

MPRA

Munich Personal RePEc Archive

Competitive Bundling

Zhou, Jidong

Yale University

12 December 2015

Online at <https://mpra.ub.uni-muenchen.de/68358/>

MPRA Paper No. 68358, posted 14 Dec 2015 17:42 UTC

Competitive Bundling*

Jidong Zhou

School of Management

Yale University

December 2015

(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 relative to separate sales pure bundling tends to raise market prices, benefit firms, and harm 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. Our analysis also sheds new light on how consumer valuation dispersion affects price competition more generally. In the regime of mixed bundling, having more than two firms raises new challenges in solving the model. We derive the equilibrium pricing conditions and 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.

Keywords: bundling, multiproduct pricing, product compatibility, oligopoly

JEL classification: D43, L13, L15

1 Introduction

Bundling is commonplace in many markets. Sometimes firms only sell packages and no individual products are available for purchase. For example, in the market for CDs,

*I am grateful for helpful comments to Simon Anderson, Mark Armstrong, Mariagiovanna Bacara, Heski Bar-Isaac, Luis Cabral, Jay Pil Choi, Ignacio Esponda, Xavier Gabaix, Phil Haile, Justin Johnson, Alessandro Lizzeri, Barry Nalebuff, Andrew Rhodes, Mike Riordan, Ran Spiegler, Gabor Virag, Glen Weyl and seminar participants in BC, Cornell, Duke-UNC, HKUST, Michigan, NYU Stern, Richmond Fed, SHUFE, SMU, Toronto, WUSTL Olin, WZB Berlin, Yale, the 13th Columbia/Duke/Northwestern/MIT IO Theory Conference, the 13th NYU IO Day, and the 2015 New Directions in Applied Microeconomics in Caltech. The research assistance of Diego Daruich is greatly appreciated.

newspapers, or cable TV (e.g., in the US), firms do not usually sell songs, articles, or TV channels separately. This is called *pure bundling*. Other examples include banking accounts, party services, and repair services tied with the product. On the other hand firms sometimes sell both a package and individual products, but the package is offered at a discounted price relative to the sum of its component prices. Relevant examples include software suites, TV-internet-phone bundles, 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.

One obvious reason for bundling is economies of scale in production, selling or buying, or complementarity in consumption. For example, traditionally it was too costly to sell newspaper articles separately. There are also other important reasons for bundling. Bundling can be a profitable price discrimination device to extract more consumer surplus (Stigler, 1968, and Adams and Yellen, 1976).¹ Bundling can also be used as a leverage device by a multiproduct firm to deter the entry of potential single-product competitors or to induce the exit of existing competitors (Whinston, 1990, and Nalebuff, 2004).²

The main anti-trust concern about bundling is that it may restrict market competition. One possible reason, as suggested by the leverage theory, is that bundling can lead to foreclosure and so a more concentrated market. Another possible reason is that even for a given market structure, bundling may relax competition and inflate market prices because it changes the space of pricing strategies. This is the main research question in the literature on competitive bundling (see, e.g., Matutes and Regibeau, 1988, and Nalebuff, 2000, for pure bundling, and Matutes and Regibeau, 1992, Anderson and Leruth, 1993, and Armstrong and Vickers, 2010, for mixed bundling). However the existing research suggests that bundle-against-bundle competition tends to be fiercer than competition with separate sales (or component pricing), and so the second possibility of inflating market prices is usually not a concern.³ Nevertheless this assessment of bundling is based on duopoly models.⁴ In this paper we will argue that considering

¹There has been a substantial body of literature which studies bundling as a price discrimination device. Most papers consider a monopoly market structure. For example, Stigler (1968), 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. In general bundling can be regarded as a nonlinear pricing scheme (see, e.g., Armstrong, 2015, for a recent survey on this topic).

²See also Choi and Stefanadis (2001), and Carlton and Waldman (2002).

³When a multiproduct firm competes with a single-product rival, if consumers have heterogeneous valuations for the additional product, bundling can create vertical product differentiation (i.e., the bundle *vs* a single product) and relax price competition. See Carbajo, de Meza and Seidmann, 1990, and Chen, 1997, for two such examples.

⁴One exception is Economides (1989). He studies competitive pure bundling when there is an arbitrary number of firms and each sells two products, but comes to the same conclusion. We will discuss this paper in detail in section 4.6.

more firms can qualitatively change our view of the impact of bundling, especially in the pure bundling case.

There are many markets where more than two firms compete with each other and adopt bundling strategies.⁵ The reason why the existing literature on competitive bundling mainly focuses on the duopoly case is partly because it has not developed a tractable enough model which can be used to study 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. This paper aims to fill this gap in the literature.

The existing works on competitive bundling use spatial models to capture product differentiation,⁶ and they often use a two-dimensional Hotelling model where consumers are distributed on a square and two multiproduct firms are located at two opposite corners. With more than two firms, however, it becomes less convenient to model product differentiation in a spatial framework. For example, if there are four firms and each sells three products, it is not obvious what spatial models will be easy to use. In this paper, we will instead adopt a multiproduct version of the random utility framework developed in Perloff and Salop (1985). Specifically, a consumer's valuation for a 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 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 converted into a two-dimensional Hotelling model (such that we can compare our results with the existing findings from the duopoly model).

Our study of how bundling affects price competition and market performance has broader implications. For example, pure bundling can be regarded as an outcome of product incompatibility. Consider a system (e.g., a computer, a smartphone, a stereo system) 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 make 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.⁷ Bundling can also arise due to shopping costs. If consumers need to incur an extra cost to source from more than one store,

⁵For example, the companies that offer the TV-internet-phone service in New York City include at least Verizon, AT&T, Time Warner, and RCN.

⁶One exception is Anderson and Leruth (1993). They use a logit model to study competitive mixed bundling in the duopoly case. Introducing product differentiation is necessary for studying competitive bundling when firms have similar cost conditions. If there is no product differentiation, prices settle at marginal costs, and so there is no meaningful scope for bundling.

⁷This is the interpretation adopted in the early works on competitive pure bundling such as Matutes and Regibeau (1988) and Economides (1989).

they have less incentive to multi-stop shop and are more likely to buy all the products they want from a single store. This is like buying the whole package from a single firm to enjoy a mixed bundling discount. If the extra shopping cost is sufficiently high, consumers are forced to behave as if they were in the pure bundling situation.

Our study is also relevant to the recent trend of unbundling in many markets (especially in online markets). For example, in the music industry nowadays consumers can download single songs from iTunes or Amazon. A similar idea is emerging in the publishing industry. For instance, Blendle, an online news platform, offers users in Netherlands and Germany access to newspaper and magazine articles on a pay-per-article basis. (A new startup CoinTent is trying to start a similar business in the US market.) In the higher education market, the rapid development of online course platforms such as Coursea is creating the possibility of unbundled higher education. Even for non-digital products, unbundling is taking place in some markets where it used to be difficult. For example, by using online platforms like Caviar and Served by Stadium, consumers can mix and match their desired dishes from different restaurants and have them delivered in one order. Unbundling benefits consumers in terms of the improved choice flexibility, but to evaluate its impact on consumer welfare we also need to understand how unbundling might change market prices. This issue is also related to the recent debate about whether US cable companies should be required to unbundle their TV packages.

In the section on pure bundling, the main message is that the number of firms can qualitatively matter for the impact of bundling (or unbundling) on prices, profits, and consumer welfare. In the duopoly case we confirm the existing findings (but in a more general setup): compared to separate sales, pure bundling intensifies price competition and lowers market prices and profits. For consumers this positive price effect often outweighs the loss from the reduced choice flexibility caused by bundling. Beyond duopoly, however, we show that under fairly general conditions the opposite is true (i.e., pure bundling raises prices, benefits firms, and harms consumers) when the number of firms is above a threshold (which can be small). This suggests that even if bundling does not influence market structure, it can be anti-competitive.

To understand these two contrasting results, first notice that pure bundling reduces consumer valuation heterogeneity and makes the distribution of consumer valuations less dispersed. Compared to the single-product valuation density function, the density function of the per-product valuation for the bundle is more peaked but has thinner tails. Intuitively, this is because finding a well-matched bundle is harder than finding a well-matched component. On the other hand, a firm's pricing decision hinges on the number of its marginal consumers who are indifferent between its product and the best product from its competitors. When there are many 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 be positioned on the right tail

of the valuation density. Since bundling yields 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.⁸ In contrast, when there are relatively few firms in the market, the average position of marginal consumers is closer to the mean. Since bundling makes the valuation density more peaked, it leads to more marginal consumers and so a more elastic demand. This induces firms to reduce their prices.

The existing research on competitive pure bundling argues that bundle-against-bundle competition is more intense than single-product competition because bundling makes a price reduction doubly profitable. (When a two-product firm reduces its price, a consumer who switches to it buys both products.) However our analysis suggests that this intuition is incomplete. Essentially it ignores the fact that bundling also changes the number of marginal consumers who will switch due to a price reduction, and this effect tends to work against the double profitability effect when there are enough firms in the market.

In the section on mixed bundling, we find that considering more than two firms raises new challenges in analysis due to the complication of the consumer choice problem. Our main contribution is to propose a method to solve the pricing game with mixed bundling, and to show that under mild conditions the equilibrium prices have simple approximations when the number of firms is large. For example, when the production cost is zero the bundle discount will be approximately equal to half of the single-product price (i.e., 50% off for the second product). In terms of the impacts of mixed bundling on profits and consumer surplus, they tend to be ambiguous in the duopoly case and depend on the distribution of consumer valuations. However with a large number of firms mixed bundling benefits consumers and harms firms under mild conditions.

We also study firms' incentives to bundle in both parts of the paper. When pure bundling is the only alternative to separate sales (e.g., when bundling is a product compatibility strategy), the number of firms matters for a firm's incentive to bundle. Bundling is the unique Nash equilibrium outcome in duopoly, but when the number of firms is above some threshold, separate sales can be an equilibrium outcome as well. In some examples separate sales is another equilibrium if and only if consumers prefer separate sales to pure bundling. When firms can choose the more flexible mixed bundling strategy, starting from separate sales each firm has a strict incentive to introduce mixed bundling, independent of the number of firms in the market. That is, when mixed bundling is feasible and costless to implement, separate sales can never be an equilibrium outcome.

Finally, our study of the benchmark case of separate sales also contributes to the

⁸More precisely, the average position of marginal consumers differs between the two regimes, and their relative distance also matters for the price comparison. That is why as we will see later there are also cases (especially when the support of the valuation distribution is unbounded) where bundling always lowers market prices.

literature on oligopolistic competition. We show that a standard log-concavity condition (which ensures the existence of pure-strategy pricing equilibrium) guarantees that market prices decline with the number of firms. This result is not new, but we offer a simple proof. We also investigate how the dispersion of consumer valuations affects price competition. This provides the foundation for the price comparison result in the pure bundling part, and it is also useful for studying the impact on price competition of any economic activities (such as information disclosure, advertising, and product design) which can change the dispersion of consumer valuations.

The rest of the paper is organized as follows: Section 2 presents the model and Section 3 analyzes the benchmark case of separate sales. Section 4 studies pure bundling, and Section 5 deals with mixed bundling. (A discussion of the related literature will be provided in each section.) We conclude in Section 6, and all omitted proofs and details 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 of bundling.) The measure of consumers is normalized to one. There are $n \geq 2$ firms, each supplying 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). Each product is horizontally differentiated across firms (e.g., each firm produces a different version of the product).⁹ 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. 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 cumulative distribution function (CDF) F with support $[\underline{x}, \bar{x}]$ (where $\underline{x} = -\infty$ and $\bar{x} = \infty$ are allowed), and for a given consumer it is realized independently across firms and products. Suppose x_i^j has a finite mean and variance and its density function f is differentiable and bounded. (In Section 4.5.1, we will consider a more general setting where a firm's m products can be asymmetric and have correlated match utilities.)¹⁰

⁹It is important to introduce product differentiation at the product level. 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.

¹⁰In the basic model, 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 which is i.i.d. across firms and consumers but has the same realization for all the m products in a firm. This is a special case of the general setting with potentially correlated match utilities in section 4.5.1.

We consider a discrete-choice framework where the incremental utility from having more than one version of a product is zero and so a consumer only wants to buy one version of each product.¹¹ We also assume that a consumer has unit demand for the desired version of each product. (Elastic demand will be discussed in Section 4.5.2.) If a consumer consumes m products with match utilities (x_1, \dots, x_m) (which can be purchased from different firms if firms are not bundling) and makes a total payment T , she obtains surplus $\sum_{i=1}^m x_i - T$.

If a firm sells its products separately, it chooses a price vector (p_1^j, \dots, p_m^j) , $j = 1, \dots, n$. If a firm adopts the pure bundling strategy, it chooses a bundle price P^j . In the first part of the paper, we assume that firms can only take one of the two selling strategies. (This is naturally the case if bundling is a product compatibility strategy or if mixed bundling is too complicated to use.) We will first study the regime of separate sales where all firms sell their products separately. We will then study the regime of pure bundling where all firms bundle their products, and compare it with the separate sales regime. Finally we will investigate firms' incentives to bundle by considering an extended game where each firm can individually choose whether to bundle its products or not. In the second part of the paper, we allow firms to use the more general mixed bundling strategy and each firm needs to specify prices P_s^j for each possible subset s of its m products. (If $m = 2$, it can be described by a pair of stand-alone prices (ρ_1^j, ρ_2^j) together with a joint-purchase discount δ^j .) In all the regimes the timing is that firms choose their prices simultaneously, and then consumers make their purchase decisions after observing all the prices and match utilities.

As often assumed in the literature on oligopolistic competition, the market is fully covered (i.e., consumers buy all the m products). This will be the case if consumers do not have outside options, or if 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 regimes of separate sales and pure bundling, we will relax this assumption in Section 4.5.2 and argue that the basic insights remain qualitatively unchanged. However, in the regime of mixed bundling this assumption is important for tractability.

¹¹This assumption is made in all the papers on competitive (pure or mixed) bundling. But it is not without loss of generality. For example, reading another article on the same subject in a different newspaper, or reading another chapter on the same topic in a different textbook, sometimes improves utility. There are works on consumer demand which extend the usual discrete choice model by allowing consumers to consume multiple versions of a product (see, e.g., Gentzkow, 2007).

3 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.¹² Suppose firm j deviates to price p' , while other firms stick to the equilibrium price p . Then the demand for firm j 's product i is

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

(In the following, whenever there is no confusion, we will suppress the integral limits \underline{x} and \bar{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 has match utility distribution $F(x)^{n-1}$ and charges p . In equilibrium the demand is $q(p) = 1/n$ since firms are symmetric to each other.

Firm j 's profit from product i is $p'q(p')$, and in equilibrium it should be maximized at $p' = p$. This yields the first-order condition for p to be the equilibrium price:

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

This condition is also sufficient for defining the equilibrium price if f is log-concave (see Caplin and Nalebuff, 1991).¹³ In the uniform distribution example with $F(x) = x$, it is easy to see that $p = 1/n$, and in the extreme value distribution example with $F(x) = e^{-e^{-x}}$ (which generates the logit model), one can check that $p = n/(n-1)$. Notice that with full market coverage, shifting the support of the match utility does not affect the equilibrium price.

In the following, we study two comparative static questions which are important for our subsequent analysis.

Price and the number of firms. The first question is: how does the equilibrium price vary with the number of firms? Let us rewrite (1) as

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

The numerator is a firm's equilibrium demand and it decreases with n . The denominator is the absolute value of a firm's equilibrium demand slope. It measures the density of a

¹²In the duopoly case Perloff and Salop (1985) have shown that the pricing game has no asymmetric equilibrium. Beyond duopoly Caplin and Nalebuff (1991) show that there is no asymmetric equilibrium in the logit model. More recently Quint (2014) proves a general result (see Lemma 1 there) which implies that our pricing game has no asymmetric equilibrium if f is log-concave.

¹³Caplin and Nalebuff (1991) provide a slightly weaker sufficient condition which requires f to be $-\frac{1}{n+1}$ -concave. Our subsequent analysis needs this to be true for any n , and when $n \rightarrow \infty$ this condition converges to zero-concavity or equivalently log-concavity.

firm's marginal consumers who are indifferent between its product and the best product among its competitors. How the denominator changes with n depends on the shape of f . For example, if the density f is increasing, it increases with n and so p must decrease with n . Conversely if f is decreasing, it decreases with n , which works against the demand size effect. However, as long as the denominator does not decrease with n too quickly, the equilibrium price decreases with n . The following result reports a sufficient condition for that.

Lemma 1 *Suppose $1 - F$ is log-concave (which is implied by the log-concavity of f). Then p defined in (1) decreases with n . Moreover, $\lim_{n \rightarrow \infty} p = 0$ if and only if $\lim_{x \rightarrow \bar{x}} \frac{f(x)}{1-F(x)} = \infty$.*

Proof. Let $x_{(n-1)}$ be the second highest order statistic of $\{x_1, \dots, x_n\}$. Let $F_{(n-1)}$ and $f_{(n-1)}$ be its CDF and density function, respectively. Using

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

we can rewrite (1) as¹⁴

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

Since $x_{(n-1)}$ increases in n in the sense of first order stochastic dominance, p decreases in n if the hazard rate $f/(1-F)$ is increasing (or equivalently, if $1-F$ is log-concave). The limit result also follows from (3) because $x_{(n-1)}$ converges to \bar{x} as $n \rightarrow \infty$. ■

Anderson, de Palma, and Nesterov (1995) is the first paper that proves this monotonicity result (see their Proposition 1). Our proof is simpler than theirs. (Their proof requires f to be log-concave, which is slightly stronger than $1-F$ being log-concave.) More recently, Quint (2014) shows that when f is log-concave, prices are strategic complements and the pricing game is supermodular in a general setting which allows for asymmetric firms and the existence of an outside option. Then the monotonicity result follows, since introducing an additional firm is the same as treating that firm as an existing one which drops its price from infinity to the equilibrium price level.¹⁵ Though less general, our method is simple and also offers a tail behavior condition for the markup to converge to zero in the limit.

¹⁴The 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 product have the same match utility. Conditional on $x_{(n-1)} = x$, the CDF of $x_{(n)}$ is $\frac{F(z)-F(x)}{1-F(x)}$ for $z \geq x$, and so its density function at $x_{(n)} = x$ is the hazard rate $\frac{f(x)}{1-F(x)}$. Integrating this according to the distribution of $x_{(n-1)}$ yields the right-hand side of (3). Dividing it by n gives the density of each firm's marginal consumers (i.e., $|q'(p)|$).

¹⁵Weyl and Fabinger (2013) make a similar observation through the lens of pass-through rate: the drop of one firm's price induces other firms to lower their prices if pass-through is below 1, and with a constant marginal cost this is true if demand is log-concave. Gabaix et al. (2015) show a similar monotonicity result when n is sufficiently large.

Notice that the log-concavity of $1 - F$ is not a necessary condition. Even if $1 - F$ is not log-concave, it is still possible that price decreases with n .¹⁶ But if $1 - F$ is log-convex (and the equilibrium price is still determined by (1)), then the same proof implies that p increases in n . The tail behavior condition for $\lim_{n \rightarrow \infty} p = 0$ is satisfied if $f(\bar{x}) > 0$. But it can be violated if $f(\bar{x}) = 0$. For instance, in the extreme value distribution example we mentioned before, the price $p = n/(n - 1)$ converges to 1 in the limit.

Price and the dispersion of consumer valuations. The second comparative static question is: if the distribution of consumer valuations becomes less “dispersed” from f to g as illustrated in Figure 1 below, how will the equilibrium price change? Intuitively, less dispersed consumer valuations mean less product differentiation across firms, and so this should intensify price competition and induce a lower market price. (This must be the case if the density g degenerates at one point such that all products become homogenous.) However, this intuition is not totally right, and g does *not* necessarily lead to a lower market price than f . As we show below, it depends on how to rank the dispersion of two random variables.

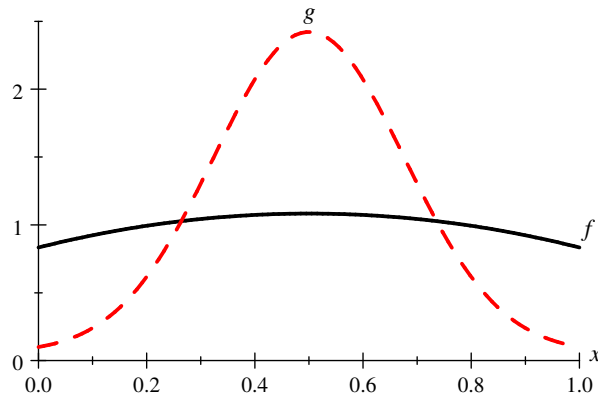


Figure 1: An example of less dispersed consumer valuations

In the literature on stochastic orders there are several possible ways to rank the dispersion of two random variables. (The classic reference on this topic is Chapter 3 in Shaked and Shanthikumar, 2007.) One of them is *convex order*. It is the most familiar one for economists because it is equivalent to a *mean-preserving spread* when two random variables have equal means.¹⁷ For example, f and g in Figure 1 can be

¹⁶One such example is the power distribution: $F(x) = x^k$ with $k \in (\frac{1}{n}, 1)$. In this example, $1 - F$ is neither log-concave nor log-convex. But one can check that the equilibrium price is $p = \frac{nk-1}{n(n-1)k^2}$ and it decreases in n .

¹⁷Let x_F and x_G be two random variables, and let F and G be their CDFs, respectively. Then x_G is smaller than x_F in the *convex order* (denoted as $x_G \leq_{\text{cx}} x_F$) if $\mathbb{E}[\phi(x_G)] \leq \mathbb{E}[\phi(x_F)]$ for any convex

ranked in this order if they have equal means.¹⁸ However, as we will see below this order usually does not ensure a clear-cut price comparison result.

Another one is *dispersive order*. A random variable x_G is said to be smaller than x_F in the dispersive order (denoted as $x_G \leq_{\text{disp}} x_F$) if $G^{-1}(t) - G^{-1}(t') \leq F^{-1}(t) - F^{-1}(t')$ for any $0 < t' \leq t < 1$, where G and F are the CDFs of x_G and x_F , respectively. (This means that the difference between any two quantiles of G is smaller than the difference between the corresponding quantiles of F .) Dispersive order ensures a clear-cut price comparison result as shown in the following result, but we will also see that it is in general a too strong condition for our bundling application.¹⁹

Lemma 2 *Consider two Perloff-Salop markets with consumer valuations denoted by x_F and x_G , respectively. Let F and G be their CDFs, f and g be their density functions, and $[\underline{x}_F, \bar{x}_F]$ and $[\underline{x}_G, \bar{x}_G]$ be their supports, respectively. Without loss of generality suppose $\mathbb{E}[x_F] = \mathbb{E}[x_G]$. Let p_k , $k = F, G$, be the equilibrium price associated with x_k . Suppose both f and g are log-concave such that the equilibrium prices are determined as in (1). (i) If x_G is less dispersed than x_F according to the dispersive order, then $p_G \leq p_F$ for any $n \geq 2$.*

(ii) However, if $f(\bar{x}_F) > g(\bar{x}_G)$, then there exists \hat{n} such that $p_G > p_F$ for $n > \hat{n}$.

Proof. Changing the integral variable from x to $t = F(x)$, we get

$$\frac{1}{p_F} = n \int_{\underline{x}_F}^{\bar{x}_F} f(x) dF(x)^{n-1} = n \int_0^1 l_F(t) dt^{n-1} ,$$

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

$$\frac{1}{p_G} = n \int_0^1 l_G(t) dt^{n-1} ,$$

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

$$p_G \leq p_F \Leftrightarrow \int_0^1 [l_F(t) - l_G(t)] dt^{n-1} \leq 0 . \quad (4)$$

(i) $x_G \leq_{\text{disp}} x_F$ if and only if $F^{-1}(t) - G^{-1}(t)$ increases in $t \in (0, 1)$. This implies that

$$\frac{dF^{-1}(t)}{dt} \geq \frac{dG^{-1}(t)}{dt} \Leftrightarrow l_F(t) \leq l_G(t) .$$

function ϕ whenever the expectations exist. When x_F and x_G have equal means, the equivalence to a mean-preserving spread is established in Theorem 3.A.1. in Shaked and Shanthikumar (2007).

¹⁸According to Theorem 3.A.44. in Shaked and Shanthikumar (2007), a sufficient condition for f to be a mean-preserving spread of g when they have equal means is that $f - g$ changes its sign twice in the order $+, -, +$. (More generally two densities ranked by convex order can cross each other many times.)

¹⁹When two random variables have equal means, dispersive order implies convex order. (See Theorem 3.B.16. in Shaked and Shanthikumar, 2007.) Ganuza and Penalva (2010) use these two ordering concepts to study information disclosure in auctions.

Therefore, $p_G \leq p_F$ follows from (4). (In particular, $x_G \leq_{\text{disp}} x_F$ implies $l_F(1) \leq l_G(1)$, or $f(\bar{x}_F) \leq g(\bar{x}_G)$.)

(ii) $f(\bar{x}_F) > g(\bar{x}_G)$ implies $l_F(1) - l_G(1) > 0$. Then

$$\lim_{n \rightarrow \infty} \int_0^1 [l_F(t) - l_G(t)] dt^{n-1} = l_F(1) - l_G(1) > 0,$$

since $l_F(t) - l_G(t)$ is bounded (given we consider bounded density functions) and the distribution t^{n-1} converges to the upper bound 1 as $n \rightarrow \infty$. Then it follows from (4) that $p_G > p_F$ when n is sufficiently large.²⁰ ■

Result (i) shows that if one distribution is less dispersed than the other in the dispersive order, the usual intuition works and less dispersed consumer valuations lead to a lower market price. Perloff and Salop (1985) show that if $x_G = \theta x_F$ with $\theta \in (0, 1)$, then $p_G < p_F$ (more precisely, $p_G = \theta p_F$). This is a special case of result (i) since $\theta x <_{\text{disp}} x$ for any random variable x and constant $\theta \in (0, 1)$. (Here “ $<_{\text{disp}}$ ” denotes a strict dispersive order.)

However, dispersive order is a relatively strong condition. When x_F and x_G have the same finite support, $x_G \leq_{\text{disp}} x_F$ requires $F^{-1}(t) - G^{-1}(t)$ increase in $t \in (0, 1)$, but this implies $F^{-1}(t) = G^{-1}(t)$ everywhere, i.e., the two random variables must be equal. This excludes many natural cases where one random variable is intuitively less dispersed than the other. For instance, the two distributions in Figure 1 cannot be ranked by the dispersive order. (When x_F and x_G have equal means and their supports are intervals, $x_G \leq_{\text{disp}} x_F$ requires that the support of x_G is a strict subset of the support of x_F , or both are infinite supports.)

Notice that $f(\bar{x}_F) > g(\bar{x}_G)$ is not compatible with $x_G \leq_{\text{disp}} x_F$ as we already see from the proof. So result (ii) indicates that if we go beyond the dispersive order, even in natural cases such as the example in Figure 1 where one distribution is intuitively less dispersed than the other, the number of firms can matter for price comparison. When there are sufficiently many firms, a less dispersed distribution can lead to a higher market price. Since this result is crucial for understanding our price comparison result in the pure bundling part, we explain its economic intuition in detail.

Let us consider the example in Figure 1 where $f(1) > g(1)$. From (2) we already know that equilibrium price equals the ratio of equilibrium demand to the negative of equilibrium demand slope. Since equilibrium demand is always $1/n$ due to firm symmetry, only equilibrium demand slope (or the density of marginal consumers) matters for price comparison. Let firm j be the firm in question. When n is large, a given consumer’s valuation for the best product among firm j ’s competitors must be close

²⁰This argument cannot be extended to the case where $f(\bar{x}_F) = g(\bar{x}_G)$ but $f > g$ for x close to the upper bounds. If $l_F(1) = l_G(1)$, then for a large n , $[l_F(t) - l_G(t)](n-1)t^{n-2}$ is close to zero everywhere (and it equals zero at $t = 1$). Then the sign of $\int_0^1 [l_F(t) - l_G(t)] dt^{n-1}$ does not necessarily depend only on the sign of $l_F(t) - l_G(t)$ at t close to 1.

to the upper bound 1 almost for sure. For that consumer to be firm j 's marginal consumer, her valuation for its product should also be close to 1. In other words, when n is large, a firm's marginal consumers should be positioned close to the upper bound no matter which density function applies. Since $f(1) > g(1)$, we deduce that a firm has fewer marginal consumers and so faces a less elastic demand when the density g applies. Therefore, when n is large, the less dispersed density g leads to a higher market price. (The intuition here is explained when n is large, but as we will see in the next section the result can hold even for a small n .) Our discussion suggests that when the number of firms is large, the tail behavior, instead of the peakedness, of the consumer valuation density matters for price comparison.²¹

Lemma 2 has its own interest in the literature on oligopolistic price competition. As well as its bundling application in the next section, it is useful for studying the impact on price competition of firm or consumer activities (such as information disclosure/acquisition, advertising, product design, and spurious product differentiation) which can change the dispersion of consumer valuations in the market.

4 Pure Bundling

Now consider the regime where all firms adopt the pure bundling strategy. We assume that consumers do not buy more than one bundle to mix and match by themselves. This is naturally the case if pure bundling is caused by product incompatibility or high shopping costs. When pure bundling is a pricing strategy, this assumption can be justified if the bundle is too expensive (e.g., due to high production costs) relative to the match utility improvement from mixing and matching. (If the unit production cost is c for each product, a sufficient condition is $c > \bar{x} - \underline{x}$.) As we will discuss in the conclusion, allowing consumers to buy multiple bundles will make the situation similar to mixed bundling.

4.1 Equilibrium prices

Let $X^j \equiv \sum_{i=1}^m x_i^j$ be the match utility of firm j 's bundle, and let P be the equilibrium bundle price. If firm j unilaterally deviates and charges P' , the demand for its bundle is

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

We divide everything by m because we want to compare the per-product bundle price P/m with the single-product price p in the benchmark regime of separate sales. Let G

²¹Gabaix et al. (2015) study the asymptotic behavior of the equilibrium price and make a similar point. By using extreme value theory they show that when the number of firms is large, the price is proportional to $[nf(F^{-1}(1 - 1/n))]^{-1}$. By noticing $\int_0^1 t dt^{n-1} = 1 - 1/n$, this can also be intuitively seen from the proof of our Lemma 2.

and g denote the CDF and density function of X^j/m , respectively. Then P/m is determined similarly as the separate sales price p , except that now a different distribution G applies:

$$\frac{1}{P/m} = n \int g(x) dG(x)^{n-1} . \quad (5)$$

Notice that g is log-concave if f is log-concave (see, e.g., Miravete, 2002). Therefore, the first-order condition (5) is also sufficient for defining the equilibrium bundle price if f is log-concave. Also notice that $1 - G$ is log-concave if $1 - F$ is log-concave. Hence, similar results as in Lemma 1 hold here.

Lemma 3 *Suppose $1 - F$ is log-concave (which is implied by the log-concavity of f). Then the bundle price P defined in (5) decreases with n . Moreover, $\lim_{n \rightarrow \infty} P = 0$ if and only if $\lim_{x \rightarrow \bar{x}} \frac{g(x)}{1-G(x)} = \infty$.*

Notice that the per-product bundle valuation X^j/m is a mean-preserving contraction of x_i^j (provided that the mean of x_i^j exists) and they have the same support. So g is less dispersed than f as illustrated in Figure 1 above. In particular, $g(\bar{x}) = 0$ even if $f(\bar{x}) > 0$ because $X^j/m = \bar{x}$ only if $x_i^j = \bar{x}$ for all $i = 1, \dots, m$. Intuitively this is because finding a well-matched bundle is much harder than finding a well-matched single product.²²

4.2 Comparing prices and profits

From (1) and (5), we can see that the comparison between separate sales and pure bundling is just a comparison between two Perloff-Salop models with two different consumer valuation distributions F and G . According to result (i) in Lemma 2, bundling reduces market price if $X^j/m \leq_{\text{disp}} x_i^j$. However, X^j/m and x_i^j often cannot be ranked by the dispersive order (e.g., when x_i^j has a finite support).²³ We will show that in the duopoly case, bundling leads to lower prices even if X^j/m and x_i^j are not ranked by the dispersive order, but if we go beyond duopoly, result (ii) in Lemma 2 implies that bundling can raise market prices.

Using the technique in the proof of Lemma 2, we have

$$\frac{P}{m} \leq p \Leftrightarrow \int_0^1 [l_F(t) - l_G(t)] t^{n-2} dt \leq 0 , \quad (6)$$

²²Formally, when $m = 2$ the density function of $(x_1^j + x_2^j)/2$ is $g(x) = 2 \int_{2x-\bar{x}}^{\bar{x}} f(2x-t) dF(t)$ for $x \geq (\underline{x} + \bar{x})/2$, so $g(\bar{x}) = 0$. A similar argument works for $m \geq 3$.

²³Given the further restriction here that X^j/m and x_i^j have the same support and mean (which means that F and G must cross each other at least once), they also cannot be ranked by the dispersive order if x_i^j has a semi-infinite support with a finite lower bound or upper bound. Hence, the only case where X^j/m and x_i^j might be ranked by the dispersive order is when the support of x_i^j is the whole real line. The only class of distributions for which this is the case is the stable distributions, as we will discuss later.

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

Proposition 1 *Suppose f is log-concave.*

- (i) *When $n = 2$, bundling reduces market prices and profits for any $m \geq 2$.*
- (ii) *For a fixed m , if $f(\bar{x}) > 0$, there exists \hat{n} such that bundling increases market prices and profits for $n > \hat{n}$ (and $\lim_{n \rightarrow \infty} \frac{P/m}{p} = \infty$). 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 , P/m decreases in m when m is large and $\lim_{m \rightarrow \infty} P/m = 0$, so there exists \hat{m} such that bundling reduces market prices and profits for $m > \hat{m}$.*

Result (i) generalizes the existing finding on how pure bundling affects market prices in duopoly. Bundling reduces price in duopoly if $\int f(x)^2 dx \leq \int g(x)^2 dx$. The intuition is more transparent when the density function f is symmetric. In that case the average position of marginal consumers is at the mean and g is more peaked than f at the mean. Each firm therefore has more marginal consumers in the case of g , and so they face a more elastic demand and charge lower prices.²⁴

Result (iii) follows from the *law of large numbers*. Let $\mu < \infty$ be the mean of x_i^j . Then X^j/m converges to μ as $m \rightarrow \infty$. In other words, with many products the per-product valuation for the bundle tends to be homogeneous across both consumers and firms. Then P/m must converge to zero.²⁵

The first part of result (ii) follows immediately from result (ii) in Lemma 2 as $f(\bar{x}) > g(\bar{x}) = 0$. Bundling generates a thinner right tail of the valuation density, and when n is large the marginal consumers are located on the right tail. Hence, bundling reduces the number of marginal consumers, and this leads to a less elastic demand and a higher market price. The limit result as $n \rightarrow \infty$ indicates that the increase of price caused by bundling can be proportionally significant. (With $f(\bar{x}) > 0$ both p and P/m converge to zero, but they converge in different speeds.) Extra work is needed to prove the cut-off result in the second part. One way to interpret the economic meaning of the condition $f(\bar{x}) > 0$ is that the equilibrium price p in the regime of separate sales converges to the marginal cost fast enough (at a speed of $1/n$).²⁶

To illustrate the cut-off result, consider two examples which satisfy all the conditions needed in result (ii). In the uniform distribution example with $f(x) = 1$, $P/m < p$ if $n \leq 6$. Figure 2(a) below describes how both prices vary with n (where the solid curve

²⁴Notice that we are calculating P/m instead of the bundle price P . So a firm's marginal consumers in the bundling case are those who will switch when the firm changes P/m by a small ε .

²⁵Nalebuff (2000) shows a similar result in a multi-dimensional Hotelling model with two firms, an arbitrary number of products, and consumers uniformly distributed inside a hypercube.

²⁶As we discussed in footnote 21, when n is large the equilibrium price p is proportional to $[nf(F^{-1}(1 - 1/n))]^{-1}$. When $f(\bar{x}) > 0$, this converges to zero at a speed of $1/n$.

is p and the dashed one is P/m). In the example with an increasing density $f(x) = 4x^3$, as described in Figure 2(b) below $P/m < p$ only if $n = 2$. These examples show that the threshold \hat{n} can be small.

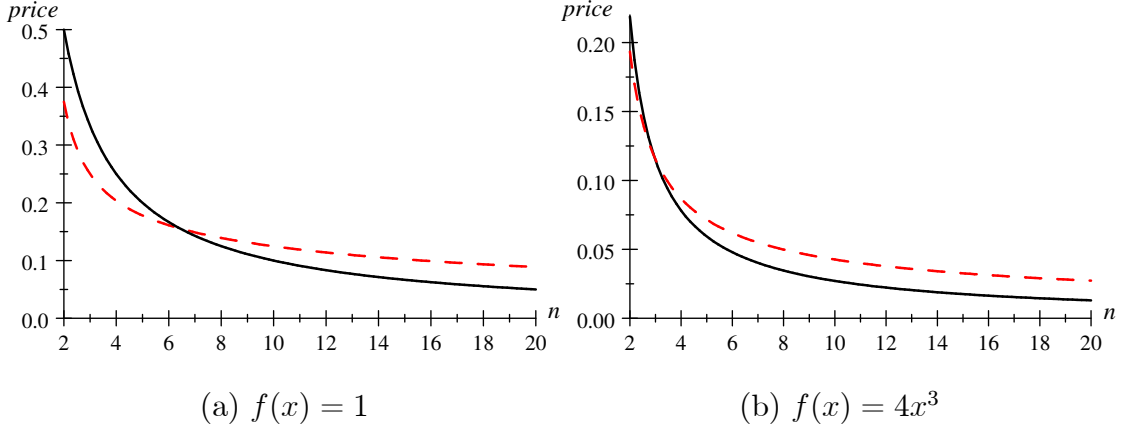


Figure 2: Price comparison with $m = 2$

Result (ii) in Proposition 1 requires $f(\bar{x}) > 0$. If $f(\bar{x}) = 0$ (where \bar{x} can be infinity), then $f(\bar{x}) = g(\bar{x})$ and the result may not hold any more. For instance, in the example of normal distribution where $\lim_{x \rightarrow \infty} f(x) = 0$, bundling always lowers market prices.

Example of normal distribution. We already know that with full market coverage, 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

$$p = \frac{\sigma}{n \int_{-\infty}^{\infty} \phi(x) d\Phi(x)^{n-1}}, \quad (7)$$

where Φ and ϕ are the CDF and density function of the standard normal distribution $\mathcal{N}(0, 1)$, respectively. The definition of X^j implies that $X^j/m \sim \mathcal{N}(0, \sigma^2/m)$. Thus, $X^j/m = x_i^j/\sqrt{m} <_{\text{disp}} x_i^j$, so result (i) in Lemma 2 implies $P/m < p$. That is, in this example bundling always reduces market prices (and so profits) regardless of n and m .²⁷ A more precise relationship between the two prices is available: Firm j 's demand in the bundling regime, when it unilaterally deviates to price P' , is

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

²⁷However, for any truncated normal distribution with a finite upper bound, result (ii) in Proposition 1 still applies. For instance, for the truncated standard normal with support $[-1, 1]$ bundling leads to a higher market price if $n > 9$.

This equals the demand for firm j 's product i in the separate sales regime when firm j charges P'/\sqrt{m} and other firms charge P/\sqrt{m} . Then we deduce that

$$\frac{P}{\sqrt{m}} = p. \quad (8)$$

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

The key feature in the normal distribution example is that x_i^j and X^j/m belong to the same class of distributions, such that the dispersive order result in Lemma 2 can apply. More generally this is a property of *stable distributions*.²⁸ Three notable examples of stable distributions are normal, Cauchy, and Lévy. Suppose x_i^j has a stable distribution with a stability parameter $\alpha \in (0, 2]$ and a location parameter zero. (Normal distribution has $\alpha = 2$, Cauchy distribution has $\alpha = 1$, and Lévy distribution has $\alpha = 1/2$.) Then using the results in Chapter 1.6 in Nolan (2015) one can show that $X^j/m = m^{\frac{1}{\alpha}-1}x_i^j$. Therefore, $X^j/m \leq_{\text{disp}} x_i^j$ and bundling reduces market prices if $\alpha \geq 1$. (In the edge case with $\alpha = 1$, i.e., with the Cauchy distribution, bundling does not affect market prices.) If $\alpha < 1$, bundling raises market prices. (Note that this does not contradict with the duopoly result in Proposition 1 because a stable distribution with $\alpha < 1$ does no longer have a log-concave density.)²⁹

Finally, we want to point out that $f(\bar{x}) > 0$ (or a finite \bar{x}) is not necessary for bundling to raise market prices even if we keep the log-concavity condition. For instance, consider the distribution with a log-concave density $f(x) = 2(1-x)$ on $[0, 1]$. In this example, $f(\bar{x}) = 0$, but numerical simulations suggest a similar price comparison result

²⁸Let x_1 and x_2 be independent copies of a random variable x . Then x is said to be *stable* if for any constants $a > 0$ and $b > 0$ the random variable $ax_1 + bx_2$ has the same distribution as $cx + d$ for some constants $c > 0$ and d .

²⁹All the discussion here is subject to the qualification that for some stable distributions, the first-order condition may not be sufficient for defining the equilibrium price or the integral in the first-order condition may not even exist.

as in Figure 2 (though the threshold \hat{n} is bigger).³⁰ There are also examples with $\bar{x} = \infty$ and a log-concave density where bundling can raise market prices. For instance, consider the generalized normal distribution with density $f(x) = \frac{\beta}{2\Gamma(1/\beta)}e^{-|x|^\beta}$, where β is the shape parameter and the support is the whole real line. (The density function is log-concave when $\beta > 1$.) This distribution becomes the standard normal when $\beta = 2$, and it converges to the uniform distribution on $[-1, 1]$ when $\beta \rightarrow \infty$. Suppose \hat{n} is the threshold in the case of uniform distribution with support $[-1, 1]$. Then for any $n > \hat{n}$, there must exist a sufficiently large β such that bundling raises market prices.

4.3 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. Since pure bundling eliminates the opportunities for consumers to mix and match, 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, then it must harm consumers since consumers suffer from both higher prices and having no opportunities to mix and match. The trickier situation is when pure bundling lowers market prices (e.g., when $n = 2$, m is large, or the distribution is normal). In that case there is a trade-off between the negative match quality effect and the positive price effect. The main message in this section is that even if bundling intensifies price competition, the negative match quality effect often dominates such that bundling harms consumers when the number of firms is above a usually small threshold.

The per-product consumer surplus in the regime of separate sales and pure bundling are respectively

$$\mathbb{E} \left[\max_j \{x_i^j\} \right] - p \quad \text{and} \quad \mathbb{E} \left[\max_j \left\{ \frac{X^j}{m} \right\} \right] - \frac{P}{m} .$$

Then pure bundling benefits consumers if and only if

$$\mathbb{E} \left[\max_j \{x_i^j\} \right] - \mathbb{E} \left[\max_j \left\{ \frac{X^j}{m} \right\} \right] < p - \frac{P}{m} . \quad (9)$$

The left-hand side (which must be positive) reflects the match quality effect, and the right-hand side is the price effect.

Proposition 2 *Suppose f is log-concave.*

(i) *For a fixed m , if $f(\bar{x}) > 0$, or if $\lim_{x \rightarrow \bar{x}} \frac{d}{dx} \left(\frac{1-F(x)}{f(x)} \right) = 0$, there exists \hat{n} such that bundling harms consumers if $n > \hat{n}$.*

³⁰In fact, it can be shown that if $f(\bar{x}) = 0$ but $f'(\bar{x}) < 0$, we still have the result that bundling raises market prices when n is above a certain threshold. The proof, though less transparent, is available upon request.

(ii) There exists n^* such that (a) for $n \leq n^*$, there exists $\hat{m}(n)$ such that bundling benefits consumers if $m > \hat{m}(n)$, and (b) for $n > n^*$, there exists $\hat{m}(n)$ such that bundling harms consumers if $m > \hat{m}(n)$.

From the price comparison result (ii) in Proposition 1, we know that if $f(\bar{x}) > 0$ bundling will raise prices (and so harm consumers) when n is sufficiently large. If $f(\bar{x}) = 0$ (e.g., when $\bar{x} = \infty$), bundling may lower prices. But the negative match quality effect will always dominate when n is sufficiently large if the second condition in result (i) holds (which is true for many often used distributions such as normal, exponential, extreme value, and logistic). This is easy to understand when $\bar{x} = \infty$. In that case, the difference between $\mathbb{E}[\max_j \{x_i^j\}]$ and $\mathbb{E}[\max_j \{\frac{X^j}{m}\}]$ can go to infinity as $n \rightarrow \infty$, while the price difference is always finite since both prices decrease with n under the log-concavity condition.

Result (ii) says that in the limit case with $m \rightarrow \infty$ we have a stronger cut-off result: pure bundling improves consumer welfare if and only if the number of firms is below some threshold. Notice that in the limit case with $m \rightarrow \infty$ we have $\lim_{m \rightarrow \infty} X^j/m = \mu$ and $\lim_{m \rightarrow \infty} P/m = 0$. Then for fixed n , bundling benefits consumers if and only if

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

The match quality effect on the left-hand side increases with n , while the price effect decreases with n . In the proof we show that (10) holds for $n = 2$ but fails for a sufficiently large n . This proves the cut-off result. Intuitively, when the number of firms increases, bundling deprives consumers of more and more opportunities to mix and match, such that eventually the match quality effect dominates. The threshold n^* is typically 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 small m it appears difficult to prove a cut-off result.³¹ 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). The threshold is also 3.³²

³¹For a finite m , we have examples where bundling harms consumers even in the duopoly case. One such example is when $m = 2$ and the distribution is exponential.

³²In this example, the harm of bundling will disappear eventually as $n \rightarrow \infty$. This is because $\lim_{n \rightarrow \infty} p = \lim_{n \rightarrow \infty} P/m = 0$ and $\lim_{n \rightarrow \infty} \mathbb{E}[\max_j \{x_i^j\}] = \lim_{n \rightarrow \infty} \mathbb{E}[\max_j \{X^j/m\}] = \bar{x}$. But for a not too large n , the harm of bundling on consumers can be significant. For example, when $n = 10$ bundling reduces consumer surplus by about 15%.

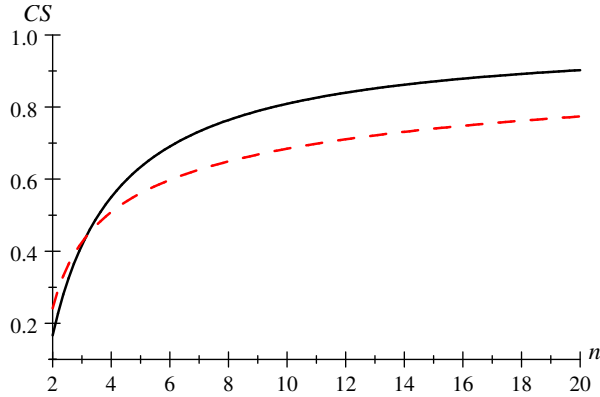


Figure 3: Consumer surplus comparison with uniform distribution and $m = 2$

A similar cut-off result holds in the normal distribution example for any $m \geq 2$.

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

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

In the Appendix, we show that

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

Then (11) simplifies to $p > \sigma$. Using (7), one can check that this holds only for $n = 2, 3$, so the threshold is also 3.

4.4 Incentive to bundle

This section studies firms' incentive to bundle. Consider an extended game where firms can choose both whether to bundle their products and what prices to set. When there are more than two products, for tractability we assume that each firm either bundles all its products or not at all and there are no finer bundling strategies (by which a firm sells some products in a package but sells others separately). The pricing game where firms adopt asymmetric partial bundling strategies is hard to analyze.

If firms can collectively choose their bundling strategies, Proposition 1 implies that they tend to choose separate sales when n is small but pure bundling when n is large. The outcome is different if firms choose their bundling strategies non-cooperatively. In the following, we will focus on the case where firms choose bundling strategies and prices simultaneously. This captures the situations where it is relatively easy to adjust the bundling strategy.

Proposition 3 *Suppose f is log-concave and firms make bundling and pricing decisions simultaneously.*

(i) *It is a Nash equilibrium that all firms choose to bundle their products and charge the bundle price P defined in (5). When $n = 2$, this is the unique (pure-strategy) Nash equilibrium if $p \neq P/m$.*

(ii) *There exists \tilde{n} such that (a) for $n \leq \tilde{n}$, there exists $\tilde{m}(n)$ such that separate sales is not a Nash equilibrium if $m > \tilde{m}(n)$, and (b) for $n > \tilde{n}$, there exists $\tilde{m}(n)$ such that separate sales is also a Nash equilibrium if $m > \tilde{m}(n)$.*

It is easy to understand that it is a Nash equilibrium that all firms bundle. This is simply because in our model if a firm unilaterally unbundles, the market situation does not change for consumers.³³ In the duopoly case, it can be further shown that neither separate sales nor asymmetric equilibria (where one firm bundles and the other does not) can arise in the market.

When there are more than two firms, one may wonder whether separate sales can be another equilibrium as well. Result (ii) says that when $m \rightarrow \infty$, this is the case if and only if n is above some threshold. The intuition of why the number of firms matters is that the more firms in the market, the worse a firm's bundle appears when it unilaterally bundles. (This is not true in the duopoly case where one firm bundling is the same as both firms bundling.) More formally, suppose that all other firms offer separate sales at price p , but firm j bundles unilaterally. Denote by

$$y_i \equiv \max_{k \neq j} \{x_i^k\} \quad (13)$$

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

$$\Pr(X^j > Y) \leq \Pr(X^j > \max_{k \neq j} \{X^k\}) = \frac{1}{n}. \quad (14)$$

The inequality is because Y is greater than $\max_{k \neq j} \{X^k\}$ stochastically, and it is strict when $n \geq 3$. Thus, without further price adjustment it cannot be profitable for firm j to unilaterally bundle.

Suppose now firm j also adjusts its price. It is more convenient to rephrase the problem into a monopoly one where a consumer's net valuation for product i is $u_i \equiv x_i - (y_i - p)$. (Here $y_i - p$ is regarded as the outside option to product i .) If firm j does not bundle, then its optimal separate sales price is p and its profit from each product is p/n . But its optimal profit when it bundles is hard to calculate in general, except in the limit case with $m \rightarrow \infty$. In this limit case, according to the *law of large numbers*

³³This argument depends on the assumptions that consumers buy all products and for each product they only buy one variant.

firm j can extract all surplus by charging a bundle price $m \times \mathbb{E}[u_i]$ and its per product profit will be $\mathbb{E}[u_i] = \mu - \mathbb{E}[y_i] + p$.³⁴ This is no greater than the separate sales profit (and so firm j has no unilateral incentive to bundle) if and only if

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

This is clearly not true for $n = 2$ (which is consistent with result (i) in Proposition 3). In the proof, we show that this is true if and only if n is above a certain threshold. (This argument assumes $\mathbb{E}[u_i] > 0$. If $\mathbb{E}[u_i] \leq 0$ (which occurs if n is sufficiently large), then firm j of course has no incentive to bundle.)

The threshold \tilde{n} in result (ii) is usually small. For instance, with a uniform distribution (15) becomes $n^2 - 4n + 2 > 0$ and so $\tilde{n} = 3$. This is the same as the threshold n^* in the consumer surplus comparison result in Proposition 2. This means that in this uniform example with a large number of products, separate sales is another equilibrium outcome if and only if consumers prefer separate sales to pure bundling. In other words, with a proper equilibrium selection the market itself can work well for consumers. The same is true in the normal distribution example. (But this is not generally true. We have examples, for instance, $f(x) = 2(1 - x)$, where $\tilde{n} \neq n^*$.)

Possibility of asymmetric equilibria. With more than two firms one may also wonder the possibility of asymmetric equilibria where some firms bundle and the others do not. An analytical investigation into this problem is hard because the pricing equilibrium when firms adopt asymmetric bundling strategies does not have a simple characterization.³⁵ However, numerical analysis can be done. 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 $p \approx 0.317$

³⁴This simplicity of optimal pricing with many products has been explored by Armstrong (1999) and Bakos and Brynjolfsson (1999). Fang and Norman (2006) have studied the profitability of pure bundling in the monopoly case with a finite number of products. They assume that the density of u_i is log-concave and symmetric. In our model, the log-concavity is guaranteed if f is log-concave, but the density of u_i is not symmetric when $n \geq 3$ (because y_i is stochastically greater than x_i). Without symmetry the Proschan (1965) result that $\sum_{i=1}^m u_i/m$ is more peaked than u_i does not hold any more. (The Proschan result has been extended in various ways, but not when the density is asymmetric.) The analysis in Fang and Norman (2006), however, crucially relies on that result. That is why their monopoly result cannot be applied to our competition model directly.

³⁵The reason is that when some firms bundle, other firms will treat their products as complements, and this will complicate the demand analysis. To see this, let us suppose firm 1 bundles while other firms do not. When firm $k \neq 1$ lowers its price for product i , some consumers will stop buying firm 1's bundle and switch to buying all products from the other firms. This will increase the demand for all of firm k 's products. (The details on the demand calculation and the first-order conditions are available, but no further analytical progress can be made.)

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. Then the situation is like all firms are bundling. Each bundling firm charges a bundle price $P = 0.5$, the third firm charges two single-product prices such that $p_1 + p_2 = 0.5$, 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 know from (14), the remaining bundling firm will have a demand less than $\frac{1}{3}$. This implies that the deviation firm will have a demand greater than $\frac{1}{3}$ and so earn a higher profit. (This argument actually does not depend on the uniform distribution and $m = 2$.)

4.5 Discussions

4.5.1 Asymmetric products and correlated valuations

We now consider a more general setting where a firm's m products are potentially asymmetric and their match utilities are potentially correlated. Let $\mathbf{x}^j = (x_1^j, \dots, x_m^j)$ be a consumer's valuations for the m products at firm j . Suppose that \mathbf{x}^j is still i.i.d. across firms and consumers, and it is distributed according to a joint CDF $F(x_1, \dots, x_m)$ with support $S \subset \mathbb{R}^m$ and a bounded joint density function $f(x_1, \dots, x_m)$. Let F_i and f_i , $i = 1, \dots, m$, be the marginal CDF and density function of x_i^j , and let $[\underline{x}_i, \bar{x}_i]$ be its support. Let G and g be the CDF and density function of X^j/m , where $X^j = \sum_{i=1}^m x_i^j$ as before, and let $[\underline{x}, \bar{x}]$ be its support. All f_i and g are log-concave if the joint density function f is log-concave. Then the equilibrium price in each regime is defined similarly as before:

$$\frac{1}{p_i} = n \int_{\underline{x}_i}^{\bar{x}_i} f_i(x) dF_i(x)^{n-1}, \quad i = 1, \dots, m; \quad \frac{1}{P/m} = \int_{\underline{x}}^{\bar{x}} g(x) dG(x)^{n-1}.$$

The limit result that $\lim_{m \rightarrow \infty} P/m = 0$ still holds as long as X^j/m converges to a deterministic value as $m \rightarrow \infty$. So for a fixed n , bundling lowers market price when m is sufficiently large. Under similar conditions as before, we also have the result that for a fixed m bundling raises market prices when n is sufficiently large. (We have not been able to extend the duopoly result in this general setting. See the online appendix for a discussion.)

Proposition 4 *Suppose f is log-concave. Suppose $S \subset \mathbb{R}^m$ is compact, strictly convex, and has full dimension. Then for a fixed m ,*

- (i) *if $f_i(\bar{x}_i) > 0$, there exists \hat{n}_i such that $P/m > p_i$ for $n > \hat{n}_i$;*
- (ii) *if $f_i(\bar{x}_i) > 0$ for all $i = 1, \dots, m$, there exists \hat{n} such that $P > \sum_{i=1}^m p_i$ for $n > \hat{n}$.*

Proof. Our conditions imply that $g(\bar{x}) = 0$ (e.g., see the proof of Proposition 1 in Armstrong, 1996).³⁶ Then the results immediately follow from result (ii) in Lemma 2.

■

The normal distribution example can also be extended to this general case. Suppose $\mathbf{x}^j \sim \mathcal{N}(\mathbf{0}, \Sigma)$, where σ_i^2 in Σ is the variance of x_i^j and σ_{ik} in Σ is the covariance of (x_i^j, x_k^j) . Then $X^j/m \sim \mathcal{N}(0, (\sum_{i=1}^m \sigma_i^2 + \sum_{i \neq k} \sigma_{ik})/m^2)$. According to formula (7), $P < \sum_{i=1}^m p_i$ if and only if $\sum_{i=1}^m \sigma_i^2 + \sum_{i \neq k} \sigma_{ik} < (\sum_{i=1}^m \sigma_i)^2$. Given $\sigma_{ik} \leq \sigma_i \sigma_k$ for any $i \neq k$, this condition must hold provided that at least one pair of (x_i^j, x_k^j) is not perfectly correlated.

4.5.2 Without full market coverage

We now return to the baseline model but relax the assumption of full market coverage. A subtle issue here is whether the m products are independent products or perfect complements (e.g., the essential components of a system). This will affect the analysis of the separate sales benchmark. If the m products are independent, consumers decide whether to buy each product separately. If the m products are perfect complements, then whether to buy a product also depends on how well-matched the 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 a consumer buys a product or bundle only if it is the best offer in the market and provides a positive surplus. To make the case interesting, let us suppose $\underline{x} \leq 0$ but the mean of x_i^j is positive, such that some consumers do not buy but the market is still active.

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

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

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

$$p = \frac{q(p)}{|q'(p)|} = \frac{[1 - F(p)^n]/n}{\underbrace{F(p)^{n-1} f(p)}_{\text{exclusion effect}} + \underbrace{\int_p^{\bar{x}} f(x) dF(x)^{n-1}}_{\text{competition effect}}}. \quad (16)$$

(If f is log-concave, this is also sufficient for defining the equilibrium price.) In equilibrium, a consumer will leave the market without purchasing product i with probability

³⁶Formally, $g(\bar{x}) = \lim_{\varepsilon \rightarrow 0} \frac{1 - G(\bar{x} - \varepsilon)}{\varepsilon}$, and our conditions about S ensure that $1 - G(\bar{x} - \varepsilon) = o(\varepsilon)$. Among the conditions, strict convexity of S excludes the possibility that the plane of $X^j/m = \bar{x}$ coincides with a part of S 's boundary, and the full dimension condition excludes the possibility that x_i^j , $i = 1, \dots, m$, are perfectly correlated.

$F(p)^n$ (i.e., when each product i has a valuation less than p). Given the symmetry of firms, the numerator in (16) is the equilibrium demand for each firm's product i . The denominator is the negative of the demand slope, and it 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 with probability $F(p)^{n-1}$), firm j acts as a monopoly. 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 p).

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

$$\frac{P}{m} = \frac{[1 - G(P/m)^n]/n}{G(P/m)^{n-1}g(P/m) + \int_{P/m}^{\bar{x}} g(x)dG(x)^{n-1}}, \quad (17)$$

where G and g are the CDF and density function of X^j/m as before.

Unlike the case with full market coverage, the equilibrium price in each regime is now implicitly determined in the corresponding first-order condition. The following result reports the condition for each first-order condition to have a unique solution. (See the online appendix for the proof.)

Lemma 4 *Suppose f is log-concave. There is a unique equilibrium price $p \in (0, p_M)$ defined in (16), where p_M is the monopoly price solving $p_M = [1 - F(p_M)]/f(p_M)$, and p decreases with n . Similar results hold for P/m defined in (17).*

For a fixed $n < \infty$, we still have $\lim_{m \rightarrow \infty} P/m = 0$ since X^j/m converges to the mean as $m \rightarrow \infty$. For a fixed $m < \infty$, if n is large, the demand size difference between the two numerators in (16) and (17) becomes negligible and so is the exclusion effect difference in the denominators. Therefore, price comparison is again determined by the comparison of $f(\bar{x})$ and $g(\bar{x})$. Intuitively, when there are many varieties in the market, almost every consumer can find something she likes and so almost no consumers will leave the market without purchasing anything. Then the situation will be close to the case with full market coverage. Consequently we have a similar result that when $f(\bar{x}) > 0$, bundling raises market prices when n is greater than a certain threshold. (We have not been able to extend the duopoly result in this setting without full market coverage.)

Figure 4 below reports the impacts of pure bundling on market prices, profits, consumer surplus, and total welfare in the uniform example with $F(x) = x$ and $m = 2$. (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 similar even if we introduce the exclusion effect of price.

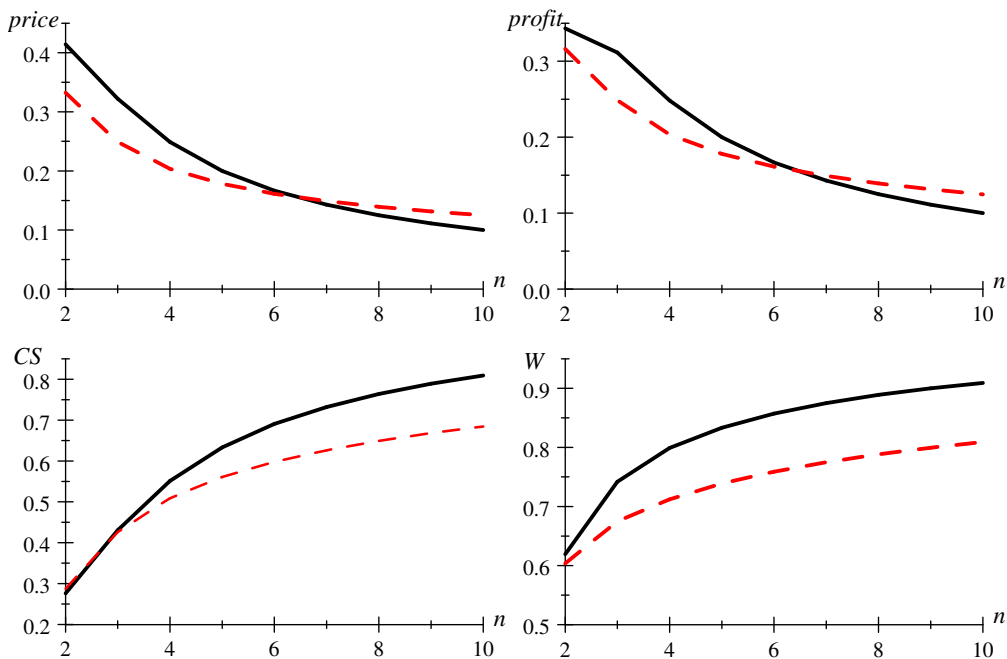


Figure 4: The impact of pure bundling—uniform distribution example without full market coverage

An alternative way to introduce the exclusion effect of price is to consider elastic demand. In the online appendix, we extend the baseline model by considering elastic consumer demand and show that the basic insights remain unchanged.

4.6 Related literature

Pure bundling or product incompatibility with product differentiation. Matutes and Regibeau (1988) initiated the study of competitive pure bundling in the context of product compatibility. They study the 2×2 case in a two-dimensional Hotelling model where consumers are uniformly distributed on a square. They show that bundling lowers market prices and profits, and it also benefits consumers if the market is fully covered. Our analysis in the duopoly case has generalized their results by considering more products and more general distributions.

Hurkens, Jeon, and Menicucci (2013) extend Matutes and Regibeau (1988) to the case with two asymmetric firms where one firm produces higher-quality products than the other. Consumers are distributed on the Hotelling square according to a symmetric distribution, and the quality premium is captured by a higher basic valuation for each product. Under certain technical assumptions Hurkens et al. show that when the quality difference is sufficiently large, pure bundling raises both firms' profits. (They do not state a formal result concerning price comparison.) Our 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 so the asymmetry between the firm and

its strongest competitor expands. This has a similar effect as increasing firm asymmetry in Hurkens, Jeon, and Menicucci (2013) and shifts the position of marginal consumers to the tail. 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 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) studies a spatial model of competitive pure bundling with an arbitrary number of 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 shows that for a regular transportation cost function, pure bundling always reduces market prices relative to separate sales. His spatial model features *local competition*: each firm is directly competing with its two neighbor rivals only (regardless of the separate sales regime or the bundling regime),³⁷ and they are always symmetric to each other no matter how many firms in total are present in the market. Conversely our random utility model features *global competition*: each firm is directly competing with all other firms. When there are more firms, each firm is effectively competing with a stronger competitor. It is this expanding asymmetry, which does not exist in Economides's spatial model, that drives our result that the impact of pure bundling can be reversed when the number of firms is above a threshold.

In a recent independent work, Kim and Choi (2015) propose an alternative $n \times 2$ spatial model. They assume that consumers are uniformly distributed and firms are symmetrically located on the surface of a torus. (Notice that firms can be symmetrically located in many possible ways in this model.) For a quadratic 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 prices and profits. This is consistent with our comparison result when $f(\bar{x}) > 0$. Compared to Economides (1989), a key difference in their model, using the insight learned from our paper, is that each firm can directly compete with more firms when the number of firms increases if we carefully select the location of firms. In this sense, their model is closer to our random utility model.

We have proposed a random utility approach to study competitive bundling. Our analysis has generated useful insights which can help us understand the discrepancy among the existing models and results. Whether a spatial model or a random utility model is more appropriate may depend on the context. But the random utility model

³⁷The consumers around the two polars should compare all firms, but they are ignored in the analysis.

is more flexible and easier to use. The analysis of pure bundling is much simpler in our framework than in the above spatial models. Our framework can also accommodate more than two products, and it can even be used to study competitive mixed bundling as we will see in the next section. Neither of them is easy to deal with in a spatial model with more than two firms. The random utility approach also accords well with econometric models of discrete consumer choice. In addition, neither Economides (1989) nor Kim and Choi (2015) investigate firms' individual incentives to bundle.

Pure bundling in auctions. Our study of competitive pure bundling is also related to the literature on multi-object auctions with bundling. Consider a private-value second-price auction where a bidder's valuation for each object is i.i.d. across objects and bidders. (The auction format does not matter given the revenue equivalence result.) Palfrey (1983) show that if there are only two bidders, selling all the objects in a package is more profitable than selling them separately, while the opposite is true when the number of bidders is sufficiently large. The revenue from selling the objects separately is equal to the sum of the second highest valuations for each product. While the revenue from selling them together in a bundle is equal to the second highest valuation for the bundle. With only two bidders the second highest valuation is the minimum one, so the revenue must be higher in the bundling case. With many bidders, however, the second highest valuation is close to the maximum one, so the revenue must be higher with separate sales. (In this limit case, bidders' information rent disappears, and so only the allocation efficiency matters for the revenue. Then bundling clearly reduces allocation efficiency.) In the two-object case, Chakraborty (1999) further show a cut-off result under certain regularity conditions.

Notice that the seller in an auction is like a consumer in our price competition model, and the agents on the other side of the market are competing for her. Hence, revenue comparison in auctions is related to consumer surplus comparison (instead of price and profit comparison) in our model. The analysis in an auction model does not directly apply to our price competition model. Competition occurs on the informed side in auctions but on the uninformed side in our price competition model. In the auction model, since we can focus on the second-price auction, the equilibrium bidding strategy is simple and all analysis can be conducted based on the order statistic of the second highest valuation. In our model as we have seen in (3) the equilibrium price is related to the second order statistic in a more complex way. For this reason, our analysis is less straightforward than in the auction case, and the main economic insight about how bundling affects price competition in our model does not exist in the bundling literature.³⁸

³⁸As we mentioned before, product information disclosure can affect consumer valuation distribution in a similar way as pure bundling does. The same analogy applies in the auction case. Among the works that study information disclosure in auctions, Board (2009) and Ganuza and Penalva (2010) are the two most related papers. Both of them show that whether or not the auctioneer benefits from

Pure bundling with homogenous products and heterogeneous costs. In the literature on competitive bundling, the standard approach that makes bundling a meaningful issue to study is to consider a market with horizontal production differentiation. We have followed that tradition, though we have adopted a random utility model instead of a spatial one. There is, however, an alternative modelling approach which considers homogeneous products with heterogeneous costs across firms. A simple setting is to assume that consumers have an identical (high) valuation for all products but the unit production cost is i.i.d. across firms and products and that each firm has private cost realizations. Competition now occurs on the informed side, and so the situation is actually like a first-price procurement auction. Since the outcome is the same as in a second-price auction, the above argument for auctions implies that consumers must benefit from bundling in duopoly but suffer when there are many firms.³⁹ Given total welfare is always lower with bundling, we can deduce that firms must suffer from bundling in the duopoly case. When the number of firms is sufficiently large, as in result (ii) of Proposition 1 we can show that the opposite is true if the cost density function is strictly positive at the lower bound.

Farrell, Monroe, and Saloner (1998) offers a model in this vein. They 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). The difference is that they assume a Bertrand price competition among firms with public cost information. Then the firm with the lowest cost wins the whole market and charges a price equal to the second lowest cost. This setting generates the same *ex ante* outcome as in the private cost setting. In the $n \times 2$ case they show a similar profit comparison result.

5 Mixed Bundling

We now turn to mixed bundling. Mixed bundling is a pricing strategy intended to screen consumers by offering more purchase options. In a competition environment, mixed bundling can induce more consumers to one-stop shop. Mixed bundling is harder to deal with than pure bundling because it leads to a more complicated pricing strategy space, especially when the number of products is large. For this reason, we will focus on the case where each firm supplies two products only. Then a mixed bundling strategy can be described by a pair of stand-alone prices (ρ_1, ρ_2) and a joint-purchase discount $\delta > 0$. If a consumer buys both products from the same firm, she pays $\rho_1 + \rho_2 - \delta$. For simplicity we will also return to the baseline model with full market coverage and i.i.d.

publicly providing more information to bidders depends on the number of bidders as in Palfrey (1983). Board (2009) points out the connection between bundling and information disclosure.

³⁹Chen and Li (2015) study bundling in a multiproduct procurement. They focus on the duopoly case, but they allow for correlated production costs in each firm and a not fully-covered market.

match utilities.

All the existing research on competitive mixed bundling focuses on the duopoly case. See, for example, Matutes and Regibeau (1992), Anderson and Leruth (1993), Thanassoulis (2007), and Armstrong and Vickers (2010). Armstrong and Vickers (2010) consider the most general setup so far in the literature. They allow for asymmetric products, correlated valuations, and the existence of an exogenous shopping cost. They also consider an elastic demand and a general nonlinear pricing schedule. Our paper is the first to consider more than two firms, but in other aspects we focus on the relatively simple setup. As we will see below, solving the mixed bundling pricing game becomes much harder when we go beyond the duopoly case. One contribution of this paper is to propose a way to solve the problem, and when the number of firms is large we also offer a simple approximation of the equilibrium prices.

In the following, we will first investigate firms' incentives to use mixed bundling. We will then characterize the symmetric pricing equilibrium with mixed bundling and examine the impact of mixed bundling relative to separate sales.

5.1 Incentive to use mixed bundling

In the pure bundling case, we have shown that whether a firm has a unilateral incentive to bundle its products depends on the number of firms in the market. However, when mixed bundling is possible, the following result shows that each firm always has an individual incentive to introduce mixed bundling, regardless of the number of firms. This implies that when mixed bundling is feasible, separate sales cannot be an equilibrium outcome.

Proposition 5 *Starting from separate sales with price defined in (1), each firm has a strict unilateral incentive to introduce mixed bundling.*

Proof. Suppose firm j unilaterally deviates from separate sales and introduces a small joint-purchase discount δ (but keeps its stand-alone price p unchanged). We aim to show that this deviation is profitable.

The negative (first-order) effect of this deviation 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 those consumers is $1/n^2$ and so the loss is δ/n^2 .

The positive effect is that now more consumers 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 and its CDF is $F(y_i)^{n-1}$. 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 .

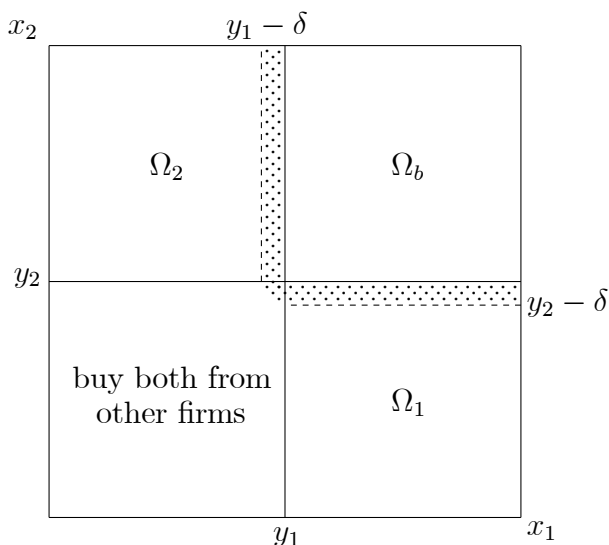


Figure 5: The impact of a small joint-purchase discount on demand

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 only the two rectangle areas matter. The measure of consumers on these two areas is $\delta[f(y_1)(1 - F(y_2)) + f(y_2)(1 - F(y_1))]$. From each of these consumers, firm j makes an extra profit $(2p - \delta) - p = p - \delta$. Therefore, conditional on (y_1, y_2) the positive (first-order) effect of introducing a small discount δ is $(p - \delta)\delta[f(y_1)(1 - F(y_2)) + f(y_2)(1 - F(y_1))]$. Integrating it over (y_1, y_2) and using the symmetry yields

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

(We have discarded the higher-order effect δ^2 . The first equality used the fact that $\int (1 - F(y_2))dF(y_2)^{n-1} = 1/n$, 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.⁴⁰

■

⁴⁰The 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.

5.2 Equilibrium prices

We aim to characterize a symmetric mixed-bundling equilibrium (ρ, δ) , where ρ is the stand-alone price for each individual product and $\delta \leq \rho$ is the joint-purchase discount.⁴¹ Suppose that all other firms use the equilibrium strategy and firm j unilaterally deviates and sets prices (ρ', δ') . Then a consumer has the following purchase options (for convenience we suppress the superscripts in firm j 's match utilities):

- (a) buy both products from firm j , in which case her surplus is $x_1 + x_2 - (2\rho' - \delta')$;
- (b) buy product 1 from firm j but product 2 from elsewhere, in which case her surplus is $x_1 + y_2 - \rho' - \rho$;
- (c) buy product 2 from firm j but product 1 from elsewhere, in which case her surplus is $y_1 + x_2 - \rho' - \rho$;
- (d) buy both products from some 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 some other firm, she will buy the best one with match utility y_i . When she buys both products from some other firms (and $n \geq 3$), there are two cases to consider: If y_1 and y_2 are from the same firm, the decision is simple and the consumer will buy both products from that firm, in which case $A = y_1 + y_2$. This occurs with probability $\frac{1}{n-1}$. With the remaining probability $\frac{n-2}{n-1}$, y_1 and y_2 are from 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

$$z \equiv \max_{k \neq j} \{x_1^k + x_2^k\} < y_1 + y_2 \quad (18)$$

is the match utility of the best bundle among all other firms. Hence, when y_1 and y_2 are from different firms, $A = \max\{z, y_1 + y_2 - \delta\}$. In sum, we have

$$A = \begin{cases} y_1 + y_2 & \text{with prob. } \frac{1}{n-1} \\ \max\{z, y_1 + y_2 - \delta\} & \text{with prob. } \frac{n-2}{n-1} \end{cases} \quad (19)$$

The relatively simple case is when $n = 2$. Then $A = y_1 + y_2$, and the surplus from the fourth option above is simply $y_1 + y_2 - (2\rho - \delta)$. The problem can then be converted 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 often used in the existing literature on competitive mixed bundling.

When $n \geq 3$, the situation is more involved. We need to deal with one more random variable z defined in (18), and moreover z is correlated with y_1 and y_2 as reported in

⁴¹If $\delta > \rho$, the bundle would be cheaper than each individual product, and so only the bundle price would matter for consumer choices. The situation would then be like pure bundling where consumers can buy multiple bundles. Of course, if the bundle price is high enough (e.g., due to high production costs), consumers will not have an incentive to buy multiple bundles.

the following lemma. Characterizing this correlation is an important step in solving the mixed bundling pricing game with more than two firms. Given the following result, the distribution of A conditional on y_1 and y_2 is fully characterized.

Lemma 5 *When $n \geq 3$, the CDF of z defined in (18), conditional on y_1, y_2 , and they being from different firms, is*

$$L(z) = \frac{F(z - y_1)F(z - y_2)}{(F(y_1)F(y_2))^{n-2}} \left(F(y_2)F(z - y_2) + \int_{z-y_2}^{y_1} F(z - x)dF(x) \right)^{n-3} \quad (20)$$

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

Proof. For a given consumer, let $I(y_i)$, $i = 1, 2$, be the identity of the firm that generates y_i . The lower bound of z is because the lowest possible match utility of the bundle from firm $I(y_i)$ is $y_i + \underline{x}$. We now calculate the conditional probability of $\max_{k \neq j} \{x_1^k + x_2^k\} < t$. This event occurs if and only if all the following three conditions are satisfied: (i) $y_1 + x_2^{I(y_1)} < t$, (ii) $x_1^{I(y_2)} + y_2 < t$, and (iii) $x_1^k + x_2^k < t$ for all $k \neq j, I(y_1), I(y_2)$. Given y_1 and y_2 , condition (i) holds with probability $F(t - y_1)/F(y_2)$, since 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 $F(t - y_2)/F(y_1)$. One can also check (with the help of a graph) that the probability that $x_1^k + x_2^k < t$ holds for one firm other than $j, I(y_1)$ and $I(y_2)$, is

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

Conditional on y_1 and y_2 , these three events are independent of each other. Therefore, the conditional probability of $\max_{k \neq j} \{x_1^k + x_2^k\} < t$ is the right-hand side of (20). ■

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

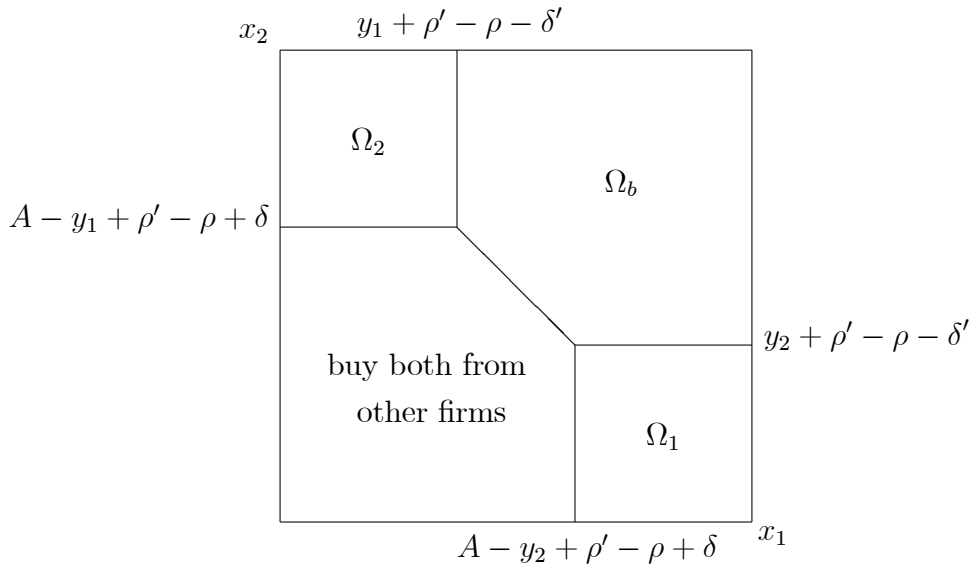


Figure 6: The pattern of consumer choice

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.

From Figure 6, we can see that the equilibrium demand for firm j 's single product i is

$$\Omega_i(\delta) \equiv \mathbb{E} [F(y_j - \delta)(1 - F(A - y_j + \delta))], \quad i = 1, 2, \quad j \neq i. \quad (21)$$

(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

$$\Omega_i(\delta) + \Omega_b(\delta) = \frac{1}{n}. \quad (22)$$

With full market coverage, all consumers 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 $\Omega_1(\delta) = \Omega_2(\delta)$.

To characterize the equilibrium prices, it is useful to introduce a few more pieces of notation:

$$\begin{aligned} \alpha(\delta) &\equiv \mathbb{E} [f(y_1 - \delta)(1 - F(A - y_1 + \delta))], \\ \beta(\delta) &\equiv \mathbb{E} [f(A - y_1 + \delta)F(y_1 - \delta)], \\ \gamma(\delta) &\equiv \mathbb{E} \left[\int_{y_1 - \delta}^{A - y_2 + \delta} f(A - x)f(x)dx \right]. \end{aligned} \quad (23)$$

(All the expectations are taken over (y_1, y_2, A) .) Then the necessary conditions for ρ and δ to be the equilibrium prices are given in the following result. (See the online appendix for the proof.)

Proposition 6 *If a symmetric (pure strategy) mixed-bundling equilibrium exists, the stand-alone price ρ and the joint-purchase discount δ must satisfy*

$$\rho = \frac{1/n + \delta(\alpha(\delta) + \gamma(\delta))}{\alpha(\delta) + \beta(\delta) + 2\gamma(\delta)}, \quad (24)$$

and

$$\rho(\beta(\delta) - \alpha(\delta)) = \Omega_1(\delta) - \delta\alpha(\delta). \quad (25)$$

Notice that substituting (24) into (25) yields an equation of δ . Once δ is solved, ρ can be calculated from (24).

Discussion: equilibrium existence. To prove the existence of a symmetric (pure-strategy) equilibrium,⁴² we need to show that (i) the system of necessary conditions

⁴²In our model, each firm's pricing strategy space is a subset of \mathbb{R}^3 and we can make it compact without loss of generality. The profit function is continuous in prices as long as the density functions are. Then Theorem 1 in Becker and Damianov (2006) implies that our model has a symmetric equilibrium (but not necessarily in pure strategy).

(24) and (25) has a solution with $\delta < \rho$, and (ii) the necessary conditions are also sufficient for defining the equilibrium prices. Unfortunately, both issues are hard to investigate in general. For the first one, we can show it when $n = 2$ or when n is sufficiently large under the log-concavity condition. For the second one, no analytical progress has been made. This issue is hard even in the duopoly case and is generally an unsolved problem in the literature on mixed bundling. Numerically we can show the existence in the uniform distribution and the normal distribution example.

In general, $\alpha(\delta)$, $\beta(\delta)$, $\gamma(\delta)$, and $\Omega_1(\delta)$ in (24) and (25) have complicated expressions. However, they are simple in the duopoly case, and they also have simple approximations when δ is small (which can be shown to be true under mild conditions when the number of firms is large). Hence, in the following we study these two cases.

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

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

In particular, when $n = 2$, $h(t)$ 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 in the regime of separate sales. So the price p in (1) 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 check that $\alpha(\delta) = \beta(\delta) = h(\delta)[1 - H(\delta)]$ and $\Omega_1(\delta) = [1 - H(\delta)]^2$. Thus, (25) simplifies to

$$\delta = \frac{1 - H(\delta)}{h(\delta)} . \tag{26}$$

(Alternatively, it can be written as $\Omega_1(\delta) + \frac{1}{2}\delta\Omega_1'(\delta) = 0$ as in Armstrong and Vickers, 2010.) If $1 - H$ is log-concave (which is implied by the log-concavity of f), this equation has a unique positive solution. Meanwhile, (24) becomes

$$\rho = \frac{\delta}{2} + \frac{1}{4(\alpha(\delta) + \gamma(\delta))} \tag{27}$$

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

When n is large, under mild conditions we can show that the system of (24) and (25) has a solution with δ close to zero. For a small δ both sides of (25) have simple approximations, and an approximated δ can be solved. The following proposition provides the details. (See the online appendix for the proof.)

Proposition 7 *Suppose f is log-concave.*

(i) *When $n = 2$, the system of (26) and (27) has a solution with $\delta < \rho$, and the bundle price is lower than in the regime of separate sales (i.e., $2\rho - \delta < 2p$).*

(ii) *Suppose $\frac{|f'(x)|}{f(x)}$ is bounded and $\lim_{n \rightarrow \infty} p = 0$, where $p = \frac{1}{nh(0)}$ is the separate-sales price in (1). When n is large, the system of (24) and (25) has a solution with $\delta \in (0, p)$ and it can be approximated as*

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

Both the stand-alone price and the bundle price are lower than in the regime of separate sales.

Result (ii) 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. The mixed bundling price scheme can then be interpreted as “50% off for the second product”.⁴³ Result (ii) also implies that when there are many firms in the market, mixed bundling tends to be pro-competitive relative to separate sales.

5.3 Impact 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 industry profit and consumer surplus.

Let $\pi(\rho, \delta)$ be the equilibrium industry profit. 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 industry profit is

$$\pi(\rho, \delta) - \pi(p, 0) = 2(\rho - p) - n\delta\Omega_b(\delta). \quad (28)$$

Let $v(\tilde{p}, \tilde{\delta})$ be the consumer surplus when the stand-alone price is \tilde{p} and the joint-purchase discount is $\tilde{\delta}$. Given full market coverage, an envelope argument implies that $v_1(\tilde{p}, \tilde{\delta}) = -2$ and $v_2(\tilde{p}, \tilde{\delta}) = n\Omega_b(\tilde{\delta})$, where the subscripts indicate partial derivatives. (This is because raising \tilde{p} by ε will make every consumer pay 2ε more, and raising the discount $\tilde{\delta}$ by ε will save ε for every consumer who buy both products from the

⁴³When there is a positive production cost c for each product, we have $\delta \approx (p - c)/2$, i.e., the bundling discount is approximately equal to half of the single product’s markup.

same firm.) Then relative to separate sales, the impact of mixed bundling on consumer surplus is

$$\begin{aligned}
v(\rho, \delta) - v(p, 0) &= v(\rho, \delta) - v(p, \delta) + v(p, \delta) - v(p, 0) \\
&= \int_p^\rho v_1(\tilde{\rho}, \delta) d\tilde{\rho} + \int_0^\delta v_2(p, \tilde{\delta}) d\tilde{\delta} \\
&= -2(\rho - p) + n \int_0^\delta \Omega_b(\tilde{\delta}) d\tilde{\delta} .
\end{aligned} \tag{29}$$

Hence, (28) and (29) provide the formulas for calculating the impacts of mixed bundling on industry profit and consumer surplus. (From these two formulas, it is also clear that mixed bundling harms total welfare given that $\Omega_b(\tilde{\delta})$ is increasing in $\tilde{\delta}$.)

In the duopoly case, 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.

6 Conclusion

This paper has offered a framework to study competitive bundling with an arbitrary number of firms. In the 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 welfare assessment of pure bundling based on a duopoly model can be misleading. In the mixed bundling part, we found that solving the pricing 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 the pure bundling part is that consumers do not buy more than one bundle. This is without loss of generality if production cost is high or if bundling is caused by product incompatibility or high shopping costs. However if production cost is relatively small and bundling is a pricing strategy, then it is possible that the bundle price is low enough such that some consumers want to buy multiple bundles to mix and match. 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. When the bundle price is P , a consumer faces two options: buy the best single bundle and pay 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 method of solving the mixed bundling game can be used to deal with this case.

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.

Finally, there are some recent empirical works on bundling. 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. These works focus on how the interaction between bundling and the vertical market structure might affect market performance. This is an interesting aspect ignored by the existing theoretical literature on bundling and deserves more exploration.

Appendix

Proof of Proposition 1: (i) The duopoly model can be converted into a two-dimensional Hotelling model. Define $d_i \equiv x_i^1 - x_i^2$, and let H and h be its CDF and density function, respectively. Since x_i^1 and x_i^2 are i.i.d., d_i has a support $[\underline{x} - \bar{x}, \bar{x} - \underline{x}]$ and is symmetric around zero. One can check that $h(0) = \int f(x)^2 dx$. Then in the duopoly case the equilibrium price in (1) can be written as

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

Notice that h is log-concave given f is log-concave.

Let \bar{H} and \bar{h} be the CDF and density function of $\sum_{i=1}^m d_i/m$. One can check that $\bar{h}(0) = \int g(x)^2 dx$. Then the per-product bundle price can be written as

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

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

(ii) Given $f(\bar{x}) > 0$, we have $f(\bar{x}) > g(\bar{x}) = 0$, and so the result for $n > \hat{n}$ follows immediately from result (ii) in Lemma 2. To prove the limit result, we use the technique in the proof of Lemma 2 and notice that

$$\lim_{n \rightarrow \infty} \frac{P/m}{p} = \lim_{n \rightarrow \infty} \frac{\int_0^1 l_F(t) dt^{n-1}}{\int_0^1 l_G(t) dt^{n-1}} .$$

Given $l_F(t)$ is bounded and $l_F(1) > 0$, the numerator converges to $l_F(1) > 0$. While the denominator converges to zero since $l_G(t)$ is bounded and $l_G(1) = 0$. Then $\lim_{n \rightarrow \infty} \frac{P/m}{p} = \infty$.

To prove the cut-off result, define

$$\lambda(n) \equiv \int_0^1 [l_F(t) - l_G(t)] t^{n-2} dt . \quad (30)$$

Under our conditions, we already know that $\lambda(2) < 0$ and $\lambda(n) > 0$ for a 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 use one version of the *Variation Diminishing Theorem* (see Theorem 3.1 in Karlin, 1968).⁴⁴ 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 $x_1 < \dots < x_k$ and $y_1 < \dots < y_k$ 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 a way to count the number of sign changes of a function. Consider a function $f(t)$ for $t \in T$ where T is an ordered set of the real line. Let

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

where the supremum is extended over all sets $t_1 \leq \dots \leq t_k$ ($t_i \in T$), k is arbitrary but finite, and $S(x_1, \dots, x_k)$ 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

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

⁴⁴Chakraborty (1999) uses this same theorem in proving a cut-off result on how bundling affects a seller's revenue in multiproduct auctions. See section 4.6 for a detailed discussion about the difference between auctions and our price competition model. The main technical difference here, compared to the proof in Chakraborty (1999), is our Lemma 6 below.

Now consider $\lambda(n)$ defined in (30). 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 a sufficiently large n , it is impossible that $\lambda(n)$ changes its sign exactly twice. Therefore, it must change its sign only once and so there exists \hat{n} such that $\lambda(n) < 0$ 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

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

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 that 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 $\frac{\partial^2}{\partial x \partial y} \log(x^y - 1)$ has the same sign as $x^y - 1 - \log x^y > 0$. (The inequality is 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 is distributed (approximately) according to the normal distribution $\mathcal{N}(\mu, \sigma^2/m)$. Then (8) implies that

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

where $p_{\mathcal{N}}$ is the separate sales price when x_i^j follows a normal distribution $\mathcal{N}(\mu, \sigma^2)$. Then it is clear that P/m decreases in m and converges to zero as $m \rightarrow \infty$.

Proof of Proposition 2: (i): The first condition simply follows from result (ii) in Proposition 1. To prove the second condition, we use two results in Gabaix et al. (2015). From their Theorem 1 and Proposition 2, we know that when $\lim_{x \rightarrow \bar{x}} \frac{d}{dx} \left(\frac{1-F(x)}{f(x)} \right) = 0$, the equilibrium price (1) in the Perloff-Salop model has the same approximation as $\mathbb{E}[x_{(n)} - x_{(n-1)}]$ when n is large, where $x_{(n)}$ and $x_{(n-1)}$ are the first and the second order statistic of a sequence of i.i.d. random variables $\{x_1, \dots, x_n\}$. In other words,

$$p \approx \mathbb{E} [x_{(n)} - x_{(n-1)}]$$

as $n \rightarrow \infty$.⁴⁵ Then the per-product consumer surplus in the regime of separate sales is

$$\mathbb{E} [x_{(n)}] - p \approx \mathbb{E} [x_{(n-1)}] .$$

Since $\lim_{x \rightarrow \bar{x}} \frac{d}{dx} \left(\frac{1-F(x)}{f(x)} \right) = 0$ implies $\lim_{x \rightarrow \bar{x}} \frac{d}{dx} \left(\frac{1-G(x)}{g(x)} \right) = 0$, we can deduce that when n is large the per-product consumer surplus in the regime of pure bundling is

$$\mathbb{E} \left[\frac{X_{(n)}}{m} \right] - \frac{P}{m} \approx \mathbb{E} \left[\frac{X_{(n-1)}}{m} \right] .$$

When n is large, it is clear that $\mathbb{E} \left[\frac{X_{(n-1)}}{m} \right] < \mathbb{E} [x_{(n-1)}]$ and so our result follows.

(ii): It suffices to show the threshold n^* exists when $m \rightarrow \infty$. As we have explained in the main text, when $m \rightarrow \infty$ bundling benefits consumers if and only if

$$\mathbb{E} \left[\max_j \{x_i^j\} \right] - \mu < p . \quad (31)$$

Since the left-hand side of (31) increases with n while the right-hand side decreases with n , we only need to prove two things: (a) condition (31) holds for $n = 2$, and (b) the opposite is true for a sufficiently large n .

Condition (b) is relatively easy to show. The left-hand side of (31) approaches $\bar{x} - \mu = \int_{\underline{x}}^{\bar{x}} F(x) dx$ as $n \rightarrow \infty$. So we need

$$\lim_{n \rightarrow \infty} p = \frac{1 - F(\bar{x})}{f(\bar{x})} < \int_{\underline{x}}^{\bar{x}} F(x) dx , \quad (32)$$

where the equality is from (3). If $\lim_{n \rightarrow \infty} p = 0$ (e.g., if $f(\bar{x}) > 0$), this is clearly the case. But we want to show this is true even if $\lim_{n \rightarrow \infty} p > 0$. Notice that log-concave f implies log-concave $1 - F$ (or decreasing $(1 - F)/f$). Then

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

⁴⁵This result implies that when the tail behavior condition holds, the Perloff-Salop price competition model is asymptotically equivalent to an auction model where firms bid for a consumer whose valuations for each product are publicly known. However, when the tail behavior condition does not hold or when n is not large enough, this equivalence result does not hold.

Since the integral term is infinity, condition (32) must hold.

We now prove condition (a). Using (1) and the notation $l(t) \equiv f(F^{-1}(t))$, we can rewrite (31) 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

$$\int_0^1 \frac{t(1-t)}{l(t)} dt \int_0^1 l(t) dt < \frac{1}{2} . \quad (33)$$

To prove this inequality, we need the following technical result:⁴⁶

Lemma 7 *Suppose $\varphi : [0, 1] \rightarrow \mathbb{R}$ is a nonnegative function such that $\int_0^1 \frac{\varphi(t)}{t(1-t)} dt < \infty$, and $r : [0, 1] \rightarrow \mathbb{R}$ is a concave density function. 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 density function, 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 density function defined on $[0, 1]$, $r_1(t) = 2t$, $r_0(t) = 2(1-t)$, and for $0 < \theta < 1$

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

(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

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

⁴⁶I am grateful to Tomás F. Móri in Budapest for helping me to prove this lemma.

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 log-concave. So the defined $r(t)$ is indeed a concave density function. (The integral in the denominator is finite since $l(t)$ is nonnegative and concave.) Then Lemma 7 implies that the left-hand side of (33) is no greater than $1/4$.

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 firm j 's product has match utility x^j . Then*

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

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

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

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 3: (i) The result that bundling is always a NE has been explained in the main text. We now prove the uniqueness in the duopoly case. We first show that it is not an equilibrium that both firms sell their products separately. Consider a hypothetical separate sales equilibrium with price p . Suppose now firm j unilaterally bundles. Then the situation is like both firms are bundling, and firm j can at least earn the same profit as before by setting a bundle price mp . It can actually do strictly better by adjusting its prices as well. Suppose firm j sets a bundle price $mp - m\varepsilon$, where $\varepsilon > 0$ is small. The negative (first-order) effect of this deviation on firm j 's profits is $\frac{m}{2}\varepsilon$. This is because half of the consumers buy from firm j when $\varepsilon = 0$,

and now they pay $m\varepsilon$ less. On the other hand, the deviation increases the demand for firm j 's bundle to

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

So the demand increases by $\varepsilon \int g(x)^2 dx$, and the positive (first-order) 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 used (5) at $n = 2$.) Therefore, the deviation is profitable if $P < mp$. Similarly, if $P > mp$, then charging a bundle price $mp + m\varepsilon$ will be a profitable deviation.

Second, we show that there are no asymmetric equilibria either. Consider a hypothetical equilibrium where firm j bundles and firm k does not. For consumers, this is like both firms are bundling. As a result, in equilibrium firm j offers a bundle price P defined in (5) 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. It can at least earn the same profit as before by charging the same prices as firm k . But it can do strictly better by offering prices $\{p_i - \varepsilon\}_{i=1}^m$, where $\varepsilon > 0$ is small. The negative (first-order) effect of this deviation on firm j 's profit is $\frac{m}{2}\varepsilon$. On the other hand, the demand for firm j 's each product increases by $\varepsilon \int f(x)^2 dx$, and so the positive (first-order) 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) at $n = 2$ and $\sum_{i=1}^m p_i = P$.) Therefore, the deviation is profitable if $P > mp$. Similarly, if $P < mp$, setting prices $\{p_i + \varepsilon\}_{i=1}^m$ will be a profitable deviation.

(ii) It suffices to show the threshold \tilde{n} exists when $m \rightarrow \infty$. As we have explained in the main text, when $m \rightarrow \infty$ a firm has no unilateral incentive to bundle if and only if

$$\left(1 - \frac{1}{n}\right)p < \int [F(x) - F(x)^{n-1}] dx . \quad (34)$$

We have known that (34) fails to hold for $n = 2$. On the other hand, we have $\lim_{n \rightarrow \infty} p < \int_{\underline{x}}^{\bar{x}} F(x) dx$ as shown in the proof of Proposition 2. Then (34) must hold when n is sufficiently large. In the following, we further show a cut-off result. Using the notation $l(t) \equiv f(F^{-1}(t))$, we rewrite (34) as

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

It suffices to show that $\Delta(n)$ decreases in n . This is not obvious given $1 - \frac{1}{n}$ is increasing in n .

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

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

On one hand, from Lemma 1 we know that $p_{n+1} < p_n$ when f is log-concave. 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 $\frac{1}{p_n} = n(n-1) \int_0^1 l(t)t^{n-2} dt$.) On the other hand, we have

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

where $\kappa(t) \equiv n(n+1)t^{n-1}(1-t)$ is a density function on $[0, 1]$, and the inequality is from Jensen's Inequality.

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

$$\begin{aligned} n(n+1) \int_0^1 l(t)\kappa(t) dt &< n^2(n^2-1) \int_0^1 l(t)t^{n-2} dt \\ \Leftrightarrow (n+1)^2 \int_0^1 l(t)t^{n-1}(1-t) dt &< (n^2-1) \int_0^1 l(t)t^{n-2} dt. \end{aligned}$$

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

References

- ADAMS, W., AND J. YELLEN (1976): "Commodity Bundling and the Burden of Monopoly," *Quarterly Journal of Economics*, 90(3), 475–498.
- ANDERSON, S., A. DE PALMA, AND Y. NESTEROV (1995): "Oligopolistic Competition and the Optimal Provision of Products," *Econometrica*, 63(6), 1281–1301.
- ANDERSON, S., AND L. LERUTH (1993): "Why Firms May Prefer Not to Price Discriminate via Mixed Bundling," *International Journal of Industrial Organization*, 11(1), 49–61.
- ARMSTRONG, M. (1996): "Multiproduct Nonlinear Pricing," *Econometrica*, 64(1), 51–75.
- (1999): "Price Discrimination by a Many-Product Firm," *Review of Economic Studies*, 66(1), 151–168.
- (2015): "Nonlinear Pricing," mimeo, Oxford.
- ARMSTRONG, M., AND J. VICKERS (2010): "Competitive Non-linear Pricing and Bundling," *Review of Economic Studies*, 77(1), 30–60.

- BAKOS, Y., AND E. BRYNJOLFSSON (1999): “Bundling Information Goods: Pricing, Profits, and Efficiency,” *Management Science*, 45(12), 1613–1630.
- BECKER, J., AND D. DAMIANOV (2006): “On the Existence of Symmetric Mixed Strategy Equilibria,” *Economics Letters*, 90(1), 84–87.
- BOARD, S. (2009): “Revealing Information in Auctions: the Allocation Effect,” *Economic Theory*, 38(1), 125–135.
- 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.
- CHEN, Y. (1997): “Equilibrium Product Bundling,” *Journal of Business*, 70(1), 85–103.
- CHEN, Y., AND J. LI (2015): “Bundled Procurement,” mimeo, Colorado.
- 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.
- GABAIX, X., D. LAIBSON, D. LI, H. LI, S. RESNICK, AND C. DE VRIES (2015): “The Impact of Competition on Prices With Numerous Firms,” mimeo, NYU.
- GANUZA, J., AND J. PENALVA (2010): “Signal Orderings Based on Dispersion and the Supply of Private Information in Auctions,” *Econometrica*, 78(3), 1007–1030.
- GENTZKOW, M. (2007): “Valuing New Goods in a Model with Complementarities: Online Newspapers,” *American Economic Review*, 97(3), 713–744.
- 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 (2015): “Optimal Compatibility in Systems Markets,” *Games and Economic Behavior*, 90, 106–118.
- 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.

- NOLAN, J. P. (2015): *Stable Distributions - Models for Heavy Tailed Data*. Birkhauser, Boston, In progress, Chapter 1 online at academic2.american.edu/~jpnolan.
- 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.
- QUINT, D. (2014): “Imperfect Competition with Complements and Substitutes,” *Journal of Economic Theory*, 152, 266–290.
- SCHMALENSEE, R. (1984): “Gaussian Demand and Commodity Bundling,” *Journal of Business*, 57(1), S211–230.
- SHAKED, M., AND G. SHANTHIKUMAR (2007): *Stochastic Orders*. Springer.
- 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.
- WEYL, G., AND M. FABINGER (2013): “Pass-Through as an Economic Tool: Principles of Incidence under Imperfect Competition,” *Journal of Political Economy*, 121(3), 528–583.
- WHINSTON, M. (1990): “Tying, Foreclosure, and Exclusion,” *American Economic Review*, 80(4), 837–859.

Online Appendix

[Not For Publication]

This online appendix contains a few proofs and the details of some discussions omitted in the paper.

The duopoly case with asymmetric products and correlated valuations. Consider the duopoly case with a general joint distribution as introduced in Section 4.5.1. Following the logic in the proof of result (i) in Proposition 1, let h_i be the density function of $x_i^1 - x_i^2$. (Notice that $x_i^1 - x_i^2$ is symmetric around zero, and h_i is log-concave if f_i is log-concave.) Then the separate sales price for product i is $p_i = \frac{1}{2h_i(0)}$. Let \bar{h} be the density function of $\frac{1}{m} \sum_{i=1}^m (x_i^1 - x_i^2)$. Then the per-product bundle price is $P/m = \frac{1}{2\bar{h}(0)}$, and so $P < \sum_{i=1}^m p_i$ if and only if

$$\frac{1}{\bar{h}(0)} < \frac{1}{m} \sum_{i=1}^m \frac{1}{h_i(0)}. \quad (35)$$

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

$$\frac{1}{m} \sum_{i=1}^m h_i(0) \leq \bar{h}(0). \quad (36)$$

It appears hard to find simple primitive conditions on the joint density function f such that (35) or (36) is satisfied.¹

The following example shows that in the duopoly case bundling can lower market price even if products are substantially asymmetric. Consider a two-product example where x_1^j is uniformly distributed on $[0, 1]$ and x_2^j is independent of x_1^j and is uniformly distributed on $[0, b]$ with $b > 1$. In duopoly, $x_1^1 - x_1^2$ and $x_2^1 - x_2^2$ have density functions

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

respectively. Then $h_1(0) = 1$ and $h_2(0) = \frac{1}{b}$. One can also check $\bar{h}(0) = \frac{2(3b-1)}{3b^2}$. The sufficient condition (36) holds only for b less than about 2.46, but the "iff" condition (35) holds for any $b > 1$. Hence, bundling lowers prices in this duopoly example. However, according to Proposition 4 bundling raises prices in this example if n is sufficiently large.

¹If the m products are symmetric and the joint density function of $\{x_i^1 - x_i^2\}_{i=1}^m$ is Schur-concave, or if the m products at each firm have independent match utilities and any two random variables $x_i^1 - x_i^2$ and $x_k^1 - x_k^2$ can be ranked according to the likelihood ratio order, then there are extensions of the Proshan (1956) result which can help prove (36). But we do not have simple primitive conditions for either of the conditions to hold. (In the symmetric and independent case, both conditions are satisfied if f is log-concave.)

Proof of Lemma 4: We only prove the results for p . (The same logic works for P .) Notice that $1 - F$ is log-concave given f is log-concave. When $p = 0$, it is clear that the left-hand side of (16) 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 (16) equals

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

(The inequality is because $f/(1 - F)$ is increasing, and the equality used $F_{(2)}(p) = F(p)^n + nF(p)^{n-1}(1 - F(p))$.) This, together with the fact that $p_M = \frac{1-F(p_M)}{f(p_M)}$, implies that (16) has a solution $p \in (0, p_M)$.

To show the uniqueness, we prove that the right-hand side of (16) decreases with p . One can verify that its derivative with respect to p is negative if and only if

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

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

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

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

To prove the result that p decreases in n , let us first rewrite (16) as

$$\frac{1}{p} = \frac{f(\bar{x}) - \int_p^{\bar{x}} f'(x)F(x)^{n-1}dx}{(1 - F(p)^n)/n} = \frac{nf(\bar{x})}{1 - F(p)^n} - \int_p^{\bar{x}} \frac{f'(x)}{f(x)} d \frac{F(x)^n - F(p)^n}{1 - F(p)^n}. \quad (37)$$

(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 log-concavity of f implies that $-\frac{f'}{f}$ 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 (37) increases with n . Therefore, the unique solution p must decrease with n .

Elastic consumer demand. We extend the baseline model by considering elastic consumer demand. Suppose each product is divisible and consumers can buy any quantity of a product, and the m products are independent products. If a consumer purchases τ_i units of product i from firm j , suppose that 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 random utility component as before and reflects product differentiation. We also assume

that firms use a linear pricing scheme for each product. Then if a consumer chooses to buy product i from firm j , she must buy all the units from it. 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 . Then $v(p_i)$ is a decreasing function, and $-v'(p_i)$ is the usual demand function.

Consider the regime of separate sales first. Let p be the (symmetric) equilibrium unit price. Suppose firm j unilaterally 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 turns out to be 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 $p = v^{-1}(s)$ 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 unilaterally deviates and offers s' , the number 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

$$-\frac{r'(s)}{r(s)} = n \int f(x) dF(x)^{n-1} . \quad (38)$$

If both $r(s)$ and $f(x)$ are log-concave, this is also sufficient for defining the equilibrium indirect utility. This equation has a unique solution if $r(s)$ is log-concave (so $-\frac{r'(s)}{r(s)}$ is increasing in s).²

Now consider the regime of pure bundling. If a firm adopts a pure bundling strategy, it requires a consumer to buy all products from it or nothing at all. But unlike in the baseline model, now a firm's pricing strategy can no longer be represented by a simple bundle price, and it has to specify a vector of prices (p_1, \dots, p_m) . (In the baseline model with unit demand, offering a bundle price is equivalent to offering a vector of single-product prices in the bundling regime.) If a consumer buys all products from this firm, her indirect utility is $\sum_{i=1}^m v(p_i)$. Suppose a firm offers an indirect utility S . Then the optimal price vector should solve the problem $\max_{\{p_i\}} \sum_{i=1}^m p_i (-v'(p_i))$ subject to $\sum_{i=1}^m v(p_i) = S$. Let us suppose this problem has a unique solution with

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

$p_i = p_k$. The optimal unit price for each product is then $v^{-1}(\frac{S}{m})$. Once we work on the indirect utility, it is similar as before that in the regime of separate sales a firm offers an indirect utility for each product, and in the regime of pure bundling it offers a bundle indirect utility.

We look for a symmetric equilibrium where each firm offers an indirect utility S . Suppose firm j unilaterally deviates to S' . Then the number of consumers who buy all products from firm j 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 $mr(\frac{S'}{m})Q(S')$, and then the first-order condition is

$$-\frac{r'(\frac{S}{m})}{r(\frac{S}{m})} = n \int g(x) dG(x)^{n-1} . \quad (39)$$

Therefore, as long as $-\frac{r'(s)}{r(s)}$ is increasing in s (or $r(s)$ is log-concave), (38) and (39) are similar as the equilibrium price conditions (1) and (5) in the baseline model with unit demand. Then our price comparison results in Proposition 1 continues to hold.

Proof of Proposition 6: To derive the equilibrium conditions for ρ and δ , let us consider the following two local 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 consumer choices: Ω_b expands because now more consumers buy both products from firm j . The marginal consumers are distributed on the shaded area.

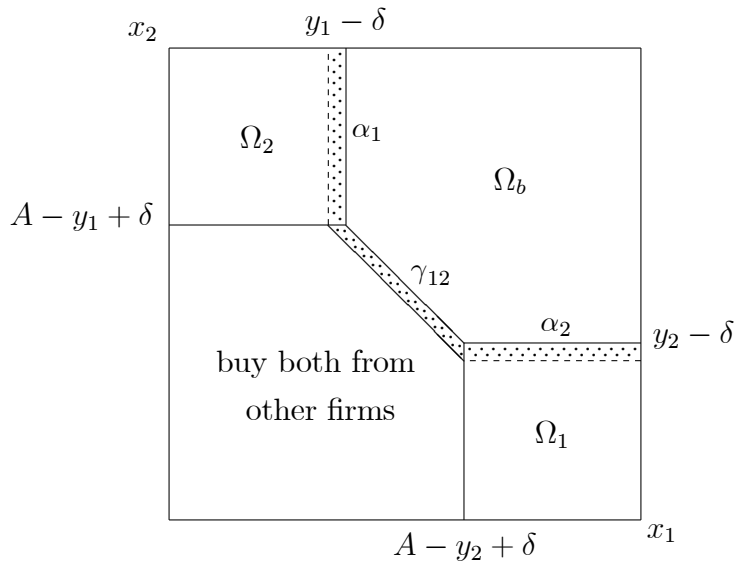


Figure 7(a): Price deviation and consumer choice I

Here α_i , $i = 1, 2$, indicates the density of marginal consumers along the line segment α_i on the graph, and 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 area (which have a measure of $\varepsilon(\alpha_1 + \alpha_2)$), they now buy one more product from firm j and so firm j makes $\rho - \delta$ extra profit 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 some other firms to buying both from firm j . So firm j makes $2\rho - \delta$ extra profit 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:

$$2(\rho - \delta)\alpha(\delta) + (2\rho - \delta)\gamma(\delta) = \Omega_b(\delta) , \quad (40)$$

where $\alpha(\delta)$ and $\gamma(\delta)$ are defined in (23) and $\Omega_b(\delta)$ is defined in (22).

Second, suppose firm j raises its stand-alone 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 consumer choices: both Ω_1 and Ω_2 shrink because now fewer consumers buy a single product from firm j .

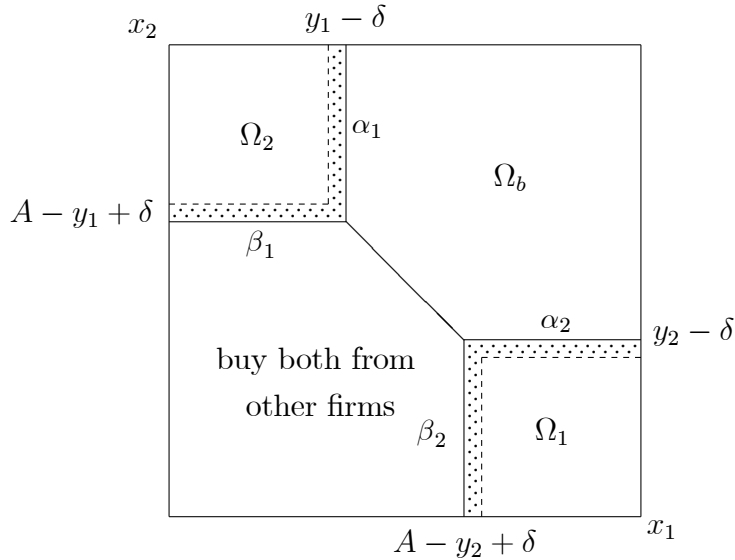


Figure 7(b): Price deviation and consumer choice II

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) yields 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 make $\rho - \delta$ extra profit from each of them. For those marginal consumers with a measure of $\varepsilon(\beta_1 + \beta_2)$, they switch from buying one product to buying nothing from firm j . So firm j loses ρ from each of them. The direct revenue effect of the deviation is that firm j earns 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:

$$(\rho - \delta)\alpha(\delta) + \Omega_1(\delta) = \rho\beta(\delta) , \quad (41)$$

where $\alpha(\delta)$ and $\beta(\delta)$ are defined in (23) and $\Omega_1(\delta)$ is defined in (21), and we have used the fact $\Omega_1(\delta) = \Omega_2(\delta)$.

Both (40) and (41) are linear in ρ . By using (22), it is straightforward to solve ρ as in (24). Substituting it into (41) yields (25) which is an equation of δ .

Proof of Proposition 7:

(i) The duopoly case. The existence of solution has been shown in the main text. To prove $\delta < \rho$, notice that it is equivalent to

$$\alpha(\delta) + \gamma(\delta) = h(\delta)[1 - H(\delta)] + 2 \int_0^\delta h(t)^2 dt < \frac{1}{2\delta} . \quad (42)$$

On one hand,

$$\int_0^\delta h(t)^2 dt < h(0)[H(\delta) - \frac{1}{2}] .$$

(This is because the log-concavity and symmetry of $h(t)$ implies that $h(t)$ is decreasing in $t > 0$ and $H(0) = \frac{1}{2}$.) On the other hand, (26) and the log-concavity of $h(t)$ imply that

$$\delta = \frac{1 - H(\delta)}{h(\delta)} < \frac{1 - H(0)}{h(0)} = \frac{1}{2h(0)} .$$

Then a sufficient condition for (42) is

$$h(\delta)[1 - H(\delta)] + h(0)[2H(\delta) - 1] < h(0) \Leftrightarrow h(\delta) < 2h(0) .$$

This is clearly true since $h(t)$ is decreasing in $t > 0$.

To prove the bundle price comparison result, notice that the bundle price in the duopoly case is $2\rho - \delta = 1/[2(\alpha(\delta) + \gamma(\delta))]$, and the bundle price in the regime of separate sales is $1/h(0)$. The former is lower if

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

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 log-concave. This is implied by the log-concavity of f .

(ii) The case with large n . The proof consists of a few steps.

Step 1: Approximate $\alpha(\delta)$, $\beta(\delta)$, $\gamma(\delta)$ and $\Omega_1(\delta)$ when δ is small.

Lemma 9 *For a given n , if $\delta \approx 0$, we have*

$$\begin{aligned}\alpha(\delta) &\approx \frac{h(0)}{n} - \left(\frac{h'(0)}{n} + \frac{h(0)^2}{n-1} \right) \delta, \\ \beta(\delta) &\approx \left(1 - \frac{1}{n} \right) h(0) + \left(\frac{h'(0)}{n} - h(0)^2 \right) \delta, \\ \gamma(\delta) &\approx \frac{nh(0)^2}{n-1} \delta, \\ \Omega_1(\delta) &\approx \frac{1}{n} \left(1 - \frac{1}{n} \right) - \frac{2h(0)}{n} \delta,\end{aligned}\tag{43}$$

where $h(0) = \int f(x)dF(x)^{n-1}$ and $h'(0) = \int f'(x)dF(x)^{n-1}$.³

Proof. We 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 (19), we have

$$\begin{aligned}\mathbb{E}[\psi(y_1, y_2, A)] &= \frac{1}{n-1} \mathbb{E}_{y_1, y_2} [\psi(y_1, y_2, y_1 + y_2)] \\ &\quad + \frac{n-2}{n-1} \mathbb{E}_{y_1, y_2} [L(y_1 + y_2 - \delta) \psi(y_1, y_2, y_1 + y_2 - \delta) + \int_{y_1 + y_2 - \delta}^{y_1 + y_2} \psi(y_1, y_2, z) dL(z)],\end{aligned}$$

where $L(z)$ is defined in (20). By integration by parts and using $L(y_1 + y_2) = 1$, we can simplify this to

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

Now let us derive the first-order approximation of $\alpha(\delta)$. (For our purpose, we do not need the higher-order approximations.) According to the formula above, we have

$$\alpha(\delta) = \mathbb{E} [f(y_1 - \delta)(1 - F(y_2 + \delta))] + \frac{n-2}{n-1} \mathbb{E}[\varphi(\delta)],\tag{44}$$

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 .

³When the support of x_i is finite and $f(\bar{x}) > 0$, 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 $\lim_{t \rightarrow 0^-} h'(t) = \int f'(x)dF(x)^{n-1}$ and $\lim_{t \rightarrow 0^+} h'(t) = \int f'(x)dF(x)^{n-1} - (n-1)f(\bar{x})^2$. We use $h'(0^-)$ in our approximations.

When $\delta \approx 0$, we have $f(y_1 - \delta) \approx f(y_1) - \delta f'(y_1)$, so

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

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

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

To approximate $\mathbb{E}[\varphi(\delta)]$, notice that $\varphi(0) = 0$ and $\varphi'(0) = f(y_1)f(y_2)$ since $L(z)$ is independent of δ and $L(y_1 + y_2) = 1$. Hence,

$$\mathbb{E}[\varphi(\delta)] \approx \delta \mathbb{E}[f(y_1)f(y_2)] = \delta h(0)^2.$$

Substituting these approximations into (44) and discarding all higher order terms yields the approximation for $\alpha(\delta)$ in (43). The other approximations can be derived similarly.

■

Step 2: When n is large, the system of (24) and (25) has a solution with a small δ .

Lemma 10 *Suppose $\frac{|f'(x)|}{f(x)}$ is bounded and $\lim_{n \rightarrow \infty} p = 0$, where $p = \frac{1}{nh(0)}$ is the separate sales price in (1). Then when n is sufficiently large, the system of (24) and (25) has a solution with $\delta \in (0, \frac{1}{nh(0)})$.*

Proof. Recall that (25) is

$$\underbrace{\frac{1/n + \delta(\alpha(\delta) + \gamma(\delta))}{\alpha(\delta) + \beta(\delta) + 2\gamma(\delta)}}_{\rho(\delta)} (\beta(\delta) - \alpha(\delta)) = \Omega_1(\delta) - \delta\alpha(\delta).$$

Denote the left-hand side by $\chi_L(\delta)$ and the right-hand side by $\chi_R(\delta)$. Notice that the assumption that $\frac{|f'(x)|}{f(x)}$ is bounded implies that $\frac{|h'(0)|}{h(0)}$ is uniformly bounded for any n .⁴

We first show that $\chi_L(0) < \chi_R(0)$. At $\delta = 0$, it is easy to verify that $\alpha(\delta) = \frac{1}{n}h(0)$, $\beta(\delta) = (1 - \frac{1}{n})h(0)$, $\gamma(\delta) = 0$ and $\Omega_1(\delta) = \frac{1}{n}(1 - \frac{1}{n})$. Then $\rho(0) = \frac{1}{nh(0)}$ and

$$\chi_L(\delta) = \frac{1}{n}(1 - \frac{2}{n}) < \chi_R(\delta) = \frac{1}{n}(1 - \frac{1}{n}).$$

Next, we show that $\chi_L(\delta) > \chi_R(\delta)$ at $\delta = \frac{1}{nh(0)}$ when n is sufficiently large. The condition $\lim_{n \rightarrow \infty} p = 0$ implies that $\delta = \frac{1}{nh(0)} \approx 0$ when n is large. Replacing δ in (43) by $\frac{1}{nh(0)}$, we have

$$\alpha(\delta) \approx \frac{h(0)}{n} - \left(\frac{h'(0)}{n} + \frac{h(0)^2}{n-1} \right) \frac{1}{nh(0)} = \frac{h(0)}{n} - \frac{h'(0)}{n^2 h(0)} - \frac{h(0)}{n(n-1)}.$$

⁴Suppose $\frac{|f'(x)|}{f(x)} < M$ for a constant $M < \infty$. Then $-Mf(x) < f'(x) < Mf(x)$, and so $-M \int f(x) dF(x)^{n-1} < \int f'(x) dF(x)^{n-1} < M \int f(x) dF(x)^{n-1}$ for any n . That is, $-Mh(0) < h'(0) < Mh(0)$ for any n , and so $\frac{|h'(0)|}{h(0)}$ is uniformly bounded.

Similarly,

$$\beta(\delta) \approx \left(1 - \frac{1}{n}\right) h(0) + \left(\frac{h'(0)}{n} - h(0)^2\right) \frac{1}{nh(0)} = \left(1 - \frac{2}{n}\right) h(0) + \frac{h'(0)}{n^2 h(0)},$$

$$\gamma(\delta) \approx \frac{nh(0)^2}{n-1} \frac{1}{nh(0)} = \frac{h(0)}{n-1},$$

and

$$\Omega_1(\delta) \approx \frac{1}{n} \left(1 - \frac{1}{n}\right) - \frac{2h(0)}{n} \frac{1}{nh(0)} = \frac{1}{n} - \frac{3}{n^2}.$$

(Notice that in each expression we just replaced δ by $\frac{1}{nh(0)}$ and no further approximations have been made.)

Notice that $\chi_L(\delta) > \chi_R(\delta)$ if and only if

$$\left[\frac{1}{n} + \delta(\alpha(\delta) + \gamma(\delta))\right] [\beta(\delta) - \alpha(\delta)] > [\Omega_1(\delta) - \delta\alpha(\delta)] [\alpha(\delta) + \beta(\delta) + 2\gamma(\delta)]. \quad (45)$$

Using the above approximations, we have

$$\alpha(\delta) + \gamma(\delta) \approx \frac{2h(0)}{n} - \frac{h'(0)}{n^2 h(0)} \quad \text{and} \quad \beta(\delta) - \alpha(\delta) \approx \left(1 - \frac{3}{n}\right) h(0) + \frac{2h'(0)}{n^2 h(0)} + \frac{h(0)}{n(n-1)}.$$

Then the left-hand side of (45) equals

$$\begin{aligned} & \left[\frac{1}{n} + \frac{1}{nh(0)} \left(\frac{2h(0)}{n} - \frac{h'(0)}{n^2 h(0)}\right)\right] \times \left[\left(1 - \frac{3}{n}\right) h(0) + \frac{2h'(0)}{n^2 h(0)} + \frac{h(0)}{n(n-1)}\right] \\ &= \left[\frac{1}{n} + \frac{2}{n^2} - \frac{1}{n^3} \frac{h'(0)}{h(0)^2}\right] \times \left[h(0) - \frac{3}{n} h(0) + \frac{2h'(0)}{n^2 h(0)} + \frac{h(0)}{n(n-1)}\right] \\ &\approx \left(\frac{1}{n} - \frac{1}{n^2}\right) h(0). \end{aligned}$$

(The final step is from discarding all higher order terms. This is valid given $\lim_{n \rightarrow \infty} \frac{1}{nh(0)} = 0$ and $\frac{|h'(0)|}{h(0)}$ is uniformly bounded for any n .)

Using the approximations, we also have

$$\Omega_1(\delta) - \delta\alpha(\delta) \approx \frac{1}{n} - \frac{4}{n^2} + \frac{1}{n^2(n-1)} + \frac{h'(0)}{n^3 h(0)^2},$$

and

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

Then the right-hand side of (45) equals

$$\left[\frac{1}{n} - \frac{4}{n^2} + \frac{1}{n^2(n-1)} + \frac{h'(0)}{n^3 h(0)^2}\right] \times \frac{nh(0)}{n-1} \approx \left(\frac{1}{n} - \frac{3}{n(n-1)}\right) h(0).$$

(The final step is again from discarding all higher order terms.) Then it is ready to see that $\chi_L(\delta) > \chi_R(\delta)$ at $\delta = \frac{1}{nh(0)}$ when n is sufficiently large. This completes the proof of the lemma. ■

Step 3: Approximate the solution to the system of (24) and (25) when n is large.

Given the system has a solution with a small δ when n is large, we can approximate each side of (25) around $\delta \approx 0$ by using (43) and discarding all higher order terms. Then one can solve

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

It is clear that $\rho < p = \frac{1}{nh(0)}$. Since n is large and $\frac{|h'(0)|}{h(0)}$ is uniformly bounded for any n , this can be further approximated as

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