

To Bundle or Not To Bundle*

Hanming Fang[†]

Peter Norman[‡]

September 2003

Abstract

Commodity bundling is studied in an environment where the dispersion of valuations unambiguously decreases when two or more goods are sold as a bundle only. Bundling is more likely to dominate separately selling the goods if marginal costs are low relative to the average valuation, or if the distribution of valuations is very peaked around the mean.

*This note unbundles results previously provided in Fang and Norman [5]. We thank Mark Armstrong, Ted Bergstrom, Larry Samuelson, Bill Sandholm, Ananth Seshadri, and Stephen Morris for comments and helpful discussions. We also thank 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. The usual disclaimer applies.

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

[‡]Department of Economics, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, WI 53706-1393. Email: pnorman@facstaff.wisc.edu

1 Introduction

Bundling, the practice of selling two or more products as a package deal, rather than selling each product separately, 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]).

The economics literature makes a distinction between *mixed bundling* and *pure bundling*. Mixed bundling refers to a pricing strategy where commodities included in a bundle can also be purchased separately, whereas pure bundling is used to describe a situation where the commodities in a bundle are not offered for sale separately.

Despite several recent contributions to the literature, our understanding of why and when bundling is a useful screening device is still very incomplete. In the earliest literature, bundling was typically seen as a way to exploit negative correlation between reservation values for different goods (see Adams and Yellen [1] and Schmalensee [14]). Since then, it has been shown that bundling can be useful also when valuations for different goods are stochastically independent. In particular, McAfee et al [8] show that allowing a monopolist to bundle leads to a strict increase in profits relative to the case where goods must be sold separately, under a condition implied by stochastic independence.

McAfee et al [8] do not consider a full mechanism design problem. In their analysis, the monopolist sells two goods and can only set three prices: one for each good, and a price for the bundle. Since (by Myerson [9] and Riley and Zeckhauser [12]) the optimal selling mechanism when goods must be provided separately is a fixed price mechanism, their result implies that there must be an element of bundling also in an optimal mechanism. However, little is known about exactly how such an optimal selling mechanism operates. Armstrong [2] demonstrates that the participation constraints bind for a set of agents with positive mass (unlike the single-dimensional case, where it may bind for the minimal type only), and solves the problem fully for a few examples. Rochet and Chone [13] shows that (unlike regular single-dimensional problems) “bunching of types” is a robust feature of the solution to the design problem. Manelli and Vincent [7] provides some conditions under which randomizations can be ruled out, which in the case where goods are binary provides a justification for the approach in McAfee et al [8].

In this note, we take a different route. Rather than attempting to solve a truly multidimensional pricing problem, we simply rule out mixed bundling. That is, we assume that if good j is included in a particular bundle, then good j cannot be sold either separately or as part of any other bundle. While this may seem ad hoc, one justification of this approach is that it is probably easier to avoid being charged with “anti-competitive mixed bundling” if no good can be demonstrated to be priced in different ways depending on which bundle the consumer chooses.

The advantage of limiting our comparison to pure bundling versus separate provision is that we are able to highlight a clear intuition for what happens when two or more goods are sold as a bundle: the variance in willingness to pay is smaller for the bundled goods. This paper provides a partial characterization of when this reduction in variance is beneficial for the monopolist, and when it is not. We restrict attention to symmetric log-concave i.i.d. densities for the various goods, which allows us to compare the “peakedness” of distributions of different bundles. For this class of distributions we show, roughly, that goods with “thin markets” (which are goods with high marginal cost of production, high dispersion in valuations, and/or low average valuations) should not be bundled, whereas bundling is more likely to be optimal if the market is “thick”.

The intuition for the result is as follows: under the conditions for “thick markets”, the optimal unbundled price is to the left of the mode of the distribution of valuations. Since the average valuation of the goods in the bundle is more peaked than an individual distribution, the demand increases if the goods are bundled and sold at the sum of the unbundled prices. Symmetrically, under the conditions for a “thin market”, the bundled price is to the right of the mode. Unbundling the goods leads to higher demand in this case. Hence, unlike results for mixed bundling, according to which the bundling instrument should be used for all goods, our analysis has something to say about how use of the bundling instrument should relate to primitives of the model.

As far as we know, the only other paper that discusses bundling as an instrument to reduce the effective dispersion in buyers’ tastes is Schmalensee [15], who considers the case with normally distributed distributions of valuations. Relying mainly on numerical methods he reaches a similar conclusion, that is, that the difference between the average valuation and the marginal cost is a crucial determinant for whether (pure) bundling dominates separate sales.

2 The Model

The underlying economic environment is the same as in McAfee et al [8], except that we allow for more than two goods. A profit maximizing monopolist sells m indivisible products indexed by $j = 1, \dots, m$, and each good is produced at a constant marginal 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 $\theta = (\theta_1, \dots, \theta_m)$, 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

$$\sum_{j=1}^m \mathbb{I}_j \theta_j - p,$$

where p is the transfer from the consumer to the seller and \mathbb{I}_j is a dummy taking on value 1 if good j is consumed and 0 otherwise. We only consider the case where θ_j and $\theta_{j'}$ are stochastically independent for each pair of goods j, j' .

2.1 The Problem

The obvious design problem for this environment is to characterize a profit maximizing selling mechanism subject to incentive compatibility and participation constraints. There have been many attempts to tackle this problem, but with limited success. The main analytical difficulty is that, because of the multidimensional typespace, there is no known methodology to transform the constraints so as to generate a tractable programming problem. It is true that if random selling mechanisms are ruled out, the problem (with binary goods) boils down to constructing a *simple pricing policy*, setting a price for each bundle and letting the consumer self-select. This is a relatively straightforward problem, but very little is known about when randomizations can be ruled out (see Manelli and Vincent [7]).

We will not attempt to solve the general design problem in this paper. In fact, we will not even consider optimal simple pricing policies. Instead, we will assume that if the monopolist sells, say, good 1 and 2 as a bundle, then the monopolist cannot at the same time sell either good 1 or good 2 as a separate good. For simplicity of language we will refer to this as *pure bundling*, even in cases where only a subset of the goods are sold as a bundle, and consider *mixed bundling* to be any selling mechanism where at least one good can be obtained in at least two different “packages”.

It may seem arbitrary to rule out mixed bundling, but anti-trust law is explicitly expressed in terms of “anti-competitive mixed bundling”. While the legal interpretation of “mixed” is unclear,

it seems reasonable to assume that it is easier to get away with pure bundling since there is then no way to demonstrate that the bundling is used as an instrument to price discriminate.¹

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, it says nothing about the probability of a given size deviation from the mean being monotonically decreasing in the sample size. Indeed, such monotone convergence fails in general, and can only be guaranteed under some rather tight restrictions on the underlying probability distribution.

To discuss such monotonicity it is useful to have some notion of “relative peakedness” of two distributions. We use a definition from Birnbaum [4]:

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 .*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, extreme value, exponential, Laplace, Weibull, and many other common parametric densities (see Bagnoli and Bergstrom [3] for further examples). Comparative peakedness of convex combinations

¹The decision by the Office of Fair Trading [10] in the UK on alleged anti-competitive mixed bundling by the British Sky Broadcasting Limited provides a clear definition of mixed bundling, which is consistent with the terminology in economics. It is explicitly stated 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.* The legal test is simply to compare marginal prices, which cannot be done in the case of pure bundling. In contrast, the interpretation of “mixed” in US law seems unclear. Microsoft was accused of “anti-competitive mixed bundling” for bundling Internet Explorer with Windows in *United States vs Microsoft*. In standard economics terminology, this would arguably be considered as pure bundling, because neither the Browser or the operation system was available separately at the time.

²Strictly speaking, Birnbaum [4] 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 [11] and drop the qualifiers.

of log-concave random variables are studied in Proschan [11], and we will apply one of his results in this paper. To avoid discussing majorizations of vectors we will not state his main result. Instead, we will directly use one of his lemmas:

Theorem 1 (Lemma 2.2 in Proschan [11]) *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 check that unimodality (which is implied by log-concavity) is necessary for Theorem 1. However, unimodality is not sufficient. To get some intuition for why log-concavity is needed, note that a *necessary* condition for the average to be more peaked than the underlying distribution is that the density at the median is greater for the average than for the underlying (symmetric) distribution.⁴ However, if the underlying distribution is flat, except for a spike at the mode, then this condition will not be satisfied (see the example on page 9). Hence, some condition is needed to rule out too abrupt changes in the density, and log-concavity does exactly that. The need for symmetry is due to the fact that the location of the peak of the density would otherwise be changing depending on the weights.

³To see this. First use the Theorem 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 more 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.

⁴More specifically, the cumulative probabilities at the median (which is also the mode and the expected value for symmetric distributions) are 1/2 for both the average and the underlying individual distributions. For the cumulative probabilities of the average to be strictly lower given than the underlying distribution at any point below the median, the probability density must be larger for the average at the median.

4 To Bundle or Not to Bundle

To be able to apply Theorem 1 we now assume that each θ_i is independently distributed according to a symmetric log-concave distribution f with expectation $\tilde{\theta} > 0$. Since we rule out any form of mixed bundling, the problem for the monopolist can be separated in two parts; i) decide which goods, if any, should be sold together as a bundle, and ii) decide on how to price each bundle.

We describe the way the monopolist packages goods together by a *bundling menu*, which is simply a partition of the set of goods produced by the monopolist. Such a menu is denoted by $B = \{B_1, \dots, B_K\}$, where each $B_k \in B$ is a subset of $\{1, \dots, m\}$ and where $B_k \cap B_{k'} = \emptyset$ for each $k \neq k'$, and where $1 \leq K \leq m$ is the number of bundles sold by the monopolist.

For any fixed bundling menu, the optimal selling mechanism is known. There is no instrument left to price discriminate between two types θ, θ' such that $\sum_{j \in B_k} \theta_j = \sum_{j \in B_k} \theta'_j$. Hence, the problem is unidimensional, and it is then well-known that restricting the monopolist to sell by a fixed price mechanism is without loss of generality (Myerson [9], Riley and Zeckhauser [12]).

Our main result is:

Proposition 1 *Suppose that each θ_j is independently and identically distributed according to a log-concave symmetric density f with expectation $\tilde{\theta}$ and support $[\underline{\theta}, \bar{\theta}]$, and that each good j is produced with constant marginal cost c_j . Then:*

1. *If $\tilde{\theta} \leq c_j < \bar{\theta}$ for each j , the profit maximizing bundle menu is $\{\{1\}, \{2\}, \dots, \{m\}\}$, that is, goods are sold separately;*
2. *If there is some c such that $c_j = c < \tilde{\theta}$ for all j and if $f(\tilde{\theta}) > \frac{1}{2(\tilde{\theta}-c)}$, then the profit maximizing bundle menu is $\{\{1, \dots, m\}\}$, that is, all goods are sold together as a single bundle.*

Proof. (Part 1) Suppose for contradiction that there is at least one bundle consisting of more than a single good in the profit maximizing bundle menu. Call that bundle B_k , let n_k be the number of goods in B_k , and let f_k denote the density for the random variable $\theta_k = \sum_{j \in B_k} \theta_j / n_k$ and denote by F_k the associated cumulative distribution. The optimal price to set for the bundle B_k is the solution to

$$\max_{p_k} \left(p_k - \sum_{j \in B_k} c_j \right) \Pr \left[\sum_{j \in B_k} \theta_j \geq p_k \right] = \max_{p_k} \left(p_k - \sum_{j \in B_k} c_j \right) \left[1 - F_k \left(\frac{p_k}{n_k} \right) \right] \quad (1)$$

It is immediate that the profit maximizing price, p_k^* , satisfies $p_k^* > \sum_{j \in B_k} c_j$ since any price $p_k \leq \sum_{j \in B_k} c_j$ generates a loss, whereas the price $p_k = \frac{n_k \tilde{\theta} + \sum_{j \in B_k} c_j}{2} > \sum_{j \in B_k} c_j$ generates a strict profit. The profit from selling B_k at price p_k^* is

$$\left(p_k^* - \sum_{j \in B_k} c_j \right) \left[1 - F_k \left(\frac{p_k^*}{n_k} \right) \right]. \quad (2)$$

Instead, consider a deviation where the monopolist sells all goods in B_k separately, charging a price $\frac{p_k^*}{n_k}$ for each good. Under this selling strategy the goods in B_k generates a profit

$$\sum_{j \in B_k} \left(\frac{p_k^*}{n_k} - c_j \right) \left[1 - F \left(\frac{p_k^*}{n_k} \right) \right] = \left(p_k^* - \sum_{j \in B_k} c_j \right) \left[1 - F \left(\frac{p_k^*}{n_k} \right) \right]. \quad (3)$$

From Theorem 1 we know that f_k is strictly more peaked than f , which since $\frac{p_k^*}{n_k} > \tilde{\theta}$ implies that $F_k \left(\frac{p_k^*}{n_k} \right) > F \left(\frac{p_k^*}{n_k} \right)$. The profit in (3) is thus strictly larger than in (2). Since B_k was an arbitrarily chosen non-trivial bundle, the result follows.

(Part 2) Suppose for contradiction that the monopolist is selling at least two bundles. Pick two arbitrary bundles, B_1 and B_2 , let n_1 and n_2 denote the cardinality of B_1 and B_2 , let f_k denote the density of $\theta_k = \sum_{j \in B_k} \theta_j / n_k$, let and F_k be the associated cumulative distribution for $k = 1, 2$. Since log-concavity is preserved under convolutions (see Karlin [6]), f_1 and f_2 are both symmetric log-concave densities with expectation $\tilde{\theta}$. This implies that the profit function

$$\left(p_k - \sum_{j \in B_k} c_j \right) \left[1 - F_k \left(\frac{p_k}{n_k} \right) \right] = (p_k - n_k c) \left[1 - F_k \left(\frac{p_k}{n_k} \right) \right] \quad (4)$$

is single-peaked in p_k for $k = 1, 2$.⁵

STEP 1: We will first establish that $p_k^* < n_k \tilde{\theta}$ for $k = 1, 2$ under the condition $f(\tilde{\theta}) \geq \frac{1}{2(\tilde{\theta} - c)}$. To see this, suppose that $p_k^* \geq n_k \tilde{\theta}$. By the single-peakedness it must then be that the profit function

⁵To 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.

is non-decreasing at $n_k \tilde{\theta}$, that is

$$\begin{aligned}
& \left. \frac{d}{dp_k} \right|_{p_k=n_k \tilde{\theta}} \left\{ (p_k - n_k c) \left[1 - F_k \left(\frac{p_k}{n_k} \right) \right] \right\} \\
&= \left[1 - F_k(\tilde{\theta}) \right] - \left[n_k \tilde{\theta} - n_k c \right] f_k(\tilde{\theta}) \frac{1}{n_k} \\
&= \frac{1}{2} - (\tilde{\theta} - c) f_k(\tilde{\theta}) \geq 0 \Leftrightarrow (\tilde{\theta} - c) \leq \frac{1}{2f_k(\tilde{\theta})}
\end{aligned} \tag{5}$$

But, $f_k(\tilde{\theta}) \geq f(\tilde{\theta})$ (otherwise there would be some $\varepsilon > 0$ such that $F_k(\theta) > F(\theta)$ for $\theta \in (\tilde{\theta} - \varepsilon, \tilde{\theta})$ and $F_k(\theta) < F(\theta)$ for $\theta \in (\tilde{\theta}, \tilde{\theta} + \varepsilon)$, which would contradict the fact that f_k is more peaked than f). Hence, (5) implies that $(\tilde{\theta} - c) \leq \frac{1}{2f(\tilde{\theta})}$, which contradicts the hypothesis of the Proposition.

STEP 2: Consider the deviation where the monopolist sells all the goods in B_1 and B_2 as a single bundle, which we label B_{1+2} , and uses (a suboptimal) random pricing mechanism where

$$p_{1+2} = \begin{cases} \frac{n_1+n_2}{n_1} p_1^* & \text{with probability } \frac{n_1}{n_1+n_2} \\ \frac{n_1+n_2}{n_2} p_2^* & \text{with probability } \frac{n_2}{n_1+n_2} \end{cases} \tag{6}$$

Let the cumulative for the random variable $\theta_{1+2} = \frac{\sum_{j \in B_2} \theta_j + \sum_{j \in B_1} \theta_j}{n_1+n_2}$ be denoted by F_{1+2} and write the profit for the monopolist under the deviation as

$$\begin{aligned}
\Pi' &= \frac{n_1}{n_1+n_2} \left[\frac{n_1+n_2}{n_1} p_1^* - (n_1+n_2)c \right] \Pr \left[\sum_{j \in B_1 \cup B_2} \theta_j \geq \frac{n_1+n_2}{n_1} p_1^* \right] \\
&+ \frac{n_2}{n_1+n_2} \left[\frac{n_1+n_2}{n_2} p_2^* - (n_1+n_2)c \right] \Pr \left[\sum_{j \in B_1 \cup B_2} \theta_j \geq \frac{n_1+n_2}{n_2} p_2^* \right] \\
&= [p_1^* - n_1 c] \left[1 - F_{1+2} \left(\frac{p_1^*}{n_1} \right) \right] + [p_2^* - n_2 c] \left[1 - F_{1+2} \left(\frac{p_2^*}{n_2} \right) \right],
\end{aligned} \tag{7}$$

which is to be compared with the profit under the original bundling menu

$$\Pi = [p_1^* - n_1 c] \left[1 - F_1 \left(\frac{p_1^*}{n_1} \right) \right] + [p_2^* - n_2 c] \left[1 - F_2 \left(\frac{p_2^*}{n_2} \right) \right] \tag{8}$$

Since $\frac{p_k^*}{n_k} < \tilde{\theta}$ for $k = 1, 2$ and since f_{1+2} is more peaked than either f_1 or f_2 we have that

$$\begin{aligned}
1 - F_{1+2} \left(\frac{p_1^*}{n_1} \right) &> 1 - F_1 \left(\frac{p_1^*}{n_1} \right) \\
1 - F_{1+2} \left(\frac{p_2^*}{n_2} \right) &> 1 - F_2 \left(\frac{p_2^*}{n_2} \right).
\end{aligned} \tag{9}$$

Hence,

$$\begin{aligned}
\Pi' &= [p_1^* - n_1c] \left[1 - F_{1+2} \left(\frac{p_1^*}{n_1} \right) \right] + [p_2^* - n_2c] \left[1 - F_{1+2} \left(\frac{p_2^*}{n_2} \right) \right] & (10) \\
/(9)/ &> [p_1^* - n_1c] \left[1 - F_1 \left(\frac{p_1^*}{n_1} \right) \right] + [p_2^* - n_2c] \left[1 - F_2 \left(\frac{p_2^*}{n_2} \right) \right] \\
/(8)/ &= \Pi,
\end{aligned}$$

which implies that the random mechanism that sells the goods in B_1 and B_2 as a single bundle generates a higher profit. Since B_1 and B_2 were arbitrary bundles, the result follows. ■

While Proposition 1 is far from a complete characterization of the optimal bundling strategy, the conditions are easily interpretable. Roughly speaking, the result says that the monopolist should sell goods that are sold to many consumers (either because they are cheap to produce, and/or the average willingness to pay is high, and/or the density at $\tilde{\theta}$ is high) as a bundle, whereas goods with a thin market should be sold separately.

Notice that Part 1 of Proposition 1 is not a contradiction of the finding in McAfee et al [8] that the bundling instrument is always useful if valuations are independent. They show that a mixed bundling mechanism, that is a mechanism where goods are sold both separately and in bundles, dominates unbundled sales. Our result is consistent with theirs because we are explicitly ruling out mixed bundling. Also notice that, since they use a local argument, McAfee et al [8] does not rule out pure bundling as the best “mixed bundling mechanism.” Part 1 of Proposition 1 therefore provides conditions under which the optimal mechanism involves non-degenerate mixing in McAfee et al [8]’s environment where mixed bundling is allowed.

5 Discussion and Examples

5.1 The Necessity of Log-concavity

This note is based on the simple idea that bundling tends to reduce the asymmetric information between the seller and the buyer. Intuition then suggests that, if the price is to the left of the mode, then the monopolist will increase sales if bundling two goods at twice the price of the optimal separate sales price. Theorem 1 implies that this is indeed the case for symmetric log-concave distributions, and we will here provide an example to show the necessity log-concavity.

Consider the case with two goods, and suppose that for $j = 1, 2$, with probability $\alpha \in (0, 1)$, θ_j is uniform $[0, 2]$, and with probability $1 - \alpha$, θ_j takes value 1. This distribution satisfies all

conditions in Theorem 1 except log-concavity. The cumulative of θ_j is

$$F(\theta_i^j) = \begin{cases} \frac{\alpha}{2}\theta_i^j & \text{for } \theta_i^j < 1 \\ (1 - \alpha) + \frac{\alpha}{2}\theta_i^j & \theta_i^j \geq 1 \end{cases} .$$

The cumulative of $(\theta_1 + \theta_2)/2$ is discontinuous at 1, and the probability that it takes exact value of 1 (its mean) is $(1 - \alpha)^2$, smaller than the probability that θ_j takes on value 1 which is $1 - \alpha$. Write F^A as the cumulative distribution of the average valuation. Since the jump at the discontinuity is smaller for the average, and the density is symmetric, we conclude that there exists $\epsilon > 0$ such that $F^A(y) > F(y)$ for any $y \in (1 - \epsilon, 1)$. Thus it is possible that, if the two goods are bundled at a price in this interval, there would be more exclusions with bundling than under separate provision. Also note that, since F^A is strictly above F over the interval $(1 - \epsilon, 1)$, we can obtain the same result also using continuous probability distributions.

5.2 The Private Good Assumption

A striking implication of the first part of Proposition 1 is that even if we consider a large number of identical goods, the monopolist may still sell them all separately. This would be true also if the monopolist were some regulated welfare maximizing entity. In contrast to this, Fang and Norman [5] show that if a welfare maximizer provides (excludable) non-rival goods, then a pure bundling mechanism can approximate first best if the numbers of goods and consumers are both large. A profit maximizing monopolist could extract almost the entire surplus from the consumers, so with a large number of excludable non-rival goods (and independent valuations), a profit maximizer would always want to sell all goods as a bundle *no matter what the distributions of willingness to pay are*.

These differences are instructive. It is always true that bundling a large number of goods virtually eliminates the asymmetric information about the willingness to pay for the bundle. The difference is that, in the private goods case, it is wasteful for society to give goods to agents with very low valuations, whereas in the public good case, the more people that enjoy the good the better. The case for bundling as an instrument to reduce the dispersion in willingness to pay is thus more clear-cut in the public goods case (studied in Fang and Norman [5]), or, more generally, in natural monopoly situations.

References

- [1] Adams, W.J. and Janet L. Yellen, “Commodity Bundling and the Burden of Monopoly.” *Quarterly Journal of Economics*, 90, 1976, 475 –98.
- [2] Armstrong, Mark, “Multiproduct Non-Linear Pricing,” *Econometrica*, 64, 51-76, 1996.
- [3] Bagnoli, Mark, and Ted Bergstrom “Log-concave Probability and its Applications” *unpublished manuscript*, 1989.
- [4] Birnbaum, Z. W., “On Random Variables with Comparable Peakedness,” *Annals of Mathematical Statistics*, 19(1), March 1948, 76-81.
- [5] Fang, Hanming and Peter Norman, “An Efficiency Rationale for Bundling of Public Goods,” *typescript*, University of Wisconsin and Yale, September 2003.
- [6] Karlin, S. “Total Positivity,” Stanford University Press, Stanford CA, 1968.
- [7] Manelli, Alejandro M. and Daniel R. Vincent, “Bundling as an Optimal Mechanism by a Multiple-Good Monopolist”, *typescript*, University of Maryland, October, 2002.
- [8] 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.
- [9] Myerson, Roger, “Optimal Auction Design,” *Mathematics of Operations Research* **6**, 1981, 58-73.
- [10] 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>
- [11] Proschan, Frank, “Peakedness of Distributions of Convex Combinations,” *Annals of Mathematical Statistics*, 36(6), December 1965, 1703-1706.
- [12] 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.

- [13] Rochet, Jean-Charles and Philippe Chone, “Ironing, Sweeping and Multidimensional Screening,” *Econometrica*, 66(4), 783-826, July 1998.
- [14] Schmalansee, Richard, “Commodity Bundling By Single-Product Monopolies,” *Journal of Law and Economics*, 15, April 1982.
- [15] Schmalansee, Richard, “Gaussian Demand and Commodity Bundling,” *Journal of Business*, 57(1), January 1984, s211-2230.