Competitive Bundling*

Jidong Zhou
School of Management
Yale University

March 2016
(First Version: February 2014)

Abstract

This paper proposes a framework for studying competitive bundling in an oligopoly market. We find that under certain conditions, relative to separate sales pure bundling raises market prices, benefits firms, and harms consumers when the number of firms is above a threshold (which can be small). 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. With mixed bundling, having more than two firms raises significant 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 a prevalent business strategy by which a multiproduct firm sells its products in a package. For example, in the market for CDs, newspapers, or cable TV, firms do not usually sell songs, articles, or TV channels separately. This is the case of pure bundling. Other examples include banking accounts, party services, and repair services tied with the product. Sometimes both the package and individual products are available for purchase, but the package is offered at a discounted price relative to the sum of its component prices. This is the case of mixed bundling. Relevant examples include software suites, TV-internet-phone bundles, season tickets, package tours, and value meals.

Since Stigler (1968) there has been substantial research in studying when bundling is a profitable strategy and how bundling affects consumer welfare and overall market efficiency. These questions are often studied in a monopoly context, but in many cases bundling occurs in markets where firms compete with each other. With competition bundling has broader interpretations. For example, pure bundling can be regarded as an outcome of product incompatibility. Consider a system (e.g., a computer, a stereo system) that consists of several essential components (e.g., hardware and software, receiver and speaker). If firms make their components incompatible with each other, then consumers have to buy the whole system from a single firm. Similarly if consumers need to incur an extra shopping cost to source from more than one firm, they will be more likely to buy all the products they want from a single firm. This paper aims to offer a framework for studying competitive bundling in an oligopoly market.

The main anti-trust concern about bundling is that it may restrict market competition. One possible reason, as suggested by the leverage theory (Whinston, 1990), is that bundling can be used by a multiproduct firm to deter the entry of potential single-product competitors (or to induce the exit of existing competitors). Another possible reason is that even for a given market structure, bundling may relax competition among multiproduct firms and inflate market prices because it changes the space of pricing strategies. The existing research on competitive bundling, however, suggests that bundle-against-bundle competition tends to be fiercer than competition

---

1 One obvious reason for bundling is economies of scale in production, selling or buying, or complementarity in consumption. This paper focuses on bundling as a price discrimination device to extract more consumer surplus.


3 See also Choi and Stefanadis (2001), Carlton and Waldman (2002), and Nalebuff (2004).

4 Another possibility is that when a multiproduct firm competes with a single-product rival, if consumers have heterogenous 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.
with separate sales, and so this danger of high prices is usually not a concern.\textsuperscript{5} This is part of the reason why the anti-trust practice concerning bundling mainly focuses on its foreclosure effect. Nevertheless this assessment of bundling is based on duopoly models.\textsuperscript{6} By developing a general oligopoly model, we will confirm the existing insight in the duopoly case under more general conditions, but more importantly we will also argue that going beyond the duopoly case can qualitatively change our view of the impact of bundling, especially in the pure bundling case.

The first contribution of the paper is to offer an oligopoly model which can be used to study both pure and mixed bundling. The existing works on competitive bundling often focus on the case with two firms and each selling two products. They use a two-dimensional Hotelling framework to model product differentiation, where consumers are distributed on a square.\textsuperscript{7} With more firms and more products, however, it becomes less convenient to use a spatial model. (For example, if there are four firms and each sells three products, it is not obvious what spatial models are easy to use.) In this paper, we give up the usual spatial approach and instead adopt a multiproduct version of the random utility framework in Perloff and Salop (1985). Specifically, a consumer’s valuations for a firm’s products are random draws from some joint distribution, and the realization is independent across firms and consumers, reflecting the idea that firms sell differentiated products and consumers have idiosyncratic tastes. This framework can accommodate any number of firms and products. 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 those from the existing duopoly models.

We start our analysis with a general investigation of how the dispersion of consumer valuations affects price competition in a single-product environment. This is because pure bundling essentially reduces consumer valuation heterogeneity, such that the per-product valuation density becomes more peaked but has thinner tails compared to the original single-product valuation distribution. Intuitively this is because finding a well-matched bundle is harder than finding a well-matched component. We show that whether a less dispersed valuation distribution leads to a lower market price depends on how to rank dispersion and often also on the number of competitors in the market.

This investigation itself contributes to the oligopoly competition literature. Besides our bundling application, it is also useful for studying the impact on price competition

\textsuperscript{5}See, e.g., Matutes and Regibeau (1988), Economides (1989), and Nalebuff (2000) for competitive pure bundling, and Matutes and Regibeau (1992), Anderson and Leruth (1993), and Armstrong and Vickers (2010) for competitive mixed bundling. See also Section 7 in Stole (2007) and Section 4 in Armstrong (2015) for surveys of this literature.

\textsuperscript{6}Economides (1989) studies competitive pure bundling with any number of firms and each selling two products, but he comes to the same conclusion. We will discuss this paper in section 4.6.

\textsuperscript{7}Introducing product differentiation is necessary for studying competitive bundling if firms have similar cost conditions. Otherwise, prices would settle at marginal costs and there would be no meaningful scope for bundling. Anderson and Leruth (1993) is the only competitive bundling paper which does not adopt a spatial model. They use a logit model to study competitive mixed bundling in duopoly.
of any economic activity (such as information disclosure, advertising, and product design) which changes the dispersion of consumer valuations.

We then show that the number of firms can qualitatively matter for the impacts of pure bundling on prices, profits, and consumer welfare. In the duopoly case we generalize the existing findings: compared to separate sales, pure bundling intensifies price competition and lowers market prices and profits under a regularity condition. 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 certain 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 for some valuation distributions). This suggests that bundling can be anti-competitive even if it does not influence market structure.

To understand these two contrasting results, notice that 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 tend to have a high valuation for its product because with a high chance 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 tends to induce 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 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 of its 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 can work against and even dominate the double profit effect.

In the section on mixed bundling, we find that considering more than two firms raises significant new challenges in analysis due to the complexity of the consumer choice problem. We propose a method to solve the pricing game with mixed bundling. When the products at the same firm are symmetric and have independent valuations, the equilibrium prices have simple approximations when the number of firms is large. Both the single-product price and the bundle price become lower relative to the sep-

\(^8\)More precisely, the average position of marginal consumers differs between the two regimes, and their relative distance also matters for the price comparison (especially when the support of the valuation distribution is unbounded). That is why as we will see later there are also cases where bundling always lowers market prices. But even in those cases we will show that bundling starts to harm consumers when the number of firms exceeds a usually small threshold.
arate sales case, and 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 are 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 con-
ditions.

We also study firms’ incentives to bundle. When pure bundling is the only alter-
native to separate sales (e.g., when bundling is a product compatibility strategy), the
number of firms matters for the incentive to bundle. Bundling is the unique Nash
equilibrium outcome in duopoly, but when the number of firms is above some thresh-
old, separate sales can be an equilibrium outcome as well. In some examples separate
sales is another equilibrium if and only if consumers obtain higher equilibrium surplus
in the regime of separate sales than pure bundling. When firms can choose the more
flexible mixed bundling strategy, we provide conditions on valuation distributions un-
der which starting from separate sales each firm has a strict incentive to introduce
mixed bundling. In that case, separate sales can never be an equilibrium outcome,
regardless of the number of firms in the market.

The rest of the paper is organized as follows: Section 2 presents the model, and
Section 3 analyzes the benchmark case of separate sales where we investigate how con-
sumer valuation dispersion affects price competition. Section 4 studies pure bundling,
starting with a relatively simple setting with symmetric products and independent
valuations and then extending the main results to the general case. Section 5 deals
with mixed bundling. (A discussion of the related literature will be provided in each
part.) We conclude in Section 6. 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 prices as markups). Each product is horizontally
differentiated across firms (e.g., each firm produces a different version of the product).
We adopt a multiproduct version of the random utility framework in Perloff and Salop
(1985) to model product differentiation. Let \( x^j_l \equiv (x^j_{1,l}, \ldots, x^j_{m,l}) \), \( j = 1, \ldots, n \), denote
the match utilities of firm \( j \)'s \( m \) products for consumer \( l \). We assume that \( x^j_l \) is i.i.d.
across consumers, which reflects, for instance, idiosyncratic consumer tastes. In the
following we therefore suppress the subscript \( l \). Suppose that \( x^j \) is also i.i.d. across
firms (so firms are \emph{ex ante} symmetric), and it is distributed according to a common
joint cumulative distribution function (CDF) \( F(x_1, \ldots, x_m) \). \( F \) has support \( S \subset \mathbb{R}^m \).
and a bounded and differentiable density function $f(x_1, \cdots, x_m)$. We assume that $S$ is of full dimension. (This excludes the possibility that some products in a firm have perfectly correlated match utilities.) Let $F_i(x)$ and $f_i(x)$, $i = 1, \cdots, m$, be the marginal CDF and density function of $x_i^j$, and let $[\underline{x}_i, \overline{x}_i]$ be its support (where $\underline{x}_i = -\infty$ and $\overline{x}_i = \infty$ are allowed). When the products are symmetric, or sometimes for notational convenience, we also use $F(x)$ and $f(x)$ to denote the marginal CDF and density function.

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. 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.) If a consumer consumes $m$ products with match utilities $(x_1, \cdots, x_m)$ (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 $p^j \equiv (p^j_1, \cdots, p^j_m)$. If a firm adopts the pure bundling strategy and sells its products in a package only, it chooses a bundle price $P^j$. In the first part of the paper, we assume that firms can only take one of these 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 or not to bundle its products. 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 $P^j_s$ for each possible subset $s$ of its $m$ products. (If $m = 2$, it can be described by a pair of stand-alone prices $(\rho^j_1, \rho^j_2)$ together with a joint-purchase discount $\delta^j$.) 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., all 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, consumers have a sufficiently high basic valuation for each product (such that the lower bound of the valuation 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 full market coverage assumption in Section 4.5 and argue that the basic insights remain

---

9This assumption is made in all the papers on competitive (pure or mixed) bundling, though it is not always 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).
qualitatively unchanged. However, in the regime of mixed bundling, as known in the literature this assumption is important for tractability.

3 Separate Sales: Revisiting Perlo¤-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 Perlo¤-Salop model where only the marginal distribution of that product’s match utilities matters. Consider the market for product \( i \), and let \( p_i \) be the (symmetric) equilibrium price.

Suppose \( j \) deviates to price \( p'_j \), while other firms stick to the equilibrium price \( p_i \). Then the demand for firm \( j \)’s product \( i \) is

\[
q_i(p'_j) = \Pr[x_i' - p'_j > \max_{k \neq j} \{x_k - p_i\}] = \int_{\mathbb{R}} [1 - F_i(x - p_i + p'_j)]dF_i(x)^{n-1}.
\]

Notice that \( F_i(x)^{n-1} \) 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 \( F_i(x)^{n-1} \) and charges price \( p_i \). In equilibrium the demand is \( q_i(p_i) = 1/n \) since firms are symmetric to each other.

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

\[
\frac{1}{p_i} = n \int f_i(x)dF_i(x)^{n-1}.
\]  

(Whenever there is no confusion we suppress the integral limits.) Henceforth, we assume that this first-order condition is also sufficient for defining the equilibrium price. This is the case, for example, when \( f_i \) is log-concave (see Caplin and Nalebuff, 1991).  

In the uniform distribution example with \( F_i(x) = x \), it is easy to see that \( p_i = 1/n \), and in the extreme value distribution example with \( F_i(x) = e^{-e^{-x}} \) (which generates the logit model), one can check that \( p_i = n/(n - 1) \). Notice that with full market coverage, shifting the support of the match utility does not affect the equilibrium price.

How does the equilibrium price vary with the number of firms? Let us rewrite (1) as

\[
p_i = \frac{q_i(p_i)}{q'_i(p_i)} = \frac{1/n}{\int f_i(x)dF_i(x)^{n-1}}.
\]  

\[\text{10}\] In 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 \( f_i \) is log-concave.

\[\text{11}\] Many often used distributions such as uniform, normal, logistic, and extreme value have a log-concave density. Caplin and Nalebuff (1991) provide a slightly weaker sufficient condition which requires \( f_i \) to be \(-1/(n + 1)\)-concave for a given \( n \).
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 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 $f_i$. For example, if the density $f_i$ is increasing, it increases with $n$ and so $p_i$ must decrease with $n$. Conversely if $f_i$ is decreasing, it decreases with $n$, which works against the demand size effect.

The following result reports a sufficient condition for $p_i$ to decrease with $n$.

**Lemma 1** Suppose $1 - F_i$ is log-concave. Then $p_i$ defined in (1) decreases with $n$. Moreover, $\lim_{n \to \infty} p_i = 0$ if and only if $\lim_{x \to \bar{x}} \frac{f_i(x)}{1 - F_i(x)} = \infty$.

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

$$f_{(n-1)}(x) = n(n-1)(1 - F_i(x))F_i(x)^{n-2}f_i(x),$$

we can rewrite (1) as

$$\frac{1}{p_i} = \int \frac{f_i(x)}{1 - F_i(x)} dF_{(n-1)}(x).$$

(3)

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

Anderson, de Palma, and Nesterov (1995) is the first paper that proves this monotonicity result (see their Proposition 1) under the assumption that $f_i$ is log-concave (which implies the log-concavity of $1 - F_i$).\(^\text{13}\) Our proof is simpler than theirs, and we also offer a tail behavior condition for the markup to converge to zero in the limit.\(^\text{14}\) Another advantage of our proof is that it also immediately implies that if $1 - F_i$ is log-convex, then $p_i$, defined in (1), increases in $n$. (But in that case it is more a concern whether the first-order condition determines the equilibrium price.)

**Price and the dispersion of consumer valuations.** One important comparative static question for our subsequent analysis is: if the distribution of consumer valuations becomes less “dispersed” from $f$ to $g$ as illustrated in Figure 1 below, how will the

\(^{12}\)The right-hand side of (3) is the density of all marginal consumers in the market. A consumer is a marginal one if her best product and second-best product have the same match utility. Conditional on $x_{(n-1)} = x$, the CDF of $x_{(n)}$ is $\frac{F_i(z) - F_i(x)}{1 - F_i(x)}$ for $z \geq x$, and so its density function at $x_{(n)} = x$ is the hazard rate $\frac{f_i(x)}{1 - F_i(x)}$. 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., $|q_i'(p)|$).

\(^{13}\)See also Weyl and Fabinger (2013), and Quint (2014) for a similar result.

\(^{14}\)The tail behavior condition for $\lim_{n \to \infty} p_i = 0$ is satisfied if $f_i(\bar{x}) > 0$, but it can be violated if $f_i(\bar{x}) = 0$. For instance, in the extreme value distribution example the price $p_i = n/(n-1)$ converges to 1 in the limit.
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.

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.\(^{15}\) For example, \( f \) and \( g \) in Figure 1 can be ranked in this order if they have equal means.\(^{16}\) 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

\(^{15}\)Let \( x_F \) 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 function \( \phi \) 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).

\(^{16}\)According 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.)
Lemma 2 Consider two Perlof-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 $[x_F, \pi_F]$ and $[x_G, \pi_G]$ be their supports, respectively. Without loss of generality suppose $\mathbb{E}[x_F] = \mathbb{E}[x_G]$. Let $p_F$ and $p_G$ be the equilibrium prices, and suppose they 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 $f(\pi_F) > g(\pi_G)$, 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

$$
\frac{1}{p_F} = n \int_{\pi_F}^{x} f(x) dF(x)^{n-1} = n \int_{0}^{1} l_F(t) dt^{n-1},
$$

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

$$
p_G \leq p_F \iff \int_{0}^{1} [l_F(t) - l_G(t)] dt^{n-1} \leq 0. \tag{4}
$$

(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

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

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 $f(\pi_F) \leq g(\pi_G)$.)

(ii) $f(\pi_F) > g(\pi_G)$ implies $l_F(1) - l_G(1) > 0$. Then

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

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 $n \to \infty$. Then it follows from (4) that $p_G > p_F$ when $n$ is sufficiently large.18

---

17 When 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.

18 This argument cannot necessarily be extended to the case where $f(\pi_F) = g(\pi_G)$ 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 $t \approx 1$. 


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. Perlo¤ 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 \( \theta x <_{\text{disp}} x \) (where “\( <_{\text{disp}} \)” 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.)

Notice that \( f(\pi_F) > g(\pi_G) \) 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 \( f(1) > g(1) \). Since equilibrium demand is always \( 1/n \) due to firm symmetry, from (2) we know that 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 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.\(^{19}\)

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 any firm or consumer activity (such as information dis-

\(^{19}\)Gabaix et al., 2015, study the asymptotic behavior of the equilibrium price in random utility models 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.
closure/acquisition, advertising, product design, and spurious product differentiation) which changes 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 $c_i$ for product $i$, a sufficient condition is $c_i > \bar{x}_i - \underline{x}_i$ for all $i = 1, \ldots, m$.) As we will discuss in the conclusion, allowing consumers to buy multiple bundles will make the situation like mixed bundling.

Let $X^j \equiv \sum_{i=1}^m x^j_i$ 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

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

It is more convenient to work on the per-product bundle price $P/m$, so we divide everything by $m$. Let $G$ and $g$ denote the CDF and density function of $X^j/m$, and let $[\underline{x}_G, \bar{x}_G]$ be its support. Then the first-order condition for $P/m$ to be the equilibrium per-product bundle price is similar to (1), except that now a different distribution $G$ applies:

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

We assume that this first-order condition is also sufficient for defining the equilibrium bundle price. This is the case, for example, when $g$ is log-concave (which is true if the joint density function $f(x_1, \ldots, x_m)$ is log-concave). Similar results as in Lemma 1 hold here.

Lemma 3 Suppose $1 - G$ is log-concave. Then $P$ defined in (5) decreases with $n$. Moreover, $\lim_{n \to \infty} P = 0$ if and only if $\lim_{x \to x_G} g(x) \frac{x}{1 - G(x)} = \infty$.

In the following, we compare the bundling regime with the benchmark regime of separate sales. For transparency, in Sections 4.1-4.3 we will focus on the “i.i.d.” case where the $m$ products are symmetric and have independent match utilities at each

---

\[ \text{Notice that } g(t) = \int_{x_{-1}}^t f(mt - \sum_{i=2}^m x_i, x_{-1}) dx_{-1}, \text{ where } x_{-1} = (x_2, \ldots, x_m). \text{ It is easy to check that the integrand is log-concave in } (t, x_{-1}) \text{ if the joint density } f \text{ is log-concave. Then the Prékopa theorem implies that } g \text{ is log-concave (and so is } 1 - G). \]
firm. In Section 4.4 we will show that all the main results derived from the i.i.d. case can be extended to the general case under certain conditions.

Whenever there is no confusion, let \( F(x) \) and \( f(x) \) denote the common marginal CDF and density function in the i.i.d. case. Then \( F(x_1, \ldots, x_m) = \prod_{i=1}^{m} F(x_i) \) and \( f(x_1, \ldots, x_m) = \prod_{i=1}^{m} f(x_i) \). Let \([\underline{x}, \overline{x}]\) be the common support (where \( \underline{x} = -\infty \) and \( \overline{x} = \infty \) are allowed), and let \( p \) be the common separate sales price for all products. All the existing works on competitive pure bundling focus on this relatively simple case (often with only two firms and two products in a Hotelling setting).

### 4.1 Comparing prices and profits

In the i.i.d. case, 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 \( g \) can be less dispersed than \( f \) as illustrated in Figure 1 above. In particular, \( g(\overline{x}) = 0 \) even if \( f(\overline{x}) > 0 \). (This is because \( X_j/m = \overline{x} \) only if \( x_i^j = \overline{x} \) for all \( i = 1, \ldots, m \).) Intuitively this is because finding a well-matched bundle is much harder than finding a well-matched single product.

From (1) and (5), we can see that the comparison between \( p \) and \( P/m \) 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). (One exception is the class of stable distributions as we will discuss later.) In the following, we will show that in the duopoly case bundling lowers 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

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

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. The following proposition reports our price and profit comparison results. (All the omitted proofs can be found in the appendix.)

**Proposition 1**

(i) When \( n = 2 \) and \( f(x) \) is log-concave, bundling reduces market prices and profits for any \( m \geq 2 \).

(ii) For a fixed \( m \), if \( f(\overline{x}) > 0 \), there exists \( \hat{n} \) such that bundling increases market prices and profits for \( n > \hat{n} \). If \( f(x) \) is further log-concave and \( l_F(t) \) and \( l_G(t) \) cross each other at most twice, then bundling decreases prices and profits if and only if \( n \leq \hat{n} \).

\[21\] Formally, when \( m = 2 \) the density function of \((x_1^2 + x_2^2)/2\) is \( g(t) = 2 \int_{2t-\overline{x}}^{\overline{x}} f(2t-x) dF(x) \) for \( t \geq (\underline{x} + \overline{x})/2 \), so \( g(\overline{x}) = 0 \). This argument also works for \( m \geq 3 \).
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.\(^{22}\) Each firm therefore has more marginal consumers in the case of \( g \), and so they face a more elastic demand and charge lower prices.

The first part of result (ii) follows immediately from result (ii) in Lemma 2 as \( f(x) > g(x) = 0 \). Bundling makes the right tail of the valuation density thinner, 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. Extra work is needed to prove the cut-off result.

The following two examples show that the threshold \( \hat{n} \) in result (ii) can be small and bundling can significantly increase the market price. In the uniform distribution example with \( f(x) = 1 \) and \( m = 2 \), \( P/m < p \) if and only if \( n \leq 6 \). Figure 2(a) below describes how both prices vary with \( n \) (where the solid curve is \( p \) and the dashed one is \( P/m \)). In the example with an increasing density \( f(x) = 4x^3 \) and \( m = 2 \), as described in Figure 2(b) below \( P/m < p \) if and only if \( n = 2 \). Another observation from these two examples is that the increase of price caused by bundling can be proportionally significant even if \( n \) is relatively large. For example, when \( n = 10 \) the per-product bundle price \( P/m \) is about 25% higher than the separate sales price \( p \) in the first example and 57% higher in the second one.\(^{23}\)

![Figure 2: Price comparison with \( m = 2 \)](image)

\( \text{Discussion: the log-concavity condition.} \) As long as equilibrium prices are determined by the first-order conditions (for which a log-concave \( f(x) \) is a sufficient but

\(^{22}\)Notice 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 adjusts \( P/m \) by a small \( \varepsilon \).

\(^{23}\)Under the tail behavior conditions in Lemmas 1 and 3, both prices converge to zero as \( n \to \infty \). But they can converge to zero at very different speeds. For instance, \( \lim_{n \to \infty} \frac{p}{P/m} = \frac{g(\bar{x})}{f(\bar{x})} = 0 \) when \( f(\bar{x}) > 0 \) and \( g(\bar{x}) = 0 \). (When the cost is positive, prices should be interpreted as markups.)
not a necessary condition), the log-concavity of \( f(x) \) is only needed for the duopoly result but not for the large-\( n \) result. It is used to ensure that \( g \) is more peaked than \( f \) (which is a stronger requirement than mean-preserving contraction). If \( f(x) \) is not log-concave, as we will show below by examples, bundling can raise market prices even in the duopoly case.

**Discussion:** the condition of \( f(\bar{x}) > 0 \). What is the economic meaning of \( f(\bar{x}) > 0 \)? One interpretation is that this condition ensures that the equilibrium price \( p \) in the regime of separate sales converges to the marginal cost fast enough (at a speed of \( 1/n \)).\(^{24}\) It is also consistent with the existing spatial models where consumer match utility must be bounded. This facilitates the comparison with the existing literature.

Also notice that \( f(\bar{x}) > 0 \) is not a necessary condition for bundling to raise market prices. In fact, it can be shown that if \( f(\bar{x}) = 0 \) but \( f'(\bar{x}) < 0 \) (which requires \( \bar{x} < \infty \)), bundling also raises market prices when \( n \) is above a certain threshold. (The proof is similar but less transparent. It is available upon request.) There are also examples with \( \bar{x} = \infty \) where bundling can raise market prices. We will show this below in the class of stable distributions. But there the density is not log-concave such that bundling raises prices even in the duopoly case. Here is an example with a log-concave density and \( \bar{x} = \infty \). Consider the exponential power distribution with density \( f(x) = \frac{\beta}{2\Gamma(\beta)} e^{-|x|^{\beta}} \), where \( \beta \) is the shape parameter and the support is the whole real line. Suppose \( \beta > 1 \) such that \( f \) is log-concave. This distribution becomes the standard normal when \( \beta = 2 \), and it converges to the uniform distribution on \([-1, 1]\) when \( \beta \to \infty \). Suppose \( \hat{n} \) is the threshold in the case of uniform distribution with support \([-1, 1]\). Then for \( n = 2 \), bundling reduces prices according to result (i) in Proposition 1, but for any \( n > \hat{n} \), there exists a sufficiently large \( \beta \) such that bundling raises market prices.

We have so far emphasized the possibility that bundling can raise market prices when \( n \) is above a threshold (which can be small). But there are many examples with \( \bar{x} = \infty \) (so \( f(\bar{x}) = 0 \)) where bundling always reduces market prices. A simple example is the normal distribution. (Other examples include, for instance, the exponential and the logistic distribution according to numerical simulations.)

**Example of normal distribution.** 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

\[
p = \frac{\sigma}{\Lambda(n)}
\]

with \( \Lambda(n) \equiv n \int_{-\infty}^{\infty} \phi(x) d\Phi(x)^{n-1} \), where \( \Phi \) and \( \phi \) are the CDF and density function of the standard normal distribution \( \mathcal{N}(0, 1) \). The definition of \( X^j \) implies

\(^{24}\) As we discussed in footnote 19, 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 \( 1/n \). If \( f(\bar{x}) = 0 \), the speed of convergence will be slower.
that $X^j/m \sim \mathcal{N}(0, \sigma^2/m)$. Thus, $X^j/m = x^j_i / \sqrt{m} \prec_{\text{disp}} x^j_i$, so result (i) in Lemma 2 implies $P/m < p$. That is, in this example bundling always reduces market prices (and so profits).\textsuperscript{25} More precisely, (7) also implies that

$$\frac{P}{m} = \frac{\sigma / \sqrt{m}}{\Lambda(n)} = \frac{p}{\sqrt{m}}.$$  \hspace{1cm} (8)

Discussion: why can the large-$n$ result in Proposition 1 fail when $\tau = \infty$? 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. Then why does the previous large-$n$ result fail? The reason is that with an unbounded support now the relative moving speed of marginal consumers matters. Let $\hat{x}_f$ and $\hat{x}_g$ denote the (average) position of marginal consumers in the separate sales and the bundling regime, respectively. A thinner right tail implies a smaller chance of having a high valuation draw, so $\hat{x}_g$ moves to the right slower than $\hat{x}_f$. With $\tau = \infty$, even if $f(x) > g(x)$ for all relatively large $x$, it is possible that $\hat{x}_f$ is far ahead of $\hat{x}_g$ when $n$ is large such that $f(\hat{x}_f) < g(\hat{x}_g)$. In that case the $g$ density leads to more marginal consumers and so a lower price. This is what happens in the normal distribution example. This is, however, impossible if $\tau$ is finite. In that case, when $n$ is large both $\hat{x}_f$ and $\hat{x}_g$ will be close to $\tau$, and so we must have $f(\hat{x}_f) > g(\hat{x}_g)$ given $f(\tau) > g(\tau)$.

The key feature in the normal distribution example which makes comparison simple is that $x^j_i$ 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. The following proposition reports a price and profit comparison result among this class of distributions.

**Proposition 2** Suppose $x^j_i$ has a standardized stable distribution $S(\alpha, \beta, 1, 0; 1)$ defined in Definition 1.8 in Nolan (2015), where $\alpha \in (0, 2]$ is the stability parameter and $\beta \in [-1, 1]$ is the skewness parameter. Then bundling reduces market prices and profits for any $n \geq 2$ if and only if $\alpha > 1$.

The only three stable distributions which have a closed formula for the density function are normal, Cauchy, and Lévy. Normal distribution has $\alpha = 2$ and $\beta = 0$, so bundling reduces market prices as we already know. Cauchy distribution has $\alpha = 1$ and $\beta = 0$. It is the edge case where bundling has no impact on market prices. When $\alpha < 1$ (of which Lévy distribution is one example with $\alpha = 1/2$), bundling raises market prices. (This does not contradict with the duopoly result in Proposition 1

\textsuperscript{25}However, for any truncated normal distribution with a finite upper bound, result (ii) in Proposition 1 applies. For instance, for the truncated standard normal with support $[-1, 1]$ bundling leads to a higher market price if $n > 9$. 

15
because a stable distribution with $\alpha < 1$ does no longer have a log-concave density.)

In sum, our discussion so far suggests: (i) If the match utility distribution has a finite support, the impact of bundling on prices and profits is reversed when the number of firms exceeds a threshold (which can be small). (ii) If the distribution has an infinite support, the details of the tail behavior also matter. In that case, although we do not have a general result, our discussion suggests that if the tail is close to the normal distribution, then bundling tends to reduce market prices.

The case with many products. Finally we briefly discuss the case with a large number of products. This case makes sense in the markets for, for example, newspapers, magazines, and TV packages if we regard each article or channel as a product. As long as $x_i^j$ has a finite mean $\mu$, the law of large numbers implies that $X^j/m$ converges to $\mu$ as $m \to \infty$. That is, with many products the per-product valuation for the bundle tends to be homogeneous across firms. This leads to a Bertrand competition and so a zero markup. Therefore, for a fixed $n$, $\lim_{m \to \infty} P/m = 0$ and so there exists $\hat{m}$ such that bundling reduces market prices and profits for $m > \hat{m}$. This gives us another case where bundling is pro-competitive and reduces market prices. In the next sections sometimes we rely on this limit case to derive analytical results.

4.2 Comparing consumer surplus and total welfare

Bundling must reduce total welfare (which is the sum of firm profits and consumer surplus) under the assumption of full market coverage. Since consumer payment is a pure transfer, total welfare only reflects the match quality between consumers and products. Pure bundling eliminates the opportunities for consumers to mix and match and so it must reduce the overall match quality. (Notice that this argument does not rely on the i.i.d. assumption. In Section 4.5 we will also argue that a similar result can hold even without the assumption of full market coverage.)

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.

\footnote{This discussion is based on the assumption that the equilibrium price is determined by the first-order condition. This may not be the case for some stable distributions. However, for the Cauchy distribution it can be numerically shown that the first-order condition determines the equilibrium price when $n \leq 5$. So our claim is not empty at least for a relatively small $n$.}

\footnote{For a large $m$, we can also claim that $P/m$ decreases in $m$ at a speed of $1/\sqrt{m}$. Suppose $x_i^j$ has a mean $\mu$ and variance $\sigma^2$. When $m$ is large, by the central limit theorem, $X^j/m$ is distributed (approximately) according to the normal distribution $N(\mu, \sigma^2/m)$. Then (8) implies that $P/m = \frac{\sigma/\sqrt{m}}{N(m)}$. See Nalebuff (2000) for a similar observation in a multi-dimensional Hotelling model with two firms and an arbitrary number of products.}
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
\[
\mathbb{E} \left[ \max_j \{ x_i^j \} \right] - \mathbb{E} \left[ \max_j \left\{ \frac{X_j}{m} \right\} \right] < p - \frac{P}{m}. \tag{9}
\]
The left-hand side (which must be positive) reflects the match quality effect, and the right-hand side is the price effect.

**Proposition 3**

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

(ii) If \( f(x) \) is log-concave, 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 \( f(x) > 0 \) bundling will raise prices (and so harm consumers) when \( n \) is sufficiently large. If \( f(x) = 0 \) (e.g., when \( x = \infty \)), 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).

Result (ii) says that in the limit case with \( m \to \infty \) we have a stronger cut-off result: pure bundling improves consumer welfare if and only if the number of firms is below some threshold. As we pointed out in the end of the last section, \( \lim_{m \to \infty} \frac{X_j}{m} = \mu \) and \( \lim_{m \to \infty} \frac{P}{m} = 0 \). Then for fixed \( n \), bundling benefits consumers in the limit case if and only if
\[
\mathbb{E} \left[ \max_j \{ x_i^j \} \right] - \mu < p. \tag{10}
\]
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 \( n \). Then the cut-off result follows. 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. 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. The harm of bundling can be significant. For example, when
$n = 10$ bundling reduces consumer surplus by about 15% and total welfare by about 11%.\footnote{In this example, the harm of bundling will eventually disappear as $n \to \infty$. This is because $\lim_{n \to \infty} p = \lim_{n \to \infty} P/m = 0$ and $\lim_{n \to \infty} \mathbb{E}[\max_j \{x^j_i\}] = \lim_{n \to \infty} \mathbb{E}[\max_j \{X^j/m\}] = \pi$.}

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 $x^j_i \sim \mathcal{N}(0, \sigma^2)$. From (8) and (9), we can see that pure bundling improves consumer surplus in this example if and only if

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

In the Appendix, we show that

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

Then (11) simplifies to $p > \sigma$. Using (7), one can check that this is equivalent to $\Lambda(n) < 1$, which holds only for $n = 2, 3$. Hence, the threshold is again 3.

In sum, 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.

### 4.3 Incentive to bundle

We now turn to 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, we assume that each firm either bundles all its products or not at all. (This rules out the possibility of using 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 4** Suppose 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) If \( f(x) \) is log-concave, 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 if a firm unilaterally unbundles, the market situation does not change for consumers. This argument of course depends on the assumptions that consumers buy all products and for each product they only buy one variant (e.g., in the market for systems). In the duopoly case, it can be further shown that there is neither symmetric equilibrium with separate sales nor asymmetric equilibrium where one firm bundles and the other does not.

When there are more than two firms, one may wonder whether separate sales can be another equilibrium as well. Result (ii) says that when \( m \to \infty \), 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

\[ y_i = \max_{k \neq j} \{ x_k \} \tag{13} \]

the match utility of the best product \( i \) among firm \( j \)’s competitors. Then firm \( j \) is as if competing with one firm that offers a bundle with match utility \( Y = \sum_{i=1}^{m} y_i \) and price \( mp \). If firm \( j \) charges the same bundle price \( mp \), its demand will be

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

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

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

This is clearly not true for $n = 2$ (which is consistent with result (i) in Proposition 4). 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 3. 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 $\tilde{n} \neq n^*$.)

Discussion: possibility of asymmetric equilibria. With more than two firms one may also wonder the possibility of asymmetric equilibria in which some firms bundle and the others do not. An analytical investigation into this problem is hard because the pricing game when firms adopt asymmetric bundling strategies does not have a simple solution. In the online appendix, we numerically study a uniform distribution example with $n = 3$ and $m = 2$. There are no asymmetric equilibria in that example.

### 4.4 Asymmetric products and dependent valuations

We now return to the general setup with a joint CDF $F(x_1, \cdots, x_m)$. We aim to generalize the main results derived in the i.i.d. case and also explore extra insights from considering asymmetric products and correlated valuations. Recall that $F_i$ and

---

29This 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.
Proposition 5 Suppose the support $S \subset \mathbb{R}^m$ is compact, strictly convex, and has full dimension. Then for a fixed $m$, if $f_i(\overline{x}_i) > 0$, there exists $\hat{n}_i$ such that $P/m > p_i$ for $n > \hat{n}_i$; if $f_i(\overline{x}_i) > 0$ for all $i = 1, \ldots, m$, there exists $\hat{n}$ such that $P > \sum_{i=1}^m p_i$ for $n > \hat{n}$.

Proof. The conditions on $S$, together with $f$ being bounded and differentiable, imply $g(\overline{x}_G) = 0$ (e.g., see the proof of Proposition 1 in Armstrong, 1996). Then the result immediately follows from result (ii) in Lemma 2.  

It is, however, more difficult to fully generalize the duopoly result in Proposition 1. Progress can be made if the match utility of firm $j$’s product $i$ takes the following form:

$$x_i^j = \theta x_0^j + (1 - \theta)(a_i \tilde{x}_i^j + b_i),$$

(16)

where $\theta \in [0,1)$, $a_i > 0$ and $b_i$ are constants, $x_0^j$ is a random variable which has the same realization across firm $j$’s all products for a given consumer but is i.i.d. across firms and consumers, and $\tilde{x}_i^j$ is i.i.d. across products, firms and consumers. The first common utility component in (16) introduces dependence among a consumer’s valuations for different products. It reflects, for instance, consumers’ brand preferences at the firm level. The second random utility component in (16) is similar as in the baseline model, but allowing $(a_i, b_i)$ to differ across products captures potential product asymmetry. (With full market coverage, $b_i$ actually does not play any role in our analysis.) The i.i.d. model corresponds to $\theta = 0$ and $a_i = a_k$ for all $i \neq k$.

Proposition 6 Suppose consumer valuations take the form in (16). Suppose $x_0^j$ and $\tilde{x}_i^j$ are independent of each other, and each has a log-concave density function. Then when $n = 2$ bundling reduces market prices and profits for any $m \geq 2$.

---

30 Notice that the large-$n$ result can hold even if consumer valuations are correlated across firms. A simple case to see this point is when a consumer’s valuation for firm $j$’s product $i$ is $v_i + x_i$, where the new component $v_i$ is a consumer’s basic valuation for product $i$, reflecting, say, the income effect. Suppose it is heterogeneous across consumers but has the same realization across firms for a given consumer. This introduces valuation correlation across firms. Under the full market coverage assumption (which can be justified if the lower bound of $v_i$ is high enough), the equilibrium prices remain unchanged in both separate sales and bundling regimes.

31 Formally, $g(\overline{x}_G) = \lim_{\varepsilon \to 0} \frac{1 - G(\overline{x}_G - \varepsilon)}{\varepsilon}$, and our conditions ensure that $1 - G(\overline{x}_G - \varepsilon) = o(\varepsilon)$. Among the conditions, strict convexity of $S$ excludes the possibility that the plane of $X^j/m = \overline{x}_G$ coincides with a part of $S$’s boundary, and the full dimension condition excludes the possibility that some products have perfectly correlated valuations. Also note that our result holds as long as $f_i(\overline{x}_i) > g(\overline{x}_G)$ even if $g(\overline{x}_G) > 0$. 

21
(In the proof of this result, we also discuss why it is hard to deal with a general joint distribution.)

Other results can also be generalized. It is easy to see that for a fixed \( n \), \( \lim_{m \to \infty} P/m = 0 \) continues to hold provided that \( X^j/m \) converges to a deterministic value as \( m \to \infty \). Both Proposition 3 concerning consumer surplus comparison and Proposition 4 concerning incentive to bundle can also be extended under mild conditions if the joint density \( f \) is log-concave.\(^32\)

Finally, we extend the normal distribution example to the general case.

**Example of normal distribution.** Suppose \( x^j \sim \mathcal{N}(0, \Sigma) \), where \( \sigma_i^2 \) in \( \Sigma \) is the variance of \( x_i^j \) and \( \sigma_{ik} \) in \( \Sigma \) is the covariance of \( (x_i^j, x_k^j) \). Then the marginal distribution of \( x_i^j \) is \( \mathcal{N}(0, \sigma_i^2) \), and \( X^j \sim \mathcal{N}(0, \sum_{i=1}^m \sigma_i^2 + \sum_{i \neq k} \sigma_{ik}) \). According to formula (7), we have

\[
p_i = \frac{\sigma_i}{\Lambda(n)} \quad \text{and} \quad P = \frac{1}{\Lambda(n)} \sqrt{\sum_{i=1}^m \sigma_i^2 + \sum_{i \neq k} \sigma_{ik}}.
\]

Therefore, \( P < \sum_{i=1}^m p_i \) if and only if

\[
\sqrt{\sum_{i=1}^m \sigma_i^2 + \sum_{i \neq k} \sigma_{ik}} < \sum_{i=1}^m \sigma_i.
\]

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. Now consider consumer surplus comparison. Using (12) and (17), one can check that bundling reduces match quality by

\[
\sum_{i=1}^m \mathbb{E} \left[ \max_j \{x_i^j\} \right] - \mathbb{E} \left[ \max_j \{X^j\} \right] = \Lambda(n) \left( \sum_{i=1}^m \sigma_i - \sqrt{\sum_{i=1}^m \sigma_i^2 + \sum_{i \neq k} \sigma_{ik}} \right),
\]

and saves consumer payment by

\[
\sum_{i=1}^m p_i - P = \frac{1}{\Lambda(n)} \left( \sum_{i=1}^m \sigma_i - \sqrt{\sum_{i=1}^m \sigma_i^2 + \sum_{i \neq k} \sigma_{ik}} \right).
\]

Given (18), bundling benefits consumers if and only if \( \Lambda(n) < 1 \), which holds only for \( n = 2, 3 \). This is exactly the same as in the i.i.d. case.

\(^32\)Result (i) in Proposition 3 continues to hold if \( f_i(\tilde{x}_i) > 0 \) for all \( i \), or if \( \lim_{x \to \infty} \frac{d}{dx} \left( \frac{1-F_i(x)}{f_i(x)} \right) = 0 \) for all \( i \). Result (ii) in Proposition 3 and Proposition 4 also continue to hold if \( X^j/m \) converges to \( \lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^m \mathbb{E}[x_i^j] < \infty \), and if for any fixed \( n \), \( \lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^m p_i < \infty \) and \( \lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^m \mathbb{E}[\max_{j \in \{1, \ldots, n\}} \{x_i^j\}] < \infty \). The proofs only need to be modified slightly. The cut-off result in Proposition 1 can be extended if the \( m \) products are symmetric and the duopoly result in Proposition 6 holds. All the details are available upon request.
**Product asymmetry.** Considering asymmetric products has a useful implication. Consider a market where each firm sells two products. Suppose product 1 is almost homogeneous across firms such that price competition in that market is fierce, while product 2 is significantly differentiated across firms. In this scenario, can the firms benefit from bundling their products and selling a package with a “medium” level of differentiation? In other words, do firms have a joint incentive to leverage the product differentiation in one market to the other? Our results tend to suggest that this is not a profitable strategy in the duopoly case or the normal distribution case. While if the marginal distributions have finite supports, this strategy can be profitable when the number of firms is sufficiently large.

**Valuation dependence.** The general setup also allows us to explore how consumer valuation dependence might affect the impact of bundling. The literature on monopoly bundling tends to suggest that bundling is a more profitable strategy when consumer valuations for different products are more negatively correlated. For instance, if a consumer’s valuations for two products are perfectly negatively correlated, the firm can extract the whole consumer surplus by pure bundling. With competition this view may not be true any more because eliminating consumer heterogeneity by bundling can also intensify price competition. But as we have discussed, whether or not bundling intensifies price competition often depends on the number of firms and the tail behavior of the valuation distribution. It is hard to study how consumer valuation dependence affects the bundle price generally.\(^{33}\) In the following, we discuss two examples.

Consider the above normal distribution example with two products. Let \(\rho \sigma_1 \sigma_2\) be the covariance of \((x_1^j, x_2^j)\), where \(\rho \in [-1, 1]\) is the usual correlation coefficient. Then from (17) it is easy to see that the bundle price increases in \(\rho\). That is, more negative correlation reduces the equilibrium bundle price and makes bundling a less attractive strategy for all the firms collectively. (But each firm’s individual incentive to bundle can become higher.)

Consider another example where with probability \(\theta \in [0, 1]\), \(x_1^j = \cdots = x_m^j\) and they are distributed according to a CDF \(F(x)\); with probability \(1 - \theta\), they are i.i.d. and follow the same distribution \(F(x)\). So \(\theta\) is an index of the degree of correlation. The marginal distribution of each \(x_i^j\) is always \(F(x)\) and so the separate sales price is independent of \(\theta\).\(^{34}\) The CDF and density function of \(X^j/m\) are respectively \(G_\theta(x) = \theta F(x) + (1 - \theta)G(x)\) and \(g_\theta(x) = \theta f(x) + (1 - \theta)g(x)\), where \(G\) and \(g\) are the CDF and density function of \(\sum_{i=1}^m x_i^j/m\) when all \(x_i^j\) are i.i.d.. The bundle price is then determined by

\[
\frac{1}{P_{\theta/m}} = n \int g_\theta(x) dG_\theta(x)^{n-1}.
\]

\(^{33}\)Reisinger (2004) and Armstrong and Vickers (2010) have discussed how consumer valuation correlation can affect the impact of competitive mixed bundling in some duopoly examples.

\(^{34}\)We do not use the dependence structure in (16) because in that case the marginal distributions (so the separate sales prices) also vary with \(\theta\). Then we cannot isolate the effect of correlation on the bundle price when \(\theta\) changes.
In the online appendix we show: (i) when \( n = 2 \), \( P_0 \) typically increases in \( \theta \); (ii) while if \( f(\overline{x}) > 0 \), \( P_0 \) decreases in \( \theta \) (i.e., the bundle price increases when the valuations become less correlated) when \( n \) is sufficiently large. In the uniform distribution example, numerical simulations suggest that \( P_0 \) starts to vary with \( n \) non-monotonically (increasing first and then decreasing) from \( n = 4 \).

Intuitively, when \( \{x_i^j\}_{i=1}^m \) become more negatively (or less positively) correlated, \( X_j/m \) becomes more concentrated around the mean and has thinner tails. As we have learned from Section 4.1, whether this leads to a lower market price or not depends on the number of firms if \( f(\overline{x}) > 0 \) and otherwise also depends on the tail behavior (e.g., in the normal distribution case this always decreases price).

### 4.5 Without full market coverage

This section aims to 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 focus on the case of independent products.

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 \( x_i \leq 0 \) but the mean of \( x^j_i \) is positive, such that some consumers do not purchase but the market is still active.

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

\[
q_i(p'_i) = \Pr[x^j_i - p'_i > \max_{k \neq j} \{0, x^k_i - p_i\}] = \int_{p'_i}^{\overline{x}_i} F(x - p'_i + p_i)^{n-1} dF_i(x) .
\]

One can check that the first-order condition for \( p_i \) to be the equilibrium price for product \( i \) is

\[
p_i = \frac{q_i(p_i)}{q'_i(p_i)} = \frac{[1 - F_i(p_i)^n]/n}{\frac{F_i(p_i)^{n-1}f_i(p_i)}{p_i} + \int_{p'_i}^{\overline{x}_i} f_i(x) dF_i(x)^{n-1}} .
\]

(We assume this is sufficient for defining the equilibrium price. This is the case, for example, if \( f_i \) is log-concave.) In equilibrium, a consumer will leave the market without purchasing product \( i \) with probability \( F_i(p_i)^n \) (i.e., when each product \( i \) has a valuation less than \( p_i \)). Given the symmetry of firms, the numerator in (19) 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 for all other firms’ product \( i \) are below \( p_i \) (which
occurs with probability $F_i(p_i)^{n-1}$, firm $j$ acts as a monopoly. Raising its price $p_i$ by $\varepsilon$ will exclude $\varepsilon f_i(p_i)$ consumers from the market. (ii) The second term is the same competition effect as before (up to the adjustment that a marginal consumer’s valuation must be greater than $p_i$).

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

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

(20)

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 first-order conditions. 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** When $f_i$ is log-concave, there is a unique equilibrium price $p_i \in (0, p_i^M)$ defined in (19), where $p_i^M$ is the monopoly price solving $p_i^M = [1 - F_i(p_i^M)]/f_i(p_i^M)$, and $p_i$ decreases with $n$. Similar results hold for $P/m$ defined in (20) if $g$ is log-concave (which is true if the joint density function $f$ is log-concave).

For a fixed $m < \infty$, if $n$ is large, the demand difference between the two numerators in (19) and (20) becomes negligible, and so is the exclusion effect difference in the denominators. Therefore, price comparison is again determined by the comparison of the tails. 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 $f_i(\pi_i) > 0$ for all $i$, bundling raises market prices when $n$ is greater than a certain threshold. (It is more difficult to extend the duopoly result.)

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.
An alternative way to introduce the exclusion effect of price is to consider elastic demand. Such an extension is presented in the online appendix, and the basic insights remain unchanged there.

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 $2 \times 2$ 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), 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 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 the valuation upper bound is finite. (Notice that in all these spatial models the consumer valuation upper bound must be finite.) 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 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.

\[^{35}\text{The consumers around the two polars should compare all firms, but they are ignored in the Economides’s analysis.}\]
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) shows 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. Bundling clearly reduces allocation efficiency and so revenue.) In the two-object case, Chakraborty (1999) further shows 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.\footnote{As 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 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.}

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
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.\textsuperscript{37} 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) have offered a model in this vein but in a different context. 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 \textit{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 competitive 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 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 ($\rho_1, \rho_2$) and a joint-purchase discount $\delta > 0$. If a consumer buys both products from the same firm, she pays $\rho_1 + \rho_2 - \delta$.\textsuperscript{38}

All the existing research on competitive mixed bundling focuses on the duopoly case. See, for example, Matutes and Regibeau (1992), Anderson and Leruth (1993), Reisinger (2004), Thanassoulis (2007), and Armstrong and Vickers (2010). Armstrong and Vickers (2010) consider the most general setup so far in the literature. They allow for the existence of an exogenous shopping cost, and also consider elastic demand and general nonlinear pricing schedules. Our paper is the first to consider more than two products.

\textsuperscript{37}Chen and Li (2015) study bundling in a multiproduct procurement. They focus on the duopoly case, but allow for correlated production costs in each firm and a not fully-covered market.

\textsuperscript{38}Even in the duopoly case it is hard to analyze mixed bundling pricing equilibrium with more than two products. The literature has made little analytical progress in that direction. (See a discussion of this issue in the appendix of Armstrong and Vickers, 2010.) One possible way to proceed is to consider simple pricing policies such as two-part tariffs, or bundle-size pricing as in Chu, Leslie, and Sorensen (2011).
firms, but in other aspects we focus on a simpler setup with unit demand and no exogenous shopping costs. As we will see below, solving the mixed bundling pricing game becomes much more challenging 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 and examine the impact of mixed bundling relative to separate sales.

5.1 Incentive to use mixed bundling

Following the existing research on mixed bundling, we examine when a firm, starting from separate sales, has a unilateral incentive to introduce a small joint-purchase discount. Recall that $y_i \equiv \max_{k \neq j}\{x^k_i\}$ denotes the match utility of the best product $i$ among all firm $j$’s competitors. The joint CDF of $(y_1, y_2)$ is $W(y_1, y_2) \equiv F(y_1^n, y_2^n)$, and let $w(y_1, y_2)$ be the associated joint density function. The marginal CDF of $y_i$ is $F_i(y) \equiv \int F(y + t_i \xi_i)dF(y)^{n-1}$, and let $h_i(y)$ be the associated marginal density function. It is easy to check that $H_i(0) = 1 - \frac{1}{n}$, and the separate sales price in (1) can be written as $p_i = [nh_i(0)]^{-1}$. The conditional CDF of $\xi_i$ is

\[H_i(\xi_i|\xi_{-i}) = \int_{-\infty}^{\xi_i} h_i(t_i|\xi_{-i})dt_i,\]

where $h_i(t_i|\xi_{-i}) = h_i(\xi_i) / h_i(\xi_{-i})$ is the conditional density function of $\xi_i$.

The following result reports sufficient conditions for each firm to have an individual incentive to use mixed bundling (when it is costless to implement). In that case separate sales cannot be an equilibrium outcome.

**Proposition 7** Starting from separate sales with prices defined in (1), each firm has a strict unilateral incentive to introduce mixed bundling if

\[n[1 - H(0, 0)] > H_1(0|0) + H_2(0|0), \quad (21)\]
(i) For a given distribution, (21) holds if \( n = 2 \), or if \( n \) is sufficiently large and
\[
\lim_{\xi_i \to -\infty} \frac{h_i(\xi_i)}{h_i(\xi_i) h_2(\xi_2)} > 0.
\]
(ii) For a given \( n \), (21) holds if \( x_1 \) and \( x_2 \) are independent, negatively dependent (in the sense that \( \Pr(x_i > a | x_{-i} > b) \) is decreasing in \( b \) for any \( a \)), or limitedly positively dependent (in the sense that \( \Pr(x_i > a | x_{-i} > b) = \Pr(x_i > a) \) for any \( a \) and \( b \), and \( \frac{d}{dz} H_i(0)[H^{-1}_i(z)] > -1 \) for \( z \in [1 - \frac{1}{n}, 1] \)).

**Proof.** Suppose firm \( j \) unilaterally deviates from separate sales and introduces a small joint-purchase discount \( \delta > 0 \) (but keeps its stand-alone prices \( p_1 \) and \( p_2 \) unchanged). Figure 5 below depicts how this small deviation affects consumer demand in the space of \((\xi_1, \xi_2)\), where \( \Omega_i, i = 1, 2 \), indicates consumers who buy only product \( i \) from firm \( j \) and \( \Omega_b \) indicates consumers who buy both products from firm \( j \).

![Figure 5: The impact of a small joint-purchase discount on demand](image)

The negative effect of the deviation is that firm \( j \) earns \( \delta \) less from the consumers who buy both products from it. In the regime of separate sales, the measure of those consumers is
\[
\Omega_b = 1 - H_1(0) - H_2(0) + H(0, 0) = \frac{2}{n} - 1 + H(0, 0),
\]
where we have used \( H_i(0) = 1 - \frac{1}{n} \). So the (first-order) loss from the deviation is \( \delta \Omega_b \).

The positive effect of the deviation is that more consumers buy both products from firm \( j \), i.e., the region \( \Omega_b \) expands as indicated on the graph. Those consumers on the two rectangle shaded areas switch from buying only one product to buying both from firm \( j \), and those on the small triangle shaded area switch from buying nothing to buying both products from firm \( j \).

Notice that the small triangle area is a second-order effect when \( \delta \) is small, so only the two rectangle areas matter. The measure of consumers on the vertical rectangle
area is \( \delta \int_0^\infty h(0, \xi_2)d\xi_2 \), and firm \( j \) now makes an extra profit \( p_1 - \delta \) from each of them. Similarly, the measure of consumers on the horizontal rectangle area is \( \delta \int_0^\infty h(\xi_1, 0)d\xi_1 \), and firm \( j \) now makes an extra profit \( p_2 - \delta \) from each of them. Thus, the (first-order) gain from the small deviation is

\[
p_1 \delta \int_0^\infty h(0, \xi_2)d\xi_2 + p_2 \delta \int_0^\infty h(\xi_1, 0)d\xi_1 = \frac{\delta}{n} [2 - H_1(0|0) - H_2(0|0)] .
\]

The equality is because

\[
p_1 \int_0^\infty h(0, \xi_2)d\xi_2 = \frac{1}{nh_1(0)} \times h_1(0) \int_0^\infty h_2(\xi_2|0)d\xi_2 = \frac{1}{n} [1 - H_2(0|0)] ,
\]

and similarly

\[
p_2 \int_0^\infty h(\xi_1, 0)d\xi_1 = \frac{1}{n} [1 - H_1(0|0)] .
\]

The small deviation is profitable if the gain is greater than the loss, which yields the condition (21).

(i) When \( n = 2 \), we have \( y_i = x_{k_i}^j, k \neq j \), and then \( H_1(0|0) = H_2(0|0) = 1/2 \). This is because conditional on \( \xi_i = 0 \) (i.e., \( x_{i}^j = x_{k_i}^j \), \( x_{j-i}^j \) and \( x_{k_i}^j \) should have the same distribution and so \( \xi_{-i} = x_{j-i}^j - x_{k_i}^j \) is symmetric around zero. Meanwhile, \( H(0, 0) \) must be strictly less than \( H_i(0) = 1/2 \). Then \( 2[1 - H(0, 0)] > 2[1 - H_i(0)] = 1 = H_1(0|0) + H_2(0|0) \). The proof for the large-\( n \) result is longer, and we present it in the appendix.

(ii) When the two products have independent valuations, \( H_i(0|0) = H_i(0) \) and \( H(0, 0) = H_1(0)H_2(0) \). Using \( H_i(0) = 1 - \frac{1}{n} \), one can easily verify that the gain is twice the loss for any \( n \). The proofs for the cases with dependent valuations are presented in the appendix.

As we did before, given other firms are selling their products separately, firm \( j \)'s problem can actually be rephrased into a monopoly one where a consumer’s net valuation for its product \( i \) is \( x_i - (y_i - p_i) \), where \( y_i - p_i \) is regarded as an outside option. Then our incentive results in (ii) are very closely related to the existing works on the profitability of bundling in a monopoly setting. (See, e.g., Long, 1984, McAfee, McMillan, and Whinston, 1989, and Chen and Riordan, 2013. The last paper offers the most general conditions so far in the literature by using a copula approach. Our proofs for the cases with dependent valuations closely follow their approach.) However, in our setting \( y_i \) is related to \( x_i \) in a particular way such that the optimal monopoly separate sales price for product \( i \) is \( p_i \) which also appears in the outside option. This additional structure leads to the result when \( n \) equals two (which has been derived in Armstrong and Vickers, 2010, in a Hotelling model), or when \( n \) is sufficiently large (which is new in the literature). (Notice that the tail behavior condition for the large-\( n \) result is not empty because it holds at least in the situations close to the independent case.)
5.2 Equilibrium prices

We now characterize a symmetric mixed-bundling equilibrium \((\rho_1, \rho_2, \delta)\), where \(\rho_i\) is the stand-alone price for product \(i\) and \(\delta\) is the joint-purchase discount.\(^{39}\) Suppose that all other firms use the equilibrium strategy and firm \(j\) unilaterally deviates and sets prices \((\rho_1', \rho_2', \delta')\). Then a consumer has the following purchase options:

(a) buy both products from firm \(j\), in which case her surplus is \(x_1 + x_2 - (\rho_1' + \rho_2' - \delta')\);
(b) buy product 1 from firm \(j\) but product 2 from elsewhere, in which case her surplus is \(x_1 + y_2 - \rho_1' - \rho_2\);
(c) buy product 2 from firm \(j\) but product 1 from elsewhere, in which case her surplus is \(y_1 + x_2 - \rho_1 - \rho_2'\);
(d) buy both products from some other firms, in which case her surplus is \(A - (\rho_1 + \rho_2 - \delta)\) (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 the consumer buys both products from some other firms, the situation is more complicated if \(n \geq 3\), because she may buy them from the same firm or two different firms.

(i) If \(y_1\) and \(y_2\) are from the same firm, the consumer will of course buy both products from that firm, in which case \(A = y_1 + y_2\). Conditional on \(y_1\) and \(y_2\), this event occurs with probability\(^{40}\)

\[
\kappa(y_1, y_2, n) = \frac{(n - 1)f(y_1, y_2)F(y_1, y_2)^{n-2}}{w(y_1, y_2)},
\]

where the denominator \(w(y_1, y_2) = \frac{\delta^2}{\partial y_1 \partial y_2} F(y_1, y_2)^{n-1}\) is the joint density function of \((y_1, y_2)\). If the two products at each firm have independent match utilities, one can check that \(\kappa(y_1, y_2, n) = 1/(n - 1)\).

(ii) With probability \(1 - \kappa(y_1, y_2, n)\), \(y_1\) and \(y_2\) are from two different firms. Then the consumer faces the trade-off between consuming better-matched products by two-stop shopping, in which case she gets surplus \(y_1 + y_2 - (\rho_1 + \rho_2)\), or enjoying the joint-purchase discount by one-stop shopping, in which case she gets surplus \(\max_{k \neq j} \{x_1^k + x_2^k\} - (\rho_1 + \rho_2 - \delta)\), where \(\max_{k \neq j} \{x_1^k + x_2^k\}\) is the match utility of the best bundle among all other firms. Hence, when \(y_1\) and \(y_2\) are from different firms, \(A = \max\{\max_{k \neq j} \{x_1^k + x_2^k\}, y_1 + y_2 - \delta\}\).

In sum, conditional on \(y_1\) and \(y_2\), we have

\[
A = \begin{cases} 
  y_1 + y_2 & \text{with prob. } \kappa(y_1, y_2, n) \\
  \max\{\max_{k \neq j} \{x_1^k + x_2^k\}, y_1 + y_2 - \delta\} & \text{with prob. } 1 - \kappa(y_1, y_2, n)
\end{cases}.
\] \,(22)

\(^{39}\)We restrict our attention to the case with \(\delta < \min\{\rho_1, \rho_2\}\). Otherwise, the bundle would be cheaper than at least one individual product.

\(^{40}\)The numerator is the probability that \(\max_{k \neq j} x_i^k = y_i\), \(i = 1, 2\), and \(y_1\) and \(y_2\) are realized in one of the \(n - 1\) firms. The denominator is the probability that \(\max_{k \neq j} x_i^k = y_i\), \(i = 1, 2\).
The simple case is when \( n = 2 \). Then \( y_1 \) and \( y_2 \) must be from the same firm and so \( A = y_1 + y_2 \) for sure. The problem can then be converted into a two-dimensional Hotelling model by using two “location” random variables \( \xi_1 = x_1 - y_1 \) and \( \xi_2 = x_2 - y_2 \). That is the model often used in the existing literature on competitive mixed bundling.

When \( n \geq 3 \), the situation is more complicated. We need to deal with one more random variable \( \max_{k \neq j} \{ x_1^k + x_2^k \} \) which is correlated with \( y_1 \) and \( y_2 \). The following result characterizes this correlation. Then the distribution of \( A \), conditional on \( y_1 \) and \( y_2 \), is fully specified.

**Lemma 5** When \( n \geq 3 \), the CDF of \( \max_{k \neq j} \{ x_1^k + x_2^k \} \), conditional on \( y_1 \), \( y_2 \), and they being from different firms, is

\[
L(z|y_1, y_2) = \frac{F_1(z - y_2|y_2) F_2(z - y_1|y_1)}{F_1(y_1|y_2) F_2(y_2|y_1)} \times \frac{1}{F(y_1, y_2)^{n-3}} \left( F(y_1, z - y_1) + \int_{z-y_1}^{y_2} \int_{z_1}^{x_2} f(x_1, x_2) dx_1 dx_2 \right)^{n-3}
\]

for \( z \in [\max\{y_1 + x_2, y_2 + x_1\}, y_1 + y_2) \), where \( F_i(y_i|y_{-i}) \) is the conditional CDF of \( y_i \).

**Proof.** For a given consumer, let \( I(y_i) \), \( i = 1, 2 \), be the identity of the firm where \( y_i \) is realized. The lower bound of \( \max_{k \neq j} \{ x_1^k + x_2^k \} \) is the fact that the lowest possible match utility of the bundle from firm \( I(y_i) \) is \( y_i + x_{-i} \). We now calculate the conditional probability of \( \max_{k \neq j} \{ x_1^k + x_2^k \} < z \). This event occurs if and only if all the following three conditions are satisfied: (i) \( y_1 + x_2^{I(y_1)} < z \), (ii) \( x_1^{I(y_2)} + y_2 < z \), and (iii) \( x_1^k + x_2^k < z \) for all \( k \neq j \), \( I(y_i) \), \( I(y_j) \). Given \( y_1 \) and \( y_2 \), condition (i) holds with probability \( F_2(z - y_1|y_1)/F_2(y_2|y_1) \), since the CDF of \( x_2^{I(y_1)} \) conditional on \( y_1 \) and \( x^{I(y_2)}_2 < y_2 \) is \( F_2(x_2|y_2)/F_2(y_2|y_1) \). Similarly, condition (ii) holds with probability \( F_1(z - y_2|y_2)/F_1(y_1|y_2) \). One can also check (with the help of a graph) that the probability that \( x_1^k + x_2^k < z \) holds for a firm other than \( j \), \( I(y_1) \) and \( I(y_2) \), is

\[
\frac{1}{F(y_1, y_2)} \left( F(y_1, z - y_1) + \int_{z-y_1}^{y_2} \int_{z_1}^{x_2} f(x_1, x_2) dx_1 dx_2 \right).
\]

(The term in the bracket is the unconditional probability that \( \{ x_1^k, x_2^k \} \) lies in the region where \( x_1^k < y_i \) and \( x_1^k + x_2^k < z \).) 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 (23). \( \blacksquare \)

Given a realization of \((y_1, y_2, A)\), Figure 6 below describes how a consumer chooses among the four purchase options. As before, \( \Omega_i \), \( i = 1, 2 \), indicates the region where the consumer buys only product \( i \) from firm \( j \), and \( \Omega_b \) indicates the region where the consumer buys both products from firm \( j \). Then integrating the area of \( \Omega_i \) over \((y_1, y_2, A)\) yields the demand function for firm \( j \)'s single product \( i \), and integrating the area of \( \Omega_b \) 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 1 is
\[ \Omega_1(\delta) \equiv \mathbb{E}[\int_{\mathcal{Z}_2}^{y_2-\delta} \int_{A-y_2+\delta}^{A-y_1} f(x_1, x_2) dx_1 dx_2] , \] (24)
and the equilibrium demand for the single product 2 is
\[ \Omega_2(\delta) \equiv \mathbb{E}[\int_{A-y_1+\delta}^{\mathcal{Z}_1} \int_{y_1-\delta}^{y_2} f(x_1, x_2) dx_1 dx_2] . \] (25)
(The expectations are taken over $(y_1, y_2, A)$.) Given full market coverage, the equilibrium demand depends only on the joint-purchase discount $\delta$ but not on the stand-alone prices $\rho_1$ and $\rho_2$. Let $\Omega_b(\delta)$ be the equilibrium demand for firm $j$’s bundle. Then
\[ \Omega_i(\delta) + \Omega_b(\delta) = \frac{1}{n} . \] (26)

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 (whose economic meanings are explained in the proof of the proposition below):
\[ \alpha_1 \equiv \mathbb{E}[\int_{A-y_1+\delta}^{\mathcal{Z}_2} f(y_1-\delta, x_2) dx_2] , \quad \alpha_2 \equiv \mathbb{E}[\int_{A-y_2+\delta}^{\mathcal{Z}_1} f(x_1, y_2-\delta) dx_1] ; \]
\[ \beta_1 \equiv \mathbb{E}[\int_{\mathcal{Z}_1}^{y_1-\delta} f(x_1, A-y_1+\delta) dx_1] , \quad \beta_2 \equiv \mathbb{E}[\int_{\mathcal{Z}_2}^{y_2-\delta} f(A-y_2+\delta, x_2) dx_2] ; \] (27)
\[ \gamma \equiv \mathbb{E}[\int_{y_1-\delta}^{A-y_2+\delta} f(x_1, A-x_1) dx_1] . \]
(All the expectations are taken over \((y_1, y_2, A)\). Notice that all \(\alpha_i, \beta_i\) and \(\gamma\) are functions of \(\delta\) and are independent of the stand-alone prices. Then the necessary conditions for \((\rho_1, \rho_2, \delta)\) to be the equilibrium prices are reported in the following result. (See the online appendix for the proof.)

**Proposition 8** If a symmetric (pure strategy) mixed-bundling equilibrium exists, the stand-alone prices \(\rho_1\) and \(\rho_2\) and the joint-purchase discount \(\delta\) must satisfy

\[
(\alpha_1 + \beta_2 + \gamma)\rho_1 + \gamma\rho_2 = \frac{1}{n} + (\alpha_1 + \gamma)\delta ,
\]

(28)

\[
(\alpha_2 + \beta_1 + \gamma)\rho_2 + \gamma\rho_1 = \frac{1}{n} + (\alpha_2 + \gamma)\delta ,
\]

(29)

and

\[
(\beta_2 - \alpha_1)\rho_1 + (\beta_1 - \alpha_2)\rho_2 = 2\Omega_1(\delta) - (\alpha_1 + \alpha_2)\delta .
\]

(30)

The first two conditions are linear in \(\rho_1\) and \(\rho_2\). From them we can solve \(\rho_1\) and \(\rho_2\) as functions of \(\delta\). Substituting them into the third condition yields an equation of \(\delta\). These equations are more complicated than they appear because all \(\alpha_i, \beta_i\) and \(\gamma\) are functions of \(\delta\).

**Discussion: equilibrium existence.** To prove the existence of a symmetric (pure-strategy) equilibrium,\(^{41}\) we need to show that (i) the system of necessary conditions (28)-(30) has a solution, 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 in the i.i.d. case 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.

To make more progress in characterizing the equilibrium prices, we focus on the i.i.d. case. Let \(F(x)\) and \(f(x)\) be the common CDF and density function. Let \(\rho\) be the common stand-alone price, and let \(\alpha = \alpha_i\) and \(\beta = \beta_i\). Then the above equilibrium conditions simplify to

\[
\rho = \frac{1/n + \delta(\alpha + \gamma)}{\alpha + \beta + 2\gamma} ,
\]

(31)

and

\[
(\beta - \alpha)\rho = \Omega_1(\delta) - \alpha\delta .
\]

(32)

Without causing confusion, let

\[
H(\xi) = \int F(y + \xi)dF(y)^{n-1} \quad \text{and} \quad h(\xi) = \int f(y + \xi)dF(y)^{n-1}
\]

\(^{41}\)The existence of a symmetric mixed-strategy equilibrium is guaranteed by Theorem 1 in Becker and Damianov (2006) if we can impose a finite upper bound on prices.
be the common CDF and density function of $\xi_i = x_i - y_i$, $i = 1, 2$. When $n = 2$, $h(\xi)$ is symmetric around zero, and so $h(-\xi) = h(\xi)$ and $H(-\xi) = 1 - H(\xi)$.

The expressions for $\alpha$, $\beta$, $\gamma$, and $\Omega_1(\delta)$ are still complicated in general. However, they are simple in the duopoly case, and they also have simple approximations when $\delta$ 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 polar cases.

In the duopoly case, $A = y_1 + y_2$ and by using the symmetry of $h$ one can check that $\alpha = \beta = h(\delta)[1 - H(\delta)]$ and $\Omega_1(\delta) = [1 - H(\delta)]^2$. Thus, (32) simplifies to

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

(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, (31) becomes

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

with $\gamma(\delta) = 2 \int_0^\delta h(t)^2 dt$. In the uniform distribution example, one can check that $\delta = 1/3$, $\rho \approx 0.57$ and the bundle price is $2\rho - \delta \approx 0.81$. In the normal distribution example, one can check that $\delta \approx 1.06$, $\rho \approx 1.85$ and the bundle price is $2\rho - \delta \approx 2.63$. 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 (31) and (32) has a solution with $\delta$ close to zero. For a small $\delta$ both sides of (32) have simple approximations, and an approximated $\delta$ can be solved. The following proposition provides the details. (See the online appendix for the proof.)

**Proposition 9** Consider the i.i.d. case and suppose $f$ is log-concave.

(i) When $n = 2$, the system of (33) and (34) has a solution with $\delta < \rho$, and the bundle price is lower than in the regime of separate sales (i.e., $2\rho - \delta \leq 2p$).

(ii) Suppose $\frac{f'(x)}{f(x)}$ is bounded and $\lim_{n \to \infty} p = 0$, where $p = [nh(0)]^{-1}$ is the separate-sales price in (1). When $n$ is large, the system of (31) and (32) 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”. Result (ii) also implies that when there are many firms in the market, mixed bundling tends to be pro-competitive relative to separate sales. This is very different from pure bundling. The reason is that with mixed bundling, firms can control the degree of bundling by adjusting the joint-purchase discount and this tends to become smaller when there are more firms. While this is not the case in the regime of pure bundling.

### 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 \( \pi(\rho_1, \rho_2, \delta) \) be the equilibrium industry profit. Then

\[
\pi(\rho_1, \rho_2, \delta) = \rho_1 + \rho_2 - n\delta\Omega_b(\delta) .
\]

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

\[
\pi(\rho_1, \rho_2, \delta) - \pi(p_1, p_2, 0) = (\rho_1 - p_1) + (\rho_2 - p_2) - n\delta\Omega_b(\delta) .
\]

(35)

Let \( v(\tilde{\rho}_1, \tilde{\rho}_2, \tilde{\delta}) \) be the consumer surplus when all firms are charging stand-alone prices \( (\tilde{\rho}_1, \tilde{\rho}_2) \) and offering a joint-purchase discount \( \tilde{\delta} \). Given full market coverage, an envelope argument implies that \( v_i(\tilde{\rho}_1, \tilde{\rho}_2, \tilde{\delta}) = -1, i = 1, 2 \), and \( v_3(\tilde{\rho}_1, \tilde{\rho}_2, \tilde{\delta}) = n\Omega_b(\tilde{\delta}) \), where the subscripts indicate partial derivatives. (This is because raising \( \tilde{\rho}_i \) by \( \varepsilon \) will make every consumer pay \( \varepsilon \) more, and raising the discount \( \tilde{\delta} \) by \( \varepsilon \) will save \( \varepsilon \) for every consumer who buy both products from the same firm.) Then relative to separate sales, the impact of mixed bundling on consumer surplus is

\[
v(\rho_1, \rho_2, \delta) - v(p_1, p_2, 0) = \int_{p_1}^{\rho_1} v_1(\tilde{\rho}_1, \rho_2, \delta) d\tilde{\rho}_1 + \int_{p_2}^{\rho_2} v_2(p_1, \tilde{\rho}_2, \delta) d\tilde{\rho}_2 + \int_0^\delta v_3(p_1, p_2, \tilde{\delta}) d\tilde{\delta} \\
= - (\rho_1 - p_1) - (\rho_2 - p_2) + n \int_0^\delta \Omega_b(\tilde{\delta}) d\tilde{\delta} .
\]

(36)

Hence, (35) and (36) 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 always harms total welfare since \( \Omega_b(\tilde{\delta}) \) is increasing in \( \tilde{\delta} \).)

Consider the i.i.d. case. If \( n = 2 \), we know that relative to separate sales the bundle is cheaper in the regime of mixed bundling, while the single products tend to

\[42\text{When there is a positive production cost } c \text{ for each product, we have } \delta \approx (p - c)/2, \text{i.e., the bundling discount is approximately equal to half of the single product’s markup.} \]
be more expensive. This often makes the impacts of mixed bundling on industry profit and consumer surplus ambiguous. For instance, in the uniform distribution example, industry profit falls from 1 to about 0.84 and consumer surplus rises from about 0.33 to about 0.44 when we move from separate sales to mixed bundling. However in the exponential distribution example, industry profit rises from 2 to about 2.07 and consumer surplus falls from 1 to about 0.78. Armstrong and Vickers (2010) have derived a sufficient condition under which mixed bundling benefits consumers and harms firms (see their Proposition 4). With our notation, the condition is \( \frac{d}{d\xi} H(\xi) \geq \frac{1}{4} \) for \( \xi \leq 0 \). The impacts become less ambiguous when \( n \) is large. Whenever the approximation results in Proposition 9 hold, all prices become lower in the regime of mixed bundling, and so relative to separate sales mixed bundling harms firms and benefits consumers.

6 Conclusion

This paper has offered a random utility model for studying competitive bundling in an oligopoly market. When there are more than two firms, this framework is more convenient to use than the spatial models in the existing literature. In the pure bundling part, we found that the number of firms can qualitatively matter for the impact of pure bundling relative to separate sales. Under certain conditions, the impacts of pure bundling on prices, profits, and consumer surplus are reversed when the number of firms exceeds some threshold (which can be small). This suggests that the welfare assessment of pure bundling based on a duopoly model can be incomplete. This is based on a general investigation of how consumer valuation dispersion affects price competition. 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 in the i.i.d. case 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.

Our study can be useful for some future research projects. First, there is a recent trend of unbundling in many markets (especially in online markets) with the emergence of some platforms. 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 affects market prices. (This issue is also relevant in the recent debate about whether US cable companies should be required to unbundle their TV packages. See, e.g., Crawford and Yurukoglu, 2012, for a nice empirical work in this direction.)

Second, in the pure bundling part it is assumed 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, it is possible that the bundle price is low enough such that some consumers buy multiple bundles in order 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 becomes 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 analyze this case.

Appendix

Proof of Proposition 1: (i) This is a special case of Proposition 6 which we will prove later.
(ii) Given $f(x) > 0$, we have $f(x) > g(x) = 0$, and so the result for $n > \hat{n}$ follows immediately from result (ii) in Lemma 2. To prove the cut-off result, define

$$\lambda(n) \equiv \int_0^1 [l_F(t) - l_G(t)]t^{n-2}dt.$$  \hspace{1cm} (37)

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).\footnote{Chakraborty (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.} Let us first introduce two concepts. A real function $K(x, y)$ of two
variables is said to be *totally positive of order* $r$ if for all $x_1 < \cdots < x_k$ and $y_1 < \cdots < y_k$ with $1 \leq k \leq r$, we have

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

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

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

where the supremum is extended over all sets $t_1 \leq \cdots \leq t_k$ ($t_i \in T$), $k$ is arbitrary but finite, and $S(x_1, \ldots, x_k)$ 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 $\mu$ is a sigma-finite regular measure defined on $Y$. Suppose $f$ is Borel-measurable and bounded, and the integral exists. Then if $K$ is totally positive of order $r$ and $S(f) \leq r - 1$, then

\[
S(\zeta) \leq S(f).
\]

Now consider $\lambda(n)$ defined in (37). Our assumption implies that $S[l_f(t) - l_y(t)] \leq 2$. 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 $\lambda(\hat{n}) < 0$ 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

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

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

\[
\begin{vmatrix}
t_1^{n_1} & t_2^{n_2} & t_3^{n_3} \\
t_1^{n_2} & t_2^{n_2} & t_3^{n_3} \\
t_1^{n_3} & t_2^{n_3} & t_3^{n_3}
\end{vmatrix} > 0.
\]
Dividing the $i$th row by $t_{i1}^{n_1}$ ($i = 1, 2, 3$) and then dividing the second column by $t_{12}^{n_2-n_1}$ and the third column by $t_{13}^{n_3-n_1}$, we can see that the determinant has the same sign as

$$
\begin{vmatrix}
1 & 1 & 1 \\
1 & r_{2}^{\delta_2} & r_{3}^{\delta_3} \\
1 & r_{2}^{\delta_2} & r_{3}^{\delta_3}
\end{vmatrix} = (r_{2}^{\delta_2} - 1)(r_{3}^{\delta_3} - 1) - (r_{2}^{\delta_3} - 1)(r_{3}^{\delta_2} - 1),
$$

where $\delta_j \equiv n_j - n_1$ and $r_j \equiv t_j/t_1$, $j = 2, 3$. Notice that $0 < r_2 < r_3$ and $r_2^2 < r_3^2$. 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 $z \neq 1$.)

**Proof of Proposition 2:** From the result (1.9) in Nolan (2015), we have $X^j = \sum_{i=1}^{m} x_i^j \sim S(\alpha, \beta, m^{\frac{1}{\alpha}}; 0; 1)$. Then result (a) in Proposition 1.17 in Nolan (2015) implies that

$$
X^j/m \sim \begin{cases} 
S(\alpha, \beta, m^{\frac{1}{\alpha}}; 0; 1) & \text{if } \alpha \neq 1 \\
S(1, \beta, 1, \frac{2}{\alpha} \beta \ln m; 1) & \text{if } \alpha = 1
\end{cases}.
$$

The same result implies that $X^j/m = m^{\frac{1}{\alpha}} x_i^j$ if $\alpha \neq 1$. Therefore, $X^j/m <_{\text{disp}} x_i^j$ if $\alpha > 1$, and $x_i^j <_{\text{disp}} X^j/m$ if $\alpha < 1$. Then from result (i) in Lemma 2 we can deduce that bundling reduces market prices (and so profits) if $\alpha > 1$, and the opposite is true if $\alpha < 1$. When $\alpha = 1$, $X^j/m$ and $x_i^j$ differ only in the location parameter. With full market coverage, this does not affect the equilibrium price. Hence, bundling has no impact on prices and profits when $\alpha = 1$.

**Proof of Proposition 3:** (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 $\lim_{x \to \infty} \frac{d}{dx} \left( \frac{1-F(x)}{f(x)} \right) = 0$, the equilibrium price (1) in the Perlof-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 $\{x_1, \ldots, x_n\}$. In other words,

$$
p \approx \mathbb{E} \left[ x(n) - x(n-1) \right],
$$

as $n \to \infty$.\(^{44}\) Then the per-product consumer surplus in the regime of separate sales is

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

\(^{44}\)This result implies that when the tail behavior condition holds, the Perlof-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 does not exists.
Since \( \lim_{x \to \pi} \frac{d}{dx} \left( \frac{1-F(x)}{f(x)} \right) = 0 \) implies \( \lim_{x \to \pi} \frac{d}{dx} \left( \frac{1-G(x)}{g(x)} \right) = 0 \), we can deduce that when \( n \) is large the per-product consumer surplus in the regime of pure bundling is
\[
E \left[ \frac{X(n)}{m} \right] - P \approx E \left[ \frac{X(n-1)}{m} \right].
\]
When \( n \) is large, it is clear that \( E \left[ \frac{X(n-1)}{m} \right] < E \left[ x_{(n-1)} \right] \) and so our result follows.

(ii): It suffices to show the threshold \( n^* \) exists when \( m \to \infty \). As we have explained in the main text, when \( m \to \infty \) bundling benefits consumers if and only if
\[
E \left[ \max_j \{ x_i^j \} \right] - \mu < p.
\]
Since the left-hand side of (38) increases with \( n \) while the right-hand side decreases with \( n \) under the log-concavity condition, we only need to prove two things: (a) condition (38) 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 (38) approaches \( \int_{x^*}^{\pi} F(x) dx \) as \( n \to \infty \). So it suffices to show
\[
\lim_{n \to \infty} p = \frac{1 - F(\pi)}{f(\pi)} < \int_{x^*}^{\pi} F(x) dx, \tag{39}
\]
where the equality is from (3). If \( \lim_{n \to \infty} p = 0 \) (e.g., if \( f(\pi) > 0 \)), this is clearly the case. But this is also true even if \( \lim_{n \to \infty} p > 0 \). Notice that log-concave \( f \) implies log-concave \( 1 - F \) (or decreasing \( (1 - F)/f \)). Then
\[
\int_{x}^{\pi} F(x) dx = \int_{x}^{\pi} \frac{1 - F(x)}{f(x)} \frac{F(x)}{1 - F(x)} dF(x)
\]
\[
> \frac{1 - F(\pi)}{f(\pi)} \int_{x}^{\pi} \frac{F(x)}{1 - F(x)} dF(x)
\]
\[
= \frac{1 - F(\pi)}{f(\pi)} \int_{0}^{1} \frac{t}{1 - t} dt.
\]
Since the integral term is infinity, condition (39) must hold.

We now prove condition (a). Notice that
\[
E \left[ \max_j \{ x_i^j \} \right] - \mu = \int (F(x) - F(x^*)) dx.
\]
(The equality is from integration by parts.) Then using (1) and the notation \( l(t) \equiv f(F^{-1}(t)) \), we can rewrite (38) as
\[
\int_{0}^{1} \frac{t - t^n}{l(t)} dt \int_{0}^{1} t^{n-2} l(t) dt < \frac{1}{n(n-1)}.
\]
When \( n = 2 \), this becomes
\[
\int_{0}^{1} \frac{t(1-t)}{l(t)} dt \int_{0}^{1} l(t) dt < \frac{1}{2}. \tag{40}
\]
To prove this inequality, we need the following technical result:45

\[45\]I am grateful to Tomás F. Móri in Budapest for helping me to prove this lemma.
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

\[
\int_{0}^{1} \frac{\varphi(t)}{r(t)} \, dt \leq \int_{0}^{1} \varphi(t) \left( \int_{0}^{1} \frac{1}{r_{\theta}(t)} \lambda(\theta) \, d\theta \right) \, dt = \int_{0}^{1} \left( \int_{0}^{1} \frac{\varphi(t)}{r_{\theta}(t)} \, dt \right) \lambda(\theta) \, d\theta \leq \sup_{1/\theta \leq 1} \int_{0}^{1} \frac{\varphi(t)}{r_{\theta}(t)} \, dt.
\]

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 \( \theta \), because its derivative

\[
\frac{1}{2} \int_{0}^{\theta} \frac{\varphi(t)}{t} \, dt - \frac{1}{2} \int_{\theta}^{1} \frac{\varphi(t)}{1-t} \, dt
\]

increases in \( \theta \). Hence, its maximum is attained at one of the endpoints of the domain \([0, 1]\). This completes the proof of the lemma. \( \blacksquare \)

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 (40) is no greater than \( 1/4 \).

**Consumer surplus comparison with normal distribution.** To prove (12), it suffices to establish the following result:
Lemma 8 Suppose firm j’s product i has match utility \( x_j^i \sim \mathcal{N}(0, \sigma^2) \). Let \( p_i \) be the separate sales price of product i as defined in (1). Then

\[
\mathbb{E} \left[ \max_j \{ x_j^i \} \right] = \frac{\sigma^2}{p_i}.
\]

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

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

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

\[
\mathbb{E} \left[ \max_j \{ x_j^i \} \right] = -\sigma^2 n \int_{-\infty}^{\infty} F(x)^{n-1} f'(x) dx
\]
\[
= \sigma^2 n(n-1) \int_{-\infty}^{\infty} F(x)^{n-2} f(x)^2 dx
\]
\[
= \frac{\sigma^2}{p}.
\]

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

Proof of Proposition 4: (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 \( mp - m\epsilon \), where \( \epsilon > 0 \) is small. Firm j’s (first-order) loss from this deviation is \( m\epsilon/2 \), since half of the consumers buy from firm j when \( \epsilon = 0 \) and now they pay \( m\epsilon \) less. On the other hand, the deviation increases the demand for firm j’s bundle to \( \Pr(X_j + m\epsilon > X_k) = \int G(x + \epsilon) dG(x) \). So the demand increases by \( \epsilon \int g(x)^2 dx \), and firm j’s (first-order) gain from the deviation is

\[
mp \times \epsilon \int g(x)^2 dx = \frac{mp}{P} \times \frac{m\epsilon}{2}.
\]

(The equality used (5) at \( n = 2 \).) Therefore, the deviation is profitable if \( mp > P \). Similarly if \( mp < P \), then raising the bundle price to \( mp + m\epsilon \) 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. Firm $j$'s loss from this deviation is $m\varepsilon/2$. On the other hand, the demand for firm $j$'s each product increases by $\varepsilon \int f(x)^2 dx$, and so firm $j$'s gain 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 $m \to \infty$. As we have explained in the main text, when $m \to \infty$ a firm has no unilateral incentive to bundle if and only if

\[
(1 - \frac{1}{n})p < \int [F(x) - F(x)^{n-1}] \, dx . \tag{41}
\]

We have known that (41) fails to hold for $n = 2$. On the other hand, we have $\lim_{n \to \infty} p < \int \frac{F(x)dx}{\tilde{x}}$ as shown in the proof of Proposition 3. Then (41) 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 (41) as

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

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

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

(The equality used $p_n = \lceil n(n-1) \int_0^1 l(t)l^{n-2}dt \rceil^{-1}$.) On the other hand, we have

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

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

\[
n^2(n^2 - 1) \int_0^1 l(t)l^{n-2}dt > n(n+1) \int_0^1 l(t)\kappa(t) dt
\]

46
\[ (n^2 - 1) \int_0^1 l(t)t^{n-2} dt > (n + 1)^2 \int_0^1 l(t)t^{n-1}(1-t) dt. \]

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

**Proof of Proposition 6.** Consider the duopoly case with \( j = 1, 2 \). Then the model can be converted into a multi-dimensional Hotelling model. Define

\[ \xi_i \equiv x_i^1 - x_i^2 = \theta \xi_0 + (1 - \theta) a_i \tilde{\xi}_i, \]

where \( \xi_0 \equiv x_0^1 - x_0^2 \) and \( \tilde{\xi}_i \equiv \tilde{x}_i^1 - \tilde{x}_i^2 \). Let \( h_i, h_0 \), and \( \tilde{h} \) be the density functions of \( \xi_i, \xi_0 \), and \( \tilde{\xi}_i \), respectively. Under our assumptions, all the density functions are log-concave and symmetric around zero. The separate sales price in (1) can be written as \( p_i = \{2h_i(0)\}^{-1} \). (If \( f_i \) is the density of \( x_i^j \), one can check that \( h_i(0) = \int f_i(x)^2 dx \).

Let \( \tilde{h} \) be the density function of

\[ \frac{1}{m} \sum_{i=1}^m \xi_i = \theta \xi_0 + \frac{1 - \theta}{m} \sum_{i=1}^m a_i \tilde{\xi}_i. \]

Then the per-product bundle price is \( P/m = [2\tilde{h}(0)]^{-1} \). Thus, \( P \leq \sum_{i=1}^m p_i \) if and only if

\[ \frac{1}{h(0)} \leq \frac{1}{m} \sum_{i=1}^m \frac{1}{h_i(0)} \] \( \quad (42) \)

Since \( \tilde{\xi}_i \) is i.i.d. across products and has a log-concave and symmetric density, Theorem 2.3 in Proschan (1965) implies that

\[ \frac{\sum_{i=1}^m a_i \tilde{\xi}_i}{\sum_{i=1}^m a_i} \text{ is more peaked than } \tilde{\xi}_i \Leftrightarrow \frac{1}{m} \sum_{i=1}^m a_i \tilde{\xi}_i \text{ is more peaked than } \tilde{a} \tilde{\xi}_i, \quad (43) \]

where \( \tilde{a} \equiv \frac{1}{m} \sum_{i=1}^m a_i \) is the average of all \( a_i \). (A random variable \( x_G \) is said to be more peaked than \( x_F \) about zero if \( \Pr(|x_G| \leq t) \geq \Pr(|x_F| \leq t) \) for any \( t \geq 0 \). The equivalence in (43) is from multiplying each of the first two random variables by \( \tilde{a} \).) In the i.i.d. setting with \( \theta = 0 \) and \( a_i = 1 \) for all \( i \), this already implies that \( \tilde{h}(0) \geq h_i(0) \), and so bundling reduces market prices.

Since \( \xi_0 \) is independent of \( \tilde{\xi}_i \) and both have a symmetric density function, (43) further implies that

\[ \frac{1}{m} \sum_{i=1}^m \xi_i = \theta \xi_0 + \frac{1 - \theta}{m} \sum_{i=1}^m a_i \tilde{\xi}_i \text{ is more peaked than } \theta \xi_0 + (1 - \theta) \tilde{a} \tilde{\xi}_i. \]

(See, e.g., Lemma 11.2 in Karlin, 1968.) Then

\[ \tilde{h}(0) \geq \eta(\tilde{a}), \]

where

\[ \eta(\tilde{a}) \equiv \frac{1}{(1 - \theta)\tilde{a}} \int \tilde{h} \left( \frac{\theta \xi}{(1 - \theta)\tilde{a}} \right) h_0(\xi) d\xi \]

47
is the peak of the density function of \( \theta \xi_0 + (1-\theta)\tilde{a} \tilde{\xi}_i \). The peak of the density function of \( \xi_i \) is \( h_i(0) = \eta(a_i) \). Therefore, a sufficient condition for (42) is

\[
\frac{1}{\eta(\tilde{a})} < \frac{1}{m} \sum_{i=1}^{m} \frac{1}{\eta(a_i)} . \tag{44}
\]

If the \( m \) products in each firm are symmetric in the sense that \( a_1 = \cdots = a_m \), or if they have independent valuations (i.e., \( \theta = 0 \)), then (44) holds trivially. More generally, (44) holds if \( [\eta(\cdot)]^{-1} \) is convex. Notice that

\[
\eta(a) = \frac{1}{(1-\theta)a} \int \tilde{h} \left( \frac{\theta \xi}{(1-\theta)a} \right) h_0(\xi) d\xi = \int \tilde{h}(\theta t) h_0((1-\theta)at) dt .
\]

(The second equality is from changing the integral variable from \( \xi \) to \( t = \frac{\xi}{(1-\theta)a} \))

Since both \( \tilde{h} \) and \( h_0 \) are log-concave, the integrand is log-concave in \( (a,t) \). Then the Prékopa Theorem implies that \( \eta(a) \) is log-concave, and so \( [\eta(a)]^{-1} \) must be convex.\(^{46}\)

This completes the proof.

**Discussion.** The key step in the proof of Proposition 6 is to compare the peakedness between \( \sum_{i=1}^{m} (x_i^1 - x_i^2)/m \) and \( x_i^1 - x_i^2 \), where \( \{x_i^1 - x_i^2\}_{i=1}^{m} \) is a sequence of random variables whose density functions are log-concave and symmetric around zero. Proschan (1965) has proved that if \( \{x_i^1 - x_i^2\}_{i=1}^{m} \) are further i.i.d., then the former is more peaked. That implies our duopoly result in the i.i.d. setting. However, according to my knowledge, only limited progress has been made in extending the Proschan theorem for non i.i.d. random variables. Chan et al. (1989) and Ma (1998) have respectively extended the result when the joint density function of \( \{x_i^1 - x_i^2\}_{i=1}^{m} \) is sign-invariant and Schur-concave,\(^{47}\) or when \( \{x_i^1 - x_i^2\}_{i=1}^{m} \) are independent and can be ranked by the likelihood ratio order. But these are strong conditions: the first condition basically requires all products to be symmetric and allows little dependence, and the second one is not much applicable when all the density functions of \( x_i^1 - x_i^2 \) are symmetric around zero. In addition, these conditions do not have simple primitive conditions based on \( f \). (Of course, both conditions are satisfied in our i.i.d. setting when \( f \) is log-concave.) Our conditions in Proposition 6 are not covered by either of them.

**Proof of Proposition 7:** We prove the other sufficient conditions for (21). We use the copula approach introduced in Chen and Riordan (2013). (A classic reference on copula is Nelson, 2006.) Let \( C(t_1, t_2) \) be the copula associated with the joint CDF \( H \) such that \( H(\xi_1, \xi_2) = C(H_1(\xi_1), H_2(\xi_2)) \). According to the Sklar’s Theorem, such a copula exists uniquely for a given joint CDF if its marginal distributions are

\(^{46}\)If a function \( f \) is log-concave, then \((f')^2 \geq f \times f''\). While for \( f^{-1} \) to be convex, we only need \( 2(f')^2 \geq f \times f'' \).

\(^{47}\)A density function \( f(z_1, \cdots, z_n) \) is sign-invariant if \( f(z_1, \cdots, z_n) = f(|z_1|, \cdots, |z_n|) \). This implies joint symmetry among the random variables \( \{z_i\}_{i=1}^{n} \). Joint symmetry requires the random variables to be uncorrelated if they have finite means.
continuous. Therefore, a joint CDF can be represented by its marginal CDF’s and a copula. A copula itself is a joint CDF on \([0,1]^2\) with uniform marginal distributions, and it captures the dependence structure of the original distribution. Let \(C_i(t_1, t_2)\) be the partial derivative with respect to \(t_i\). Let \(d(t) \equiv C(t, t)\) be the diagonal section of \(C\), and it is increasing and uniformly continuous on \([0,1]\). The following properties on copula are useful:

(a) \(C(t_1, 0) = C(0, t_2) = 0\);
(b) \(C(t_1, 1) = t_1\) and \(C(1, t_2) = t_2\);
(c) \(C_i(t_1, t_2)\) is the conditional distribution of \(t_{-i}\) given \(t_i\);
(d) \(\max\{0, 2t - 1\} \leq d(t) \leq t\).

We first claim that (21) is equivalent to

\[
1 - d(t) > (1 - t)d'(t) \text{ at } t = 1 - \frac{1}{n}. \tag{45}
\]

The definition of copula and \(H_i(0) = 1 - \frac{1}{n}\) imply that \(H(0, 0) = C(H_1(0), H_2(0)) = d(t)\) at \(t = 1 - \frac{1}{n}\). Using the fact

\[
h(\xi_1, \xi_2) = C_{12}(H_1(\xi_1), H_2(\xi_2))h_1(\xi_1)h_2(\xi_2) \tag{46}
\]

and property (a), one can check that \(H_1(0|0) = C_2(t, t)\) and \(H_2(0|0) = C_1(t, t)\) at \(t = 1 - \frac{1}{n}\). Then (21) can be written as \(n(1 - d(t)) > C_1(t, t) + C_2(t, t)\) at \(t = 1 - \frac{1}{n}\) which is equivalent to (45).

The large-\(n\) result. Let \(t = 1 - \varepsilon\) with \(\varepsilon \approx 0\). Then Taylor expansion, together with \(d(1) = 1\) and \(d'(1) = 2\) (both of which are from property (b)) imply that \(1 - d(t) \approx 2\varepsilon - \frac{1}{2}d''(1)\varepsilon^2\) and \((1 - t)d'(t) \approx 2\varepsilon - d''(1)\varepsilon^2\). The former is greater whenever \(d''(1) > 0\). Notice that \(d''(1) = C_{11}(1, 1) + 2C_{12}(1, 1) + C_{22}(1, 1) = 2C_{12}(1, 1)\) since \(C_{ii}(1, 1) = 0\) (which is again from property (b)). So \(d''(1) > 0\) if and only if \(C_{12}(1, 1) > 0\), which is equivalent to the condition stated in the proposition according to (46).

The negative dependence result. Since \(\Pr(x_i > a|x_{-i} > b)\) decreases in \(b\) for any \(a\), for any given realization of \((y_i, y_{-i})\) we have \(\Pr(x_i > a + y_i|x_{-i} > b + y_{-i})\) decreases in \(b\). Then \(\Pr(\xi_i > a|\xi_{-i} > b)\) decreases in \(b\) for any \(a\). (This is called “right tail decreasing” in Nelson, 2006.) Corollary 5.2.6. in Nelson (2006) then implies that for any \(t \in (0, 1)\) we have

\[
C_i(t, t) < \frac{t - C(t, t)}{1 - t}, \quad i = 1, 2.
\]

So \((1 - t)d'(t) < 2(t - d(t))\). Then a sufficient condition for (45) is

\[
1 - d(t) \geq 2(t - d(t)) \iff d(t) \geq 2t - 1.
\]

This is always true given property (d).

The positive dependence result. As in the proof of Proposition 3 in Chen and Riordan (2013), (45) can be rewritten as

\[
1 - 2t + d(t) + \int_t^1 (1 - z)C_{11}(z, t)dz + \int_t^1 (1 - z)C_{22}(t, z)dz > 0 \text{ at } t = 1 - \frac{1}{n}. \tag{47}
\]
(This can be verified by using integration by parts and property (b).) Given $\Pr(x_i > a | x_{-i} > b) \geq \Pr(x_i > a)$ (which is called “positive quadrant dependence” in Nelson, 2006), we have $F(x_1, x_2) \geq F_1(x_1)F_2(x_2)$. This implies that $H(\xi_1, \xi_2) \geq H_1(\xi_1)H_2(\xi_2)$ and so $d(t) \geq t^2$ for any $t$. Also notice that $C_1(z, t) = H_2(H_1^{-1}(t)|H_1^{-1}(z)) = H_2(0|H_1^{-1}(z))$ at $t = 1 - \frac{1}{n}$. Then our condition on the conditional distribution implies that $C_{11}(z, t) > -1$ for $z \geq t = 1 - \frac{1}{n}$. Similarly, $C_{22}(t, z) > -1$ for $z \geq t = 1 - \frac{1}{n}$. Then the left-hand side of (47) is strictly greater than $(1-t)^2 - 2 \int_t^1 (1-z)dz = 0$.

References


Online Appendix
[Not For Publication]

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

The details of the discussion of possible asymmetric equilibria in Section 4.3. We first explain why it is difficult to get analytical solution in a pricing game where firms adopt asymmetric bundling strategies. The reason is that when some firms bundle, other firms will treat their separate products as complements, which 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. Numerical analysis can be done, but no further analytical progress can be made.

We now explain why there are no asymmetric equilibria in the uniform distribution example with \( n = 3 \) and \( m = 2 \). The first possible asymmetric equilibrium is that one firm bundles and the other two do not. In this hypothetical equilibrium, we can numerically show that the bundling firm charges \( P \approx 0.513 \) and earns a profit about 0.176, and the other two firms charge a separate price \( p \approx 0.317 \) and each earns a profit about 0.208. But if the bundling firm unbundles and charges the same separate price as the other two firms, it will have a demand \( \frac{1}{3} \) from each product, and its profit will rise to \( \frac{1}{3} \times 0.317 \times 2 \approx 0.211 \).

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. (Notice that this argument does not depend on the uniform distribution and \( m = 2 \).)

The details of the example with dependent valuations in Section 4.4. (i)
When \( n = 2 \), the bundle price is given by

\[
\frac{1}{P_{\theta}/m} = 2 \int g_{\theta}(x)^2 dx,
\]

where \( g_{\theta}(x) = \theta f(x) + (1 - \theta)g(x) \). One can check that

\[
\frac{d}{d\theta} \int g_{\theta}(x)^2 dx = \int (f^2 - g^2) dx - (1 - 2\theta) \int (f - g)^2 dx. \quad (48)
\]
(We have suppressed the variable $x$ in each integral grand for convenience.) So $P_\theta$ increases in $\theta$ if and only if $(48)$ is negative. In the duopoly case with the log-concavity assumption, we have $P/m \leq p$, which is equivalent to $\int f^2 dx \leq \int g^2 dx$. So the first term in $(48)$ must be nonpositive. Then clearly $(48)$ is negative if $\theta \leq \frac{1}{2}$. Also notice that $(48)$ is increasing $\theta$, and it equals $2 \int (f - g) f dx$ at $\theta = 1$. Therefore, if $\int (f - g) f dx \leq 0$ (which is true, for example, for the uniform and the normal distribution), then $(48)$ must be negative and so $P_\theta$ increases in $\theta$.

(ii) Given $f(\pi) > 0$ and $g(\pi) = 0$, we have $g_\theta(\pi) = \theta f(\pi) > 0$ and it clearly increases in $\theta$. Following the logic in the proof of Lemma 2(ii) we have

$$\lim_{n \to \infty} n P_\theta = \frac{m}{g_\theta(\pi)}.$$ 

Therefore, $P_\theta$ decreases in $\theta$ when $n$ is sufficiently large.

**Proof of Lemma 4:** We only prove the results for $p_i$. (The same logic works for $P_\theta$.) Notice that $1 - F_i$ is log-concave if $f_i$ is log-concave. In the following, we suppress the subscript $i$ for notational simplicity.

When $p = 0$, it is clear that the left-hand side of (19) is less than the right-hand side. We can also show the opposite is true when $p = p^M$. By using the second highest order statistic as in the proof of Lemma 1, the right-hand side of (19) equals

$$\frac{1 - F(p)^n}{nf(p)^{n-1} f(p) + \int_0^x \frac{f(x)}{1-F(x)} dF_{n-1}(x)} < \frac{1 - F(p)^n}{nf(p)^{n-1} f(p) + \frac{f(p)}{1-F(p)} (1 - F_{n-1}(p)))} = \frac{1 - F(p)}{f(p)}.$$ 

(The inequality is because $f/(1 - F)$ is increasing, and the equality used $F_{n-1}(p) = F(p)^n + nF(p)^{n-1}(1 - F(p)).$) This, together with the fact that $p^M = \frac{1-F(p^M)}{f(p^M)}$, implies that (19) has a solution $p \in (0, p^M)$.

To show the uniqueness, we prove that the right-hand side of (19) decreases in $p$. One can verify that its derivative with respect to $p$ is negative if and only if

$$f'(p)(1 - F(p)^n) + nf(p) \left( F(p)^{n-1} f(p) + \int_0^x f(x) dF(x)^{n-1} \right) > 0.$$ 

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

$$n \int_0^\pi f(x) dF(x)^{n-1} > (1 - F(p)^n) \frac{f(p)}{1-F(p)} - n f(p) F(p)^{n-1}.$$ 

The left-hand side equals $\int_0^\pi \frac{f(x)}{1-F(x)} dF_{n-1}(x)$, and the right-hand side equals $\frac{f(p)}{1-F(p)} (1 - F_{n-1}(p)))$. Therefore, the inequality holds since $f/(1 - F)$ is increasing.

To prove the result that $p$ decreases in $n$, let us first rewrite (19) as

$$\frac{1}{p} = \frac{f(\pi) - \int_0^\pi f'(x) F(x)^{n-1} dx}{(1 - F(p)^n)/n} = \frac{n f(\pi)}{1-F(p)^n} - \int_0^\pi \frac{f'(x)}{f(x)} \frac{F(x)^n - F(p)^n}{1-F(p)^n}. \quad (49)$$
(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 (49) increases with \( n \). Therefore, the unique solution \( p \) must decrease with \( n \).

The details of the discussion about elastic consumer demand in Section 4.5. We extend our model with full market coverage by considering elastic consumer demand. Suppose each product is divisible, and the \( m \) products are independent. For notational simplicity, we focus on the i.i.d. case. If a consumer purchases \( \tau_i \) 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 \( \tau_i \) 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 buys 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 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 an 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

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

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

\[
-\frac{r'(s)}{r(s)} = n \int f(x) dF(x)^{n-1}.
\] (50)

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 (such that
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 single bundle price, and it has to specify a vector of prices \((p_1, \cdots, p_m)\). (In the unit demand model with bundling, offering a bundle price \(P\) is equivalent to offering a vector of single-product prices such that \(\sum_{i=1}^{m} p_i = P\).) If a consumer buys all products from this firm, her indirect utility is \(\sum_{i=1}^{m} v(p_i)\). Suppose instead a firm offers an indirect utility \(S\). Then the optimal price vector should solve the problem:

\[
\max_{p_i} \sum_{i=1}^{m} p_i \quad \text{subject to} \quad \sum_{i=1}^{m} v(p_i) = S
\]

Let us suppose this problem has a unique solution with \(p_i(S) = p_k(S)\). The optimal unit price for each product is then \(\frac{1}{m} S\).

We look for an equilibrium where each firm offers an indirect utility \(S\). Suppose firm \(j\) unilaterally deviates to \(S_0\). 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 \(mr(\frac{S}{m})Q(S')\). Then the first-order condition is

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

Therefore, if \(-\frac{r'(s)}{r(s)}\) is increasing in \(s\) (or if \(r(s)\) is log-concave), (50) and (51) are similar as the equilibrium price conditions (1) and (5) in the unit demand model. Then our price comparison results in Proposition 1 continues to hold.

**Proof of Proposition 8:** To derive the necessary conditions for \(\rho_1\), \(\rho_2\) and \(\delta\) to be the equilibrium prices, let us consider the following local deviations:

First, suppose firm \(j\) unilaterally raises its joint-purchase discount to \(\delta' = \delta + \varepsilon\) (where \(\varepsilon > 0\) is small) while keeps its stand-alone prices unchanged. Then conditional on \((y_1, y_2, A)\), Figure 7(a) below describes how this small deviation affects consumer choices: \(\Omega_b\) expands because now more consumers buy both products from firm \(j\). The marginal consumers who change their purchase behavior are distributed on the shaded area.

\[\text{In the case with a linear demand function } -v'(p) = 1 - p, \text{ one can check that } r(s) = \sqrt{2s} - 2s \text{ (which is concave) and } -\frac{r'(s)}{r(s)} \text{ increases from } -\infty \text{ to } \infty \text{ when } s \text{ varies from } 0 \text{ to } 1/2.\]
Figure 7(a): Price deviation and consumer choice I

Here

$$\tilde{\alpha}_1 = \int_{A-y_1+\delta}^{y_1-\delta} f(y_1-\delta, x_2)dx_2 \quad \text{and} \quad \tilde{\alpha}_2 = \int_{A-y_2+\delta}^{x_1} f(x_1, y_2-\delta)dx_1$$

are the densities of marginal consumers along the line segments $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ on the graph, respectively. And

$$\tilde{\gamma} = \int_{y_1-\delta}^{A-y_2+\delta} f(x_1, A-x_1)dx_1$$

is the density of marginal consumers along the diagonal line segment on the graph. Integrating them over $(y_1, y_2, A)$ yields the previously introduced notation: $E[\tilde{\alpha}_i] = \alpha_i$ and $E[\tilde{\gamma}] = \gamma$. For the marginal consumers on the vertical shaded area (which has a measure of $\varepsilon \tilde{\alpha}_1$), they switch from buying only product 2 to buying both products from firm $j$, and so firm $j$ makes $\rho_1 - \delta$ extra profit from each of them. Similarly, firm $j$ makes $\rho_2 - \delta$ extra profit from each of the marginal consumers on the horizontal shaded area (which has a measure of $\varepsilon \tilde{\alpha}_2$). For those marginal consumers on the diagonal shaded area (which has a measure of $\varepsilon \tilde{\gamma}$), they switch from buying both products from some other firms to buying both from firm $j$. So firm $j$ makes $\rho_1 + \rho_2 - \delta$ extra profit from each of them. The only negative effect of the deviation is that those consumers on $\Omega_b$ who were already purchasing both products at firm $j$ now pay $\varepsilon$ less. The sum of all these effects integrated over $(y_1, y_2, A)$ should be equal to zero in equilibrium. This yields the following first-order condition:

$$\alpha_1(\rho_1 - \delta) + \alpha_2(\rho_2 - \delta) + \gamma(\rho_1 + \rho_2 - \delta) = \Omega_b(\delta), \quad (52)$$

where $\alpha_i$ and $\gamma$ are defined in (27) and $\Omega_b(\delta)$ is defined in (26).

Second, suppose firm $j$ unilaterally raises its stand-alone price $\rho_1$ to $\rho'_1 = \rho_1 + \varepsilon$ and its joint-purchase discount to $\delta' = \delta + \varepsilon$ (such that its bundle price remains unchanged). Figure 7(b) below describes how this small deviation affects consumer choices: $\Omega_1$ shrinks because now fewer consumers buy a single product 1 from firm $j$. 

5
Here

\[ \tilde{\beta}_2 = \int_{\tilde{\alpha}_2}^{y_2-\delta} f(A - y_2 + \delta, x_2) dx_2 \]

is the density of marginal consumers along the line segment \( \tilde{\beta}_2 \) on the graph. Integrating it over \( (y_1, y_2, A) \) yields the previously introduced notation: \( \mathbb{E}[\tilde{\beta}_2] = \beta_2 \). For those marginal consumers on the horizontal shaded area (which has a measure of \( \varepsilon \tilde{\alpha}_2 \)), they switch from buying only product 1 to buying both products from firm \( j \). So firm \( j \) make \( \rho_2 - \delta \) extra profit from each of them. For those marginal consumers on the vertical shaded area (which has a measure of \( \varepsilon \tilde{\beta}_2 \)), they switch from buying product 1 to buying nothing from firm \( j \). So firm \( j \) loses \( \rho_1 \) from each of them. The direct revenue effect of the deviation is that firm \( j \) earns \( \varepsilon \) more from each consumer on \( \Omega_1 \). The sum of these effects integrated over \( (y_1, y_2, A) \) should be equal to zero in equilibrium. This yields the second first-order condition:

\[ \alpha_2(\rho_2 - \delta) + \Omega_1(\delta) = \beta_2 \rho_1, \]  

where \( \alpha_2 \) and \( \beta_2 \) are defined in (27) and \( \Omega_1(\delta) \) is defined in (24).

Third, suppose firm \( j \) slightly raises its stand-alone price \( \rho_2 \) to \( \rho_2' = \rho_1 + \varepsilon \) and its joint-purchase discount to \( \delta' = \delta + \varepsilon \) (such that its bundle price remains unchanged). Then \( \Omega_2 \) shrinks as described in Figure 7(c) below.
Figure 7(c): Price deviation and consumer choice III

Here

\[
\tilde{\beta}_1 = \int_{x_1}^{y_1 - \delta} f(x_1, A - y_1 + \delta) dx_1
\]

is the density of marginal consumers along the line segment \( \tilde{\beta}_1 \) on the graph. Integrating it over \((y_1, y_2, A)\) yields the previously introduced notation: \( \mathbb{E}[\tilde{\beta}_1] = \beta_1 \). A similar argument as before yields the third first-order condition:

\[
\alpha_1 (\rho_1 - \delta) + \Omega_2(\delta) = \beta_1 \rho_2 ,
\]

where \( \alpha_1 \) and \( \beta_1 \) are defined in (27) and \( \Omega_2(\delta) \) is defined in (25).

Using the fact \( \Omega_1(\delta) + \Omega_2(\delta) = \frac{1}{2} \) in (26), we can derive (28) from (52) and (53), and (29) from (52) and (54). Adding (53) to (54) and using \( \Omega_1(\delta) = \Omega_2(\delta) \) yields (30).

**Proof of Proposition 9:**

(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

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

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, (33) 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 (55) 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 \( 2\rho - \delta = \frac{1}{2(x_0 + \gamma)} \), and the bundle price in the regime of separate sales is \( 1/h(0) \). The former is lower if
\[
\alpha + \gamma = h(\delta)[1 - H(\delta)] + 2\int_0^\delta h(t)^2dt \geq \frac{1}{2}h(0).
\]
Notice that the equality holds at \( \delta = 0 \). So it suffices to show that the left-hand side is increasing in \( \delta \). 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 \( \alpha, \beta, \gamma \) and \( \Omega_1(\delta) \) when \( \delta \) is small.

**Lemma 9** For a given \( n \), if \( \delta \approx 0 \), we have
\[
\alpha \approx \frac{h(0)}{n} - \left( \frac{h'(0)}{n} + \frac{h(0)^2}{n - 1} \right) \delta,
\]
\[
\beta \approx \left( 1 - \frac{1}{n} \right) h(0) + \left( \frac{h'(0)}{n} - h(0)^2 \right) \delta,
\]
\[
\gamma \approx \frac{n h(0)^2}{n - 1} \delta,
\]
\[
\Omega_1(\delta) \approx \frac{1}{n} \left( 1 - \frac{1}{n} \right) - \frac{2h(0)}{n} \delta,
\]
where \( h(0) = \int f(x)dF(x)^{n-1} \) and \( h'(0) = \int f'(x)dF(x)^{n-1}. \)

**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 (22), we have
\[
\mathbb{E}[\psi(y_1, y_2, A)] = \frac{1}{n - 1} \mathbb{E}_{y_1, y_2}[\psi(y_1, y_2, y_1 + y_2)]
\]
\[
+ \frac{n - 2}{n - 1} \mathbb{E}_{y_1, y_2}[L(y_1 + y_2 - \delta)\psi(y_1, y_2, y_1 + y_2 - \delta) + \int_{y_1 + y_2 - \delta}^{y_1 + y_2} \psi(y_1, y_2, z)dL(z)],
\]
where \( L(z) \) is defined in (23). By integration by parts and using \( L(y_1 + y_2) = 1 \), we can simplify this to
\[
\mathbb{E}[\psi(y_1, y_2, A)] = \mathbb{E}_{y_1, y_2}[\psi(y_1, y_2, y_1 + y_2)] - \frac{n - 2}{n - 1} \mathbb{E}_{y_1, y_2} \int_{y_1 + y_2 - \delta}^{y_1 + y_2} \frac{\partial}{\partial z} \psi(y_1, y_2, z)dL(z).}
\]

\footnote{When the support of \( x_1 \) is finite and \( f(\overline{x}) > 0 \), the density of \( x_1 - y_1 \) has a kink at zero such that \( h'(0) \) is not well defined. However, one can check that \( \lim_{t \to 0^+} h'(t) = \int f'(x)dF(x)^{n-1} \) and \( \lim_{t \to 0^-} h'(t) = \int f'(x)dF(x)^{n-1} - (n - 1)f(\overline{x}). \) We use \( h'(0^-) \) in our approximations.}
Now let us derive the first-order approximation of $\alpha$. (For our purpose, we do not need the higher-order approximations.) According to the formula above, we have

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

(57)

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

and the expectations are taken over $y_1$ and $y_2$.

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

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

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 $\delta$ 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 (57) and discarding all higher order terms yields the approximation for $\alpha$ in (56). The other approximations can be derived similarly. \hfill \blacksquare

**Step 2:** When $n$ is large, the system of (31) and (32) has a solution with a small $\delta$.

**Lemma 10** Suppose $\frac{f'(x)}{f(x)}$ is bounded and $\lim_{n \to \infty} p = 0$, where $p = \frac{1}{nh(0)}$ is the separate sales price in (1). Then when $n$ is sufficiently large, the system of (31) and (32) has a solution with $\delta \in (0, \frac{1}{nh(0)})$.

**Proof.** Recall that (32) is

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

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 $\frac{\delta h(0)}{n h(0)}$ is uniformly bounded for any $n$.

---

3Suppose $\frac{f'(x)}{f(x)} < M$ for a constant $M < \infty$. Then $-M f(x) < f'(x) < M f(x)$, 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, $-M h(0) < h'(0) < M h(0)$ for any $n$, and so $\frac{h'(0)}{h(0)}$ is uniformly bounded.
We first show that $\chi_L(0) < \chi_R(0)$. At $\delta = 0$, it is easy to verify that $\alpha = \frac{1}{n}h(0)$, $\beta = (1 - \frac{1}{n})h(0)$, $\gamma = 0$ and $\Omega_1(0) = \frac{1}{n} \left( 1 - \frac{1}{n} \right)$. Then

$$\chi_L(0) = \frac{1}{n} \left( 1 - \frac{2}{n} \right) < \chi_R(0) = \frac{1}{n} \left( 1 - \frac{1}{n} \right).$$

Next, we show that $\chi_L(\delta) > \chi_R(\delta)$ at $\delta = 0$ when $n$ is sufficiently large. The condition $\lim_{n \to \infty} p = 0$ implies that $\delta = \frac{1}{nh(0)} \approx 0$ when $n$ is large. Replacing $\delta$ in (56) by $\frac{1}{nh(0)}$, we have

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

Similarly,

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

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

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 $\delta$ by $\frac{1}{nh(0)}$ and no further approximations have been made.)

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

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

Using the above approximations, we have

$$\alpha + \gamma \approx \frac{2h(0)}{n} - \frac{h'(0)}{n^2h(0)}$$

and

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

Then the left-hand side of (58) equals

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

$$\approx \left( \frac{1}{n} - \frac{1}{n^2} \right) h(0).$$

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

Using the approximations, we also have

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

\[ \alpha + \beta + 2 \gamma \approx \frac{h(0)}{n} - \frac{h'(0)}{n^2 h(0)} - \frac{h(0)}{n(n-1)} + \left(1 - \frac{2}{n}\right) h(0) + \frac{h'(0)}{n^2 h(0)} + \frac{2h(0)}{n-1} \]

\[ = \left(1 - \frac{1}{n}\right) h(0) + \frac{h(0)}{n-1} \left(2 - \frac{1}{n}\right) \]

\[ = \frac{nh(0)}{n-1} \]

Then the right-hand side of (58) equals

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

(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 (31) and (32) when \( n \) is large.**

Given the system has a solution with a small \( \delta \) when \( n \) is large, we can approximate each side of (32) around \( \delta \approx 0 \) by using (56) and discarding all higher order terms. Then one can solve

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

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

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