

To bundle or not to bundle

Hanming Fang*

and

Peter Norman**

Comparing monopoly bundling with separate sales is relatively straightforward in an environment with a large number of goods. We show that results similar to those for the asymptotic case can be obtained in the more realistic case with a given finite number of goods, provided that the distributions of valuations are symmetric and log-concave.

When I go to the grocery store to buy a quart of milk, I don't have to buy a package of celery and a bunch of broccoli. . . . I don't like broccoli. (U.S. Senator John McCain, in an interview on cable TV rates published in the *Washington Post*, C1, March 24, 2004)

1. Introduction

■ Bundling, the practice of selling two or more products as a package deal, is a common phenomenon in markets where sellers have market power. It is sometimes possible to rationalize bundling by complementarities in technologies or in preferences. However, it has long been understood that bundling may be a profitable device for price discrimination, even when the willingness to pay for one good is unaffected by whether or not other goods in the bundle are consumed, and when no costs are saved through bundling (Adams and Yellen, 1976; Schmalensee, 1982). While the earliest literature of bundling typically understood it as a way to exploit negative correlation between valuations for different goods, McAfee, McMillan, and Whinston (1989) show that mixed bundling, which refers to a selling strategy where each good can be purchased either as a separate good or as part of a bundle, leads to a strict increase in profits relative to fully separate sales, provided that a condition on the joint distribution of valuations is satisfied. Importantly, the distributional condition holds generically and is implied by stochastic independence, so the profit-improving role of mixed bundling has nothing to do with exploiting negative correlations of valuation distributions.

In this article we rule out mixed bundling by assumption and focus on the comparison of

* Duke University, SHUFE and CEMA; hanming.fang@duke.edu.

** University of North Carolina, Chapel Hill; normanp@email.unc.edu.

The first version of this article unbundled some material that was previously reported in Fang and Norman (2004). We are especially grateful to Ted Bergstrom and Larry Samuelson, who suggested an idea that inspired the example in Section 5. We also thank Editor Joseph Harrington, two anonymous referees, Mark Armstrong, Martin Hellwig, Rustam Ibragimov, Stephen Morris, Bill Sandholm, Ananth Seshadri, seminar participants at New York University, Ohio State University, University of Essex, University of Illinois, University of Pittsburgh, University of Wisconsin, Yale, SITE 2003 Summer Workshop at Stanford, and the SED 2003 meetings in Paris for comments and helpful discussions. The usual disclaimer applies.

pure bundling and separate sales. By pure bundling, we refer to the case that any good is sold either as an item in a larger bundle or as a separate item, but not both. Of course, mixed bundling does occur in the real world. For example, in many markets it is possible to buy access to cable TV at one price, high-speed Internet access at one price, and a bundle consisting of both cable TV and high-speed Internet access at a price that is lower than the sum of the component prices.¹ We offer three reasons for our focus on pure bundling.

First, McAfee, McMillan, and Whinston (1989) showed that, generically, any multiproduct monopolist should offer to sell all of the goods in mixed bundles. This powerful result does make some of the crude bundling schemes that we observe in the real world rather puzzling. For example, the question of why ESPN is available as a component of a bundle while championship boxing matches tend to be available only on a pay-per-view basis cannot be answered.

Second, in some cases technological reasons may make mixed bundling infeasible or too costly to implement. For example, in the context of bundling computer programs it does not seem farfetched to assume that selling components separately would require substantial extra programming costs in order to guarantee compatibility of the components with older softwares, costs that could be avoided if the new programs are bundled.

Third, in some cases the practice of mixed bundling is more likely than pure bundling to get in trouble with antitrust laws, which is explicitly expressed in terms of “anticompetitive mixed bundling.”² Of course in general, the legal interpretation of “mixed” is unclear. But in a recent case in the United Kingdom, the decision by the Office of Fair Trading (2002) on the alleged anticompetitive mixed bundling by the British Sky Broadcasting Limited explicitly stated that “[m]ixed bundling refers to a situation where two or more products are offered together at a price less than the sum of the individual product prices—i.e., there are discounts for the purchase of additional products.” This test, which compares marginal prices, requires that a product can be bought both as a bundle and as a separate good. Thus it has no bite at all when the monopolist uses pure bundling.³

In this article we obtain a rather intuitive characterization for when a multiproduct monopolist should bundle and when it should sell the goods separately in order to maximize its profits. To some extent our characterization confirms (mainly) numerical results in Schmalensee (1984), namely, the higher the marginal cost and the lower the mean valuation, the less likely that bundling dominates separate sales. When limiting our comparison to pure bundling and separate sales, we are able to highlight a clear intuition for what happens when two or more goods are sold as a bundle. The key effect driving all the results is that the variance in the relevant willingness to pay is reduced when goods are bundled. We shall provide a partial characterization for when this reduction in variance is beneficial for the monopolist and when it is not.

The crucial idea is that bundling makes the tails of the distribution of willingness to pay thinner. However, what we need is a rather strong notion of what “thinner tails” mean. Specifically, we need to be able to conclude that for a given per-good price below (respectively, above) the mean, bundling increases (respectively, reduces) the probability of trade. This can be rephrased by saying that the average valuation is more peaked than the underlying distributions. Notice that the law of large numbers can be used to reach this conclusion if there are sufficiently many goods available, but for a given finite number of goods, counterexamples are easy to construct. We therefore need to make some distributional assumptions. Indeed, assuming that valuations

¹ However, Crawford (2004), in his study of bundling in cable television, finds that mixed bundling of channels is quite uncommon in that industry.

² Our model is not suited to study “anticompetitiveness” to the extent that it refers to preemptive pricing strategies to limit entry of competitors, because in our model there is no threat of entry for the monopolist (see Nalebuff (2004) for a more suitable model). However, consumer advocates arguing for introducing “à la carte” pricing for cable TV stations are explicitly concerned about how bundling improves the possibilities for surplus extraction.

³ The Microsoft case is a counterexample—the failure to provide the browser separately was used as evidence of anticompetitive behavior. However, had the Windows operating system and the Internet browser been two new products rather than upgrades of existing products (with a history of being thought of as different programs), it would seem difficult to make an argument for unbundling.

are distributed in accordance to symmetric and log-concave densities, we can use a result from Proschan (1965) to unambiguously rank distributions in terms of relative peakedness.

Under these distributional assumptions, bundling reduces the effective dispersion in the buyers' valuations. This reduction of valuation dispersion is to the advantage of the monopolist when a good should be sold with high probability (either because costs are low or because valuations tend to be high). In such cases, we show that bundling increases the monopolist's profits. The reduction of taste dispersion may be to the disadvantage of the monopolist when the goods have only a thin market (either because the costs are high or because valuations tend to be low). Indeed in such cases the monopolist is better off relying on the right tail of the distribution and selling all goods separately.

The idea that "bundling reduces dispersion" has been around for a long time, and there is even some emerging empirical evidence supporting it as a motivation to bundle (see Crawford, 2004). What is largely missing in the literature, however, are results that establish reasonably general conditions to explain bundling as a profit-maximizing selling strategy. The most related article is Schmalensee (1984), who considers the case with normally distributed distributions of valuations (which belongs to the class we consider). Relying mainly on numerical methods, he reaches a similar conclusion. Recently, Ibragimov (2005) has developed a related characterization relying on a generalization of the result in Proschan (1965).

In the context of "information goods" (goods with zero marginal costs), Bakos and Brynjolfsson (1999) and, more recently, Geng, Stinchcombe, and Whinston (2005), used a similar idea to argue that bundling is better than separate sales. While both sets of authors assume zero marginal costs, the main difference with our article is that they focus on results for large numbers of goods. Though we also prove some asymptotic results, our main contribution is to provide conditions under which we can obtain analogous results for the finite-good case.

The remainder of the article is structured as follows. Section 2 presents the model. Section 3 introduces the statistics notion of peakedness. Section 4 presents the asymptotic results in an environment with a large number of goods. Section 5 provides our main analysis for the finite-good case. Finally, Section 6 concludes. All proofs are relegated to Appendix A.

2. The model

■ The underlying economic environment is the same as in McAfee, McMillan, and Whinston (1989), except that we allow for more than two goods. A profit-maximizing monopolist sells K indivisible products indexed by $j = 1, \dots, K$, and good j is produced at a constant unit cost c_j . A representative consumer is interested in buying at most one unit of each good and is characterized by a vector of valuations $\theta = (\theta_1, \dots, \theta_K)$, where θ_j is interpreted as the consumer's valuation of good j . The vector θ is private information to the consumer, and the utility of the consumer is given by

$$\sum_{j=1}^K I_j \theta_j - p,$$

where p is the transfer from the consumer to the seller and I_j is a dummy taking on value one if good j is consumed and zero otherwise. Valuations are assumed stochastically independent, and we let F_j denote the marginal distribution of θ_j . Hence, $\prod_{j=1}^K F_j(\theta_j)$ is the cumulative distribution of θ .

3. Peakedness of convolutions of log-concave densities

■ A rough interpretation of the law of large numbers is that the distribution of the average of a random sample gets more and more concentrated around the population mean as the sample size grows. However, the law of large numbers does not imply that the probability of a given size deviation from the mean is monotonically decreasing in the sample size. In general, no such monotone convergence can be guaranteed.

To discuss such monotonicity, a notion of “relative peakedness” of two distributions is needed. We use a definition from Birnbaum (1948):

Definition 1. Let X_1 and X_2 be real random variables. Then X_1 is said to be more peaked than X_2 if

$$\Pr[|X_1 - E(X_1)| \geq t] \leq \Pr[|X_2 - E(X_2)| \geq t]$$

for all $t \geq 0$. If the inequality is strict for all $t > 0$, we say that X_1 is strictly more peaked than X_2 .⁴

A random variable is said to be log-concave if the logarithm of the probability density function is concave. This is a rather broad set of distributions that includes the uniform, normal, logistic, extreme value, exponential, Laplace, Weibull, and many other common parametric densities (see Bagnoli and Bergstrom (2005) for further examples). Proschan (1965) studies comparative peakedness of convex combinations of log-concave random variables, and we will apply one of his results here. To avoid discussing majorization theory, we will use his key lemma directly rather than his main result.

Theorem 1 (See Proschan, 1965). Suppose that X_1, \dots, X_m are independently and identically distributed random variables with a symmetric log-concave density f . Fix $(w_3, \dots, w_m) \geq 0$ with $\sum_{i=3}^m w_i < 1$. Then the random variable

$$w_1 X_1 + (1 - w_1 - \sum_{i=3}^m w_i) X_2 + \sum_{i=3}^m w_i X_i$$

is strictly more peaked as w_1 increases from zero to $(1 - \sum_{i=3}^m w_i)/2$.

A corollary of this result is that $\sum_{i=1}^m X_i/m$ is strictly more peaked as m increases.⁵ That is, the probability of a given size deviation from the population average is indeed monotonically decreasing in sample size for the class of symmetric log-concave distributions. It is rather easy to construct discrete examples to verify that unimodality (which is implied by log-concavity) is necessary for Theorem 1. However, unimodality is not sufficient. An example that clarifies the role of log-concavity is considered in Section 5. The role of the symmetry assumption is simply to avoid the location of the peak to depend on the weights.

4. To bundle or not to bundle many goods

■ Since Theorem 1 may be viewed as a result establishing monotone convergence to a law of large numbers, it is useful to first consider the implications of bundling a large number of goods. This analysis is not particularly innovative, and it is only meant to establish a benchmark for the results in Section 5. The basic ideas are similar to Armstrong (1999) and Bakos and Brynjolfsson (1999); also, a careful analysis of a more general specification of consumer preferences (that allows the valuation for the good to decline in the number of goods consumed) can be found in Geng, Stinchcombe, and Whinston (2005).

Let $\{j\}_{j=1}^\infty$ be a sequence of goods, where each good j can be produced at a constant marginal cost c_j . In the absence of the bundling instrument, the maximized profit from sales of good j is thus given by

$$\Pi_j = \max_{p_j} (p_j - c_j)[1 - F_j(p_j)]. \quad (1)$$

⁴ Strictly speaking, Birnbaum (1948) uses a local definition of peakedness where the expectations are replaced with arbitrary points in the support. For our purposes, only “peakedness around the mean” is relevant, so we follow Proschan (1965) and drop the qualifiers.

⁵ To see this, we can first use Theorem 1 to conclude that weights $\mathbf{w}_1 = (1/m, \dots, 1/m)$ result in a more peaked distribution than from $\mathbf{w}_2 = ((m-2)/[m(m-1)], 1/(m-1), 1/m, \dots, 1/m)$. By the same token, $\mathbf{w}_3 = ((m-3)/[m(m-1)], 1/(m-1), 1/(m-1), 1/m, \dots, 1/m)$ is less peaked than \mathbf{w}_2 . Continuing recursively all the way up to $\mathbf{w}_m = (0, 1/(m-1), \dots, 1/(m-1))$, we have a sequence of m random variables with decreasing peakedness.

Assume instead that the monopolist has monopoly rights to the first K goods in the sequence and sells them as a bundle. That is, the monopolist posts a single price p and the consumer must choose between purchasing all the goods at price p or none at all. Assuming that there is a uniform upper bound σ^2 such that $\text{Var}\theta_j \leq \sigma^2$ for every j , we know from Chebyshev's inequality that for each $\varepsilon > 0$, there exists $K(\varepsilon) < \infty$ such that, for any $K > K(\varepsilon)$,

$$\Pr \left[\left| \sum_{j=1}^K \theta_j - \sum_{j=1}^K E\theta_j \right| \leq \varepsilon K \right] \geq 1 - \varepsilon. \quad (2)$$

Inequality (2) implies that

- (i) $\Pr[\sum_{j=1}^K \theta_j \geq p] \geq 1 - \varepsilon$ if $p \leq \sum_{j=1}^K E\theta_j - \varepsilon K$;
- (ii) $\Pr[\sum_{j=1}^K \theta_j \geq p] \leq \varepsilon$ if $p \geq \sum_{j=1}^K E\theta_j + \varepsilon K$.

In words, charging a price that on a per-good basis is just slightly below the average expected valuation ensures that the bundle will be sold almost surely. On the other hand, a price that exceeds the average expected valuation ever so slightly implies that almost all types will decide not to buy. This observation can be used to provide a simple sufficient condition for when separate sales dominates bundling in the case of many goods.

Proposition 1. Suppose that for every j , $\text{Var}\theta_j \leq \sigma^2$ and $\Pi_j \geq \underline{\Pi}$ for some finite σ^2 and $\underline{\Pi} > 0$ (where Π_j is defined in (1)). Also, suppose that $\sum_{j=1}^K E\theta_j \leq \sum_{j=1}^K c_j$ for every K . Then, selling all goods separately is better than selling all goods as a single bundle whenever $K > \sigma^2/\underline{\Pi}^2$.

The condition says that if the profit from separate sales is not negligible for every good (which is the condition $\Pi_j \geq \underline{\Pi}$) and if costs exceed the sum of the expected valuations (which is the condition $\sum_{j=1}^K E\theta_j \leq \sum_{j=1}^K c_j$), then the monopolist is better off selling the goods separately. The idea is that if goods are bundled, the monopolist must charge a price above the sum of the costs in order to make a profit. But if many goods are bundled, such price inevitably leads to negligible sales. An alternative interpretation of this result is that, under the conditions of the proposition, $\sigma^2/\underline{\Pi}^2$ is an upper bound on the number of goods in each bundle for a profit-maximizing monopolist.

Except for the restriction that the maximized profit from selling a single good is uniformly bounded away from zero for all j , Proposition 1 is expressed in terms of exogenous parameters. Unfortunately, such a "clean" characterization is impossible for the case when $\sum_{j=1}^K E\theta_j > \sum_{j=1}^K c_j$. The reason is that although the profits under bundling can be bound tightly from above and below, the profits from separate sales depend crucially on the shape of the distribution of valuations. Therefore any reasonably general condition for when bundling dominates asymptotically must be expressed in the (endogenous) nonbundling profits.

Proposition 2. Suppose that for every j , $E\theta_j \leq \mu$ and $\text{Var}\theta_j \leq \sigma^2$ for some finite μ and σ^2 . Also, suppose that there exists $\delta > 0$ such that

$$0 \leq \sum_{j=1}^K \Pi_j \leq \sum_{j=1}^K E\theta_j - \sum_{j=1}^K c_j - \delta K \quad (3)$$

for every K (where Π_j is defined in (1)). Then there exists K^* such that selling all goods as a single bundle is better than separate sales for every $K \geq K^*$.

The proposition is an immediate consequence of the fact that the bundling profit can be made close to $\sum_j E\theta_j - \sum_j c_j$, but a proof is in the Appendix for completeness. For comparison with the results in Section 5, it is useful to observe that a sufficient (but not necessary) condition for (3) is that if p_j^* solves (1), then $p_j^* < E\theta_j$ for every j .

It is also useful to remark that the uniform bound on the expected valuation is needed to rule out examples of the following nature: assume that θ_j is uniformly distributed on $[j - 1, j + 1]$ and $c_j = 0$ for each j . Condition (3) is satisfied for every K , since the optimal monopoly price for good j is $p_j = j - 1$ for each $j \geq 3$. However, the profit *per good* explodes as K tends to infinity, implying that even a negligible probability of the consumer's rejecting the bundle could be more important than the increase in profit conditional on selling the bundle.

5. To bundle or not to bundle in the finite case

■ **An example.** To demonstrate how small numbers in general can overturn the intuition from the asymptotic results, we consider an example with two goods, $j = 1, 2$, each produced at zero marginal cost. Assume that the valuation for each good j is distributed in accordance with cumulative density F over $[0, 2]$ defined as

$$F(\theta_j) = \begin{cases} \frac{\alpha}{2}\theta_j & \text{for } \theta_j < 1 \\ (1 - \alpha) + \frac{\alpha}{2}\theta_j & \text{for } \theta_j \geq 1. \end{cases} \quad (4)$$

This cumulative distribution can be thought of as the result of drawing θ_j from a uniform $[0, 2]$ distribution with probability α and setting $\theta_j = 1$ with probability $1 - \alpha$. In the case of separate sales, we first note that if $\alpha = 1$, then the optimal price is clearly to set $p_j = 1$ for $j = 1, 2$. But for $\alpha < 1$, $p_j = 1$ continues to be the optimal price, since the probability mass is moved to valuation 1 without changing the distribution of θ_j conditional on $\theta_j \neq 1$. Hence, the maximized profit in the case of separate sales is given by

$$\Pi_1 = \Pi_2 = \frac{\alpha}{2} + (1 - \alpha) = 1 - \frac{\alpha}{2},$$

and the total profits from separate sales $\Pi_1 + \Pi_2 = 2 - \alpha$. Next, consider the case with the two goods being bundled. The optimal price for the bundle is then the solution to

$$\max_p p \Pr[\theta_1 + \theta_2 \geq p] = \max_x 2x \Pr\left[\frac{\theta_1 + \theta_2}{2} \geq x\right],$$

where the change in variable allows us to transform the question as to whether separate sales or bundling is better into a comparative-statics exercise with respect to the cumulative distribution of valuations.

Denote by G_B the cumulative density of the average valuation $(\theta_1 + \theta_2)/2$. Clearly, G_B has mean one and a smaller variance than F , but—and this is the crucial feature of the example— G_B is not unambiguously more peaked than F . This follows immediately from the fact that the probability that $(\theta_1 + \theta_2)/2$ is exactly equal to one is $(1 - \alpha)^2$, whereas the probability that θ_j is exactly equal to one is $1 - \alpha$. It follows that there exists a range $[0, t^*]$ where

$$\Pr\left[\frac{\theta_1 + \theta_2}{2} - 1 \leq -t\right] = G_B(1 - t) > F(1 - t) = \Pr[\theta_j - 1 \geq -t]$$

for $t \in [0, t^*]$. Hence, G_B and F cannot be compared in terms of relative peakedness. The implication of this for the comparison between bundling and separate sales is that the construction that worked in the asymptotic case—pricing the bundle just below the expected value—will reduce rather than increase sales. However, this does not prove that bundling is worse, since (i) a price slightly above the expectation leads to higher sales under bundling than under separate sales, and (ii) a sufficiently large reduction in price from the expected value also leads to higher sales under bundling than under separate sales.

To obtain the explicit comparison of profits under bundling and separate sales, let p_B denote the profit-maximizing price for the bundled good and let Π_B be the associated profit. There are three possibilities:

Case 1. $p_B = 2$. By symmetry of G_B , it follows that

$$\Pr\left[\frac{\theta_1 + \theta_2}{2} < 1\right] = \frac{1 - \Pr\left[\frac{\theta_1 + \theta_2}{2}\right]}{2} = \frac{1 - (1 - \alpha)^2}{2}.$$

Hence, the probability of selling the bundle is

$$(1 - \alpha)^2 + \frac{1 - (1 - \alpha)^2}{2} = \frac{\alpha^2}{2} + (1 - \alpha).$$

The profit is thus given by $\Pi_B = 2 - \alpha - \alpha(1 - \alpha) < 2 - \alpha = \Pi_1 + \Pi_2$. Hence, if $p_B = 2$ is the best price for the bundled good, the monopolist is strictly better off selling the goods separately.

Case 2. $p_B < 2$. If α is close to one, then this will indeed lead to an increase in profits. However, one can show that if α is sufficiently small, any price strictly below two will generate lower profits than the maximized profits under separate sales.⁶ The idea is as follows. To be able to make a larger profit than $\Pi_1 + \Pi_2 = 2 - \alpha$, it is necessary to sell at a price $p_B > 2 - \alpha$. The smaller is α , the closer to two this price is; and for α small, such a price is in the range where $G_B((2 - \alpha)/2) > F((2 - \alpha)/2)$. But this implies that any price for the bundled goods in the interval $(2 - \alpha, 2)$ is worse than selling the goods separately at price $(2 - \alpha)/2$ each.

Case 3. $p_B > 2$. As $\alpha \rightarrow 0$, the probability of selling the bundle at such a price $p_B > 2$ goes to zero, so for α sufficiently small this can be ruled out as well.

Summing up, we have an example (when α is small) where if the monopolist had access to a large number of goods with valuations being independently and identically distributed in accordance with the distribution (4), it would be possible to almost fully extract the surplus from the consumer by selling all goods as a single bundle. Nevertheless, with only two goods, separate sales does better than bundling.

Easier examples can be constructed, but (4) has been chosen for a reason. Standard continuity arguments can be used to extend the example to the case where θ_j is distributed uniformly on $[0, 2]$ with probability α and distributed with, say, a normal distribution with mean 1 and variance σ^2 with probability $1 - \alpha$. If σ^2 and α are both sufficiently small, separate sales dominate bundling. Notice that this is so despite the fact that the distribution is symmetric, unimodal, smooth, and generated as a mixture of two (different) log-concave densities with identical means. This may seem inconsistent with Propositions 4 and 5 below, but mixtures of log-concave densities are not necessarily log-concave (see An, 1998).

□ **Bundling with symmetric log-concave densities.** We now assume that each θ_j is independently and identically distributed according to a symmetric log-concave probability density f with expectation $\tilde{\theta} > 0$. Any form of mixed bundling is ruled out by assumption. The problem for the monopolist can therefore be separated in two parts:

- (i) Decide how to package the goods into different bundles. Because we rule out by default mixed bundling, this packaging decision is the same as partitioning the set of goods in what we will refer to as a bundling menu. Following Palfrey (1983), we denote such a bundling menu by $\mathcal{B} = \{B_1, \dots, B_M\}$, where each $B_i \in \mathcal{B}$ is a subset of $\{1, \dots, m\}$ and where $B_i \cap B_{i'} = \emptyset$ for each $i \neq i'$, and M is the number of bundles sold by the monopolist. The menu $\{\{1\}, \{2\}, \dots, \{K\}\}$ corresponds with separate sales, and the

⁶ The details of these calculations are available upon request from the authors.

menu $\{\{1, \dots, K\}\}$ describes the other extreme case where all goods are sold as a single bundle.

- (ii) For each bundle, construct the optimal pricing rule. This is a single-dimensional problem (since the consumer either gets the bundle or not, any two types θ and θ' with $\sum_{j \in B_i} \theta_j = \sum_{j \in B_i} \theta'_j$ must be treated symmetrically). By standard results (see Myerson, 1981; Riley and Zeckhauser, 1983), there is therefore no further loss of generality in restricting the monopolist to fixed-price mechanisms for each bundle.

We are now in a position to prove an analogue of Proposition 1 that is valid also in the finite case.

Proposition 3. Suppose that each θ_j is independently and identically distributed according to a symmetric log-concave density f that is strictly positive on support $[\underline{\theta}, \bar{\theta}]$ and has expectation $\bar{\theta}$. Assume that each good j is produced at unit cost c_j , where $c_j < \bar{\theta}$. Let \mathcal{B}^* be the optimal bundling menu for the monopolist. Then there exists no $B_i \in \mathcal{B}^*$ with more than a single good such that $\sum_{j \in B_i} E\theta_j \leq \sum_{j \in B_i} c_j$.

While the assumptions are obviously much more restrictive, Proposition 3 provides a close analogue to Proposition 1. It is worth noting that the “nontriviality assumption” $c_j < \bar{\theta}$ in Proposition 3 is analogous to the condition that $\Pi_j > \underline{\Pi}$ in Proposition 1. Thus the only difference between Propositions 1 and 3 is whether separate sales is compared with a large bundle or a bundle of any finite size.

The link between the large-numbers analysis and the finite case is somewhat weaker in the case with unit costs below the expected value. The result is as follows.

Proposition 4. Suppose that each θ_j is independently and identically distributed according to a symmetric log-concave density f that is strictly positive on support $[\underline{\theta}, \bar{\theta}]$ and has expectation $\bar{\theta}$. Furthermore, assume that the unit cost is $c_j = c$ for each good j . Let the (unique) profit-maximizing price in the case of separate sales be given by p^* and the (unique) profit-maximizing price when all goods are sold as a single bundle be p_B^* . Then

- (i) if $p^* \leq \tilde{\theta}$, it is profit maximizing to sell all goods as a single bundle, and
(ii) if $p_B^* \geq K\tilde{\theta}$, it is profit maximizing to sell all goods separately.

Proof. See the Appendix.

Even though the conditions for Proposition 4 are stated in terms of the endogenous prices, it is possible to find sufficient conditions on primitives for $p^* \leq \tilde{\theta}$ and $p_B^* \geq K\tilde{\theta}$. In fact, maintaining the symmetric log-concavity assumption on the density f , the necessary and sufficient condition on the primitives for $p^* \leq \tilde{\theta}$ is $f(\tilde{\theta}) \geq 1/[2(\tilde{\theta} - c)]$. To see this, note that the profit from selling a single good at price p is given by

$$(p - c)[1 - F(p)],$$

which can be shown to be a single-peaked function of p when f is log-concave.⁷ However,

$$\left. \frac{d}{dp} (p - c)[1 - F(p)] \right|_{p=\tilde{\theta}} = 1 - F(\tilde{\theta}) - (\tilde{\theta} - c)f(\tilde{\theta}) = \frac{1}{2} - (\tilde{\theta} - c)f(\tilde{\theta}),$$

which is nonpositive if and only if $f(\tilde{\theta}) \geq 1/[2(\tilde{\theta} - c)]$. Thus, the profit-maximizing single-good price p^* is no larger than $\tilde{\theta}$ if and only if $f(\tilde{\theta}) \geq 1/[2(\tilde{\theta} - c)]$.

It is also easy to see that a sufficient condition for $p_B^* \geq K\tilde{\theta}$ is $c_j \in [\tilde{\theta}, \bar{\theta}]$ for all j . To

⁷ To see this, note that the profit function is increasing in p whenever $c - [p + [(1 - F(p))/f(p)]] > 0$ and decreasing when the inequality is reversed. Since log-concavity implies that $p + [1 - F(p)]/f(p)$ is strictly increasing, we conclude that the profit function is strictly single-peaked.

see this, note that the profit from selling all goods as a single bundle at price p_B is given by $(p_B - \sum_{j=1}^K c_j)[1 - G_B(p_B/K)]$, where G_B is the CDF of $\theta^B = \sum_{j=1}^K \theta_j/K$. Obviously, the optimal price for the whole bundle B must satisfy $p_B^* \geq \sum_{j=1}^K c_j \geq K\bar{\theta}$, because otherwise the profit would be negative.⁸ However, in many cases $p_B^* \geq K\bar{\theta}$ holds even when $c_j < \bar{\theta}$ (see the next subsection for numerical examples).

Propositions 3 and 4, together with the comparative-statics properties of monopoly pricing, also have some natural implications as to which types of goods we should expect to see bundled. For example, the optimal monopoly price for a single good is increasing in its unit cost of production; thus the condition $p^* \leq \bar{\theta}$ is less likely to be satisfied as c increases. As a result, Propositions 3 and 4 imply that a monopolist is less likely to provide more costly goods in bundles. Similarly, shifting the distribution of θ_j to the right, or replacing F with a (log-concave symmetric distribution) F' with the same mode that is strictly less peaked than F , also leads to an increase in the optimal monopoly price. Thus, such changes will lead bundling to be less profitable than separate sales.

It is also worth commenting that Proposition 4 only considers the case where the unit costs are identical. The reason for this is simple: when unit costs of production vary across goods, bundling has a distinct disadvantage relative to separate sales in that it does have the flexibility in terms of adjusting the price for a particular good to its production cost. That is, bundling two goods with different unit costs has negative consequences for productive efficiency. Of course this disadvantage is also present in the asymptotic analysis, but there the monopolist can extract almost the full surplus leading to the relatively clean condition (3) that applies even for heterogeneous-cost goods. In the finite-good case, while bundling does increase revenue when the average price is to the left of the mode of the distribution, if costs are different, the change in profit depends on a nontrivial tradeoff between the increase in revenue and the loss in productive efficiency.

We would like to note that the example in the opening of this section satisfies all conditions in the statement except log-concavity; thus log-concavity cannot be dropped from the statement of the result of Proposition 4.⁹

Finally, we would like to point out that even for the case where all goods being sold are produced at the same unit cost, the characterization in Proposition 4 is incomplete because it is quite possible that $p^* > \bar{\theta}$ and $p_B^* < K\bar{\theta}$, in which case Proposition 4 is silent on whether bundling or separate sales maximizes the monopolist's profits. Proposition 5 below shows that in such cases the optimal bundling strategy must be either full bundling or separate sales, even though we do not have a complete characterization of which is better. In the next subsection we report results from numerical analysis for Gaussian demand to further shed light on these cases.

Proposition 5. Suppose that each θ_j is independently and identically distributed according to a symmetric log-concave density f that is strictly positive on support $[\underline{\theta}, \bar{\theta}]$ and has expectation $\bar{\theta}$. Furthermore, assume that the unit cost is given by $c_j = c$ for each good j . Then either full bundling or separate sales is profit maximizing.

An interesting corollary of Proposition 5 is as follows. Suppose that a monopolist with k goods (whose valuations are independently and identically distributed according to a symmetric log-concave density f and unit costs are the same) finds it profit maximizing to fully bundle the k goods. Then a $(k + 1)$ -good monopolist, where the additional good has the same valuation distribution f and unit cost, will also find it optimal to bundle all $k + 1$ goods. This comparative-statics result is *a priori* not obvious, but it follows straightforwardly from Proposition 5. The reason is simple: Proposition 5 establishes that the $(k + 1)$ -good monopolist only needs to compare the profits from the full bundling and separate sales. Suppose to the contrary that separate sales were optimal for the $(k + 1)$ -good monopolist. But this must imply that separate sales would have been

⁸ In fact, one can show a similar result when the constant marginal costs c_j are not equal across the goods: if $\bar{\theta} \leq c_j < \bar{\theta}$ for all j , then all goods should be sold separately (see Fang and Norman (2004) for details).

⁹ Log-concavity of the valuation distribution is a sufficient condition to rule out "too abrupt" changes in the density, which was the culprit for the results in the example above.

optimal for the k -good monopolist as well, a contradiction. To see this, if a $(k+1)$ -good monopolist finds separate sales optimal, it means that $(k+1)\Pi$, where Π is the monopoly profit from selling a single good separately, is higher than the profits under all other bundling menus, including a bundling menu that consists of a k -good bundle and a single good. This alternative bundling menu will generate, at their respective optimal prices, a profit that equals to $\Pi_{B(k)} + \Pi$, where $\Pi_{B(k)}$ is the monopoly profit from the k -good bundle under its optimal price. But this implies that $k\Pi$ is higher than $\Pi_{B(k)}$, i.e., a k -good monopolist will receive higher profit from selling k goods separately than from selling all the k goods in a single bundle. Because Proposition 5 tells us that for a k -good monopolist, other partial bundling options are always dominated by either full bundle or separate sales, we conclude that separate sales would be the optimal selling strategy for the k -good monopolist, a contradiction to our original hypothesis that the k -good monopolist prefers full bundle.

□ **Numerical analysis for Gaussian demand.** Proposition 5 informs us that even in cases where $p^* > \tilde{\theta}$ and $p_B^* < K\tilde{\theta}$, the monopolist will choose either full bundling or separate sales. However, we do not have a general characterization of when the case $p^* > \tilde{\theta}$ and $p_B^* < K\tilde{\theta}$ will occur and, in such cases, when full bundling will dominate separate sales. In this subsection we report numerical results to shed some light on these questions. Following Schmalensee (1984), we consider the case that the valuation distributions for the goods follow Gaussian distributions. Suppose that a monopolist has K goods, and the valuation for each good is independently drawn from $N(\mu, \sigma^2)$. Let c be the constant unit cost for the goods. Gaussian demand clearly satisfies the symmetric log-concavity density assumption postulated in our analysis. The primitives of the numerical analysis are simply the four parameters (μ, σ, c, K) . In what follows, we set $c = .01$ and fix K at 5 or 15.

We will first briefly describe the numerical analysis.¹⁰ Let $\Phi(\cdot)$ be the CDF of a standard Gaussian distribution $N(0, 1)$. Then the demand for a k -good bundle, if the valuation for each good is independently drawn from $N(\mu, \sigma^2)$, at a bundle price p is given by

$$D(p, \mu, \sigma, k) = 1 - \Phi\left(\frac{p - k\mu}{\sqrt{k}\sigma}\right),$$

and the profit from selling the bundle at price p is

$$\Pi(p, \mu, \sigma, k, c) = (p - kc)D(p, \mu, \sigma, k).$$

In this subsection we use the following notation:¹¹

$$p_k^*(\mu, \sigma; c) = \arg \max_{\{p\}} \Pi(p, \mu, \sigma, k, c)$$

$$\Pi_k^*(\mu, \sigma; c) = \max_{\{p\}} \Pi(p, \mu, \sigma, k, c).$$

For our purposes, we will focus on the single-good monopoly price $p_1^*(\mu, \sigma; c = .01)$, the full bundle monopoly price $p_K^*(\mu, \sigma; c = .01)$, and the difference in profits between separate sales $K\Pi_1^*(\mu, \sigma; c = .01)$ and full bundling $\Pi_K^*(\mu, \sigma; c = .01)$.

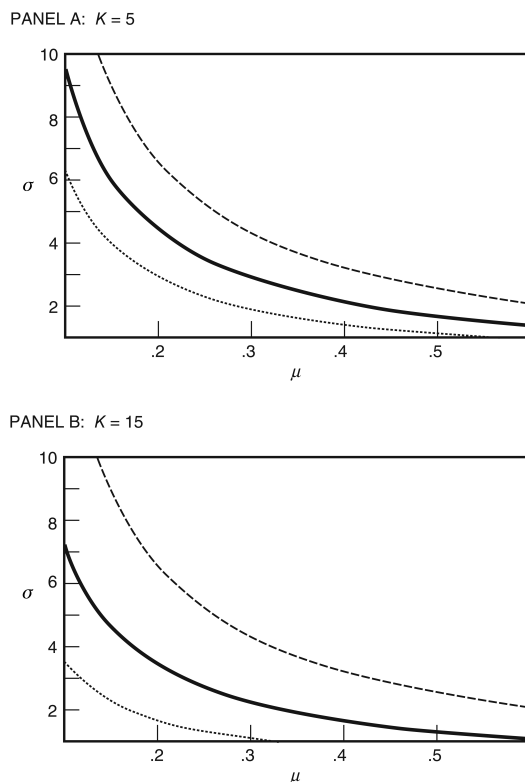
Figure 1 graphically illustrates our numerical results, where panels A and B are respectively for the case $K = 5$ and $K = 15$. In each panel, the outer dashed curve is the combination of (μ, σ) for which the single-good optimal monopolist price $p^* = \tilde{\theta} = \mu$ (thus, in the notation above, it is the contour plot of $p_1^*(\mu, \sigma; c = .01) - \mu = 0$ in the (μ, σ) -space);¹² the inner dashed curve is the combination of (μ, σ) for which the optimal monopolist price for the full bundle $p_B^* = K\tilde{\theta}$ (thus,

¹⁰ The program used in the numerical analysis is available from the authors upon request.

¹¹ The Gaussian distribution has the property that the maximization problem has a unique solution.

¹² The contour plots are shown only for $(\mu, \sigma) \in [0.1, 0.6] \times [1, 10]$.

FIGURE 1



in the notation above, it is the contour plot of $p_K^*(\mu, \sigma; c = .01) - K\mu = 0$ in the (μ, σ) -space). The region of (μ, σ) between the two dashed curves is where $p^* > \tilde{\theta}$ and $p_B^* < K\tilde{\theta}$. The solid line in each panel depicts the combination of (μ, σ) for which full bundling and separate sales at their respectively optimal prices generate the same profit (thus, in the notation above, it is the contour plot of $K\Pi_1^*(\mu, \sigma; c = .01) - \Pi_K^*(\mu, \sigma; c = .01) = 0$ in the (μ, σ) -space). Full bundling (respectively, separate sales) is profit maximizing in the region to the right (respectively, to the left) of the solid curve.

We adopt the Schmalensee (1984) convention to refer to regularities from numerical analysis by the italicized adverb *apparently*. The first pattern to notice is that *apparently* separate sales are profit maximizing when σ is small, and particularly in conjunction with small μ . Second, $p_B^* > K\tilde{\theta}$ (the region to the left of the inner dashed curve) *apparently* frequently holds despite the fact that $c < \tilde{\theta}$; moreover, the region of (μ, σ) in which $p_B^* > K\tilde{\theta}$ is *apparently* shrinking as K increases. This is not surprising in the light of our discussion of the corollary of Proposition 5. Third, for every μ , there is an interval of σ , $(\underline{\sigma}, \overline{\sigma})$, where $p^* > \tilde{\theta}$ and $p_B^* < K\tilde{\theta}$ for all (μ, σ) as long as $\sigma \in (\underline{\sigma}, \overline{\sigma})$. *Apparently*, both the lower bound $\underline{\sigma}$ and upper bound $\overline{\sigma}$ are decreasing in μ . Similarly, for every μ , there is (apparently) a threshold $\hat{\sigma}$ such that full bundling dominates separate sales for (μ, σ) if and only if $\sigma > \hat{\sigma}$. Fourth, as K increases, the region of (μ, σ) in which $p^* > \tilde{\theta}$ and $p_B^* < K\tilde{\theta}$ expands. Moreover, the region of (μ, σ) in which full bundling dominates (the area to the right of the solid curve) also expands, as predicted by the corollary of Proposition 5.

6. Conclusion

■ Many articles on bundling, in particular in the more recent literature, take a “purist” mechanism-design approach to the problem. These articles allow a monopolist to design selling mechanisms that consist of a mapping from vectors of valuations to probabilities to consume

each of the goods and a transfer rule. The problem is then to find the optimal mechanism for the monopolist, subject to incentive and participation constraints. While this in principle is a more satisfactory setup for studying the pros and cons of bundling than the approach in our article, the obvious downside is that the problem is generally rather intractable. Hence, except for a few qualitative features, we know very little about the solution to this problem.

In this article we restrict the monopolist to choose between pure bundling and separate sales. We show that results that are similar to the asymptotic results can be obtained in the more realistic case with a given finite number of goods, provided that the distributions of valuations are symmetric and log-concave. Our results confirm intuition obtained from Schmalensee's (1984) numerical analysis.

Appendix A

■ The proofs of Propositions 1–5 follow.

Proof of Proposition 1. In order not to make a negative profit, the price of the bundle must exceed the costs. Since $\sum_{j=1}^K E\theta_j \leq \sum_{j=1}^K c_j$, we can therefore formulate the monopolist's maximization problem as

$$\Pi_B(K) = \max_{\varepsilon \geq 0} \left[\underbrace{\sum_{j=1}^K E\theta_j + \varepsilon K}_{=p} - \sum_{j=1}^K c_j \right] \Pr \left[\sum_{j=1}^K \theta_j \geq \underbrace{\sum_{j=1}^K E\theta_j + \varepsilon K}_{=p} \right].$$

Using Chebyshev's inequality,

$$\begin{aligned} \Pr \left[\sum_{j=1}^K \theta_j \geq \sum_{j=1}^K E\theta_j + \varepsilon K \right] &\leq \Pr \left[\left| \sum_{j=1}^K \theta_j - \sum_{j=1}^K E\theta_j \right| \leq \varepsilon K \right] \\ &\leq \frac{\text{Var} \left(\sum_{j=1}^K \theta_j \right)}{(\varepsilon K)^2} \leq \frac{K\sigma^2}{(\varepsilon K)^2} = \frac{\sigma^2}{K\varepsilon^2}. \end{aligned}$$

Moreover, $\sum_{j=1}^K E\theta_j + \varepsilon K - \sum_{j=1}^K c_j \leq \varepsilon K$, so

$$\max_{\varepsilon \geq 0} \left[\sum_{j=1}^K E\theta_j + \varepsilon K - \sum_{j=1}^K c_j \right] \Pr \left[\sum_{j=1}^K \theta_j \geq \sum_{j=1}^K E\theta_j + \varepsilon K \right] \leq \max_{\varepsilon \geq 0} \left[\varepsilon K \min \left\{ \frac{\sigma^2}{K\varepsilon^2}, 1 \right\} \right],$$

where the term $\min\{\sigma^2/(K\varepsilon^2), 1\}$ comes from observing that a probability is always less than one (if ε is sufficiently small, the bound from Chebyshev's inequality is useless). We observe that $\sigma^2/(K\varepsilon^2) \leq 1$ if and only if $\varepsilon \geq \sqrt{\sigma^2/K}$, so

$$\varepsilon K \min \left\{ \frac{\sigma^2}{K\varepsilon^2}, 1 \right\} = \begin{cases} \varepsilon K & \text{if } \varepsilon \leq \sqrt{\frac{\sigma^2}{K}} \\ \frac{\sigma^2}{\varepsilon} & \text{if } \varepsilon > \sqrt{\frac{\sigma^2}{K}}, \end{cases}$$

implying that $\max_{\varepsilon \geq 0} [\varepsilon K \min\{\sigma^2/(K\varepsilon^2), 1\}] = \sqrt{\sigma^2 K}$. We conclude that $\Pi_B(K) - \sum_{j=1}^K \Pi_j \leq \sqrt{\sigma^2 K} - K\underline{\Pi} < 0$ for every $K > \sigma^2/\underline{\Pi}^2$. *Q.E.D.*

Proof of Proposition 2. Suppose the monopolist charges a price $p = \sum_{j=1}^K E\theta_j - \delta K/2$ for the full bundle. Then

$$\Pr \left[\sum_{j=1}^K \theta_j \leq p \right] = \Pr \left[\sum_{j=1}^K \theta_j - \sum_{j=1}^K E\theta_j \leq -\frac{\delta K}{2} \right]$$

$$\begin{aligned} &\leq \Pr \left[\left| \sum_{j=1}^K \theta_j - \sum_{j=1}^K E\theta_j \right| \geq \frac{\delta K}{2} \right] \leq \frac{4\text{Var} \left(\sum_{j=1}^K \theta_j \right)}{\delta^2 K^2} \\ &\leq \frac{4\sigma^2}{\delta^2 K}. \end{aligned}$$

Hence, the profit is at least

$$\left(1 - \frac{4\sigma^2}{\delta^2 K} \right) \left(\sum_{j=1}^K E\theta_j - \frac{\delta K}{2} - \sum_{j=1}^K c_j \right),$$

whereas the profit from separate sales (by assumption) is at most $\sum_{j=1}^K E\theta_j - \sum_{j=1}^K c_j - \delta K$. Hence, the difference between the bundling profit and the profit from separate sales is at least

$$\begin{aligned} &\left(1 - \frac{4\sigma^2}{\delta^2 K} \right) \left(\sum_{j=1}^K E\theta_j - \frac{\delta K}{2} - \sum_{j=1}^K c_j \right) - \left(\sum_{j=1}^K E\theta_j - \sum_{j=1}^K c_j - \delta K \right) \\ &= \frac{\delta K}{2} - \frac{4\sigma^2}{\delta^2 K} \left(\sum_{j=1}^K E\theta_j - \frac{\delta K}{2} - \sum_{j=1}^K c_j \right). \end{aligned}$$

Under the assumption that there exists μ such that $E\theta_j < \mu$ for every j , the expression above is positive for K large enough. *Q.E.D.*

Proof of Proposition 3. Suppose for contradiction that the monopolist offers a bundle B_i with more than one good for which $\sum_{j \in B_i} E\theta_j \leq \sum_{j \in B_i} c_j$. Let $n_i \geq 2$ denote the number of goods in bundle B_i , and let g_i and G_i denote the probability density and the cumulative density of the random variable $\theta^i \equiv \sum_{j \in B_i} \theta_j / n_i$. The optimal price of the bundle B_i , denoted by p^{i*} , solves

$$\max_{p^i} \left(p^i - \sum_{j \in B_i} c_j \right) \Pr \left[\sum_{j \in B_i} \theta_j \geq p^i \right] = \max_{p^i} \left(p^i - \sum_{j \in B_i} c_j \right) \left[1 - G_i \left(\frac{p^i}{n_i} \right) \right]. \quad (\text{A1})$$

Log-concavity of f implies log-concavity of g_i , which in turn implies that (A1) has a unique solution p_i^* . Moreover, it must be the case that $p^{i*} > \sum_{j \in B_i} c_j$, since any price less than or equal to $\sum_{j \in B_i} c_j$ yields a nonpositive profit, whereas any price in the interval $(\sum_{j \in B_i} c_j, n_i \tilde{\theta})$ yields a strictly positive profit.

Now consider a deviation where the monopolist sells all the goods in the bundle B_i separately at price p^{i*}/n_i per good. By Theorem 1, g_i is strictly more peaked than the underlying density f . Since $p^{i*}/n_i > \sum_{j \in B_i} c_j/n_i \geq \sum_{j \in B_i} E\theta_j/n_i = \tilde{\theta}$, we have $G_i(p^{i*}/n_i) > F(p^{i*}/n_i)$. Hence,

$$\begin{aligned} \sum_{j \in B_i} \left(\frac{p^{i*}}{n_i} - c_j \right) \left[1 - F \left(\frac{p^{i*}}{n_i} \right) \right] &= \left(p^{i*} - \sum_{j \in B_i} c_j \right) \left[1 - F \left(\frac{p^{i*}}{n_i} \right) \right] \\ &> \left(p^{i*} - \sum_{j \in B_i} c_j \right) \left[1 - G_i \left(\frac{p^{i*}}{n_i} \right) \right], \end{aligned}$$

showing that unbundling the goods in B_i increases the profit for the monopolist. *Q.E.D.*

Proof of Proposition 4. The essence of the proof is that bundling at a constant per-good price leads to higher sales if and only if the per-good price is below the mode of the distribution.

Part (i). Suppose, for contradiction, that there are in the monopolist optimal bundling menu at least two bundles labelled B_1 and B_2 . For $i = 1, 2$, let n_i denote the number of goods in B_i , and let g_i (respectively G_i) denote the PDF (respectively the CDF) of $\theta^i \equiv \sum_{j \in B_i} \theta_j / n_i$. Because log-concavity is preserved under convolutions (see, e.g., Karlin, 1968), g_1 and g_2 are both symmetric log-concave densities with expectation $\tilde{\theta}$. Thus, if the monopolist charges p^i for bundle B_i , the profit function

$$\left(p^i - \sum_{j \in B_i} c_j \right) \left[1 - G_i \left(\frac{p^i}{n_i} \right) \right] = (p^i - n_i c) \left[1 - G_i \left(\frac{p^i}{n_i} \right) \right]$$

is single-peaked in p^i for $i = 1, 2$ (see footnote 7 for the proof of single-peakedness). Let p^{i*} denote the optimal price of bundle B_i . We observe the following.

Claim A1. $p^{i*}/n_i \leq \tilde{\theta}$.

To see this, first note that if B_i contains a single good, the claim immediately follows from the stated condition $p^* \leq \tilde{\theta}$. If B_i contains more than a single good, suppose to the contrary that $p^{i*}/n_i > \tilde{\theta}$. Due to single-peakedness of the profit function, $p^{i*}/n_i > \tilde{\theta}$ implies that

$$\left. \frac{d}{dp^i} \right|_{p^i = n_i \tilde{\theta}} \left\{ (p^i - n_i c) \left[1 - G_i \left(\frac{p^i}{n_i} \right) \right] \right\} = 1 - G_i(\tilde{\theta}) - (\tilde{\theta} - c)g_i(\tilde{\theta}) = \frac{1}{2} - (\tilde{\theta} - c)g_i(\tilde{\theta}) > 0. \quad (\text{A2})$$

However, the condition $p^* \leq \tilde{\theta}$ implies that

$$\left. \frac{d}{dp} \right|_{p=\tilde{\theta}} (p - c)[1 - F(p)] = 1 - F(\tilde{\theta}) - (\tilde{\theta} - c)f(\tilde{\theta}) = \frac{1}{2} - (\tilde{\theta} - c)f(\tilde{\theta}) \leq 0. \quad (\text{A3})$$

But (A2) and (A3) together imply that $g_i(\tilde{\theta}) < f(\tilde{\theta})$. However, Theorem 1 implies that θ^i is strictly more peaked than the underlying distribution, which in turn implies that $g_i(\tilde{\theta}) > f(\tilde{\theta})$. The claim thus follows from the contradiction.

Now consider a deviation where the monopolist sells all the goods in B_1 and B_2 as a single bundle, labelled as \hat{B} . Furthermore, consider the random pricing mechanism \hat{p} where

$$\hat{p} = \begin{cases} \frac{n_1 + n_2}{n_1} p^{1*} & \text{with probability } \frac{n_1}{n_1 + n_2} \\ \frac{n_1 + n_2}{n_2} p^{2*} & \text{with probability } \frac{n_2}{n_1 + n_2}. \end{cases} \quad (\text{A4})$$

Denote by \hat{G} and \hat{g} respectively the CDF and PDF of $\hat{\theta} = \sum_{j \in B_1 \cup B_2} \theta_j / (n_1 + n_2)$. The profit from sales of the bundle \hat{B} is then

$$\begin{aligned} \hat{\Pi} &= \frac{n_1}{n_1 + n_2} \left[\frac{n_1 + n_2}{n_1} p^{1*} - (n_1 + n_2)c \right] \Pr \left[\sum_{j \in B_1 \cup B_2} \theta_j \geq \frac{n_1 + n_2}{n_1} p^{1*} \right] \\ &\quad + \frac{n_2}{n_1 + n_2} \left[\frac{n_1 + n_2}{n_2} p^{2*} - (n_1 + n_2)c \right] \Pr \left[\sum_{j \in B_1 \cup B_2} \theta_j \geq \frac{n_1 + n_2}{n_2} p^{2*} \right] \\ &= (p^{1*} - n_1 c) \left[1 - \hat{G} \left(\frac{p^{1*}}{n_1} \right) \right] + (p^{2*} - n_2 c) \left[1 - \hat{G} \left(\frac{p^{2*}}{n_2} \right) \right]. \end{aligned} \quad (\text{A5})$$

First, suppose that $p^{1*}/n_1 < \tilde{\theta}$. Then, since \hat{G} is strictly more peaked than G_1 , it follows that $\hat{G}(p^{1*}/n_1) < G_1(p^{1*}/n_1)$. Moreover, since $p^{2*}/n_2 \leq \tilde{\theta}$, we have that $\hat{G}(p^{2*}/n_2) \leq G_2(p^{2*}/n_2)$. Combining with (A5), we obtain

$$\hat{\Pi} > (p^{1*} - n_1 c) \left[1 - G_1 \left(\frac{p^{1*}}{n_1} \right) \right] + (p^{2*} - n_2 c) \left[1 - G_2 \left(\frac{p^{2*}}{n_2} \right) \right].$$

Hence, the bundle \hat{B} generates a higher profit than the sum of the profits from B_1 and B_2 . This is true for the analogous case where $p^{2*}/n_2 < \tilde{\theta}$ and $p^{1*}/n_1 < \tilde{\theta}$. The only remaining case to consider is if $p^{1*}/n_1 = p^{2*}/n_2 = \tilde{\theta}$. In this case the profit from selling \hat{B} at price $\hat{p} = (n_1 + n_2)\tilde{\theta}$ is the same as the sum of profits from selling B_1 and B_2 as separate bundles. However, for $p^{i*} = n_i \tilde{\theta}$ to be optimal, it is necessary that

$$\left. \frac{d}{dp^i} \right|_{p^i = n_i \tilde{\theta}} \left\{ (p^i - n_i c) \left[1 - G_i \left(\frac{p^i}{n_i} \right) \right] \right\} = \frac{1}{2} - (\tilde{\theta} - c)g_i(\tilde{\theta}) = 0.$$

This in turn implies that

$$\left. \frac{d}{d\hat{p}} \right|_{\hat{p} = (n_1 + n_2)\tilde{\theta}} \left\{ [\hat{p} - (n_1 + n_2)c] \left[1 - \hat{G} \left(\frac{\hat{p}}{n_1 + n_2} \right) \right] \right\} = \frac{1}{2} - (\tilde{\theta} - c)\hat{g}(\tilde{\theta}) < 0$$

because \hat{g} is strictly more peaked than g_i . Thus a small decrease in the price of the joint bundle \hat{B} from $(n_1 + n_2)\tilde{\theta}$ will lead to profits that are strictly higher than the sum from selling B_1 and B_2 as separate bundles.

Part (ii). Suppose for contradiction that there is at least one bundle, labelled B_k , consisting of more than a single good in the profit-maximizing bundle menu. Let n_k be the number of goods in B_k , and let g_k and G_k respectively be the PDF

and CDF of $\theta^k \equiv \sum_{j \in B_k} \theta_j / n_k$. Using reasoning similar to that in the proof of Claim A1, one can show this establishes that if $p_B^* \geq K\tilde{\theta}$, then the optimal bundle price for B_k , denoted by p^{k*} , must satisfy $p^{k*} \geq n_k\tilde{\theta}$. Once this is established, we can now consider a deviation where the monopolist sells all goods in B_k separately, charging a price p^{k*}/n_k for each good. Under this selling strategy, the profits from selling the goods in B_k are

$$\sum_{j \in B_k} \left(\frac{p^{k*}}{n_k} - c_j \right) \left[1 - F \left(\frac{p^{k*}}{n_k} \right) \right] = \left(p^{k*} - \sum_{j \in B_k} c_j \right) \left[1 - F \left(\frac{p^{k*}}{n_k} \right) \right]. \quad (A6)$$

From Theorem 1 we know that g_k is strictly more peaked than f . Thus if $p^{k*}/n_k > \tilde{\theta}$, we have $G_k(p^{k*}/n_k) > F(p^{k*}/n_k)$. Thus the profit in (A6) is strictly larger than the profit from the bundle B_k at price p^{k*} . For the case in which $p^{k*} = n_k\tilde{\theta}$, the same argument as in the last step of the proof of part (i) can be used to show that the profit from selling the goods in B_k separately can be strictly increased if one is to marginally increase the single-good price from $\tilde{\theta}$. *Q.E.D.*

Proof of Proposition 5. Suppose not. Let the optimal bundling menu be given by $\mathcal{B} = \{B_1, \dots, B_M\}$, where $2 \leq M \leq K - 1$. Without loss of generality assume that $n_1 \geq 2$ and $n_1 \geq n_2 \geq \dots \geq n_M$. Let $\bar{i} \geq 1$ be the highest index such that $n_{\bar{i}} \geq 2$, and let p^{i*} be the optimal price for bundle B_i . First observe that for all $i \leq \bar{i}$, it must be the case that $p^{i*} \leq n_i\tilde{\theta}$. Otherwise, by the same argument as in the proof of Proposition 3, the monopolist would increase sales and therefore profits by selling the goods in B_i separately at price p^{i*}/n_i each for all $i \leq \bar{i}$. Thus we can apply the same argument as that in the proof of part (i) of Proposition 4 to show that creating a single bundle that includes all goods in the bundles $B_1, \dots, B_{\bar{i}}$ can strictly increase the profit for the monopolist. Hence the only remaining possibility we need consider is that there is one nontrivial bundle, which we will label \hat{B} , and that the rest of the goods are sold separately.

Assume without loss that $\hat{B} = \{1, \dots, k\}$ and that goods $k+1, \dots, K$ are sold separately. Let $\Pi_{\hat{B}}$ denote the profit from sales of the bundle \hat{B} and \hat{p}^* be the profit-maximizing monopoly price for bundle \hat{B} . Arguments above imply that $\hat{p}^* \leq k\tilde{\theta}$. Let Π denote the maximized profit from selling a good if sold separately and Π_B denote the maximized profit if all goods are bundled. The presumed optimality of the bundling menu $\mathcal{B} = \{\hat{B}, \{k+1\}, \dots, \{K\}\}$ implies that

$$\Pi_{\hat{B}} + (K - k)\Pi \geq \Pi_B, \quad (A7)$$

$$\Pi_{\hat{B}} + (K - k)\Pi \geq K\Pi, \quad (A8)$$

where the first inequality says that the bundling menu \mathcal{B} is more profitable than selling all goods in a single bundle B , and the second inequality says that \mathcal{B} is more profitable than selling all goods separately. But applying the argument in Proposition 4, we can show that it must be the case that $\hat{p}^* \leq k\tilde{\theta}$. This in turn must imply the following.

Claim A2. $\Pi_B/K > \Pi_{\hat{B}}/k$.

Claim A2 can be proved as follows. By definition of Π_B , we have

$$\frac{\Pi_B}{K} = \max_{p_B} \left(\frac{p_B}{K} - c \right) \left[1 - G_B \left(\frac{p_B}{K} \right) \right],$$

where G_B is the CDF of $\sum_{j=1}^K \theta_j / K$. Thus (by setting p_B above to \hat{p}^*K/k)

$$\frac{\Pi_B}{K} \geq \left(\frac{\hat{p}^*}{k} - c \right) \left[1 - G_B \left(\frac{\hat{p}^*}{k} \right) \right].$$

From Theorem 1, G_B is strictly more peaked than \hat{G} , the CDF of $\sum_{j=1}^k \theta_j / k$; and since $\hat{p}^*/k \leq \tilde{\theta}$, we have

$$\left(\frac{\hat{p}^*}{k} - c \right) \left[1 - G_B \left(\frac{\hat{p}^*}{k} \right) \right] > \left(\frac{\hat{p}^*}{k} - c \right) \left[1 - \hat{G} \left(\frac{\hat{p}^*}{k} \right) \right] \equiv \frac{\Pi_{\hat{B}}}{k}.$$

Therefore,

$$\frac{\Pi_B}{K} > \frac{\Pi_{\hat{B}}}{k}. \quad (A9)$$

Now we show that the inequalities (A7)–(A9) cannot simultaneously hold. To see this, if (A7) and (A9) hold, then

$$\Pi_{\hat{B}} + (K - k)\Pi \geq \Pi_B > \frac{K}{k}\Pi_{\hat{B}},$$

which after simplification implies that

$$k\Pi > \Pi_{\hat{B}},$$

which in turn contradicts (A8). Thus, the only possible optimal bundling menu is full bundling or fully separate sales. *Q.E.D.*

Appendix B

■ **Calculations for the example in Section 5.** We want to show that if the monopolist sets a price $p_B < 2$, then, if α is sufficiently small, the profit is lower than the maximized profit under separate sales, $\Pi_1 + \Pi_2 = 2 - \alpha$. We begin with a simple observation.

Claim B1. It must be the case that $p_B > 2 - \alpha$ in order for the profit under bundling to exceed $2 - \alpha$.

This is obvious, since p_B would be the profit if the consumer would buy for sure.

Next, observe that for $p < 1$ we have that

$$\begin{aligned} 1 - G_B(p) &= \Pr\left[\frac{\theta_1 + \theta_2}{2} \geq p\right] \\ &= \underbrace{(1 - \alpha)^2}_{\Pr[\theta_1, \theta_2] = (1, 1)} + \underbrace{2(1 - \alpha)\alpha}_{\Pr[\theta_1 = 1 \cap \theta_2 \neq 1] + \Pr[\theta_1 \neq 1 \cap \theta_2 = 1]} [1 - \bar{F}(2p - 1)] + \underbrace{\alpha^2}_{\Pr[\theta_1 \neq 1 \cap \theta_2 \neq 1]} [1 - G(p)], \end{aligned}$$

where \bar{F} is the CDF of the underlying uniform distribution over $[0, 2]$ and

$$G(p) = \begin{cases} \frac{p^2}{2} & \text{on } [0, 1] \\ 1 - \frac{(2 - p)^2}{2} & \text{on } [1, 2]. \end{cases}$$

By Claim B1, we can restrict our attention to values of $p_B \in (2 - \alpha, 2)$, which is equivalent to restricting to per-good average price $p \in (1 - \alpha/2, 1)$. Since $\alpha \in [0, 1]$, $(1 - \alpha/2, 1)$ is a subset of $(1/2, 1)$. For any $1/2 < p < 1$, the monopolist's profit from selling a bundle at a bundle price of $2p$ receives profit

$$2p[1 - G_B(p)] = 2p \left[(1 - \alpha)^2 + 2(1 - \alpha)\alpha \left(\frac{3 - 2p}{2} \right) + \alpha^2 \left(1 - \frac{p^2}{2} \right) \right].$$

On the other hand, the monopolist's profit from selling the two goods at price p for each receives profit

$$2p[1 - F(p)] = 2p \left(1 - \frac{\alpha}{2} p \right).$$

Define

$$\begin{aligned} \Delta(p) &= G_B(p) - F(p) \\ &= 1 - \left[(1 - \alpha)^2 + 2(1 - \alpha)\alpha \left(\frac{3 - 2p}{2} \right) + \alpha^2 \left(1 - \frac{p^2}{2} \right) \right] - \frac{\alpha}{2} p \\ &= (1 - \alpha)^2 + 2\alpha(1 - \alpha) + \alpha^2 - \left[(1 - \alpha)^2 + 2(1 - \alpha)\alpha \left(\frac{3 - 2p}{2} \right) + \alpha^2 \left(1 - \frac{p^2}{2} \right) \right] - \frac{\alpha}{2} p \\ &= \alpha \left[(1 - \alpha)(2p - 1) + \alpha \frac{p^2}{2} - \frac{1}{2} p \right]. \end{aligned}$$

Note that $2p\Delta(p)$ measures the difference in the monopolist's profit between selling each good separately at a price of p for each good and selling the bundle at a price of $2p$. Thus if $\Delta(p)$ is positive, then the monopolist increases its profit by selling the goods separately at half the price of the bundled good; if $\Delta(p)$ is negative, then the profit under bundling is higher.

Next, we show that $\Delta(p)$ is monotonic on $(1/2, 1)$. Differentiating $\Delta(p)$, we have that

$$\frac{d\Delta(p)}{dp} = \alpha \left[2(1 - \alpha) + \alpha p - \frac{1}{2} \right] = \alpha \left[\frac{3}{2} + \alpha(p - 2) \right]$$

$$> \alpha \left[\frac{3}{2} + \alpha \left(\frac{1}{2} - 2 \right) \right] = \alpha \frac{3}{2} (1 - \alpha) > 0.$$

Hence we have the following.

Claim B2. $\Delta(p)$ is strictly increasing on $(1/2, 1)$.

We know from Claim B1 that we only need to consider $p_B > 2 - \alpha$, which corresponds to an average price $p > 1 - \alpha/2$. Evaluating $\Delta(p)$ at $p = 1 - \alpha/2$, we have that

$$\begin{aligned} \Delta \left(1 - \frac{\alpha}{2} \right) &= \alpha \left\{ (1 - \alpha) \left[2 \left(1 - \frac{\alpha}{2} \right) - 1 \right] + \alpha \left[1 - \frac{\left(2 - \left(1 - \frac{\alpha}{2} \right) \right)^2}{2} \right] - \frac{1}{2} \left(1 - \frac{\alpha}{2} \right) \right\} \\ &= \alpha \left\{ (1 - \alpha)^2 + \alpha \left[1 - \frac{\left(1 + \frac{\alpha}{2} \right)^2}{2} \right] - \frac{1}{2} \left(1 - \frac{\alpha}{2} \right) \right\}. \end{aligned}$$

Hence,

$$\begin{aligned} \Delta \left(1 - \frac{\alpha}{2} \right) &\geq 0 \Leftrightarrow \\ (1 - \alpha)^2 + \alpha \left[1 - \frac{\left(1 + \frac{\alpha}{2} \right)^2}{2} \right] - \frac{1}{2} \left(1 - \frac{\alpha}{2} \right) &\geq 0 \Leftrightarrow \\ (1 - \alpha) - \alpha(1 - \alpha) + \alpha - \alpha \frac{\left(1 + \frac{\alpha}{2} \right)^2}{2} - \frac{1}{2} \left(1 - \frac{\alpha}{2} \right) &\geq 0 \Leftrightarrow \\ \frac{1}{2} - \alpha(1 - \alpha) - \alpha \frac{\left(1 + \frac{\alpha}{2} \right)^2}{2} + \frac{\alpha}{4} &\geq 0. \end{aligned}$$

We conclude as follows.

Claim B3. $\Delta(1 - \alpha/2) > 0$ for α sufficiently small.

To sum up:

- (i) Claim B1 shows that bundling at $p_B < 2 - \alpha$ is dominated by separate sales.
- (ii) Claim B2 shows that bundling at any price on the interval $(2 - \alpha, 2)$ leads to lower sales than separate sales, provided that α is small enough.
- (iii) Claim B3 shows that bundling at $p_B = 2 - \alpha$ also leads to lower sales than separate sales if α is small enough.

Together, these claims imply that for α sufficiently small, there exists no price $p_B < 2$ for the bundled good that gives a higher payoff than $\Pi_1 + \Pi_2$.

References

- ADAMS, W.J. AND YELLEN, J.L. "Commodity Bundling and the Burden of Monopoly." *Quarterly Journal of Economics*, Vol. 90 (1976), pp. 475–498.
- AN, M.Y. "Logconcavity Versus Logconvexity: A Complete Characterization." *Journal of Economic Theory*, Vol. 80 (1998), pp. 350–369.
- ARMSTRONG, M. "Price Discrimination by a Many-Product Firm." *Review of Economic Studies*, Vol. 66 (1999), pp. 151–168.
- BAGNOLI, M. AND BERGSTROM, T. "Log-Concave Probability and Its Applications." *Economic Theory*, Vol. 26 (2005), pp. 445–469.
- BAKOS, Y. AND BRYNJOLFSSON, E. "Bundling Information Goods: Pricing, Profits and Efficiency." *Management Science*, Vol. 45 (1999), pp. 1613–1630.
- BIRNBAUM, Z.W. "On Random Variables with Comparable Peakedness." *Annals of Mathematical Statistics*, Vol. 19 (1948), pp. 76–81.
- CRAWFORD, G. "The Discriminatory Incentives to Bundle in the Cable Television Industry." Working Paper, University of Arizona, 2004.
- FANG, H. AND NORMAN, P. "To Bundle or Not to Bundle." Cowles Foundation Discussion Paper no. 1440, Yale University, 2003.

- AND ———. “An Efficiency Rationale for Bundling of Public Goods.” Cowles Foundation Discussion Paper no. 1441, Yale University, 2003.
- GENG, X.J., STINCHCOMBE, M.B., AND WHINSTON, A.B. “Bundling Information Goods of Decreasing Value.” *Management Science*, Vol. 51 (2005), pp. 662–667.
- IBRAGIMOV, R. “Optimal Bundling Decisions for Complements and Substitutes with Heavy Tail Distributions.” Working Paper, Harvard University, 2005.
- KARLIN, S. *Total Positivity*. Stanford, Calif.: Stanford University Press, 1968.
- MCAFEE, R.P., McMILLAN, J., AND WHINSTON, M.D. “Multiproduct Monopoly, Commodity Bundling, and Correlation of Values.” *Quarterly Journal of Economics*, Vol. 104 (1989), pp. 371–384.
- MYERSON, R. “Optimal Auction Design.” *Mathematics of Operations Research*, Vol. 6 (1981), pp. 58–73.
- NALEBUFF, B. “Bundling as an Entry Barrier.” *Quarterly Journal of Economics*, Vol. 119 (2004), pp. 159–187.
- OFFICE OF FAIR TRADING. *BSkyB decision dated 17 December 2002: rejection of applications under section 47 of the Competition Act 1998 by ITV Digital in Liquidation and NTL*. London, UK, July 2003. Available at: www.offt.gov.uk/business/competition+act/decisions/bskyb.htm.
- PALFREY, T.R. “Bundling Decisions by a Multiproduct Monopolist with Incomplete Information.” *Econometrica*, Vol. 51 (1983), pp. 463–484.
- PROSCHAN, F. “Peakedness of Distributions of Convex Combinations.” *Annals of Mathematical Statistics*, Vol. 36 (1965), pp. 1703–1706.
- RILEY, J. AND ZECKHAUSER, R. “Optimal Selling Strategies: When to Haggle, When to Hold Firm.” *Quarterly Journal of Economics*, Vol. 98 (1983), pp. 267–289.
- SCHMALENSEE, R. “Commodity Bundling by Single-Product Monopolies.” *Journal of Law and Economics*, Vol. 15 (1982), pp. 67–71.
- . “Gaussian Demand and Commodity Bundling.” *Journal of Business*, Vol. 57 (1984), pp. S211–S230.