from combining Example 4 with part (i). ☐

I.6. Proof of Theorem 4

Proof. Let F be an exchangeable distribution where item valuations can take only two points $\{a, b\}$ where $a < b$, and suppose the market is α-mixed. Recall for distributions supported on two points the optimal discriminatory pricing uses prices $\vec{p} = (a, a, \ldots, a)$, $(b, b, \ldots, b)$ or a mixed pricing where exactly one price (since F is exchangeable it doesn't matter which price) is low $(a, b, b, \ldots, b)$. If either $(a, a, \ldots, a)$ or $(b, b, \ldots, b)$ is the the optimal discriminatory pricing given F, then $\mathcal{R}_{SP} = \mathcal{R}_{DP}$ and the claim follows automatically. Suppose $\mathcal{R}_{DP} > \mathcal{R}_{SP}$, then the optimal pricing is the mixed strategy and, under a mixed pricing, a discriminatory selling strategy always sells the item. Further we restrict ourselves to 1OPQ strategies that always sell the item, thus we may normalize the support of F to $\{1, 1 + \delta\}$ without changing the ratio $\frac{\mathcal{R}_{DP}}{\mathcal{R}_{1OPQ}}$.

Define U to be the random variable supported on $\{1, \ldots, N\}$ such that $\mathbb{P}(U = i) = \mathbb{P}\left(V^{(i)} = 1 + \delta, V^{(i+1)} = 1\right)$ where $V^{(i)}$ is the i^{th} highest order statistic of F. In words, U is the random variable for how many of the N valuations are equal to $1 + \delta$. Recall since F is exchangeable, if $U = i$ then all $\binom{N}{i}$ arrangements of valuations over the N items are equally likely (thus knowing the distribution of U is equivalent to knowing F in a two point setting). Let $u_i := \mathbb{P}(U = i)$. Conditioning on U we can compute $\mathcal{R}_{DP}$ as

$$\mathcal{R}_{DP} = E[\mathcal{R}_{DP}(1, 1 + \delta, \ldots, 1 + \delta) \,|U = i] = u_0 + \sum_{i=1}^N u_i \left(\frac{i}{N} + \frac{N - i}{N} (1 + \delta) \right) = 1 + \delta \sum_{i=1}^N \frac{N - i}{N} u_i \tag{7}$$

where the second equality follows from Eq. (1).

For each $i = 1, \ldots, N$, we shall lower bound the 1OPQ pricing strategies $(p, p_N) = (1 + \frac{N-i}{N}\delta, 1)$. To analyze the revenue from such a pricing, first note a pessimistic customers always buy an item at price $1 + \frac{N-i}{N}\delta$ for any 1OPQ pricing $(1 + \frac{N-i}{N}\delta, 1)$ as long as $U \neq 0, N$, and the opaque product otherwise. Thus a pessimistic customer has expected revenue $1 + (1 - u_0 - u_N)\frac{N-i}{N}\delta$.

For risk-neutral customers and 1OPQ pricing $(1 + \frac{N-i}{N}\delta, 1)$, customers will purchase an item at price $1 + \frac{N-i}{N}\delta$ if $0 < U \leq i, U \neq N$, and the opaque product otherwise. Thus a risk-neutral customer has expected revenue $1 + \frac{N-i}{N}\delta \sum_{j=1}^{\min(N-1,i)} u_j$. Putting them together we have

$$\begin{aligned}
\mathcal{R}_{1OPQ} &\geq 1 + \max_{i\in[N]} \alpha(1 - u_0 - u_N)\frac{N-i}{N}\delta + (1-\alpha)\frac{N-i}{N}\delta \sum_{j=1}^{\min(N-1,i)} u_j \\
&= 1 + \max_{i\in[N]} \frac{N-i}{N}\delta\,(1 - u_0 - u_N)\left(\alpha + \frac{1-\alpha}{1 - u_0 - u_N}\sum_{j=1}^{\min(N-1,i)} u_j\right). \\
&\geq 1 + \max_{i\in[N]} \frac{N-i}{N}\delta \sum_{j=1}^{\min(N-1,i)} u_j \\
&= 1 + \max_{i\in[N]} \frac{N-i}{N}\delta \sum_{j=1}^{i} u_j
\end{aligned} \tag{8}$$

where the second inequality follows from noting Eq. (8) is an increasing function of α and then plugging in $\alpha = 0$. The second equality follows since $i = N$ is not the maximizer. Let $u_i' = \frac{u_i}{1-u_0}$ so that $\sum_{i=1}^{N} u_i' = 1$. Then

$$\begin{aligned}
\frac{\mathcal{R}_{DP}}{\mathcal{R}_{1OPQ}} \leq \frac{1 + \delta\sum_{i=1}^{N}\frac{N-i}{N}u_i}{1 + \delta\max_{i\in[N]}\frac{N-i}{N}\sum_{j=1}^{i}u_j} &= \frac{1 + \delta(1-u_0)\sum_{i=1}^{N}\frac{N-i}{N}u_i'}{1 + \delta(1-u_0)\max_{i\in[N]}\frac{N-i}{N}\sum_{j=1}^{i}u_j'} \\
&\leq \frac{1 + \frac{(1-u_0)\delta}{N}H_{N-1}}{1 + \frac{(1-u_0)\delta}{N}}.
\end{aligned}$$

The first inequality follows from Eq. (7) and Eq. (8). The second inequality follows from applying Lemma 7 (under the appropriate change of variables i.e. relabeling $x_i \to x_{N-i}$) where H_{N-1} is the $(N-1)^{th}$ harmonic number. Lastly note that

$$\frac{\mathcal{R}_{DP}}{\mathcal{R}_{1OPQ}} \leq \frac{\mathcal{R}_{DP}}{\mathcal{R}_{SP}(1+\delta)} \leq \frac{1 + \delta(1-u_0)}{(1-u_0)(1+\delta)} \leq \frac{1 + \delta(1-u_0)}{(1-u_0)\delta}.$$

The first inequality follows from observing that $\mathcal{R}_{1OPQ}$ earns as much as a SP strategy with price $1+\delta$, which has expected revenue $(1-u_0)(1+\delta)$. The second inequality follows from a simple upper bound on Eq. (7). Define $C = (1-u_0)\delta$ and putting it all together we have,

$$\frac{\mathcal{R}_{DP}}{\mathcal{R}_{1OPQ}} \leq \max_{N\in\mathbb{N}, C>0} \min\left(\frac{1+C}{C}, \frac{1 + C\frac{1+H_{N-1}}{N}}{1 + \frac{C}{N}}\right) \tag{9}$$

which can be checked to be maximized when $N = 7, C = \frac{2}{29}\left(5 + 4\sqrt{65}\right)$ yielding a ratio $\geq .719$.

I.7. Proof of Theorem 5.

Proof. We divide the proof of Theorem 2 into two parts. First we show that when customers are purely pessimistic, $\mathcal{R}_{1OPQ} \leq 2\mathcal{R}_{SP}$. Second we show that when customers are purely risk-neutral $\mathcal{R}_{1OPQ} \leq$

$\left(2-\frac{1}{N}\right)\mathcal{R}_{SP}$. To obtain the result, we relax 1OPQ to observe X_α and price pessimistic and risk-neutral customers separately. Using the previously mentioned results, we get that $\mathcal{R}_{1OPQ} \le \left(\alpha \cdot 2 + (1-\alpha)\left(2-\frac{1}{N}\right)\right)\mathcal{R}_{SP}$ which is equivalent to the desired result.

First we prove that $\mathcal{R}_{1OPQ} \le 2\mathcal{R}_{SP}$ when customers are all pessimistic ($\alpha = 1$). Fix a distribution F and let (p, p^N) be an optimal solution corresponding to $\mathcal{R}^F_{1OPQ}$. Then

$$\begin{aligned}
\mathcal{R}_{1OPQ} &= \mathcal{R}_{1OPQ}(p,p^N) \\
&= p\mathbb{P}(\max_i\{V_i - p\} \ge \min_i\{V_i - p^N, 0\}) + p^N\mathbb{P}(V^N - p^N > \max_i\{V_i - p\} \cap V^N - p^N \ge 0) \\
&\le p\mathbb{P}(\max_i\{V_i - p\} \ge 0) + p^N\mathbb{P}(\max_i\{V_i - p^N\} \ge 0) \\
&= \mathcal{R}_{SP}(p) + \mathcal{R}_{SP}(p^N) \\
&\le 2\mathcal{R}_{SP}.
\end{aligned}$$

The second equation follows from the definition of $\mathcal{R}_{1OPQ}(p,p^N)$ and breaks ties by choosing to buy an item versus an opaque product. The first inequality follows from increasing the size of the event being measured. The second inequality follows from the fact that p and p^N are feasible solutions to SP. For tightness, see Example 5.

Now we focus on the case when customers are all risk-neutral ($\alpha = 0$) and show $\mathcal{R}_{1OPQ} \le (2-\frac{1}{N})\mathcal{R}_{SP}$. Fix a distribution F and let (p, p^N) be an optimal solution corresponding to $\mathcal{R}^F_{1OPQ}$. Our proof breaks into two cases depending on the relative gap between p and p^N, corresponding to $p^N \ge \frac{1}{N}p$ (Case 1) and $p^N < \frac{1}{N}p$ (Case 2). Fig. 17 illustrates the geometric difference in the two cases for $N = 2$ items.

Case 1: Recall in this case, $p^N \ge \frac{1}{N}p$. We first define q_A, q_B, q_C, q_D to be the probabilities corresponding to the following disjoint events under F, namely

$$\begin{aligned}
q_A &= \mathbb{P}(\max\{V_i\} - p \ge \frac{\sum V_i}{N} - p^N, \max\{V_i\} \ge p) \\
q_B &= \mathbb{P}(\max\{V_i\} - p < \frac{\sum V_i}{N} - p^N, \max\{V_i\} \ge p) \\
q_C &= \mathbb{P}(\max\{V_i\} < p, \frac{\sum V_i}{N} \ge p^N) \\
q_D &= \mathbb{P}(\max\{V_i\} < p, \frac{\sum_{i=1}^N V_i}{N} < p^N).
\end{aligned}$$

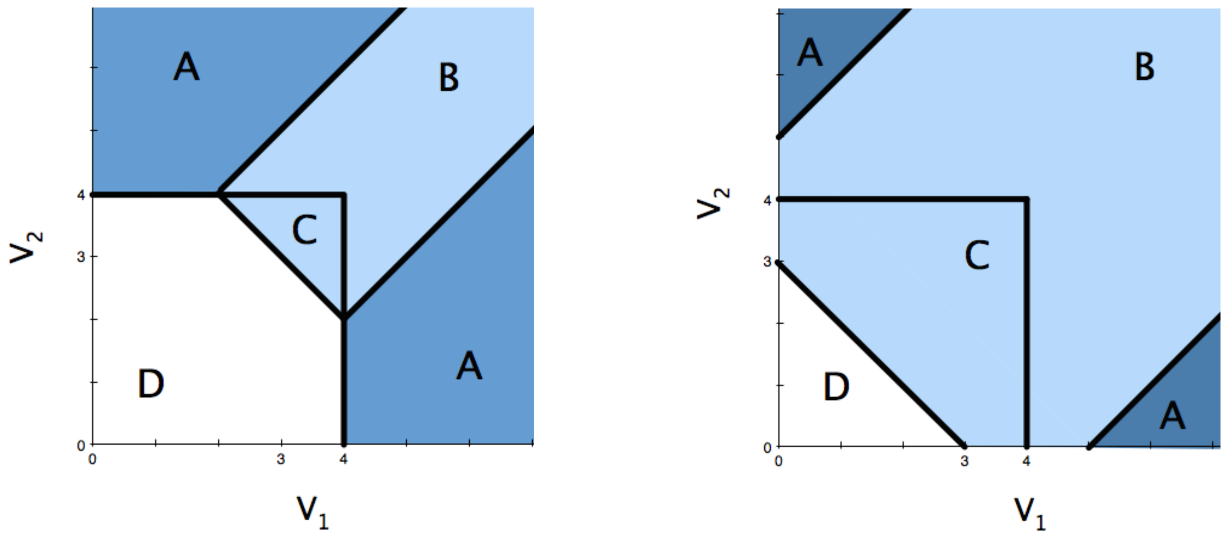
Note that $q_A + q_B + q_C + q_D = 1$. Using these probabilities, we have that $\mathcal{R}^F_{1OPQ}(p,p^N) = pq_A + p^N(q_B + q_C)$. Further, we can express the revenues from the single pricing approximately as $\mathcal{R}^F_{SP}(p) = p(q_A + q_B)$ and $\mathcal{R}^F_{SP}(p^N) = p^N\mathbb{P}(\max\{V_i\} \ge p^N) \ge p^N(q_A + q_B + q_C)$. Thus we have

$$\frac{\mathcal{R}^F_{1OPQ}}{\mathcal{R}^F_{SP}} \le \frac{\mathcal{R}^F_{1OPQ}(p,p^N)}{\max\{\mathcal{R}^F_{SP}(p), \mathcal{R}^F_{SP}(p^N)\}} \tag{10}$$

$$\le \frac{pq_A + p^N(q_B + q_C)}{\max\{p(q_A + q_B), p^N(q_A + q_B + q_C)\}} \tag{11}$$

$$\le \max_{\substack{a,b,c,d \ge 0 \\ a+b+c+d=1 \\ x \ge y \ge \frac{x}{N} \ge 0}} \frac{xa + y(b+c)}{\max\{x(a+b), y(a+b+c)\}}. \tag{12}$$

Figure 17 Visualization of the two cases for the proof of Theorem 5



(a) The case where $p^2 \geq \frac{1}{2}p$. (b) The case where $p^2 < \frac{1}{2}p$.

Note. These figures demonstrate the partition of the valuation space for a risk-neutral customer when $N = 2$. In both figures, $p = 4$. The left figure corresponds to a small discount for the opaque product and the right figure corresponds to a large discount for the opaque product. The four letters denote different buying behaviors of the customer under a (p, p^N) 1OPQ strategy and a SP strategy with price p.

The first inequality follows from the fact that p and p^N are feasible for SP. The second inequality follows from the previous discussion. The third inequality follows from the fact that $(p, p^N, q_A, q_B, q_C, q_D)$ is a feasible solution to the optimization problem in Eq. (12), which we denote by OPT. Lemma 8, proved separately, shows that $OPT \leq 2 - \frac{1}{N}$. Combining Lemma 8 with Equations Eq. (10))-Eq. (12) completes the proof for the case of $p^N \geq \frac{1}{N}p$.

Case 2: Recall in this case, $p^N < \frac{1}{N}p$, where (p, p^N) are optimal prices corresponding to $\mathcal{R}^F_{1OPQ}$. We partition the valuation space under F according to the events

$$\begin{aligned}
E_0 &= \{\max\{V_i\} < p^N\} \\
E_1 &= \{p^N \leq \max\{V_i\} < p\} \\
E_2 &= \{p \leq \max\{V_i\} < \frac{N}{N-1}(p - p^N)\} \\
E_3 &= \{\frac{N}{N-1}(p - p^N) \leq \max\{V_i\}\}.
\end{aligned}$$

We upper bound the revenue from single opaque selling using this partition. Customers lying in E_0 do not generate any revenue. The revenue generated by customers lying in E_1 is at most $p^N \mathbb{P}(E_1)$ since they never consider buying an item at price p. The revenue generated by customers lying in E_3 is at most $p\mathbb{P}(E_3)$ since the best case scenario is that they all buy an item at price p. Lemma 9, proved separately, shows that the customers lying in E_2 buy the opaque product at price p^N. Combining the previous arguments shows that

$$\mathcal{R}_{1OPQ}(p, p^N) \leq p\mathbb{P}(E_3) + p^N(\mathbb{P}(E_2) + \mathbb{P}(E_1)). \tag{13}$$

Now suppose for contradiction that $\mathcal{R}_{1OPQ}(p,p^N) = \mathcal{R}_{1OPQ} > (2-\frac{1}{N})\mathcal{R}_{SP}$. Then the following two inequalities must also hold:

$$\mathcal{R}_{1OPQ}(p,p^N) > (2-\frac{1}{N})\mathcal{R}_{SP}\left(\frac{N}{N-1}(p-p^N)\right) = (2-\frac{1}{N})\left(\frac{N}{N-1}(p-p^N)\mathbb{P}(E_3)\right) \tag{14}$$

$$\mathcal{R}_{1OPQ}(p,p^N) > (2-\frac{1}{N})\mathcal{R}_{SP}\left(p^N\right) = (2-\frac{1}{N})(p^N)(1-\mathbb{P}(E_0)) \tag{15}$$

by the optimality of $\mathcal{R}_{SP}$. Now define $\delta' = (p-p^N)/p$. Then combining the Eq. (14) and Eq. (15) with Eq. (13) and dividing by p yields

$$1-\mathbb{P}(E_0)-\delta'(1-\mathbb{P}(E_3)-\mathbb{P}(E_0)) = 1-\delta'-(1-\delta')\mathbb{P}(E_0)+\delta'\mathbb{P}(E_3) > \frac{2N-1}{N-1}\delta'\mathbb{P}(E_3) \tag{16}$$

$$1-\mathbb{P}(E_0)-\delta'(1-\mathbb{P}(E_3)-\mathbb{P}(E_0)) = 1-\delta'-(1-\delta')\mathbb{P}(E_0)+\delta'\mathbb{P}(E_3) > \frac{2N-1}{N}(1-\delta')(1-\mathbb{P}(E_0)) \tag{17}$$

Rearranging Eq. (16) yields

$$1-\delta'-(1-\delta')\mathbb{P}(E_0) > \frac{N}{N-1}\delta'\mathbb{P}(E_3)$$

and rearranging Eq. (17) yields

$$1-\delta'-(1-\delta')\mathbb{P}(E_0) < \frac{N}{N-1}\delta'\mathbb{P}(E_3),$$

which is a contradiction and thus $\mathcal{R}_{1OPQ} \leq \frac{2N-1}{N}\mathcal{R}_{SP}$. For tightness when $N=2$, see Example 6. $\square$

J. Auxiliary Lemmas

LEMMA 6. *For any $x_1,\dots,x_N \geq 0$, $\sum_i x_i \leq K$, and constants $C_1, C_2 \geq 0$,*

$$\frac{C_1 + C_2\sum_i x_i}{C_1 + C_2\max_i ix_i} \leq \frac{C_1 + C_2 K H_N}{C_1 + C_2 K}. \tag{18}$$

Proof. We first claim that the left hand side of Eq. (18), viewed as an optimization problem over all feasible $\vec{x}$ is maximized when $ix_i = jx_j$ for all i,j. To prove this, suppose $\vec{x} \in [0,K]^N$ maximizes the left hand side of Eq. (18) and suppose the claim does not hold. Let $i = \arg\max_k kx_k$ and $j = \arg\min_k kx_k$, so by assumption $ix_i > jx_j$. Define y_i, y_j as solutions to the following system of two linear equations:

$$y_i + y_j = x_i + x_j$$

$$iy_i = jy_j$$

This yields $iy_i = jy_j = \frac{ij(x_i+x_j)}{i+j}$ which can be rewritten as $iy_i = iy_i = \frac{i}{i+j}jx_j + \frac{j}{i+j}ix_i$, a weighted average of ix_i and jx_j. Consider the maximal solution with components x_i, x_j replaced by y_i, y_j. Since $y_i + y_j = x_i + x_j$, the numerator in the l.h.s. of Eq. (18) is unchanged. However since $\max\{iy_i, jy_j\} < ix_i$, the denominator strictly decreases (in the case of many indices's that maximize kx_k, iterating the argument at most $N-1$ times yields a strict reduction) contradicting the optimality of $\vec{x}$. Thus Eq. (18) is maximized when $ix_i = jx_j \forall i,j$. Solving for the worst case $\vec{x}$ gives $x_i = \frac{x_1}{i}$, and plugging in gives

$$\max_{\vec{x}} \frac{C_1 + C_2\sum_i x_i}{C_1 + C_2\max_i ix_i} \leq \frac{C_1 + C_2 x_1 \sum_{i=1}^{N}\frac{1}{i}}{C_1 + C_2 x_1} \leq \frac{C_1 + C_2 K H_N}{C_1 + C_2 K}. \tag{19}$$

$\square$

LEMMA 7. *For any $x_1, \ldots, x_{N-1} \geq 0$, such that $\sum_i x_i \leq 1$, and constant $C \geq 0$,*

$$\frac{1 + C\sum_{i=1}^{N-1} ix_i}{1 + C\max_i i\sum_{k=i}^{N-1} x_k} \leq \frac{1 + CH_{N-1}}{1 + C}. \tag{20}$$

Proof. We first claim that the left hand side of Eq. (20), as a function of $\vec{x}$, is maximized when $i\sum_{k=i}^{N-1} x_k \leq (i+1)\sum_{k=i+1}^{N-1} x_k$ for all $i \leq N-2$. To prove this, let $\vec{x}$ be the maximizing vector and suppose the claim does not hold. For notational convenience, define $S_i = \sum_{k=i}^{N-1} x_k$, and let j be the smallest index such that $jS_j > (j+1)S_{j+1}$. Note by subtracting jS_{j+1} from both sides, it follows that j satisfies $jx_j > S_{j+1}$. Define a new vector $\vec{y}$ that is the same as $\vec{x}$ except for y_j, y_{j+1} which is the solution to the following system of equations:

$$jy_j = y_{j+1} + S_{j+2}$$
$$y_j + y_{j+1} = x_j + x_{j+1}.$$

Note that this has a solution where $y_j < x_j$ and $y_{j+1} > x_{j+1}$. We shall show that $\vec{y}$ results in a higher ratio than x, contradicting the optimality of $\vec{x}$. Since $y_j + y_{j+1} = x_j + x_{j+1}$ and $y_{j+1} > x_{j+1}$, then $jy_j + (j+1)y_{j+1} > jx_j + (j+1)x_{j+1}$ which implies that the numerator of Eq. (20) is strictly increased under $\vec{y}$. Next we argue that the denominator does not change. To see this, first observe that $i\sum_{k=i}^{N-1} y_k = i\sum_{k=i}^{N-1} x_k$ for all $i \neq j+1$. Since $y_j + y_{j+1} = x_j + x_{j+1}$ and $jy_j = y_{j+1} + S_{j+2}$, then $j\sum_{k=j}^{N-1} x_k = j(y_j + y_{j+1}) + jS_{j+2} = (j+1)y_{j+1} + (j+1)S_{j+2}$ and thus the denominator is unchanged under $\vec{y}$. Thus $\vec{y}$ has an increased the value of Eq. (20), contradicting the maximality of $\vec{x}$.

Thus we may assume w.l.o.g $kS_k \leq (k+1)S_{k+1}$ for any set $x_1, \ldots, x_{N-1}$ that maximizes Eq. (20), which in turn implies $kx_k \leq S_{k+1} \leq S_k$ for all $k \leq N$. Thus

$$\frac{1 + C\sum_{i=1}^{N-1} ix_i}{1 + C\max_i i\sum_{k=i}^{N-1} x_k} \leq \frac{1 + C\sum_{i=1}^{N-1} S_i}{1 + C\max_i iS_i} \leq \frac{1 + CS_1H_{N-1}}{1 + CS_1} \leq \frac{1 + CH_{N-1}}{1 + C}. \tag{21}$$

The first inequality follows from $kx_k \leq S_{k+1} \leq S_k$, the second inequality follows from Lemma 6, and the third in+equality from the fact that $S_1 = \sum_{k=1}^{N-1} x_i \leq 1$. $\square$

LEMMA 8. *$OPT \leq 2 - \frac{1}{N}$.*

Proof. First note for any optimal solution $v^* = (x, y, a, b, c, d)$ we may assume w.l.o.g. that $d = 0$ (if $d > 0$ consider v' where $(a', b', c') = (a, b, c)/(1-d)$ and d' = 0). We may also scale x to 1 w.l.o.g. Then OPT is the solution of

$$\begin{aligned} &\max \frac{a + y(1-a)}{z} \\ &\text{s.t } z \geq y \\ &\quad z \geq a + b \\ &\quad 0 \leq \frac{1}{N} \leq y \leq 1 \\ &\quad a + b + c = 1 \\ &\quad y, a, b, c \geq 0 \end{aligned}$$

At optimality either $z = y \geq a+b$ or $z = a+b \geq y$. Suppose $z=y$, then the objective becomes $1+\frac{a-ay}{y}$, which is maximized when a is maximal. Thus the constraint $a+b \leq y$ forces $b=0$ and $a=y$. Subbing in, then $OPT = \max_{y\geq \frac{1}{N}} 1+(1-y)\frac{y}{y} = 2-\frac{1}{N}$. Similarly, suppose $z = a+b \geq y$ which implies $c \leq 1-y$, then the objective becomes $\frac{a+y(b+c)}{1-c}$ which is maximized when $c = 1-y$, $a=y$, and $b=0$. Thus $OPT = \max_{y\geq\frac{1}{N}} \frac{y+y(1-y)}{y} = \max_{y\geq\frac{1}{N}} 2-y = 2-\frac{1}{N}$. $\square$

Lemma 9. *Any customer $\vec{v} \in E_2$ buys the opaque product in the 1OPQ strategy (p, p^N).*

Proof. Suppose a customer draws valuation $(v_1, v_2, \ldots, v_N) \in E_2$, then $\max_i\{v_i\} = p+k$, for some $k \in [0, \frac{N}{N-1}(p-p^N)-p)$. Then

$$\max_i\{v_i\} - p = k = \frac{k}{N} + \frac{(N-1)k}{N} < \frac{k}{N} + p - p^N - p\frac{N-1}{N} = \frac{k+p}{N} - p^N = \frac{\max_i v_i}{N} - p^N \leq \frac{\sum_i v_i}{N} - p^N$$

where the first inequality follows from the definition of k, and the second inequality follows from the fact that $\sum_i v_i \geq \max_i v_i$. Thus the utility from any item at price p, $\max_i\{v_i\}-p$, is less than $\frac{\sum_i v_i}{N} - p^N$, the utility from the opaque product. We also note that the utility from the opaque product is nonnegative, since $\frac{\sum_i v_i}{N} - p^N \geq \frac{k+p}{N} - p^N \geq 0$, where the last inequality follows from the case assumption $p^N < \frac{1}{N}p$. $\square$

Lemma 10. *Let F be a distribution over m customer types. Let $\vec{p}$ be a revenue optimal pricing and suppose that $p_1 \geq p_2 \geq \ldots \geq p_N$. Then for all $i \in [N]$, either there exists a type j such that $p_i = v_{j,i} - \max_{k>i}\{(v_{j,k}-p_k)^+\}$, or no customer type buys item i.*

Proof. Let $\vec{p}$ be the optimal prices with $p_1 \geq \ldots \geq p_N$, and let $\tilde{v}_{j,i} = v_{j,i} - \max_{k>i}(v_{j,k}-p_k)^+$. Suppose for a contradiction that there exists an item i (choose largest index if there are multiple options) that is purchased by a customer type, but $p_i \notin \{\tilde{v}_{1,i}, \ldots, \tilde{v}_{m,i}\}$. Call the customer type that purchases item i as type j, and if there are multiple options select the type with smallest $\tilde{v}_{\cdot,i}$. Note that $p_i < \tilde{v}_{j,i}$, since the reverse inequality implies type j would prefer to buy a different item (or no item) based on the definition of $\tilde{v}_{j,i}$.

We now consider an alternate pricing scheme $\vec{p}\,'$ where all prices are the same except $p'_i = \tilde{v}_{j,i}$, which is a price increase. Under $\vec{p}\,'$, clearly all customers who purchased an item other than i will still purchase that item due to the price increase of i. The type j customer will buy an item with index at most i, since his favorite among the items with index greater than or equal to i is i under the new pricing (recall ties go the higher priced item), i.e., $v_{j,i} - p'_i = \max_{k>i}(v_{j,k}-p_k)^+$. Now consider a type $l \neq j$ that also purchased i under the pricing $\vec{p}$. Then $v_{l,i} - p'_i = \tilde{v}_{l,i} + \max_{k>i}(v_{l,k}-p_k)^+ - p'_i \geq \tilde{v}_{j,i} + \max_{k>i}(v_{l,k}-p_k)^+ - p'_i = \max_{k>i}(v_{l,k}-p_k)^+$, where the inequality follows since j was chosen to have the smallest $\tilde{v}_{\cdot,i}$. Thus l, like j, also prefers an item with index i or lower. Therefore, $\vec{p}\,'$ is a pricing with strictly better revenue, resulting in a contradiction of the optimality of $\vec{p}$. $\square$