informs
http://pubsonline.informs.org/journal/msom

MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 24, No. 4, July–August 2022, pp. 1939–1958
ISSN 1523-4614 (print), ISSN 1526-5498 (online)

Distribution-Free Pricing

Hongqiao Chen,[a] **Ming Hu,**[b] **Georgia Perakis**[c]

[a] School of Management & Engineering, Nanjing University, Nanjing 210093, China; [b] Rotman School of Management, University of Toronto, Toronto, Ontario M5S 3E6, Canada; [c] Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Contact: chq@nju.edu.cn, https://orcid.org/0000-0002-6776-6954 (HC); ming.hu@rotman.utoronto.ca, https://orcid.org/0000-0003-0900-7631 (MH); georgiap@mit.edu, https://orcid.org/0000-0002-0888-9030 (GP)

Received: March 5, 2020
Revised: July 28, 2020; September 22, 2021
Accepted: October 6, 2021
Published Online in Articles in Advance: January 20, 2022

https://doi.org/10.1287/msom.2021.1055

Abstract. ***Problem definition*:** We study a monopolistic robust pricing problem in which the seller does not know the customers' valuation distribution for a product but knows its mean and variance. ***Academic/practical relevance*:** This minimal requirement for information means that the pricing managers only need to be able to answer two questions: How much will your targeted customers pay on average? To measure your confidence in the previous answer, what is the standard deviation of customer valuations? ***Methodology*:** We focus on the maximin profit criterion and derive distribution-free upper and lower bounds on the profit function. ***Results*:** By maximizing the tight profit lower bound, we obtain the optimal robust price in closed form as well as its distribution-free, worst-case performance bound. We then extend the single-product result to study the robust pure bundle pricing problem where the seller only knows the mean and variance of each product, and we provide easily verifiable, distribution-free, sufficient conditions that guarantee the pure bundle to be more robustly profitable than à la carte (i.e., separate) sales. We further derive a distribution-free, worst-case performance guarantee for a heuristic scheme in which customers choose between buying either a single product or a pure bundle. Moreover, we generalize separate sales and pure bundling to a scheme called clustered bundling that imposes a price for each part (i.e., cluster) of a partition of all products and allows customers to choose one or multiple parts (i.e., clusters), and we provide various algorithms to compute clustered bundling heuristics. In parallel, most of our results hold for the minimax relative regret criterion as well. ***Managerial implications*:** The robust price for a single product is in closed form under the maximin profit or minimax relative regret criterion and hence, is easily computable. Its interpretation can be easily explained to pricing managers. We also provide efficient algorithms to compute various mixed bundling heuristics for the multiproduct problem.

Funding: H. Chen was supported by the National Natural Science Foundation of China [Grants 71971108 and 72171110] and the China Scholarship Council [Grant 201504910540]. M. Hu was supported by the Natural Sciences and Engineering Research Council of Canada [Grants RGPIN-2015-06757 and RGPIN-2021-04295].
Supplemental Material: The online appendix is available at https://doi.org/10.1287/msom.2021.1055.

Keywords: pricing • robust optimization • performance guarantee • bundling • separate sales • clustering

1. Introduction

We study a monopoly pricing problem in which a seller has to decide on and commit to a price with only limited knowledge of the customer valuation distribution. This problem arises when the seller launches an innovative product with limited market information or when the seller needs to offer a personalized price to a one-time customer about whom the seller knows little. Similar problems also occur when a firm launches a new subscription service, with a bundle of goods and services, for which detailed information about customer valuations may be hard to obtain. Such pricing problems with limited information about customers are extremely challenging yet crucial for the seller's profitability. Despite the lack of detailed information, the seller is likely to know how to answer the following two questions. First, how much will the targeted customers pay on average? Second, to measure the confidence level of the previous answer, what is the standard deviation of customer valuations? Technically, these two questions can tease out information about the first and second moments of the customer valuation distribution. With answers to those two questions about the demand, as well as the known cost information on the supply side, we solve for the robust pricing problem under various criteria. Under the focal criterion referred to as the maximin profit criterion, the seller sets a price to maximize its profit in the worst scenario among all customer valuation distributions that share the same mean and standard deviation. Under an alternative criterion referred to as the minimax relative (or absolute) regret criterion, the seller minimizes the maximum relative (or absolute) regret because of not knowing the true distribution as opposed to possessing this information. In any case, we search for the most robust pricing solution.

Under the maximin profit criterion, we provide a robust price that is conservative in protecting the seller against the worst possible profit. The lower bound on the profit for the robust price is (asymptotically) achievable; hence, the robust price is the *best* against all odds. Furthermore, this optimal robust price is in closed form and hence, easily computable. It depends only on the mean and standard deviation of the customer valuation distribution, teased out from the two questions, and on the product's marginal cost. (Hereafter, we refer to the knowledge of just the mean and variance about the valuation distribution as "limited information.") We show that the smaller the ratio of the standard deviation to the mean of the customer valuation distribution (i.e., the coefficient of variation (CV)), the higher the optimal robust price and the better its performance guarantee. The higher the marginal cost, the higher the optimal robust price and the worse its performance guarantee. That is, on the demand side, the CV of the valuation distribution provides a relative measure of the concentration of the customer valuations. The seller needs to set a lower optimal robust price if the valuation distribution is more dispersed. On the supply side, the marginal cost plays a critical role in determining how aggressive the seller should be when uncertain about customer valuation. With only limited information about the demand side, there is less room for a seller to make a mistake (i.e., in the range of profitable feasible prices) when the production cost is higher. As a result, the performance guarantee of the optimal robust price becomes worse. The optimal robust price under the minimax relative regret criterion has a similar closed form, and hence, it has the same monotonicity properties as those under the focal maximin profit criterion, although the latter is shown to be lower than the former.

We further investigate the benefit of having less or more information than the mean and standard deviation of the valuation distribution. If the seller knows only the mean and range of the valuation distribution, the worst case that results in the largest standard deviation is a two-point valuation distribution with the two ends of the range as the possible realizations. Then, the robust pricing problem can be reduced to the one studied in our base model. Moreover, we can obtain an optimal robust price with a sharper performance guarantee if the seller also knows that the optimal price is below the mean valuation, the valuation distribution is symmetric or unimodal, or the third moment information is revealed. However, having such additional information may be of limited marginal benefit. The intuition is that it may not help eliminate the worst case of a bipolar valuation distribution.

A seller may encounter the same robust pricing problem in selling multiple products. In this situation, the seller knows only the mean and standard deviation of the customer valuation distribution for each product. The seller may sell the products separately (as "separate sales") or may choose to sell all products in a pure bundle (referred to as "pure bundling") by setting a bundle price. Under the focal maximin profit criterion, we apply the optimal robust price from the single-product problem and compare the performance of pure bundling with that of separate sales in the limited information setting. The idea that "bundling reduces dispersion" and improves profitability is not new (see, e.g., Adams and Yellen 1976). However, previous formal statements of the idea rely on specific assumptions about valuation distributions. In contrast, we provide easily verifiable sufficient conditions under which a bundle is guaranteed to be more profitable than separate sales without making any distributional assumption. In particular, we show that as long as the cost structure in terms of profit margins and the demand characteristics in terms of CVs are the same for all products, the bundle is more robustly profitable than separate sales, regardless of the detailed valuation distributions and how they are correlated. We also prove that the optimal robust bundle price performs better than separate sales if the CV of the bundle valuation is smaller. This can be achieved when the bundle is reasonably large in spite of having nonzero marginal costs. We illustrate the performance of our robust bundle price with three practical examples: the streaming subscription service, Spotify, for selling digital goods; a leading information technology (IT) company offering a personalized bundle of hardware and software to commercial customers; and a theater company selling season tickets to patrons who have correlated valuations for plays. Furthermore, we derive a distribution-free, worst-case performance guarantee for pricing under a heuristic scheme in which customers choose between buying either a single product or a pure bundle. Lastly, we generalize separate sales and pure bundling to a novel scheme, which we refer to as clustered bundling that imposes a price for each part (referred to as a cluster) of a partition of all products and allows customers to choose one or multiple parts (i.e., clusters), and provide various algorithms to compute such clustered bundling heuristics. In parallel, all of our results except those of Section 4.2, in the multiproduct problem, hold for the minimax relative regret criterion as well.

1.1. Contributions

We make the following contributions. First, we provide a performance guarantee under the focal maximin profit criterion for a practical robust pricing problem with limited information and a general marginal cost. The optimal robust price is in closed form (hence, easily computable) and achieves a *tight* performance lower bound. We also show that having additional information beyond the first and second moments of the valuation distribution may be of little additional value in protecting the seller from the worst-case scenario. Second, we fully characterize the optimal robust price that minimizes the maximum relative or absolute regret, which is shown to be higher than that under the maximin profit criterion, whereas the optimal robust price under the minimax relative regret criterion has a closed form similar to that under the focal maximin profit criterion. This result demonstrates how various distribution-free objectives may result in different optimal robust pricing decisions even with the same information. Third, for the multiproduct problem, we give theoretical support to the profitability of the pure bundle over separate sales without making any detailed distributional assumptions. That serves as a formalization of the idea that "bundling reduces dispersion" in a distribution-free sense, without the need to resort to any limiting argument on the bundle size. Finally, we provide various heuristics for the mixed bundling problem. In particular, we derive a distribution-free, worst-case performance guarantee for a heuristic scheme in which customers choose between buying either a single product or a pure bundle. We also generalize separate sales and pure bundling to clustered bundling and provide various algorithms to compute clustered bundling heuristics. Because of the similarity between the closed forms of the optimal robust price under the maximin profit and minimax relative regret criteria, our results on comparing pure bundling and separate sales derived for the former criterion hold for the latter one as well.

2. Literature Review

Our work belongs to the stream of research on robust pricing. This stream studies a *one-shot* pricing decision with limited market information. This problem is more appropriate in practice when the firm has to commit to a price before the selling season starts or has to quote a personalized price to a one-time customer. In those situations, the firm has no opportunity to learn the market and update the price, but it has to make a decision with the limited information it has about customers. Thiele (2008) studies a monopolistic pricing problem in which the seller knows the number of potential buyers and their random valuation support. Cohen et al. (2021) study a robust pricing model without knowing the specific form of the demand function. The authors derive analytical performance bounds for a variety of demand functions with only information about the marginal cost and the *maximum price* (i.e., choke price) at which no customers would want to buy. In contrast, we impose limited information assumptions directly on the customer valuation instead of on the demand function.[1] In a similar fashion to our paper, Bergemann and Schlag (2011) consider a customer valuation distribution that is unknown but is assumed to be in a neighborhood of a given distribution. In contrast, we assume that the pricing managers only know the mean and standard deviation and have no further detailed information about the valuation distribution. In the operations literature, this information structure is often referred to as "distribution free." Similarly, Bergemann and Schlag (2008) consider the pricing problem in which the seller only knows the support of the possible valuations and has no further distributional information. However, the authors do not give a performance guarantee of the optimal robust price as we do. We show that the robust pricing problem with information about the valuation mean and support can be transformed into our problem. Azar and Micali (2013) study revenue maximization for auctions of digital goods with zero marginal costs and give a posted price mechanism that maximizes revenue in the worst case under the knowledge of the mean and variance of each bidder's marginal distribution. In contrast, we study a robust pricing problem with nonnegative marginal costs.

More recently, Carroll (2017) studies a robust mechanism design problem in which a principal wants to screen an agent along several dimensions of private information. With the application of his results to the monopoly pricing of multiple products with private customer valuations, the author shows that, if the principal knows the marginal valuation distribution of each product but does not know the joint distribution, separate sales are the most robust solution even when mixed bundling is allowed, in the sense of achieving the highest worst-case expected profit among all joint distributions consistent with the known marginals. If the seller knows only the mean and variance of the marginal distribution of each product, for a *given* correlation structure, we show that selling a bundle is robustly more profitable than separate sales under some mild conditions. If the correlation structure turns out to be the worst for bundling, in which case all products are positively correlated, bundling is as effective as separate sales, which is consistent with Carroll (2017). However, bundling is robustly more profitable than separate sales for correlation structures other than the worst case, which restores the robust profitability of bundling over separate sales in a more general sense. Koçyiğit et al. (2021) study a robust multiproduct pricing problem with a rectangular uncertainty set for customer valuations by minimizing the maximum absolute regret and show that the optimal robust selling mechanism is to sell goods separately. In contrast, we consider such a problem with moments information under the maximin profit and minimax relative regret criteria and show when pure bundling is more robust than separate sales.

The papers most closely related to our work are Kos and Messner (2015) and Carrasco et al. (2018). Both papers study revenue maximization of a single product with the product cost assumed away and adopt the maximin revenue criterion, similar to our maximin profit criterion. Kos and Messner (2015) assume the information structure of knowing the range of the mean valuation and the upper bound of the valuation support. The authors provide the optimal robust deterministic price and randomized price distribution. Our work assumes that the seller knows precisely the first two (or three) moments of the valuation distribution and focuses on a deterministic posted pricing policy for a nonnegative product cost. In a more general setting than ours, Carrasco et al. (2018) study the optimal mechanism design problem of a seller who is partially informed about the distribution of the buyer's valuations by only knowing its first N moments. The authors characterize the optimal mechanism for $N = 2$ that contains an allocation rule and a transfer function, and also, they obtain an optimal robust price with the product cost being zero.[2] We complement Carrasco et al. (2018) by focusing on the optimal robust (simple-to-implement posted) price with a nonzero product cost and obtain a tight lower profit bound. We also solve the robust pricing problem under alternative criteria and compare various optimal robust prices. We then extend the single-product result to study the robust pricing problems of selling multiple products through separate sales, pure bundling, or mixed bundling heuristic schemes.

We apply our optimal robust price to the (mixed) bundle pricing problem. In the economics literature, it is well known that bundling can be an effective way for a multiproduct monopolist to increase profits. The early economics papers focus on a two-product monopoly problem because of the difficulty of analyzing cases of multiple products. The work on bundling with more than two products emerges in the literature on information goods, which have small or even no marginal costs. In particular, Bakos and Brynjolfsson (1999) consider a monopolist selling a large number of information goods with *zero* marginal costs where the consumer valuations of those goods are independent and identically distributed (i.i.d.). By the weak law of large numbers, they show that "bundling very large numbers of unrelated information goods can be surprisingly profitable" (Bakos and Brynjolfsson 1999, p. 1913). We calculate a closed-form robust price for pure bundling of a fixed bundle size with an easily computable performance guarantee. Our approach and results go beyond zero marginal costs and the i.i.d. valuation assumption. We also do not require any limiting argument on the bundle size (see, e.g., Abdallah 2019, Abdallah et al. 2021).

Because it is difficult to compute the optimal mixed bundling prices with a number of, say n, products (with $2^n - 1$ prices to be decided), it is desirable to offer a simple scheme with a small number of prices to determine yet one that still captures most of the optimal profit of the mixed bundling. Babaioff et al. (2014) consider a monopolist seller who sells n heterogeneous products to a buyer; they show that selling each product separately or all products together as a bundle can achieve at least a fraction of one-sixth of the optimal mixed bundling pricing. Chu et al. (2011) propose a pricing scheme called bundle-size pricing (BSP), where the price for a bundle option depends only on the size of the bundle; they show that the optimal BSP scheme tends to yield profits close to the optimal profit of mixed bundling as calculated by numerical experiments. Li et al. (2022) show that the BSP problem can be approximated by a convex optimization problem using only the first and second moments of customer valuations; then, they numerically solve the convex optimization problem for bundle pricing heuristics. In contrast, we provide provable performance guarantees for our pricing heuristics. Ma and Simchi-Levi (2015) introduce a bundling mechanism called "pure bundling with disposal for cost" (PBDC). Their main purpose is to provide a theoretical guarantee for the performance of PBDC that holds for arbitrary independent distributions by using techniques from the mechanism design literature. In particular, it is shown that either PBDC or individual sales will obtain at least 1/5.2 of the optimal profit among all incentive-compatible and individually rational mechanisms.

In economics, there are papers that compare pure bundling and separate sales (see, e.g., Chakraborty 1999, Fang and Norman 2006). These papers make restrictive assumptions either about the number of products or about the detailed valuation distributions. For example, Fang and Norman (2006) show that bundling is more profitable than separate sales with a given finite number of goods, provided that the distributions of valuations are i.i.d. and follow a symmetric and log-concave distribution. Furthermore, they assume that all products have the same marginal cost and that the optimal separate-sales price is no more than the mean valuation for any product. In contrast, without making any detailed distributional assumptions, we impose easily verifiable sufficient conditions on the marginal costs and the means and variances of valuation distributions, under which selling a bundle is guaranteed to generate higher profits than selling products separately in a robust sense. One can view our results as formalizing the idea that "bundling reduces dispersion" in a robust sense because essentially, our sufficient conditions on the valuation distributions of individual products amount to requiring the CV of the bundle valuation to be less than those of individual products.

Finally, our work is related to robust decision making in operations, such as the optimal pricing or inventory decision under three robust criteria; one is to maximize the seller's worst-case utility, and the other two are to minimize the seller's worst-case relative/absolute regret. In the newsvendor context, Gallego and Moon (1993) consider the maximin profit criterion, whereas Perakis and Roels (2008) adopt the minimax absolute regret criterion. In the pricing context, Cohen et al. (2021) use the maximin profit criterion, whereas Allouah et al. (2021) study a robust pricing problem with sample valuations under the minimax relative regret criterion. In our setting, we analyze and compare all three criteria.

3. The Model

Consider a monopoly pricing problem. The seller sells a product to a market of customers with heterogenous valuations. The valuation (or willingness to pay) V of a randomly selected customer in the targeted market follows a distribution with cumulative distribution function (c.d.f.) $F(v)$ over a nonnegative support, its probability density function $f(v)$, mean μ, and standard deviation σ. The seller incurs a constant marginal cost c. For any price p, the corresponding profit is

$$\pi(p;F) = (p-c)\mathsf{P}(V \geq p) = (p-c)(1 - \lim_{v \to p^-} F(v)).$$

In the case of a continuous distribution F, $\pi(p;F) = (p-c)(1-F(p))$. Our results are generic to accommodate continuous and discrete distributions $F(\cdot)$.

Our model setup is the same as in the classical monopoly pricing problem, except that the seller does not know the customers' exact valuation distribution $F(v)$ when it sets the price. When an innovative product or service is being sold, although the exact valuation distribution may be hard to obtain, the mean and standard deviation information can be much easier to estimate. It essentially amounts to asking the pricing managers the following two questions.

(i) How much will your targeted customers pay on average? (information about μ)

(ii) How sure are you? (information about σ)

Thus, we assume that the seller has an idea about the finite first and second moments of the valuation distribution (i.e., μ and $\sigma > 0$).[3] (If $\sigma = 0$, the demand is completely deterministic, and the optimal price is simply μ.)

We denote by $\delta \equiv \sigma/\mu$ the CV of the valuation distribution and by $\gamma \equiv c/\mu$ the ratio of the marginal cost to the mean valuation. Because the term $(1/\gamma - 1)$ relates to the potential profit margin of the product, a low (high) value of γ corresponds to a high (low) potential profit margin. The unit-free measure δ is on the demand side, and the unit-free measure γ is on the supply side. We further define another unit-free measure that combines the measures on both the demand and supply sides: for $\sigma > 0$, denote

$$\tau \equiv \frac{1-\gamma}{\delta} = \frac{\mu - c}{\sigma}.$$

The following assumption is the only one we need for our results.

Assumption (P). $\mu \geq c$ (*i.e.*, $\gamma \leq 1$).

This assumption is satisfied for digital goods whose reproduction/distribution may incur almost zero costs (i.e., $c \approx 0$). In general, this assumption is innocuous because μ is the average valuation of the *targeted* market. Suppose it does not hold. Then, even if the seller can perfectly price discriminate every customer *within* the targeted valuation range such that each one pays at their own valuation, the expected profit is negative. This suggests that it is beneficial for the seller to exclude some very low-valuation customers from its targeted market. Assumption (P) results in $\tau \geq 0$.

3.1. The Focal Maximin Profit Criterion

Let $\mathscr{F}$ be the class of all (continuous or discrete) distributions that share the same mean $\mu \geq c$ and standard deviation $\sigma > 0$ and are supported on a nonnegative real line. In particular, $\mathscr{F} \equiv \{F(v), v \geq 0 \mid \mathsf{E}[V] = \mu, \mathsf{E}[V^2] = \mu^2 + \sigma^2\}$. Armed with this set, our goal is to solve for an optimal robust price

$$p^* = \arg\max_p \{\min_{F \in \mathscr{F}} \pi(p;F)\}$$

and provide a worst-case performance guarantee for this optimal robust price p^*, under the focal *maximin profit criterion* (see, e.g., Bergemann and Schlag 2011, Kos and Messner 2015, Carrasco et al. 2018, Cohen et al. 2021). This criterion protects the seller against the worst-case scenario in terms of the absolute profit value, which could be a *survival* measure for the seller. We focus on this maximin profit criterion and consider two alternative criteria in Section 5. We will show that the results obtained under the maximin profit criterion carry over to the minimax relative regret criterion because their optimal robust prices have similar closed-form expressions.

To provide a distribution-free performance guarantee (in terms of the primitives c, μ, and σ), we first derive a distribution-free upper bound on the optimal profit. Then, we provide a *tight* distribution-free lower bound for any given price p based on which we obtain an optimal robust price p^* by maximizing the lower bound as a function of p. As a result, the performance lower bound of the optimal robust price p^* can be again expressed in terms of the system primitives. Lastly, we show the performance guarantee of this optimal robust price by comparing the distribution-free lower and upper bounds.

Theorem 1 (Distribution-Free Bounds). *We have the following bounds.*

(a) (*Upper bound*) *For all* $F \in \mathscr{F}$, $\max_p \pi(p;F) \le \mu - \frac{\tau^2}{1+\tau^2} c \equiv U$.

(b) (*Lower bound*) *For all* $F \in \mathscr{F}$, $\max_p \pi(p;F) \ge \pi(p^* = \mu - k^*\sigma; F) \ge (\mu - c)(1 - \frac{3}{2\tau}k^*) \equiv L$, *where*

$$k^* \equiv \sqrt[3]{\tau + \sqrt{\tau^2+1}} + \sqrt[3]{\tau - \sqrt{\tau^2+1}} \ge 0.$$

Moreover, the lower bound is asymptotically achievable by a series of two-point distributions:

$$V = \begin{cases} \mu - k_\epsilon \sigma & \text{with probability } \dfrac{1}{1+k_\epsilon^2}, \\ \mu + \dfrac{1}{k_\epsilon}\sigma & \text{with probability } 1 - \dfrac{1}{1+k_\epsilon^2}, \end{cases}$$

with $k_\epsilon = k^* + \epsilon$, *as* $\epsilon \searrow 0$.

Theorem 1(b) shows the lower bound is "tight" in an asymptotic sense. This tightness notion follows from Bertsimas and Popescu (2005, definition 1.2). Specifically, the lower bound can be asymptotically achievable by a series of two-point distributions. In a different setting, Gallego and Moon (1993) model the robust newsvendor problem as that of finding the size of order that maximizes the expected profit against the worst possible demand distribution. In their paper, only the mean μ and the variance σ^2 of the demand are known without any further assumptions about the form of the distribution. The authors show that the worst possible distribution is a two-point distribution (see Gallego and Moon 1993, remark 2). Theorem 1(b) suggests that the worst possible distribution for our model is also a two-point distribution, analogous to the robust newsvendor problem. In the worst case, the seller faces two distinct customer segments, one with a valuation below μ and the other with a valuation above μ.

Taking the ratio of the distribution-free lower bound L to the upper bound U, we have the following distribution-free performance guarantee in closed form. From the sketch of the proof of Theorem 1, we can see that U is, in fact, also an upper bound on the total profit when every profitable customer (whose valuation is more than c) pays their own valuation. Hence, this upper bound U is also an upper bound on the total profit of fully personalized pricing. As a result, the following performance guarantee is also one on the optimality of using a single price benchmarked against fully personalized pricing. Elmachtoub et al. (2021) study such a problem with the seller knowing the customer valuation distribution, whereas Cohen et al. (2022) study the problem of customized pricing for different groups under fairness constraints.

Theorem 2 (Robust Price and Performance Guarantee). *The optimal robust price is* $p^*(c,\mu,\sigma) = \mu - k^*\sigma$ *with the safety factor* $k^* = \sqrt[3]{\tau + \sqrt{\tau^2+1}} + \sqrt[3]{\tau - \sqrt{\tau^2+1}} \ge 0$, *which achieves the following performance guarantee*: *For all* $F \in \mathscr{F}$,

$$\frac{\pi(p^*;F)}{\max_p \pi(p;F)} \ge \frac{1 - \frac{3}{2\tau}k^*}{1 + \frac{\gamma}{(1-\gamma)(1+\tau^2)}} \equiv \rho.$$

The form of the optimal robust price is analogous to the optimal order quantity for a newsvendor facing a normally distributed demand distribution $N(\mu, \sigma)$: $Q^* = \mu + z^*\sigma$, where $z^* = \Phi^{-1}$(critical fractile) and $\Phi^{-1}(x)$ is the inverse function of the c.d.f. of the standard normal distribution $N(0, 1)$. The safety factor z^* determines the additional amount of inventory the newsvendor needs to hold beyond the expected demand to cope with demand uncertainty. Analogously, in our robust pricing setting, k^*, as a *safety factor*, determines the discount the seller needs to offer below the expected valuation to cope with valuation uncertainty. In the newsvendor setting, being conservative means having an order quantity *beyond* the mean demand, whereas in the robust pricing setting, being conservative means charging a price *below* the mean valuation. It is easy to see that the safety factor k^*, only depending on τ, is increasing in τ. That is, the higher the τ value (i.e., the lower the potential profit margin or the larger the variability in the valuation distribution), the larger the safety factor in robust pricing should be and hence, the lower the optimal robust price.

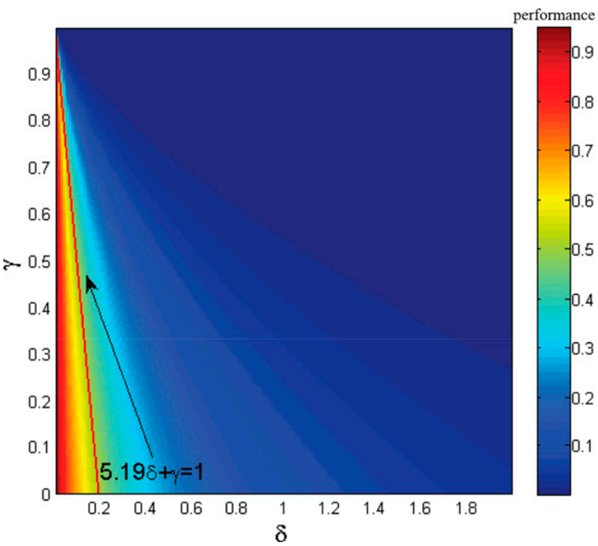
Corollary 1 (Monotonicity Properties). *The following monotonicity properties hold.*

(a) *The performance guarantee* $\rho(c,\mu,\sigma) = \rho(\gamma,\delta)$ *is strictly decreasing in* γ *and* δ.

(b) *The optimal robust price* $p^*(c,\mu,\sigma) = \mu - k^*\sigma$ *is strictly increasing in* c *and* μ *and is strictly decreasing in* σ.

Corollary 1(a) first says that the performance guarantee ρ is determined by the two unit-free measures γ and δ. The measure γ captures the potential profitability of making a sale, and the measure δ quantifies the variability in the valuation distribution of the targeted market. Corollary 1(a) further says that the proposed optimal robust price is guaranteed to perform better when the marginal cost is lower or the valuation distribution has lower variability. The former statement suggests that the optimal robust price is more likely to perform well for digital goods than physical products because digital goods tend to have lower reproduction and distribution costs. The latter statement, which is intuitive, suggests that the more accurate market information the seller has, the higher profits can be guaranteed. Figure 1 shows the guaranteed performance as a function of γ and δ, with γ ranging from zero to one and δ ranging from zero to two. As the

Figure 1. (Color online) Performance Guarantee



upper bound U may not be tight, the actual performance of the optimal robust price can be better than the displayed value. From Figure 1, we see that the provable performance is good when γ and δ are relatively small. For example, to the left of the line defined by the equation $5.19\delta + \gamma = 1$, the provable performance guarantee is more than 50%. In Section 4, we will apply the single-product optimal robust price to the bundling pricing problem and show that the performance of a robust heuristic for the pure bundle can be very good for even large values of δ of individual products when the bundle is medium-sized.

Corollary 1(b) implies that if there is more uncertainty in the valuation distribution (measured by its CV $= \delta = \sigma/\mu$), the optimal robust price should be more aggressive (i.e., lower). Bergemann and Schlag (2011) obtain similar insights under a different formulation of the uncertainty set.

Our optimal robust price is designed against the worst possible situation selected by the nature. Next, we will demonstrate its performance for five commonly used distributions: exponential, uniform, truncated normal, truncated logit, and log-normal distributions. These distributions are the same distributions adopted by Chu et al. (2011) to represent various valuation distributions when the performance of a heuristic bundle pricing policy was tested. Because we know that the true valuation distribution is an exponential or uniform distribution, their optimal price can be easily expressed in closed form. Then, we can compare the profit of our optimal robust price without knowing the true distribution to the optimal profit under the knowledge of the true distribution. In particular, we can obtain tighter bounds on the performance of the optimal robust price for those specific distributions as follows.

Proposition 1.

(a) *For an exponential distribution with parameter λ as the underlying distribution, the profit performance of the optimal robust price is $(1-(k_e(\gamma)+\gamma))e^{(k_e(\gamma)+\gamma)}$, where $k_e(\gamma)$ is the unique real solution to $k^3 + 3k = 2(1-\gamma)$. In particular, if $\gamma = 0$, the profit of our heuristic captures at least 73.31% of the optimal profit under the knowledge of the true distribution.*

(b) *For a uniform distribution on $[0, a]$ as the underlying distribution, the profit performance of the optimal robust price is $(1-k_u(\gamma)/\sqrt{3}-\gamma)(1+k_u(\gamma)/\sqrt{3})/(1-\gamma/2)^2$, where $k_u(\gamma)$ is the unique real solution to $k^3 + 3k = 2\sqrt{3}(1-\gamma)$. In particular, if $\gamma = 0$, the profit of our heuristic captures at least 72.61% of the optimal profit under the knowledge of the true distribution.*

See Proposition S.1 in the online supplement for analogous results of Proposition 1 for normal, logit, and log-normal valuation distributions that do not have a closed form characterization of the optimal price. Figure 2 shows the performance of the optimal robust price for five distributions: exponential distribution ($\delta = 1$), uniform distribution ($\delta = 1/\sqrt{3}$), truncated normal distribution with mean 0 and standard deviation 0.5 for the original distribution before truncation, truncated logit distribution with mean 0.125 and scale parameter 0.25 for the original distribution, and log-normal distribution with mean 1 and standard deviation 0.5. The latter three distributions serve as a benchmark for the following numerical studies that adopt parameters of magnitude similar to Chu et al. (2011, table 2). Figure 3(a) shows the performance for truncated normal distributions with $\overline{\sigma} = 0.5$ and $\overline{\mu}$ ranging from −1 to 2.5 for the original distribution. Figure 3(b) shows the performance for truncated logit distributions with $\overline{\sigma} = 0.25$ and $\overline{\mu}$ ranging from 0 to 2.5. Figure 3(c) shows the performance for log-normal distributions with $\overline{\mu} = 1$ and $\overline{\sigma}$ ranging from 0.5 to 2. From these examples, we see that the optimal

Figure 2. (Color online) Performance for Five Distributions

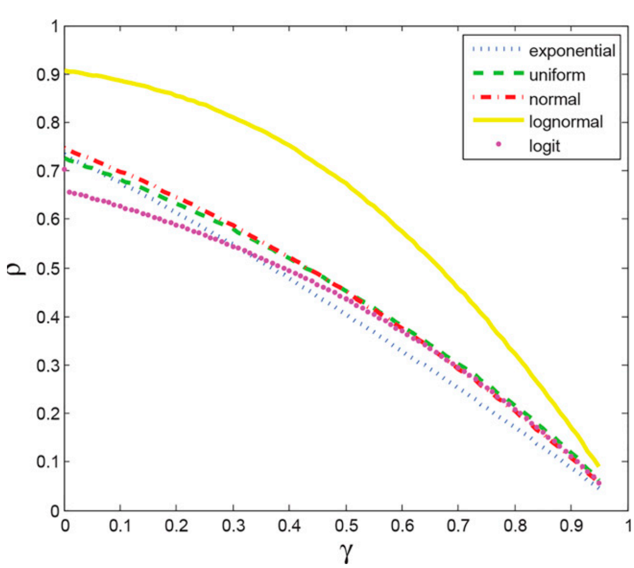
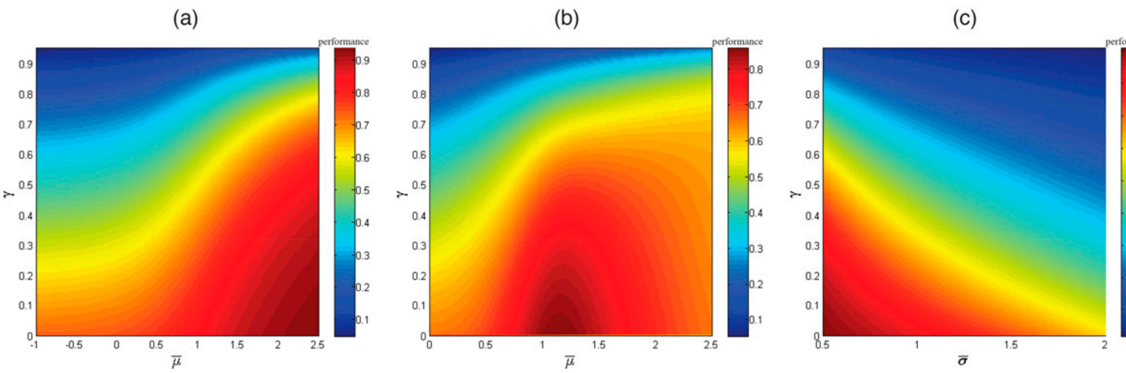


Figure 3. (Color online) Performance for Commonly Used Distributions



Notes. (a) Truncated normal distribution with $\overline{\mu}$=0.5 and $\overline{\mu}$ ranging from −1 to 2.5. (b) Truncated logit distribution with $\overline{\sigma}$=0.25 and $\overline{\mu}$ ranging from 0 to 2.5. (c) Log-normal distribution with $\overline{\mu}$=1 and $\overline{\sigma}$ ranging from 0.5 to 2.

robust price performs decently for commonly used distributions when γ and δ are not too large.

3.2. Alternative Information Structures

We study the situation where the seller may have less or more information than the mean and variance of the valuation distribution that we assumed in the base model. The general approach is similar to that in the base model. That is, we obtain a price heuristic by maximizing a distribution-free lower bound. As the lower bound may not be tight for these alternative information structures, in this subsection we no longer refer to a price heuristic as the optimal robust price.

3.2.1. Valuation Support. Suppose the seller only knows the mean μ and the support $[v_L, v_U]$ of the customer valuation. The latter is a form of information commonly seen in the robust optimization literature. With these two pieces of information, we can find the scenario that has the largest standard deviation. Because the larger the standard deviation, the worse the performance guarantee for that standard deviation, we can then apply our robust price formula from the base model to the largest standard deviation to get a performance guarantee against the worst-case scenario.

Proposition 2. *Denote by $\mathscr{F}_r$ the class of valuation distributions that share the same mean μ and support $[v_L, v_U]$. For all $F \in \mathscr{F}_r$, the price heuristic $p_r^* = p^*\left(c, \mu, \sigma = \sqrt{\mu(v_L + v_U) - v_L v_U - \mu^2}\right)$ achieves the performance guarantee $\rho\left(\gamma, \delta = \sqrt{\mu(v_L + v_U) - v_L v_U - \mu^2}/\mu\right)$, where $p^*(c, \mu, \sigma)$ is specified in Theorem 2 and $\rho(\gamma, \delta)$ is introduced in Corollary 1.*

In practice, the pricing managers may often have more information, or a hunch, about the valuation distribution than we assumed in the base model. In the rest of this subsection, we focus on how much benefit these additional pieces of information might bring.

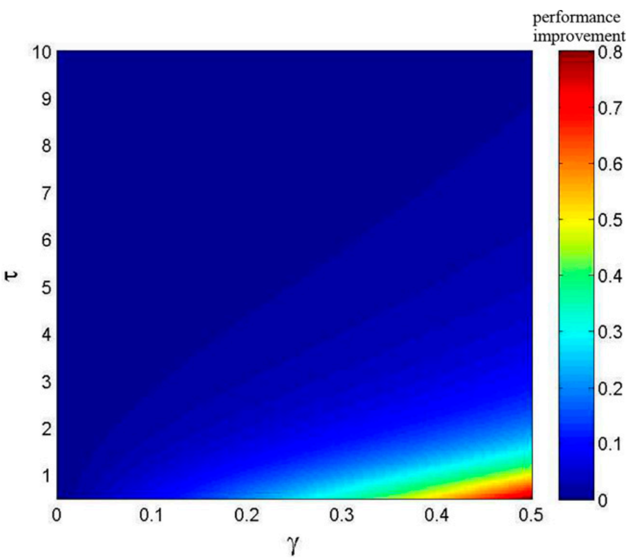
3.2.2. Optimal Price Lower Than Mean. So far, we have obtained a performance guarantee under the robust price $p^* = \mu - k^*\sigma$ that is no more than the mean valuation μ. For those valuation distributions whose optimal price is known to be also no more than μ (which is assumed in Fang and Norman 2006), our robust price heuristic can achieve a provably sharper performance guarantee.

Proposition 3. *Denote $\mathscr{F}_l = \{F \in \mathscr{F} \mid \min\{\arg\max_p \pi(p; F)\} \le \mu\}$. For all $F \in \mathscr{F}_l$, the price heuristic $p^* = \mu - k^*\sigma$ achieves a strictly sharper performance guarantee $\min\{1 - k^*/\tau, 1 - 1/[(k^*)^2 + 1]\}$, of which the relative improvement beyond the performance guarantee in Theorem 2 is no less than $\gamma/[(1-\gamma)(1+\tau^2)]$.*

Proposition 3 shows that there is a benefit in knowing that one of the optimal prices is no more than the mean valuation.[4] The relative improvement in performance guarantee with the *same* price heuristic p^* is at least $\gamma/[(1-\gamma)(1+\tau^2)]$, which is increasing in γ and δ. Thus, the benefit of such additional information to improving the performance guarantee tends to be higher if the potential profit margin is lower or the variability of the valuation distribution is higher. This is because in those situations, the original performance guarantee can be poor, and there is much room for improvement with extra information. Figure 4 displays the relative improvement lower bound $\gamma/[(1-\gamma)(1+\tau^2)]$. From this figure, we observe that for a large part of the parameter space (e.g., when τ is large enough or γ is small enough), the performance improvement tends to be less than 10%.

3.2.3. Distribution Symmetry or Unimodality. In Online Appendix A, we impose two restrictions on the

Figure 4. (Color online) Performance Improvement If the Optimal Price Is Known to Be Lower Than μ



valuation distributions, symmetry, or unimodality (i.e., there is one primary market segment), and show that within these special classes of distributions, we can obtain a strictly sharper performance guarantee. However, we also show that such a performance improvement is not significant. Moreover, when we impose a restrictive assumption of symmetry on the valuation distribution, the benefit of doing so can be little beyond knowing the value of the median.

3.2.4. Third Moment. In the base model, we obtain a performance guarantee with the knowledge of the first and second moments of the valuation distribution (see Theorem 2). Now, we investigate the benefit of knowing in addition the third moment of the valuation distribution. Recall that the seller knows the first moment $M_1 = \mu$ and the second moment $M_2 = \mathsf{E}(V^2) = \mu^2 + \sigma^2$ of the valuation distribution. Now, suppose the seller also knows the third moment $M_3 = \mathsf{E}(V^3)$.

Bertsimas and Popescu (2005, theorem 3.3) show that the following bound on the distribution involving the third moment is tight:

$$\mathsf{P}(V < (1-k)M_1) \leq 1 - \frac{(\delta^2+k)^3}{(\eta + (\delta^2+1)(\delta^2+k))(\eta + (\delta^2+k)^2)}, \qquad (1)$$

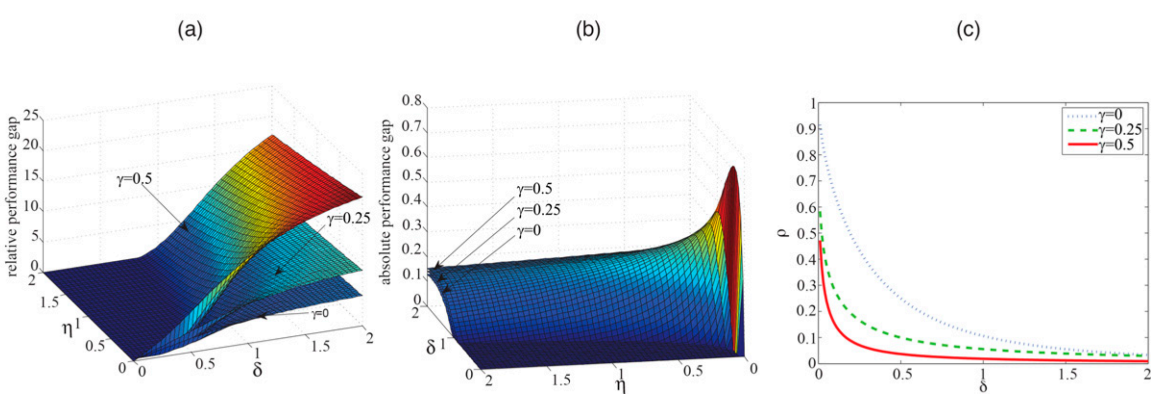
where $\eta = (M_1 M_3 - M_2^2)/M_1^4$. Select a price heuristic in the form of $p = (1-k)M_1$. Then, the profit can be written as

$$\begin{aligned}\pi(p;F) &= (p-c)\mathsf{P}(V \geq p) = ((1-k)M_1 - c)\mathsf{P}(V \geq (1-k)M_1)\\ &\geq (\mu - c)\left(1 - \frac{k}{1-\gamma}\right)\frac{(\delta^2+k)^3}{(\eta + (\delta^2+1)(\delta^2+k))(\eta + (\delta^2+k)^2)}\\ &\equiv (\mu - c)\phi(k),\end{aligned}$$

where the inequality is because of (1). With the additional third moment information, the seller could maximize $\phi(k)$ to obtain a price heuristic. As a result, the benefit of knowing M_3 can be measured by the relative performance gap $(\mu - c)\max_k \phi(k)/L - 1$, where L is the guaranteed profit with the information about the first and second moments only (see Theorem 1(b)). The relative performance improvement $(\mu - c)\max_k \phi(k)/L - 1$ depends on the system primitives through γ, δ, and M_3.

The maximizer of $\phi(k)$, as a root to a quartic equation, has a closed-form expression but is very unwieldy. For simplicity, we illustrate the benefit of knowing the third moment numerically. In particular, we fix $\gamma = 0, 0.25, 0.5$ and display the relative and absolute gap as a function of δ and η; see Figure 5(a) and 5(b). As a benchmark, Figure 5(c) shows the performance of the base model with only the first and second moment information. We make the following observations from the comparisons. First, the majority of the absolute gaps is no more than 20%. Second, the benefit of knowing M_3 is decreasing in η. That is, if η

Figure 5. (Color online) Comparison of the Third Moment Model and Base Model



Notes. (a) The relative gap. (b) The absolute gap. (c) Benchmark: base model.

Published in Manufacturing & Service Operations Management on January 20, 2022 as DOI: 10.1287/msom.2021.1055.
This article has not been copyedited or formatted. The final version may differ from this version.

is higher and hence, the skewness is larger, the benefit of knowing M_3 is lower. This is because within the class of distributions that have a larger skewness, the mass of the valuation distribution is more concentrated on the left, and as a result, the price heuristic $p^* = \mu - k^*\sigma$ without knowing M_3 tends to have already performed well. Third, when the valuation distribution is more variable, knowing the third moment provides more benefit by pinning down a subclass of distributions. Lastly, as γ increases, the benefit of knowing M_3 increases significantly. This is consistent with the understanding that the performance guarantee in Theorem 2 decreases in γ (see Corollary 1(a)). Hence, with a larger value of γ, there is more room for improvement by knowing M_3.

4. Multiproduct Robust Pricing

In this section, we generalize the single-product base model to consider the distribution-free pricing problem in selling multiple products. The seller knows only the mean and standard deviation for each product but does not know the customers' exact valuation distribution for each product. Suppose customer valuations for n products follow n distributions; for each product $i = 1, 2, \ldots, n$, its valuation distribution has the mean μ_i and standard deviation σ_i. We define the CV of product i as $\delta_i = \sigma_i/\mu_i$. The marginal cost for product i is c_i. As in Assumption (P), we assume $\mu_i \geq c_i$ for all i. We assume the valuation distributions across products are independent and consider the correlated valuation distributions in Section 4.1.2. Moreover, customer valuation for a set of products is additive (i.e., a customer's valuation for any subset of the products is the sum of its valuations for all products in the subset).

For any $S \subseteq [n] \equiv \{1, 2, \ldots, n\}$, denote by p_S, c_S, and V_S the price, cost, and customers' random valuation of the combination of products in subset S, respectively. Then, $c_S = \sum_{i\in S} c_i$ and $V_S = \sum_{i \in S} V_i$, where V_i is the (random) customer valuation for each product $i \in [n]$ with c.d.f. $F_i(\cdot)$. Let $\mathbf{p}^{\mathscr{M}} = (p_S, S \in \mathscr{M})$ be the price vector of (bundling) assortment $\mathscr{M}$ that is a subset of the power set of $[n]$ (i.e., $\mathscr{M} \subseteq \mathcal{P}([n])$). For example, if $\mathscr{M} = \{[n]\}$, the only product offering is a pure bundle. The firm sets $\mathbf{p}^{\mathscr{M}}$, and each customer chooses subsets of $\mathscr{M}$ such that any product is purchased at most once and its total surplus is maximized. For a given bundling scheme $\mathscr{M}$, the profit function is $\sum_{S\in\mathscr{M}} (p_S - c_S) q_S(\mathbf{p}^{\mathscr{M}})$, where $q_S(\mathbf{p}^{\mathscr{M}})$ is the market share for $S \in \mathscr{M}$ and a given price vector $\mathbf{p}^{\mathscr{M}}$. In a fashion similar to the single-product problem, for a given bundling scheme $\mathscr{M}$, we focus on the robust pricing problem under the maximin profit criterion:

$$\max_{\mathbf{p}^{\mathscr{M}}} \{ \min_{F_i \in \mathscr{F}_i, \forall i \in [n]} \sum_{S \in \mathscr{M}} (p_S - c_S) q_S(\mathbf{p}^{\mathscr{M}}) \},$$

where $\mathscr{F}_i \equiv \{F_i(v), v \geq 0 \mid \mathsf{E}[V_i] = \mu_i, \mathsf{E}[V_i^2] = \mu_i^2 + \sigma_i^2\}$. By Theorem 1(a), an upper bound of the profit for a single product i is $\mu_i - \tau_i^2 c_i/(1+\tau_i^2)$. Thus, an upper bound of the optimal profit for the multiproduct problem with full knowledge of valuation distributions is $\sum_{i=1}^n [\mu_i - \tau_i^2 c_i/(1+\tau_i^2)]$.[5] If we can derive a lower bound of the worst-case profit for a bundling policy $\mathscr{M}$, we can obtain its performance guarantee by taking the ratio of the lower and upper bounds. As the upper bound of the multiproduct problem is likely loose, the true performance guarantee can be much better than the displayed value. Again, as the derived worst-case profit lower bound of a bundling policy tends not to be tight, in this section we only deal with robust bundle pricing heuristics.

4.1. À La Carte vs. Pure Bundle

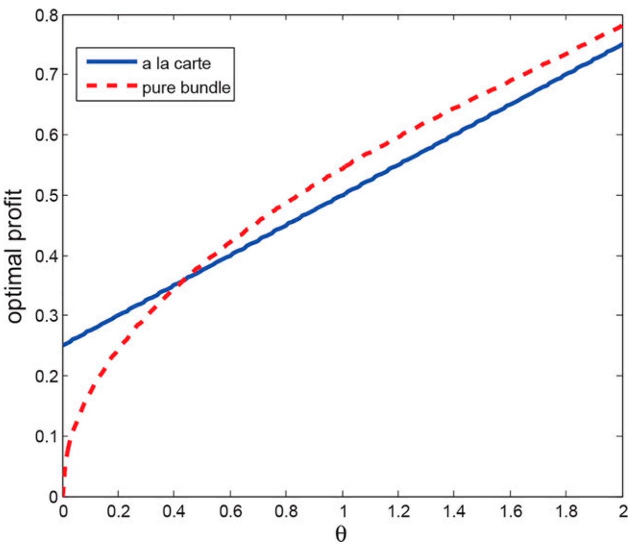
There are two typical ways to sell multiple products: à la carte (i.e., separate sales) with $\mathscr{M} = \mathcal{P}([n])$ and $p_S = \sum_{i \in S} p_i$ for any $S \in \mathscr{M}$ for which the seller decides on each individual product's price p_i, $i \in [n]$, or a pure bundle with $\mathscr{M} = \{[n]\}$ for which the seller decides on the bundle price p_b: that is, either to set a price for each individual product so customers can buy each one separately[6] or to set a price for the bundle of all products so customers buy either a pure bundle or nothing. The latter practice is becoming common as consumer buying habits trend toward simpler and more trouble-free experiences. For both selling mechanisms, we can adopt the performance guarantee result obtained for a single product to study the multiproduct problems. In particular, because the seller only knows the first and second moments of the valuation distribution for each product, it can apply our robust price for each and obtain a performance guarantee. The seller can also compute the first and second moments of the valuation distribution for the pure bundle of all products, assuming their valuations are independent or there is some known correlation among them, and obtain the performance guarantee for a robust bundle pricing heuristic. Then, we can compare the performance guarantees of selling separately and as a pure bundle. It is often unclear whether it is more profitable to sell separately or as a bundle for given known valuation distributions.

Example 1 (À La Carte vs. Bundle with Known Distributions). Consider two products. Consumers' valuations for the two products are known and independently distributed uniform random variables on $[0,1]$ and $[0,\theta]$, respectively, where $\theta > 0$. Assume the marginal cost is zero. Figure 6 shows the optimal profits of à la carte sales and pure bundle for various values of θ ranging from zero to two. Although the bundle tends to outperform à la carte sales, that is not always true. In other words, à la carte sales can perform better for specific valuation distributions.

From the robust pricing perspective, when the seller knows only the first and second moments μ_i and σ_i of

Published in Manufacturing & Service Operations Management on January 20, 2022 as DOI: 10.1287/msom.2021.1055.
This article has not been copyedited or formatted. The final version may differ from this version.

Figure 6. (Color online) À La Carte Sales vs. Pure Bundle



the valuation distributions V_i, we can make more general statements about when it is more profitable to sell a bundle. To be isolated from the supply-side effects, we make the following assumption throughout Section 4.1.

Assumption (S) (Equal Margin). $\frac{c_1}{\mu_1} = \frac{c_2}{\mu_2} = \cdots = \frac{c_n}{\mu_n} = \gamma$.

Assumption (S) allows products to have distinct valuation distributions on the demand side but requires them to have the same profit margin on the supply side. This may not be restrictive for some product lines in which profit margins tend not to vary much for different products.

4.1.1. Bundling as a Robust Device. For à la carte sales, the robust price heuristic for product i is $p^*(c_i, \mu_i, \sigma_i)$, where the mapping $p^*(c, \mu, \sigma)$ is specified in Theorem 2. By Theorem 1(b), the lower bound on each individual product's profit achieved by our robust price is tight. For pure bundling, the robust bundle price heuristic is $p^*\left(\sum_i c_i, \sum_i \mu_i, \sqrt{\sum_i \sigma_i^2}\right)$, and the lower bound on the bundle's profit may not be tight because this lower bound comes from treating the total valuation of all products as a single random variable and then applying Chebyshev's inequality. Nevertheless, if the profit lower bound of the bundle (which may not be achievable) is higher than that of separate sales (which is tight), the pure bundle generates higher profits than separate sales in a robust sense.

Proposition 4. *If* $\sqrt{\sum_i \sigma_i^2}/\sum_i \mu_i \le \min_i \delta_i$, *selling a bundle is guaranteed to generate higher profits than selling products separately.*

Proposition 4 says that to guarantee that the bundle is more robustly profitable, it is enough to require that the CV of the bundle, $\sqrt{\sum_i \sigma_i^2}/\sum_i \mu_i$, be less than the CVs of all products. In general, it could be computationally cumbersome to verify whether a bundle is more profitable than à la carte sales even for given distributions. Proposition 4 provides easily verifiable conditions under which the bundle is guaranteed to be more profitable when the detailed distribution information, other than the means and variances, is not available. We now present some special cases of this proposition.

Corollary 2. *For the following cases, selling a bundle is guaranteed to generate higher profits than selling products separately*: (a) $\sigma_1/\mu_1 = \sigma_2/\mu_2 = \cdots = \sigma_n/\mu_n = \delta$; (b) $\mu_1 = \mu_2 = \cdots = \mu_n = \mu$ *and* $\sqrt{\sum_i \sigma_i^2} \le n \min_i \sigma_i$ (in particular $\max_i \sigma_i / \min_i \sigma_i \le \sqrt{n}$); (c) $\sigma_1 = \sigma_2 = \cdots = \sigma_n = \sigma$ *and* $\sqrt{n} \max_i \mu_i \le \sum_i \mu_i$ *(in particular,* $\max_i \mu_i / \min_i \mu_i \le \sqrt{n}$).

Corollary 2(a) says that as long as valuation distributions for different products have the same CV (a first-order measure of variability), it is more robust to sell them in a bundle. This is not trivial as it appears. In Example 1, all uniform random variables on $[0, \theta]$ for $\theta > 0$ do share the same CV $= 1/\sqrt{3}$. However, given that their specific forms of distributions are uniform, it is not always more profitable to sell a bundle. Nevertheless, lacking the detailed information about the distributions, one would find it more robustly profitable to offer a bundle instead of à la carte sales. The benefit comes from a reduced CV $= \delta/\sqrt{n}$ in the valuation for the bundle. In part (b), one sufficient condition for $\sqrt{\sum_i \sigma_i^2} \le n \min_i \sigma_i$ is $\max_i \sigma_i / \min_i \sigma_i \le \sqrt{n}$. Similarly, in part (c), one sufficient condition for $\sqrt{n} \max_i \mu_i \le \sum_i \mu_i$ is $\max_i \mu_i / \min_i \mu_i \le \sqrt{n}$. Thus, parts (b) and (c) imply that if the valuation distributions have the same mean (standard deviation) and the ratio of the maximum and minimum of standard deviations (means) is bounded by the square root of the number of products, it is more robustly profitable to sell a bundle. Corollary 2 can be generalized as follows, without requiring characteristics of valuation distributions across products to be the same.

Corollary 3. *For the following cases, selling a bundle is guaranteed to generate higher profits than selling products separately*: (a) $\max_i \delta_i / \min_i \delta_i \le \sum_i \mu_i / \sqrt{\sum_i \mu_i^2}$; (b) $\max_i \mu_i / \min_i \mu_i \le n \min_i \sigma_i / \sqrt{\sum_i \sigma_i^2}$; (c) $\max_i \sigma_i / \min_i \sigma_i \le \sum_i \mu_i / (\sqrt{n} \max_i \mu_i)$.

Corollary 3 implies that the insights obtained from Corollary 2 for equal system primitives still hold, as long as those primitives do not vary much. Next, we study the benefit of a large bundle.

Proposition 5 (Large Bundle: Asymptotic Optimality). *Suppose all n products have the same mean μ, standard deviation σ, and marginal cost c but can have different distributions. For any $\epsilon \in (0,1]$, there exists a threshold n^** $\left(\text{if } \gamma = 0,\ n^* = (3/\epsilon)^2(3/\epsilon - 3)\delta^2/4\right)$ *such that for a bundle whose size is more than n^*, the robust bundle price $p_b^* = p^*(nc, n\mu, \sqrt{n}\sigma)$ achieves an ϵ-optimality.*[7] *Moreover, n^* is decreasing in ϵ.*

The standard argument for the benefits of bundling a large number of products, such as Bakos and Brynjolfsson (1999), resorts to the law of large numbers, which requires all random valuations for products to be independent and *identically* distributed. See Abdallah (2019) for a recent development along this line. In fact, without this restrictive limiting argument, it is difficult, if not impossible, to extend the result to heterogeneously distributed valuations (which may be correlated; see Proposition 6). From a robust perspective, Proposition 5 indeed confirms the benefit of a large bundle regardless of the detailed valuation distributions, as long as they have the same first- and second-order statistical characteristics.

With n products all sharing the same mean and standard deviation (hence, the same CV = δ but not the same distribution), when n is large enough, the standard deviation of the bundle $\sqrt{\sum_i \sigma_i^2}/\sum_i \mu_i = \delta/\sqrt{n}$ will be sufficiently small. In Figure 7, we fix $\delta = 1$, a relatively large variability in customers' valuation, and let the bundle size range from 100 to 850 and γ range from 0.5 to 0.8. We see that the bundle performance increases in the bundle size n and decreases in the profit margin measure γ. Moreover, by Proposition 5, when $\gamma = 0$, the minimum near-optimal bundle size is $n^* = (3/\epsilon)^2(3/\epsilon - 3)/(4\tau^2)$. Figure 8 displays n^* as a function of τ by fixing $\epsilon = 7\%, 8\%, 9\%, 10\%$. For any given τ, the minimum near-optimal bundle size decreases in ϵ. In addition, the marginal value of a larger bundle size decreases. When $\tau = 3$, the minimum near-optimal bundle size is 675 for 0.1-optimality. However, in order to achieve 0.07-optimality, the minimum near-optimal bundle size has to be more than 2,000.

Figure 7. (Color online) Benefit of Large-Bundle Pricing

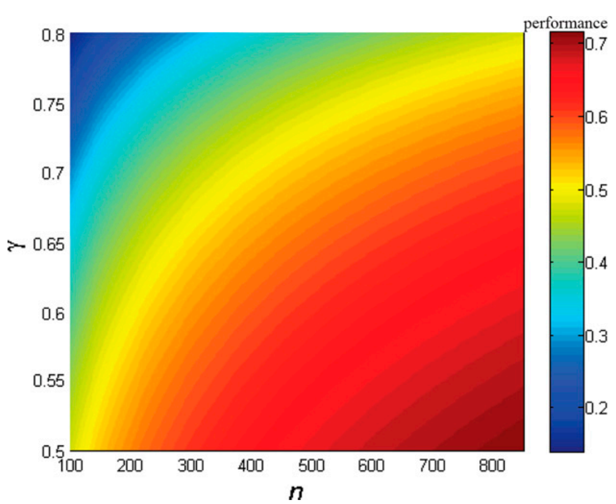
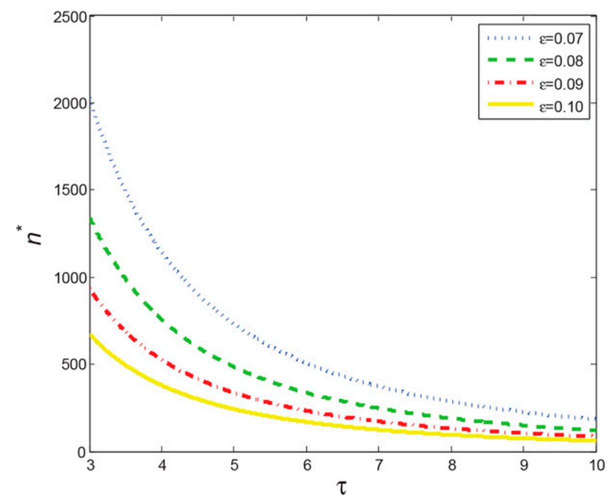


Figure 8. (Color online) Minimum Bundle Size for ϵ-Optimality



4.1.2. Correlated Values. So far, we assume that customer valuations for all products are independent. Our results continue to hold for nonpositive correlated values and under some conditions, even for positive correlated values.

Proposition 6 (Nonpositive Correlation). *If the correlation between valuations of any two products is nonpositive, Proposition 4, Corollaries 2 and 3, and Proposition 5 still hold.*

Recall from Corollary 1(a) that the performance guarantee decreases in the CV of the valuation distribution. Thus, nonpositive valuation correlations among products boost the performance of the bundle because they reduce the CV of the bundle valuation, whereas the performance of selling separately is unaffected by valuation correlations. By Corollary 1(b), the robust bundle price is higher under nonpositive valuation correlations than under independent valuations.

Proposition 7. *If $\sigma_1/\mu_1 = \sigma_2/\mu_2 = \cdots = \sigma_n/\mu_n = \delta$, selling a bundle is guaranteed to generate higher profits than selling products separately, regardless of valuation correlations across products.*

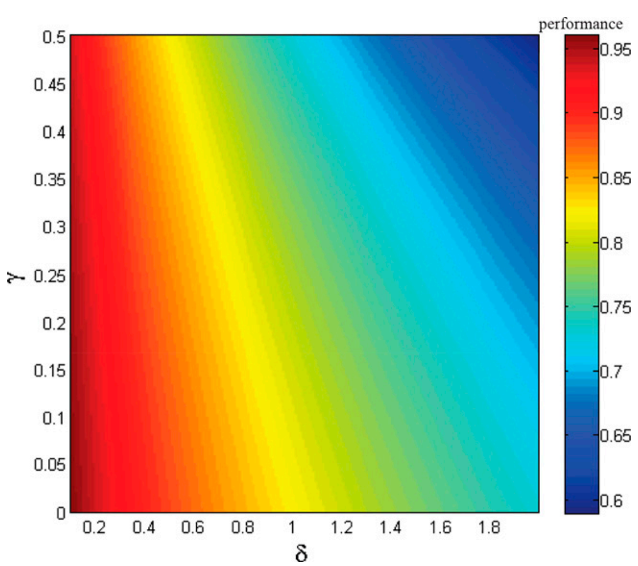
In the bundling literature, it is well known that if product valuations are known, independent and identically distributed, profits are higher under pure bundling than under separate sales. Moreover, when the known valuation distributions are correlated, profits are higher under pure bundling than under separate sales if and only if the correlation between the values for the products is negative or is sufficiently weak if positive (see, e.g., Belleflamme 2006). In other words, if the correlations are sufficiently positive, separate sales can be more profitable for given and known correlated valuation distributions. However, Proposition 7 confirms the benefit of bundling in a robust sense; as long as the cost structure in terms of profit margins and the demand characteristics in terms of CVs are the same across all products, the bundle is more robustly profitable than separate sales, regardless of the detailed valuation distributions and whether they are negatively or positively correlated. This is because, to protect against the worst case of valuation distributions, the robust bundle price heuristic and its performance guarantee depend only on the means, variances, and covariances of the valuation distributions, which determine the CV of the bundle valuation, and moreover, because this bundle valuation's CV under correlations is weakly smaller than those of the individual products. If the correlation structure turns out to be the worst for bundling, in which case all products are positively correlated, bundling is as effective as separate sales, which is consistent with Carroll (2017). However, bundling is robustly more profitable than separate sales for correlation structures other than the worst case, which restores the robust profitability of bundling over separate sales in a more general sense.

4.1.3. Practical Examples. In the following discussion, we will illustrate the benefit of our robust bundle price with three practical examples, each with its own distinctive features. In the first example, the bundle contains only digital goods that can be considered as having zero marginal costs, and the bundle is large. Both features contribute to the very good performance of our robust bundle price. In the second example, the bundle contains a mix of goods with zero and nonzero marginal costs, and the bundle is medium-sized. In the last example, the bundle is small, and customer valuations are notably correlated.

4.1.3.1. Spotify. Spotify is a music streaming service provider. Its premium-tier service gives users unlimited music streaming at a cost of \$9.99 per month. Spotify pays royalties for all of the listening that takes place through its service by distributing nearly 70% of its revenues to the copyright holders. A single play on Spotify Premium was worth on average about 0.68 cents in royalties (Smith 2015). The subscription service of Spotify is in fact a bundle of zero-marginal-cost digital goods (i.e., $c \approx 0$). We can estimate the average number of songs a subscriber listens to per month: $\frac{9.99\times 70\%}{0.0068} \approx 1028$. In general, it would be very hard to price such bundles, which allow customers to choose the songs they want. The subscribers on average choose and listen to about 1,000 songs a month, according to their own heterogeneous preferences. Fortunately, our distribution-free performance guarantee for a bundle only requires estimates of three primitives (i.e., the valuation heterogeneity in a song δ, the profit margin measure γ, and the bundle size n). We assume all songs have roughly the same δ, which may be close to reality because customers choose to play their favorite songs. In Figure 9, we illustrate the performance guarantee of the robust bundle price by having δ range from 0.1 to 2, γ range from 0 to 0.5, and n equal 1,000. We see that for Spotify, which has $\gamma \approx 0$, a bundle of size 1,000 would guarantee a performance of about 83% for valuation $\delta = 1$, which is typically considered as "random," and about 75% for valuation $\delta = 2$, which is typically considered to be far more than "random." Moreover, even for the extreme case where $\gamma = 0.5$, which typically corresponds to a profit margin for physical goods and $\delta = 2$, the performance guarantee exceeds 60%.

4.1.3.2. IT Service Provider. In Xue et al. (2015), the authors study the pricing problem for personalized bundles of goods and services. They test their model by using empirical data from a leading IT service provider. We adopt the parameters from this paper and use them as inputs to test our robust bundle price. A typical bundle from this IT company consists of 100 components, which are a mix of software and hardware. The marginal cost of noncustomized commercial software can be considered as zero (i.e., $\gamma_1 = 0$ for software). Following

Figure 9. (Color online) Performance for Spotify



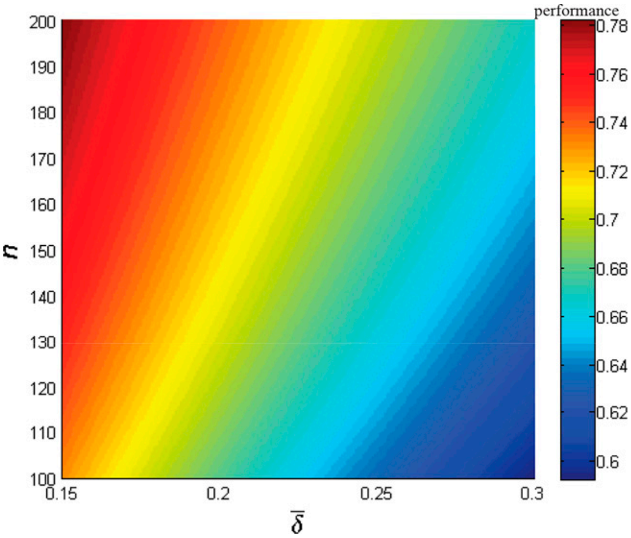
a working paper version of Xue et al. (2015), we choose $\gamma_2 = 55.7\%$ for hardware. In general, 70% of the components of a personalized bundle are software. The average CV of customer valuation for a component varies from 15% to 30%. Because the authors do not differentiate the CVs between software and hardware, we can consider the CVs of valuations for software and hardware to be the same, but they have different means and standard deviations. That is, if we denote means and standard deviations of software and hardware by μ_1, μ_2, σ_1, and σ_2, respectively, $\sigma_1/\mu_1 = \sigma_2/\mu_2 = \bar{\delta}$. Then, assuming independence across valuations for different products, the CV of valuations for the whole bundle is at

$$\delta = \frac{\sqrt{0.7\sigma_1^2 + 0.3\sigma_2^2}}{0.7\mu_1 + 0.3\mu_2} = \frac{\bar{\delta}\sqrt{0.7\mu_1^2 + 0.3\mu_2^2}}{0.7\mu_1 + 0.3\mu_2}$$

$$= \frac{\bar{\delta}\sqrt{0.7\left(\frac{\mu_1}{\mu_2}\right)^2 + 0.3}}{0.7\frac{\mu_1}{\mu_2} + 0.3} \le \frac{\bar{\delta}}{\sqrt{0.3}}.$$

By Corollary 1(b), the performance guarantee is decreasing in γ and δ. Hence, to be conservative, we set the CV for the whole bundle $\delta = \bar{\delta}/\sqrt{0.3}$, where $\bar{\delta} \in [15\%, 30\%]$. Moreover, because the bundle is a mix of software with $\gamma_1 = 0$ and hardware with $\gamma_2 = 55.7\%$, the ratio of the average cost and mean for the bundle, γ, should be between 0% and 55.7%. To be conservative, we choose $\gamma = 55.7\%$. With the primitives estimated as above, Figure 10 shows the results of applying our robust bundle price to a personalized bundle that may have about 100–200 components. We observe that in all the instances tested with those very conservative inputs, our robust bundle price has a performance guarantee of more than 60%.

Figure 10. (Color online) Performance for the IT Service Provider



4.1.3.3. TheatreWorks. TheatreWorks is a theater company based in Palo Alto, California. It offers a full eight-play season subscription package. Chu et al. (2011) use the data of eight different plays or musicals from TheatreWorks to estimate consumer valuations. Unlike the previous two examples, consumer valuations of eight plays or musicals are correlated. The authors assume that the joint distribution of consumer valuations is an eight-dimensional bimodal normal distribution, with the estimated values listed in Chu et al. (2011, table 8A). Rather than assuming a specific joint distribution, we apply our robust bundle price with $\mu = 164.7568$ and $\sigma = 210.5679$, imputed from the data. As a result, $\delta = 1.2780$. Our heuristic guarantees a profit of \$33.47. With Chu et al. (2011) obtaining the optimal profit of \$63.67 under the knowledge of the true valuation distribution, our bundle price guarantees a 0.5257-optimality without the knowledge of any detailed distribution information.

4.2. Heuristic Mixed Bundling: A Single Product or Pure Bundle

The mixed bundling problem with all the combinations of subsets of products priced and available for purchase is generally challenging. We consider a heuristic bundling scheme that allows customers to choose from buying only one of the products or a pure bundle (i.e., the purchase option for any customer is chosen from the set $\mathscr{M} = \{\{1\}, \{2\}, \ldots, \{n\}, [n]\}$). Thus, the seller has $n + 1$ number of prices to set. When $n = 2$, such a bundling scheme boils down to mixed bundling.

We follow the same idea as the single-product robust pricing problem to derive a performance guarantee for this heuristic bundling scheme. We first derive a lower bound by using the the inequality of Olkin and Pratt (i.e., the multivariate case of Chebyshev's inequality; see Marshall and Olkin 1960)[8]. Then, we derive a distribution-free heuristic pricing policy by maximizing the lower bound. With the profit upper bound $\sum_{i=1}^{n}[\mu_i - \tau_i^2 c_i/(1+\tau_i^2)]$, we can obtain a performance guarantee for the heuristic policy.

Proposition 8. *Let $i^* = \arg\max_i\{L_i\}$, where L_i is the optimal value of Problem (2):*

$$\begin{aligned}\max_{k_1>0,\ldots,k_n>0,k_b>0} \quad & (\mu_i - c_i - k_i\sigma_i)\left(1 - \prod_{i=1}^{n}\frac{1}{1+k_i^2}\right) \\ & + \left(\sum_{j\neq i}(\mu_j - c_j) - \bar{k}_i\sqrt{\sum_{j\neq i}\sigma_j^2}\right)T, \\ \text{s.t.} \quad & \mu_i - c_i - k_i\sigma_i \le \mu_j - c_j - k_j\sigma_j, \quad \forall j \neq i, \\ & k_i\sigma_i + \bar{k}_i\sqrt{\sum_{j\neq i}\sigma_j^2} = k_b\sqrt{\sum_{i=1}^{n}\sigma_i^2}, \quad \bar{k}_i > 0,\end{aligned} \tag{2}$$

with

$$T = 1 - \frac{1}{(n+1)^2}\left(\sqrt{u} + \sqrt{n}\sqrt{(n+1)\left(\sum_{i=1}^n \frac{1}{\bar{k}_i^2} + \frac{1}{k_b^2}\right) - u}\right)^2,$$

$$u = \sum_{i=1}^n \frac{1}{\bar{k}_i^2} + \frac{1}{k_b^2} + 2\left(\sum_{i=1}^n \sum_{j<i} \frac{\rho_{ij}}{\bar{k}_i \bar{k}_j} + \sum_{i=1}^n \frac{\rho_{n+1,i}}{k_b \bar{k}_i}\right),$$

$\rho_{n+1,i} = \sqrt{\sum_{j\neq i}\sigma_j^2}/\sqrt{\sum_{i=1}^n \sigma_i^2}$, *and* $\rho_{ij} = \sum_{k\neq i,j}\sigma_k^2/\left(\sqrt{\sum_{k\neq i}\sigma_k^2}\sqrt{\sum_{k\neq j}\sigma_k^2}\right)$. *The heuristic bundling scheme, given by* $p_i^* = \mu_i - k_i^*\sigma_i$, $\forall i$ *and* $p_b^* = \sum_{i=1}^n \mu_i - k_b^*\sqrt{\sum_{i=1}^n \sigma_i^2}$, *where* k_i^*, $\forall i$, *and* k_b^* *are the optimal solutions to Problem* (2) *for* $i = i^*$, *achieves a distribution-free lower bound* L_{i^*}.

When all products have the same parameters, they are "symmetric" in a robust sense (i.e., the robust pricing heuristic derived in Proposition 8 degenerates to a pricing policy that charges the same price for any single product), which simplifies the problem. We summarize this observation as follows.

Corollary 4. *Suppose* $\mu_1 = \mu_2 = \cdots = \mu_n$, $\sigma_1 = \sigma_2 = \cdots = \sigma_n$, *and* $c_1 = c_2 = \cdots = c_n$. *Derive a pricing policy for the heuristic bundling scheme as* $p_i^* = \mu_i - k_i^*\sigma_i$ *and* $p_b^* = \sum_{i=1}^n \mu_i - k_b^*\sqrt{\sum_{i=1}^n \sigma_i^2}$, *where* k_i^* *and* k_b^* *are the optimal solutions to the following optimization problem*:

$$\max_{\sqrt{n}k_b > k_i > 0} \; (\mu_i - c_i - k_i\sigma_i)\left(1 - \frac{1}{(1+k_i^2)^n}\right) + ((n-1)(\mu_i - c_i) - (\sqrt{n}k_b - k_i)\sigma_i)T, \tag{3}$$

where

$$T = 1 - \frac{1}{(n+1)^2}\left(\frac{\sqrt{n(n-1)}}{\bar{k}_i} + \frac{1}{k_b} + \sqrt{n}\sqrt{\frac{2n}{\bar{k}_i^2} + \frac{n}{k_b^2} - \frac{2\sqrt{n(n-1)}}{k_b \bar{k}_i}}\right)^2$$

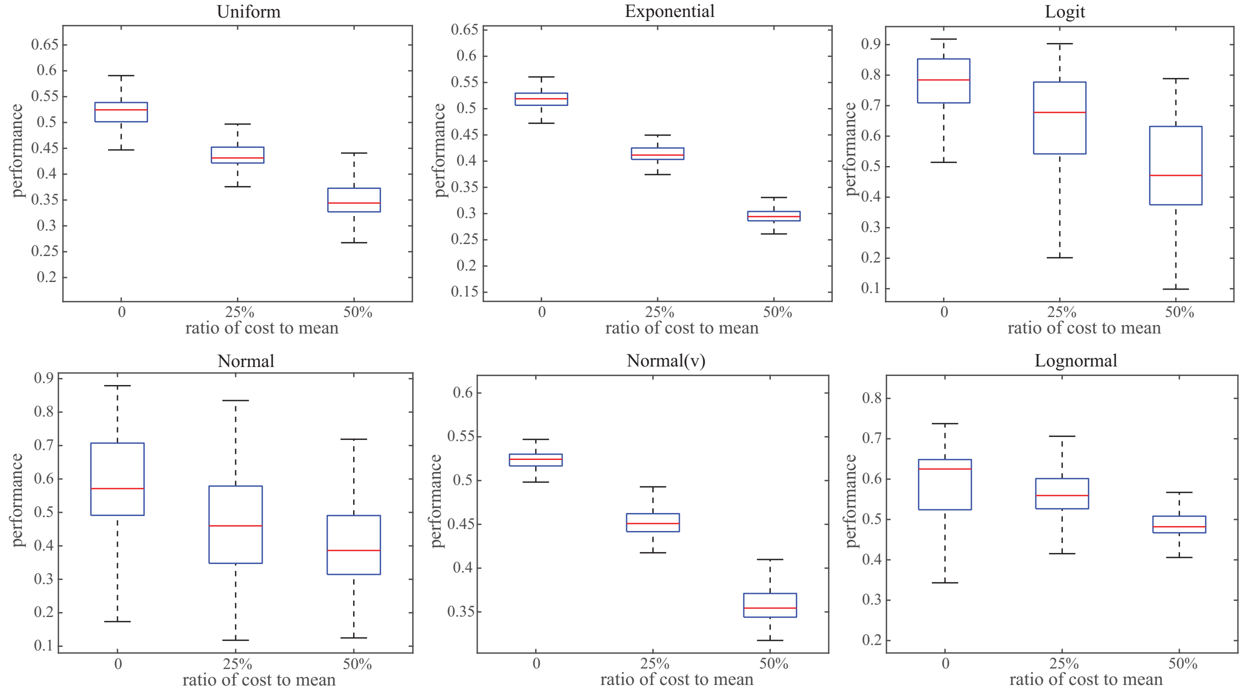
and $k_i + \sqrt{n-1}\bar{k}_i = \sqrt{n}k_b$. *The optimal value of* (3) *is a distribution-free lower bound for the heuristic bundling scheme.*

To show the performance of the heuristic in Proposition 8, we assume that $n = 2$ (for which our heuristic bundling scheme is mixed bundling) and that both products have the same parameters, and we use the taste distributions from Chu et al. (2011, table 2). Figure 11 displays the box plots of various percentiles of the heuristic's performances benchmarked against the optimal profit under full information. As expected, when the marginal cost increases, the performance decreases. Even though the performance may not be stellar, at least Proposition 8 provides a way to compute a heuristic that performs reasonably and has a provable performance guarantee.

4.3. Clustered Bundling

In this subsection, we partition n products into k disjoint parts (i.e., clusters; $[n] = \cup_l S_l$) and charge a bundle price for each cluster S_l. We refer to this as the scheme of *clustered bundling* with $\mathscr{M} = \mathcal{P}(\{S_l, \forall l\})$ as a special case of mixed bundling, which imposes a price for each part (i.e., cluster) of a partition of all products and allows customers to choose one or multiple parts

Figure 11. (Color online) Performance of Heuristic Mixed Bundling Pricing



(i.e., clusters). The idea behind this is that within a cluster that we could leverage Computer Science's clustering algorithms to identify, products tend to be similar in nature, and our robust bundle heuristic for that cluster of products should perform well.

4.3.1. Mean-Ranked Clustered Bundling. We first propose a naive clustered bundling heuristic, which places the products with similar means into the same cluster. Without loss of generality, let $\mu_1 \le \mu_2 \le \cdots \le \mu_n$. We propose the following heuristic: all products are grouped into k clusters by order (i.e., any product in the ith cluster has a smaller index than that of any product in the $(i+1)$th cluster). We call this heuristic *mean-ranked clustered bundling*. To find the optimal mean-ranked clustered bundling, we develop Algorithm OA.1 in the online appendix. For a sufficiently large number of products, searching for the optimal mean-ranked clustered bundling may require significant computational efforts, as we show in Proposition OA.5 in the online appendix that the computational complexity of Algorithm OA.1 is exponential. Then, we develop two polynomial time algorithms to generate a heuristic mean-ranked clustered bundling: fixed radius clustering (FRC) and top-down/bottom-up clustering (TBC) algorithms. We show that both heuristics are more profitable than separate sales and the pure bundle, or reduce to one of them (see the online appendix for more details).

4.3.2. General Clustered Bundling. We next treat (c_i, μ_i, σ_i) as the features of product i. It is natural to place those products with a similar feature vector into the same cluster. Thus, we propose a practical heuristic of general clustered bundling (GCB). First, for any given $k = 1, 2, \ldots, n$, we divide all products into k clusters by a commonly used clustering algorithm (e.g., k-means clustering); then, we calculate the lower bound on the profit of each bundle in the form of Theorem 1(b), and lastly, we find the optimal k^* such that the total profit lower bound is maximized. It is easy to see this GCB heuristic reduces to separate sales when $k^* = n$ and the pure bundle when $k^* = 1$. Thus, this GCB heuristic has a weakly better performance guarantee than separate sales and the pure bundle. The pseudocode for the algorithm is given in Algorithm 1.

Algorithm 1 (GCB)
1. **Input**: $x_i = (c_i, \mu_i, \sigma_i)$, where $i = 1, 2, \ldots, n$
2. **for** $k = 1, 2, \ldots, n$ **do**
3. Using a clustering algorithm to divide $\{1, 2, \ldots, n\}$ into k clusters
4. $L_m^k \leftarrow$ the lower bound on the profit of the mth cluster, $m = 1, 2, \ldots, k$
5. $L_k \leftarrow \sum_m L_m^k$
6. **end for**
7. Let $k^* = \arg\max_k L_k$
8. **Output**: The total profit lower bound L_{k^*} and the clustered bundling heuristic corresponding to the k^* clusters

4.3.3. Comparison of Various Policies. To compare the performance guarantees of separate sales, the pure bundle, FRC, TBC, and GCB, we fix the ratio of cost to mean as $\gamma = 0$ or $\gamma = 0.25$ for each product. For $n = 10$, 30, or 50, we generate 1,000 scenarios,

Figure 12. (Color online) Performance Guarantees for Five Policies

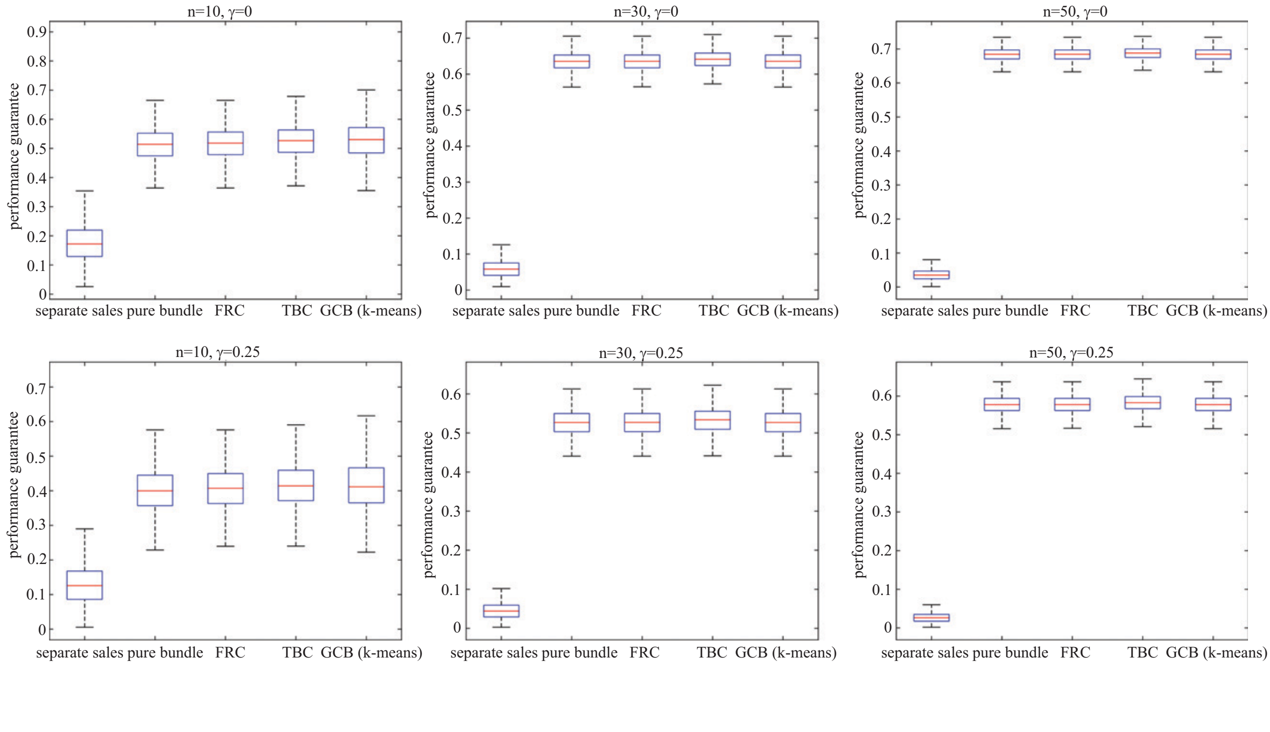
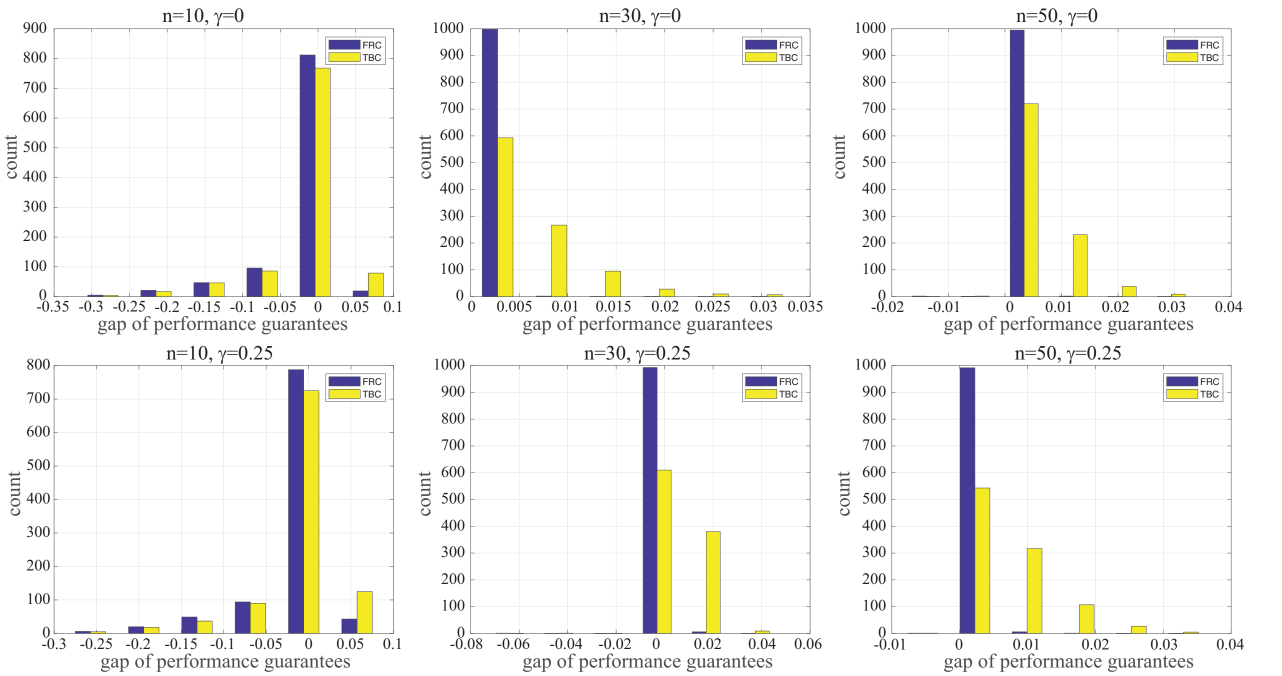


Figure 13. (Color online) Gap of Performance Guarantees Between FRC and GCB and Between TBC and GCB



and each scenario contains n products with μ_i independently generated from the uniform distribution $U[0, 10]$ and σ_i independently generated from the uniform distribution $U[0.1, 5]$. Figure 12 displays the box plots of the performance guarantees (i.e., the sum of the profit lower bounds of each cluster divided by the sum of the profit upper bounds of each cluster) across five policies, where step 3 of Algorithm 1 adopts the celebrated k-means clustering. It is observed from Figure 12 that the gap in the performance guarantees between separate sales and the other four bundling policies becomes larger when n increases. Additionally, the gap in the performance guarantees between the pure bundle and our three heuristics becomes smaller when n increases, which implies the pure bundle performs well for large values of n. Furthermore, Figure 13 plots the empirical distributions of the gap of performance guarantees between FRC and GCB and between TBC and GCB across 1,000 scenarios. On one hand, when n is relatively small, GCB may be better or worse than FRC and TBC. If GCB does better, the gap can be as large as 0.3; however, the gap is no more than 0.1 if FRC or TBC is better. On the other hand, TBC dominates the other two heuristics when n is relatively large.

5. Alternative Robust Criteria

Previously, we have focused on the maximin profit criterion. In this section, we consider two alternative performance criteria and then compare the optimal robust prices under all three criteria. Given a price p, if the true customer valuation distribution is F, then the relative regret because of not knowing the true distribution can be measured by $1 - \pi(p; F)/\max_z \pi(z; F)$. Denote by $\bar{p}$ the optimal price that minimizes the worst relative regret, referred to as the *minimax relative regret criterion*: that is,

$$\bar{p} = \arg\min_p \{\max_{F\in\mathscr{F}} 1 - \pi(p; F)/\max_z \pi(z; F)\}.$$

The regret in its absolute term because of not knowing the true distribution can be measured by $\max_z \pi(z; F) - \pi(p; F)$. Denote by $\tilde{p}$ the optimal price that minimizes the worst absolute regret, referred to as the *minimax absolute regret criterion*: that is, $\tilde{p} = \arg\min_p \{\max_{F\in\mathscr{F}} [\max_z \pi(z; F) - \pi(p; F)]\}$.

Under the minimax relative regret criterion, the optimal robust price and the corresponding worst relative regret can be characterized as follows.

Theorem 3 (Minimax Relative Regret Criterion). *The optimal price that minimizes the worst relative regret is $\bar{p} = \mu - \bar{k}\sigma$, where*

$$\bar{k} = \sqrt[3]{\tau/2 + \sqrt{(\tau/2)^2 + (2/3)^3}} + \sqrt[3]{\tau/2 - \sqrt{(\tau/2)^2 + (2/3)^3}} \geq 0,$$

and the corresponding worst relative regret is $1/(1 + \bar{k}^2)$. Moreover, the worst relative regret is asymptotically achievable by a series of two-point distributions:

$$V = \begin{cases} \mu - k_\eta \sigma & \text{with probability } \dfrac{1}{1 + k_\eta^2}, \\ \mu + \dfrac{1}{k_\eta}\sigma & \text{with probability } 1 - \dfrac{1}{1 + k_\eta^2}, \end{cases}$$

with $k_\eta = \bar{k} + \eta$, as $\eta \searrow 0$.

Theorem 3 shows that the optimal robust price $\overline{p}$ under the minimax relative regret criterion has a closed form similar to the optimal price p^* under the maximin profit criterion. Moreover, the worst relative regret for $\overline{p}$ can also be asymptotically achievable by a series of two-point distributions. Next, we compare the optimal prices under the three robust criteria.

Theorem 4 (Three Robust Criteria). *The optimal robust prices under the three criteria are ordered as $p^* < \overline{p} < \tilde{p}$, where the characterization of $\tilde{p}$ is given in the proof.*

The optimal robust price p^* achieves the highest profit for the seller against the worst possible distribution, $\overline{p}$ minimizes the maximum relative regret (i.e., one minus the relative profit gap because of not knowing the true distribution), and $\tilde{p}$ minimizes the maximum absolute regret (i.e., the absolute profit gap because of not knowing the true distribution). Theorem 4 shows that among the three criteria, the maximin profit criterion leads to the most aggressive (i.e., lowest) price. This is because the optimal prices under the two alternative criteria aim at a profit gap and are tolerant of less aggressive behavior because the optimal profits $\max_z \pi(z; F)$ under the knowledge of the true extreme distributions are also low. In contrast, as a survival measure aimed at the absolute profit value, the optimal price under the maximin profit criterion has to be more aggressive. The comparison demonstrates the importance of identifying an appropriate objective in robust price optimization, as different criteria result in distinct optimal robust prices even under the same information structure. Moreover, with the maximin profit criterion resulting in a lower optimal robust price, it leads to a higher market penetration level and higher social welfare than the two alternative criteria.[9] Higher market penetration can be preferred by a for-profit firm, and higher social welfare can be appealing to a socially responsible seller.

The alternative optimal robust prices $\overline{p}$ and $\tilde{p}$ may not perform well for the absolute profit measure. For example, consider the two-point distribution F^o defined in Theorem 3 with $k_\eta = \overline{k} + \eta$, where η is sufficiently small. Figure 14(a) displays the relative profit gap between p^* and $\overline{p}$, $[\pi(p^*; F^o) - \pi(\overline{p}; F^o)]/\pi(\overline{p}; F^o)$, under the two-point distribution F^o, and shows that the relative gap is always more than one when τ ranges from 0.5 to 2. Moreover, this relative gap is decreasing in τ. Figure 14(b) illustrates the relative profit gap between p^* and $\tilde{p}$ under F^o, $[\pi(p^*; F^o) - \pi(\tilde{p}; F^o)]/\pi(\tilde{p}; F^o)$, and shows that this relative gap is increasing in δ and γ, always more than zero, and more than one for most cases.

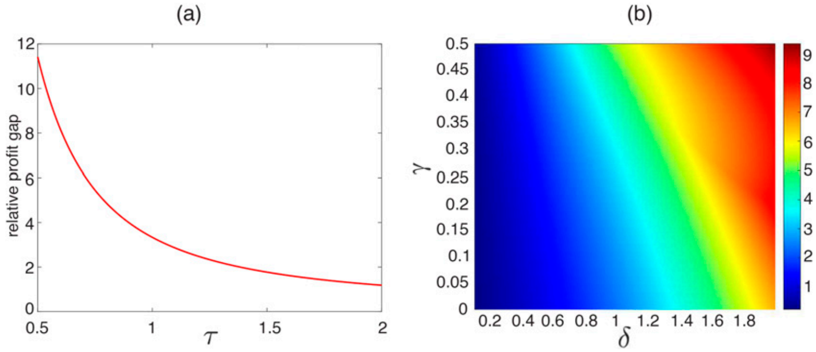
In a fashion similar to the multiproduct robust pricing problem under the maximin profit criterion, we can define the multiproduct robust pricing problem under the minimax relative regret criterion as

$$\min_{\overline{\mathbf{p}}^{\mathscr{M}}} \left\{ \max_{F_i \in \mathscr{F}_i,\, \forall i \in [n]} 1 - \frac{\sum_{S \in \mathscr{M}} (\overline{p}_S - c_S) q_S(\overline{\mathbf{p}}^{\mathscr{M}})}{\max_{p_S} \sum_{S \in \mathscr{M}} (p_S - c_S) q_S(\mathbf{p}^{\mathscr{M}})} \right\}. \quad (4)$$

In view of the similarity between the closed forms of the optimal robust price under the maximin profit and minimax relative regret criteria, we will compare separate sales and the pure bundle under the minimax relative regret criterion. For separate sales, $\mathscr{M} = \mathcal{P}([n])$ and $\overline{p}_S = \sum_{i \in S} \overline{p}_i$ for any $S \in \mathscr{M}$, and the seller decides on each individual product's price $\overline{p}_i$, $i \in [n]$. For the pure bundle, $\mathscr{M} = \{[n]\}$, and the seller decides on the bundle price $\overline{p}_b$.

Under the minimax relative regret criterion, we say pure bundling is guaranteed to perform better than separate sales if the optimal objective value of (4) under pure bundling is smaller than that under separate sales. All the results in the comparison of pure bundling and separate sales under the maximin profit

Figure 14. (Color online) Performance of Alternative Robust Prices for the Absolute Profit Measure



Notes. (a) The performance of $\overline{p}$ that is optimal under the minimax relative regret criterion. (b) The performance of $\tilde{p}$ that is optimal under the minimax absolute regret criterion.

criterion in Section 4.1 carry over to the minimax relative regret criterion, which we summarize as follows.

Proposition 9. (Bundling as a Robust Device). *Consider the minimax relative regret criterion. If any condition in Propositions 4, 6, and 7 or Corollaries 2 and 3 is satisfied, selling a bundle is guaranteed to perform better than selling products separately. Moreover, the asymptotic optimality of a large bundle in Proposition 5 holds.*

As results in Section 4.2 rely on the multivariate case of Chebyshev's inequality, they cannot be extended to the minimax relative regret criterion. Nevertheless, the idea of clustered bundling and the proposed heuristics in Section 4.3 are sufficiently generic to account for the minimax relative regret criterion.

6. Concluding Remarks

In this paper, we provide a robust price for the distribution-free pricing problem of a single product under two robust criteria. The robust prices (see Theorems 2 and 3) are in closed form and hence, are efficiently computable. Their interpretation can be easily explained to pricing managers and has a natural connection with the newsvendor ordering decision that is widely taught in the Master of Business Administration classrooms. Both robust prices also have a provable performance guarantee and are indeed the *optimal* robust price for a single product under the associated robust criterion. We extend the single-product result to the multiple-product setting by studying separate sales, the pure bundle, and other bundling heuristics, such as clustered bundling. Given their simplicity and practical values, we expect the results to provide a tool in situations where prices need to be determined frequently for different one-time customers of whom the seller has little knowledge.

Acknowledgments

The authors thank the department editor, Morris Cohen, the anonymous associate editor, and two anonymous reviewers for their guidance and constructive comments.

Endnotes

[1] If the random customer valuation V is known, the demand function may be obtained as $d(p) = M \cdot \mathsf{P}(V \geq p)$, where M is the potential market size.

[2] Unfortunately, the optimal robust price expression in Carrasco et al. (2018, p. 257), which corresponds to the problem with $c = 0$, is incorrect.

[3] Sometimes, the seller may only be sure about the range of the mean and standard deviation of the valuation distribution. By Corollary 1(a), it is conservative to let μ be the lower bound of the mean and σ be the upper bound of the standard deviation.

[4] When $c = 0$, if the valuation distribution is realized as some common distributions, such as truncated normal, truncated logit, or log-normal, for the realized distributions, the optimal price is indeed less than the mean valuation μ.

[5] Abdallah et al. (2021) show that a profit upper bound for the multi-product pricing problem is the expected profit under perfect price discrimination (i.e., $\sum_{i=1}^{n} \mathsf{E}[(V_i - c_i)^+]$, where $\mathsf{E}[(V_i - c_i)^+]$ is the expected profit under perfect price discrimination for single product i).

[6] Given that the price for each product is set separately, there is no discount for buying a set of products. Hence, the optimal strategy for a customer is to buy every product that generates a nonnegative surplus.

[7] That is, the ratio of the guaranteed profit achieved by our robust bundle price p_b^* to the optimal profit achieved by the optimal mixed bundling prices with $\mathscr{M} = \mathcal{P}([n])$, for any given valuation distributions of n products, is more than $1 - \epsilon$.

[8] Rujeerapaiboon et al. (2018) provide multivariate Chebyshev's inequalities for products of symmetric nonnegative random variables.

[9] For any given random valuation V, the market penetration level $\mathsf{P}(V \geq p)$ and the social welfare $\mathsf{E}[V - c \mid V \geq p]\mathsf{P}(V \geq p)$ are decreasing in p.

References

Abdallah T (2019) On the benefit (or cost) of large-scale bundling. *Production Oper. Management* 28(4):955–969.

Abdallah T, Asadpour A, Reed J (2021) Large-scale bundle size pricing: A theoretical analysis. *Oper. Res.* 69(4):1158–1185.

Adams WJ, Yellen JL (1976) Commodity bundling and the burden of monopoly. *Quart. J. Econom.* 90(3):475–498.

Allouah A, Bahamou A, Besbes O (2021) Pricing with samples. Preprint, submitted February 25, http://dx.doi.org/10.2139/ssrn.3334650.

Azar PD, Micali S (2013) Parametric digital auctions. Kleinberg R, ed. *Proc. 4th Conf. Innovations Theoretical Comput. Sci.* (Association for Computing Machinery, Berkeley, CA), 231–232.

Babaioff M, Immorlica N, Lucier B, Weinberg SM (2014) A simple and approximately optimal mechanism for an additive buyer. Barak B, ed. *Proc. 55th Annual IEEE Sympos. Foundations Comput. Sci.* (Institute of Electrical and Electronics Engineers, Philadelphia), 21–30.

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

Belleflamme P (2006) Versioning information goods. Illing G, Peitz M, eds. *Industrial Organization and the Digital Economy* (MIT Press, Cambridge, MA), 175–206.

Bergemann D, Schlag K (2008) Pricing without priors. *J. Eur. Econom. Assoc.* 6(2-3):560–569.

Bergemann D, Schlag K (2011) Robust monopoly pricing. *J. Econom. Theory* 146(6):2527–2543.

Bertsimas D, Popescu I (2005) Optimal inequalities in probability theory: A convex optimization approach. *SIAM J. Optim.* 15(3): 780–804.

Carrasco V, Luz VF, Kos N, Messner M, Monteiro PK, Moreira H (2018) Optimal selling mechanisms under moment conditions. *J. Econom. Theory* 177(2018):245–279.

Carroll G (2017) Robustness and separation in multidimensional screening. *Econometrica* 85(2):453–488.

Chakraborty I (1999) Bundling decisions for selling multiple objects. *Econom. Theory* 13(3):723–733.

Chu CS, Leslie P, Sorensen A (2011) Bundle-size pricing as an approximation to mixed bundling. *Amer. Econom. Rev.* 101(1): 263–303.

Cohen MC, Elmachtoub AN, Lei X (2022) Price discrimination with fairness constraints. *Management Sci.* Forthcoming.

Cohen MC, Perakis G, Pindyck RS (2021) A simple rule for pricing with limited knowledge of demand. *Management Sci.* 67(3): 1608–1621.

Elmachtoub AN, Gupta V, Hamilton ML (2021) The value of personalized pricing. *Management Sci.* 67(10): 6055–6070.

Fang H, Norman P (2006) To bundle or not to bundle. *RAND J. Econom.* 37(4):946–963.

Gallego G, Moon I (1993) The distribution free newsboy problem: Review and extensions. *J. Oper. Res. Soc.* 44(8):825–834.

Koçyiğit Ç, Rujeerapaiboon N, Kuhn D (2021) Robust multidimensional pricing: Separation without regret. *Math. Programming*, ePub ahead of print January 19, https://doi.org/10.1007/s10107-021-01615-4.

Kos N, Messner M (2015) Selling to the mean. Preprint, submitted June 19, http://dx.doi.org/10.2139/ssrn.2632014.

Li X, Sun H, Teo C-P (2022) Convex optimization for bundle size pricing problem. *Management Sci*. Forthcoming.

Ma W, Simchi-Levi D (2015) Reaping the benefits of bundling under high production costs. Preprint, submitted December 8, http://dx.doi.org/10.2139/ssrn.2700394.

Marshall A, Olkin I (1960) Multivariate Chebyshev inequalities. *Ann. Math. Statist.* 31(4):1001–1014.

Perakis G, Roels G (2008) Regret in the newsvendor model with partial information. *Oper. Res.* 56(1):188–203.

Rujeerapaiboon N, Kuhn D, Wiesemann W (2018) Chebyshev inequalities for products of random variables. *Math. Oper. Res.* 43(3):887–918.

Smith E (2015) How the cash flows in Spotify streams. *Wall Street Journal* (June 3), https://www.wsj.com/articles/how-the-cash-flows-in-spotify-streams-1433376207.

Thiele A (2008) Pricing with uncertain customer valuations. *J. Revenue Pricing Management* 7(2):196–206.

Xue Z, Wang Z, Ettl M (2015) Pricing personalized bundles: A new approach and an empirical study. *Manufacturing Service Oper. Management* 18(1):51–68.