

To Bundle or Not To Bundle*

Hanming Fang[†]

Peter Norman[‡]

First version: September 3, 2003

This Version: June 8, 2005

Abstract

Comparing monopoly bundling with separate sales is relatively straightforward in an environment with a large number of goods. In this paper 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.

Keywords: Monopoly Pricing; Bundling; Logconcavity; Peakedness.

JEL Classification Number: L11, L12.

*The first version of this paper unbundled some material that was previously provided in Fang and Norman [9]. We are especially grateful to Ted Bergstrom and Larry Samuelson who suggested an idea that inspired the example in Section 5.1. We also thank 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.

[†]Department of Economics, Yale University, P.O. Box 208264, New Haven, CT 06520-8264, hanming.fang@yale.edu

[‡]Department of Economics, University of British Columbia, 909-1873 East Mall, Vancouver, BC V6T1Z1, penorman@interchange.ubc.ca

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. (US Senator John McCain in an interview on Cable TV rates published in the Washington Post, 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 other goods in the bundle are consumed or not, and when no costs are saved through bundling (Adams and Yellen [1]).

In the earliest literature, bundling was typically seen as a way to exploit negative correlation between valuations for different goods (see Adams and Yellen [1] and Schmalensee [21]). Since then, it has been shown that bundling is useful also when valuations for different goods are stochastically independent. In particular, McAfee et al [14] 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 the case where goods are only offered separately, provided that a condition on the joint distribution of valuations is satisfied. The distributional condition holds generically, and is implied by stochastic independence, so the logic has nothing to do with exploiting correlations. An obvious downside with the generality of this result is that the analysis is silent on which goods should be bundled. All goods should be bundled in one way or the other!

In this paper, we show that if mixed bundling is ruled out (by assumption), then we obtain a rather intuitive characterization for when the monopolist should bundle and when separate sales is a better idea. To some extent this characterization confirms (mainly) numerical results in Schmalensee [22], namely, the higher the marginal cost and the lower is the mean valuation, the less likely that bundling dominates separate sales.

When limiting our comparison to pure bundling versus separate provision 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. In our paper we 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” means. Specifically, we need to be able to conclude that for a given per-good price below (above) the mean, bundling increases (reduces) the probability of trade. This can be rephrased as saying that the average valuation is more peaked than the underlying distributions. Notice that the law of large numbers can be used to make 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, and, assuming that valuations are distributed in accordance to symmetric and log-concave densities, we can use a result from Proschan [19] to unambiguously rank distributions in terms of relative peakedness.¹

Bundling therefore reduces the effective dispersion in the buyers’ tastes under our distributional assumptions. Sometimes, that is when a good should be sold with high probability (either because costs are low or because valuations tend to be high) this is to the advantage for the monopolist. At other times, that is for goods with a thin market, the monopolist is better off relying on the right tail of the distribution and bundling is a bad idea.

The idea that “bundling reduces dispersion” has been around for a long time, and there is even some emerging empirical evidence supporting this as a motivation to bundle (see Crawford [7]). What is largely missing in the literature however are results at a reasonable level of generality that establishes conditions for when bundling can be explained in this way. Schmalensee [22] considers the case with normally distributed distributions of valuations (which belongs to the class we consider), and, relying mainly on numerical methods, he reaches a similar conclusion. Recently, Ibragimov [12] has developed a related characterization relying on a generalization of the result in Proschan [19].

In the context of “information goods” (goods with zero marginal costs) Bakos and Brynjolfsson [5] used a similar idea to argue that bundling is better than separate sales. More recently, Geng et al [10] shows that even the case with zero marginal costs can yield surprising results. While both Bakos and Brynjolfsson [5] and Geng et al [10] assume zero marginal costs, the main difference with our paper is that they focus on results for large numbers of goods. We also prove some asymptotic results, but our preference structure is much simpler, and the results are accordingly straightforward. Instead, our contribution is that we can get results for the case with a given finite number of goods that are similar in spirit to the large numbers analysis.

¹Recently, Ibragimov [11] has generalized the result in Proschan [19].

2 The Model

The underlying economic environment is the same as in McAfee et al [14], 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 consumers' 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 1 if good j is consumed and 0 otherwise. Valuations are assumed stochastically independent and we let F_j denote the marginal distribution of θ_j . Hence, $\times_{j=1}^K F_j(\theta_j)$ 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 [6]:

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*

$$\Pr [|x_1 - \mathbf{E}(x_1)| \geq t] \leq \Pr [|x_2 - \mathbf{E}(x_2)| \geq t]$$

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 .²

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,

²Strictly speaking, Birnbaum [6] 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 [19] and drop the qualifiers.

extreme value, exponential, Laplace, Weibull, and many other common parametric densities (see Bagnoli and Bergstrom [4] for further examples). Comparative peakedness of convex combinations of log-concave random variables are studied in Proschan [19], and we will apply one of his results in this paper. To avoid discussing majorization theory we will use his key lemma directly rather than his main result.

Theorem 1 (Lemma 2.2 in Proschan [19]) *Let f be a symmetric log-concave density. Suppose that x_1, \dots, x_m are independently distributed with density f , fix $(w_3, \dots, w_m) \geq \mathbf{0}$ with $\sum_{i=3}^m w_i < 1$. Then*

$$w_1 x_1 + \left(1 - w_1 - \sum_{i=3}^m w_i\right) x_2 + \sum_{i=3}^m w_i x_i$$

is strictly increasing in peakedness as w_1 increases from 0 to $\frac{1 - \sum_{i=3}^m w_i}{2}$.

A corollary of this result is that $\frac{1}{m} \sum_{i=1}^m x_i$ is strictly increasing in peakedness in m , 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.³ 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 5.1. 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 is only meant to establish a benchmark to compare the results in Section 5 with. The basic ideas are similar to Armstrong [3] and Bakos and Brynjolfsson [4], and 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 et al [10].

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

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

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

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 that, for each $\varepsilon > 0$ there exists $K(\varepsilon) < \infty$ such that

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

for any $K > K(\varepsilon)$. Obviously, (2) implies that;

1. $\Pr \left[\sum_{j=1}^K \theta_j \geq p \right] \geq 1 - \varepsilon$ if $p \leq \sum_{j=1}^K \mathbf{E}\theta_j - \varepsilon K$
2. $\Pr \left[\sum_{j=1}^K \theta_j \geq p \right] \leq \varepsilon$ if $p \geq \sum_{j=1}^K \mathbf{E}\theta_j + \varepsilon K$

In other words, charging a price which 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 which 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 there exists σ^2 such that $\text{Var}\theta_j \leq \sigma^2$ for every j and $\delta > 0$ such that $\Pi_j \geq \delta$ for every j (where Π_j is defined in (1)). Also, suppose that $\sum_{j=1}^K \mathbf{E}\theta_j \leq \sum_{j=1}^K c_j$ for every K . Then, selling all goods separately is better than selling all goods as a single bundle whenever $K > \sigma^2/\delta^2$.⁴*

The condition says that if the profit from separate sales is non-negligible for every good and if costs exceeds the sum of the expected valuations, 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

⁴Instead of viewing the proposition as a comparison between separate sales and bundling all goods, one may also interpret the result as saying that σ^2/δ^2 is an upper bound on the number of goods in each bundle.

sum of the costs in order to make a profit. But, if many goods are bundled, such price inevitably leads to negligible sales.

Except for the rather innocuous restriction that $\Pi_j \geq \delta$ for each 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, while 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. Any reasonably general condition for when bundling dominates asymptotically therefore must be expressed in the (endogenous) non-bundling profits.

Proposition 2 *Suppose that there exists μ, σ^2 such that $E\theta_j \leq \mu$ and $\text{Var}\theta_j \leq \sigma^2$ for every j .⁵ 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 j (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 5.2 it may be observed 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 for the discussion below to observe that if we make the further regularity assumptions that $E\theta_j$ and $\text{Var}\theta_j$ are bounded away from zero, then (3) is automatically satisfied if $c_j = 0$ for every j .

4.1 The Inefficiency of Bundling

The reader should note that all we have done is to compare the profits of two possible selling strategies. Except for in very special circumstances, both separate sales and full bundling are economically inefficient. The inefficiency of bundling should be obvious: consumers will purchase

⁵The 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 monopoly price 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 rejecting the bundle could be more important than the increase in profit conditional on selling the bundle.

goods for which their valuation is below the unit cost of production in order to be able to consume other goods that they value highly.

In the case with many goods, Armstrong [3] has shown that a two-part tariff can be used to obtain an approximately efficient allocation, where the monopolist almost fully extracts the surplus. This mechanism sets a tariff T for the right to purchase any good the monopolist sells at marginal cost. With T/K set a little bit below the average consumer surplus (where the average is taken over the K goods), almost all types are willing to pay the tariff T . In general, this mechanism dominates the pure bundling mechanism both in terms of the profit for the monopolist and the social surplus generated. The one exception is when all the unit costs are zero—in which case the two mechanisms coincide. While this observation is trivial it does suggest that the place to search for an efficiency enhancing role of bundling is in the context of non-rival goods, or, more generally, in natural monopoly situations.⁶

5 To Bundle or Not to Bundle in the Finite Case

5.1 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,

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

This cumulative distribution is most easily 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 to set $p = 1$. But, for $\alpha < 1$, $p = 1$ continues to be the optimal price, since 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

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

⁶To make this problem interesting, there has to be a non-trivial decision as to whether a good should be produced or not. See Fang and Norman [9] for an analysis.

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 point of the change in variable is that the question as to whether separate sales or bundling is better is transformed into a comparative statics exercise with respect to the cumulative distribution of valuations.

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

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

for $t \in [0, t^*]$. Hence, F_A and F cannot be compared in terms of relative peakedness. For the comparison between bundling and separate sales, the implication of this 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 doesn't *prove* that bundling is worse since (i) a price slightly above the expectation leads to higher sales than when goods are sold separately, (ii) a sufficiently large reduction in price relative from the expected value also leads to higher sales than when goods are sold separately.

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 F_A 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 this 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 1, then this will indeed lead to an increase in profits. However, if α is sufficiently small, any price strictly below 2 will generate lower profits than the maximized

profits under separate sales. This is shown formally in Appendix B. The idea is that 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 2 this price is, and for α small such a price is in the range where $F_A\left(\frac{2-\alpha}{2}\right) > F\left(\frac{2-\alpha}{2}\right)$. But this implies that any price for the bundled good on the interval $(2 - \alpha, 2)$ is worse than selling the goods separately at price $\frac{2-\alpha}{2}$ each.

Case 3: $p_B > 2$. Finally, we need to consider $p_B > 2$. However, as $\alpha \rightarrow 0$ the probability of selling the bundle at such a price 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 to 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 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 dominates bundling. Notice that this is despite the fact that the distribution is symmetric, unimodal, smooth, and generated as a mixture of two (different) logconcave densities with identical means. However, mixtures of logconcave densities are not necessarily logconcave (see Section 3.4 in An [2]).

5.2 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;

1. Decide how to package the goods into different bundles, which, with mixed bundling ruled out, is the same as partitioning the set of goods produced by the monopolist in what we refer to as a *bundling menu*. Following Palfrey [17], we denote such a bundling menu by $B = \{B_1, \dots, B_M\}$, where each $B_i \in B$ 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 $B = \{\{1\}, \{2\}, \dots, \{K\}\}$ corresponds

with separate sales and $B = \{\{1, \dots, K\}\}$ describes the other extreme case where all goods are sold as a single bundle.

2. 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, 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 [15], Riley and Zeckhauser [20]) 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 which is strictly positive on support $[\underline{\theta}, \bar{\theta}]$ and has expectation $\tilde{\theta}$. Assume that each good j is produced at unit cost c_j , where $c_j < \bar{\theta}$. Let B^* be the optimal bundling menu for the monopolist. Then, there exists no $B_i \in B^*$ such that $\sum_{j \in B_i} \mathbf{E}\theta_j \leq \sum_{j \in B_i} c_j$.*

Proof. Suppose for contradiction that the monopolist offers a bundle B_i for which $\sum_{j \in B_i} \mathbf{E}\theta_j \leq \sum_{j \in B_i} c_j$. Let n_i denote the number of goods in bundle B_i , and let f_i and F_i denote the probability density and the cumulative density of the random variable $\theta^i \equiv \sum_{j \in B_i} \frac{\theta_j}{n_i}$. The optimal price of the bundle B_i solves

$$\max_{p_i} \left(p_i - \sum_{j \in B_i} c_j \right) \Pr \left[\sum_{j \in B_i} \theta_j \geq p_i \right] = \max_{p_i} \left(p_i - \sum_{j \in B_i} c_j \right) \left[1 - F_i \left(\frac{p_i}{n_i} \right) \right]. \quad (5)$$

Log-concavity of f implies log-concavity of f_i , which in turn implies that (5) has a unique solution p_i^* . Moreover, $p_i^* > \sum_{j \in B_i} c_j$, since any price less than or equal to $\sum_{j \in B_i} c_j$ gives a weakly negative profit, whereas any price on the interval $(\sum_{j \in B_i} c_j, n_i \bar{\theta})$ gives a strictly positive profit. Instead, consider a deviation where the monopolist is selling all the goods in the bundle B_i separately at price $\frac{p_i^*}{n_i}$. By Theorem 1, f_i is strictly more peaked than the underlying density f , which since $\frac{p_i^*}{n_i} > \frac{\sum_{j \in B_i} c_j}{n_i} \geq \tilde{\theta}$ implies that $F_i \left(\frac{p_i^*}{n_i} \right) > F \left(\frac{p_i^*}{n} \right)$. Hence,

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

showing that unbundling the goods in B_i increases the profit for the monopolist. ■

While the assumptions obviously are much more restrictive, Proposition 3 provides a close analogue to Proposition 1. It may be noted that the “non-triviality assumption” $c_j < \bar{\theta}$ is analogous to the condition that $\Pi_i > \delta$ in Proposition 1, so that the only difference between the results 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:

Proposition 4 *Suppose that each θ_j is independently and identically distributed according to a symmetric log-concave density f which is strictly positive on support $[\underline{\theta}, \bar{\theta}]$ and has expectation $\tilde{\theta}$. Furthermore, assume that the unit cost is given by $c_j = c < \tilde{\theta}$ 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;*

1. *If $p^* \leq \tilde{\theta}$, all goods should be sold as a single bundle.⁷*
2. *If $p_B^* \geq K\tilde{\theta}$, all goods should be sold separately.*

While phrased in terms of endogenous prices, the comparative statics properties of monopoly pricing are rather straightforward. The monopoly price is increasing in the unit cost of production. Shifting the distribution of θ_j to the right or replacing F with a (log-concave symmetric distribution) G with the same mode that is strictly less peaked than F also leads to an increase in the monopoly price. Taken together, Propositions 3 and 4 therefore have some rather natural implications as to which type of goods we should expect to see bundled.

The reader may notice that Proposition 4 only considers the case where the unit costs are identical. The reason for this is simply that a disadvantage with bundling is lost flexibility in terms of adjusting the pricing to the unit costs of production when they vary across goods. That is, bundling two goods with different unit costs has negative consequences for productive efficiency. This is true also in the asymptotic analysis, but there the monopolist can extract almost the full surplus leading to the relatively clean condition (3), which applies even when costs are different for different goods. In the finite case, bundling does increase revenue when the average price is to the left of the mode of the distribution, but, if costs are different, the change in profit depends on a non-trivial trade-off between the increase in revenue and the loss in productive efficiency.

⁷A sufficient condition for $p^* \leq \tilde{\theta}$ is that $f(\tilde{\theta}) > \frac{1}{2(\tilde{\theta}-c)}$ (see Fang and Norman [8]).

Also note that the example in Section 5.1 demonstrates that log-concavity cannot be dropped from the statement of the result of Proposition 4; the example satisfies all conditions in the statement except log-concavity.⁸

Finally, it is worthwhile to point out that, even for the case with all goods being sold at the same unit cost, the characterization in Proposition 4 is incomplete. It is quite possible that $p^* > \tilde{\theta}$ and $p_B^* < K\tilde{\theta}$, in which case Proposition 4 is silent on whether bundling or separate sales maximizes the monopolist's profits. We return to this case in Section 5.2.2.

5.2.1 The 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.

Proof. (Part 1) Suppose for contradiction that there are least two bundles in the monopolist optimal bundling menu and label these B_1 and B_2 . Let n_1 and n_2 denote the number of goods in B_1 and B_2 respectively and let f_i and F_i denote the density and cumulative density of $\theta_i = \sum_{j \in B_i} \frac{\theta_j}{n_i}$ for $i = 1, 2$. Log-concavity is preserved under convolutions (Karlin [13]), implying that f_1 and f_2 are both symmetric log-concave densities with expectation $\tilde{\theta}$. This implies that the profit function

$$\left(p_i - \sum_{j \in B_i} c_j \right) \left[1 - F_i \left(\frac{p_i}{n_i} \right) \right] = (p_i - n_i c) \left[1 - F_i \left(\frac{p_i}{n_i} \right) \right] \quad (6)$$

is single-peaked in p_k for $k = 1, 2$.⁹ Let p_i^* denote the optimal price of bundle B_i . We observe;

Claim 1 $\frac{p_i^*}{n_i} \leq \tilde{\theta}$.

To see this, first note that if B_i contains a single good, the claim is immediate. For B_i containing more than a single good, note that, due to single-peakedness of the profit function, $\frac{p_i^*}{n_i} > \tilde{\theta}$ implies

⁸Logconcavity 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 in Section 5.1.

⁹To see this, let F be a distribution with log-concave density. Then, the profit function is increasing in p whenever

$$c - \left[p + \frac{1 - F(p)}{f(p)} \right] > 0$$

and decreasing when the inequality is reversed. Since log-concavity implies that $p + \frac{1 - F(p)}{f(p)}$ is strictly increasing, we conclude that the profit function is strictly single-peaked.

that

$$\begin{aligned} & \left. \frac{d}{dp_i} \right|_{p_i=n_i\tilde{\theta}} \left\{ (p_i - n_i c) \left[1 - F_i \left(\frac{p_i}{n_i} \right) \right] \right\} \\ &= 1 - F_i(\tilde{\theta}) - (\tilde{\theta} - c) f_i(\tilde{\theta}) = \frac{1}{2} - (\tilde{\theta} - c) f_i(\tilde{\theta}) > 0, \end{aligned} \quad (7)$$

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

$$\begin{aligned} & \left. \frac{d}{dp} \right|_{p_i=\tilde{\theta}} (p - c) [1 - F(p)] = \\ & 1 - F(\tilde{\theta}) - (\tilde{\theta} - c) f(\tilde{\theta}) = \frac{1}{2} - (\tilde{\theta} - c) f(\tilde{\theta}) \leq 0. \end{aligned} \quad (8)$$

But (7) and (8) together implies that $f_i(\tilde{\theta}) < f(\tilde{\theta})$. However, Theorem 1 implies that θ_i is strictly more peaked than the underlying distribution, which in turn implies that $f_i(\tilde{\theta}) > f(\tilde{\theta})$. Hence, the claim follows.

Now, consider a deviation where the monopolist sells all the goods in B_1 and B_2 as a single bundle. Label this bundle B_{1+2} . Furthermore, consider the (suboptimal if $\frac{p_1^*}{n_1} \neq \frac{p_2^*}{n_2}$) random pricing mechanism where

$$p_{1+2} = \begin{cases} \frac{n_1+n_2}{n_1} p_1^* & \text{with probability } \frac{n_1}{n_1+n_2} \\ \frac{n_1+n_2}{n_2} p_2^* & \text{with probability } \frac{n_2}{n_1+n_2} \end{cases} \quad (9)$$

Denote by F_{1+2} the cumulative of $\theta_{1+2} = \frac{\sum_{j \in B_1 \cup B_2} \theta_j}{n_1+n_2}$. The profit from sales of the bundle B_{1+2} is then

$$\begin{aligned} \Pi_{1+2} &= \frac{n_1}{n_1+n_2} \left[\frac{n_1+n_2}{n_1} p_1^* - (n_1+n_2)c \right] \Pr \left[\sum_{j \in B_1 \cup B_2} \theta_j \geq \frac{n_1+n_2}{n_1} p_1^* \right] \\ &+ \frac{n_2}{n_1+n_2} \left[\frac{n_1+n_2}{n_2} p_2^* - (n_1+n_2)c \right] \Pr \left[\sum_{j \in B_1 \cup B_2} \theta_j \geq \frac{n_1+n_2}{n_2} p_2^* \right] \\ &= (p_1^* - n_1 c) \left[1 - F_{1+2} \left(\frac{p_1^*}{n_1} \right) \right] + (p_2^* - n_1 c) \left[1 - F_{1+2} \left(\frac{p_2^*}{n_2} \right) \right]. \end{aligned} \quad (10)$$

Suppose first that $\frac{p_1^*}{n_1} < \tilde{\theta}$. Then, since F_{1+2} is strictly more peaked than F_1 , it follows that $F_{1+2} \left(\frac{p_1^*}{n_1} \right) < F_1 \left(\frac{p_1^*}{n_1} \right)$. Also, since $\frac{p_2^*}{n_2} \leq \tilde{\theta}$ we have that $F_{1+2} \left(\frac{p_2^*}{n_2} \right) \leq F_1 \left(\frac{p_2^*}{n_2} \right)$. Combining with (10) it is immediate that

$$\Pi_{1+2} > (p_1^* - n_1 c) \left[1 - F_1 \left(\frac{p_1^*}{n_1} \right) \right] + (p_2^* - n_1 c) \left[1 - F_2 \left(\frac{p_2^*}{n_2} \right) \right].$$

Hence, the bundle B_{1+2} generates a higher profit than the sum of the profits from B_1 and B_2 . The only remaining case is if $\frac{p_1^*}{n_1} = \frac{p_2^*}{n_2} = \tilde{\theta}$. In this case the profit from selling B_{1+2} at price

$p_1^* + p_2^* = (n_1 + n_2)\tilde{\theta}$ is the same as B_1 and B_2 as separate bundles. However, for $p_i^* = n_i\tilde{\theta}$ to be optimal, it is necessary that

$$\frac{d}{dp_i} \Big|_{p_i=n_i\tilde{\theta}} \left[(p_i - n_i c) \left[1 - F_i \left(\frac{p_i}{n_i} \right) \right] \right] = \frac{1}{2} - (\tilde{\theta} - c) f_i(\tilde{\theta}) = 0,$$

which implies that

$$\begin{aligned} & \frac{d}{dp} \Big|_{p=(n_1+n_2)\tilde{\theta}} \left[(p - (n_1 + n_2) c) \left[1 - F_i \left(\frac{p}{n_1 + n_2} \right) \right] \right] \\ &= \frac{1}{2} - (\tilde{\theta} - c) f_{1+2}(\tilde{\theta}) < 0 \end{aligned}$$

since f_{1+2} is strictly more peaked than f_1 and f_2 . Hence, a small decrease in the price leads to a profit from selling B_{1+2} that is strictly higher than the profits from selling B_1 and B_2 as separate bundles.

(Part 2) **Sketch.** The argument is the same as for Part 1, but in reverse. Using the same style of reasoning as in Claim 1 one establishes that any bundle B_i has to sell at price $p_i^* \geq n_i\tilde{\theta}$ if $p_B^* \geq K\tilde{\theta}$. Once this is established, the strict ordering in terms of relative peakedness is used to show that breaking up any bundle and selling the goods separately leads to a strict increase in profits if $p_i^* > n_i\tilde{\theta}$. For the case with $p_i^* = n_i\tilde{\theta}$ one shows that, if the goods are sold separately, the profit function must be strictly increasing in p at $\tilde{\theta}$. ■

5.2.2 The Case with $p^* > \tilde{\theta}$ and $p_B^* < K\tilde{\theta}$

As we have already observed, there is a hole in the characterization in Proposition 4. We end this Section by observing that, while we don't know whether separate sales or full bundling is better, at least we can show that the optimal bundling strategy must be one of these two.

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

Proof. Suppose not. Let the optimal bundling menu be given by $B = \{B_1, \dots, B_M\}$, where $2 \leq M \leq K - 1$. Without loss of generality assume that $n_1 \geq 2$ and $n_1 \geq n_2 \geq \dots \geq n_M$. We then observe that for B_1 to be part of the optimal bundling menu, it is necessary that $p_1^* \leq n_1\tilde{\theta}$. 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_1 separately at price $\frac{p_1^*}{n_1}$ each. Moreover, let $i^* \geq 1$ be the highest index such that $n_{i^*} \geq 2$. For the same reason as for bundle B_1 , $p_i^* \leq n_i \tilde{\theta}$ for every $i \leq i^*$. We can therefore apply the same argument as in Proposition 4 to argue that creating a single bundle of bundles B_1, \dots, B_{i^*} increases the profit for the monopolist. Hence the only remaining possibility is that there is one non-trivial bundle, which we will label B_1^* , and that the rest of the goods are sold separately.

Assume without loss that $B_1^* = \{1, \dots, k\}$ and that goods $k+1, \dots, K$ are sold separately. Let Π_1^* denote the profit from sales of the bundle B_1^* and p_1^* be the profit maximizing monopoly price for bundle B_1^* . Arguments above imply that $p_1^* \leq k\tilde{\theta}$. Let Π denote the maximized profit when a good is sold separately. Finally, let Π_B denote the maximized profit if all goods are bundled. Optimality of the bundling menu implies that

$$\begin{aligned} \Pi_1^* + (K - k)\Pi &\geq \Pi_B \\ \Pi_1^* + (K - k)\Pi &\geq K\Pi \end{aligned} \tag{11}$$

But applying the argument in Proposition 4, we can show that when $p_1^* \leq k\theta$,¹⁰

$$\frac{\Pi_B}{K} > \frac{\Pi_1^*}{k}.$$

Hence, if the first inequality in (11) holds, then

$$\begin{aligned} \Pi_1^* + (K - k)\Pi &\geq \Pi_B > \frac{K}{k}\Pi_1^* \Rightarrow \\ (K - k)\Pi &> \frac{K - k}{k}\Pi_1^* \Leftrightarrow \\ k\Pi &> \Pi_1^*, \end{aligned}$$

¹⁰By definition of Π_B , we have

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

where F_A is the CDF of $\sum_{j=1}^K \theta_j / K$. Thus, (by setting p_B above to $p_1^* K / k$),

$$\frac{\Pi_B}{K} \geq \left(\frac{p_1^*}{k} - c \right) \left[1 - F_A \left(\frac{p_1^*}{k} \right) \right].$$

From Theorem 1 F_A is strictly more single peaked than F_1 , the CDF of $\sum_{j=1}^k \theta_j / k$, and since $p_1^* / k \leq \tilde{\theta}$, we have

$$\left(\frac{p_1^*}{k} - c \right) \left[1 - F_A \left(\frac{p_1^*}{k} \right) \right] > \left(\frac{p_1^*}{k} - c \right) \left[1 - F_1 \left(\frac{p_1^*}{k} \right) \right] \equiv \frac{\Pi_1^*}{k}.$$

which contradicts the second inequality in (11). Symmetrically, if the second inequality in (11) is satisfied, that means that $\Pi \leq \frac{\Pi_1^*}{k}$, so

$$\Pi_1^* + (K - k)\Pi \leq \Pi_1^* + \frac{(K - k)}{k}\Pi_1^* = \frac{K}{k}\Pi_1^* < \Pi_B,$$

violating the first inequality in (11). ■

6 Discussion

Many papers on bundling, in particular in the more recent literature, take a “purist” mechanism design approach to the problem. These papers allow a monopolist to design selling mechanisms, which 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 paper, 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.

Other papers in the literature, notably McAfee et al [14], study a simpler problem, where the monopolist is restricted to simple pricing policies. While more restrictive than a full mechanism design approach, this is still less restrictive than our set of available instruments. Any given good can be sold *both* as part of a bundle and as a separate good, a practice that is referred to a *mixed bundling* in the literature. This obviously makes lots of sense: 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 price of the components. Hence, we do observe mixed bundling in the real world.

The main result in McAfee et al [14] is that, generically, the monopolist should always offer a non-trivial mixed bundling scheme. While obviously a powerful result, this has the arguably unpleasant implication that any monopolist selling more than a single good should offer to sell all their goods as part of a bundle. Put, differently, the result does not tell us about which kind of goods we should expect to see bundled. Moreover, the rather crude real world bundling schemes that we observe are rather puzzling in the light of McAfee et al [14]: the question as to why ESPN is available as a component of a bundle when championship boxing games tend to be available only on a pay-per-view basis cannot be answered.

Our approach, which is to assume that any good can only be part of a single bundle, is obviously ad hoc, but it gives us a reasonable hypothesis about the difference between ESPN and Championship boxing games. ESPN programming is presumably cheaper for the local cable provider, implying that it is more likely that it is a good where bundling leads to higher profits.

We do not pretend to have any general justification for ruling out mixed bundling. However, in some cases one can imagine technological reasons; there may simply be some fixed costs involved in the creation of a bundle. For example, in the context of bundling computer programs it doesn't seem farfetched to assume that selling components separately would require some extra programming work to guarantee compatibility of the components with older software that could be avoided if the new programs are bundled.

Another justification is that anti-trust law is explicitly expressed in terms of “anti-competitive mixed bundling”. While the legal interpretation of “mixed” is unclear it does seem somewhat reasonable that there may be instances where it is easier to get away with pure bundling than with mixed bundling.¹¹¹² In fact, in a recent case in the UK, the legal interpretation of mixed bundling seems to coincide with the terminology used in the Economics literature. The decision by the Office of Fair Trading [18] in the UK on alleged anti-competitive mixed bundling by the British Sky Broadcasting Limited explicitly states that: *mixed 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, therefore has no bite at all when bundling is of the form considered in this paper.

References

- [1] Adams, William J. and Janet L. Yellen, “Commodity Bundling and the Burden of Monopoly.” *Quarterly Journal of Economics*, 90, 1976, 475 –98.

¹¹In our model, there is no threat of entry for the monopolist. To the extent that “anti-competitive” refers to pricing strategies with the intent of keeping competitors out, our model is therefore more or less useless (see Nalebuff [16] for a more suitable model). However, consumer advocates arguing for introducing “a la carte” pricing for Cable TV stations are explicitly concerned about how bundling improves the possibilities for surplus extraction.

¹²The Microsoft case is a direct counterexample (the failure to provide the browser separately was used as evidence of anti-competitive behavior). However, had it 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.

- [2] An, Mark Yuying, “Logconcavity versus Logconvexity: A Complete Characterization.” *Journal of Economic Theory*, **80**, 1998, 350-369.
- [3] Armstrong, Mark, “Price Discrimination by a Many-Product Firm,” *Review of Economic Studies* **66**, 1999, 151-168.
- [4] Bagnoli, Mark, and Ted Bergstrom “Log-concave Probability and its Applications” *Economic Theory* 26(2), August 2005, 445 - 469 .
- [5] Bakos, Yannis, and Eric Brynjolfsson, “Bundling Information Goods: Pricing, Profits and Efficiency,” *Management Science* 45(12), December 1999, 1613-1630.
- [6] Birnbaum, Zygmunt W, “On Random Variables with Comparable Peakedness,” *Annals of Mathematical Statistics*, 19(1), March 1948, 76-81.
- [7] Crawford, Gregory, “The Discriminatory Incentives to Bundle in the Cable Television Industry,” *typescript*, University of Arizona, 2004.
- [8] Fang, Hanming and Peter Norman, “To Bundle or Not to Bundle,” Cowles Foundation Discussion Paper No. 1440, 2003.
- [9] Fang, Hanming and Peter Norman, “An Efficiency Rationale for Bundling of Public Goods,” *typescript*, University of Wisconsin-Madison and Yale University, November 2004.
- [10] Geng, Xianjun, Maxwell B. Stinchcombe and Andrew B. Whinston “Bundling Information Goods of Decreasing Value,” *Management Science* 51(4), April 2005, 662-667.
- [11] Ibragimov, Rustam, “Peakedness of Majorization Properties of Heavy Tail Distributions”, *typescript*, Yale University, 2004.
- [12] Ibragimov, Rustam, “Optimal Bundling Decisions for Complements and Substitutes with Heavy Tail Distributions”, *typescript*, Harvard University, 2005.
- [13] Karlin, Samuel, *Total Positivity*. Stanford University Press, Stanford CA, 1968.
- [14] McAfee, R. Preston, John McMillan and Michael D. Whinston, “Multiproduct Monopoly, Commodity Bundling, and Correlation of Values.” *Quarterly Journal of Economics*, 104, 1989, 371–84.

- [15] Myerson, Roger, “Optimal Auction Design,” *Mathematics of Operations Research* **6**, 1981, 58-73.
- [16] Nalebuff, Barry, “Bundling as an Entry Barrier,” *Quarterly Journal of Economics*, 119(1), 2004, 159-187.
- [17] Palfrey, Thomas, R. “Bundling Decisions by a Multiproduct Monopolist with Incomplete Information,” *Econometrica*, 51(2), March 1983, 463-484.
- [18] 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*, July 2003, London, UK. Available at:
<http://www.offt.gov.uk/business/competition+act/decisions/bskyb.htm>
- [19] Proschan, Frank, “Peakedness of Distributions of Convex Combinations,” *Annals of Mathematical Statistics*, 36(6), December 1965, 1703-1706.
- [20] Riley, John and Richard Zeckhauser, “Optimal Selling Strategies: When to Haggle, when to Hold Firm,” *Quarterly Journal of Economics*, **98**(2), May 1983, 267-289.
- [21] Schmalansee, Richard, “Commodity Bundling By Single-Product Monopolies,” *Journal of Law and Economics*, 15, April 1982.
- [22] Schmalansee, Richard, “Gaussian Demand and Commodity Bundling,” *Journal of Business*, 57(1), January 1984, s211-2230.

A Appendix A: Proofs of Results in Section 4.

Proof of Proposition 1.

Proof. 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 monopolists maximization problem as

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

Using Chebyshev's inequality,

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

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

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

where the term $\min \left\{ \frac{\sigma^2}{K\varepsilon^2}, 1 \right\}$ 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 $\frac{\sigma^2}{K\varepsilon^2} \leq 1$ if and only if $\varepsilon \geq \sqrt{\frac{\sigma^2}{K}}$, 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 $\max_{\varepsilon \geq 0} \varepsilon K \min \left\{ \frac{\sigma^2}{K\varepsilon^2}, 1 \right\} = \sqrt{\sigma^2 K}$. We conclude that $\Pi_B(K) - \sum_{j=1}^K \Pi_j \leq \sqrt{\sigma^2 K} - \delta K < 0$ for every $K > \frac{\sigma^2}{\delta^2}$. \blacksquare

Proof of Proposition 2.

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

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

Hence, the profit is at least $\left[1 - \frac{4\sigma^2}{\delta^2 K} \right] \left[\sum_{j=1}^K \mathbb{E}\theta_j - \frac{\delta K}{2} - \sum_{j=1}^K c_j \right]$, whereas the profit from separate sales (by assumption) is at most $\sum_{j=1}^K \mathbb{E}\theta_j - \sum_{j=1}^K c_j - \delta K$. Hence, the difference between the

bundling profit and the profit from separate sales is at least

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

Under the assumption that there exists μ such that $\sum_{j=1}^K \mathbb{E}\theta_j < \mu$ for every j the expression above is positive for K large enough. \blacksquare

B Appendix B: Calculations For Example 5.1

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 $p_B > 2 - \alpha$ 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

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

where 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 $p \in (1 - \frac{\alpha}{2}, 1)$. Since $\alpha \in [0, 1]$, $(1 - \frac{\alpha}{2}, 1)$ is a subset of $(\frac{1}{2}, 1)$. For any $\frac{1}{2} < p < 1$, the monopolist's profit from selling bundle at a bundle-price of $2p$ receives profit:

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

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

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

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; and if $\Delta(p)$ is negative, then the profit under bundling is higher.

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

$$\begin{aligned}\frac{d\Delta(p)}{dp} &= \alpha\left[2(1 - \alpha) + \alpha p - \frac{1}{2}\right] = \alpha\left[\frac{3}{2} + \alpha(p - 2)\right] \\ &> \alpha\left[\frac{3}{2} + \alpha\left(\frac{1}{2} - 2\right)\right] = \alpha\frac{3}{2}(1 - \alpha) > 0.\end{aligned}$$

Hence;

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

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

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

Hence,

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

We conclude:

Claim B3 $\Delta\left(1 - \frac{\alpha}{2}\right) > 0$ for α sufficiently small.

To sum up:

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

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