Competitive Bundling*

Jidong Zhou
Department of Economics
NYU Stern School of Business

December 2014
(First Version: February 2014)

Abstract

This paper proposes a model of competitive bundling with an arbitrary number of firms. In the regime of pure bundling, we find that under fairly general conditions, relative to separate sales pure bundling raises market prices, improves profits and harms consumers when the number of firms is above a threshold. This is in contrast to the findings in the duopoly case on which the existing literature often focuses. In the regime of mixed bundling, having more than two firms raises new challenges in solving the model. We derive the equilibrium conditions, and we show that when the number of firms is large, the equilibrium prices have simple approximations and mixed bundling is generally pro-competitive relative to separate sales. Firms' incentives to bundle are also investigated in both regimes.

Keywords: bundling, product compatibility, oligopoly, multiproduct pricing
JEL classification: D43, L13, L15

1 Introduction

Bundling is commonplace in the market. Sometimes firms sell their products in packages only and no individual products are available for purchase. For example, in the market for CDs, newspapers, books, or TV packages (e.g., in the US), firms

*I thank Mark Armstrong, Heski Bar-Isaac, Luis Cabral, Jay Pil Choi, Xavier Gabaix, Phil Haile, Alessandro Lizzeri, Barry Nalebuff, Andrew Rhodes, Ran Spiegler, Gabor Virag, Glen Weyl and seminar participants in Duke-UNC, HKUST, Michigan, NYU Stern, SHUFE, Toronto, WZB in Berlin, Yale SOM, and the 13th Columbia/Duke/Northwestern/MIT IO Theory Conference for their helpful comments. The research assistance by Diego Daruich is greatly appreciated.

usually do not sell songs, articles, chapters, or TV channels separately.[1] This is called *pure bundling*. Sometimes firms offer both the package and individual products, but the package is offered at a discount relative to its components. Relevant examples include software suites, TV-internet-phone, season tickets, package tours, and value meals. This is called *mixed bundling*. In many cases, bundling occurs in markets where firms compete with each other.

With competition, bundling has a broader interpretation. For example, pure bundling can be the outcome of product incompatibility. Consider a system (e.g., a computer, a stereo system, a smartphone) that consists of several components (e.g., hardware and software, receiver and speaker). If firms make their components incompatible with each other (e.g., by not adopting a common standard, or by making it very costly to disassemble the system), then consumers have to buy the whole system from a single firm and cannot mix and match to assemble a new system by themselves.[2] Bundling can also be the consequence of shopping costs. If consumers need to incur an extra cost to visit more than one grocery store, for instance, then buying all desired products from a single store can save from the extra shopping cost, which is like buying the package to enjoy the discount in mixed bundling. And if the shopping cost is sufficiently high, the situation will be like pure bundling.

The obvious motivation for bundling is economies of scale in production, selling or buying, or complementarity in consumption. For example, in the traditional market it is perhaps too costly to sell newspaper articles separately. There are other important reasons for bundling. For instance, pure bundling can reduce consumer valuation heterogeneity and so facilitate firms extracting consumer surplus (Stigler, 1968). Mixed bundling can be a profitable price discrimination device by offering purchase options to screen consumers (Adams and Yellen, 1976). Bundling can also be used as a leverage device by a multiproduct firm to deter the entry of potential competitors or induce the exit of existing competitors (Whinston, 1990, and Nalebuff, 2004).[3]

The main anti-trust concern about bundling is that it may restrict market competition. One possible reason is that as suggested by the leverage theory bundling

[1] The situation, however, is changing with the development of the online market. For instance, consumers nowadays can download single songs from iTune or Amazon. Some websites like www.CengageBrain.com sell e-chapters of textbooks, and individual electronic articles in many academic journals are also available for purchase.

[2] This is actually the leading interpretation taken by early works on competitive pure bundling such as Matutes and Regibeau (1988).

[3] See also Choi and Stefanadis (2001), and Carlton and Waldman (2002). Without changing the market structure, bundling by a multiproduct firm may also help segment the market and relax the price competition with a single product firm (Carbajo, de Meza and Seidmann, 1990, and Chen, 1997).

can change the market structure and make it more concentrated. Another possible reason is that even if the market structure is given, bundling may relax competition among firms and inflate market prices. One purpose of this paper is to revisit the second issue in a more general setup than the existing literature. The economics literature has extensively studied bundling in the monopoly case.[4] There is also research on competitive bundling. However, the existing works on competitive bundling often focus on the case with two firms and each selling two products (see, e.g., Matutes and Regibeau, 1988, for pure bundling, and Matutes and Regibeau, 1992, and Armstrong and Vickers, 2010, for mixed bundling).[5]

In reality, there are many markets where more than two firms compete with each other and adopt (pure or mixed) bundling strategies. (For example, the companies that offer the TV-internet-phone service in New York City include at least Verizon, AT&T, Time Warner and RCN.) The literature, however, has not developed a general model which is suitable for studying both pure and mixed bundling with an arbitrary number of firms. This has limited our understanding of how the degree of market concentration might affect firms' incentives to bundle and the impact of bundling on market performance relative to separate sales (or the impact of banning bundling). This paper aims to provide such a framework for studying competitive bundling. We will show that considering more than two firms can shed new light on firms' incentives to bundle and how bundling affects firms and consumers. For example, in the pure bundling case, having more than two firms can reverse the impact of bundling relative to separate sales. This suggests that the insights we have learned from the existing duopoly models can be incomplete, and the number of firms qualitatively matters for the welfare assessment of bundling.

The existing papers on competitive bundling with two firms and two products use the two-dimensional Hotelling model where consumers are distributed on a square and firms are located at two opposite corners. With more than two firms, it is no longer convenient to model product differentiation in a spatial framework, especially when there are also more than two products.[6] In this paper, we adopt

[4]For example, Schmalensee (1984) and Fang and Norman (2006) study the profitability of pure bundling relative to separate sales, and Adams and Yellen (1976), Long (1984), McAfee, McMillan, and Whinston (1989), and Chen and Riordan (2013) study the profitability of mixed bundling.

[5]Two notable exceptions are Economides (1989), and Kim and Choi (2014). Both papers study pure bundling in the context of product compatibility when there are more than two firms, each selling two products. We will discuss the detailed relationship with them in section 3.6. Some recent empirical works on bundling also deal with the case with more than two firms. See, e.g., Crawford and Yurukoglu (2012) for bundling in the cable TV industry, and Ho, Ho, and Mortimer (2012) for bundling in the video rental industry. They focus on how the interaction between bundling and the vertical market structure might affect market performance. This is an interesting dimension ignored by the existing theory literature on bundling.

[6]Introducing differentiation at the product level is important for studying competitive bundling

the random utility framework in Perloff and Salop (1985) to model product differentiation. Specifically, a consumer's valuation for a firm's product is a random draw from some distribution, and its realization is independent across firms and consumers. This reflects, for example, the idea that firms sell products with different styles and consumers often have idiosyncratic tastes. This framework is flexible enough to accommodate any number of firms and products, and in the case with two firms and two products it can be rephrased into a two-dimensional Hotelling model.

Section 2 presents the model and analyzes the benchmark case of separate sales. With separate sales, firms compete on each product separately in our setting and the situation is like we have several separate Perloff-Salop models. We study two comparative static questions. First, we show that a standard logconcavity condition (which ensures the existence of pure strategy pricing equilibrium) guarantees that market prices decline with the number of firms. This is true even if we relax the usual assumption of full market coverage. Second, we argue that when the number of firms is large, the tail behavior, instead of the peakedness, of the valuation distribution determines the equilibrium price. In particular, less dispersed consumer valuations can lead to higher market prices when the number of firms is sufficiently large. Both results have their own interest in the price competition literature, and the latter result provides the foundation for the price comparison result in the pure bundling case.

Section 3 studies competitive pure bundling. We show that in the duopoly case pure bundling intensifies competition and leads to lower prices and profits compared to separate sales. This generalizes the result in the existing literature which assumes two products and also often a particular consumer valuation distribution. Moreover, for consumers this positive price effect typically outweighs the negative match quality effect (which is caused by the loss of opportunities to mix and match), such that bundling tends to benefit consumers. However, under fairly general conditions, the results are reversed (i.e., pure bundling raises prices, benefits firms and harms consumers) when the number of firms is above some threshold (which can be small).

The intuition consists of two arguments. First, bundling reduces the dispersion of consumer valuation (in terms of per-product valuation for the bundle). Compared to the density function of the valuation for a single product, the density function of the per-product valuation for the bundle is more peaked but with a thinner tail. Intuitively, this is because finding a very well matched bundle is much harder than finding a very well matched single product. Second, a firm's pricing decision hinges

if firms have the same cost conditions. If there is no product differentiation, prices will settle at the marginal costs anyway and so there will be no meaningful scope for bundling. If differentiation is only at the firm level, consumers will one-stop shop even without bundling, which is not realistic in many markets and also makes the study of competitive bundling less interesting.

on the number of its marginal consumers who are indifferent between its product and the best product among its competitors, and it determines demand elasticity. When there are a large number of firms, a firm's marginal consumers should have a high valuation for its product because their valuation for the best rival product is high. In other words, they tend to position on the right tail of the valuation density. Since bundling generates a thinner tail than separate sales, it leads to fewer marginal consumers and so a less elastic demand. This induces firms to raise their prices. But when there are relatively few firms in the market, the average position of marginal consumer is close to the mean valuation. Then bundling leads to more marginal consumers and so a more elastic demand. This induces firms to reduce their prices.

We also study firms' incentives to bundle. When firms can choose between separate sales and pure bundling, it is always a Nash equilibrium that all firms bundle if consumers buy all products. (This is simply because if one firm unilaterally unbundles, the market situation does not change.) In the duopoly case, we further show that this is the unique equilibrium. However, when the number of firms is above some threshold, separate sales can be an equilibrium as well. In many examples separate sales is another equilibrium if and only if consumers prefer separate sales to pure bundling. In the end of Section 3, we extend the pure bundling analysis in various ways which include asymmetric products and correlated valuations, a market without full market coverage, and elastic demand. There we also discuss in details the relationship with a few closely related papers on competitive pure bundling.

Section 4 studies competitive mixed bundling. According to our knowledge, all the existing papers on competitive mixed bundling deal with the duopoly case. This paper is the first to consider the case with more than two firms. We first show that in our setting with full market coverage, starting from separate sales each firm has a strict incentive to introduce mixed bundling. This implies that when mixed bundling is feasible and costless to implement, separate sales cannot be an equilibrium outcome. Solving the pricing game with mixed bundling is significantly harder when there are more than two firms. However, we are able to characterize the equilibrium conditions, and we can also show that under mild conditions the equilibrium prices have simple approximations when the number of firms is large. For example, when the number of firms is large and the production cost is zero, the joint-purchase discount will be approximately half of the stand-alone price (i.e., 50% off for the second product). In terms of the impacts of mixed bundling on profits and consumer surplus, it is usually ambiguous in the duopoly case. But when there are a large number of firms, under mild conditions mixed bundling benefits consumers and harms firms. We conclude in Section 5, and all omitted proofs are presented in the Appendix.

2 The Model

Consider a market where each consumer needs $m \geq 2$ products. (They can be m independent products, or m components of a system, depending on the interpretation we will take below for bundling.) The measure of consumers is normalized to one. There are $n \geq 2$ firms, and each firm supplies all the m products. The unit production cost of any product is normalized to zero (so we can regard the price below as the markup).[7] Each product is horizontally differentiated across firms (e.g., each firm produces a different version of the product). We adopt the random utility framework in Perloff and Salop (1985) to model product differentiation. Let $x_{i,k}^j$ denote the match utility of firm j's product i for consumer k. We assume that $x_{i,k}^j$ is i.i.d. across consumers, which reflects, for instance, idiosyncratic consumer tastes. So in the following we suppress the subscript k. We consider a setting with symmetric firms and products: x_i^j is distributed according to a common cdf F with support $[\underline{x}, \overline{x}]$ (where $\underline{x} = -\infty$ and $\overline{x} = \infty$ are allowed), and it is realized independently across firms and products. Suppose the corresponding pdf f is continuous, and x_i^j has a finite mean $\mu < \infty$. In section 3.5.1, we will consider a more general setting where the m products in each firm can be asymmetric and have correlated match utilities.

We consider a discrete-choice framework where each consumer only buys one version of each product, i.e., the incremental utility from having more than one version of a product is zero.[8] Moreover, in the basic model we also assume that consumers have unit demand for the version of a product which she wants to buy. (Elastic demand will be considered in section 3.5.3.) If a consumer consumes m products with match utilities $(x_1, \cdots, x_m)$ (which can be from different firms if consumers are not restricted to buy all products from a single firm) and makes a total payment T, she obtains surplus $\sum_{i=1}^m x_i - T$.[9]

The space of firm pricing strategies differs across the regimes we are going to investigate. In the benchmark regime of separate sales, each firm sells its products

[7] This paper assumes away the possible effect of bundling on production costs.

[8] This assumption is made in all the papers on competitive (pure or mixed) bundling. But clearly it is not innocuous. For example, reading another different newspaper article on the same story, or reading another chapter on the same topic in a different textbook usually improves utility, unless it is too costly for the reader to do that. There are works on consumer demand which extend the usual discrete choice model by allowing consumers to buy multiple versions of a product (see, e.g., Gentzkow, 2007.)

[9] For simplicity, we have assumed away possible differentiation at the firm level. This can be included, for example, by assuming that a consumer's valuation for firm j's product i is $u^j + x_i^j$, where u^j is another random variable and is i.i.d. across firms and consumers but has the same realization for all products in a firm. This is a special case of the general setting with potentially correlated match utilities in section 3.5.1.

separately and they choose price vectors $(p_1^j, \cdots, p_m^j)$, $j = 1, \cdots, n$. In the regime of pure bundling, each firm sells its products in a package only and they choose bundle prices P^j. In the regime of mixed bundling, each firm specifies prices P_s^j for every possible subset s of its m products. (If $m = 2$, then firm j's pricing strategy can be simply described as a pair of stand-alone prices (ρ_1^j, ρ_2^j) together with a joint-purchase discount δ^j.) The pricing strategy is the most general in the mixed bundling regime. (In the story of product incompatibility, however, the only relevant pricing strategies are separate sales and pure bundling.) In all the regimes the timing is that firms choose their prices simultaneously, and then consumers make their purchase decisions after observing all prices and match utilities. Since firms are *ex ante* symmetric, we will focus on symmetric pricing equilibrium.

As often assumed in the literature on oligopolistic competition, the market is fully covered in equilibrium. That is, each consumer buys all the m products. This will be the case if consumers do not have outside options, or on top of the above match utilities x_i^j, consumers have a sufficiently high basic valuation for each product (or if the lower bound of match utility $\underline{x}$ is high enough). Alternatively, we can consider a situation where the m products are essential components of a system for which consumers have a high basic valuation. In the regime of pure bundling, we will relax this assumption in section 3.5.2 and argue that the basic insights about the impacts of pure bundling remain qualitatively unchanged. However, in the mixed bundling part this assumption is important for tractability.

In the regime of pure bundling, two additional assumptions are made. First, consumers do not buy more than one bundle. This can be justified if the bundle is too expensive (e.g., due to high production costs) relative to the match utility difference across firms. (If the unit production cost is c for each product, a sufficient condition will be $c > \overline{x} - \underline{x}$.) This assumption is also naturally satisfied if we interpret pure bundling as an outcome of product incompatibility or high shopping costs as discussed before. (See more discussion about this assumption in the conclusion section.) Second, when there are more than two products (i.e., $m \geq 3$), we assume that each firm either bundles all its products or not at all, and there are no finer bundling strategies (e.g., bundling products 1 and 2 but selling product 3 separately). This assumption excludes the possible situations where firms bundle their products in asymmetric ways. (The pricing games in those asymmetric situations are hard to analyze.)[10]

[10] In the monopoly case, it is possible to deal with finer bundling strategies (see, e.g., Palfrey, 1983, and Fang and Norman, 2006).

2.1 Separate sales: revisiting Perloff-Salop model

This section studies the benchmark regime of separate sales. Since firms compete on each product separately, the market for each product is a Perloff-Salop model. Consider the market for product i, and let p be the (symmetric) equilibrium price.[11] Suppose firm j deviates and charges p', while other firms stick to the equilibrium price p. Then the demand for firm j's product i is

$$q(p') = \Pr[x_i^j - p' > \max_{k \neq j}\{x_i^k - p\}] = \int_{\underline{x}}^{\overline{x}} [1 - F(x - p + p')] dF(x)^{n-1} \ .$$

(In the following, whenever there is no confusion, we will suppress the integral limits $\underline{x}$ and $\overline{x}$.) Notice that $F(x)^{n-1}$ is the cdf of the match utility of the best product i among the $n-1$ competitors. So firm j is as if competing with one firm which supplies a product with match utility distribution $F(x)^{n-1}$ (but in equilibrium they have to charge the same price p). Firm j's profit from product i is $p'q(p')$, and one can check that the first-order condition for p to be the equilibrium price is

$$\frac{1}{p} = n \int f(x) dF(x)^{n-1} \ . \tag{1}$$

This first-order condition is also sufficient for defining the equilibrium price if f is logconcave (see, e.g., Caplin and Nalebuff, 1991). A simple observation is that given the assumption of full market coverage, shifting the support of the match utility does not affect the equilibrium price.

Two comparative static exercises are useful for our subsequent analysis. The first question is: how does the equilibrium price vary with the number of firms? The equilibrium condition (1) can be rewritten as

$$p = \frac{q(p)}{|q'(p)|} = \frac{1/n}{\int f(x) dF(x)^{n-1}} \ . \tag{2}$$

The numerator is a firm's equilibrium demand and it must decrease with n. The denominator is the absolute value of a firm's equilibrium demand slope. It captures the density of a firm's marginal consumers who are indifferent between its product i and the best product i among its competitors. How the denominator changes with n depends on the shape of f. For example, if the density function f is increasing, it increases with n and so p must decrease with n. While if f is decreasing, it then decreases with n, which works against the demand size effect. However, as long as $|q'(p)|$ does not decrease with n at a speed faster than $1/n$, the demand elasticity (at any price p) increases with n such that the equilibrium price decreases with n. The following result reports a sufficient condition for that.

[11] In the duopoly case, there is no asymmetric pricing equilibrium (see, e.g., section 4 in Perloff and Salop, 1985). Beyond the duopoly case, whether there are asymmetric equilibria or not is still an open question in general. But in the logit model where x_i^j follows an extreme value distribution, the unique equilibrium is the symmetric equilibrium (see, e.g., Caplin and Nalebuff, 1991).

Lemma 1 *Suppose $1-F$ is logconcave (which is implied by logconcave f). Then p defined in (1) decreases with n. Moreover, $\lim_{n\to\infty} p = 0$ if and only if $\lim_{x\to\overline{x}} \frac{f(x)}{1-F(x)} = \infty$.*

Proof. Let $x_{(2)}$ be the second highest order statistics of $\{x_1, \cdots, x_n\}$. Let $F_{(2)}$ and $f_{(2)}$ be its cdf and pdf, respectively. Using

$$f_{(2)}(x) = n(n-1)(1-F(x))F(x)^{n-2}f(x) \ ,$$

we can rewrite (1) as[12]

$$\frac{1}{p} = \int \frac{f(x)}{1-F(x)} dF_{(2)}(x) \ . \tag{3}$$

Since $x_{(2)}$ increases with n in the sense of first order stochastic dominance, a sufficient condition for p to be decreasing in n is the hazard rate $f/(1-F)$ being increasing (or equivalently, $1-F$ being logconcave). The limit result as $n \to \infty$ also follows from (3) since $x_{(2)}$ converges to $\overline{x}$ as $n \to \infty$. ■

This result improves the relevant discussion in Perloff and Salop (1985). (They investigated the same comparative static question but did not derive a simple condition for p to decrease in n.) One special case is the exponential distribution which has a constant hazard rate $f/(1-F)$. In that case the price is independent of the number of firms. Nevertheless, the logconcavity of $1-F$ is not a necessary condition. That is, even if $1-F$ is not logconcave, it is still possible that price decreases with n (if the equilibrium price is determined by (1)).[13] The tail behavior condition for $\lim_{n\to\infty} p = 0$ is satisfied if $f(\overline{x}) > 0$. But it can be violated if $f(\overline{x}) = 0$. One example is the extreme value distribution with $F(x) = e^{-e^{-x}}$ (which generates the logit model). Both f and $1-F$ are strictly logconcave in this example. But one can check that $p = \frac{n}{n-1}$, which decreases with n and approaches to 1 in the limit.

The second comparative static question is: if the distribution of match utility becomes more concentrated as illustrated in Figure 1 below (where the density function becomes more "peaked" from the solid to the dashed one), how will the equilibrium price change? Intuitively, a more peaked density as in Figure 1 means less dispersed consumer valuations, or less product differentiation across firms. So this should intensify price competition and induce a lower market price.

[12] This rewriting has an economic interpretation. The right-hand side of (3) is the density of all marginal consumers in the market. A consumer is a marginal one if her best product and second best one have the same match utility. Conditional on $x_{(2)} = x$, the cdf of $x_{(1)}$ is $\frac{F(z)-F(x)}{1-F(x)}$ for $z \geq x$. So its pdf at $x_{(1)} = x$ is the hazard rate $\frac{f(x)}{1-F(x)}$. Integrating this according to the distribution of $x_{(2)}$ yields the right-hand side of (3). Dividing it by n gives the density of one firm's marginal consumers (i.e., $|q'(p)|$).

[13] One such example is the power distribution: $F(x) = x^k$ with $k \in (\frac{1}{n}, 1)$. In this example, $1-F$ is neither logconcave nor logconvex. But one can check that the equilibrium price is $p = \frac{nk-1}{n(n-1)k^2}$ and it decreases in n. If $1-F$ is logconvex and the equilibrium price is determined by (1), then p increases with n.

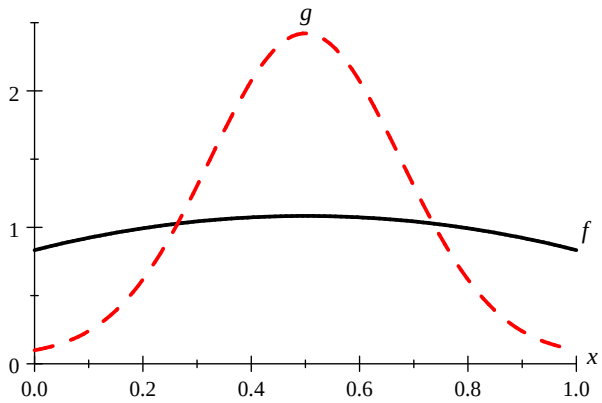


Figure 1: An example of more concentrated match utility distribution

This must be the case if the dashed distribution degenerates at a certain point (say, its mean) such that all products become homogenous. However, except for this limit case, a distribution concentration as in Figure 1 does not necessarily lead to a lower market price. Let f be the solid density in Figure 1, and let g be the dashed one. (Suppose both have the support $[0, 1]$.) We already knew from (2) that in either case the equilibrium price equals the ratio of the equilibrium demand to the negative of the equilibrium demand slope. Since the equilibrium demand is always $1/n$ due to firm symmetry, only the equilibrium demand slopes (or the densities of marginal consumers) matter for the price comparison. When n is large, a given consumer's valuation for the best product among a firm's competitors is close to the upper bound 1 almost for sure. So for that consumer to be this firm's marginal consumer, her valuation for its product should also be close to 1. In other words, when n is large, the position of a firm's marginal consumers should be close to the upper bound no matter which valuation density function applies. Since $f(1) > g(1)$, we deduce that each firm has a lower density of marginal consumers (or a smaller demand slope) when the density function is g. Therefore, when n is large, the more peaked density g should induce a higher market price than f. This suggests that when the number of firms is large, the tail behavior, instead of the peakedness, of the match utility density function determines the equilibrium price.[14]

Lemma 2 *Consider two continuous and bounded valuation density functions f and g with supports $[\underline{x}_f, \overline{x}_f]$ and $[\underline{x}_g, \overline{x}_g]$, respectively. Let p_i, $i = f, g$, be the equilibrium prices associated with the two densities, and suppose they are determined as in (1). Then if $\overline{x}_f < \infty$ and $f(\overline{x}_f) > g(\overline{x}_g)$, there exists $\hat{n}$ such that $p_f < p_g$ for $n > \hat{n}$.*

[14] Gabaix et al. (2013) study the asymptotic behavior of the equilibrium price in random utility models and make the same point. By using extreme value theory, they show that when the number of firms is large, markups are proportional to $[nf(F^{-1}(1 - 1/n))]^{-1}$. Intuitively, by noticing $\int_0^1 t dt^{n-1} = 1 - 1/n$, this can also be seen from the proof of Lemma 2 below.

Proof. Let F and G be the corresponding cdf's. By changing the integral variable from x to $t = F(x)$, we have

$$\frac{1}{p_f} = n \int_{\underline{x}_f}^{\overline{x}_f} f(x) dF(x)^{n-1} \Rightarrow \frac{1}{p_f} = n \int_0^1 l_f(t) dt^{n-1} \ ,$$

where $l_f(t) \equiv f(F^{-1}(t))$, and t^{n-1} is a cdf on $[0,1]$. Similarly, we have

$$\frac{1}{p_g} = n \int_0^1 l_g(t) dt^{n-1} \quad ,$$

where $l_g(t) \equiv g(G^{-1}(t))$. Then

$$p_f < p_g \Leftrightarrow \int_0^1 [l_f(t) - l_g(t)] dt^{n-1} > 0 \ .$$

Given f and g are bounded, so is $l_f(t) - l_g(t)$. Given $f(\overline{x}_f) > g(\overline{x}_g)$, we have $l_f(1) - l_g(1) = f(\overline{x}_f) - g(\overline{x}_g) > 0$. Then

$$\lim_{n \to \infty} \int_0^1 [l_f(t) - l_g(t)] dt^{n-1} = l_f(1) - l_g(1) > 0$$

as the distribution t^{n-1} converges to the upper bound 1 as $n \to \infty$. This implies that $p_f < p_g$ when n is sufficiently large. ■

Two technical points are worth mentioning: (i) As we have mentioned before, if g is completely degenerated at some point, then $p_g = 0 < p_f$ and so Lemma 2 does not hold. (In that case, g is no longer bounded and the above proof does not apply.) (ii) The above result cannot necessarily be extended to the case where $f(\overline{x}_f) = g(\overline{x}_g)$ but $f > g$ for x close to the upper bounds. First of all, it is now possible that $l_f(t) < l_g(t)$ for t close to 1.[15] Second, even if $l_f(t) > l_g(t)$ for t close to 1, given $l_f(1) = l_g(1)$, now for a large but finite n, $[l_f(t) - l_g(t)](n-1)t^{n-2}$ is close to zero everywhere (and it equals zero at $t = 1$), and so the sign of $\int_0^1 [l_f(t) - l_g(t)] dt^{n-1}$ does not necessarily only depend on the sign of $l_f(t) - l_g(t)$ for t close to 1.

An implication of Lemma 2 is that in the Perloff-Salop model a mean-preserving contraction of the match utility distribution may actually *increase* the market price. As we will see in next section, this observation is crucial for understanding the price comparison result between separate sales and pure bundling.

3 Pure Bundling

3.1 Equilibrium prices

Now consider the regime where all firms adopt the pure bundling strategy. Denote by $X^j \equiv \sum_{i=1}^m x_i^j$ the match utility of firm j's bundle. Then if firm j charges a

[15] For example, when f and g are pdf's of two normal distributions and g has a smaller variance, we actually have $l_f(t) < l_g(t)$ for $t \in (0,1)$.

bundle price P' while other firms charge the equilibrium price P, the demand for j's bundle is

$$Q(P') = \Pr[X^j - P' > \max_{k \neq j}\{X^k - P\}] = \Pr[\frac{X^j}{m} - \frac{P'}{m} > \max_{k \neq j}\{\frac{X^k}{m} - \frac{P}{m}\}] \ .$$

Let G and g denote the cdf and pdf of X^j/m, respectively. Then the equilibrium per-product bundle price P/m is determined similarly as the separate sales price p in (1), except that now we use a different distribution G. That is,

$$\frac{1}{P/m} = n \int g(x) dG(x)^{n-1} \ . \tag{4}$$

Notice that g is logconcave if f is logconcave (see, e.g., Miravete, 2002). Then the first-order condition (4) is sufficient for defining the equilibrium bundle price if f is logconcave. Also notice that $1 - G$ is logconcave if $1 - F$ is logconcave. Hence, similar results as in Lemma 1 also hold here.

Lemma 3 *Suppose $1 - F$ is logconcave (which is implied by logconcave f). Then the bundle price P defined in (4) decreases with n. Moreover, $\lim_{n\to\infty} P = 0$ if and only if $\lim_{x\to\overline{x}} \frac{g(x)}{1-G(x)} = \infty$.*

Notice that the per-product bundle valuation X^j/m is a mean-preserving contraction of x_i^j with the same support. So g is more peaked than f and it also has a thinner tail, as illustrated in Figure 1 above. In particular, $g(\overline{x}) = 0$ even if $f(\overline{x}) > 0$. This is simply because $X^j/m = \overline{x}$ only if $x_i^j = \overline{x}$ for all $i = 1, \cdots, m$, or intuitively this is because finding a very well matched bundle is much less likely than finding a very well matched single product.[16] This also implies that the tail behavior condition for $\lim_{n\to\infty} P = 0$ is less likely to be satisfied, compared to the case of separate sales.[17]

3.2 Comparing prices and profits

From (1) and (4), we can see that the comparison between separate sales and pure bundling is just a comparison between two Perloff-Salop models with different match utility distributions F and G (where the latter is a mean-preserving contraction of the former with the same support). Bundling leads to lower prices if $P/m \leq p$. Using the technique in the proof of Lemma 2, we have

$$\frac{P}{m} \leq p \Leftrightarrow \int_0^1 [l_f(t) - l_g(t)] t^{n-2} dt \leq 0 \ , \tag{5}$$

[16] More formally, when $m = 2$ the pdf of $(x_1^j + x_2^j)/2$ is $g(x) = 2\int_{2x-\overline{x}}^{\overline{x}} f(2x - t) dF(t)$ for $x \geq (\underline{x} + \overline{x})/2$, so $g(\overline{x}) = 0$. A similar argument works for $m \geq 3$.

[17] Even if both P/m and p converge to zero as $n \to \infty$, P/m converges much slower than p if $f(\overline{x}) > 0$. In that case we have $\lim_{n\to\infty} \frac{p}{P/m} = \frac{g(\overline{x})}{f(\overline{x})} = 0$.

where $l_f(t) \equiv f(F^{-1}(t))$ and $l_g(t) \equiv g(G^{-1}(t))$. Given full market coverage, profit comparison is the same as price comparison.

Proposition 1 *Suppose f is logconcave. (i) When $n = 2$, bundling reduces market prices and profits for any $m \geq 2$.*
(ii) For a fixed $m < \infty$, if f is bounded and $f(\bar{x}) > 0$, there exists $\hat{n}$ such that bundling increases market prices and profits for $n > \hat{n}$. If f is further such that $l_f(t)$ and $l_g(t)$ cross each other at most twice, then bundling decreases prices and profits if and only if $n \leq \hat{n}$.
(iii) For a fixed $n < \infty$, $\lim_{m\to\infty} P/m = 0$ and so there exists $\hat{m}$ such that bundling reduces market prices and profits for $m > \hat{m}$.

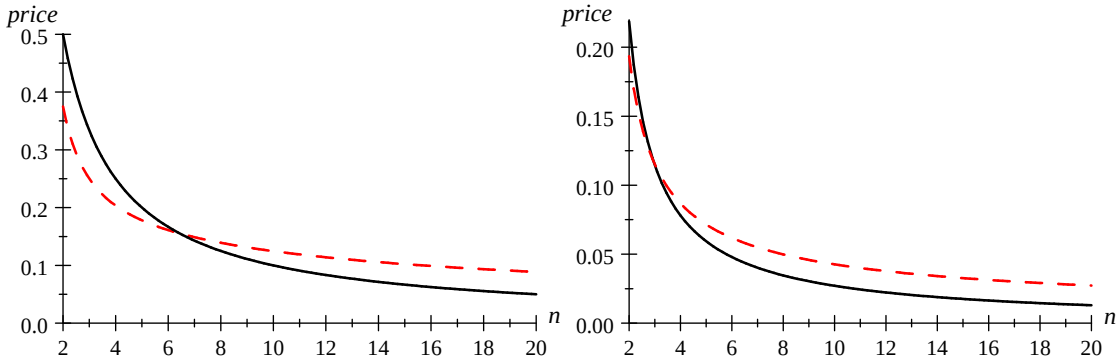
Result (i) generalizes the observation in the existing literature about how pure bundling affects market prices in duopoly. The intuition of this result is more transparent when the density function f is symmetric. In that case, the average position of marginal consumers in a duopoly is at the mean, and g is more peaked at the mean than f. So there are more marginal consumers (or a larger demand slope) in the case of g. This induces firms to charge a lower price in the bundling regime.[18] Result (iii) simply follows from the *law of large numbers*. Given the assumption that x_i^j has a finite mean μ, X^j/m converges to μ as $m \to \infty$. In other words, the per-product valuation for the bundle becomes homogeneous across both consumers and firms. So P/m converges to zero.[19]

Result (ii) that pure bundling can soften price competition is more surprising. The result for large n is from Lemma 2 since $f(\overline{x}) > g(\overline{x}) = 0$. Bundling leads to a Perloff-Salop model where the (per-product) valuation density has a thinner right tail. When n is large, the marginal consumers mainly locate on the right tail, and so there are fewer marginal consumers in the bundling case. This leads to a less elastic demand and a higher market price. To illustrate, consider two examples which satisfy all the conditions in result (ii). In the uniform distribution example with $f(x) = 1$, $P/m < p$ when $n \leq 6$ and $P/m > p$ when $n > 6$. Figure 2(a) below describes how both prices vary with n (where the solid curve is p and the dashed one is P/m). In the example with an increasing density $f(x) = 4x^3$, as described in

[18] Notice that we are dealing with P/m instead of the bundle price P directly. So when we mention the measure of marginal consumers in the bundling case, it is actually m times the real measure of marginal consumers who are indifferent between a firm's bundle and the best bundle among the competitors. This is because reducing P/m by ε is equivalent to reducing P by $m\varepsilon$.

[19] Notice that the valuation for the whole bundle X^j has a larger variance when m increases, and the bundle price P increases in m. From the proof of result (iii) we can see that when m is large P increases in m in a speed of $\sqrt{m}$. A similar result has been shown in Nalebuff (2000) which considered a multi-dimensional Hotelling model with two firms and an arbitrary number of products.

Figure 2(b) below $P/m < p$ only when $n = 2$ and $P/m > p$ whenever $n > 2$. This second example indicates that the threshold $\hat{n}$ can be small. Another observation from these two examples is that the increase of price caused by bundling can be significant even if n is relatively large (e.g., when n is around 20).[20]



(a) $f(x) = 1$ (b) $f(x) = 4x^3$

Figure 2: Price comparison with $m = 2$

Result (ii) requires $f(\overline{x}) > 0$. If $f(\overline{x}) = 0$ (where $\overline{x}$ can be infinity), then $f(\overline{x}) = g(\overline{x})$ and Lemma 2 may not hold as we discussed before. For instance, in the following example of normal distribution where $\lim_{x\to\infty} f(x) = 0$, bundling always lowers market prices.

Example of normal distribution. As we have observed before, shifting the support of the match utility distribution does not affect the equilibrium price. So let us normalize the mean to zero and suppose $x_i^j \sim \mathcal{N}(0, \sigma^2)$. Then the separate sales price defined in (1) is

$$p = \frac{\sigma}{n \int_{-\infty}^{\infty} \phi(x) d\Phi(x)^{n-1}} , \tag{6}$$

where Φ and ϕ are the cdf and pdf of the standard normal distribution $\mathcal{N}(0, 1)$, respectively.[21] The definition of X^j implies that $X^j/m \sim \mathcal{N}(0, \sigma^2/m)$, and so X^j/m is the same random variable as $x_i^j/\sqrt{m}$. Hence, the demand function in the bundling case is

$$Q(P') = \Pr[\frac{X^j}{m} - \frac{P'}{m} > \max_{k\neq j}\{\frac{X^k}{m} - \frac{P}{m}\}] = \Pr[x_i^j - \frac{P'}{\sqrt{m}} > \max_{k\neq j}\{x_i^k - \frac{P}{\sqrt{m}}\}] .$$

Therefore,

$$\frac{P}{\sqrt{m}} = p \Rightarrow \frac{P}{m} = \frac{p}{\sqrt{m}} < p . \tag{7}$$

[20] If $\lim_{n\to\infty} p = \lim_{n\to\infty} P/m = 0$, the difference between p and P/m will vanish as $n \to \infty$, but they can converge to zero at different speeds as we pointed out in footnote 17.

[21] One can check this result by using the fact that $x_i^j = \sigma\tilde{x}_i^j$, where $\tilde{x}_i^j$ has the standard normal distribution.

That is, with a normal distribution, bundling always reduces market prices (and so profits) regardless of n and m.

In this normal distribution example, bundling also makes the right tail thinner (i.e., $g(x) < f(x)$ for relatively large x) and the (average) position of marginal consumers also moves to the right as n increases. However, with unbounded support now the relative moving speed matters. The density tail is higher in the separate sales case and so it is more likely in that case to have a high valuation draw. This means that the position of marginal consumers moves to the right faster in the separate sales regime than in the bundling regime. Hence, for large n even if $f(x) > g(x)$, it is possible that $f(\hat{x}_f) < g(\hat{x}_g)$ where $\hat{x}_f$ is the position of marginal consumer in the separate sales regime and $\hat{x}_g$ is the position of marginal consumers in the bundling regime. (This is indeed the case in this normal distribution example.) This cannot happen if the upper bound is finite and $f(\overline{x}) > 0 = g(\overline{x})$. In that case, when n is large both $\hat{x}_f$ and $\hat{x}_g$ will be very close to $\overline{x}$ anyway and so we must have $f(\hat{x}_f) > g(\hat{x}_g)$. But in the case with an infinite upper bound, $\hat{x}_f$ and $\hat{x}_g$ can still be sufficiently far away from each other even if both move to infinity.

However, there are also examples where $f(\overline{x}) = 0$ and result (ii) still holds. For instance, in the example with a decreasing density $f(x) = 2(1-x)$ for $x \in [0,1]$, numerical analysis suggests a similar price comparison pattern as in Figure 2 (though the threshold $\hat{n}$ is bigger).

3.3 Comparing consumer surplus and total welfare

With full market coverage, consumer payment is a pure transfer and so total welfare (which is the sum of firm profits and consumer surplus) only reflects the match quality between consumers and products. In either regime, price is the same across firms in a symmetric equilibrium and so it does not distort consumer choices. Since bundling eliminates the opportunity to mix and match for consumers, it must reduce match quality and so total welfare.

However, the comparison of consumer surplus can be more complicated. If pure bundling increases market prices, it must harm consumers. From Proposition 1, we know this is the case when $f(\overline{x}) > 0$ and n is sufficiently large. The trickier situation is when pure bundling decreases market prices. Then there is a trade-off between the negative match quality effect and the positive price effect. Our analysis below suggests that the positive price effect dominates (and so consumers benefit from bundling) only if the number of firms is relatively small. Intuitively, this is because when there are more firms, bundling eliminates more mixing-and-matching opportunities and so the negative match quality effect becomes more significant.

Denote by v a consumer's expected surplus in the regime of separate sales. Then

the per-product consumer surplus is

$$\frac{v}{m} = \mathbb{E}\left[\max_j\{x_i^j\}\right] - p \ . \tag{8}$$

Denote by V a consumer's expected surplus in the regime of pure bundling. Then the per-product consumer surplus is

$$\frac{V}{m} = \mathbb{E}\left[\max_j\left\{\frac{X^j}{m}\right\}\right] - \frac{P}{m} \ . \tag{9}$$

We can get an analytical result in the limit case with $m \to \infty$**.** Recall that μ is the mean of x_i^j. We already knew that $\lim_{m\to\infty} X^j/m = \mu$ and $\lim_{m\to\infty} P/m = 0$. So

$$\lim_{m\to\infty} \frac{V}{m} = \mu \ .$$

Therefore, when $m \to \infty$, pure bundling improves consumer welfare if and only if

$$\mathbb{E}\left[\max_j\{x_i^j\}\right] - \mu < p \ . \tag{10}$$

With separate sales, consumers enjoy better matched goods (which is reflected by the left-hand side), but they also pay more (which is reflected by the right-hand side). It is clear that the match quality effect increases with n, while the price effect decreases with n given the logconcavity condition. In the proof of Proposition 2 below, we show that (10) holds for $n = 2$ but fails for a sufficiently large n. Therefore, in this limit case with $m \to \infty$ pure bundling improves consumer welfare if and only if the number of firms is below some threshold.

Proposition 2 *Suppose f is logconcave. (i) For a fixed $m < \infty$, if f is bounded and $f(\bar{x}) > 0$, there exists $\hat{n}$ such that bundling harms consumers if $n > \hat{n}$.*
(ii) If $m \to \infty$, there exists n^* such that pure bundling benefits consumers if and only if $n \leq n^*$.

The threshold n^* in result (ii) can be small. For example, in the uniform distribution case with $F(x) = x$, condition (10) simplifies to $n^2 - 3n - 2 < 0$, which holds only for $n \leq 3$.

For a finite m, it is difficult to prove more analytical results beyond result (i). However, numerical analysis suggests a similar threshold result as in the limit case with $m \to \infty$. Figure 3 below describes how consumer surplus varies with n in the uniform distribution case when $m = 2$ (where the solid curve is for separate sales, and the dashed one is for bundling). Again, the threshold is $n^* = 3$. Moreover, notice that in this example the negative effect when n is relatively large is more significant than the positive effect when n is small.[22]

[22]The difference will disappear eventually as $n \to \infty$, because in this example $\lim_{n\to\infty} p = \lim_{n\to\infty} P/m = 0$ and $\lim_{n\to\infty} \mathbb{E}[\max_j\{x_i^j\}] = \lim_{n\to\infty} \mathbb{E}[\max_j\{X^j/m\}] = \bar{x}$.

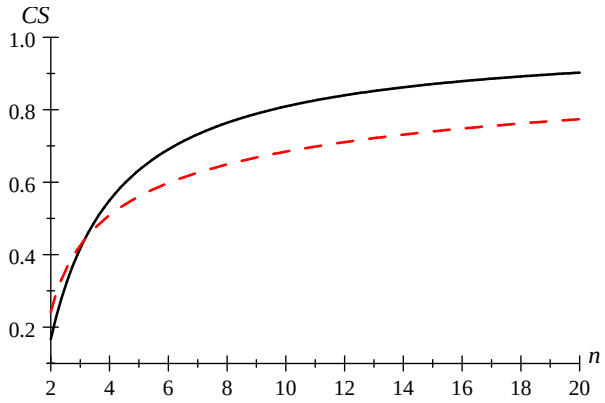


Figure 3: Consumer surplus comparison with uniform distribution and $m = 2$

In the normal distribution example, we can get analytical results for any m. A similar threshold result holds and the threshold is independent of m.

Example of normal distribution. Suppose $x_i^j \sim \mathcal{N}(0, \sigma^2)$. Using the definitions of consumer surplus in (8) and (9) and the result (7), we can see that pure bundling improves consumer surplus if and only if

$$\mathbb{E}\left[\max_j\{x_i^j\}\right] - \mathbb{E}\left[\max_j\left\{\frac{X^j}{m}\right\}\right] < p[1 - \frac{1}{\sqrt{m}}] \ . \tag{11}$$

(The left-hand side is the match quality effect, and the right-hand side is the price effect.) In the Appendix, we show that

$$\mathbb{E}\left[\max_j\{x_i^j\}\right] = \frac{\sigma^2}{p} \ , \quad \mathbb{E}\left[\max_j\left\{\frac{X^j}{m}\right\}\right] = \frac{1}{\sqrt{m}}\frac{\sigma^2}{p} \ . \tag{12}$$

Then (11) simplifies to $p > \sigma$. Using (6), one can check that this holds only for $n = 2, 3$, and so the threshold is again $n^* = 3$.

3.4 Incentive to bundle

We now turn to firms' incentive to bundle. Consider an extended game where firms can choose both bundling strategies and prices. But suppose that firms can choose only between separate sales and pure bundling. We first investigate the case where firms choose bundling strategies and prices simultaneously. This tends to describe the situation when bundling is a pricing strategy and so can be easily adjusted.

Simultaneous choices. We start with the case with two firms only. In this duopoly case, if one firm chooses to bundle, the situation will be like both firms bundling.

Proposition 3 *Suppose $n = 2$ and firms make bundling and pricing decisions simultaneously. Suppose the equilibrium price p in the regime of separate sales is defined in (1) and the equilibrium price P in the regime of pure bundling is defined*

in (4). Then whenever $p \neq P/m$, the unique (pure-strategy) Nash equilibrium is that both firms choose to bundle and charge a bundle price P.

Proof. First of all, both firms bundling is a Nash equilibrium outcome. This is simply because if a firm unilaterally unbundles, the market situation does not change for consumers.[23]

Second, it is not an equilibrium outcome that both firms adopt separate sales. Consider the hypothetical equilibrium where both firms sell their products separately at price p. Now suppose firm j unilaterally bundles. It can at least earn the same profit as before by setting a bundle price mp. But it can do strictly better by adjusting prices as well. Suppose firm j sets a bundle price $mp - m\varepsilon$, where ε is a small positive number. The negative (first-order) effect of this deviation on firm j's profits is $\frac{m}{2}\varepsilon$. (Half of the consumers buy from firm j when $\varepsilon = 0$, and now they pay $m\varepsilon$ less.) The demand for firm j's bundle becomes

$$\Pr(X^j + m\varepsilon > X^k) = \int G(x+\varepsilon)dG(x) \ ,$$

where $k \neq j$. So the demand increases by $\varepsilon \int g(x)^2 dx$. This means that the positive effect of the deviation on firm j's profit is

$$mp \times \varepsilon \int g(x)^2 dx = \frac{mp}{P} \times \frac{m}{2}\varepsilon \ .$$

(The equality is because of (4) for $n = 2$.) Therefore, the deviation is profitable if $P < mp$. Similarly, one can show that if $P > mp$, then charging a bundle price $mp + m\varepsilon$ will be a profitable deviation.

Finally, there are no asymmetric equilibria where one firm bundles and the other does not. Consider a hypothetical equilibrium where firm j bundles and firm k does not. For consumers, this situation is the same as both firms bundling. So in equilibrium it must be the case that firm j offers a bundle price P defined in (4) with $n = 2$, and firm k offers individual prices $\{p_i\}_{i=1}^m$ such that $\sum_{i=1}^m p_i = P$. Suppose now firm j unbundles and offers prices $\{p_i - \varepsilon\}_{i=1}^m$, where ε is a small positive number. The negative (first-order) effect of this deviation on firm j's profit is $\frac{m}{2}\varepsilon$. But the demand for firm j's each product increases by $\varepsilon \int f(x)^2 dx$. So the positive effect is

$$\sum_{i=1}^m p_i \times \varepsilon \int f(x)^2 dx = \frac{P}{mp} \times \frac{m}{2}\varepsilon \ .$$

(The equality used (1) for $n = 2$ and $\sum_{i=1}^m p_i = P$.) Therefore, the proposed deviation is profitable if $P > mp$. Similarly, if $P < mp$, setting prices $\{p_i + \varepsilon\}_{i=1}^m$ will be a profitable deviation for firm j. ■

[23] This argument depends on the assumption that consumers buy all products but for each product they have unit demand and do not buy more than one variant.

The situation is more complicated when there are more than two firms. But a simple observation is that it is still a Nash equilibrium that all firms bundle for the same reason as before. However, now separate sales may also be an equilibrium outcome. Given all other firms offer separate sales at price p, if one firm, say, firm j unilaterally bundles and sets a bundle price mp, the demand for its bundle will be strictly less than $1/n$ when $n \geq 3$. Formally, let

$$y_i \equiv \max_{k \neq j} \{x_i^k\} \tag{13}$$

denote the maximum match utility of product i among firm j's competitors. Then firm j is as if competing with *one* firm that offers a bundle with match utility $Y \equiv \sum_{i=1}^{m} y_i$ and price mp. If firm j charges the same bundle price mp, its demand will be $\Pr(X^j > Y) \leq \Pr(X^j > \max\{X^k\}_{k \neq j}) = 1/n$. The inequality is strict for $n \geq 3$. Thus, without further adjusting its prices it is unprofitable for firm j to unilaterally bundle.

Suppose now firm j adjusts its bundle price to $m(p - \varepsilon)$. Its demand is then

$$\Pr\left(\frac{X^j}{m} + \varepsilon > \frac{Y}{m}\right) . \tag{14}$$

Firm j has no incentive to bundle unilaterally if the optimal deviation profit is less than mp/n. But in general it is hard to calculate the optimal deviation profit and do the comparison.

To make progress, let us consider the limit case with $m \to \infty$. We then have $\lim_{m\to\infty} X^j/m = \mu$ and $\lim_{m\to\infty} Y/m = \mathbb{E}[y_i]$, and so (14) is a step function. If ε is greater than $\mathbb{E}[y_i] - \mu$, firm j will attract all consumers. Otherwise, it gets zero demand. Therefore, the optimal (per product) deviation profit is $\max\{0, p - (\mathbb{E}[y_i] - \mu)\}$. This is less than the (per product) profit p/n in the regime of separate sales (such that firm j has no incentive to bundle unilaterally) if and only if

$$(1 - \frac{1}{n})p < \mathbb{E}[y_i] - \mu = \int \left[F(x) - F(x)^{n-1}\right] dx . \tag{15}$$

This is clearly not true for $n = 2$ (consistent with Proposition 3). On the other hand, we have $\int_{\underline{x}}^{\overline{x}} F(x)dx > \lim_{n\to\infty} p$ as shown in the proof of Proposition 2. Then (15) must hold when n is sufficiently large. The following proposition reports a cut-off result:

Proposition 4 *Suppose f is logconcave and $m \to \infty$. Then there exists $\tilde{n}$ such that separate sales is also a Nash equilibrium outcome if and only if $n > \tilde{n}$.*

This result implies that the extended game with bundling choices can have two symmetric equilibria when n is sufficiently large. To illustrate the magnitude of $\tilde{n}$, let us consider the uniform distribution example. Then condition (15) becomes

$n^2 - 4n + 2 > 0$, and it holds for $n > 3$ (i.e., $\tilde{n} = 3$). This is the same as the threshold n^* in consumer surplus comparison reported in result (ii) of Proposition 2. Therefore, in this uniform example (with a large number of products), separate sales is also an equilibrium outcome if and only if consumers prefer separate sales to pure bundling. In other words, with a proper equilibrium selection the market can work well for consumers. (The same is also true for the exponential distribution and the normal distribution.)[24]

With more than two firms, there may also exist asymmetric equilibria where some firms bundle and the others do not. A general investigation into this problem is hard because the pricing equilibrium when firms adopt asymmetric bundling strategies does not have a simple characterization.[25] But numerical analysis is feasible. Let us illustrate by a uniform example with $n = 3$ and $m = 2$. In this example, we can claim that there are no asymmetric equilibria.

The first possible asymmetric equilibrium is that one firm bundles and the other two do not. In this hypothetical equilibrium, the bundling firm charges $P \approx 0.513$ and earns a profit about 0.176, and the other two firms charge a separate price $p \approx 0.317$ and each earns a profit about 0.208. But if the bundling firm unbundles and charges the same separate price as the other two firms, it will have a demand $\frac{1}{3}$ and its profit will rise to about 0.211. The second possible asymmetric equilibrium is that two firms bundle and the third one does not. This hypothetical equilibrium is like all firm are bundling. Then each bundling firm charges a bundle price $P = 0.5$, the third firm charges $p_1 + p_2 = 0.5$, and each firm has market share $\frac{1}{3}$. But if one bundling firm unbundles and offers the same separate prices as the third firm, as we already knew before the remaining bundling firm will have a demand less than $\frac{1}{3}$, and this implies that the deviation firm will have a demand higher than $\frac{1}{3}$. This improves its profit.

Sequential choices. Now suppose that firms make their bundling choices first, and then engage in price competition after observing the bundling outcome. (This is usually the case when bundling is a product design strategy.) First of all, as before all firms bundling is a Nash equilibrium outcome. The issue of whether separate sales is also an equilibrium is more complicated. In the duopoly case, if f is logconcave,

[24] There exists examples (e.g., $f(x) = 2(1-x)$) where $\tilde{n} \neq n^*$. But usually the difference is small.

[25] The main reason is that the firms which bundle their products make each other firm treat their products as complements. This complicates the demand calculation. To understand this, let us consider a situation where firm 1 bundles while other firms do not. Suppose firm $k \neq 1$ lowers its product 1's price. Now some consumers will stop buying firm 1's bundle and switch to buying all products from other firms. This will increase the demand for firm k's *all* products. In other words, for any firm $k \neq 1$ reducing one product's price will increase the demand for its other products as well. The details on the demand functions and the first-order conditions in this case are available upon request.

then bundling leads to a lower price and profit as we have shown in Proposition 1. Then no firm has a unilateral incentive to deviate from separate sales. That is, separate sales is a Nash equilibrium outcome as well.

When there are more than two firms, due to the complication of the pricing game when firms adopt asymmetric bundling strategies, no general results are available. In the uniform example with $n = 3$ and $m = 2$, we can numerically show that separate sales is another equilibrium, but there are no asymmetric equilibria. If all three firms sell their products separately, each firm charges a price $p = \frac{1}{3}$ and each firm's profit is $\frac{2}{9} \approx 0.222$. Now if one firm, say, firm 1 deviates and bundles, then in the asymmetric pricing game, firm 1 charges a bundle price $P \approx 0.513$ and the other two firms charge a separate price $p \approx 0.317$ for each product. Firm 1's profit drops to about 0.176 (and each other firm's profit drops to about 0.208). So no firm wants to bundle unilaterally. This also implies that one firm bundling and the other two not is not an equilibrium, because all firms benefit if the bundling firm unbundles. The last possibility is that two firms bundle and the other does not. That situation is like all firms bundling, and each firm's profit is about 0.167. But if one bundling firm unbundles, its profit will rise to 0.208. Hence, there are no asymmetric equilibria.

3.5 Discussions

3.5.1 Asymmetric products and correlated distributions

We now consider a more general setting where the m products within each firm are potentially asymmetric and their match utilities are potentially correlated. Let $\mathbf{x}^j = (x_1^j, \cdots, x_m^j)$ be a consumer's valuations for the m products at firm j. Suppose $\mathbf{x}^j$ is i.i.d. across firms and consumers (so firms are still *ex ante* symmetric), and it is distributed according to a common joint cdf $F(x_1, \cdots, x_m)$ with support $S \subset \mathbb{R}^m$ and a continuous joint pdf $f(x_1, \cdots, x_m)$. Let F_i and f_i, $i = 1, \cdots, m$, be the marginal cdf and pdf of x_i^j, and let $[\underline{x}_i, \overline{x}_i]$ be its support. Let $[\underline{x}, \overline{x}]$ be the support of X^j/m, the average per-product match utility of the bundle. Let G and g be its cdf and pdf as before.

In the regime of separate sales, let p_i denote the (symmetric) equilibrium price for product i. Since the competition is still separate across products, the same formula in (1) applies as long as we use the marginal distributions:

$$\frac{1}{p_i} = n \int_{\underline{x}_i}^{\overline{x}_i} f_i(x) dF_i(x)^{n-1} \ .$$

In the regime of pure bundling, the average per-product bundle price P/m is still determined as in (4):

$$\frac{1}{P/m} = \int_{\underline{x}}^{\overline{x}} g(x) dG(x)^{n-1} \ .$$

As shown in the following proposition, we still have the result that bundling raises market prices when the number of firms is above some threshold.

Proposition 5 *Suppose f is continuous and bounded. Suppose $S \subset \mathbb{R}^m$ is compact, strictly convex, and has full dimension. Then*
(i) if $f_i(\overline{x}_i) > 0$, there exists $\hat{n}_i$ such that $P/m > p_i$ for $n > \hat{n}_i$;
(ii) if $f_i(\overline{x}_i) > 0$ for all $i = 1, \cdots, m$, there exists $\hat{n}$ such that $P > \sum_{i=1}^m p_i$ for $n > \hat{n}$.

Proof. Our conditions imply that $g(\overline{x}) = 0$ (e.g., see the proof of Proposition 1 in Armstrong (1996)).[26] Then our results immediately follow from our Lemma 2. ■

The limit result that $\lim_{m\to\infty} P/m = 0$ still holds as long as X^j/m converges as $m \to \infty$. The trickier question is whether we still have the result that in the duopoly case bundling lowers market prices (i.e., $P < \sum_{i=1}^m p_i$ when $n = 2$)? Following the proof of result (i) in Proposition 1, let $d_i \equiv x_i^1 - x_i^2$ and let h_i be its pdf. (Notice that d_i is symmetric around zero, and h_i is logconcave if f_i is logconcave.) Then the separate sales price for product i is $p_i = \frac{1}{2h_i(0)}$. Let $\bar{h}$ be the pdf of $\frac{1}{m}\sum_{i=1}^m d_i$. Then the per-product bundle price is $P/m = \frac{1}{2\bar{h}(0)}$. The condition for $P < \sum_{i=1}^m p_i$ is then

$$\frac{1}{\bar{h}(0)} < \frac{1}{m}\sum_{i=1}^m \frac{1}{h_i(0)} \ . \tag{16}$$

Jensen's Inequality implies that the right-hand side is greater than $\left(\frac{1}{m}\sum_{i=1}^m h_i(0)\right)^{-1}$. Therefore, a sufficient condition for $P < \sum_{i=1}^m p_i$ is

$$\frac{1}{m}\sum_{i=1}^m h_i(0) \leq \bar{h}(0) \ . \tag{17}$$

If the m products at each firm are symmetric (but they can have correlated match utilities), the sufficient condition (17) can be proved under certain general conditions by an extension of Theorem 2.3 in Proschan (1965) (see Chan, Park, and Proschan, 1989).

If the m products at each firm are asymmetric, the sufficient condition (17) may not hold, but we have not found examples where the "iff" condition (16) is violated. Let us illustrate by a two-product example where x_1^j is uniformly distributed on $[0, 1]$ and x_2^j is uniformly distributed on $[0, k]$ with $k > 1$. Then d_1 and d_2 have pdf's

$$h_1(x) = \begin{cases} 1 + x & \text{if } x \in [-1, 0] \\ 1 - x & \text{if } x \in [0, 1] \end{cases} \quad \text{and} \quad h_2(x) = \begin{cases} \frac{1}{k}(1 + \frac{x}{k}) & \text{if } x \in [-k, 0] \\ \frac{1}{k}(1 - \frac{x}{k}) & \text{if } x \in [0, k] \end{cases} \ ,$$

[26] More formally, $g(\overline{x}) = \lim_{\varepsilon\to 0} \frac{1-G(\overline{x}-\varepsilon)}{\varepsilon}$, and our conditions ensure $1-G(\overline{x}-\varepsilon) = o(\varepsilon)$. Among the conditions, strict convexity of S excludes the possibility that the plane of $X^j/m = \overline{x}$ coincides with some part of S's boundary, and S being of full dimension excludes the possibility that x_i^j, $i = 1, \cdots, m$, are perfectly correlated.

respectively. Therefore, $h_1(0) = 1$, $h_2(0) = \frac{1}{k}$, and one can also check $\bar{h}(0) = \frac{2(3k-1)}{3k^2}$. The sufficient condition (17) holds only for k less than about 2.46, but the "iff" condition (16) holds for all $k > 1$.[27]

3.5.2 Without full market coverage

We now relax the assumption of full market coverage. For expositional convenience, let us return to the baseline case with symmetric products and independent match utilities. (The analysis below can be extended to the general setup as in the previous section.) A subtle issue here is whether the m products are independent products or perfect complements. This will affect the analysis in the benchmark of separate sales. If the m products are independent products, consumers decide whether to buy each product separately. While if the m products are perfect complements (e.g., they are essential components of a system), then whether to buy a product also depends on how well matched other products are. (With full market coverage, this distinction does not matter.) In the following, we consider the case of independent products for simplicity.

Suppose now x_i^j denotes the whole valuation for firm j's product i, and a consumer will buy a product or bundle only if the best offer in the market provides a positive surplus. Without loss of generality let $\underline{x} \leq 0$, and we also assume that x_i^j has a mean greater than the production cost which is normalized to zero.

In the regime of separate sales, if firm j deviates and charges p' for its product i, then the demand for its product i is

$$q(p') = \Pr[x_i^j - p' > \max_{k \neq j}\{0, x_i^k - p\}] = \int_{p'}^{\overline{x}} F(x_i^j - p' + p)^{n-1} dF(x_i^j) \ .$$

One can check that the first-order condition for p to be the (symmetric) equilibrium price is

$$p = \frac{q(p)}{|q'(p)|} = \frac{[1 - F(p)^n]/n}{\underbrace{F(p)^{n-1} f(p)}_{\text{exclusion effect}} + \underbrace{\int_p^{\overline{x}} f(x) dF(x)^{n-1}}_{\text{competition effect}}} \ . \tag{18}$$

(If f is logconcave, this is also sufficient for defining the symmetric equilibrium price.) In equilibrium, the measure of consumers who leave the market without purchasing product i is $F(p)^n$. (This is the probability that each firm's product i has a valuation less than the price p.) Given the symmetry of firms, the demand for each firm's product i is thus the numerator in (18). The demand slope in the denominator now has two parts: (i) The first term is the standard market exclusion effect: when the valuations of all other firms' product i are below p (which occurs

[27] If $0 < k < 1$, one can check $\bar{h}(0) = 2 - \frac{2}{3}k$. Then the sufficient condition (17) holds only for k greater than about 0.41, but the "iff" condition (16) holds for all $0 < k < 1$.

with probability $F(p)^{n-1}$), firm j will play as a monopoly. Then raising its price p by ε will exclude $\varepsilon f(p)$ consumers from the market. (ii) The second term is the same competition effect as in the case with full market coverage (up to the adjustment that a marginal consumer's valuation must be greater than the price p).

Similarly, in the bundling case, the equilibrium per-product price P/m is determined by the first-order condition:

$$\frac{P}{m} = \frac{[1 - G(P/m)^n]/n}{G(P/m)^{n-1} g(P/m) + \int_{P/m}^{\overline{x}} g(x) dG(x)^{n-1}} , \tag{19}$$

where G and g are the cdf and pdf of X^j/m as before.

Unlike the case with full market coverage, now the equilibrium price in each regime is implicitly determined in the first-order condition. The following result reports the condition for each first-order condition has a unique solution.

Lemma 4 *Suppose f is logconcave. There is a unique equilibrium price $p \in (0, p_M)$ defined in (18), where p_M is the monopoly price and solves $p_M = [1-F(p_M)]/f(p_M)$, and p decreases with n. Similar results hold for P/m defined in (19).*

According to our knowledge, neither the existence result nor the monotonicity result has been proved before when there is no full market coverage.

Similar results concerning the price comparison hold when n is large or when m is large. For a fixed $n < \infty$, we still have $\lim_{m\to\infty} P/m = 0$ since X^j/m converges as $m \to \infty$. For a fixed $m < \infty$, if n is large, then the demand size difference between the two numerators in (18) and (19) is negligible, and so is the exclusion effect difference in the denominators. So price comparison is again determined by the comparison of $f(\overline{x})$ and $g(\overline{x})$ (by applying the same logic as in Lemma 2). Intuitively, when there are a large number of varieties in the market, almost every consumer can find something she likes and so almost no consumers will leave the market without purchasing anything. The situation will be then close to the case with full market coverage. Thus, we have a similar result that when $f(x)$ is (uniformly) bounded and $f(\overline{x}) > 0$, bundling raises market prices when n is greater than a certain threshold. The duopoly case without full market coverage is harder to deal with, and we have not been able to prove a result similar to result (i) in Proposition 1.

Figure 4 below reports the impacts of pure bundling on market prices, profits, consumer surplus and total welfare in the uniform example with $F(x) = x$. (The solid curves are for separate sales, and the dashed ones are for pure bundling.) They are qualitatively similar as those in the case with full market coverage. In particular, the total welfare result is the same even if we introduce the exclusion effect of price.

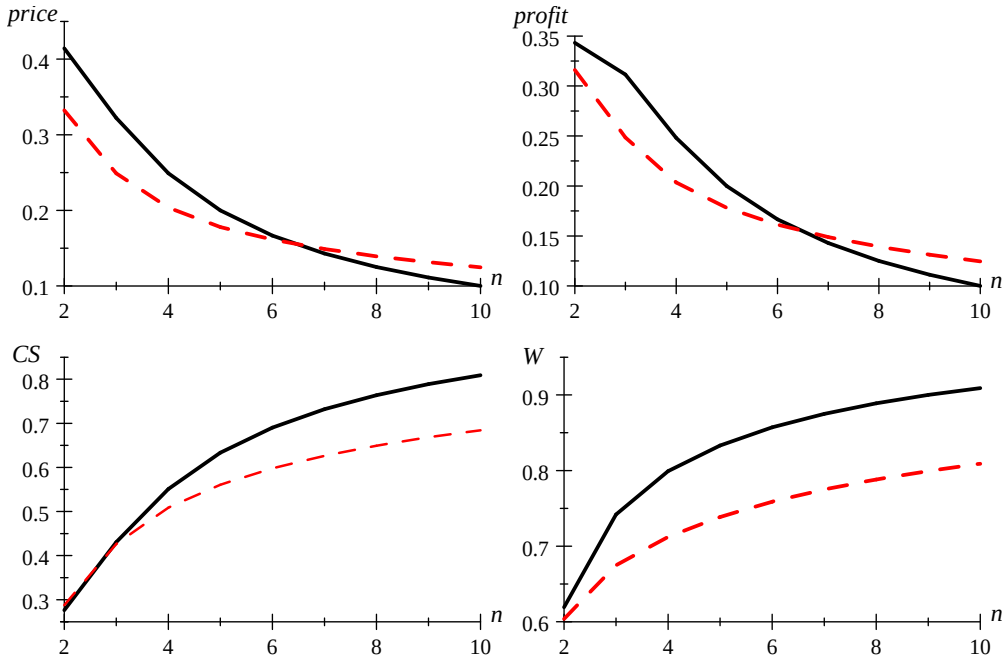


Figure 4: The impact of pure bundling—uniform example without full market coverage

3.5.3 Elastic demand

An alternative way to introduce the exclusion effect of price is to consider elastic demand. Suppose each product is divisible and consumers can buy any quantity of a product. In other aspects, for convenience let us focus on the baseline model. As in the previous section we also assume that the m products are independent to each other. In this setting with elastic demand, if a firm adopts pure bundling strategy, it requires a consumer to buy all products from it or nothing at all.

If a consumer consumes τ_i units of product i from firm j, she obtains utility $u(\tau_i) + x_i^j$, where $u(\tau_i)$ is the basic utility from consuming τ_i units of product i and x_i^j is the match utility at the product level as before. We assume that a consumer has to buy all units of a product from the same firm, and firms use linear pricing policies for each product. Denote by $v(p_i) \equiv \max_{\tau_i} u(\tau_i) - p_i\tau_i$ the indirect utility function when a consumer optimally buys product i at unit price p_i. It is clear that $v(p_i)$ is decreasing and $-v'(p_i)$ is the usual demand function.

Let p be the (symmetric) equilibrium unit price for product i in the regime of separate sales. Suppose firm j deviates and charges p'. Then the probability that a consumer will buy product i from firm j is

$$q(p') = \Pr[v(p') + x_i^j > \max_{k \neq j}\{v(p) + x_i^k\}] \ .$$

Firm j's profit from product i is then $-v'(p')p'q(p')$. It is more convenient to work on indirect utility directly. We then look for a symmetric equilibrium where each firm offers indirect utility s. Given $v(p)$ is monotonic in p, there is a one-to-one

correspondence between p and s. When a firm offers indirect utility s, it must be charging a price $v^{-1}(s)$ and the optimal quantity a consumer will buy is $-v'(v^{-1}(s))$. Denote by $r(s) \equiv v^{-1}(s)(-v'(v^{-1}(s)))$ the per-consumer profit when a firm offers indirect utility s. If firm j deviates and offers s', then the measure of consumers who choose to buy from it is

$$q(s') = \Pr[s' + x_i^j > \max_{k \neq j}\{s + x_i^k\}] = \int [1 - F(x + s - s')] dF(x)^{n-1} .$$

Then firm j's profit from its product i is $r(s')q(s')$. The first-order condition for s to be the equilibrium indirect utility is

$$-\frac{r'(s)}{r(s)} = n \int f(x) dF(x)^{n-1} .$$

If both $r(s)$ and $f(x)$ are logconcave, this is also sufficient for defining the equilibrium indirect utility. This equation has a unique solution when $r(s)$ is logconcave (so $-\frac{r'(s)}{r(s)}$ is increasing in s). Once we solve s, we can back out a unique equilibrium price $v^{-1}(s)$.[28]

In the regime of pure bundling, if a firm offers a vector of prices $(p_1, \cdots, p_m)$ and a consumer buys all products from it, then the indirect utility is $\sum_{i=1}^m v(p_i)$. If a firm offers an indirect utility S, then the optimal prices should solve the problem $\max_{\{p_i\}} \sum_{i=1}^m p_i(-v'(p_i))$ subject to $\sum_{i=1}^m v(p_i) = S$. Suppose this problem has a unique solution with all p_i being equal to each other. The optimal unit price is then $v^{-1}(\frac{S}{m})$. We look for a symmetric equilibrium where each firm offers an indirect utility S. Suppose firm j deviates to S'. Then the measure of consumers who buy all products from it is

$$Q(S') = \Pr[S' + X^j > \max_{k \neq j}\{S + X^k\}] = \Pr[\frac{S'}{m} + \frac{X^j}{m} > \max_{k \neq j}\{\frac{S}{m} + \frac{X^k}{m}\}] .$$

Firm j's profit is $mr(\frac{S'}{m})Q(S')$. So the first-order condition is

$$-\frac{r'(\frac{S}{m})}{r(\frac{S}{m})} = n \int g(x) dG(x)^{n-1} .$$

It is then clear that our previous price comparison results still hold here as long as $r(s)$ is logconcave (or equivalently $-\frac{r'(s)}{r(s)}$ is increasing in s).

3.6 Related literature on competitive pure bundling

Matutes and Regibeau (1988) initiated the study of competitive pure bundling in the context of product compatibility. They study the duopoly case in a two-dimensional

[28] In the case with a linear demand function $-v'(p) = 1 - p$, one can check that $r(s) = \sqrt{2s} - 2s$ (which is actually concave) and $-\frac{r'(s)}{r(s)}$ increases from $-\infty$ to ∞ when s varies from 0 to $\frac{1}{2}$.

Hotelling model where consumers are uniformly distributed on a square. They showed that bundling induces lower market prices and profits, and it also benefits consumers if the market is fully covered. We have derived similar results in the duopoly case under more general conditions.

Hurkens, Jeon, and Menicucci (2013) extended Matutes and Regibeau (1988) to the case with two asymmetric firms where one firm produces higher quality products than the other. The quality premium is modelled by a higher basic valuation for each product. Firm asymmetry is clearly relevant in the real market. For a general symmetric consumer distribution, under certain technical assumptions they show that when the quality difference is sufficiently large, pure bundling raises both firms' profits.[29] (The same should be true for prices though they did not state a formal result.) Our price and profit comparison results beyond duopoly have a similar intuition as theirs. In our model, for each given consumer a firm is competing with the best product among its competitors. When the number of firms increases, the best rival product improves and the asymmetry between the firm and its strongest competitor expands. (This shifts the position of the firm's marginal consumers to the right tail, and bundling makes the right tail thinner and leads to fewer marginal consumers.[30]) These two papers are complementary in the sense that they point out that either firm asymmetry or having more (symmetric) firms can reverse the usual result that pure bundling intensifies price competition. However, to accommodate more firms and more products we have adopted a different modelling approach. Our model is also more general in other aspects. For example, we can allow for potentially asymmetric products with correlated valuations, and we can also allow for a not fully covered market or elastic demand.

In the context of product compatibility in systems markets, Economides (1989) proposed a $n \times 2$ spatial model of competitive pure bundling with more than two firms and each selling two products. Specifically, consumers are distributed uniformly on the surface of a *sphere* and firms are symmetrically located on a great circle (in the spirit of the Salop circular city model). He showed that for a general transportation cost function pure bundling reduces market prices relative to separate sales. Although there is no simple comparison between the two modelling approaches, his spatial model features *local competition*: each firm is directly competing with its two neighbor rivals only (no matter in the separate sales regime or the bundling regime), and they are always symmetric to each other no matter

[29] For a medium quality difference, one firm earns more and the other earns less. See also Hahn and Kim (2012) for a similar model with uniform consumer distribution.

[30] The same argument works for the superior firm in Hurkens, Jeon, and Menicucci (2013). For the superior firm in their model, the position of its marginal consumers shifts to the left tail instead when the quality difference increases. But bundling also makes the left tail thinner when the distribution is symmetric.

how many firms in total are present in the market. While in our random utility model, there is a *global competition* and each firm is directly competing with all other firms. When the number of firms increases, each firm is effectively competing with a stronger "competitor". This expanding asymmetry, which does not exist in Economides's spatial model, drives our result that the impact of pure bundling is reversed when the number of firms is above a threshold.

A recent independent work by Kim and Choi (2014) proposed an alternative $n \times 2$ spatial model to study product compatibility. They assume that consumers are uniformly distributed on the surface of a *torus*, and firms are symmetrically located on the same surface. (In this model there are many possible ways to symmetrically locate firms.) For a quardratic transportation cost function, they show that when there are four or more firms in the market, there exists at least one symmetric location of firms under which making the products incompatible across firms raises market price and profit. This is consistent with our price comparison result. Compared to Economides (1989), a difference in their model is that more direct competition among firms is allowed when products are incompatible, and each firm competes with more firms when the number of firms increases.[31] This makes the model closer to ours.[32]

Whether a spatial model or a random utility model is more appropriate for study competitive pure bundling may depend on the context. The random utility model, however, seems more flexible and easier to use. For example, the analysis of the pure bundling regime is simpler in the random utility framework. It can also accommodate more than two products, and it can even be used to study competitive mixed bundling as we will see in next section. Neither of them is easy to deal with in a spatial model with more than two firms. The random utility approach also

[31] More specifically, a torus is just a Cartesian product of two circles. So in the regime of compatible products, for each product a firm still competes only with two neighbor rivals. While in the regime of incompatible products, a firm competes with more neighbor rivals as n increases. With the quardratic transportation cost, the density of marginal consumers in the first case actually increases in n linearly such that price decreases in an order of $1/n^2$. While the density of marginal consumers in the second case can decrease in n as long as we carefully choose the location of firms. That is why incompatibility can lead to higher market prices when n is above a certain threshold.

[32] Farrell, Monroe, and Saloner (1998) compare profitability of two forms of vertical organization of industry: open organization (which can be interpreted as separate sales) *vs* closed organization (which can be interpreted as pure bundling). They provide an alternative model with homogenous products but with heterogeneous costs (or qualities). They assume that Bertrand price competition takes place after cost uncertainty resolves and becomes public information. Then the firm with the lowest cost wins the market and charges a price equal to the second lowest cost. They show a similar result in the $n \times 2$ case that open organization generates higher prices and profits than closed organization in the duopoly case, but the opposite is true when the number of firms is large enough. (If cost realizations remain each firm's private information, their model will be closer to ours.)

accords well with econometric models of discrete consumer choice. In addition, neither Economides (1989) nor Kim and Choi (2014) investigated firms' individual incentives to bundle.

Our study of competitive pure bundling is also related to the literature on multi-object auctions with bundling. Palfrey (1983) showed that in the case with two bidders, the seller benefits from selling all the objects in a package, and the opposite is true when the number of bidders is sufficiently large. Chakraborty (1999) explicitly showed that in the two-object case, bundling benefits the seller if and only if the number of bidders is less than some threshold. The seller in an auction is like the consumer in our price competition model, and the agents on the other side of the market are competing for them. So their result about the profitability of bundling for the seller is related to our result concerning consumer surplus comparison (instead of price comparison).[33] In an auction, the profit from selling the objects separately is equal to the sum of the second highest valuations for each product. While the profit from selling them together in a bundle is equal to the second highest valuation for the bundle (which is the sum of the valuations for each separate product). In the case with two bidders, the second highest valuation is the minimum valuation, so the latter profit must be higher. In the case with a large number of bidders, the second highest valuation is close to the maximum valuation, so the former profit must be higher.

4 Mixed Bundling

In this section, we study competitive mixed bundling. All the existing research on competitive mixed bundling focuses on the duopoly case. See, for example, Matutes and Regibeau (1992), Thanassoulis (2007), and Armstrong and Vickers (2010). Armstrong and Vickers (2010) considered the most general setup in the literature so far: they allow for asymmetric products and correlated valuations, and they also consider elastic demand and a general nonlinear pricing schedule. This paper is the first to consider the case with more than two firms. As we will see below, solving the mixed bundling pricing game with more than two firms involves extra complications. One contribution of this paper is to propose a way to solve the problem, and when the number of firms is large we also show that the equilibrium prices have relatively simple approximations.

For tractability, we focus on the baseline model with full market coverage. We

[33] In terms of analysis, an important difference is that competition occurs on the uninformed side in our price competition model but on the informed side in auctions. This leads to a different mathematical treatment of equilibrium characterization. In the auction case, due to the revenue equivalence result, we can focus on the second-price auction and then the equilibrium bidding strategy is simple. This makes the analysis simpler than in our price competition model.

also assume that each firm supplies two products only and they have i.i.d. match utilities.[34] If a firm adopts a mixed bundling strategy, it offers a pair of stand-alone prices (ρ_1, ρ_2) and a joint-purchase discount $\delta > 0$. (So if a consumer buys both products from this firm, she pays $\rho_1 + \rho_2 - \delta$.) In the following, we will first argue that starting from separate sales with price p defined in (1), each firm has a unilateral incentive to introduce mixed bundling. We will then characterize the symmetric pricing equilibrium with mixed bundling and examine the impacts of mixed bundling relative to separate sales.

4.1 Incentive to use mixed bundling

Suppose all other firms are selling their products separately at price p defined in (1). Suppose firm j introduces a small joint-purchase discount $\delta > 0$ but keeps the stand-alone price p unchanged. Will this small deviation improve firm j's profit? The negative (first-order) effect is that firm j earns δ less from those consumers who buy both products from it. In the regime of separate sales the measure of such consumers is $1/n^2$, so this loss is δ/n^2.

The positive effect is that more consumers will now buy both products from firm j. Recall that

$$y_i \equiv \max_{k \neq j} \{x_i^k\}$$

denotes the match utility of the best product i among all other firms. For a given realization of (y_1, y_2) from other firms, Figure 5 below depicts how the small deviation affects consumer demand, where Ω_i, $i = 1, 2$, indicates consumers who buy only product i from firm j and Ω_b indicates consumers who buy both products from firm j. (We have suppressed the superscripts in firm j's match utilities.)

[34] It is not difficult to extend the analysis below to the case with asymmetric products and correlated valuations. Dealing with more than two product is more difficult because of the resulted complication of the pricing strategy space.

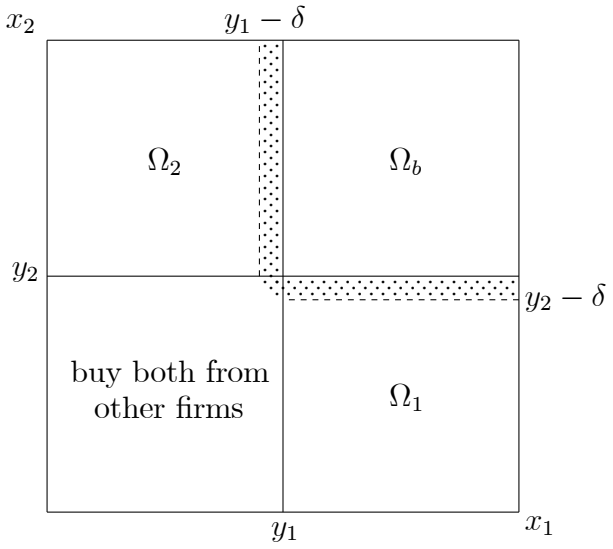


Figure 5: The impact of a joint-purchase discount on demand

When firm j introduces the small discount δ, the region of Ω_b expands but both Ω_1 and Ω_2 shrink accordingly. The shaded area indicates the increased measure of consumers who buy both products from firm j: those on the two rectangle areas switch from buying only one product to buying both from firm j, and those on the small triangle area switch from buying nothing to buying both products from firm j.

Notice that the small triangle area is a second-order effect when δ is small. So the positive (first-order) effect is only from the consumers on the two rectangle areas. The measure of them is $\delta[f(y_1)(1-F(y_2)) + f(y_2)(1-F(y_1))]$. From each of these consumers, firm j originally made a profit p but now makes a profit $2p-\delta$. Therefore, conditional on (y_1, y_2), the positive (first-order) effect of introducing a small δ is $(p-\delta)\delta[f(y_1)(1-F(y_2)) + f(y_2)(1-F(y_1))]$. Discarding higher-order effects, integrating it over (y_1, y_2) and using the symmetry yield

$$2p\delta \iint f(y_1)(1-F(y_2))dF(y_2)^{n-1}dF(y_1)^{n-1} = \frac{2}{n}p\delta \int f(y_1)dF(y_1)^{n-1} = \frac{2\delta}{n^2} .$$

(Note that the cdf of y_i is $F(y_i)^{n-1}$. The first equality used $\int(1-F(y_2))dF(y_2)^{n-1} = 1/n$, and the second one used the definition of p in (1).) Thus, the benefit is twice the loss. As a result, the proposed deviation is indeed profitable. (The spirit of this argument is similar to McAfee, McMillan, and Whinston, 1989, and Armstrong and Vickers, 2010. The former deals with a monopoly model and the latter deals with a duopoly model.)

Proposition 6 *Starting from separate sales with p defined in (1), each firm has a strict unilateral incentive to introduce mixed bundling.*

The implication of this result is that if implementing mixed bundling is feasible and costless, separate sales cannot be an equilibrium outcome. Unlike the pure

bundling case where whether a firm has a unilateral incentive to deviate from separate sales depends on the number of firms, here each firm always has an incentive to do so.

4.2 Equilibrium prices

We then turn to characterize the symmetric equilibrium (ρ, δ), where ρ is the stand-alone price for each individual product and $\delta \leq \rho$ is the joint-purchase discount.[35] Suppose all other firms use the equilibrium strategy, and firm j unilaterally deviates and sets (ρ', δ'). Then a consumer faces the following options:

- buy both products from firm j, in which case her surplus is $x_1 + x_2 - (2\rho' - \delta')$
- buy product 1 at firm j but product 2 elsewhere, in which case her surplus is $x_1 + y_2 - \rho' - \rho$
- buy product 2 at firm j but product 1 elsewhere, in which case her surplus is $y_1 + x_2 - \rho' - \rho$
- buy both products from other firms, in which case her surplus is $A - (2\rho - \delta)$ (where A will be defined below)

When the consumer buys only one product, say, product i from other firms, she will buy the one with the highest match utility y_i. When she buys both products from other firms (and $n \geq 3$), there are actually two subcases, depending on whether she buys them from the same firm or not. When y_1 and y_2 are from the same firm, the consumer buys both products at this firm, and so $A = y_1 + y_2$. This occurs with probability $\frac{1}{n-1}$. With the remaining probability $\frac{n-2}{n-1}$, y_1 and y_2 are from two different firms. Then the consumer faces the trade-off between consuming better matched products by two-stop shopping and enjoying the joint-purchase discount by one-stop shopping. In the former case, she has surplus $y_1 + y_2 - 2\rho$, and in the latter case she has surplus $z - (2\rho - \delta)$, where

$$z \equiv \max_{k \neq j} \{x_1^k + x_2^k\} \tag{20}$$

is the match utility of the best bundle among firm j's competitors. (Note that $z < y_1 + y_2$.) Hence, when y_1 are y_2 are from two different firms, $A = \max\{z, y_1 + y_2 - \delta\}$. In sum, we have

$$A = \begin{cases} y_1 + y_2 & \text{with prob. } \frac{1}{n-1} \\ \max\{z, y_1 + y_2 - \delta\} & \text{with prob. } \frac{n-2}{n-1} \end{cases}. \tag{21}$$

[35] If $\delta > \rho$, then the bundle would be cheaper than each individual product and only the bundle price would matter for consumer choice. That situation would be like pure bundling where consumers can buy multiple bundles.

The relatively simple case is when $n = 2$. Then the surplus from the fourth option is simply $y_1 + y_2 - (2\rho - \delta)$ (i.e., $A = y_1 + y_2$). The problem can then be rephrased into a two-dimensional Hotelling model by using two "location" random variables $x_1 - y_1$ and $x_2 - y_2$. That is the model in the existing literature on competitive mixed bundling.

When $n \geq 3$, the situation is more complicated. We need to deal with one more random variable z, and moreover z is correlated with y_1 and y_2 as reported in the following lemma.

Lemma 5 *When $n \geq 3$, the cdf of z defined in (20) conditional on y_1, y_2 and they being from different firms is*

$$L(z) = \frac{F(z - y_1)F(z - y_2)}{(F(y_1)F(y_2))^{n-2}} \left(F(y_2)F(z - y_2) + \int_{z-y_2}^{y_1} F(z - x)dF(x) \right)^{n-3} \tag{22}$$

for $z \in [\max\{y_1, y_2\} + \underline{x}, y_1 + y_2)$.

With this result, the distribution of A conditional on y_1 and y_2 can be fully characterized. This completes the description of the consumer's choice problem.

Given a realization of (y_1, y_2, A), the following graph describes how a consumer chooses among the above four options:

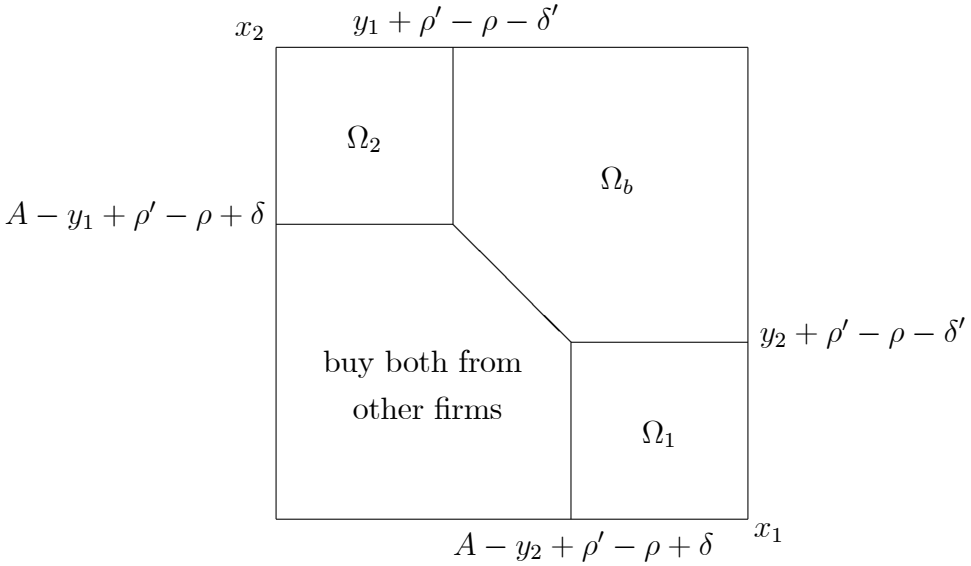


Figure 6: Price deviation and consumer choice I

As before, Ω_i, $i = 1, 2$, indicates the region where the consumer buys only product i from firm j, and Ω_b indicates the region where the consumer buys both products from firm j. Then integrating the area of Ω_i over (y_1, y_2, A) yields the demand function for firm j's *single* product i, and integrating the area of Ω_b over (y_1, y_2, A) yields the demand function for firm j's bundle.

What is useful for our analysis below is the equilibrium demand for firm j's products. From Figure 6, we can see that

$$\Omega_i(\delta) \equiv \mathbb{E}\left[F(y_j - \delta)(1 - F(A - y_j + \delta))\right], \, i = 1, 2, \; j \neq i,$$

is the *equilibrium* demand for firm j's single product i. (The expectation is taken over (y_1, y_2, A). Given full market coverage, the equilibrium demand depends only on the joint-purchase δ but not on the stand-alone price ρ.) Let $\Omega_b(\delta)$ be the equilibrium demand for firm j's bundle. Then

$$\Omega_i(\delta) + \Omega_b(\delta) = \frac{1}{n} \, . \tag{23}$$

With full market coverage, all consumers will buy product i. Since all firms are *ex ante* symmetric, the demand for each firm's product i (either from single product purchase or from bundle purchase) must be equal to $1/n$. This also implies that $\Omega_1(\delta) = \Omega_2(\delta)$.

It is useful to introduce a few more pieces of notation:

$$\begin{aligned}
\alpha(\delta) &\equiv \mathbb{E}\left[f(y_1 - \delta)\left(1 - F(A - y_1 + \delta)\right)\right], \\
\beta(\delta) &\equiv \mathbb{E}\left[f(A - y_1 + \delta)F(y_1 - \delta)\right], \\
\gamma(\delta) &\equiv \mathbb{E}[\int_{y_1 - \delta}^{A - y_2 + \delta} f(A - x)f(x)dx] \, .
\end{aligned}$$

(All the expectations are taken over (y_1, y_2, A).) The economic meanings of $\alpha(\delta)$, $\beta(\delta)$ and $\gamma(\delta)$ will be clear soon. They will help describe the marginal effect of a small price deviation by firm j on its profit.

To derive the first-order conditions for ρ and δ, let us consider the following two specific deviations: First, suppose firm j raises its joint-purchase discount to $\delta' = \delta + \varepsilon$ while keeps its stand-alone price unchanged. Then conditional on (y_1, y_2, A), Figure 7(a) below describes how this small deviation affects consumers' choices: Ω_b expands because now more consumers buy both products from firm j. The marginal consumers are distributed on the shaded area.

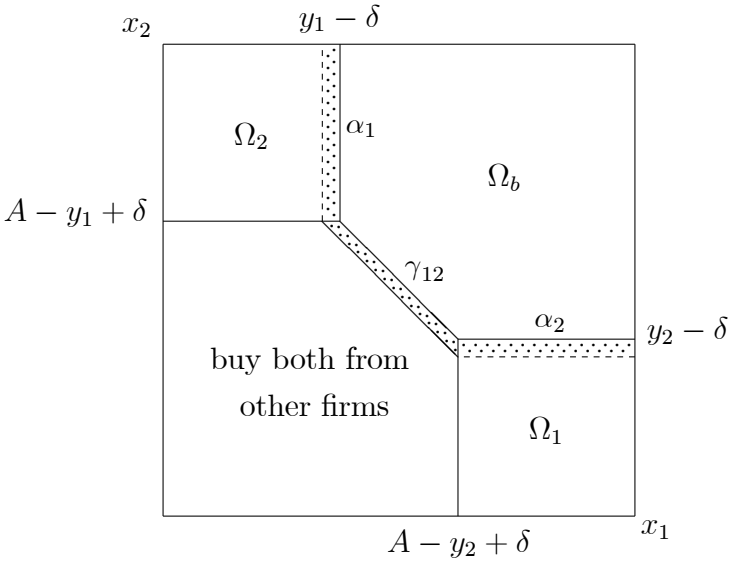


Figure 7(a): Price deviation and consumer choice II

Here α_i, $i = 1, 2$, indicates the density of marginal consumers along the line segment α_i on the graph, and so it equals $f(y_i - \delta)(1 - F(A - y_i + \delta))$. And γ_{12} indicates the density of marginal consumers along the diagonal line segment on the graph, and it equals $\int_{y_1-\delta}^{A-y_2+\delta} f(A - x)f(x)dx$. Integrating them over (y_1, y_2, A) yields the previously introduced notation: $\mathbb{E}[\alpha_1] = \mathbb{E}[\alpha_2] = \alpha(\delta)$ and $\mathbb{E}[\gamma_{12}] = \gamma(\delta)$. For the marginal consumers on the horizontal and the vertical shaded areas (which have a measure of $\varepsilon(\alpha_1 + \alpha_2)$), they now buy one more product from firm j and so firm j makes $\rho - \delta$ more money from each of them. For those marginal consumers on the diagonal shaded area (which has a measure of $\varepsilon\gamma_{12}$), they switch from buying both products from other firms to buying both from firm j. So firm j makes $2\rho - \delta$ more money from each of them. The only negative effect of this deviation is that those consumers on Ω_b who were already purchasing both products at firm j now pay ε less. Integrating the sum of all these effects over (y_1, y_2, A) should be equal to zero in equilibrium. This yields the following first-order condition:

$$2(\rho - \delta)\alpha(\delta) + (2\rho - \delta)\gamma(\delta) = \Omega_b(\delta) \ . \tag{24}$$

Second, suppose firm j raises its stand-along price to $\rho' = \rho + \varepsilon$ and its joint-purchase discount to $\delta' = \delta + 2\varepsilon$ (such that its bundle price remains unchanged). Figure 7(b) below describes how this small deviation affects consumers' choices: both Ω_1 and Ω_2 shrink because now fewer consumers buy a single product from firm j.

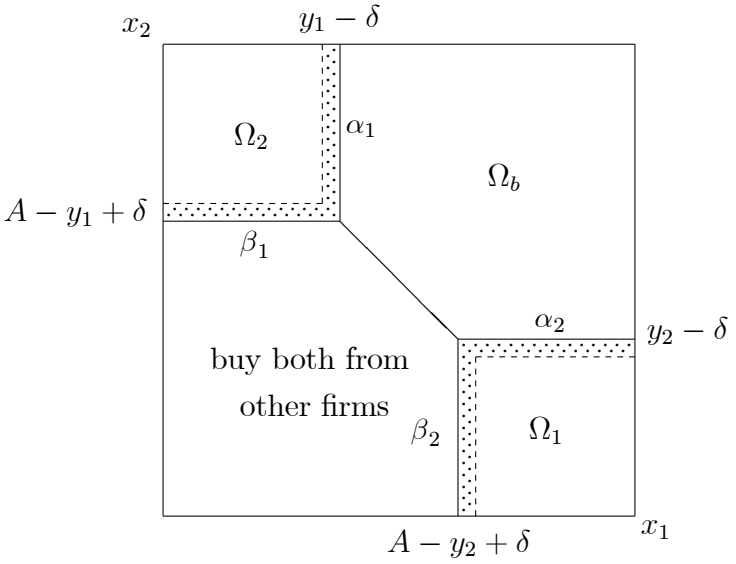


Figure 7(b): Price deviation and consumer choice III

Here, β_i, $i = 1, 2$, indicates the density of marginal consumers along the line segment β_i on the graph, and it equals $f(A-y_i+\delta)F(y_i-\delta)$. Integrating them over (y_1, y_2, A) leads to the notation $\beta(\delta)$ introduced before: $\mathbb{E}[\beta_1] = \mathbb{E}[\beta_2] = \beta(\delta)$. For those marginal consumers with a measure of $\varepsilon(\alpha_1 + \alpha_2)$, they switch from buying only one product to buying both from firm j. So firm j makes $\rho - \delta$ more money from each of them. For those marginal consumers with a measure of $\varepsilon(\beta_1 + \beta_2)$, they switch from from buying one product from firm j to buying both products from other firms. So firm j loses ρ from each of them. The direct revenue effect of the deviation is that firm j makes ε more from each consumer on Ω_1 and Ω_2 who were originally buying a single product from it. Integrating the sum of these effects over (y_1, y_2, A) should be equal to zero in equilibrium. This yields another first-order condition:

$$(\rho - \delta)\alpha(\delta) + \Omega_1(\delta) = \rho\beta(\delta) \ . \tag{25}$$

(We have used $\Omega_1(\delta) = \Omega_2(\delta)$.)

Both (24) and (25) are linear in ρ, and by using (23) it is straightforward to solve ρ as a function of δ:

$$\rho(\delta) = \frac{1/n + \delta(\alpha(\delta) + \gamma(\delta))}{\alpha(\delta) + \beta(\delta) + 2\gamma(\delta)} \ . \tag{26}$$

Substituting this into (25) yields an equation of δ:

$$\rho(\delta)\left(\beta(\delta) - \alpha(\delta)\right) = \Omega_1(\delta) - \delta\alpha(\delta) \ . \tag{27}$$

The complication of (26) and (27) comes from the fact that $\alpha(\delta)$, $\beta(\delta)$, $\gamma(\delta)$ and $\Omega_1(\delta)$ usually do not have simple expressions.[36] However, they are simple in the

[36] Another complication in mixed bundling is the second-order conditions. It is hard to analytically investigate whether the first-order conditions are also sufficient for defining a symmetric equilibrium. This issue exists even in the duopoly case (except for some specific distributions) and is an unsolved problem in general in the whole literature on mixed bundling.

duopoly case, and they also have relatively simple approximations when δ is small (which, as we will show below, is the case under mild technical conditions when the number of firms is large). Hence, in the following, we study these two polar cases.

For convenience, let $H(\cdot)$ be the cdf of $x_i - y_i$. Then

$$H(t) = \int_{\underline{x}}^{\overline{x}} F(x+t)dF(x)^{n-1}; \quad h(t) = \int_{\underline{x}}^{\overline{x}} f(x+t)dF(x)^{n-1} \ .$$

In particular, when $n = 2$, the pdf $h(t)$ is symmetric around zero, and so $h(-t) = h(t)$ and $H(-t) = 1 - H(t)$. One can also check that for any $n \geq 2$, $H(0) = 1 - \frac{1}{n}$. Notice that $h(0)$ is the density of marginal consumers for each firm. So the price (1) in the regime of separate sales can be written as

$$p = \frac{1}{nh(0)} \ .$$

In the duopoly case, $A = y_1 + y_2$ and by using the symmetry of h one can readily check that $\alpha(\delta) = \beta(\delta) = h(\delta)[1 - H(\delta)]$ and $\Omega_1(\delta) = [1 - H(\delta)]^2$. Thus, (27) simplifies to[37]

$$\delta = \frac{1 - H(\delta)}{h(\delta)} \ . \tag{28}$$

If $1 - H$ is logconcave (which is implied by the logconcavity of f), this equation has a unique positive solution. Meanwhile, (26) becomes

$$\rho = \frac{\delta}{2} + \frac{1}{4(\alpha(\delta) + \gamma(\delta))} \tag{29}$$

with $\gamma(\delta) = 2\int_0^\delta h(t)^2 dt$. In the uniform example, one can check that $\delta = 1/3$, $\rho \approx 0.572$ and the bundle price is $2\rho - \delta \approx 0.811$. Compared to the regime of separate sales where $p = 0.5$, each single product is now more expensive but the bundle is cheaper. In the normal distribution example, one can check that $\delta \approx 1.063$, $\rho \approx 1.846$ and the bundle price is $2\rho - \delta \approx 2.629$. The same observation holds.

When n is large, we conjecture that the system of (26) and (27) has a solution with δ close to zero. When δ is small, as shown in the proof of Proposition 7 below, for any given n the solution to (26) and (27) can be approximated as

$$\rho \approx \frac{1}{nh(0)} \frac{1 + \delta h(0)}{1 + \frac{n}{n-1}\delta h(0)}; \quad \delta \approx \frac{1}{\frac{2h'(0)}{h(0)} + \frac{2n^2-3n+2}{n^2-n} nh(0)} \tag{30}$$

where $h'(0) = \int f'(x)dF(x)^{n-1}$. Under mild conditions this δ is indeed close to zero as n is large. Then the approximations can be further simplified for large n.

All the main results are summarized in the following proposition.

[37] Alternatively, (28) can be written as $\Omega_1(\delta) + \frac{1}{2}\delta\Omega_1'(\delta) = 0$. This is the formula derived in Armstrong and Vickers (2010).

Proposition 7 *(i) In a symmetric mixed bundling equilibrium, the joint-purchase discount δ solves (27) and the stand-alone price ρ is given by (26).*
(ii) When $n = 2$, δ solves (28) and ρ is given by (29). The bundle price is lower than in the regime of separate sales (i.e., $2\rho - \delta < 2p$) if f is logconcave.
(iii) Suppose $\frac{|f'(x)|}{f(x)}$ is bounded and $\lim_{n\to\infty} p = 0$. When n is large, the system of (26) and (27) has a solution which can be approximated as

$$\rho \approx p; \quad \delta \approx \frac{p}{2} .$$

The bundle price is lower than in the regime of separate sales.

Result (iii) says that when n is large, the stand-alone price is approximately equal to the price in the regime of separate sales, and the joint-purchase discount is approximately half of the stand-alone price. So the mixed bundling price scheme can be interpreted as "50% off for the second product". However, this interpretation should not be taken too literally. When there is a positive production cost c for each product, we have $\delta \approx (p - c)/2$, i.e., the bundling discount is approximately half of the profit margin from selling a single product. When the number of firms is sufficiently large, it is not surprising the situation will be close to separate sales. The stand-alone price should be close to marginal cost, and so firms have no much room to offer a discount. But our hope in doing this approximation exercise is that the results can be informative for cases with a reasonably large n.

4.3 Impacts of mixed bundling

Given the assumption of full market coverage, total welfare is determined by the match quality between consumers and products. Since the joint-purchase discount induces consumers to one-stop shop too often, mixed bundling must lower total welfare relative to separate sales. In the following, we discuss the impacts of mixed bundling on profit and consumer surplus.

Let $\pi(\rho, \delta)$ be the industry profit in the symmetric mixed bundling equilibrium. Then

$$\pi(\rho, \delta) = 2\rho - n\delta\Omega_b(\delta) .$$

Every consumer buys both products, but those who buy both from the same firm pay δ less. Thus, relative to separate sales the impact of mixed bundling on profit is

$$\pi(\rho, \delta) - \pi(p, 0) = 2(\rho - p) - n\delta\Omega_b(\delta) . \tag{31}$$

Let $v(\tilde{\rho}, \tilde{\delta})$ be the consumer surplus when the stand-alone price is $\tilde{\rho}$ and the joint-purchase discount is $\tilde{\delta}$. Given full market coverage, we have $v_1(\tilde{\rho}, \tilde{\delta}) = -2$ and $v_2(\tilde{\rho}, \tilde{\delta}) = n\Omega_b(\tilde{\delta})$, where the subscripts indicate partial derivatives. This is because raising $\tilde{\rho}$ by ε will make every consumers pay 2ε more, and raising the discount $\tilde{\delta}$ by

ε will save ε for every consumer who buy both products from the same firm. Then relative to separate sales, the impact of mixed bundling on consumer surplus is

$$\begin{aligned} v(\rho,\delta)-v(p,0) &= v(\rho,\delta)-v(p,\delta)+v(p,\delta)-v(p,0) \\ &= \int_p^{\rho} v_1(\tilde{\rho},\delta)d\tilde{\rho}+\int_0^{\delta} v_2(p,\tilde{\delta})d\tilde{\delta} \\ &= -2(\rho-p)+n\int_0^{\delta}\Omega_b(\tilde{\delta})d\tilde{\delta}\ . \end{aligned} \tag{32}$$

Hence, (31) and (32) provide the formulas for calculating the welfare impacts of mixed bundling. (From these two formulas, it is also clear that mixed bundling harms total welfare given $\Omega_b(\tilde{\delta})$ is increasing in $\tilde{\delta}$.)

In the duopoly case, the impacts of mixed bundling on profits and consumer surplus are ambiguous. But Armstrong and Vickers (2010) have derived a sufficient condition under which mixed bundling benefits consumers and harms firms. (With our notation, the condition is $\frac{d}{dt}\frac{H(t)}{h(t)} \geq \frac{1}{4}$ for $t \leq 0$.) In the case with a large number of firms, as long as our approximation results in Proposition 7 hold, this must be the case.

5 Conclusion

This paper has offered a model to study competitive bundling with an arbitrary number of firms. In pure bundling part, we found that the number of firms qualitatively matters for the impact of pure bundling relative to separate sales. Under fairly general conditions, the impacts of pure bundling on prices, profits and consumer surplus are reversed when the number of firms exceeds some threshold (and the threshold can be small). This suggests that the assessment of bundling based on a duopoly model can be misleading. In the mixed bundling part, we found that solving the price equilibrium with mixed bundling is significantly more challenging when there are more than two firms. We have proposed a method to characterize the equilibrium prices, and we have also shown that they have simple approximations when the number of firms is large. Based on the approximations, we argue that mixed bundling is generally pro-competitive when the number of firms is large.

One assumption in our pure bundling model is that consumers do not buy more than one bundle. This is without loss of generality if bundling is caused by product incompatibility or high shopping costs. However, if bundling is purely a pricing strategy and if the production cost is relatively small, then it is possible that the bundle price is low enough in equilibrium such that some consumers want to buy multiple bundles in order to mix and match by themselves. Buying multiple bundles is not uncommon, for instance, in the markets for textbooks or newspapers. With the possibility of buying multiple bundles, the situation is actually similar to mixed

bundling. For example, consider the case with two products. If the bundle price is P, then a consumer faces two options: buy the best single bundle and pay price P, or buy two bundles to mix and match and pay $2P$ (suppose the unused products can be disposed freely). For consumers, this is the same as in a regime of mixed bundling with a stand-alone price P for each product and a joint-purchase discount P. Our analysis of mixed bundling can be modified to deal with this case.[38]

In the mixed bundling part, we have focused on the case with only two products. When the number of products increases, the pricing strategy space becomes more involved and we have not found a relatively transparent way to solve the model. One possible way to proceed is to consider simple pricing policies such as two-part tariffs. In addition, as we have seen, consumers face lots of purchase options in the regime of mixed bundling with many firms. The choice problem is complicated and ordinary consumers might get confused and make mistakes. Studying mixed bundling with boundedly rational consumers is also an interesting research topic.

Appendix

Proof of Proposition 1: (i) The duopoly model can be converted into a two-dimensional Hotelling setting. Let $d_i \equiv x_i^1 - x_i^2$, and let H and h be its cdf and pdf, respectively. Since x_i^1 and x_i^2 are i.i.d., d_i has support $[\underline{x} - \overline{x}, \overline{x} - \underline{x}]$ and it is symmetric around the mean zero. In particular, $h(0) = \int f(x)^2 dx$. Since x_i^j has a logconcave density, d_i has a logconcave density too. Then the equilibrium price p in the regime of separate sales is given by

$$\frac{1}{p} = 2h(0) \ .$$

Let $\bar{H}$ and $\bar{h}$ be the cdf and pdf of $\sum_{i=1}^m d_i/m$. Then the equilibrium price in the regime of pure bundling is given by

$$\frac{1}{P/m} = 2\bar{h}(0) \ .$$

Given that d_i is logconcave and symmetric, $\sum_{i=1}^m d_i/m$ is more concentrated than each d_i in the sense $\Pr(|\sum_{i=1}^m d_i/m| \leq t) \geq \Pr(|d_i| \leq t)$ for any $t \in [0, \overline{x} - \underline{x}]$. (See, e.g., Theorem 2.3 in Proschan, 1965.) This implies that $\bar{h}(0) \geq h(0)$, and so $P/m \leq p$.

[38] If we introduce shopping costs in the benchmark of separate sales (i.e., if a consumer needs to incur some extra costs when she sources from more than one firm), the situation will also be similar to mixed bundling. The shopping cost will play the role of the joint-purchase discount.

(ii) Given $f(\overline{x}) > g(\overline{x}) = 0$, the first part of the result follows from Lemma 2 immediately. To prove the second part, consider

$$\lambda(n) \equiv \int_0^1 [l_f(t) - l_g(t)]t^{n-2}dt \ . \tag{33}$$

When f is logconcave, bounded, and $f(\overline{x}) > 0$, we already knew that $\lambda(2) < 0$ and $\lambda(n) > 0$ for sufficiently large n. In the following, we show that $\lambda(n)$ changes its sign only once if $l_f(t)$ and $l_g(t)$ cross each other at most twice.

We need to use one version of the *Variation Diminishing Theorem* (see Theorem 3.1 in Karlin, 1968). Let us first introduce two concepts. A real function $K(x, y)$ of two variables is said to be *totally positive of order* r if for all $x_1 < \cdots < x_k$ and $y_1 < \cdots < y_k$ with $1 \leq k \leq r$, we have

$$\begin{vmatrix} K(x_1, y_1) & \cdots & K(x_1, y_k) \\ \vdots & & \vdots \\ K(x_k, y_1) & \cdots & K(x_k, y_k) \end{vmatrix} \geq 0 \ .$$

We also need to introduce one way to count the number of sign changes of a function. Consider a function $f(t)$ for $t \in A$ where A is an ordered set of the real line. Let

$$S(f) \equiv \sup S[f(t_1), \cdots, f(t_k)] \ ,$$

where the supremum is extended over all sets $t_1 \leq \cdots \leq t_k$ ($t_i \in A$), k is arbitrary but finite, and $S(x_1, \cdots, x_k)$ is the number of sign changes of the indicated sequence, zero terms being discarded.

Theorem 1 (Karlin,1968) *Consider the following transformation*

$$\zeta(x) = \int_Y K(x, y)f(y)d\mu(y) \ ,$$

where $K(x, y)$ is a two-dimensional Borel-measurable function and μ is a sigma-finite regular measure defined on Y. Suppose f is Borel-measurable and bounded, and the integral exists. Then if K is totally positive of order r and $S(f) \leq r - 1$, then

$$S(\zeta) \leq S(f) \ .$$

Now consider $\lambda(n)$ defined in (33). Our assumption implies that $S[l_f(t) - l_g(t)] \leq 2$. The lemma below proves that $K(t, n) = t^{n-2}$ is totally positive of order 3. Therefore, the above theorem implies that $S(\lambda) \leq 2$. That is, $\lambda(n)$ changes its sign at most twice as n varies. Given $\lambda(2) < 0$ and $\lambda(n) > 0$ for sufficiently large n, it is impossible that $\lambda(n)$ changes its sign exactly twice. Therefore, it must change its sign only once, and so $\lambda(n) < 0$ if and only if $n \leq \hat{n}$.

Lemma 6 *Let $t \in (0,1)$ and $n \geq 2$ be integers. Then t^{n-2} is strictly totally positive of order 3.*

Proof. We need to show that for all $0 < t_1 < t_2 < t_3 < 1$ and $2 \leq n_1 < n_2 < n_3$, we have $t_1^{n_1-2} > 0$,

$$\begin{vmatrix} t_1^{n_1-2} & t_1^{n_2-2} \\ t_2^{n_1-2} & t_2^{n_2-2} \end{vmatrix} > 0 \quad \text{and} \quad \begin{vmatrix} t_1^{n_1-2} & t_1^{n_2-2} & t_1^{n_3-2} \\ t_2^{n_1-2} & t_2^{n_2-2} & t_2^{n_3-2} \\ t_3^{n_1-2} & t_3^{n_2-2} & t_3^{n_3-2} \end{vmatrix} > 0 \ .$$

The first two inequalities are easy to check. The third one is equivalent to

$$\begin{vmatrix} t_1^{n_1} & t_1^{n_2} & t_1^{n_3} \\ t_2^{n_1} & t_2^{n_2} & t_2^{n_3} \\ t_3^{n_1} & t_3^{n_2} & t_3^{n_3} \end{vmatrix} > 0 \ .$$

Dividing the i_{th} row by $t_i^{n_1}$ $(i = 1, 2, 3)$ and then dividing the second column by $t_1^{n_2-n_1}$ and the third column by $t_1^{n_3-n_1}$, we can see the determinant has the same sign as

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & r_2^{\delta_2} & r_2^{\delta_3} \\ 1 & r_3^{\delta_2} & r_3^{\delta_3} \end{vmatrix} = (r_2^{\delta_2} - 1)(r_3^{\delta_3} - 1) - (r_2^{\delta_3} - 1)(r_3^{\delta_2} - 1) \ ,$$

where $\delta_j \equiv n_j - n_1$ and $r_j \equiv t_j/t_1$, $j = 2, 3$. Notice that $0 < \delta_2 < \delta_3$ and $1 < r_2 < r_3$. To show that the above expression is positive, it suffices to show that $x^y - 1$ is log-supermodular for $x > 1$ and $y > 0$. One can check that the cross partial derivative of $\log(x^y - 1)$ has the same sign as $x^y - 1 - \log x^y$. This must be strictly positive because $x^y > 1$ and $\log z < z - 1$ for $z \neq 1$. ■

(iii) Suppose x_i^j has a mean μ and variance σ^2. When m is large, by *the central limit theorem*, $X^j/m = \sum_{i=1}^m x_i^j$ distributes (approximately) according to a normal distribution $\mathcal{N}(\mu, \sigma^2/m)$. Then (7) implies that

$$\frac{P}{m} \approx \frac{p_{\mathcal{N}}}{\sqrt{m}} \ ,$$

where $p_{\mathcal{N}}$ is the separate sales price when x_i^j follows a normal distribution $\mathcal{N}(\mu, \sigma^2)$ and is independent of m. Then $\lim_{m\to\infty} \frac{P}{m} = 0$ and so $P/m < p$ for a sufficiently large m.

Proof of Proposition 2: Result (i) simply follows from Proposition 1. To prove result (ii), notice that the left-hand side of (10) increases with n while the right-hand side decreases with n given f is logconcave. So it suffices to prove two things: (a) condition (10) holds for $n = 2$, and (b) the opposite is true for a sufficiently large n.

Condition (b) is relatively easy to show. If $\lim_{n\to\infty} p = 0$, this is clearly true. But we already knew that $\lim_{n\to\infty} p$ can be strictly positive even if f is logconcave

(this requires $f(\overline{x}) = 0$). The left-hand side of (10) approaches $\overline{x} - \mu = \int_{\underline{x}}^{\overline{x}} F(x)dx$ as $n \to \infty$. So we need to show that

$$\lim_{n\to\infty} p = \frac{1 - F(\overline{x})}{f(\overline{x})} < \int_{\underline{x}}^{\overline{x}} F(x)dx \ , \tag{34}$$

where the equality is implied by (3). (When $f(\overline{x}) > 0$ or $\overline{x} = \infty$, this is obviously true.) Note that logconcave f implies logconcave $1 - F$ (or decreasing $(1 - F)/f$). So

$$\begin{aligned}
\int_{\underline{x}}^{\overline{x}} F(x)dx &= \int_{\underline{x}}^{\overline{x}} \frac{1 - F(x)}{f(x)} \frac{F(x)}{1 - F(x)} dF(x) \\
&> \frac{1 - F(\overline{x})}{f(\overline{x})} \int_{\underline{x}}^{\overline{x}} \frac{F(x)}{1 - F(x)} dF(x) \\
&= \frac{1 - F(\overline{x})}{f(\overline{x})} \int_0^1 \frac{t}{1 - t} dt \ .
\end{aligned}$$

The integral term is infinity, so condition (34) must hold.

We now prove condition (a). Using (1) and the notation $l(t) \equiv f(F^{-1}(t))$, we can rewrite (10) as

$$\int_0^1 \frac{t - t^n}{l(t)} dt \int_0^1 t^{n-2} l(t)dt < \frac{1}{n(n-1)} \ .$$

When $n = 2$, this becomes

$$\int_0^1 \frac{t(1-t)}{l(t)} dt \int_0^1 l(t)dt < \frac{1}{2} \ . \tag{35}$$

To prove this inequality, we need the following technical result:[39]

Lemma 7 *Suppose $\varphi : [0,1] \to \mathbb{R}$ is a nonnegative function such that $\int_0^1 \frac{\varphi(t)}{t(1-t)} dt < \infty$, and $r : [0,1] \to \mathbb{R}$ is a concave pdf. Then*

$$\int_0^1 \frac{\varphi(t)}{r(t)} dt \leq \max\left(\int_0^1 \frac{\varphi(t)}{2t} dt, \int_0^1 \frac{\varphi(t)}{2(1-t)} dt \right) \ .$$

Proof. Since r is a concave pdf, it is a mixture of triangular distributions and admits a representation of the form

$$r(t) = \int_0^1 r_\theta(t) \lambda(\theta) d\theta \ ,$$

where $\lambda(\cdot)$ is a pdf defined on $[0,1]$, $r_1(t) = 2t$, $r_0(t) = 2(1-t)$, and for $0 < \theta < 1$

$$r_\theta(t) = \begin{cases} 2\dfrac{t}{\theta} & \text{if } 0 \leq t < \theta \\ \\ 2\dfrac{1-t}{1-\theta} & \text{if } \theta \leq t \leq 1 \end{cases} \ .$$

[39] I am grateful to Tomás F. Móri in Budapest for helping me to prove this lemma.

(See, for instance, Example 5 in Csiszár and Móri, 2004.)

By Jensen's Inequality we have

$$\frac{1}{r(t)} = \frac{1}{\int_0^1 r_\theta(t)\lambda(\theta)d\theta} \leq \int_0^1 \frac{1}{r_\theta(t)}\lambda(\theta)d\theta \ .$$

Then

$$\int_0^1 \frac{\varphi(t)}{r(t)}dt \leq \int_0^1 \varphi(t)\left(\int_0^1 \frac{1}{r_\theta(t)}\lambda(\theta)d\theta\right)dt = \int_0^1 \left(\int_0^1 \frac{\varphi(t)}{r_\theta(t)}dt\right)\lambda(\theta)d\theta \leq \sup_{1\leq\theta\leq 1}\int_0^1 \frac{\varphi(t)}{r_\theta(t)}dt \ .$$

Notice that

$$\int_0^1 \frac{\varphi(t)}{r_\theta(t)}dt = \frac{\theta}{2}\int_0^\theta \frac{\varphi(t)}{t}dt + \frac{1-\theta}{2}\int_\theta^1 \frac{\varphi(t)}{1-t}dt \ .$$

This is a convex function of θ, because its derivative

$$\frac{1}{2}\int_0^\theta \frac{\varphi(t)}{t}dt - \frac{1}{2}\int_\theta^1 \frac{\varphi(t)}{1-t}dt$$

increases in θ. Hence, its maximum is attained at one of the endpoints of the domain $[0,1]$. This completes the proof of the lemma. ■

Now let $\varphi(t) = t(1-t)$ and

$$r(t) = \frac{l(t)}{\int_0^1 l(t)dt} \ .$$

Notice that $l(t)$ is concave when f is logconcave. So the defined $r(t)$ is indeed a concave pdf. (The integral in the denominator is finite since $l(t)$ is nonnegative and concave.) Then Lemma 7 implies that the left-hand side of (35) is actually no greater than $1/4$.[40]

Consumer surplus comparison with normal distribution. To prove (12), it suffices to establish the following result:

Lemma 8 *Consider a sequence of i.i.d. random variables $\{x^j\}_{j=1}^n$ with $x^j \sim \mathcal{N}(0,\sigma^2)$. Let p be the separate sales price as defined in (1) when each product's match utility follows the distribution of x^j. Then*

$$\mathbb{E}\left[\max_j \left\{x^j\right\}\right] = \frac{\sigma^2}{p} \ .$$

[40] For the exponential density $f(x) = e^{-x}$, the left-hand side of (35) equals $1/4$. So our result is not tight. However, if f is not logconcave, it is easy to find counterexamples. For instance, (35) fails to hold for a power distribution $F(x) = x^k$ with k close to $1/2$, or for a Weibull distribution $F(x) = 1 - e^{-x^k}$ with a small $k \in (0,1)$.

Proof. Let $F(\cdot)$ denote the cdf of x^j. Then the cdf of $\max_j \{x^j\}$ is $F(\cdot)^n$, and so

$$\mathbb{E}\left[\max_j \{x^j\}\right] = \int_{-\infty}^{\infty} x dF(x)^n = n\int_{-\infty}^{\infty} xF(x)^{n-1}f(x)dx \ .$$

For a normal distribution with zero mean, we have $f'(x) = -xf(x)/\sigma^2$. Therefore,

$$\begin{aligned}\mathbb{E}\left[\max_j \{x^j\}\right] &= -\sigma^2 n\int_{-\infty}^{\infty} F(x)^{n-1}f'(x)dx \\ &= \sigma^2 n(n-1)\int_{-\infty}^{\infty} F(x)^{n-2}f(x)^2 dx \\ &= \frac{\sigma^2}{p} \ .\end{aligned}$$

(The second step is from integration by parts, and the last step used (1).) ■

Proof of Proposition 4: Using the notation $l(t) \equiv f(F^{-1}(t))$, we rewrite (15) as

$$\Delta(n) \equiv p(1-\frac{1}{n}) - \int_0^1 \frac{t-t^{n-1}}{l(t)}dt < 0 \ .$$

In the main text, we have argued that this does not hold for $n = 2$ but holds when $n \to \infty$. So it suffices to show that $\Delta(n)$ decreases in n. But this is not obvious given $1-\frac{1}{n}$ increases in n. (For example, when the match utility distribution is close to the exponential, p is almost constant in n, and so $p(1-\frac{1}{n})$ increases in n.)

Let p_n denote the separate sales price when there are n firms. Then we have

$$\Delta(n+1) - \Delta(n) = p_{n+1}\frac{n}{n+1} - p_n\frac{n-1}{n} - \int_0^1 \frac{t^{n-1}(1-t)}{l(t)}dt \ .$$

From Lemma 1, we know that $p_{n+1} < p_n$ when f is logconcave. So

$$p_{n+1}\frac{n}{n+1} - p_n\frac{n-1}{n} < \frac{p_n}{n(n+1)} = \frac{1}{n^2(n^2-1)\int_0^1 l(t)t^{n-2}dt} \ .$$

(The equality used $\frac{1}{p_n} = n(n-1)\int_0^1 l(t)t^{n-2}dt$.) Let $\kappa(t) \equiv n(n+1)t^{n-1}(1-t)$. One can check that $\kappa(t)$ is a pdf on $[0,1]$. Then we have

$$\int_0^1 \frac{t^{n-1}(1-t)}{l(t)}dt = \frac{1}{n(n+1)}\int_0^1 \frac{\kappa(t)}{l(t)}dt > \frac{1}{n(n+1)\int_0^1 l(t)\kappa(t)dt} \ .$$

(The inequality is from Jensen's Inequality.)

Therefore, $\Delta(n+1) - \Delta(n) < 0$ if

$$n(n+1)\int_0^1 l(t)\kappa(t)dt < n^2(n^2-1)\int_0^1 l(t)t^{n-2}dt$$

$$\Leftrightarrow (n+1)^2\int_0^1 l(t)t^{n-1}(1-t)dt < (n^2-1)\int_0^1 l(t)t^{n-2}dt \ .$$

Since $t(1-t) \leq \frac{1}{4}$ for $t \in [0,1]$, this condition holds if $\frac{n+1}{4} < n-1$, which is true for any $n \geq 2$.

Proof of Lemma 4: We only prove the results for p. (The same logic works for P since the logconcavity of f implies the logconcavity of g.) Notice that $1-F$ is logconcave given f is logconcave. When $p = 0$, it is clear that the left-hand side of (18) is less than the right-hand side. We can also show the opposite is true when $p = p_M$. By using the second order statistic as in the proof of Lemma 1, the right-hand side of (18) equals

$$\frac{1-F(p)^n}{nF(p)^{n-1}f(p) + \int_p^{\overline{x}} \frac{f(x)}{1-F(x)} dF_{(2)}(x)} < \frac{1-F(p)^n}{nF(p)^{n-1}f(p) + \frac{f(p)}{1-F(p)}(1-F_{(2)}(p))} = \frac{1-F(p)}{f(p)} .$$

(The inequality is because $f/(1-F)$ is increasing, and the equality used the cdf of the second order statistic $F_{(2)}(p) = F(p)^n + nF(p)^{n-1}(1-F(p))$.) Then the fact that $p_M = \frac{1-F(p_M)}{f(p_M)}$ implies the result we want. This shows that (18) has a solution $p \in (0, p_M)$.

To show the uniqueness, we prove that the right-hand side of (18) decreases with p. One can verify that its derivative with respect to p is negative if and only if

$$f'(p)(1-F(p)^n) + nf(p)\left(F(p)^{n-1}f(p) + \int_p^{\overline{x}} f(x)dF(x)^{n-1}\right) > 0 .$$

Using $(1-F)f' + f^2 > 0$ (which is implied by the logconcavity of $1-F$), one can check that the above inequality holds if

$$n\int_p^{\overline{x}} f(x)dF(x)^{n-1} > (1-F(p)^n)\frac{f(p)}{1-F(p)} - nf(p)F(p)^{n-1} .$$

The left-hand side equals $\int_p^{\overline{x}} \frac{f(x)}{1-F(x)} dF_{(2)}(x)$, and the right-hand side equals $\frac{f(p)}{1-F(p)}(1-F_{(2)}(p))$. So the inequality is implied by $\frac{f(x)}{1-F(x)} > \frac{f(p)}{1-F(p)}$ for $x > p$.

To prove the second comparative static result, let us first rewrite (18) as

$$\frac{1}{p} = \frac{f(\overline{x}) - \int_p^{\overline{x}} f'(x)F(x)^{n-1}dx}{(1-F(p)^n)/n} = \frac{nf(\overline{x})}{1-F(p)^n} - \int_p^{\overline{x}} \frac{f'(x)}{f(x)} d\frac{F(x)^n - F(p)^n}{1-F(p)^n} . \quad (36)$$

(The first step is from integration by parts.) First of all, one can show that $\frac{n}{1-F(p)^n}$ increases with n. Second, the logconcavity of f implies $-\frac{f'}{f}$ is increasing. Third, notice that $\frac{F(x)^n - F(p)^n}{1-F(p)^n}$ is cdf of the highest order statistic of $\{x_i\}_{i=1}^n$ conditional on it being greater than p, and so it increases in n in the sense of first-order stochastic dominance. These three observations imply that the right-hand side of (36) increases with n. So the unique solution p must decrease with n.

Proof of Lemma 5: For convenience, let $I(y_i)$, $i = 1, 2$, be the identity of the firm where y_i is realized. The lower bound of z is because the lowest possible bundle

match utility at firm $I(y_i)$ is $y_i + \underline{x}$. We now calculate the conditional probability of $\max_{k \neq j}\{x_1^k + x_2^k\} < z$. This event occurs if and only if all the following three conditions are satisfied: (i) $y_1 + x_2^{I(y_1)} < z$, (ii) $x_1^{I(y_2)} + y_2 < z$, and (iii) $x_1^k + x_2^k < z$ for all $k \neq j, I(y_1), I(y_2)$. Given y_1 and y_2, condition (i) holds with probability $F(z - y_1)/F(y_2)$, as the cdf of $x_2^{I(y_1)}$ conditional on $x_2^{I(y_1)} < y_2$ is $F(x)/F(y_2)$. Similarly, condition (ii) holds with probability $F(z - y_2)/F(y_1)$. One can also check (by resorting to a graph, for example) that the probability that condition (iii) holds for a firm other than $I(y_1)$ and $I(y_2)$ is

$$\frac{1}{F(y_1)F(y_2)}\left(F(y_2)F(z-y_2) + \int_{z-y_2}^{y_1} F(z-x)dF(x)\right) .$$

Conditional on y_1 and y_2, all the above three events are independent of each other, so multiplying these probabilities yields (22).

Proof of Proposition 7: We first prove the price comparison result in (ii). From (28) and (29), we know that the bundle price in the duopoly case is $2\rho - \delta = 1/[2(\alpha + \gamma)]$ (the variable δ in $\alpha(\delta)$ and $\gamma(\delta)$ has been suppressed). And the bundle price in the regime of separate sales is $1/h(0)$. The former is lower if

$$\alpha + \gamma = h(\delta)[1 - H(\delta)] + 2\int_0^{\delta} h(t)^2 dt \geq \frac{h(0)}{2} .$$

Notice that the equality holds at $\delta = 0$. So it suffices to show that the left-hand side is increasing in δ. Its derivative is $h(\delta)^2 + h'(\delta)[1 - H(\delta)]$. This is positive if $h/(1 - H)$ is increasing or equivalently if $1 - H$ is logconcave. This is implied by the logconcavity of f.

We now prove result (iii). The key step is to show the following lemma:

Lemma 9 *If δ is small, then the system of (26) and (27) has a solution which can be approximated as*

$$\rho \approx \frac{1}{nh(0)}\frac{1 + \delta h(0)}{1 + \frac{n}{n-1}\delta h(0)}; \quad \delta \approx \frac{1}{\frac{2h'(0)}{h(0)} + \frac{2n^2-3n+2}{n^2-n}nh(0)} . \tag{37}$$

Proof. When $\delta \approx 0$, one can check that

$$\begin{aligned}
\alpha(\delta) &\approx \frac{1}{n}h(0) - \delta\left(\frac{1}{n}h'(0) + \frac{1}{n-1}h(0)^2\right) , \\
\beta(\delta) &\approx \left(1 - \frac{1}{n}\right)h(0) + \delta\left(\frac{1}{n}h'(0) - h(0)^2\right) , \\
\gamma(\delta) &\approx \frac{n\delta}{n-1}h(0)^2 , \\
\Omega_1(\delta) &\approx \frac{1}{n}\left(1 - \frac{1}{n}\right) - \frac{2\delta}{n}h(0) ,
\end{aligned} \tag{38}$$

where $h(0) = \int f(x)dF(x)^{n-1}$ and $h'(0) = \int f'(x)dF(x)^{n-1}$.[41]

To do the approximation, let us first explain how to calculate $\mathbb{E}[\psi(y_1, y_2, A)]$ for a given function $\psi(y_1, y_2, A)$, where the expectation is taken over (y_1, y_2, A). Using (21), we have

$$\mathbb{E}[\psi(y_1, y_2, A)] = \frac{1}{n-1}\mathbb{E}_{y_1,y_2}\left[\psi(y_1, y_2, y_1 + y_2)\right]$$

$$+\frac{n-2}{n-1}\mathbb{E}_{y_1,y_2}[L(y_1 + y_2 - \delta)\psi(y_1, y_2, y_1 + y_2 - \delta) + \int_{y_1+y_2-\delta}^{y_1+y_2} \psi(y_1, y_2, z)dL(z)] \ ,$$

where $L(z)$ is defined in (22). By integration by parts and using $L(y_1 + y_2) = 1$, we get

$$\mathbb{E}\left[\psi(y_1, y_2, A)\right] = \mathbb{E}_{y_1,y_2}\left[\psi(y_1, y_2, y_1 + y_2)\right] - \frac{n-2}{n-1}\mathbb{E}_{y_1,y_2}[\int_{y_1+y_2-\delta}^{y_1+y_2} \frac{\partial}{\partial z}\psi(y_1, y_2, z)L(z)dz] \ .$$

Now let us illustrate the approximation of $\alpha(\delta)$. (Others can be done similarly.) According to the formula provided above, we have

$$\alpha(\delta) = \mathbb{E}\left[f(y_1 - \delta)(1 - F(y_2 + \delta))\right] + \frac{n-2}{n-1}\mathbb{E}[\varphi(\delta)] \ , \tag{39}$$

where

$$\varphi(\delta) = \int_{y_1+y_2-\delta}^{y_1+y_2} f(y_1 - \delta)f(z - y_1 + \delta)L(z)dz \ ,$$

and the expectations are taken over y_1 and y_2. When $\delta \approx 0$, we have $f(y_1 - \delta) \approx f(y_1) - \delta f'(y_1)$, so

$$\mathbb{E}\left[f(y_1 - \delta)\right] \approx \int f(y_1)dF(y_1)^{n-1} - \delta \int f'(y_1)dF(y_1)^{n-1} = h(0) - \delta h'(0) \ .$$

We also have $1 - F(y_2 + \delta) \approx 1 - F(y_2) - \delta f(y_2)$, so

$$\mathbb{E}\left[(1 - F(y_2 + \delta))\right] \approx \int (1 - F(y_2))dF(y_2)^{n-1} - \delta \int f(y_2)dF(y_2)^{n-1} = \frac{1}{n} - \delta h(0) \ .$$

The integral term $\varphi(\delta)$ looks more complicated, but $\varphi(0) = 0$ and $\varphi'(0) = f(y_1)f(y_2)$ by noticing that $L(z)$ is independent of δ and $L(y_1 + y_2) = 1$. Hence,

$$\mathbb{E}[\varphi(\delta)] \approx \delta h(0)^2 \ .$$

Substituting these approximations into (39) and discarding all higher order terms yields the approximation for $\alpha(\delta)$.

Substituting the approximations (38) into (26) and (27) and discarding all higher order terms, one can solve ρ and δ as in (37). ■

[41] When the support of x_i is finite and $f(\overline{x}) > 0$, the density of $x_i - y_i$ has a kink at zero such that $h'(0)$ is not well defined. However, one can check that $\lim_{t\to 0^-} h'(t) = \int f'(x)dF(x)^{n-1}$ (whenever $n \geq 3$) and $\lim_{t\to 0^+} h'(t) = \int f'(x)dF(x)^{n-1} - (n-1)f(\overline{x})^2$. So we are using $h'(0^-)$ in our approximations.

When $\frac{|f'(x)|}{f(x)}$ is bounded, it is easy to see that $\frac{|h'(0)|}{h(0)}$ is bounded for any n. Since $p = \frac{1}{nh(0)}$, $\lim_{n\to\infty} p = 0$ implies that $\lim_{n\to\infty} nh(0) = \infty$. Therefore, these two technical conditions imply that the approximated δ in (37) will indeed approach zero as $n \to \infty$. Then for a large n, from (37) we can immediately see that

$$\rho \approx \frac{1}{nh(0)};\ \delta \approx \frac{1}{2nh(0)}\ .$$

References

Adams, W., and J. Yellen (1976): "Commodity Bundling and the Burden of Monopoly," *Quarterly Journal of Economics*, 90(3), 475–498.

Armstrong, M. (1996): "Multiproduct Nonlinear Pricing," *Econometrica*, 64(1), 51–75.

Armstrong, M., and J. Vickers (2010): "Competitive Non-linear Pricing and Bundling," *Review of Economic Studies*, 77(1), 30–60.

Caplin, A., and B. Nalebuff (1991): "Aggregation and Imperfect Competition: On the Existence of Equilibrium," *Econometrica*, 59(1), 25–59.

Carbajo, J., D. de Meza, and D. Seidmann (1990): "A Strategic Motivation for Commodity Bundling," *Journal of Industrial Economics*, 38(3), 283–298.

Carlton, D., and M. Waldman (2002): "The Strategic Use of Tying to Preserve and Create Market Power in Evolving Industries," *RAND Journal of Economics*, 33(2), 194–220.

Chakraborty, I. (1999): "Bundling Decisions for Selling Multiple Objects," *Economic Theory*, 13(3), 723–733.

Chan, W., H. Dong, and F. Proschan (1989): "Peakedness of Weighted Average of Jointly Distributed Random Variables," in *Contribution to Probability and Statistics: Essays in Honor of Ingram Olkin*, ed. by L. Gleser. Springer, New York.

Chen, Y. (1997): "Equilibrium Product Bundling," *Journal of Business*, 70(1), 85–103.

Chen, Y., and M. Riordan (2013): "Profitability of Product Bundling," *International Economic Review*, 54(1), 35–57.

Choi, J., and C. Stefanadis (2001): "Tying, Investment, and the Dynamic Leverage Theory," *RAND Journal of Economics*, 32(1), 52–71.

Crawford, G., and A. Yurukoglu (2012): "The Welfare Effects of Bundling in Multichannel Television Markets," *American Economic Review*, 102(2), 643–685.

Csiszár, V., and T. Móri (2004): "The Convexity Method of Proving Moment-Type Inequalities," *Statistics and Probability Letters*, 66(3), 303–313.

Economides, N. (1989): "Desirability of Compatibility in the Absence of Network Externalities," *American Economic Review*, 79(5), 1165–1181.

Fang, H., and P. Norman (2006): "To Bundle or Not To Bundle," *RAND Journal of Economics*, 37(4), 946–963.

Farrell, J., H. Monroe, and G. Saloner (1998): "The Vertical Organization of Industry: Systems Competition versus Component Competition," *Journal of Economics and Management Strategy*, 7(2), 143–182.

Gentzkow, M. (2007): "Valuing New Goods in a Model with Complementarities: Online Newspapers," *American Economic Review*, 97(3), 713–744.

Hahn, J., and S. Kim (2012): "Mix-and-Match Compatibility in Asymmetric System Markets," *Journal of Institutional and Theoretical Economics*, 168(2), 311–338.

Ho, K., J. Ho, and J. Mortimer (2012): "The Use of Full-Line Forcing Contracts in the Video Rental Industry," *American Economic Review*, 102(2), 686–719.

Hurkens, S., D. Jeon, and D. Menicucci (2013): "Dominance and Competitive Bundling," mimeo, TSE.

Karlin, S. (1968): *Total Positivity*. Stanford University Press.

Kim, S., and J. Choi (2014): "Optimal Compatibility in Systems Markets," mimeo.

Long, J. (1984): "Comments on 'Gaussian Demand and Commodity Bundling'," *Journal of Business*, 57(1), S235–246.

Matutes, C., and P. Regibeau (1988): "Mix and Match: Product Compatibility Without Network Externalities," *RAND Journal of Economics*, 19(2), 221–234.

——— (1992): "Compatibility and Bundling of Complementary Goods in a Duopoly," *Journal of Industrial Economics*, 40(1), 37–54.

McAfee, P., J. McMillan, and M. Whinston (1989): "Multiproduct Monopoly, Commodity Bundling, and Correlation of Values," *Quarterly Journal of Economics*, 104(2), 371–383.

MIRAVETE, E. (2002): "Preserving Log-Concavity Under Convolution: Comment," *Econometrica*, 70(3), 1253–1254.

NALEBUFF, B. (2000): "Competing Against Bundles," in *Incentives, Organization, and Public Economics*, ed. by G. Myles, and P. Hammond. Oxford University Press.

——— (2004): "Bundling as an Entry Barrier," *Quarterly Journal of Economics*, 119(1), 159–187.

PALFREY, T. (1983): "Bundling Decisions by a Multiproduct Monopolist with Incomplete Information," *Econometrica*, 51(2), 463–483.

PERLOFF, J., AND S. SALOP (1985): "Equilibrium with Product Differentiation," *Review of Economic Studies*, 52(1), 107–120.

PROSCHAN, F. (1965): "Peakedness of Distributions of Convex Combinations," *Annals of Mathematical Statistics*, 36(6), 1703–1706.

SCHMALENSEE, R. (1984): "Gaussian Demand and Commodity Bundling," *Journal of Business*, 57(1), S211–230.

STIGLER, G. (1968): "A Note on Block Booking," in *The Organization of Industry*, ed. by G. Stigler. University of Chicago Press.

THANASSOULIS, J. (2007): "Competitive Mixed Bundling and Consumer Surplus," *Journal of Economics & Management Strategy*, 16(2), 437–467.

WHINSTON, M. (1990): "Tying, Foreclosure, and Exclusion," *American Economic Review*, 80(4), 837–859.