

Optimal Bundling Strategies Under Heavy-Tailed Valuations

Rustam Ibragimov

Department of Economics, Harvard University, Cambridge, Massachusetts 02138,
ibragim@fas.harvard.edu

Johan Walden

Haas School of Business, University of California, Berkeley, Berkeley, California 94720,
walden@haas.berkeley.edu

We develop a framework for the optimal bundling problem of a multiproduct monopolist, who provides goods to consumers with private valuations that are random draws from a distribution with heavy tails. We show that in the Vickrey auction setting, the buyers prefer separate provision of the goods to any bundles. We also provide a complete characterization of the optimal bundling strategies for a monopolist producer, who provides goods for profit-maximizing prices. For products with low marginal costs, the seller's optimal strategy is to provide goods separately when consumers' valuations are heavy-tailed and in a single bundle when valuations are thin-tailed. These conclusions are reversed for goods with high marginal costs. For simplicity, we use a specific class of independent and identically distributed random variables, but our results can be generalized to include dependence, skewness, and the case of nonidentical one-dimensional distributions.

Key words: optimal bundling strategies; multiproduct monopolist; Vickrey auction; profit-maximizing prices; heavy-tailed valuations

History: Received December 14, 2008; accepted July 21, 2010, by Jagmohan S. Raju, marketing. Published online in *Articles in Advance* October 11, 2010.

1. Introduction

In December 2007, two tickets to a Led Zeppelin reunion concert in London were sold in a charity auction. The face value of the two tickets was in total £250, and when the winning bid turned out to be £83,000—332 times the face value—this made headlines across the world. Kenneth Donnell, 25, who bought the tickets, stated in interviews that he had wanted to see his father's (sic!) favorite band live for years and that he had been sober when he joined the auction. Although the demand for tickets to the concert had been overwhelmingly higher than the supply, it is plausible to assume that most buyers' valuations of the tickets were far lower than what Mr. Donnell paid.

The case of Mr. Donnell and the Led Zeppelin tickets is just one example of very diverse private valuations observed in markets for cultural and sport events as well as in those for antiques and collectibles and online auctions and marketplaces such as eBay and StubHub. In these markets, bundling of goods is common practice, and a natural question is then what the consumers' and the seller's preferences over bundles are when the buyers have diverse private valuations.

The problem of optimal bundling strategy has received much attention over the last quarter of a century in the marketing and economics literature (see, e.g., the review in Stremersch and Tellis 2002 and references therein). However, the importance of the distribution of consumer valuations has, to the best of our knowledge, not been emphasized.

In this paper, we analyze the optimal bundling strategy for a multiproduct monopolist when the distribution of consumer valuations is *heavy-tailed*. We do this for two situations. In the first, the seller chooses how to bundle a given set of goods and sell the bundles in different auctions. In the second, he or she produces and provides the bundles for profit-maximizing prices. We focus on the analysis of pure bundling with one set of bundles offered for sale as opposed to mixed bundling, in which consumers can choose among all possible sets of bundles (see Adams and Yellen 1976, McAfee et al. 1989).

In the auction case, our main contribution is to complement and generalize the previous literature, e.g., Palfrey (1983), to the case of heavy-tailed valuations. Palfrey (1983) showed that in the case of two buyers, the seller will prefer to bundle the products. The two buyers, in contrast, unanimously prefer separate auctions to any other bundling decision.

Palfrey (1983) further showed that with bounded valuations, if there are more than two buyers, they will *never* unanimously prefer separate auctioning of the goods. This paper demonstrates that, on the contrary, with extremely heavy-tailed distributions, the buyers *always* unanimously prefer separate auctioning. The key distinction between the main results in Palfrey (1983) and ours is the distributional assumption on consumers' valuations.

In the case of profit-maximizing prices, the results of previous literature are completely reversed when valuations are extremely heavy-tailed. For instance, the results in Bakos and Brynjolfsson (1999) and Fang and Norman (2006) indicate that with thin-tailed valuations such as those with log-concave distributions, the optimal strategy for a multiproduct monopolist is to bundle goods with low marginal costs and to separately sell products with high marginal costs. We show that, to the contrary, under extremely heavy-tailed valuations the monopolist prefers bundling goods with high marginal costs and separately providing goods with low marginal costs. However, the results in the thin-tailed case in the previous literature continue to hold for moderately heavy-tailed valuations.

The main reason why the results are so different under heavy-tailed valuations is the following. Under *thin-tailed* valuations, consumers' valuations per good for a bundle typically have a lower spread, measured by variance, relative to the valuations for individual goods (see the discussion in Palfrey 1983, Schmalensee 1984, Salinger 1995, Bakos and Brynjolfsson 1999, and Fang and Norman 2006). Similarly, under *moderately* heavy-tailed reservation prices, the consumers' valuations per good for bundles have less spread relative to the valuations for component products, as measured by their *peakedness*.¹ Under *extremely* heavy-tailed valuations, this property is reversed: in this case, the spread of reservation prices per product for bundles, as measured by peakedness, is greater than that of valuations for components.² In the auction setting, given the relatively high spread of valuations of the bundled goods, the potential upside for the seller is then very high. Therefore, because the actual price is based on the second highest bid, the most important thing for the seller is to increase the chances that multiple buyers with high valuation bid in the same auction, which is achieved

by bundling. The argument is reversed for the buyers. Similar arguments can be made for the results with a monopolist producer, as elaborated upon in the paper.

The rest of this paper is organized as follows. In §2, we discuss related literature. In §3, we discuss heavy-tailed distributions and introduce some notation. Section 4 contains our main results, which, for tractability, are given in a rather special setting with independent and identically distributed (i.i.d.) valuations and so-called stable distributions. In §5, we discuss how the results can be generalized, and §6 concludes. All proofs, as well as some more detailed discussions, are left to the appendix.

2. Related Literature

Many studies have emphasized that bundling decisions of a monopolist providing two goods depend on correlations between consumers' valuations for the products (see Adams and Yellen 1976, McAfee et al. 1989, Schmalensee 1984, Salinger 1995), the degrees of complementarity and substitutability between the goods (e.g., Dansby and Conrad 1984, Lewbel 1985, Venkatesh and Kamakura 2003), and the marginal costs for the products (see, among others, Salinger 1995, Venkatesh and Kamakura 2003).

Most of these studies on bundling have focused, however, on prescribed distributions for valuations in the case of two products and their packages, such as bivariate uniform or Gaussian distributions, and only a few general results are available for larger bundles (e.g., Palfrey 1983; Bakos and Brynjolfsson 1999, 2000a, b; Fang and Norman 2006; Chu et al. 2010). For instance, Palfrey (1983) obtained characterizations of the monopolist's and buyers' preferences over bundled Vickrey auctions with valuations concentrated on a finite interval. In a related paper, Chakraborty (1999) obtained characterizations of optimal bundling strategies for a monopolist providing two independently priced goods on Vickrey auctions under a regularity condition on quantiles of bidders' valuations. As follows from Proschan's (1965) results given by Proposition B.1 in this paper, this regularity condition is satisfied for symmetric valuations with log-concave densities.³

Bakos and Brynjolfsson (1999) investigated optimal bundling decisions for a multiproduct monopolist providing large bundles of independently priced goods with zero marginal costs (information goods) for profit-maximizing prices to consumers whose valuations belong to a class that includes, again by

¹ The terms "reservation prices" and "valuations" are used as synonyms in this paper, in accordance with the well-established tradition in the bundling literature.

² The arguments in this paper are based on peakedness and majorization results for heavy-tailed distributions recently obtained in Ibragimov (2005, 2007). Appendices B.1 and B.2 provide a review of these results.

³ From Theorems B.1 and B.2, in Appendix B.2, it further follows that the regularity condition is also satisfied for moderately heavy-tailed valuations, but it does not hold for extremely heavy-tailed valuations. Therefore, Chakraborty's (1999) analysis cannot be applied if consumers' valuations are extremely heavy-tailed.

Proschan (1965), reservation prices with log-concave densities symmetric about the mean.⁴ Among other results, Bakos and Brynjolfsson (1999) showed that for this class of valuations, if the seller prefers bundling a certain number of goods to selling them separately and if the optimal price per good for the bundle is less than the mean valuation, then bundling any greater number of goods will further increase the seller's profits, compared to the case where the additional goods are sold separately. According to the result, in the above settings, a form of superadditivity for bundling decisions holds; that is, the benefits to the seller grow as the number of goods in the bundle increases.

Recently, Fang and Norman (2006) showed that a multiproduct monopolist providing bundles of independently priced goods to consumers with valuations with log-concave densities prefers selling them separately to any other bundling decision if the marginal costs of all the products are greater than the mean valuation; under some additional distributional assumptions, the seller prefers providing the goods as a single bundle to any other bundling decision if the marginal costs of the goods are identical and are less than the mean reservation price.

Chu et al. (2010) focus on the analysis of near optimality of bundle-size pricing where the prices for bundles depend (only) on their size. They also provide a range of numerical experiments for different cost scenarios and distributional assumptions on consumers' valuations, including exponential, logit, uniform, multivariate normal, and multivariate log-normal distributions, and an empirical analysis of pricing schemes for a theater company offering tickets for eight different plays or musicals and their packages.

Hitt and Chen (2005) and Wu et al. (2008) have focused on the analysis of customized bundling of information goods, a pricing strategy under which consumers can choose a certain quantity of goods sold for a fixed price. The results in these papers, in particular, show that under some commonly used assumptions, the mixed-bundle problem can be reduced to customized bundling. They further demonstrate how the customized-bundle solution is affected by heterogeneity and correlations in customers' valuations and by complementarity or substitutability among the goods sold.

⁴ In particular, the assumptions are satisfied for valuations with a finite support $[\underline{v}, \bar{v}]$ distributed as the truncation $XI(|X - \mu| < h)$, $h > 0$, of an arbitrary random variable X with a log-concave density symmetric about $\mu = (\underline{v} + \bar{v})/2$, where $h = (\bar{v} - \underline{v})/2$ and $I(\cdot)$ is the indicator function (see also Remark 2 in An 1998).

3. Heavy-Tailed Distributions

Previous literature has thus focused on valuations with *thin-tailed* distributions, such as those with log-concave densities or with a bounded support (see Appendix B.2 for the definition and a review of properties of log-concave distributions). The introduction of heavy-tailed distributions to the social sciences dates back to Mandelbrot (1963) (see also the papers in Mandelbrot 1997 and Fama 1965), who pioneered the study of heavy-tailed distributions with tails declining as $x^{-\alpha}$, $\alpha > 0$; i.e.,

$$P(|X| > x) \sim x^{-\alpha}. \quad (1)$$

Here, $f(x) \sim g(x)$ means that $c_1g(x) \leq f(x) \leq c_2g(x)$ for some constants $c_1 > 0$, $c_2 > 0$ for large x .

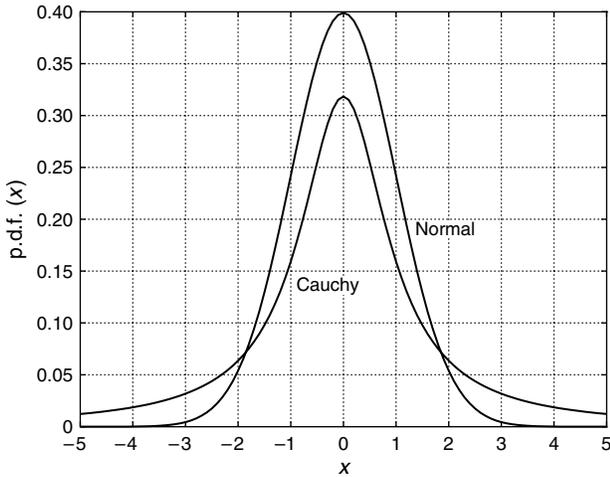
A random variable (r.v.) X with a distribution that satisfies (1) has finite moments $E|X|^p$ of order $p < \alpha$. However, the moments are infinite for $p \geq \alpha$. A r.v. X is said to be *thin-tailed* if its moments of all orders are finite: $E|X|^p < \infty$ for all $p > 0$. It is *heavy-tailed* if it follows a power law distribution (1). It is *moderately heavy-tailed* if it satisfies (1) with $1 < \alpha < \infty$, and it is *extremely heavy-tailed* if $\alpha < 1$. Distributions with log-concave densities, for which several general results in the optimal bundling literature exist, have finite moments of all orders and are therefore thin-tailed.

It has been documented in numerous studies that the time series encountered in many fields in marketing, economics, and finance are indeed heavy-tailed (see, among others, the discussion in Embrechts et al. 1997, Rachev et al. 2005, Gabaix et al. 2006, Gabaix 2009, Ibragimov 2009, Ibragimov et al. 2009, and references therein). In Appendix A, we provide a summary of this literature. In Figure 1, we compare the so-called Cauchy distribution with the thin-tailed standard normal distribution. The Cauchy distribution (with the location parameter $\mu = 0$ and the scale parameter $\sigma = 1$) has the probability density function (p.d.f.) $f(x) = 1/(\pi(1 + x^2))$ and satisfies (1) with $\alpha = 1$. It is therefore at the boundary between the classes of moderately heavy-tailed and extremely heavy-tailed distributions.

An important wide class of heavy-tailed distributions satisfying (1) is generated by scale mixtures of normal and other thin-tailed variables. Scale mixtures of normals include, for instance, the student- t distributions with arbitrary degrees of freedom (see the discussion in §5.2), the double exponential distribution, and the logistic distribution as well as symmetric stable distributions that are closed under summation and portfolio formation (see the next section).

As discussed in many works in the literature reviewed in Appendix A, in many aspects, under extreme observations and pronounced heterogeneity, heavy-tailed frameworks outperform those based on

Figure 1 Probability Density Function of Standardized Normal Distribution and Cauchy Distribution



Notes. Cauchy distribution has heavy tails and is less peaked than normal distribution. It is a power law with $\alpha = 1$.

thin-tailed distributions such as log-concave or super-exponential (e.g., normal) ones.

We note in passing that even for extremely heavy-tailed distributions, the *sample moments* of any order of i.i.d. draws will, of course, always be finite. However, these may not converge or even diverge to infinity when the number of draws increases. For example, for a distribution with the tail exponent $\alpha < 1$, the expectation (the first moment) is infinite. The average of any number of i.i.d. draws from such distribution will exist, but its limit as the number of draws increases will be infinity rather than a constant. For the Cauchy distribution, the average of any number of draws is also Cauchy distributed and therefore does not converge regardless of the number of draws.⁵ This also means that it is difficult to use finite samples to “prove” that a distribution is truly heavy-tailed: Extreme observations could be due to heavy-tailed distributions but could also be outliers in a sample from a distribution that is thin-tailed for large x . The difficulty of empirically separating these two possibilities is well known (see, e.g., Mandelbrot 1997, Perline 2005, and references therein) and is outside of the scope of this paper. However, the large range in observed private valuations, exemplified by Mr. Donnell and the Led Zeppelin tickets, is consistent with heavy-tailed private valuations and therefore of interest to study.

To illustrate the main ideas and to simplify the presentation of the main results in this paper, we first

⁵ For i.i.d. draws X_1, X_2, \dots from a Cauchy distribution with the tail exponent $\alpha = 1$, the average $\bar{X}_n = (1/n) \sum_{i=1}^n X_i$ has a Cauchy distribution with $\alpha = 1$ as well (see (4)) so that the limit of \bar{X}_n as the number n of draws increases is random (more precisely, \bar{X}_n converges in distribution to a Cauchy distributed r.v. as $n \rightarrow \infty$).

model heavy-tailedness using the framework of independent stable distributions, that is, distributions satisfying power-law relation (1) with $\alpha \in (0, 2)$. The stable distributions provide a fairly restricted class but are extremely convenient to analyze. They provide natural extensions of the Gaussian law because they are the only possible limits for appropriately normalized and centered sums of i.i.d. r.v.’s. This property is useful in representing heavy-tailed marketing, economic, and financial data as cumulative outcomes of market agents’ decisions in response to information they possess. In addition, stable distributions are flexible to accommodate both heavy-tailedness and skewness in data. In §5, we discuss how to generalize the results to much larger classes of distributions of private valuations, including dependence, power laws (1) with tail indices $\alpha > 2$, skewness, boundedness, and the case of nonidentical one-dimensional distributions.

3.1. Stable Distributions

For $0 < \alpha \leq 2$, $\sigma > 0$, $\beta \in [-1, 1]$, and $\mu \in \mathbf{R}$, we denote by $S_\alpha(\sigma, \beta, \mu)$ the stable distribution with the characteristic exponent (index of stability) α , the scale parameter σ , the symmetry index (skewness parameter) β , and the location parameter μ . That is, $S_\alpha(\sigma, \beta, \mu)$ is the distribution of a r.v. X with the characteristic function (c.f.)

$$E(e^{iyX}) = \begin{cases} \exp\{i\mu y - \sigma^\alpha |y|^\alpha (1 - i\beta \text{sign}(y) \tan(\pi\alpha/2))\}, & \alpha \neq 1; \\ \exp\{i\mu y - \sigma |y| (1 + (2/\pi)i\beta \text{sign}(y) \ln |y|)\}, & \alpha = 1; \end{cases} \quad (2)$$

$y \in \mathbf{R}$, where $i^2 = -1$ and $\text{sign}(y)$ is the sign of y defined by $\text{sign}(y) = 1$ if $y > 0$, $\text{sign}(0) = 0$, and $\text{sign}(y) = -1$ otherwise (relation (2) is one of possible parametrizations for stable distributions). In what follows, we write $X \sim S_\alpha(\sigma, \beta, \mu)$ if the r.v. X has the stable distribution $S_\alpha(\sigma, \beta, \mu)$. We also denote $\mathbf{R}_+ = [0, \infty)$.

A closed-form expression for the density $f(x)$ of a stable distribution is available in the following cases (and only in those cases): normal densities that correspond to the case $\alpha = 2$; the previously mentioned Cauchy densities $f(x) = \sigma / [\pi(\sigma^2 + (x - \mu)^2)]$, $x \in \mathbf{R}$, with $\alpha = 1$ and $\beta = 0$; and the densities $f(x) = (\sigma/(2\pi))^{1/2} \exp[-\sigma/(2(x - \mu))](x - \mu)^{-3/2}$, $x > \mu$; $f(x) = 0$, $x \leq \mu$, of Lévy distributions with $\alpha = 1/2$, $\beta = 1$, and their reflected versions (the case $\alpha = 1/2$ and $\beta = -1$). Although normal and Cauchy distributions are symmetric about μ , Lévy distributions are concentrated on the semiaxis $[\mu, \infty)$.

The stable distributions have different shapes depending on the parameters. The index of stability α characterizes the heaviness (the rate of decay) of the tails of stable distributions. The distribution of a stable r.v. $X \sim S_\alpha(\sigma, \beta, \mu)$ with $\alpha \in (0, 2)$ obeys the power law (1), and thus the p th absolute moments $E|X|^p$ of X are finite if $p < \alpha$ and are infinite otherwise. Therefore, stable distributions with $\alpha < 1$ are extremely heavy-tailed, with $1 < \alpha < 2$ they are moderately heavy-tailed, and with $\alpha = 2$ they are normal and, thus, thin-tailed.

For parametrization (2), in the case $\alpha > 1$, the location parameter μ is the mean of the stable distribution. One has

$$X - \mu \sim S_\alpha(\sigma, \beta, 0) \quad (3)$$

if $X \sim S_\alpha(\sigma, \beta, \mu)$. The symmetry index β characterizes the skewness of the distribution. The stable distributions with $\beta = 0$ are symmetric about the location parameter μ , as in the case of normal and Cauchy densities. Similar to the Lévy case $S_{1/2}(\sigma, 1, \mu)$, all stable distributions with $\beta = \pm 1$ and $\alpha \in (0, 1)$ (and only they) are one-sided; the support of these distributions is the semiaxis $[\mu, \infty)$ for $\beta = 1$ and is $(-\infty, \mu]$ for $\beta = -1$. In particular, the stable r.v.'s $X \sim S_\alpha(\sigma, 1, \mu)$, $\alpha \in (0, 1)$, are positive for $\mu \geq 0$. The parameter σ is a generalization of the concept of standard deviation; it coincides with the scaled standard deviation in the special case of normal distributions ($\alpha = 2$).

For two r.v.'s X and Y , we write $X =^d Y$ if X and Y have the same distribution.

Stable distributions are closed under portfolio formation, e.g., under summation. In particular, for i.i.d. stable r.v.'s $X_i \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \neq 1$, $i = 1, \dots, n$, and all $a_i \geq 0$, $i = 1, \dots, n$, $\sum_{i=1}^n a_i \neq 0$,

$$\frac{\sum_{i=1}^n a_i X_i}{(\sum_{i=1}^n a_i^\alpha)^{1/\alpha}} =^d X_1 \quad (4)$$

(see the monographs by Zolotarev 1986, Embrechts et al. 1997, and Uchaikin and Zolotarev 1999 for a detailed review of properties of stable distributions). Formula (4) also holds for $X_i \sim S_1(\sigma, 0, \mu)$.

4. A Framework for Modeling Optimal Bundling

We consider a setting with a single seller providing m goods to n consumers. Let $M = \{1, 2, \dots, m\}$ be the set of goods sold on the market, and let $J = \{1, 2, \dots, n\}$ denote the set of buyers. Let 2^M stand for the set of all subsets of M . As in Palfrey (1983), the seller's bundling decisions \mathcal{B} are defined as partitions of the set of items M into a set of subsets, $\{B_1, \dots, B_l\} = \mathcal{B}$; the subsets $B_s \in 2^M$, $s = 1, \dots, l$ are referred to as bundles. That is, $B_s \neq \emptyset$ for $s = 1, \dots, l$; $B_s \cap B_t = \emptyset$

for $s \neq t$, $s, t = 1, \dots, l$; and $\bigcup_{s=1}^l B_s = M$ (see Palfrey 1983, Bakos and Brynjolfsson 1999, Fang and Norman 2006). It is assumed that the seller can offer one (and only one) partition \mathcal{B} for sale on the market (this is referred to as pure bundling; see Adams and Yellen 1976). We denote by $\underline{\mathcal{B}} = \{\{1\}, \{2\}, \dots, \{m\}\}$ and $\overline{\mathcal{B}} = \{1, 2, \dots, m\}$ the bundling decisions corresponding, respectively, to the cases where the goods are sold separately (that is, on separate auctions or using unbundled sales) and as a single bundle M .

For a bundle $B \in 2^M$, we write $\text{card}(B)$ for the number of elements in B and denote by π_B the seller's profit resulting from selling the bundle, with the convention that $\pi_B = 0$ if the bundle is not sold. For a bundling decision $\mathcal{B} = \{B_1, \dots, B_l\}$, we write $\Pi_{\mathcal{B}}$ for the seller's total profit resulting from following \mathcal{B} , that is, $\Pi_{\mathcal{B}} = \sum_{s=1}^l \pi_{B_s}$.

A risk-neutral seller prefers (strictly prefers) a bundling decision \mathcal{B}_1 to a bundling decision \mathcal{B}_2 ex ante if $E\Pi_{\mathcal{B}_1} \geq E\Pi_{\mathcal{B}_2}$ (respectively, if $E\Pi_{\mathcal{B}_1} > E\Pi_{\mathcal{B}_2}$), where E denotes the expectation operator. The seller prefers a bundling decision \mathcal{B}_1 to a bundling decision \mathcal{B}_2 ex post if $\Pi_{\mathcal{B}_1} \geq \Pi_{\mathcal{B}_2}$ (a.s.), that is, if $P(\Pi_{\mathcal{B}_1} \geq \Pi_{\mathcal{B}_2}) = 1$. More generally, if the seller has an increasing utility of wealth function $U: \mathbf{R}_+ \rightarrow \mathbf{R}$, then she prefers (strictly prefers) a bundling decision \mathcal{B}_1 to a bundling decision \mathcal{B}_2 if $EU(\Pi_{\mathcal{B}_1}) \geq EU(\Pi_{\mathcal{B}_2})$ (respectively, if $EU(\Pi_{\mathcal{B}_1}) > EU(\Pi_{\mathcal{B}_2})$). The setting with a concave function U represents the case of a risk-averse seller. This paper focuses on characterizations of the seller and buyers' ex ante preferences over bundles of goods sold.

Consumers' preferences over the bundles $B \in 2^M$ are determined by their valuations (reservation prices) $v(B)$ for the bundles and, in particular, by their valuations $X_i = v(\{i\})$ for goods $i \in M$ (when the goods are sold separately), which are referred to as stand-alone valuations. Consumers' valuations for bundles of goods are assumed to be additive in those of component goods,

$$v(B) = \sum_{i \in B} v(\{i\}) = \sum_{i \in B} X_i, \quad (5)$$

and their utilities from consuming goods in $\mathcal{B} = \{B_1, \dots, B_l\}$ are given by

$$v(\mathcal{B}) = \sum_{s=1}^l v(B_s) = \sum_{s=1}^l \sum_{i \in B_s} v(\{i\}) = \sum_{s=1}^l \sum_{i \in B_s} X_i = \sum_{i=1}^m X_i. \quad (6)$$

If additivity conditions (5) and (6) hold, then the products provided by the monopolist are said to be *independently priced* (see Venkatesh and Kamakura 2003).

In the case where the valuations for bundles are nonnegative, $v(B) \geq 0$, $B \in 2^M$, it is said that the goods

in M and their bundles satisfy the *free disposal* condition. The free disposal condition is particularly important in the case of information goods and in the economics of the Internet (see Bakos and Brynjolfsson 1999; 2000a, b). In §4.2, the valuations $v(B)$ are allowed to be negative. This corresponds to the situation where the goods have negative value to some consumers (e.g., articles espousing certain political views, advertisements, or pornography in the case of information goods; see Bakos and Brynjolfsson 1999).

For our main results presented in the next two sections, X_i , $i \in M$, denote i.i.d. r.v.'s representing the distribution of consumers' valuations for goods $i \in M$ that determine their reservation prices for bundles.

For $j \in J$, the j th consumer's valuations for goods in M are assumed to be \tilde{X}_{ij} , $i \in M$, where $\tilde{X}^{(j)} = (\tilde{X}_{1j}, \dots, \tilde{X}_{mj})$, $j \in J$, are independent copies of the vector (X_1, \dots, X_m) , and her reservation prices $v_j(B)$ for bundles $B \in 2^M$ of goods in M are given by $v_j(B) = \sum_{i \in B} \tilde{X}_{ij}$. The seller is assumed to know only the distribution of consumers' reservation prices for goods in M and their bundles. The valuations $v_j(B)$ for bundles $B \in 2^M$ are known to buyer j ; however, the buyer has only the same incomplete information about the other consumers' reservation prices as does the seller (see Palfrey 1983).

4.1. Optimal Bundled Auctions with Heavy-Tailed Valuations

Let us first consider the case in which the goods in M and their bundles are provided by a seller through Vickrey auctions (see Palfrey 1983). The Vickrey auctions are separate and independent, one per bundle. In this setting, the buyers submit simultaneous sealed bids for bundles of goods. The highest bid wins the auction and pays the seller the second highest bid. It is well known that in such a setup, under additivity conditions (5) and (6), a dominant strategy for each bidder is to bid her true valuations for goods and their bundles.

Let $j \in J$ and let $\tilde{x}^{(j)} = (\tilde{x}_{1j}, \dots, \tilde{x}_{mj}) \in \mathbf{R}_+^m$. If a bundle B consisting of independently priced goods is offered for sale in a Vickrey auction, then the expectation of the surplus $S_j(B, \tilde{x}^{(j)})$ to consumer j with the values of stand-alone valuations $\tilde{X}^{(j)} = \tilde{x}^{(j)}$ and induced valuations for bundles $v_j(B) = \sum_{i \in B} \tilde{x}_{ij}$, $B \in 2^M$, is (see Palfrey 1983)

$$\begin{aligned}
 & ES_j(B, \tilde{x}^{(j)}) \\
 &= P\left(\max_{s \in J, s \neq j} v_s(B) < v_j(B)\right) \\
 &\quad \cdot \left(v_j(B) - E\left(\max_{s \in J, s \neq j} v_s(B) \mid \max_{s \in J, s \neq j} v_s(B) < v_j(B)\right)\right),
 \end{aligned}$$

where $v_s(B) = \sum_{i \in B} \tilde{X}_{is}$, $B \in 2^M$, $s \in J$, $s \neq j$. If the seller follows a bundling decision $\mathcal{B} = \{B_1, \dots, B_l\}$,

then the expectation of the surplus $S_j(\mathcal{B}, \tilde{x}^{(j)})$ to the j th buyer with the vector of stand-alone valuations $\tilde{X}^{(j)} = \tilde{x}^{(j)}$ is $ES_j(\mathcal{B}, \tilde{x}^{(j)}) = \sum_{s=1}^l ES_j(B_s, \tilde{x}^{(j)})$. The j th buyer with $\tilde{X}^{(j)} = \tilde{x}^{(j)}$ is said to prefer (strictly prefer) a bundling decision \mathcal{B}_1 to a bundling decision \mathcal{B}_2 , ex ante, if $ES_j(\mathcal{B}_1, \tilde{x}^{(j)}) \geq ES_j(\mathcal{B}_2, \tilde{x}^{(j)})$ (respectively, if $ES_j(\mathcal{B}_1, \tilde{x}^{(j)}) > ES_j(\mathcal{B}_2, \tilde{x}^{(j)})$). If all buyers $j \in J$ (strictly) prefer a bundling decision \mathcal{B}_1 to a bundling decision \mathcal{B}_2 ex ante for almost all realizations of their valuations $\tilde{X}^{(j)}$, it is said that buyers *unanimously* (strictly) prefer \mathcal{B}_1 to \mathcal{B}_2 ex ante. More precisely, buyers unanimously prefer (strictly prefer) a partition \mathcal{B}_1 to a partition \mathcal{B}_2 if, for all $j \in J$, $P[E(S_j(\mathcal{B}_1, \tilde{X}^{(j)}) \mid \tilde{X}^{(j)}) \geq E(S_j(\mathcal{B}_2, \tilde{X}^{(j)}) \mid \tilde{X}^{(j)})] = 1$ (respectively, $P[E(S_j(\mathcal{B}_1, \tilde{X}^{(j)}) \mid \tilde{X}^{(j)}) > E(S_j(\mathcal{B}_2, \tilde{X}^{(j)}) \mid \tilde{X}^{(j)})] = 1$), where, as usual, $E(\cdot \mid \tilde{X}^{(j)})$ stands for the expectation conditional on $\tilde{X}^{(j)}$.⁶

In accordance with the assumption of nonnegativity of bids and consumers' valuations usually imposed in the auction theory, we focus on the case where consumers' valuations for goods and bundles provided are nonnegative and model them using the framework of positive (extremely heavy-tailed) stable r.v.'s (see §3).

Theorem 4.1 shows that consumers unanimously prefer (ex ante) separate provision of goods on Vickrey auctions to any other bundling decision in the case of an arbitrary number of buyers, if their valuations are extremely heavy-tailed. In the case of more than two buyers, these results are reversals of those given by Theorem 6 in Palfrey (1983), from which it follows that if consumers' valuations are concentrated on a finite interval, then the buyers never unanimously prefer separate provision auctions (Theorem 4.1 does not contradict Theorem 6 in Palfrey 1983 because the support of heavy-tailed distributions in Theorem 4.1 is the infinite positive semiaxis \mathbf{R}_+).

THEOREM 4.1. *Suppose that the stand-alone valuations X_i , $i \in M$, for goods in M are i.i.d. positive stable r.v.'s such that $X_i \sim S_\alpha(\sigma, 1, 0)$ with $\alpha < 1$. Then buyers unanimously strictly prefer (ex ante) $\underline{\mathcal{B}}$ (that is, n separate auctions) to any other bundling decision.*

REMARK 4.1. Using property (3), similar to the proof of Theorem 4.1, it is straightforward to show that the theorem continues to hold for i.i.d. stable stand-alone valuations $X_i \geq \mu > 0$ such that $X_i \sim S_\alpha(\sigma, 1, \mu)$ (see §3.1).

The intuition behind the results given by Theorem 4.1 is a reversal of the intuition for the results

⁶ Clearly, in the case of discretely distributed valuations X_i , $i \in M$, consumers unanimously prefer \mathcal{B}_1 to \mathcal{B}_2 ex ante if each of them prefers \mathcal{B}_1 to \mathcal{B}_2 for all but a finite number of realizations of their stand-alone valuations.

in Palfrey (1983). In the case of extreme heavy-tailedness, consumers' valuations per good for bundles become less concentrated about the mean as the size of bundles increases (see the discussion in the introduction and Appendix B.1 and the results in Appendix B.2). Buyers who are on the upper tail of the distributions for the goods are more likely to win separate auctions, and the next highest bidder is likely to have relatively lower valuations than in the case of a bundled auction. Therefore, contrary to the case of bounded valuations (see the discussion preceding Theorem 5 in Palfrey 1983), the winner of the auction is likely to prefer separate provision of the products.

We note in passing that in this paper's setting with private values and the assumptions that valuations for each good as well as each agent are independently distributed (so that interdependent valuations and affiliated signals—see Milgrom and Weber 1982, Krishna 2002—are ruled out). English ascending auctions over bundles are weakly equivalent to Vickrey auctions. Therefore, the results given by Theorem 4.1 continue to hold for English auctions as well.

As shown by Palfrey (1983), in Vickrey auctions with independently priced goods and an arbitrary number of bidders, the total surplus (that is, the sum of the seller's profit and buyers' surplus) is always maximized in the case when the goods are provided in separate auctions. The results in Palfrey (1983) imply that under nonnegative valuations for individual goods and additive valuations for bundles, the seller prefers a single bundled Vickrey auction to any other bundling decision, if there are two buyers. The two buyers, on the other hand, unanimously prefer separate provision of items. The results for the two-buyer setting in Palfrey (1983) hold regardless of valuation distributions and therefore also for heavy-tailed valuations.

In the two-buyer setting, our results on consumers' preferences under heavy-tailed valuations are in accordance with Palfrey's. For more than two buyers, however, our results on the buyers' preferences differ from Palfrey's, who showed that (in the case of more than two buyers with bounded valuations) the buyers *never* unanimously prefer separate auctioning of the goods. By contrast, Theorem 4.1 shows that with heavy-tailed valuations they *always* unanimously prefer separate auctioning.

Palfrey's Theorem 2 shows that the seller always prefers to sell the goods in a single bundle when there are two buyers. This result holds regardless of distributions and therefore also when valuations are heavy-tailed. With more than two buyers, it is an open question what the optimal strategy for the seller is when distributions are heavy-tailed.

4.2. Optimal Bundling with Heavy-Tailed Valuations and Profit-Maximizing Prices

We turn to the case in which the prices for goods on the market and their bundles are set by the monopolist. To simplify the presentation of the results and their arguments, we assume that the marginal costs c_i of goods in M are identical: $c_i = c$, $i \in M$; however, extensions are possible for the case of arbitrary c_i . Suppose that the seller can provide bundles B of goods in M for prices per good $p \in [0, p_{\max}]$, where p_{\max} is the (regulatory) maximum price, with the convention that p_{\max} can be infinite. For a bundle of goods $B \in 2^M$, denote by p_B the profit-maximizing price per good for the bundle so that the seller's expected profit from producing and selling bundles of B s (at the price p_B per good) is

$$E(\pi_B) = nk(p_B - c)P(v(B) \geq kp_B),$$

where $k = \text{card}(B)$. We focus on the pure bundling case. The profit-maximizing price per good in the bundle is

$$\begin{aligned} p_B &= \arg \max_{p \in [0, p_{\max}]} (p - c)P(v(B) \geq kp) \\ &= \arg \max_{p \in [0, p_{\max}]} (p - c)P\left(\sum_{i=1}^k X_i \geq kp\right). \end{aligned}$$

Such optimization problems become much more complex in the mixed bundling case due to additional constraints in maximization for the buyers who can choose among many sets of bundles. We assume that $c < p_{\max}$ so that all bundles of goods in M are offered for sale. Clearly, in the case $c_i = c$ for all $i \in M$, the values of p_B are the same for all bundles B that consist of the same number $\text{card}(B)$ of goods. That is, $p_B = p_{B'}$ if $\text{card}(B) = \text{card}(B')$. Denote by \bar{p} the profit-maximizing price per good in the case where all the goods in M are sold as a single bundle and by \underline{p} the profit maximizing price of each good $i \in M$ under unbundled sales. That is, $\bar{p} = p_B$ with $B = M$, and $\underline{p} = p_B$ with $B = \{i\}$, $i \in M$.

Theorems 4.2 and 4.3 characterize the optimal bundling strategies for a multiproduct monopolist with an arbitrary degree of heavy-tailedness of valuations for goods in M . From Theorem 4.2 it follows that if consumers' reservation prices are moderately heavy-tailed, then the patterns in seller's optimal bundling strategies are the same as in the case of independently priced goods with log-concavely distributed (thin-tailed) valuations (see Bakos and Brynjolfsson 1999, Fang and Norman 2006, and the discussion in the introduction to this paper).

THEOREM 4.2. *Let $\mu \in \mathbf{R}$. Suppose that the stand-alone valuations X_i , $i \in M$, for goods in M are i.i.d. r.v.'s such that $X_i \sim S_\alpha(\sigma, \beta, \mu)$ with $\alpha > 1$. The risk-neutral seller*

strictly prefers \bar{B} to any other bundling decision (that is, the goods are sold as a single bundle) if $p < \mu$. The risk-neutral seller strictly prefers \underline{B} to any other bundling decision (that is, the goods are sold separately) if $\bar{p} > \mu$.

Theorem 4.3 shows that the patterns in the solutions to the seller's optimal bundling problem in Theorem 4.2 are reversed if consumers' valuations are extremely heavy-tailed.

THEOREM 4.3. Let $\mu \in \mathbf{R}$ and $p_{\max} < \infty$. Suppose that the stand-alone valuations X_i , $i \in M$, for goods in M are i.i.d. r.v.'s such that $X_i \sim S_\alpha(\sigma, \beta, \mu)$ with $\alpha < 1$. The risk-neutral seller strictly prefers \underline{B} to any other bundling decision (that is, the goods are sold separately) if $\bar{p} < \mu$. The risk-neutral seller strictly prefers \bar{B} to any other bundling decision (that is, the goods are sold as a single bundle) if $\underline{p} > \mu$.

REMARK 4.2. Analogues of Theorems 4.2 and 4.3 hold for expected utility comparisons for a risk-averse seller as long as her risk aversion is not too high. Specifically, because the preferences over bundling decisions in Theorems 4.2 and 4.3 are strict, they will also hold for a slightly risk-averse seller. For a severely risk-averse seller, however, the results in Theorems 4.2 and 4.3 may not hold (see also Theorem 3 in Ibragimov and Walden 2007 for a discussion of diversification decisions in the value at risk versus expected utility framework).

REMARK 4.3. From property (4) it follows that in the case $\alpha = 1$, $\beta = 0$, $P(\sum_{i=1}^k X_i \geq kp) = P(X_1 \geq p)$ for all $1 \leq k \leq m$, and, consequently, $p_B = p = \bar{p}$ and $E(\Pi_{\underline{B}}) = E(\Pi_{\bar{B}}) = E(\Pi_{\bar{B}})$ for all bundling decisions \mathcal{B} . Thus, for $\alpha = 1$, $\beta = 0$, the seller is indifferent among all bundling decisions in Theorems 4.2 and 4.3.

Similar to the argument based on variance in Bakos and Brynjolfsson (1999), the underlying intuition for Theorem 4.2 is that for moderately heavy-tailed distributions of reservation prices and the marginal costs of goods on the right of the mean valuation, bundling decreases profits because it reduces concentration (peakedness) of the valuation per good and thereby decreases the fraction of buyers with valuations for bundles greater than their total marginal costs (this is implied by the results in Appendix B.2; see also the discussion in the introduction and Appendix B.1). For the identical marginal costs of goods less than the mean valuation, bundling is likely to increase profits.

On the other hand, similar to Vickrey auctions in §4.1, the results in Theorem 4.3 are driven by the fact that in the case of extremely heavy-tailed reservation prices, concentration and peakedness of the valuations per good in bundles decrease with their size (see the introduction and Appendices B.1 and B.2). Therefore, bundling of goods in the case of extremely heavy-tailed valuations and marginal costs of goods

higher than the mean reservation price increases the fraction of buyers with reservation prices for bundles greater than their total marginal costs and thereby leads to an increase in the monopolist's profit. This effect is reversed in the case of the identical marginal costs on the left of the mean valuation.

The assumptions of Theorem 4.3 are satisfied, in particular, for positive stable valuations (stand-alone reservation prices) $X_i \sim S_\alpha(\sigma, 1, \mu)$, $i \in M$, where $\mu \geq 0$, $\sigma > 0$, and $\alpha \in (0, 1)$, for which the free disposal condition holds, including the Lévy distributions $S_{1/2}(\sigma, 1, \mu)$ with $\alpha = 1/2$ and $\mu \geq 0$. The condition $p_{\max} < \infty$ in Theorem 4.3 is necessary because otherwise the monopolist would set an infinite price for each bundle of goods under extremely heavy-tailed distributions of consumers' valuations considered in the theorem.

The general approach to the analysis of bundling strategies presented in this paper is directly applicable to bundling of information goods (that is, goods with zero marginal costs) in the case when, similar to the framework in Geng et al. (2005), consumers' valuations for infinite bundles $\{1, 2, 3, \dots\}$ of them are given by $v = \sum_{t=1}^{\infty} a_t X_t$, where a_t are nonnegative numbers reflecting the average valuations among the products and X_t are r.v.'s that capture the heterogeneity among the buyers. (As discussed in Geng et al. 2005, the above model involving a_t 's eventually decreasing to zero corresponds to the situation where consumers' average valuations for information goods decline with the number of goods consumed, as is typically the case for websites, online entertainment, weather forecasts, music, and news.) Furthermore, the framework presented in this paper allows one to obtain complete characterizations of optimal bundling in the case of valuations with arbitrary heavy-tailedness in such a setting. For instance, suppose that $X_t \sim S_\alpha(\sigma, 1, \mu)$, $t \geq 1$, where $\mu \geq 0$, $\sigma > 0$, and $\alpha \in (0, 1)$, are positive stable r.v.'s (e.g., one can take X_t to have the Lévy distributions $S_{1/2}(\sigma, 1, \mu)$ with $\alpha = 1/2$ and $\mu \geq 0$). Further, suppose that the monopolist can provide bundles $M \in \{1, 2, \dots\}$ of goods for prices that do not exceed $(\sum_{t \in M} a_t) p_{\max}$, where p_{\max} is some regulatory value. Let $\sum_{t=1}^{\infty} a_t^\alpha < \infty$ and thus $\sum_{t=1}^{\infty} a_t < \infty$. Then, similar to the proof of Theorem 4.3, it follows that bundling of the information goods in the above framework dominates their unbundled sales.⁷

⁷Indeed, the monopolist's profit from selling the information goods separately is $\pi_1 = (\sum_{t=1}^{\infty} a_t) \max_{\bar{p} \in [0, p_{\max}]} \bar{p} P(X_1 > \bar{p})$. Her profit from selling the goods as a single bundle is $\pi_2 = (\sum_{t=1}^{\infty} a_t) \max_{\bar{p} \in [0, p_{\max}]} \bar{p} P(\sum_{t=1}^{\infty} a_t X_t > (\sum_{t=1}^{\infty} a_t) \bar{p})$. Using (4) we further get that $\pi_2 = (\sum_{t=1}^{\infty} a_t) \max_{\bar{p} \in [0, p_{\max}]} \bar{p} P(X_1 > (\sum_{t=1}^{\infty} a_t) \bar{p} / (\sum_{t=1}^{\infty} a_t^\alpha)^{1/\alpha})$. Because, for $\alpha < 1$, $\sum_{t=1}^{\infty} a_t^\alpha > (\sum_{t=1}^{\infty} a_t)^\alpha$, we conclude that $\pi_2 > \pi_1$, as claimed.

5. Extensions

Our distributional assumption—i.i.d. stable valuations—has, so far, been very special. As discussed in the literature (see, e.g., the references in §2), consumers' valuations observed in many real-world markets are dependent and heterogeneous across goods and population. This is certainly true in markets for antiques and collectibles, TV channels, digital music, and other information goods. All the results obtained in this paper continue to hold for much broader classes of distributions than the i.i.d. stable distributions. Specifically, they hold for convolutions of stable distributions as well as for a wide class of multivariate distributions for which marginals are dependent and may be nonidentical and, in addition, may have finite variances, unlike stable distributions and their convolutions. Below, we give a short outline of these generalizations. For a thorough description of the methodology, see Ibragimov (2005, 2007).

5.1. Convolutions of Stable Distributions

Denote by $\overline{\mathcal{CS}}$ the class of distributions that are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with characteristic exponents $\alpha \in (1, 2]$ and $\sigma > 0$. Here and below, \mathcal{CS} stands for “convolutions of stable”; the overline indicates that convolutions of stable distributions with indices of stability *greater* than the threshold value of one are taken. That is, $\overline{\mathcal{CS}}$ consists of distributions of r.v.'s X such that for some $k \geq 1$, $X = Y_1 + \dots + Y_k$, where Y_i , $i = 1, \dots, k$, are independent r.v.'s such that $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in (1, 2]$, $\sigma_i > 0$, $i = 1, \dots, k$.

Further, $\overline{\mathcal{CSLC}}$ stands for the class of convolutions of distributions from the class $\overline{\mathcal{CS}}$ and the class \mathcal{LC} of symmetric log-concave distributions (see Appendix B.2). That is, $\overline{\mathcal{CSLC}}$ is the class of convolutions of symmetric distributions that are either log-concave or stable with characteristic exponents greater than one. \mathcal{CSLC} is the abbreviation of “convolutions of stable and log-concave.” In other words, $\overline{\mathcal{CSLC}}$ consists of distributions of r.v.'s X such that $X = Y_1 + Y_2$, where Y_1 and Y_2 are independent r.v.'s with distributions belonging to \mathcal{LC} or $\overline{\mathcal{CS}}$.

Finally, we denote by $\underline{\mathcal{CS}}$ the class of distributions that are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with indices of stability $\alpha \in (0, 1)$ and $\sigma > 0$. The underline indicates considering stable distributions with indices of stability *less* than the threshold value 1. That is, $\underline{\mathcal{CS}}$ consists of distributions of r.v.'s X such that for some $k \geq 1$, $X = Y_1 + \dots + Y_k$, where Y_i , $i = 1, \dots, k$, are independent r.v.'s such that $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in (0, 1)$, $\sigma_i > 0$, $i = 1, \dots, k$.⁸

Clearly, $\overline{\mathcal{CS}} \subset \overline{\mathcal{CSLC}}$ and $\mathcal{LC} \subset \overline{\mathcal{CSLC}}$. In what follows, we write $X \sim \mathcal{LC}$ (respectively, $X \sim \overline{\mathcal{CSLC}}$, $X \sim \overline{\mathcal{CS}}$ or $X \sim \underline{\mathcal{CS}}$) if the distribution of the r.v. X belongs to the class \mathcal{LC} (respectively, $\overline{\mathcal{CSLC}}$, $\overline{\mathcal{CS}}$ or $\underline{\mathcal{CS}}$). Similar to the stable case with $\alpha > 1$, the distributions of r.v.'s X in the classes $\overline{\mathcal{CSLC}}$ and $\overline{\mathcal{CS}}$ are moderately heavy-tailed in the sense that they have finite means: $E|X| < \infty$. In contrast, similar to the stable case with $\alpha < 1$, the distributions of r.v.'s X in $\underline{\mathcal{CS}}$ are extremely heavy-tailed in the sense that their first moments are infinite: $E|X| = \infty$.

From Theorems 3.1 and 3.2 in Ibragimov (2007), it follows that inequalities (B4) hold for i.i.d. r.v.'s $X_i \sim \overline{\mathcal{CSLC}}$ and inequalities (B5) hold for i.i.d. r.v.'s $X_i \sim \underline{\mathcal{CS}}$. Similar to the proof of the main results in this paper, this implies that Theorem 4.1 continues to hold for i.i.d. valuations $X_i \sim \underline{\mathcal{CS}}$, Theorem 4.2 continues to hold for i.i.d. valuations X_i such that $X_i - \mu \sim \overline{\mathcal{CSLC}}$, and Theorem 4.3 continues to hold for i.i.d. valuations X_i such that $X_i - \mu \sim \underline{\mathcal{CS}}$.

5.2. Dependence, Nonidentical Distributions, and Skewness

In addition to the convolution classes $\overline{\mathcal{CSLC}}$ and $\underline{\mathcal{CS}}$ in §5.1, all the results in the paper continue to hold for convolutions of *dependent* r.v.'s with joint α -symmetric distributions and their analogues with nonidentical marginals (see Fang et al. 1990 and the review in §2 in Ibragimov 2007). An n -dimensional distribution is called α -symmetric if its c.f. can be written as $\phi((\sum_{i=1}^n |y_i|^\alpha)^{1/\alpha})$, where $\phi: \mathbf{R}_+ \rightarrow \mathbf{R}$ is a continuous function (with $\phi(0) = 1$) and $\alpha > 0$. An important property of α -symmetric distributions is that, similar to stable laws, they satisfy property (4). The class of α -symmetric distributions contains, as a subclass, spherical distributions corresponding to the case $\alpha = 2$ (see Fang et al. 1990, p. 184). Spherical distributions, in turn, include such examples as Kotz type, multinormal, multivariate t , and multivariate spherically symmetric α -stable distributions (Fang et al. 1990, Chap. 3). Spherically symmetric stable distributions have characteristic functions $\exp[-\lambda(\sum_{i=1}^n y_i^2)^\gamma]$, $0 < \gamma \leq 2$, and are thus examples of α -symmetric distributions with $\alpha = 2$ and $\phi(x) = \exp(-x^\gamma)$. For any $0 < \alpha \leq 2$, the class of α -symmetric

same α . However, in general, this does not hold in the case of convolutions of stable distributions with *different* indices of stability. Therefore, the class $\overline{\mathcal{CS}}$ of convolutions of symmetric stable distributions with *different* indices of stability $\alpha \in (1, 2]$ is wider than the class of *all* symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with $\alpha \in (1, 2]$ and $\sigma > 0$. Similarly, the class $\underline{\mathcal{CS}}$ is wider than the class of *all* symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with $\alpha \in (0, 1)$ and $\sigma > 0$. It should also be noted that the class $\overline{\mathcal{CSLC}}$ is wider than the class of (twofold) convolutions of log-concave distributions with stable distributions $S_\alpha(\sigma, 0, 0)$ with $\alpha \in (1, 2]$ and $\sigma > 0$.

⁸ A linear combination of independent stable r.v.'s with the *same* characteristic exponent α also has a stable distribution with the

distributions includes distributions of risks X_1, \dots, X_n that have the common factor representation

$$(X_1, \dots, X_n) = (ZY_1, \dots, ZY_n), \quad (7)$$

where $Y_i \sim S_\alpha(\sigma, 0, 0)$ are i.i.d. symmetric stable r.v.'s with $\sigma > 0$ and the index of stability α and $Z \geq 0$ is a nonnegative r.v. independent of Y_i 's (see Fang et al. 1990, p. 197). Although the dependence structure in model (7) alone is restrictive, convolutions of such vectors provide a natural framework for modeling of random environments with different common shocks Z , such as macroeconomic or political ones, that affect all risks X_i (see Andrews 2003). In the case $Z = 1$ (a.s.), model (7) represents vectors with i.i.d. symmetric stable components that have c.f.'s $\exp[-\lambda \sum_{i=1}^n |t_i|^\alpha]$, which are particular cases of c.f.'s of α -symmetric distributions with $\phi(x) = \exp(-\lambda x^\alpha)$.

As discussed in §2 in Ibragimov (2007), convolutions of α -symmetric distributions exhibit both heavy-tailedness in marginals and dependence among them. For instance, the class of convolutions of models (7) with $\alpha < 1$ has extremely heavy-tailed marginal distributions with infinite means. On the other hand, convolutions of such models with $1 < \alpha \leq 2$ can have marginals with power moments finite up to a certain positive order (or finite exponential moments) depending on α and the choice of the r.v.'s Z . For instance, convolutions of models (7) with $1 < \alpha < 2$ and $E|Z| < \infty$ have finite means but infinite variances; however, marginals of such convolutions have infinite means if the r.v.'s Z satisfy $E|Z| = \infty$. Moments $E|ZY_i|^p$, $p > 0$, of marginals in models (7) with $\alpha = 2$ (that correspond to normal r.v.'s Y_i) are finite if and only if $E|Z|^p < \infty$. In particular, all marginal power moments in models (7) with $\alpha = 2$ are finite if $E|Z|^p < \infty$ for all $p > 0$. Similarly, marginals of spherically symmetric (that is, two-symmetric) distributions range from extremely heavy-tailed to thin-tailed ones. For example, marginal moments of spherically symmetric α -stable distributions with c.f.'s $\exp[-\lambda(\sum_{i=1}^n t_i^2)^\gamma]$, $0 < \gamma < 2$, are finite if and only if their order is less than γ . Marginal moments of a multivariate t -distribution with k degrees of freedom, which is an example of a spherical distribution, are finite if and only if the order of the moments is less than k . These distributions were used in a number of works to model heavy-tailedness phenomena with moments up to some order (see, among others, Praetz 1972, Blattberg and Gonedes 1974, Glasserman et al. 2002).

Let Φ stand for the class of c.f. generators ϕ such that $\phi(0) = 1$, $\lim_{t \rightarrow \infty} \phi(t) = 0$, and the function $\phi'(t)$ is concave. Consider random vectors (X_1, \dots, X_n) with

dependent components that satisfy one of the following conditions:

(C1) (X_1, \dots, X_n) is a sum of i.i.d. random vectors (Y_{1j}, \dots, Y_{nj}) , $j = 1, \dots, k$, where (Y_{1j}, \dots, Y_{nj}) has an absolutely continuous α -symmetric distribution with $\phi_j \in \Phi$ and $\alpha_j \in (0, 2]$;

(C2) (X_1, \dots, X_n) is a sum of i.i.d. random vectors $(Y_{1j}, \dots, Y_{nj}) = (Z_j V_{1j}, \dots, Z_j V_{nj})$, $j = 1, \dots, k$, in (7), where $V_{ij} \sim S_{\alpha_j}(\sigma_j, 0, 0)$, $i = 1, \dots, n$, $j = 1, \dots, k$, with $\sigma_j > 0$ and $\alpha_j \in (0, 2]$ and Z_j are positive absolutely continuous r.v.'s independent of V_{ij} .

From Theorems 3.1 and 3.2 in Ibragimov (2007), it follows that inequalities (B4) hold for random vectors (X_1, \dots, X_n) that satisfy (C1) or (C2) with $\alpha_j \in (1, 2]$ and inequalities (B5) hold for random vectors (X_1, \dots, X_n) that satisfy (C1) or (C2) with $\alpha_j \in (0, 1)$. Similar to the proofs in Appendix B.3, these results imply that Theorem 4.1 continues to hold under assumptions (C1) or (C2) with $\alpha_j \in (0, 1)$, Theorem 4.2 holds if the random vector $(X_1 - \mu, \dots, X_n - \mu)$ satisfies (C1) or (C2) with $\alpha_j \in (1, 2]$, and Theorem 4.3 holds if $(X_1 - \mu, \dots, X_n - \mu)$ satisfies (C1) or (C2) with $\alpha_j \in (0, 1)$.

We note that all the results in the paper are available for the case of *skewed* distributions, including skewed stable distributions (such as, for instance, extremely heavy-tailed Lévy distributions with $\alpha = 1/2$ concentrated on the positive semiaxis). In addition, as follows from Proposition 3.1 in Ibragimov (2007), for independent not necessarily identically distributed r.v.'s $X_i \sim S_\alpha(\sigma_i, \beta, \mu)$, inequalities (B4) hold if $\alpha > 1$ and $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{n+1} > 0$ and inequalities (B5) hold if $\alpha < 1$ and $\sigma_{n+1} \geq \sigma_n \geq \dots \geq \sigma_1 > 0$. Similar to the proof of the results in this paper, this implies their extensions to the case of nonidentically distributed consumers' valuations. In addition, using Remark B.2, one can further obtain analogues of the results in the bounded case and truncations of heavy-tailed distributions. Analogously, it is not difficult to obtain generalizations of the results to the case of distributions with dependent, possibly skewed, bounded and not necessarily identically distributed marginals, including convolutions of shifted and scaled α -symmetric distributions and their truncations.

Thus, our model allows for a unified analysis of the effects of all the main distributional properties of consumers' valuations on optimal bundling strategies for a multiproduct monopolist, including heavy-tailedness, dependence, skewness, boundedness, and the case of nonidentical one-dimensional distributions.

6. Concluding Remarks

As shown in this paper, the optimal bundling decision for a multiproduct seller of goods depends fundamentally on the distribution of consumer valuations

together with the supply of goods (fixed, high marginal cost of production, or low marginal cost of production). Previous studies have focused on the thin-tailed case, but there is a priori no reason to believe that the distribution of consumers' valuations is always thin-tailed. Our results complement previous literature by showing that the optimal bundling decision may be different when consumer valuations are heavy-tailed.

Acknowledgments

The authors thank the National University of Singapore Risk Management Institute for support and Ganesh Iyer for valuable comments. Rustam Ibragimov gratefully acknowledges research support provided by the National Science Foundation Grant SES-0820124. The authors also thank the editor, Gérard Cachon; the department editor, Jagmohan Raju; an associate editor; and the referees for many helpful comments and suggestions.

Appendix A. Heavy-Tailed Distributions in Marketing, Economics, and Finance

Extreme events and heavy-tailedness phenomena are especially pronounced in markets characterized by the *nobody knows*, the *winner takes all*, and the *success breeds success* principles with high uncertainty in individual demands and in the success or failure of new products, such as markets for technological innovations and information goods and creative (e.g., motion picture, music, and book publishing) industries (see, among others, the discussion and reviews in Shapiro and Varian 1999, De Vany 2004, Frank and Cook 1995, Anderson 2006, Eliashberg et al. 2006, Taleb 2007, and Gaffeo et al. 2008).

We mention a sample of estimates of tail indices α in heavy-tailed models (1) for demand characteristics reported in the empirical literature for different markets: $\alpha = 1.2$ for book demand on Amazon.com and $\alpha = 1.1$ for a sample of sales on Amazon.com for books not readily available through brick-and-mortar retailers (the point estimates given by the inverse of $\beta = 0.8$ in log-log regression (1) in Ghose and Sundararajan 2006 and the inverse of $\beta = 0.9$ in log-log regression (12) in Brynjolfsson et al. 2003); $\alpha \in (0.9, 1.5)$ for book sales and revenues in Italy (Gaffeo et al. 2008); $\alpha \in (0.2, 1.7)$ for market shares of different brands and categories of foods in the United States (the inverses of the estimates of the parameter b in Table 1 in Kohli and Sah 2006; see also Kalyanaram et al. 1995, who report the square-root relation between the market share and order of entry in markets for prescription antiulcer drugs and certain packaged consumer goods); $\alpha \in (1.5, 2.3)$, $\alpha \in (1.3, 1.8)$, $\alpha \in (1.1, 1.7)$, and $\alpha \in (2.1, 2.6)$ for, respectively, box office revenues, rates of return, profits, and losses for motion pictures with different ratings in North America (De Vany and Walls 2002, 2004; and §10 in De Vany 2004); $\alpha = 0.6, 0.8$ depending on the estimator used for 1998 motion picture revenues in the United Kingdom (Collins et al. 2002); $\alpha = 0.45$ for Rock and Roll performers' revenues in the United States and $\alpha = 0.55$ for revenues of their promoters (Connolly and Krueger 2006); $\alpha \in (2.9, 3.8)$ for citations of U.S. technology patents (Bentley et al. 2004). Silverberg and Verspagen (2007) report the tail

indices α around 0.6–1 for financial returns from technological innovations and around 3–5 for patent citations.

Appendix B. Proofs

B.1. Probabilistic Foundations for the Main Results

The proof of the results in this paper is based on general results on peakedness properties of convolutions of distributions and majorization phenomena for tail probabilities of linear combinations of r.v.'s presented in Appendix B.2. These properties and phenomena were first analyzed, under the assumptions of log-concavity of distributions, in the seminal paper by Proschan (1965) that found applications in the study of many problems in statistics, econometrics, economic theory, and other fields (see the discussion in Ibragimov 2005). The proof of the main results in this paper is based on analogues of the results in Proschan (1965) in the case of heavy-tailed distributions recently obtained by Ibragimov (2007) and also presented in Ibragimov (2005). To our knowledge, the results in Ibragimov (2005, 2007) are the first ones in the literature that give extensions of those in Proschan (1965) to the case of heavy-tailedness and their reversals for general classes of distributions. These results provide the key to the analysis of bundling problems under heavy-tailed valuations in this paper. Aside from the analysis of optimal bundling strategies considered in this paper, the majorization results obtained in Ibragimov (2005, 2007) have many other applications. These applications include the study of efficiency of linear estimators and monotone consistency of the sample mean, robustness of the model of demand-driven innovation and spatial competition over time, and value at risk analysis as well as that of inheritance models in mathematical evolutionary theory (see Ibragimov 2005). The proof of the results reviewed in Appendix B.2 overcomes the main technical difficulties in the analysis of optimal bundling in this paper and other applications discussed above.

B.2. Majorization and Peakedness Properties of Log-Concave and Stable Distributions

We say that a r.v. X with density $f: \mathbf{R} \rightarrow \mathbf{R}$ and the convex distribution support $\Omega = \{x \in \mathbf{R}: f(x) > 0\}$ is log-concavely distributed if $\log f(x)$ is concave in $x \in \Omega$; that is, if for all $x_1, x_2 \in \Omega$, and any $\lambda \in [0, 1]$,

$$f(\lambda x_1 + (1 - \lambda)x_2) \geq (f(x_1))^\lambda (f(x_2))^{1-\lambda} \quad (\text{B1})$$

(see An 1998). A distribution is said to be log-concave if its density f satisfies (B1).

If a r.v. X is log-concavely distributed, then its density has at most an exponential tail; that is, $f(x) = O(\exp(-\lambda x))$ for some $\lambda > 0$, as $x \rightarrow \infty$ and all the power moments $E|X|^p$, $p > 0$, of the r.v. exist (see Corollary 1 in An 1998). Marshall and Olkin (1979) and An (1998) provide surveys of many other properties of log-concave distributions.⁹

⁹ Examples of log-concave distributions include the normal distribution, the uniform density, the exponential density, the Gamma distribution $\Gamma(\alpha, \beta)$ with the shape parameter $\alpha \geq 1$, the Beta distribution $B(a, b)$ with $a \geq 1$ and $b \geq 1$, and the Weibull distribution $\mathcal{W}(\gamma, \alpha)$ with the shape parameter $\alpha \geq 1$.

Denote by $\mathcal{L}\mathcal{C}$ the class of symmetric log-concave distributions.¹⁰

DEFINITION B.1 (MARSHALL AND OLKIN 1979). Let $a, b \in \mathbf{R}^n$. The vector a is said to be majorized by the vector b , written $a < b$, if $\sum_{i=1}^k a_{[i]} \leq \sum_{i=1}^k b_{[i]}$, $k = 1, \dots, n-1$, and $\sum_{i=1}^n a_{[i]} = \sum_{i=1}^n b_{[i]}$, where $a_{[1]} \geq \dots \geq a_{[n]}$ and $b_{[1]} \geq \dots \geq b_{[n]}$ denote components of a and b in decreasing order.

The relation $a < b$ implies that the components of the vector b are more diverse than are those of a (see Marshall and Olkin 1979). In this context, it is easy to see that the following relations hold (see Marshall and Olkin 1979, p. 7):

$$\left(\sum_{i=1}^n a_i/n, \dots, \sum_{i=1}^n a_i/n \right) < (a_1, \dots, a_n) < \left(\sum_{i=1}^n a_i, 0, \dots, 0 \right), \quad a \in \mathbf{R}_+^n, \quad (\text{B2})$$

for all $a \in \mathbf{R}_+^n$. In particular,

$$\begin{aligned} & (1/(n+1), \dots, 1/(n+1), 1/(n+1)) \\ & < (1/n, \dots, 1/n, 0), \quad n \geq 1. \end{aligned} \quad (\text{B3})$$

DEFINITION B.2 (MARSHALL AND OLKIN 1979). A function $\phi: A \rightarrow \mathbf{R}$ defined on $A \subseteq \mathbf{R}^n$ is called *Schur-convex* (respectively, *Schur-concave*) on A if $(a < b) \implies (\phi(a) \leq \phi(b))$ (respectively, $(a < b) \implies (\phi(a) \geq \phi(b))$) for all $a, b \in A$. If, in addition, $\phi(a) < \phi(b)$ (respectively, $\phi(a) > \phi(b)$) whenever $a < b$ and a is not a permutation of b , then ϕ is said to be *strictly Schur-convex* (respectively, *strictly Schur-concave*) on A .

DEFINITION B.3 (MARSHALL AND OLKIN 1979, p. 372). A r.v. X is more peaked about $\mu \in \mathbf{R}$ than is Y if $P(|X - \mu| > x) \leq P(|Y - \mu| > x)$ for all $x \geq 0$. If these inequalities are strict whenever the two probabilities are not both 0 or both 1, then the r.v. X is strictly more peaked about μ than is Y . A r.v. X is said to be (strictly) less peaked about μ than is Y if Y is (strictly) more peaked about μ than is X .

In the case $\mu = 0$, we simply say that the r.v. X is (strictly) more peaked than Y .

Roughly speaking, a r.v. X is more peaked about $\mu \in \mathbf{R}$ than is Y , if the distribution of X is more concentrated about μ than is that of Y .

Proschan (1965) obtained the following well-known result concerning majorization and peakedness properties of tail probabilities of linear combinations of log-concavely distributed r.v.'s:

PROPOSITION B.1 (PROSCHAN 1965; SEE ALSO THEOREM 12.J.1 IN MARSHALL AND OLKIN 1979). If X_1, \dots, X_n are i.i.d. r.v.'s such that $X_i \sim \mathcal{L}\mathcal{C}$, then the function $\psi(a, x) = P(\sum_{i=1}^n a_i X_i > x)$ is strictly Schur-convex in $a = (a_1, \dots, a_n) \in \mathbf{R}_+^n$ for $x > 0$ and is strictly Schur-concave in $a = (a_1, \dots, a_n) \in \mathbf{R}_+^n$ for $x < 0$.

Clearly, from Proposition B.1 it follows that under its assumptions, $\sum_{i=1}^n a_i X_i$ is strictly more peaked than $\sum_{i=1}^n b_i X_i$ if $a < b$ and a is not a permutation of b .

Theorems B.1 and B.2 in this section give analogues of Proposition B.1 for heavy-tailed stable r.v.'s. These

propositions follow from more general results obtained in Ibragimov (2005, 2007) (see Theorems 4.3 and 4.4 in Ibragimov 2005, Theorems 3.1 and 3.2 in Ibragimov 2007, and the discussion and §5 in Ibragimov and Walden 2007).

According to Theorem B.1 below, peakedness and majorization properties of linear combinations of r.v.'s with moderately heavy-tailed stable distributions are the same as in the case of log-concave distributions in Proschan (1965).

THEOREM B.1 (IBRAGIMOV 2005, 2007). Proposition B.1 holds if X_1, \dots, X_n are i.i.d. r.v.'s such that $X_i \sim S_\alpha(\sigma, \beta, 0)$ with $\alpha > 1$.

As follows from Theorem B.2, peakedness properties given by Proposition B.1 and Theorem B.1 above are reversed in the case of r.v.'s with extremely heavy-tailed stable distributions.

THEOREM B.2 (IBRAGIMOV 2005, 2007). If X_1, \dots, X_n are i.i.d. r.v.'s such that $X_i \sim S_\alpha(\sigma, \beta, 0)$ with $\alpha < 1$, then the function $\psi(a, x)$ in Proposition B.1 is strictly Schur-concave in $(a_1, \dots, a_n) \in \mathbf{R}_+^n$ for $x > 0$ and is strictly Schur-convex in $(a_1, \dots, a_n) \in \mathbf{R}_+^n$ for $x < 0$.

From Theorem B.1 it follows that similar to the class $\mathcal{L}\mathcal{C}$ covered by Proposition B.1, $\sum_{i=1}^n a_i X_i$ is strictly more peaked than is $\sum_{i=1}^n b_i X_i$ for moderately heavy-tailed X_i 's, if $a < b$ and a is not a permutation of b . However, according to Theorem B.2, if $a < b$ and a is not a permutation of b , then $\sum_{i=1}^n a_i X_i$ is strictly less peaked than is $\sum_{i=1}^n b_i X_i$ for extremely heavy-tailed X_1, \dots, X_n .

REMARK B.1. Using comparisons (B3), from Theorems B.1 and B.2 it follows that the following inequalities hold for all $x > 0$ and $n \geq 2$:

$$P\left(\frac{1}{n+1} \sum_{i=1}^{n+1} X_i > x\right) < P\left(\frac{1}{n} \sum_{i=1}^n X_i > x\right) < P(X_1 > x) \quad (\text{B4})$$

for i.i.d. r.v.'s $X_i \sim S_\alpha(\sigma, \beta, 0)$ with $\alpha > 1$,

$$P(X_1 > x) < P\left(\frac{1}{n} \sum_{i=1}^n X_i > x\right) < P\left(\frac{1}{n+1} \sum_{i=1}^{n+1} X_i > x\right) \quad (\text{B5})$$

for i.i.d. r.v.'s $X_i \sim S_\alpha(\sigma, \beta, 0)$ with $\alpha < 1$.

REMARK B.2. Using truncation arguments similar to Ibragimov and Walden (2007), it is not difficult to show that analogues of Theorems B.1 and B.2 also hold for bounded r.v.'s X_i . For instance, consider i.i.d. r.v.'s X_i , $i = 1, \dots, N$, given by truncations of stable r.v.'s. That is, let $X_i = Y_i I(|Y_i| < K)$, $K > 0$, where Y_i , $i = 1, \dots, N$, are i.i.d. stable r.v.'s: $Y_i \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \in (0, 2]$ (evidently, $X_i \rightarrow Y_i$ in distribution as $K \rightarrow \infty$). Then, with given $\alpha > 1$, $x > 0$, and $N \geq 1$, there exists a sufficiently large truncation value $K = K(\alpha, x, N)$ such that inequalities (B4) hold for all $n = 1, \dots, N-1$. Similarly, with given $\alpha < 1$, $x > 0$, and $N \geq 1$, inequalities (B5) hold for all $n = 1, \dots, N-1$ and a sufficiently large truncation bound $K = K(\alpha, x, N)$.

B.3. Proof of Theorems 4.1–4.3

PROOF OF THEOREM 4.1. Suppose that X_i , $i \in M$, are i.i.d. positive stable r.v.'s such that $X_i \sim S_\alpha(\sigma, 1, 0)$ with $\alpha < 1$. Let $j \in J$ and let the vector $\tilde{X}^{(j)}$ of the j th buyer's reservation prices for goods in M take a value $\tilde{x}^{(j)} = (\tilde{x}_{1j}, \dots, \tilde{x}_{mj}) \in \mathbf{R}_+^m$, $(\tilde{x}_{1j}, \dots, \tilde{x}_{mj}) \neq (0, 0, \dots, 0)$. Consider any bundle $B \in 2^M$

¹⁰ $\mathcal{L}\mathcal{C}$ stands for "log-concave."

with $\text{card}(B) = k \geq 2$. The j th buyer's reservation price for the bundle is $v_j(B) = \sum_{i \in B} \tilde{x}_{ij}$. Denote $H_k(x) = P(\sum_{i=1}^k X_i \leq x)$, $x \geq 0$. Similar to Palfrey (1983), we get that the expected surplus to the buyer when B is offered for sale is

$$ES_j(B, \tilde{x}^{(j)}) = \int_0^{v_j(B)} (H_k(x))^{n-1} dx \\ = k \int_0^{v_j(B)/k} (H_k(kx))^{n-1} dx. \quad (B6)$$

On the other hand, the expected surplus to consumer j when good $i \in B$ is offered for sale separately is $ES_j(\{i\}, \tilde{x}^{(j)}) = \int_0^{\tilde{x}_{ij}} (H_1(x))^{n-1} dx$. By (B5), $H_k(kx) < H_1(x)$ for all $x > 0$. This together with (B6) implies

$$ES_j(B, \tilde{x}^{(j)}) < k \int_0^{v_j(B)/k} (H_1(x))^{n-1} dx \quad (B7)$$

if $v_j(B) > 0$. Because the function $(H_1(y))^{n-1}$ is increasing in $y \in \mathbf{R}_+$, from Theorem 3.C.1 in Marshall and Olkin (1979) we get that the function $F(y_1, \dots, y_k) = \sum_{i=1}^k \int_0^{y_i} (H_1(x))^{n-1} dx$ is Schur-convex in $(y_1, \dots, y_k) \in \mathbf{R}_+^k$. Therefore, from majorization comparisons (B2), it follows that $F(y_1, \dots, y_k) \geq F(\sum_{i=1}^k y_i/k, \dots, \sum_{i=1}^k y_i/k)$ for all $(y_1, \dots, y_k) \in \mathbf{R}_+^k$ (see also the proof of Theorem 5 in Palfrey 1983). In particular,

$$k \int_0^{v_j(B)/k} (H_1(x))^{n-1} dx \leq \sum_{i \in B} \int_0^{\tilde{x}_{ij}} (H_1(x))^{n-1} dx \\ = \sum_{i \in B} ES_j(\{i\}, \tilde{x}^{(j)}). \quad (B8)$$

From (B7) and (B8) we get

$$ES_j(B, \tilde{x}^{(j)}) < \sum_{i \in B} ES_j(\{i\}, \tilde{x}^{(j)}) \quad (B9)$$

if $v_j(B) > 0$ (clearly, (B9) holds as equality if $v_j(B) = 0$). By (B9), we have that if the seller follows a bundling decision $\mathcal{B} = \{B_1, \dots, B_l\}$ such that $\text{card}(B_s) = k_s$, $s = 1, \dots, l$, and $k_t \geq 2$ for at least one $t \in \{1, \dots, l\}$, then the expected surplus $ES_j(\mathcal{B}, \tilde{x}^{(j)})$ to buyer j satisfies $ES_j(\mathcal{B}, \tilde{x}^{(j)}) = \sum_{s=1}^l ES_j(B_s, \tilde{x}^{(j)}) < \sum_{i=1}^m ES_j(\{i\}, \tilde{x}^{(j)}) = ES_j(\underline{\mathcal{B}}, \tilde{x}^{(j)})$. \square

PROOFS OF THEOREMS 4.2 AND 4.3. Let $\mu \in \mathbf{R}$ and let $c < p_{\max} < \infty$. Suppose that the valuations X_i , $i \in M$, are i.i.d. r.v.'s such that $X_i \sim S_\alpha(\sigma, \beta, \mu)$ with $\alpha < 1$. We will show that the seller's profit maximizing bundling decision is $\underline{\mathcal{B}}$ if the prices per good $p_B < \mu$ for all bundles $B \in 2^M$ and is $\bar{\mathcal{B}}$ if $p_B > \mu$ for all $B \in 2^M$. Let $\underline{\pi}$ denote the monopolist's profit from each good under their separate provision, namely, $\underline{\pi} = \pi_{B_i}$ with $B_i = \{i\}$, $i \in M$: $E(\underline{\pi}) = n(p - c)P(X_1 \geq p)$.

Suppose that $p_B < \mu$ for all $B \in 2^M$. From property (3) and comparisons (B5) applied to the stable r.v.'s $-X_i$ it follows that for any bundle $B \in 2^M$ with the number of goods $\text{card}(B) = k \geq 2$, $E(\pi_B) = nk(p_B - c) \cdot P(v(B) \geq kp_B) = nk(p_B - c)P(\sum_{i=1}^k X_i \geq kp_B) < nk(p_B - c) \cdot P(X_1 \geq p_B) \leq kE(\underline{\pi})$. This implies that for any bundling decision $\mathcal{B} = \{B_1, \dots, B_l\}$ such that $\text{card}(B_s) = k_s$, $s = 1, \dots, l$, and $k_t \geq 2$ for at least one $t \in \{1, \dots, l\}$,

$$E(\Pi_{\mathcal{B}}) = \sum_{s=1}^l E(\pi_{B_s}) < \sum_{s=1}^l k_s E(\underline{\pi}) = mE(\underline{\pi}) = E(\Pi_{\underline{\mathcal{B}}}). \quad (B10)$$

Suppose now that $p_B > \mu$ for all $B \in 2^M$. Then using property (3) and relations (B5) we get that for any bundle $B \in 2^M$ with $\text{card}(B) = k \leq m - 1$, $E(\pi_B) = nk(p_B - c)P(\sum_{i \in B} X_i \geq kp_B) < nk(p_B - c)P(\sum_{i=1}^m X_i \geq mp_B)$. Therefore, for any bundling decision $\mathcal{B} = \{B_1, \dots, B_l\}$ such that $\text{card}(B_s) = k_s$, $s = 1, \dots, l$, and $k_t \leq m - 1$ for at least one $t \in \{1, \dots, l\}$,

$$E(\Pi_{\mathcal{B}}) = \sum_{s=1}^l E(\pi_{B_s}) < n \sum_{s=1}^l k_s (p_{B_s} - c) P\left(\sum_{i=1}^m X_i \geq mp_{B_s}\right) \\ \leq \sum_{s=1}^l (k_s/m) E(\Pi_{\bar{\mathcal{B}}}) = E(\Pi_{\bar{\mathcal{B}}}). \quad (B11)$$

From (B10) and (B11) we get that the profit maximizing bundling decision is $\underline{\mathcal{B}}$ if $p_B < \mu$ for all $B \in 2^M$ and is $\bar{\mathcal{B}}$ if $p_B > \mu$ for all $B \in 2^M$.

Let us show that the condition that $p_B > \mu$ for all $B \in 2^M$ holds if $p > \mu$. Indeed, suppose that this not the case and there exists a bundle $B \in 2^M$ with $\text{card}(B) = k \geq 2$ and $p_B \leq \mu$. Then, as above, we get $kE(\underline{\pi}) = nk(p - c)P(X_1 \geq p) < nk(p - c)P(\sum_{i=1}^k X_i \geq kp) \leq E(\pi_B)$. On the other hand, $E(\pi_B) = nk(p_B - c)P(\sum_{i=1}^k X_i \geq kp_B) < nk(p_B - c)P(X_1 \geq p_B) \leq kE(\underline{\pi})$, which is a contradiction. Similarly, we get that $\bar{p} < \mu$ implies that $p_B < \mu$ for all $B \in 2^M$. This completes the proof of Theorem 4.3. Theorem 4.2 could be proven in a similar way, with the use of inequalities (B4) instead of relations (B5). \square

References

- Adams, W. J., J. L. Yellen. 1976. Commodity bundling and the burden of monopoly. *Quart. J. Econom.* **90**(3) 475–498.
- An, M. Y. 1998. Logconcavity versus logconvexity: A complete characterization. *J. Econom. Theory* **80**(2) 350–369.
- Anderson, C. 2006. *The Long Tail*. Hyperion, New York.
- Andrews, D. W. K. 2003. Cross-section regression with common shocks. *Econometrica* **73**(5) 1551–1585.
- Bakos, Y., E. Brynjolfsson. 1999. Bundling information goods: Pricing, profits and efficiency. *Management Sci.* **45**(12) 1613–1630.
- Bakos, Y., E. Brynjolfsson. 2000a. Aggregation and disaggregation of information goods: Implications for bundling, site licensing and micropayment systems. B. Kahin, H. R. Varian, eds. *Internet Publishing and Beyond: The Economics of Digital Information and Intellectual Property*. MIT Press, Cambridge, MA, 114–137.
- Bakos, Y., E. Brynjolfsson. 2000b. Bundling and competition on the internet. *Marketing Sci.* **19**(1) 63–82.
- Bentley, R. A., M. W. Hahn, S. J. Shennan. 2004. Random drift and culture change. *Proc. Roy. Soc. B* **271**(1547) 1443–1450.
- Blattberg, R. C., N. J. Gonedes. 1974. Comparison of stable and student distributions as statistical models for stock prices. *J. Bus.* **47**(2) 244–280.
- Brynjolfsson, E., Y. Hu, M. D. Smith. 2003. Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management Sci.* **49**(11) 1580–1596.
- Chakraborty, I. 1999. Bundling decisions for selling multiple objects. *Econom. Theory* **13**(3) 723–733.
- Chu, C. S., P. Leslie, A. Sorensen. 2010. Bundle-size pricing as an approximation to mixed bundling. *Amer. Econom. Rev.* Forthcoming.
- Collins, A., C. Hand, M. C. Snell. 2002. What makes a blockbuster? Economic analysis of film success in the United Kingdom. *Managerial Decision Econom.* **23**(6) 343–354.

- Connolly, M., D. Krueger. 2006. Rockonomics: The economics of popular music. V. Ginsburgh, D. Thorsby, eds. *Handbook of the Economics of Art and Culture*. North-Holland, Amsterdam, 667–719.
- Dansby, R. E., C. Conrad. 1984. Commodity bundling. *Amer. Econom. Rev.* **74**(2) 377–381.
- De Vany, A. 2004. *Hollywood Economics: How Extreme Uncertainty Shapes the Film Industry*. Routledge, London.
- De Vany, A., D. Walls. 2002. Does Hollywood make too many R-rated movies? Risk, stochastic dominance, and the illusion of expectation. *J. Bus.* **75**(3) 425–51.
- De Vany, A., D. Walls. 2004. Motion picture profit, the stable Paretian hypothesis, and the curse of the superstar. *J. Econom. Dynam. Control* **28**(6) 1035–1057.
- Eliashberg, J., A. Elberse, M. A. A. M. Leenders. 2006. The motion picture industry: Critical issues in practice, current research, and new research directions. *Marketing Sci.* **25**(6) 638–661.
- Embrechts, P., C. Klüppelberg, T. Mikosch. 1997. *Modelling Extremal Events for Insurance and Finance*. Springer, New York.
- Fama, E. 1965. The behavior of stock market prices. *J. Bus.* **38**(1) 34–105.
- Fang, H., P. Norman. 2006. To bundle or not to bundle. *RAND J. Econom.* **37**(4) 946–963.
- Fang, K.-T., S. Kotz, K. W. Ng. 1990. *Symmetric Multivariate and Related Distributions*. Chapman and Hall, New York.
- Frank, R. H., P. J. Cook. 1995. *The Winner-Take-All Society*. Free Press, New York.
- Gabaix, X. 2009. Power laws in economics and finance. *Annual Rev. Econom.* **1** 255–293.
- Gabaix, X., P. Gopikrishnan, V. Plerou, H. E. Stanley. 2006. Institutional investors and stock market volatility. *Quart. J. Econom.* **121**(2) 461–504.
- Gaffeo, E., A. E. Scorcu, L. Vici. 2008. Demand distribution dynamics in creative industries: The market for books in Italy. *Inform. Econom. Policy* **20** 257–268.
- Geng, X., M. B. Stinchcombe, A. B. Whinston. 2005. Bundling information goods of decreasing value. *Management Sci.* **51**(4) 662–667.
- Ghose, A., A. Sundararajan. 2006. Evaluating pricing strategy using e-commerce data: Evidence and estimation challenges. *Statist. Sci.* **21**(2) 131–142.
- Glasserman, P., P. Heidelberger, P. Shahabuddin. 2002. Portfolio value-at-risk with heavy-tailed risk factors. *Math. Finance* **12**(3) 239–269.
- Hitt, L. M., P.-y. Chen. 2005. Bundling with customer self-selection: A simple approach to bundling low-marginal-cost goods. *Management Sci.* **51**(10) 1481–1493.
- Ibragimov, R. 2005. New majorization theory in economics and martingale convergence results in econometrics. Ph.D. dissertation, Yale University, New Haven, CT.
- Ibragimov, R. 2007. Efficiency of linear estimators under heavy-tailedness: Convolutions of α -symmetric distributions. *Econometric Theory* **23**(3) 501–517.
- Ibragimov, R. 2009. Heavy tailed densities. S. N. Durlauf, L. E. Blume, eds. *The New Palgrave Dictionary of Economics Online*. Palgrave Macmillan, New York. http://www.dictionarofeconomics.com/article?id=pde2009_H000191.
- Ibragimov, R., J. Walden. 2007. The limits of diversification when losses may be large. *J. Banking Finance* **31**(8) 2551–2569.
- Ibragimov, R., D. Jaffee, J. Walden. 2009. Nondiversification traps in catastrophe insurance markets. *Rev. Financial Stud.* **22**(3) 959–993.
- Kalyanaram, G., W. T. Robinson, G. L. Urban. 1995. Order of market entry: Established empirical generalizations, emerging empirical generalizations, and future research. *Marketing Sci.* **14**(3) G212–G221.
- Kohli, R., R. Sah. 2006. Some empirical regularities in market shares. *Management Sci.* **52**(11) 1792–1792.
- Krishna, V. 2002. *Auction Theory*. Academic Press, San Diego.
- Lewbel, A. 1985. Bundling of substitutes or complements. *Internat. J. Indust. Organ.* **3**(1) 101–107.
- Mandelbrot, B. 1963. The variation of certain speculative prices. *J. Bus.* **36**(4) 394–419.
- Mandelbrot, B. 1997. *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk*. Springer-Verlag, New York.
- Marshall, A. W., I. Olkin. 1979. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.
- McAfee, R., J. McMillan, M. Whinston. 1989. Multiproduct monopoly, commodity bundling and correlation of values. *Quart. J. Econom.* **114**(2) 371–384.
- Milgrom, P., R. Weber. 1982. A theory of auctions and competitive bidding. *Econometrica* **50**(5) 1089–1122.
- Palfrey, T. R. 1983. Bundling decisions by a multiproduct monopolist with incomplete information. *Econometrica* **51**(2) 463–483.
- Perline, R. 2005. Strong, weak and false inverse power laws. *Statist. Sci.* **20**(1) 68–88.
- Praetz, P. 1972. The distribution of share price changes. *J. Bus.* **45**(1) 49–55.
- Proschan, F. 1965. Peakedness of distributions of convex combinations. *Ann. Math. Statist.* **36**(6) 1703–1706.
- Rachev, S. T., C. Menn, F. J. Fabozzi. 2005. *Fat-Tailed and Skewed Asset Return Distributions: Implications for Risk Management, Portfolio Selection, and Option Pricing*. Wiley, Hoboken, NJ.
- Salinger, M. A. 1995. A graphical analysis of bundling. *J. Bus.* **68**(1) 85–98.
- Schmalensee, R. L. 1984. Gaussian demand and commodity bundling. *J. Bus.* **57**(1) S211–S230.
- Shapiro, C., H. R. Varian. 1999. *Information Rules: A Strategic Guide to the Network Economy*. Harvard Business School Press, Boston.
- Silverberg, G., B. Verspagen. 2007. The size distribution of innovations revisited: An application of extreme value statistics to citation and value measures of patent significance. *J. Econometrics* **139**(2) 318–339.
- Stremersch, S., G. J. Tellis. 2002. Strategic bundling of products and prices: A new synthesis for marketing. *J. Marketing* **66**(1) 55–72.
- Taleb, N. N. 2007. *The Black Swan: The Impact of Highly Improbable*. Random House, New York.
- Uchaikin, V. V., V. M. Zolotarev. 1999. *Chance and Stability: Stable Distributions and Their Applications*. VSP, Utrecht, The Netherlands.
- Venkatesh, R., W. Kamakura. 2003. Optimal bundling and pricing under a monopoly: Contrasting complements and substitutes from independently valued products. *J. Bus.* **76**(2) 211–231.
- Wu, S.-y., L. M. Hitt, P.-y. Chen, G. Anandalingam. 2008. Customized bundle pricing for information goods: A nonlinear mixed-integer programming approach. *Management Sci.* **54**(3) 608–622.
- Zolotarev, V. M. 1986. *One-Dimensional Stable Distributions*. American Mathematical Society, Providence, RI.