Competitive Bundling*

Jidong Zhou
School of Management
Yale University
October 2015
(First Version: February 2014)

Abstract

This paper proposes a model of competitive bundling with an arbitrary number of firms. In the regime of pure bundling, we find that relative to separate sales pure bundling tends to raise market prices, improves profits and harms consumers when the number of firms is above a threshold. This is in contrast to the findings in the duopoly case on which the existing literature often focuses. Our results also shed new light on how consumer valuation dispersion affects price competition in general. In the regime of mixed bundling, having more than two firms raises new challenges in solving the model. We derive the equilibrium 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 in both regimes.

Keywords: bundling, product compatibility, oligopoly, multiproduct pricing

JEL classification: D43, L13, L15

1 Introduction

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

*I am grateful for their comments to Simon Anderson, Mark Armstrong, Mariagiovanna Bacca, Heski Bar-Isaac, Luis Cabral, Jay Pil Choi, Ignacio Esponda, Xavier Gabaix, Phil Haile, Justin Johnson, Alessandro Lizzeri, Barry Nalebuff, Andrew Rhodes, Mike Riordan, Ran Spiegler, Gabor Virag, Glen Weyl and seminar participants in Cornell, Duke-UNC, HKUST, Michigan, NYU Stern, Richmond Fed, SHUFE, Toronto, WUSTL Olin, WZB Berlin, Yale SOM, the 13th Columbia/Duke/Northwestern/MIT IO Theory Conference, and the 13th NYU IO Day. The research assistance by Diego Daruich is greatly appreciated.
for CDs, newspapers, books, or cable TV (e.g., in the US), firms usually do not sell songs, articles, chapters, or TV channels separately. This is called pure bundling. Other examples include banking account packages, party services, buffets, repairing service tied with the main product, and even education programs can be regarded as an example of pure bundling. Sometimes firms offer both the package and individual products, but the package is offered at a discount relative to its components. Relevant examples include software suites, TV-internet-phone, season tickets, package tours, and value meals. This is called mixed bundling. In many cases, bundling occurs in markets where firms compete with each other.

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

The obvious motivation for bundling is economies of scale in production, selling or buying, or complementarity in consumption. For example, in the traditional market it is perhaps too costly to sell newspaper articles separately. There are other less obvious but important reasons for bundling. For instance, pure bundling can reduce consumer valuation heterogeneity and so facilitate firms extracting consumer surplus (Stigler, 1968). Mixed bundling can be a profitable price discrimination device by

---

1 The situation, however, is changing with the development of the online market. For instance, consumers nowadays can download single songs from iTune or Amazon. Some websites like www.CengageBrain.com sell e-chapters of textbooks, and single electronic articles in many academic journals are also available for purchase. More recently, some broadcasters such as HBO and CBS are providing service to consumers directly via Internet, and this will perhaps affect cable TV companies' bundling strategies.

2 There is actually some debate recently about whether college education should be unbundled. See, for example, the article “What if you could buy a college education a la carte, like buying dinner?” on ThinkProgress.org on August 24, 2015, and the article “The unbundling of higher education” on www.knewton.com/blog on February 26, 2014. (The idea of unbundling higher education is not entirely new. See, e.g., Wang, 1975.)

3 This is actually the leading interpretation adopted in the early works on competitive pure bundling. See, e.g., Matutes and Regibeau (1988).
offering purchase options to screen consumers (Adams and Yellen, 1976). Bundling can also be used as a leverage device by a multiproduct firm to deter the entry of potential competitors or induce the exit of existing competitors (Whinston, 1990, and Nalebuff, 2004).\(^4\)

The main anti-trust concern about bundling is that it may restrict market competition. One possible reason is, as suggested by the leverage theory, bundling can change the market structure and make it more concentrated. Another possible reason is that even if the market structure is given, bundling may relax competition and inflate market prices because it changes the pricing strategy space. One purpose of this paper is to revisit the second possibility in a more general setup than the existing literature. The economics literature has extensively studied bundling in a monopoly market.\(^5\) There is also some research on competitive bundling. However, the existing works on competitive bundling often focus on the case with two firms and each selling two products (see, e.g., Matutes and Regibeau, 1988, for pure bundling, and Matutes and Regibeau, 1992, and Armstrong and Vickers, 2010, for mixed bundling).\(^6\)

However, there are many markets where more than two firms compete with each other and adopt (pure or mixed) bundling strategies. (For example, the companies that offer the TV-internet-phone service in New York City include at least Verizon, AT&T, Time Warner and RCN.) The literature has not developed a general model which can be used to study both pure and mixed bundling with an arbitrary number of firms. This has limited our understanding of how the degree of market concentration might affect firms’ incentives to bundle and the impact of bundling on market performance relative to separate sales. This paper aims to provide such a framework for studying competitive bundling. We will show that considering an arbitrary number of firms can shed new light on firms’ incentives to bundle and how bundling affects firms and

\(^4\)See also Choi and Stefanadis (2001), and Carlton and Waldman (2002). Without changing the market structure, bundling by a multiproduct firm may also help segment the market and relax the price competition with its single product rivals (Carbajo, de Meza and Seidmann, 1990, and Chen, 1997).

\(^5\)For example, Schmalensee (1984) and Fang and Norman (2006) study the profitability of pure bundling relative to separate sales, and Adams and Yellen (1976), Long (1984), McAfee, McMillan, and Whinston (1989), and Chen and Riordan (2013) study the profitability of mixed bundling.

\(^6\)Two notable exceptions are Economides (1989), and Kim and Choi (2015). Both papers study pure bundling in the context of product compatibility when there are more than two firms in the market and each sells two products. We will discuss in detail the relationship with those two papers in section 4.6. Some recent empirical works on bundling also deal with the case with more than two firms. See, e.g., Crawford and Yurukoglu (2012) for bundling in the cable TV industry, and Ho, Ho, and Mortimer (2012) for bundling in the video rental industry. They focus on how the interaction between bundling and the vertical market structure might affect market performance. This is an interesting dimension ignored by the existing theory literature on bundling.
consumers. For example, in the pure bundling case, having more than two firms can reverse the impact of bundling relative to separate sales. This suggests that the insights we have learned from the existing duopoly models can be incomplete, and the number of firms qualitatively matters for the welfare assessment of bundling.

The existing papers on competitive bundling with two firms and two products use the two-dimensional Hotelling model where consumers are distributed on a square and firms are located at two opposite corners.\(^7\) With more than two multiproduct firms, it is no longer convenient to model product differentiation in a spatial framework.\(^8\) In this paper, we adopt the random utility framework in Perloff and Salop (1985) to model product differentiation. Specifically, a consumer’s valuation for a particular product is a random draw from some distribution, and its realization is independent across firms and consumers. This reflects, for example, the idea that firms sell products with different styles and consumers have idiosyncratic tastes. This framework is flexible enough to accommodate any number of firms and products, and in the case with two firms and two products it can be rephrased into the standard two-dimensional Hotelling model.

Section 2 presents the model and section 3 analyzes the benchmark case of separate sales. With separate sales, firms compete in each product market separately and each market is like a Perloff-Salop model. We study two comparative static questions. First, we show that a standard log-concavity condition (which ensures the existence of pure-strategy pricing equilibrium) guarantees that market prices decline with the number of firms. This is true even if we relax the usual assumption of full market coverage. Second, we study how a change of consumer valuation dispersion affects market price. We argue that when the number of firms is large, the tail behavior, instead of the peakedness, of the valuation distribution determines the equilibrium price. In particular, a less dispersed consumer valuation distribution (which is often interpreted as less product differentiation) can lead to higher market prices. Both results have their own interest in the literature on oligopolistic competition. The first result is not totally new, but our proof is simple. The second result provides the foundation for the main price comparison result in the pure bundling case, and as we will discuss more it is also useful for studying the impact on market price of any economic activities (such as information disclosure,\(^7\) One exception is Anderson and Leruth (1993). They use a random utility model (more precisely, a logit model) to study competitive mixed bundling in the duopoly case. However, different from the framework in this paper, the random utility component in their model is on the package level, instead of on the product level.\(^8\) Introducing differentiation at the product level is important for studying competitive bundling if firms have similar cost conditions. If there is no product differentiation, prices will settle at the marginal costs anyway and so there will be no meaningful scope for bundling. If differentiation is only at the firm level, consumers will one-stop shop even without bundling, which is not realistic in many markets and also makes the study of competitive bundling less interesting.)
advertising, and product design) which might change consumer valuation dispersion in the market.

Section 4 studies competitive pure bundling. We show that in the duopoly case pure bundling intensifies competition and leads to lower prices and profits compared to separate sales. This generalizes the result in the existing literature which considers two products only and also often assumes a particular consumer valuation distribution. Moreover, for consumers this positive price effect typically outweighs the negative match quality effect (which is caused by the loss of opportunities to mix and match), such that bundling tends to benefit consumers. However, under fairly general conditions, the results are reversed (i.e., pure bundling raises prices, benefits firms and harms consumers) when the number of firms is above some threshold (which can be small).

To understand these results, notice first that bundling makes the distribution of consumer valuation (in terms of the average per-product valuation) less dispersed. Compared to the single product density function, the density function of the per-product valuation for the bundle is more peaked but has thinner tails. Intuitively, this is because finding a well (or badly) matched bundle is harder than finding a well (or badly) matched component. On the other hand, a firm’s pricing decision hinges on the number of its marginal consumers who are indifferent between its product and the best product from its competitors. When there are many firms, a firm’s marginal consumers should have a high valuation for its product because their valuation for the best rival product is high. In other words, they tend to position on the right tail of the valuation density. Since bundling generates a thinner tail than separate sales, it often leads to fewer marginal consumers and so a less elastic demand. This induces firms to raise their prices. In contrast, when there are relatively few firms in the market, the average position of marginal consumer is close 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.

We also study firms’ incentives to bundle. When firms can choose between separate sales and pure bundling, it is always a Nash equilibrium that all firms bundle if consumers buy all products. (This is simply because if one firm unilaterally unbundles, the market situation does not change.) In the duopoly case, we further show that this is the unique equilibrium. However, when the number of firms is above some threshold, separate sales can be an equilibrium as well. In many examples separate sales is an-

\footnote{More precisely, the average position of the marginal consumers differs between the two regimes, and their relative distance also matters for the comparison. That is why there are also cases (especially when the valuation distribution is unbounded) where bundling leads to lower prices even if there are many firms in the market. In the analysis, we will derive conditions under which the price increase argument works.}
other equilibrium if and only if consumers prefer separate sales to pure bundling. In the end of Section 4, we extend the model in various directions which include asymmetric products and correlated valuations, a market without full market coverage, and elastic demand, and we also further discuss the relationship with a few closely related papers on competitive pure bundling.

Section 5 studies competitive mixed bundling. According to our knowledge, all the existing papers on competitive mixed bundling deal with the duopoly case. This paper is the first to consider the case with more than two firms. We first show that in our model starting from separate sales each firm has a strict incentive to introduce mixed bundling. That is, when mixed bundling is feasible and costless to implement, separate sales cannot be an equilibrium outcome. This is consistent with the findings in the monopoly and the duopoly case. However, solving the pricing game with mixed bundling is significantly harder when there are more than two firms. We are able to characterize the equilibrium conditions, and we can also show that under mild conditions the equilibrium prices have simple approximations when the number of firms is large. For example, when the number of firms is large and the production cost is zero, the joint-purchase discount will be approximately half of the stand-alone price (i.e., 50% off for the second product). In terms of the impacts of mixed bundling on profits and consumer surplus, they are usually ambiguous in the duopoly case. While with a large number of firms mixed bundling benefits consumers and harms firms under mild conditions. We conclude in Section 6, and all omitted proofs and details are presented in the Appendix.

2 The Model

Consider a market where each consumer needs $m \geq 2$ products. (They can be $m$ independent products, or $m$ components of a system, depending on the interpretation we will take below for bundling.) The measure of consumers is normalized to one. There are $n \geq 2$ firms, and each firm supplies all the $m$ products. The unit production cost of any product is normalized to zero (so we can regard the price below as the markup). Each product is horizontally differentiated across firms (e.g., each firm produces a different version of the product). We adopt the random utility framework in Perloff and Salop (1985) to model product differentiation. Let $x_{i,k}^j$ denote the match utility of firm $j$’s product $i$ for consumer $k$. We assume that $x_{i,k}^j$ is i.i.d. across consumers, which reflects, for instance, idiosyncratic consumer tastes. In the following we therefore suppress the subscript $k$. We consider a setting with symmetric firms and products: $x_i^j$ is distributed according to a common cdf $F$ with support $[\underline{x}, \overline{x}]$ (where $\underline{x} = -\infty$ and $\overline{x} = \infty$ are allowed), and it is realized independently across firms and products. Sup-
pose the corresponding pdf $f$ is continuous and bounded, and $x_i^j$ has a finite mean and variance. (In section 4.5, we will consider a more general setting where the $m$ products in each firm can be asymmetric and have correlated match utilities.)

We consider a discrete-choice framework where a consumer only buys one version of each product, i.e., the incremental utility from having more than one version of a product is zero.$^{10}$ Moreover, in the basic model we also assume that a consumer has unit demand for the desired version of a product. (Elastic demand will be considered in section 4.5.) If a consumer consumes $m$ products with match utilities $(x_1, \cdots, x_m)$ (which can be from different firms if consumers are not restricted to buy all products from a single firm) and makes a total payment $T$, she obtains surplus $\sum_{i=1}^m x_i - T.^{11}$

A firm’s pricing strategy space differs across the regimes we will investigate. In the benchmark regime of separate sales, each firm sells its products separately and so they choose price vectors $(p_j^1, \cdots, p_j^m)$, $j = 1, \cdots, n$. In the regime of pure bundling, each firm sells its products in a package only and they choose bundle prices $P_j$. In the regime of mixed bundling, each firm specifies prices $P_s^j$ for every possible subset $s$ of its $m$ products. (If $m = 2$, then firm $j$’s pricing strategy can be simply described as a pair of stand-alone prices $(p_1^j, p_2^j)$ together with a joint-purchase discount $\delta^j$.) The pricing strategy is the most general in the mixed bundling regime. (In the story of product incompatibility, however, the only relevant pricing strategies are separate sales and pure bundling.) In all the regimes the timing is that firms choose their prices simultaneously, and then consumers make their purchase decisions after observing all prices and match utilities. Since firms are ex ante symmetric, we will focus on symmetric pricing equilibrium.

As often assumed in the literature on oligopolistic competition, the market is fully covered in equilibrium. That is, each consumer buys all the $m$ products. This will be the case if consumers do not have outside options, or on top of the above match utilities $x_i^j$, consumers have a sufficiently high basic valuation for each product (or if the lower bound of match utility $x$ is high enough). Alternatively, we can consider a situation where the $m$ products are essential components of a system for which consumers have

---

$^{10}$This assumption is made in all the papers on competitive (pure or mixed) bundling. But clearly it is not innocuous. For example, reading another different newspaper article on the same story, or reading another chapter on the same topic in a different textbook usually improves utility. There are works on consumer demand which extend the usual discrete choice model by allowing consumers to buy multiple versions of a product (see, e.g., Gentzkow, 2007).

$^{11}$For simplicity, we have assumed away possible differentiation at the firm level. This can be included, for example, by assuming that a consumer’s valuation for firm $j$’s product $i$ is $u^j + x_i^j$, where $u^j$ is another random variable which is i.i.d. across firms and consumers but has the same realization for all products in a firm. This is a special case of the general setting with potentially correlated match utilities in section 4.5.1.
a high basic valuation. In the regimes of separate sales and pure bundling, we will relax this assumption in section 4.5 and argue that the basic insights concerning the impact of pure bundling remain qualitatively unchanged. However, in the regime of mixed bundling this assumption is important for tractability.

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

3 Separate Sales: Revisiting Perloff-Salop Model

This section studies the benchmark regime of separate sales. Since firms compete on each product separately, the market for each product is a Perloff-Salop model. Consider the market for product $i$, and let $p$ be the (symmetric) equilibrium price.\(^{12}\) Suppose firm $j$ deviates and charges $p'$, while other firms stick to the equilibrium price $p$. Then the demand for firm $j$’s product $i$ is

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

(In the following, whenever there is no confusion, we will suppress the integral limits $\bar{x}$ and $x$.) Notice that $F(x)^{n-1}$ is the cdf of the match utility of the best product $i$ among the $n-1$ competitors. So firm $j$ is as if competing with one firm which has match utility distribution $F(x)^{n-1}$ and charges the equilibrium price $p$. Firm $j$’s deviation profit from product $i$ is $p'q(p')$, and in equilibrium it should be maximized at $p' = p$. Notice that the equilibrium demand is $q(p) = 1/n$ since all firms equally share the market. Then

\(^{12}\)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 $f$ is log-concave.
one can check that the first-order condition for \( p \) to be the equilibrium price is

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

(1)

This first-order condition is also sufficient for defining the equilibrium price if \( f \) is log-concave (see Caplin and Nalebuff, 1991).\(^{13}\) A simple observation is that given the assumption of full market coverage, shifting the support of the match utility does not affect the equilibrium price.

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

**Price and the number of firms.** The first question is: how does the equilibrium price vary with the number of firms? The equilibrium condition (1) can be rewritten as

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

(2)

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

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

**Proof.** Let \( x_{(n-1)} \) be the second highest order statistic of \( \{x_1, \ldots, x_n\} \). Let \( F_{(n-1)} \) and \( f_{(n-1)} \) be its cdf and pdf, respectively. Using

\[
f_{(n-1)}(x) = n(n-1)(1 - F(x))F(x)^{n-2}f(x) \ .
\]

\(^{13}\)Caplin and Nalebuff (1991) derive a weaker sufficient condition which is \( f \) being \( \frac{1}{n+1} \)-concave. Our subsequent analysis needs this to be true for any \( n \), and when \( n \to \infty \) this condition becomes zero-concavity or equivalently log-concavity.
we can rewrite (1) as
\[ p = \frac{1}{p} = \int \frac{f(x)}{1 - F(x)} \, dF_{(n-1)}(x). \]  
(3)

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

Perloff and Salop (1985) studied the same comparative static question but did not find a simple answer. Anderson, de Palma, and Nesterov (1995) is the first paper that proved this monotonicity result (see their Proposition 1). They did it under the condition of \( f \) being log-concave (which is slightly stronger than \( 1 - F \) being log-concave). Our proof here is simpler than theirs. More recently, Quint (2014) shows that the log-concavity of \( f \) ensures that prices are strategic complements and the pricing game is supermodular in a general setting which allows for, for example, asymmetric firms and the existence of an outside option. Then this monotonicity result follows since introducing an additional firm is the same as treating that firm as an existing firm which drops its price from infinity to the new equilibrium price level. Though less general, our method is simple and it also offers a simple tail behavior condition for the markup to approach zero in the limit.

One special case is the exponential distribution which has a constant hazard rate \( f/(1 - F) \). In that case the price is independent of the number of firms. Nevertheless, the log-concavity of \( 1 - F \) is not a necessary condition. That is, even if \( 1 - F \) is not log-concave, it is still possible that price decreases with \( n \) (if the equilibrium price is determined by (1)). The tail behavior condition for \( \lim_{n \to \infty} p = 0 \) is satisfied if \( f(\pi) > 0 \). But it can be violated if \( f(\pi) = 0 \). One example is the extreme value distribution with \( F(x) = e^{-e^{-x}} \) (which generates the logit model). Both \( f \) and \( 1 - F \)

\[ 1 - F \text{ neither log-concave nor log-convex.} \]

\[ 1 - F \text{ is log-convex and the equilibrium price is determined by (1), then } p \text{ increases with } n. \]

---

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

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

16 One such example is the power distribution: \( F(x) = x^k \) with \( k \in (\frac{1}{n}, 1) \). In this example, \( 1 - F \) is neither log-concave nor log-convex. But one can check that the equilibrium price is \( p = \frac{n^{1-1/k}}{n^{n-1/k}} \) and it decreases in \( n \). If \( 1 - F \) is log-convex and the equilibrium price is determined by (1), then \( p \) increases with \( n \).
are strictly log-concave in this example. But one can check that \( p = \frac{n}{n-1} \), and it decreases with \( n \) and approaches to 1 in the limit.

**Price and the consumer valuation dispersion.** The second comparative static question is: if the distribution of consumer valuation becomes less “dispersed” from \( f \) to \( g \) as illustrated in Figure 1 below, how will the equilibrium price change? Intuitively, less dispersed consumer valuations mean less product differentiation across firms, and so this should intensify price competition and induce a lower market price. This must be the case if the density \( g \) degenerates at one point such that all products become homogeneous. However, as we will see below, except for this limit case \( g \) does not necessarily lead to a lower market price than \( f \).

![Figure 1: An example of less dispersed consumer valuation distribution](image)

In the literature on stochastic orders there are several possible ways to rank the dispersion of random variables. (The classic reference on this topic is Chapter 3 in Shaked and Shanthikumar, 2007.) One of them is called convex order. It is the most familiar one for economists because when two random variables have equal means it is equivalent to mean-preserving spread.\(^{17}\) For example, \( f \) and \( g \) in Figure 1 can be ranked in this order if they have equal means.\(^{18}\) However, as we will see below this order is usually not sufficient for a clear-cut price comparison result. Another one is called dispersive order. A random variable \( x_G \) is said to be smaller than \( x_F \) in the

\(^{17}\)Let \( x_F \) and \( x_G \) be two random variables, and let \( F \) and \( G \) be their cdf’s, respectively. Then \( x_G \) is smaller than \( x_F \) in the convex order (denoted as \( x_G \preceq_c x_F \)) 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 mean-preserving spread is established in Theorem 3.A.1. in Shaked and Shanthikumar (2007).

\(^{18}\)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 \(+, -, +\). (In general, two densities ranked by convex order can cross each other many times.)
dispersive order (denoted as \( x_G \preceq_{\text{disp}} x_F \)) 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 cdf’s 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 \( F \).) This order will ensure a clear-cut price comparison result as shown in the following result, but it is often a too strong condition for the applications we will discuss below.\(^{19}\)

**Lemma 2** Consider two Perloff-Salop markets with consumer valuation \( x_F \) and \( x_G \), respectively. Let \( F \) and \( G \) be their cdf’s, \( f \) and \( g \) be their bounded pdf’s, and \([\underline{x}_F, \overline{x}_F]\) and \([\underline{x}_G, \overline{x}_G]\) be their supports, respectively. Without loss of generality suppose \( \mathbb{E}[x_F] = \mathbb{E}[x_G] \). Let \( p_k, k = F,G, \) be the equilibrium price associated with \( x_k \). Suppose both \( f \) and \( g \) are log-concave such that the equilibrium prices are determined as in (1).

(i) If \( x_G \) is less dispersed than \( x_F \) according to the dispersive order, then \( p_G \leq p_F \) for any \( n \geq 2 \).

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

**Proof.** Changing the integral variable from \( x \) to \( t = F(x) \), we have

\[
\frac{1}{p_F} = n \int_{\underline{x}_F}^{\overline{x}_F} f(x) dF(x)^{n-1} \iff \frac{1}{p_F} = n \int_0^1 l_F(t) dt^{n-1},
\]

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

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

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

\[
\frac{dF^{-1}(t)}{dt} \geq \frac{dG^{-1}(t)}{dt} \iff f(F^{-1}(t)) \leq g(G^{-1}(t)).
\]

That is, \( l_F(t) \leq l_G(t) \). Therefore, \( p_G \leq p_F \) follows from (4).

(ii) \( l_F(t) - l_G(t) \) is bounded since both \( f \) and \( g \) are bounded. \( f(\overline{x}_F) > g(\overline{x}_G) \) implies \( l_F(1) - l_G(1) > 0 \). Then

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

\(^{19}\)When two random variables have equal means, dispersive order is a stronger condition than convex order. (See Theorem 3.B.16. in Shaked and Shanthikumar, 2007.) Ganuza and Penalva (2010) use these two orders to study information disclosure in auctions.
as 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. ■

Result (i) shows that if one density is less dispersed than the other in the dispersive order, the usual intuition works and we can predict that less dispersed consumer valuations lead to a lower market price. Perloff and Salop (1985) showed 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 just a special case of result (i) since $\theta x \leq_{\text{disp}} x$ for any random variable $x$ and constant $\theta \in (0, 1)$. 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. That is, the two random variables must be equal. This excludes many natural cases where one random variable is intuitively more dispersed than the other. For instance, the two distributions in Figure 1 cannot be ranked by dispersive order. (When $x_F$ and $x_G$ have equal means and their supports are intervals, $x_G \leq_{\text{disp}} x_F$ requires that the support of $x_G$ is a strict subset of the support of $x_F$, or both are infinite supports.) For this reason, as we will see this order often does not help in our bundling application.

Result (ii) shows that if we go beyond the dispersive order, even in natural cases such as the example in Figure 1 where one distribution is less dispersed than the other, the number of firms can matter for price comparison. In particular, when there are sufficiently many firms in the market, a less dispersed distribution can lead to a higher market price. When $g$ is less dispersed than $f$, we usually perceive that $g$ is more peaked but has thinner tails as in Figure 1. So $f(\pi_F) > g(\pi_G)$ is a reasonable case to consider. (From the proof of result (i) we know that this is not compatible with $x_G \leq_{\text{disp}} x_F$.) Since this result is crucial for understanding our price comparison result in next section, we explain its economic intuition in detail. Let us consider the example in Figure 1 where $f(1) > g(1)$. From (2) we already know that equilibrium price equals the ratio of equilibrium demand to the negative of equilibrium demand slope. Since equilibrium demand is always $1/n$ due to firm symmetry, only equilibrium demand slopes (or the densities of marginal consumers) matter for the price comparison. When $n$ is large, a given consumer’s valuation for the best product among a firm’s $n - 1$ competitors is close to the upper bound 1 almost for sure. Therefore, for that consumer to be this firm’s marginal consumer, her valuation for its product should also be close to 1. In other words, when $n$ is large, the position of a firm’s marginal consumers should be close to the upper bound no matter which density function applies. Since $f(1) > g(1)$, we deduce that each 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. This result suggests that when the number of firms
is large, the tail behavior, instead of the peakedness, of the densities matter for price comparison.\textsuperscript{20,21} The intuition here is given when $n$ is large, but as we will see in next section the threshold $\hat{n}$ can be actually small.

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

4 Pure Bundling

4.1 Equilibrium prices

Now consider the regime where all firms adopt the pure bundling strategy. Denote by $X^j = \sum_{i=1}^m x_i^j$ the match utility of firm $j$’s bundle. Then if firm $j$ charges a bundle price $P'$ while other firms charge the equilibrium price $P$, the demand for $j$’s bundle is

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

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

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

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

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

\textsuperscript{20}Gaiax et al. (2013) 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, markups are proportional to $[n f(F^{-1}(1 - 1/n))]^{-1}$. By noticing $\int_0^1 t dt^n - 1 = 1 - 1/n$, this can also be intuitively seen from the proof of Lemma 2 above.

\textsuperscript{21}Result (ii) 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 $n$, $[l_F(t) - l_G(t)](n - 1)n^{-2}$ is now 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 $l_F(t) - l_G(t)$ for $t$ close to $1$. 

14
Notice that the per-product bundle valuation $X^j/m$ is a mean-preserving contraction of $x^j$ and they have the same support.\footnote{This is true as long as the mean of $x^j$ exists. See, for example, p. 127 in Shaked and Shanthikumar (2007) for a formal proof.} So $g$ is less dispersed than $f$ as illustrated in Figure 1 above. In particular, $g(\bar{x}) = 0$ even if $f(\bar{x}) > 0$. This is simply because $X^j/m = \bar{x}$ only if $x^j = \bar{x}$ for all $i = 1, \cdots, m$, or intuitively this is because finding a well matched bundle is much harder than finding a well matched single product.\footnote{Formally, when $m = 2$ the pdf of $(x_1^j + x_2^j)/2$ is $g(x) = 2\int_{2x-\bar{x}}^{\infty} f(2x-t)dF(t)$ for $x \geq (\bar{x} + \bar{z})/2$, so $g(\bar{x}) = 0$. A similar argument works for $m \geq 3$.}

4.2 Comparing prices and profits

From (1) and (5), we can see that the comparison between separate sales and pure bundling is just a comparison between two Perloff-Salop models with different match utility distributions $F$ and $G$. According to result (i) in Lemma 2, bundling reduces market price if $X^j/m \leq \text{disp} x^j_i$. However, $X^j/m$ and $x^j_i$ usually cannot be ranked by the dispersive order. This is the case as we pointed out before if $x^j_i$ has a finite support. Given the further restriction here that $X^j/m$ and $x^j_i$ share the same support and mean (which means that $F$ and $G$ must cross each other at least once), this is also the case if $x^j_i$ has a semi-infinite support with a finite lower or upper bound. The only case where $X^j/m \leq \text{disp} x^j_i$ might hold is when the support of $x^j_i$ is the whole real line. One such example is the normal distribution as we will discuss more later. The following result indicates that bundling leads to lower prices in duopoly even if $X^j/m$ and $x^j_i$ are not ranked by the dispersive order, but if we go beyond duopoly it will be often the case that bundling raises market price if the number of firms is large enough.

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

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

(6)

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

**Proposition 1** Suppose $f$ is log-concave. (i) When $n = 2$, bundling reduces market prices and profits for any $m \geq 2$.

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

(iii) For a fixed $n$, $P$ increases in $m$ at a speed of $\sqrt{m}$ when $m$ is large and $\lim_{m \to \infty} P/m = 0$, so there exists $\hat{m}$ such that bundling reduces market prices and profits for $m > \hat{m}$.
Result (i) generalizes the observation in the existing literature 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 of this result 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 at the mean than \( f \). Thus, there are more marginal consumers (and so the demand is more elastic) in the case of \( g \).\(^{24}\) This induces firms to charge a lower price in the bundling regime. In Lemma 2, we did not find a simple stochastic order condition beyond dispersive order which ensures \( p_G \leq p_F \) when \( n = 2 \). This is an open question in general, but result (i) here suggests that this is the case if \( x_G \) is a sample mean of \( x_F \).

Result (iii) follows from the law of large numbers. Given the assumption that \( x_i^j \) has a finite mean \( \mu \), \( X^j/m \) converges to \( \mu \) as \( m \to \infty \). In other words, the per-product valuation for the bundle becomes homogeneous across both consumers and firms. Then \( P/m \) must converge to zero.\(^{25}\)

Result (ii) that pure bundling can soften price competition is perhaps more surprising. The result for large \( n \) follows from result (ii) in Lemma 2 since \( f(\bar{x}) > g(\bar{x}) = 0 \). Bundling makes the (per-product) valuation density have a thinner right tail, and when \( n \) is large the marginal consumers mainly locate on the right tail. In other words, bundling reduces the number of marginal consumers. This leads to a less elastic demand and a higher market price. The cut-off result is, however, harder to prove.\(^{26}\)

To illustrate the cut-off result, consider two examples which satisfy all the conditions in result (ii). In the uniform distribution example with \( f(x) = 1 \), \( P/m < p \) when \( n \leq 6 \) and \( P/m > p \) when \( n > 6 \). Figure 2(a) below describes how both prices vary with \( n \) (where the solid curve is \( p \) and the dashed one is \( P/m \)). In the example with an increasing density \( f(x) = 4x^3 \), as described in Figure 2(b) below \( P/m < p \) only when \( n = 2 \) and \( P/m > p \) whenever \( n > 2 \). These examples show that the threshold \( \hat{n} \) can be small. Another observation from these two examples is that the increase of price caused by bundling can be proportionally significant even if \( n \) is relatively large. For example, when \( n = 20 \) the bundling price is about 80% higher than the separate sales

\(^{24}\)Notice that we are dealing with \( P/m \) instead of the bundle price \( P \) directly. When we mention the measure of marginal consumers in the bundling case, it is actually \( m \) times the real measure of marginal consumers who are indifferent between a firm’s bundle and the best bundle among the competitors. This is because reducing \( P/m \) by \( \varepsilon \) is equivalent to reducing \( P \) by \( m\varepsilon \).

\(^{25}\)The result that \( P \) increases in \( m \) at a speed of \( \sqrt{m} \) when \( m \) is large is simply from the central limit theorem. A similar result has been shown by Nalebuff (2000) who considers a multi-dimensional Hotelling model with two firms, an arbitrary number of products, and consumers uniformly distributed inside the hypercube.

\(^{26}\)The proof does not rely on the bundling story. In fact, the same cut-off result holds in the context of Lemma 2 if \( p_G < p_F \) when \( n = 2 \) and \( f(\bar{x}_F) > g(\bar{x}_G) \).
price in the first example and 110% higher in the second one.\(^{27}\)

\[\begin{array}{cc}
(a) & (b)
\end{array}\]

\begin{align*}
\text{Figure 2: Price comparison with } m = 2
\end{align*}

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

**Example of normal distribution.** We already know that under the assumption of full market coverage shifting the support of the match utility distribution does not affect the equilibrium price. So let us normalize the mean to zero and suppose \(x^j_i \sim \mathcal{N}(0, \sigma^2)\). Then the separate sales price defined in (1) is

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

where \(\Phi\) and \(\phi\) are the cdf and pdf of the standard normal distribution \(\mathcal{N}(0, 1)\), respectively.\(^{28}\) The definition of \(X^j\) implies that \(X^j/m \sim \mathcal{N}(0, \sigma^2/m)\). Thus, \(X^j/m = x^j_i / \sqrt{m} \leq \text{disp} x^j_i\) (and the relationship is strict), so result (i) in Lemma 2 implies \(P/m < p\). That is, in this example bundling always reduces market prices (and so profits) regardless of \(n\) and \(m\).\(^{29}\) A more precise relationship between the two prices is available: Firm \(j\)’s demand in the bundling regime, when it

\(^{27}\) It can be the case that both \(p\) and \(P/m\) converge to zero as \(n \to \infty\) (e.g., in the uniform distribution example). But when \(f(\bar{x}) > 0\), they converge to zero at different speeds. More precisely, in that case we have \(\lim_{n \to \infty} \frac{p}{P/m} = \frac{g(\bar{x})}{f(\bar{x})} = 0\) given \(g(\bar{x}) = 0\).

\(^{28}\) One can check this result by using the fact that \(x^j_i = \sigma \tilde{x}^j_i\), where \(\tilde{x}^j_i\) has the standard normal distribution.

\(^{29}\) 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 \(n > 9\).
unilaterally deviates to price $P'$, is

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

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

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

In this normal distribution example, bundling also makes the right tail thinner (i.e., $g(x) < f(x)$ for relatively large $x$) and the (average) position of marginal consumers also moves to the right as $n$ increases. However, with unbounded support now the relative moving speed matters. The density tail is higher in the separate sales case and so it is more likely in that case to have a high valuation draw. This means that the position of marginal consumers moves to the right faster in the separate sales regime than in the bundling regime. Hence, for large $n$ even if $f(x) > g(x)$, it is possible that $f(\hat{x}_f) < g(\hat{x}_g)$ where $\hat{x}_f$ denotes the (average) position of marginal consumer in the separate sales regime and $\hat{x}_g$ denotes the (average) position of marginal consumers in the bundling regime. This cannot happen if the upper bound is finite and $f(\overline{x}) > 0 = g(\overline{x})$.

In that case, when $n$ is large both $\hat{x}_f$ and $\hat{x}_g$ will be close to $\overline{x}$ and so we must have $f(\hat{x}_f) > g(\hat{x}_g)$. Nevertheless, in the case with an infinite upper bound, even if both $\hat{x}_f$ and $\hat{x}_g$ move to infinity they can still be sufficiently far away from each other such that $f(\hat{x}_f) < g(\hat{x}_g)$ becomes possible.

The key feature in the normal distribution example is that the sample mean $X^j/m$ belongs to the same class of distributions as $x^j_i$, such that the dispersive order result in Lemma 2 can apply. In general, this is the property of stable distributions. Three notable examples of stable distributions are normal, Cauchy and Lévy. Suppose $x^j_i$ has a stable distribution with a stability parameter $\alpha \in (0, 2]$ and a location parameter $\delta = 0$. (Normal distribution has $\alpha = 2$, Cauchy distribution has $\alpha = 1$, and Lévy distribution has $\alpha = 1/2$.) Then it can be shown that $X^j/m = m^{\frac{\alpha}{\alpha - 1}} x^j_i$. (The above normal distribution result is a special case of this. The details are available upon request.) Therefore, $X^j/m \leq_{\text{disp}} x^j_i$ (and so bundling reduces market prices if the first-order condition defines the equilibrium price) if and only if $\alpha \geq 1$.\(^{31}\)

\(^{30}\)Let $x_1$ and $x_2$ be independent copies of a random variable $x$. Then $x$ is said to be stable if for any constants $a > 0$ and $b > 0$ the random variable $ax_1 + bx_2$ has the same distribution as $cx + d$ for some constants $c > 0$ and $d$. A good reference on stable distributions is Nolan (2015).

\(^{31}\)In the edge case with $\alpha = 1$ (e.g., the Cauchy distribution), bundling does not affect market prices. When $\alpha < 1$, bundling raises market prices for any $n$. This, however, does not contradict with the duopoly result in Proposition 1 because the distribution in this case is no longer log-concave. All the
Another example in which bundling always reduces market price is the exponential distribution. Although the dispersive order argument does not work any more, there is an alternative way to compare the prices. We already know that as the exponential distribution has a constant hazard rate, the separate sales price $p$ in this example is independent of $n$. While $X^j/m$, the per-product valuation for the bundle, has a strictly log-concave density, so $P/m$ strictly decreases in $n$. We know from result (i) in Proposition 1 that in this example bundling reduces market price in duopoly. Then we can deduce that it must be the case for any $n \geq 2$.

However, these examples do not suggest that $f(\tau) > 0$ (or a finite $\tau$) is necessary for the result that bundling can raise market price. For instance, consider the distribution with a log-concave density $f(x) = 2(1-x)$ and support $x \in [0, 1]$. In this example, $f(\tau) = 0$ but numerical simulations suggest a similar price comparison result as in Figure 2 (though the threshold $\hat{n}$ becomes bigger). There are also examples with $\tau = \infty$ and a log-concave density where bundling can raise market price. For instance, consider the generalized normal distribution with density $f(x) = \frac{\beta}{2(1/\beta)^{-|x|\beta}},$ where $\beta$ is the shape parameter and the support is the whole real line. (The density function is log-concave when $\beta > 1$.) This distribution becomes the standard normal when $\beta = 2$, and it converges to the uniform distribution on $[-1, 1]$ when $\beta \rightarrow \infty$. Suppose $\hat{n}$ is the threshold in the case of uniform distribution with support $[-1, 1]$. Then for any $n > \hat{n}$, there exists sufficiently large $\beta$ such that bundling raises market price.

Other possible applications. There are many other firm or consumer activities (such as information disclosure, advertising, product design, and consumer information acquisition) which can also change the dispersion of consumer valuations in a similar way as bundling. Our results in Lemma 2 and Proposition 1 can offer useful insights about how those activities might affect price competition.32

Here we briefly discuss a model of spurious product differentiation (in the spirit of Spiegler, 2006) which is isomorphic to our model. Suppose firms sell a homogenous product, and all consumers have an identical valuation $u$ for this product. Consumers discussion here is subject to the qualification that for some stable distributions the first-order condition may not be sufficient for defining the equilibrium price, or the integral in the first-order condition may not even exist.

32There are works that study firms’ individual incentive to disclose product information or conduct other activities that change the dispersion of consumer valuations. Lewis and Sappington (1994) and Johnson and Myatt (2006) consider a monopoly model and show that the firm will provide either full or zero information. Ivanov (2013) extends Johnson and Myatt (2006) and studies equilibrium information disclosure in an oligopoly model with price competition. He focuses on investigating when full information disclosure is an equilibrium and shows that this is the case when the number of firms is sufficiently large. But he does not study how the change of consumer valuation dispersion affects market prices and welfare.
do not know this true valuation, but they can independently observe $m$ noisy signals $(x_1, \cdots, x_m)$. Suppose each signal $x_i$ is an independent draw from a continuous distribution with mean $u$. Suppose consumers are non-Bayesian and they take the average face values of the signals, $\frac{1}{m} \sum_{i=1}^{m} x_i$, as the true valuation for the product. (Spiegler, 2006, considers the case with $m = 1$ and a binary signal realization.) Here the non-Bayesian behavior of consumers causes artificial product differentiation among firms such that the market price is above the cost. When $m$ increases, the distribution of consumer valuation will become more concentrated around the true valuation $u$. In particular, if $m \to \infty$ consumers will eventually find out the true valuation, and market competition will drive the markup down to zero. Therefore, in this model we can regard $m$ as an index of consumer rationality. Our price comparison results suggest that whether or not higher consumer rationality intensifies competition often depends on the number of firms in the market. In particular, beyond the duopoly case an improvement of consumer rationality can actually relax price competition and harm consumers.

4.3 Comparing consumer surplus and total welfare

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

However, the comparison of consumer surplus can be more complicated. If pure bundling increases market prices, it must harm consumers. From Proposition 1, we know this is the case when $f(\bar{p}) > 0$ and $n$ is sufficiently large. The trickier situation is when pure bundling lowers market prices (e.g., when $n = 2$, $m$ is large, or the distribution is normal or exponential). Then there is a trade-off between the negative match quality effect and the positive price effect. The main message in this section is that even if bundling intensifies price competition, the negative match quality effect often dominates such that bundling harms consumers when the number of firms is above a usually small threshold.

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

$$\mathbb{E} \left[ \max_j x_j^i \right] - p \quad \text{and} \quad \mathbb{E} \left[ \max_j \left\{ \frac{X_j}{m} \right\} \right] - \frac{P_m}{m}.$$ 

Then bundling benefits consumers if and only if

$$\mathbb{E} \left[ \max_j x_j^i \right] - \mathbb{E} \left[ \max_j \left\{ \frac{X_j}{m} \right\} \right] < p - \frac{P_m}{m}. \quad (9)$$
The left-hand side (which must be positive) reflects the match quality effect and the right-hand side is the price effect.

The following proposition reports the analytical results we have for consumer surplus comparison.

**Proposition 2** Suppose $f$ is log-concave. (i) For a fixed $m$, if $f$ is bounded and $f(x) > 0$, or if $\lim_{x \to -\infty} \frac{d}{dx} \left( \frac{1-F(x)}{f(x)} \right) = 0$, there exists $\hat{n}$ such that bundling harms consumers if $n > \hat{n}$.

(ii) 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 $m > \hat{m}(n)$.

The first condition in result (i) follows the price comparison result (ii) in Proposition 1, and it requires a finite $\bar{x}$. When $\bar{x} = \infty$ the second condition covers many often used distributions such as normal, exponential, extreme value, and logistic. Intuitively, this is because when $\bar{x} = \infty$ the difference between $\mathbb{E} \left[ \max_j \{ x^i_j \} \right]$ and $\mathbb{E} \left[ \max_j \left\{ \frac{X^i}{m} \right\} \right]$ can go to infinity as $n \to \infty$, while the price difference is always finite since both prices decrease with $n$ under the log-concavity condition.

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

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

It is clear that the match quality effect increases with $n$ (because with more firms bundling eliminates more mixing-and-matching opportunities), while the price effect decreases with $n$. What we show in the proof is that (10) holds for $n = 2$ but fails for a sufficiently large $n$. This leads to the cut-off result. 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.\footnote{We have examples where bundling harms consumers even in the duopoly case. One such example is when $m = 2$ and the distribution is exponential.} Figure 3 below describes how consumer surplus varies with $n$ in the uniform distribution case when $m = 2$ (where the solid curve is for separate sales, and the dashed one is for bundling). Again, the threshold is $n^* = 3$. The graph also suggests that the harm of bundling on consumers can be significant (e.g., when $n = 10$, bundling reduces consumer surplus by about...}
Figure 3: Consumer surplus comparison with uniform distribution and $m = 2$

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

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

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

In the Appendix, we show that

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

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

### 4.4 Incentive to bundle

This section studies firms’ incentive to bundle. Consider an extended game where firms can choose both bundling strategies and prices. But suppose that firms can choose only between separate sales and pure bundling. If firms can collectively choose their bundling strategy, then Proposition 1 indicates that the outcome often depends on the number of firms: they prefer separate sales when $n$ is small but pure bundling when $n$ is large.

---

34In this example, the harm of bundling will disappear eventually as $n \to \infty$. This is because $\lim_{n \to \infty} p = \lim_{n \to \infty} P/m = 0$ and $\lim_{n \to \infty} E[\max_j \{ x_i^j \}] = \lim_{n \to \infty} E[\max_j \{ X^j/m \}] = \bar{x}$. 

22
The outcome is different if firms choose their bundling strategies non-cooperatively. We mainly focus on the case where firms choose bundling strategies and prices simultaneously. This captures the situations where it is relatively easy to adjust the bundling strategy.

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

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

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

It is easy to understand that all firms bundling is a Nash equilibrium. This is simply because in our model if a firm unilaterally unbundles the market situation does not change for consumers. In the duopoly case, it can be further shown that separate sales cannot be an equilibrium outcome, and there are also no asymmetric equilibria where one firm bundles and the other does not.

When there are more than two firms, one may wonder whether separate sales can become another equilibrium outcome. 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 more inferior a firm’s bundle appears when it unilaterally bundles while its rivals do not. (This is different from the argument in the duopoly case where one firm bundling forces the other firm to bundle as well.) More formally, suppose that all other firms offer separate sales at price $p$, but firm $j$ unilaterally bundles. Denote by

$$y_i \equiv \max_{k \neq j} \{x^k_i\}$$  

(13)

the maximum match utility of product $i$ among firm $j$’s competitors. Then firm $j$ is as if competing with one firm that offers a bundle with match utility $Y \equiv \sum_{i=1}^m y_i$ and price $mp$. If firm $j$ charges the same bundle price $mp$, its demand will be $Pr(X^j > Y) \leq Pr(X^j > \max_{k \neq j} \{X^k\}) = 1/n$. The inequality is strict for $n \geq 3$. Thus, without further price adjustment it cannot be profitable for firm $j$ to unilaterally bundle.

Suppose now firm $j$ can also adjust 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 $y_i - p$ can be regarded as the outside option to product $i$.) If

---

35This argument depends on the assumption that consumers buy all products but for each product they do not buy more than one variant.
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 for the limit case with \( m \to \infty \).\(^{36}\) In the 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 if and only if

\[
(1 - \frac{1}{n})p < \mathbb{E}[y_i] - \mu = \int [F(x) - F(x)^{n-1}] \, dx. \quad (14)
\]

This is clearly not true for \( n = 2 \) (which is consistent with result (i)). In the proof, we show that this is true if and only if \( n \) is above a certain threshold.

The threshold \( \hat{n} \) in result (ii) is usually small. For instance, with a uniform distribution (14) becomes \( n^2 - 4n + 2 > 0 \) and so \( \hat{n} = 3 \). This is the same as the threshold \( n^* \) in the consumer surplus comparison result in Proposition 2. This means that in this uniform example with a large number of products, separate sales is another equilibrium outcome if and only if consumers prefer separate sales to pure bundling. In other words, with a proper equilibrium selection the market can work well for consumers. This is also true for the exponential and the normal distribution.\(^{37}\)

Asymmetric equilibrium. With more than two firms one may also wonder the possibility of asymmetric equilibrium where some firms bundle and the others do not. An analytical investigation into this problem is hard because the pricing equilibrium when firms adopt asymmetric bundling strategies does not have a simple characterization.\(^{38}\) However, numerical analysis can be done. Let us illustrate by a uniform example with \( n = 3 \) and \( m = 2 \). In this example, we can claim that there are no asymmetric equilibria.

The first possible asymmetric equilibrium is that one firm bundles and the other

---

\(^{36}\)This 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 (1965) result has been extended in various ways, but not when the density is asymmetric.) The analysis in Fang and Norman (2006), however, relies on that result. That is why we cannot apply their results to our model.

\(^{37}\)However, there exist examples (e.g., \( f(x) = 2(1 - x) \)) where \( \hat{n} \neq n^* \).

\(^{38}\)The reason is that the firms which bundle their products make each other firm treats their products as complements. This complicates the demand calculation. To understand this, let us suppose firm 1 bundles while other firms do not. When firm \( k \neq 1 \) lowers its product \( i \)'s price, some consumers will stop buying firm 1’s bundle and switch to buying all products from other firms. This will increase the demand for firm \( k \)'s all products. (The details on the demand calculation and the first-order conditions are available upon request, but no further analytical progress can be made.)
two do not. In this hypothetical equilibrium, the bundling firm charges $P \approx 0.513$ and earns a profit about 0.176, and the other two firms charge a separate price $p \approx 0.317$ and each earns a profit about 0.208. But if the bundling firm unbundles and charges the same separate price as the other two firms, it will have a demand $\frac{1}{3}$ and its profit will rise to about 0.211. The second possible asymmetric equilibrium is that two firms bundle and the third one does not. This hypothetical equilibrium is like all firm are bundling. Then each bundling firm charges a bundle price $P = 0.5$, the third firm charges $p_1 + p_2 = 0.5$, and each firm has market share $\frac{1}{3}$. But if one bundling firm unbundles and offers the same separate prices as the third firm, as we already knew the remaining bundling firm will have a demand less than $\frac{1}{3}$, and this implies that the deviation firm will have a demand greater than $\frac{1}{3}$. This improves its profit.

**Sequential choices.** Suppose now that firms make their bundling choices first, and then engage in price competition after observing the bundling outcome. First of all, as before all firms bundling is a Nash equilibrium outcome. The issue of whether separate sales is also an equilibrium is more complicated. In the duopoly case, if $f$ is log-concave, then bundling leads to a lower price and profit as we have shown in Proposition 1. Then no firm has a unilateral incentive to deviate from separate sales. That is, separate sales is a Nash equilibrium outcome as well.

When there are more than two firms, due to the complication of the pricing game when firms adopt asymmetric bundling strategies, no general results are available. In the uniform example with $n = 3$ and $m = 2$, we can numerically show that separate sales is another equilibrium, but there are no asymmetric equilibria. If all three firms sell their products separately, each firm charges a price $p = \frac{1}{3}$ and each firm’s profit is $\frac{2}{5} \approx 0.222$. Now if one firm, say, firm 1 deviates and bundles, then in the asymmetric pricing game, firm 1 charges a bundle price $P \approx 0.513$ and the other two firms charge a separate price $p \approx 0.317$ for each product. Firm 1’s profit drops to about 0.176 (and each other firm’s profit drops to about 0.208). So no firm wants to bundle unilaterally. This also implies that one firm bundling and the other two not is not an equilibrium, because all firms benefit if the bundling firm unbundles. The last possibility is that two firms bundle and the other does not. That situation is like all firms bundling, and each firm’s profit is about 0.167. But if one bundling firm unbundles, its profit will rise to 0.208. Hence, there are no asymmetric equilibria.

### 4.5 Discussions

#### 4.5.1 Asymmetric products and correlated distributions

We now consider a more general setting where the $m$ products within each firm are potentially asymmetric and their match utilities are potentially correlated. Let $x^j =
$(x^j_1, \ldots, x^j_m)$ be a consumer’s valuations for the $m$ products at firm $j$. Suppose $x^j$ is still i.i.d. across firms and consumers, and it is distributed according to a common joint cdf $F(x_1, \ldots, x_m)$ with support $S \subset \mathbb{R}^m$ and a continuous joint pdf $f(x_1, \ldots, x_m)$. Let $F_i$ and $f_i$, $i = 1, \ldots, m$, be the marginal cdf and pdf of $x^j_i$, and let $[x_i, \pi_i]$ be its support. Let $G$ and $g$ be the cdf and pdf of $X^j/m$, the average per-product match utility of the bundle, and let $[x, \pi]$ be its support. All $f_i$ and $g$ are log-concave if the joint pdf $f$ is log-concave.

In the regime of separate sales, since competition is still separate across products, the equilibrium price for product $i$ is determined the same as in formula (1) as long as we use the corresponding marginal distribution:

$$\frac{1}{P_i} = n \int_{x_i}^{\pi_i} f_i(x) dF_i(x)^{n-1}.$$ 

In the regime of pure bundling, the average per-product bundle price is still determined as in (5):

$$\frac{1}{P/m} = \int_x^{\pi} g(x) dG(x)^{n-1}.$$ 

The limit result that $\lim_{m \to \infty} P/m = 0$ still holds as long as $X^j/m$ converges to a deterministic value as $m \to \infty$. So for a fixed $n$, bundling lowers market price when $m$ is sufficiently large. Under similar conditions as before, we also have the result that for a fixed $m$ bundling raises market prices when $n$ is sufficiently large.

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

(i) if $f_i(\pi_i) > 0$, there exists $\hat{n}_i$ such that $P/m > p_i$ for $n > \hat{n}_i$;

(ii) if $f_i(\pi_i) > 0$ for all $i = 1, \ldots, m$, there exists $\hat{n}$ such that $P > \sum_{i=1}^m p_i$ for $n > \hat{n}$.

**Proof.** Our conditions imply that $g(\pi) = 0$ (e.g., see the proof of Proposition 1 in Armstrong, 1996). Then the results immediately follow from Lemma 2.

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

---

39 More formally, $g(\pi) = \lim_{\epsilon \to 0} \frac{1 - G(\pi - \epsilon)}{\epsilon}$, and our conditions ensure $1 - G(\pi - \epsilon) = o(\epsilon)$. Among the conditions, strict convexity of $S$ excludes the possibility that the plane of $X^j/m = \pi$ coincides with some part of $S$’s boundary, and $S$ being of full dimension excludes the possibility that $x^j_i$, $i = 1, \ldots, m$, are perfectly correlated.
What is less clear is whether bundling still lowers market price in the duopoly case. Following the logic in the proof of result (i) in Proposition 1, let \( h_i \) be the pdf of \( x_i^1 - x_i^2 \). (Notice that \( x_i^1 - x_i^2 \) is symmetric around zero, and \( h_i \) is log-concave if \( f_i \) is log-concave.) Then the separate sales price for product \( i \) is \( p_i = \frac{1}{2h_i(0)} \). Let \( \tilde{h} \) be the pdf of \( \frac{1}{m} \sum_{i=1}^m (x_i^1 - x_i^2) \). Then the per-product bundle price is \( P = \frac{1}{2\tilde{h}(0)} \), and \( P < \sum_{i=1}^m p_i \) if and only if

\[
\frac{1}{h(0)} < \frac{1}{m} \sum_{i=1}^m \frac{1}{h_i(0)} .
\]

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

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

Unfortunately, we are unable to find simple primitive conditions on the joint pdf \( f \) such that (15) or (16) is satisfied.\(^{40}\) The following example shows that in the duopoly case bundling can still lower market price even if products are substantially asymmetric.

**Example of asymmetric products.** Consider a two-product example where \( x_1^1 \) is uniformly distributed on \([0, 1]\) and \( x_2^j \) is independent of \( x_1^j \) and is uniformly distributed on \([0, b]\) with \( b > 1 \). In duopoly, \( x_1^1 - x_1^2 \) and \( x_2^1 - x_2^2 \) have pdf’s

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

respectively. Then \( h_1(0) = 1 \) and \( h_2(0) = \frac{1}{b} \). One can also check \( \tilde{h}(0) = \frac{2(3b-1)}{3b^2} \).

The sufficient condition (16) holds only for \( b \) less than about 2.46, but the “iff” condition (15) holds for any \( b > 1 \). Hence, bundling lowers prices in duopoly. However, according to Proposition 4 bundling will raise prices in this example when \( n \) is sufficiently large.

### 4.5.2 Without full market coverage

We now return to the symmetric and independent case but relax the assumption of full market coverage. (The analysis below, though, can be extended to the general setup as

\(^{40}\)If the \( m \) products are symmetric and the joint pdf of \( \{x_i^1 - x_i^2\}_{i=1}^m \) is Schur-concave, or if the \( m \) products at each firm have independent match utilities and any two random variables \( x_i^1 - x_i^2 \) and \( x_k^1 - x_k^2 \) can be ranked according to the likelihood ratio order, then there are extensions of the Proschan (1956) result which can help prove (16). But we do not have simple primitive conditions for either of the conditions to hold. (In the symmetric and independent case, both conditions are satisfied if \( f \) is log-concave.)
in the previous section.) A subtle issue here is whether the $m$ products are independent products or perfect complements. This will affect the analysis in the separate sales benchmark. If the $m$ products are independent products, consumers decide whether to buy each product separately. While if the $m$ products are perfect complements, then whether to buy a certain product also depends on how well matched other products are. (With full market coverage, this distinction does not matter.) In the following, we consider the case of independent products for simplicity.

Suppose now $x^j_i$ denotes the whole valuation for firm $j$’s product $i$, and a consumer will buy a product or bundle only if the best offer in the market provides a positive surplus. To make the case interesting, let us suppose $\mu \leq 0$ but $\mu > 0$, such that some consumers do not buy but the whole market is still active.

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

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

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

$$p = \frac{\frac{q(p)}{|q'(p)|}}{[1 - F(p)^n]/n} = \frac{[1 - F(p)^n]/n}{F(p)^{n-1} f(p) + \int_{p'}^{\pi} f(x) dF(x)^{n-1}}. \tag{17}$$

(If $f$ is log-concave, this is also sufficient for defining the equilibrium price.) In equilibrium, a consumer will leave the market without purchasing product $i$ with probability $F(p)^n$ (i.e., when each product $i$ has a valuation less than $p$). Given the symmetry of firms, the numerator in (17) is the demand for each firm’s product $i$ in equilibrium. The denominator is the negative of the demand slope but it now has two parts: (i) The first term is the standard market exclusion effect: when the valuations of all other firms’ product $i$ are below $p$ (which occurs with probability $F(p)^{n-1}$), firm $j$ will play as a monopoly. Then raising its price $p$ by $\varepsilon$ will exclude $\varepsilon f(p)$ consumers from the market. (ii) The second term is the same competition effect as in the case with full market coverage (up to the adjustment that a marginal consumer’s valuation must be greater than the price $p$).

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

$$\frac{P}{m} = \frac{[1 - G(P/m)^n]/n}{G(P/m)^{n-1} g(P/m) + \int_{F/m}^{\pi} g(x) dG(x)^{n-1}}, \tag{18}$$

where $G$ and $g$ are the cdf and pdf of $X^j/m$ as before.
Unlike the case with full market coverage, now the equilibrium price in each regime is implicitly determined in the first-order conditions. The following result reports the condition for each first-order condition to have a unique solution.

**Lemma 4** Suppose $f$ is log-concave. There is a unique equilibrium price $p \in (0, p_M)$ defined in (17), where $p_M$ is the monopoly price solving $p_M = [1 - F(p_M)]/f(p_M)$, and $p$ decreases with $n$. Similar results hold for $P/m$ defined in (18).

Similar results concerning the price comparison hold when $m$ is large or $n$ is large. For a fixed $n < \infty$, we still have $\lim_{m \to \infty} P/m = 0$ since $X^j/m$ converges as $m \to \infty$. For a fixed $m < \infty$, if $n$ is large the demand size difference between the two numerators in (17) and (18) is negligible, and so is the exclusion effect difference in the denominators. Therefore, price comparison is again determined by the comparison of $f(\pi)$ and $g(\pi)$. Intuitively, when there are a large number of varieties in the market, almost every consumer can find something she likes and so almost no consumers will leave the market without purchasing anything. Then the situation will be close to the case with full market coverage. Thus we have a similar result that when $f(x)$ is bounded and $f(\pi) > 0$, bundling raises market prices when $n$ is greater than a certain threshold. The duopoly case without full market coverage is hard to deal with, and we have not been able to prove a result similar to result (i) in Proposition 1.

Figure 4 below reports the impacts of pure bundling on market prices, profits, consumer surplus and total welfare in the uniform example with $F(x) = x$. (The solid curves are for separate sales, and the dashed ones are for pure bundling.) They are qualitatively similar as those in the case with full market coverage. In particular, the total welfare result is similar even if we introduce the exclusion effect of price.
Figure 4: The impact of pure bundling—uniform distribution example without full market coverage

An alternative way to introduce the exclusion effect of price is to consider elastic demand. In the Appendix, we extend the baseline model by considering elastic consumer demand and show that the basic insights from the baseline model still hold.

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 studied the $2 \times 2$ case in a two-dimensional Hotelling model where consumers are uniformly distributed on a square. They showed that bundling induces lower market prices and profits, and it also benefits consumers if the market is fully covered. Our analysis in the duopoly case has generalized their results by considering more products and more general distributions.

Hurkens, Jeon, and Menicucci (2013) extended Matutes and Regibeau (1988) to the case with two asymmetric firms where one firm produces higher-quality products than the other. 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. Firm asymmetry is clearly relevant in the real market. Under certain technical conditions they showed that when the quality difference is sufficiently large, pure bundling raises both firms’ profits. (They did not state a formal result on price comparisons.) 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 shifts the position of marginal consumers to the tail, and bundling makes the tail thinner and so leads to fewer marginal consumers and less elastic demand. 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) proposed 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 showed that for a general transportation cost function pure bundling reduces market prices relative to separate sales. His spatial model features local competition: each firm is directly competing with its two neighbor rivals only (no matter in the separate sales regime or the bundling regime),\footnote{In fact, for the consumers around the two polars they should compare all firms, but the analysis in Economides (1989) ignores these consumers.} and they are always symmetric to each other no matter how many firms in total are present in the market. While 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.

A recent independent work by Kim and Choi (2015) proposed an alternative $n \times 2$ spatial model. They assume that consumers are uniformly distributed on the surface of a torus, and firms are symmetrically located on the same surface. (There are many possible ways to symmetrically locate firms.) For a quadratic transportation cost function, they showed that when there are four or more firms in the market, there exists at least one symmetric location of firms under which making the products incompatible across firms raises market price and profit. This is consistent with our comparison result. Compared to Economides (1989), a key difference in their model, according to 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 locations of firms. In this sense, their model is closer to our random utility model.

Whether a spatial model or a random utility model is more appropriate for study competitive pure bundling may depend on the context. The random utility model, however, seems more flexible and easier to use. The analysis of the pure bundling regime is much simpler in our framework than in the 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 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) investigated firms’ individual incentives to bundle.

Pure bundling in auctions. Our study of competitive pure bundling is also related to the literature on multi-object auctions with bundling. Consider a private-value second-price auction where bidders’ valuations for each object are i.i.d. across objects and bidders. (The auction format does not matter given the revenue equivalence result.) Palfrey (1983) showed that if there are only two bidders, selling all the objects in a package is more profitable for the seller 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 valuation, so the revenue must be higher in the bundling case. With many bidders, however, the second highest valuation is close to the maximum valuation, so the revenue must be higher with separate sales. (In this limit bidders’ information rent disappears and so only the allocation efficiency matters for the revenue, and bundling clearly reduces allocation efficiency.) In the two-object case, Chakraborty (1999) further showed a cut-off result under certain regularity conditions.

Notice that the seller in an auction is like consumers in our price competition model, and the agents on the other side of the market are competing for them. 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 based on the order statistic of the second highest valuation. While in our model as we have seen in (3) the equilibrium price is related to the second order statistic in a more complex way. This makes our analysis less straightforward than in the auction case. This also partially explains, for example, why revenue comparison in auctions with two or many bidders is very general but its counterpart, i.e., consumer surplus comparison, in our model is less so.\footnote{As we pointed out before in the end of section 4.2, product information disclosure can affect consumer valuation distribution in a similar way as pure bundling. The same analogy applies in the auction case. Among the works which study information disclosure in auctions, Board (2009) and Gauza and Penalva (2010) are the two most related papers. Both of them showed that whether or not the auctioneer benefits from publicly providing more information to bidders depends on the number of bidders as in the above auction papers. Board (2009) pointed out the connection between bundling and information disclosure.}

\textit{Pure bundling with homogenous products and heterogeneous costs.} In the literature on competitive bundling, to make bundling a meaningful issue to study the standard approach 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 the identical valuation for all products but the unit production cost is i.i.d. across firms and products, and each firm has private cost realizations. Since competition now occurs on the informed side, the situation is like a first-price procure-
ment 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. 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) is a related paper in this direction. 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. From the ex ante perspective, this setting generates the same outcome as in the private cost setting. In the $n \times 2$ case they show a similar profit comparison result.

5 Mixed Bundling

In this section, we study competitive mixed bundling. All the existing research on competitive mixed bundling focuses on the duopoly case. See, for example, Matutes and Regibeau (1992), Anderson and Leruth (1993), Thanassoulis (2007), and Armstrong and Vickers (2010). Armstrong and Vickers (2010) considered the most general setup so far in the literature: they allow for asymmetric products and correlated valuations, and they also consider elastic demand and a general nonlinear pricing schedule. This paper is the first to consider the case with more than two firms. As we will see below, solving the mixed bundling pricing game becomes much harder when we go beyond the duopoly case. One contribution of this paper is to propose a way to solve the problem, and when the number of firms is large we also show that the equilibrium prices have relatively simple approximations.

For tractability, we focus on the baseline model with full market coverage. We also assume that each firm supplies two products only and they have i.i.d. match utilities.\footnote{It is not difficult to extend the analysis below to the case with asymmetric products and correlated valuations. Dealing with the case with more than two products is more difficult because of the resulted complication of the pricing strategy space.} If a firm adopts a mixed bundling strategy, it offers a pair of stand-alone prices $(\rho_1, \rho_2)$ and a joint-purchase discount $\delta > 0$. (Then if a consumer buys both products from this firm, she pays $\rho_1 + \rho_2 - \delta$.) In the following, we will first argue that starting from separate sales with price $p$ defined in (1), each firm has a unilateral incentive to introduce mixed bundling. We will then characterize the symmetric pricing equilibrium.
with mixed bundling and examine the impacts of mixed bundling relative to separate sales.

5.1 Incentive to use mixed bundling

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

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

\[
y_i \equiv \max_{k \neq j} \{x_i^k\}
\]

denotes the match utility of the best product \( i \) among all other firms. For a given realization of \((y_1, y_2)\) from other firms, Figure 5 below depicts how the small deviation affects consumer demand, where \( \Omega_i, i = 1, 2 \), indicates consumers who buy only product \( i \) from firm \( j \) and \( \Omega_0 \) indicates consumers who buy both products from firm \( j \). (We have suppressed the superscripts in firm \( j \)'s match utilities.)

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

When firm \( j \) introduces the small discount \( \delta \), the region of \( \Omega_0 \) expands but both \( \Omega_1 \) and \( \Omega_2 \) shrink accordingly. The shaded area indicates the increased measure of consumers who buy both products from firm \( j \): those on the two rectangle areas switch from buying only one product from \( j \) to buying both from it, and those on the small triangle area switch from buying nothing from firm \( j \) to buying both products from it.
Notice that the small triangle area is a second-order effect when \( \delta \) is small, so the positive (first-order) effect of introducing the small joint-purchase discount is only from the consumers on the two rectangle areas. The measure of them is \( \delta [f(y_1)(1 - F(y_2)) + f(y_2)(1 - F(y_1))] \). From each of these consumers, firm \( j \) originally made a profit \( p \) but now makes a profit \( 2p - \delta \). Therefore, conditional on \((y_1, y_2)\), the positive (first-order) effect of introducing a small \( \delta \) is \( (p - \delta) \delta [f(y_1)(1 - F(y_2)) + f(y_2)(1 - F(y_1))] \). Discarding higher-order effects, integrating it over \((y_1, y_2)\) and using the symmetry yield

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

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

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

The implication of this result is that if implementing mixed bundling is feasible and costless, separate sales cannot be an equilibrium outcome. Unlike the pure bundling case where whether a firm has a unilateral incentive to deviate from separate sales depends on the number of firms, here each firm always has an incentive to do so.

### 5.2 Equilibrium prices

We now characterize the symmetric mixed-bundling equilibrium \((\rho, \delta)\), where \( \rho \) is the stand-alone price for each individual product and \( \delta \leq \rho \) is the joint-purchase discount.\(^{44}\)

Suppose all other firms use the equilibrium strategy, and firm \( j \) unilaterally deviates and sets \((\rho', \delta')\). Then a consumer faces the following options:

- buy both products from firm \(j\), in which case her surplus is \( x_1 + x_2 - (2\rho' - \delta') \)
- buy product 1 at firm \(j\) but product 2 elsewhere, in which case her surplus is \( x_1 + y_2 - \rho' - \rho \)
- buy product 2 at firm \(j\) but product 1 elsewhere, in which case her surplus is \( y_1 + x_2 - \rho' - \rho \)

\(^{44}\)If \( \delta > \rho \), then the bundle would be cheaper than each individual product and only the bundle price would matter for consumer choices. That situation would be like pure bundling where consumers can buy multiple bundles.
• buy both products from other firms, in which case her surplus is \( A - (2\rho - \delta) \) (where \( A \) will be defined below)

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

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

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

When \( n \geq 3 \), the situation is more involved. We need to deal with one more random variable \( z \) defined in (19), and moreover \( z \) is correlated with \( y_1 \) and \( y_2 \) as reported in the following lemma.

**Lemma 5** When \( n \geq 3 \), the cdf of \( z \) defined in (19) conditional on \( y_1, y_2 \) and they being from different firms is

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

for \( z \in \left[ \max\{y_1, y_2\} + z, y_1 + y_2 \right) \).

With this result, the distribution of \( A \) conditional on \( y_1 \) and \( y_2 \) can be fully characterized.

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

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

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

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

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

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

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

$$\alpha(\delta) \equiv \mathbb{E}[f(y_1 - \delta)(1 - F(A - y_1 + \delta))],$$

$$\beta(\delta) \equiv \mathbb{E}[f(A - y_1 + \delta)F(y_1 - \delta)],$$

$$\gamma(\delta) \equiv \mathbb{E}[\int_{y_1-\delta}^{A-y_2+\delta} f(A-x)f(x)dx].$$
(All the expectations are taken over \((y_1, y_2, A)\).) The economic meanings of \(\alpha(\delta)\), \(\beta(\delta)\) and \(\gamma(\delta)\) will be clear soon. They will help describe the marginal effect of a small price deviation by firm \(j\) on its profit.

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

![Figure 7(a): Price deviation and consumer choice II](image)

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

\[
2(\rho - \delta)\alpha(\delta) + (2\rho - \delta)\gamma(\delta) = \Omega_b(\delta) .
\]
Second, suppose firm $j$ raises its stand-alone price to $\rho' = \rho + \varepsilon$ and its joint-purchase discount to $\delta' = \delta + 2\varepsilon$ (such that its bundle price remains unchanged). Figure 7(b) below describes how this small deviation affects consumers’ choices: both $\Omega_1$ and $\Omega_2$ shrink because now fewer consumers buy a single product from firm $j$.

![Figure 7(b): Price deviation and consumer choice III](image)

Here, $\beta_i, i = 1, 2,$ indicates the density of marginal consumers along the line segment $\beta_i$ on the graph, and it equals $f(A - y_i + \delta)F(y_i - \delta)$. Integrating them over $(y_1, y_2, A)$ leads to the notation $\beta(\delta)$ introduced before: $\mathbb{E}[\beta_1] = \mathbb{E}[\beta_2] = \beta(\delta)$. For those marginal consumers with a measure of $\varepsilon(\alpha_1 + \alpha_2)$, they switch from buying only one product to buying both from firm $j$. So firm $j$ earns $\rho - \delta$ more from each of them. For those marginal consumers with a measure of $\varepsilon(\beta_1 + \beta_2)$, they switch from buying one product from firm $j$ to buying both products from other firms. So firm $j$ loses $\rho$ from each of them. The direct revenue effect of the deviation is that firm $j$ earns $\varepsilon$ more from each consumer on $\Omega_1$ and $\Omega_2$ who were originally buying a single product from it. Integrating the sum of these effects over $(y_1, y_2, A)$ should be equal to zero in equilibrium. This yields another first-order condition:

$$\alpha(\delta) + \Omega_1(\delta) = \rho \beta(\delta).$$

(24)

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

Both (23) and (24) are linear in $\rho$, and by using (22) it is straightforward to solve $\rho$ as a function of $\delta$:

$$\rho(\delta) = \frac{1/n + \delta(\alpha(\delta) + \gamma(\delta))}{\alpha(\delta) + \beta(\delta) + 2\gamma(\delta)}. \quad (25)$$

Substituting this into (24) yields an equation of $\delta$:

$$\rho(\delta) (\beta(\delta) - \alpha(\delta)) = \Omega_1(\delta) - \delta \alpha(\delta). \quad (26)$$

39
Discussion: equilibrium existence. To prove the existence of a symmetric (pure-strategy) equilibrium, we need to show (i) the system of (25) and (26) has a solution with \( \delta < \rho \), and (ii) the first-order conditions are sufficient for defining the equilibrium. Unfortunately, both conditions are hard to investigate in general. For the first one, we can show it when \( n = 2 \) or when \( n \) is sufficiently large. For the second one, little analytical progress can be made. This issue exists even in the duopoly case (except for some specific distributions) and is an unsolved problem in the literature on mixed bundling in general. Numerically we can show that the existence is not a problem in the uniform distribution and the normal distribution example.

The complication of (25) and (26) comes from the fact that \( \alpha(\delta), \beta(\delta), \gamma(\delta) \) and \( \Omega_1(\delta) \) usually do not have simple expressions. However, they are simple in the duopoly case, and they also have relatively simple approximations when \( \delta \) is small (which, as we will show below, is often the case when the number of firms is large). Hence, we study these two cases in the following.

For convenience, let \( H(\cdot) \) be the cdf of \( x_i - y_i \). Then
\[
H(t) = \int_{-\infty}^{t} F(x + t)dF(x)^{n-1}; \quad h(t) = \int_{-\infty}^{t} f(x + t)dF(x)^{n-1}.
\]
In particular, when \( n = 2 \), the pdf \( h(t) \) is symmetric around zero, and so \( h(-t) = h(t) \) and \( H(-t) = 1 - H(t) \). One can also check that for any \( n \geq 2 \), \( H(0) = 1 - \frac{1}{n} \). Notice that \( h(0) \) is the density of marginal consumers for each firm. So the price (1) in the regime of separate sales can be written as
\[
p = \frac{1}{nh(0)}.
\]

In the duopoly case, \( A = y_1 + y_2 \) and by using the symmetry of \( h \) one can check that \( \alpha(\delta) = \beta(\delta) = h(\delta)[1 - H(\delta)] \) and \( \Omega_1(\delta) = [1 - H(\delta)]^2 \). Thus, (26) simplifies to
\[
\delta = \frac{1 - H(\delta)}{h(\delta)}.
\]  

---

45In our model, each firm’s pricing strategy space is a subset of \( \mathbb{R}^3 \) and we can make it compact without loss of generality. The profit function is continuous in prices as long as the density functions are. Then Theorem 1 in Becker and Damianov (2006) implies that our model has a symmetric equilibrium (but not necessarily in pure strategy).

46We only did the numerical exercise when \( n \) is no greater than 10. In our numerical calculations we need to approximate some triple integrals. When \( n \) is greater than 10, the numerical approximations become slow and rather imprecise.

47Alternatively, (27) can be written as \( \Omega_1(\delta) + \frac{1}{2} \delta \Omega_1'(\delta) = 0 \). This is the formula derived in Armstrong and Vickers (2010).
If $1 - H$ is log-concave (which is implied by the log-concavity of $f$), this equation has a unique positive solution. Meanwhile, (25) 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 example, one can check that $\delta = 1/3$, $\rho \approx 0.572$ and the bundle price is $2\rho - \delta \approx 0.811$. Compared to the regime of separate sales where $p = 0.5$, each single product is now more expensive but the bundle is cheaper. In the normal distribution example, one can check that $\delta \approx 1.063$, $\rho \approx 1.846$ and the bundle price is $2\rho - \delta \approx 2.629$. The same observation holds.

When $n$ is large, under mild conditions we can show that the system of (25) and (26) has a solution with $\delta$ close to zero. For a small $\delta$ both sides of (26) have simple approximations, and an approximated $\delta$ can be solved.

**Proposition 6**

(i) If a symmetric mixed bundling equilibrium exists, the stand-alone price $\rho$ and the joint-purchase discount $\delta$ satisfy (25) and (26).

(ii) When $n = 2$, $\delta$ solves (27) and $\rho$ is given by (28). The bundle price is lower than in the regime of separate sales (i.e., $2\rho - \delta < 2p$) if $f$ is log-concave.

(iii) Suppose $f'(x)/f(x)$ is bounded and $\lim_{n \to \infty} p = 0$, where $p = \frac{1}{nh(0)}$ is the separate sales price in (1). When $n$ is large, the system of (25) and (26) has a solution which can be approximated as

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

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

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

---

48When $n$ is large, one may wonder whether our approximation results could also be achieved by considering a simpler (asymmetric) duopoly model where one firm supplies superior products with match utilities $y_i$, $i = 1, 2$, but in equilibrium it is required to charge the same prices as the other inferior firm. We can show that this is actually not true.
5.3 Impacts of mixed bundling

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

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

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

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

$$\pi(\rho, \delta) - \pi(\rho, 0) = 2(\rho - p) - n\delta\Omega_b(\delta). \quad (29)$$

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

$$v(\rho, \delta) - v(p, 0) = v(\rho, \delta) - v(p, \delta) + v(p, \delta) - v(p, 0)$$

$$= \int_{\rho}^{p} v_1(\tilde{\rho}, \delta)d\tilde{\rho} + \int_{0}^{\delta} v_2(p, \tilde{\delta})d\tilde{\delta}$$

$$= -2(\rho - p) + n\int_{0}^{\delta} \Omega_b(\tilde{\delta})d\tilde{\delta}. \quad (30)$$

Hence, (29) and (30) provide the formulas for calculating the welfare impacts of mixed bundling. (From these two formulas, it is also clear that mixed bundling harms total welfare given $\Omega_b(\delta)$ is increasing in $\delta$.)

In the duopoly case, the impacts of mixed bundling on profits and consumer surplus are ambiguous. But Armstrong and Vickers (2010) have derived a sufficient condition under which mixed bundling benefits consumers and harms firms. (With our notation, the condition is $\frac{d}{dt} \frac{H(t)}{V(t)} \geq \frac{1}{4}$ for $t \leq 0$. In the case with a large number of firms, as long as our approximation results in Proposition 6 hold, this must be the case.

6 Conclusion

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

One assumption in our pure bundling model is that consumers do not buy more than one bundle. This is without loss of generality if bundling is caused by product incompatibility or high shopping costs. However, if bundling is purely a pricing strategy and if the production cost is relatively small, then it is possible that the bundle price is low enough in equilibrium such that some consumers want to buy multiple bundles in order to mix and match by themselves. Buying multiple bundles is not uncommon, for instance, in the markets for textbooks or newspapers. With the possibility of buying multiple bundles, the situation is actually similar to mixed bundling. For example, consider the case with two products. If the bundle price is $P$, then a consumer faces two options: buy the best single bundle and pay price $P$, or buy two bundles to mix and match and pay $2P$ (suppose the unused products can be disposed freely). For consumers, this is the same as in a regime of mixed bundling with a stand-alone price $P$ for each product and a joint-purchase discount $P$. Our analysis of mixed bundling can be modified to deal with this case.\footnote{If we introduce shopping costs in the benchmark of separate sales (i.e., if a consumer needs to incur some extra costs when she sources from more than one firm), the situation will also be similar to mixed bundling. The shopping cost will play the role of the joint-purchase discount.}

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

\textbf{Appendix: Omitted Proofs and Details}
Proof of Proposition 1: (i) The duopoly model can be converted into a two-dimensional Hotelling setting. Let \( d_i \equiv x_1^i - x_2^i \), and let \( H \) and \( h \) be its cdf and pdf, respectively. Since \( x_1^i \) and \( x_2^i \) are i.i.d., \( d_i \) has support \([\bar{x} - \bar{x}, \bar{x} - \bar{x}]\) and it is symmetric around the mean zero. In particular, \( h(0) = \int f(x)^2 \, dx \). Since \( x_j^i \) has a log-concave density, \( d_i \) has a log-concave density too. Then the equilibrium price \( p \) in the regime of separate sales is given by

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

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

\[
\frac{1}{P/m} = 2\tilde{h}(0) .
\]

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

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

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

When \( f \) is log-concave, bounded, and \( f(\bar{x}) > 0 \), we already know that \( \lambda(2) < 0 \) and \( \lambda(n) > 0 \) for sufficiently large \( n \). In the following, we show that \( \lambda(n) \) changes its sign only once if \( l_F(t) \) and \( l_G(t) \) cross each other at most twice.

We need to use one version of the Variation Diminishing Theorem (see Theorem 3.1 in Karlin, 1968).

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

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

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

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

Chakraborty (1999) used this same theorem to prove a cut-off result on the seller’s revenue in the context of bundling in auctions. But his result does not apply here directly as we discuss in section 4.6.
where the supremum is extended over all sets \( t_1 \leq \cdots \leq t_k \) \((t_i \in A)\), \( k \) is arbitrary but finite, and \( S(x_1, \cdots, x_k) \) is the number of sign changes of the indicated sequence, zero terms being discarded.

**Theorem 1** (Karlin, 1968)  
Consider the following transformation  
\[
\zeta(x) = \int_Y K(x, y) f(y) d\mu(y),
\]
where \( K(x, y) \) is a two-dimensional Borel-measurable function and \( \mu \) 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 (31). Our assumption implies that \( S[l_f(t) - l_g(t)] \leq 2 \). The lemma below proves that \( K(t, n) = t^{n-2} \) is totally positive of order 3. Therefore, the above theorem implies that \( S(\lambda) \leq 2 \). That is, \( \lambda(n) \) changes its sign at most twice as \( n \) varies. Given \( \lambda(2) < 0 \) and \( \lambda(n) > 0 \) for sufficiently large \( n \), it is impossible that \( \lambda(n) \) changes its sign exactly twice. Therefore, it must change its sign only once, and so \( \lambda(n) < 0 \) if and only if \( n \leq \hat{n} \).

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

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

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

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

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

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

\[
\begin{vmatrix}
  1 & 1 & 1 \\
  1 & r_1^{\delta_1} & r_1^{\delta_1} \\
  1 & r_1^{\delta_2} & r_1^{\delta_2}
\end{vmatrix} = (r_2^{\delta_2} - 1)(r_3^{\delta_3} - 1) - (r_2^{\delta_3} - 1)(r_3^{\delta_2} - 1),
\]

45
where \( \delta_j = n_j - n_1 \) and \( r_j = t_j/t_1, j = 2, 3 \). Notice that \( 0 < \delta_2 < \delta_3 \) and \( 1 < r_2 < r_3 \). To show that the above expression is positive, it suffices to show that \( x^y - 1 \) is log-supermodular for \( x > 1 \) and \( y > 0 \). One can check that the cross partial derivative of \( \log(x^y - 1) \) has the same sign as \( x^y - 1 - \log x^y \). This must be strictly positive because \( x^y > 1 \) and \( \log z < z - 1 \) for \( z \neq 1 \).  

(iii) Suppose \( x_j^i \) has a mean \( \mu \) and variance \( \sigma^2 \). When \( m \) is large, by the central limit theorem, \( X_j^i/m \) is distributed (approximately) according to the normal distribution \( N(\mu, \sigma^2/m) \). Then (8) implies that

\[
\frac{P}{m} \approx \frac{p_N}{\sqrt{m}},
\]

where \( p_N \) is the separate sales price when \( x_j^i \) follows a normal distribution \( N(\mu, \sigma^2) \). Then \( P \) increases in \( m \) at speed \( \sqrt{m} \) and so \( \lim_{m \to \infty} P/m = 0 \).

Proof of Proposition 2: Result (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 \( \lim_{x \to x^*} \frac{d}{dx} \left( \frac{1-F(x)}{f(x)} \right) = 0 \), the equilibrium price (1) in the Perlof-Salop model has the same approximation as \( E[x(n) - x(n-1)] \) 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,

\[
p \approx E\left[ x(n) - x(n-1) \right]
\]

as \( n \to \infty \).\(^51\) Then the per-product consumer surplus in the regime of separate sales is 

\[
E\left[ x(n) \right] - p \approx E\left[ x(n-1) \right].
\]

Since the above tail behavior condition also implies \( \lim_{x \to x^*} \frac{d}{dx} \left( \frac{1-G(x)}{g(x)} \right) = 0 \), we can deduce that when \( n \) is large the per-product consumer surplus in the regime of bundling is

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

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

Result (ii): It suffices to show the threshold \( n^* \) exists when \( m \to \infty \). Since the left-hand side of (10) increases with \( n \) while the right-hand side decreases with \( n \), we only need to prove two things: (a) condition (10) holds for \( n = 2 \), and (b) the opposite is true for a sufficiently large \( n \).

\(^{51}\)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 result does not apply. (Otherwise, our analysis could be greatly simplified.)
Condition (b) is relatively easy to show. The left-hand side of (10) approaches 
$\frac{1 - F(x)}{f(x)}$ as $n \to \infty$. So we need
$$
\lim_{n \to \infty} p = \frac{1 - F(x)}{f(x)} < \int_{\frac{1}{x}}^{\frac{F(x)}{f(x)}} dx ,
$$
where the equality is from (3). If $\lim_{n \to \infty} p = 0$ (e.g., if $f(\frac{x}{x}) > 0$), this is clearly the case. But we want to show this even if $\lim_{n \to \infty} p > 0$. Notice that log-concave $f$ implies log-concave $1 - F$ (or decreasing $(1 - F)/f$). Then
$$
\int_{\frac{1}{x}}^{\frac{x}{f(x)}} dx = \int_{\frac{1}{x}}^{\frac{F(x)}{f(x)}} \frac{F(x)}{1 - F(x)} dF(x)
$$
$$
= \int_{\frac{1}{x}}^{\frac{F(x)}{f(x)}} \frac{1 - F(x)}{f(x)} \int_{\frac{1}{x}}^{\frac{F(x)}{f(x)}} dF(x)
$$
$$
= \int_{\frac{1}{x}}^{\frac{F(x)}{f(x)}} \frac{1 - F(x)}{f(x)} \int_{0}^{1} \frac{t}{1 - t} dt .
$$
The integral term is infinity, so condition (32) must hold.

We now prove condition (a). Using (1) and the notation $l(t) \equiv f(F^{-1}(t))$, we can rewrite (10) as
$$
\int_{0}^{1} \frac{t - t^n}{l(t)} dt \int_{0}^{1} t^{n-2} l(t) dt < \frac{1}{n(n - 1)} .
$$
When $n = 2$, this becomes
$$
\int_{0}^{1} \frac{t(1 - t)}{l(t)} dt \int_{0}^{1} l(t) dt < \frac{1}{2} .
$$
To prove this inequality, we need the following technical result:

**Lemma 7** Suppose $\varphi : [0, 1] \to \mathbb{R}$ is a nonnegative function such that $\int_{0}^{1} \frac{\varphi(t)}{l(t)} dt < \infty$, and $r : [0, 1] \to \mathbb{R}$ is a concave pdf. Then
$$
\int_{0}^{1} \varphi(t) r(t) dt \leq \max \left( \int_{0}^{1} \frac{\varphi(t)}{2t} dt, \int_{0}^{1} \frac{\varphi(t)}{2(1 - t)} dt \right) .
$$

**Proof.** Since $r$ is a concave pdf, it is a mixture of triangular distributions and admits a representation of the form
$$
r(t) = \int_{0}^{1} r_\theta(t) \lambda(\theta) d\theta ,
$$
where $\lambda(\cdot)$ is a pdf defined on $[0, 1]$, $r_1(t) = 2t$, $r_0(t) = 2(1 - t)$, and for $0 < \theta < 1$
$$
r_\theta(t) = \begin{cases} 
2 \frac{t}{\theta} & \text{if } 0 \leq t < \theta \\
2 \frac{1 - t}{1 - \theta} & \text{if } \theta \leq t \leq 1
\end{cases} .
$$

---

52I am grateful to Tomás F. Móri in Budapest for helping me to prove this lemma.
(See, for instance, Example 5 in Csiszár and Móri, 2004.)

By Jensen’s Inequality we have
\[
\frac{1}{r(t)} = \frac{1}{\int_0^1 r_\theta(t) \lambda(\theta) d\theta} \leq \int_0^1 \frac{1}{r_\theta(t)} \lambda(\theta) d\theta.
\]

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

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

This is a convex function of \( \theta \), 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.

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 pdf. (The integral in the denominator is finite since \( l(t) \) is nonnegative and concave.) Then Lemma 7 implies that the left-hand side of (33) is no greater than \( 1/4 \).

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

**Lemma 8** Consider a sequence of i.i.d. random variables \( \{x^j\}_{j=1}^n \) with \( x^j \sim \mathcal{N}(0, \sigma^2) \). Let \( p \) be the separate sales price as defined in (1) when firm \( j \)'s product has match utility \( x^j \). Then
\[
\mathbb{E} \left[ \max_j \{x^j\} \right] = \frac{\sigma^2}{p}.
\]

**Proof.** Let \( F(\cdot) \) denote the cdf of \( x^j \). Then the cdf of \( \max_j \{x^j\} \) is \( F(\cdot)^n \), and so
\[
\mathbb{E} \left[ \max_j \{x^j\} \right] = \int_{-\infty}^{\infty} x dF(x)^n = n \int_{-\infty}^{\infty} x F(x)^{n-1} f(x) dx.
\]
For a normal distribution with zero mean, we have $f'(x) = -xf(x)/\sigma^2$. Therefore,

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

$$
= \sigma^2 n (n - 1) \int_{-\infty}^{\infty} F(x)^{n-2} f(x)^2 dx
$$

$$
= \frac{\sigma^2}{p}.
$$

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

**Proof of Proposition 3**: (i) The result that bundling is always a NE has been explained in the main text. Here we show the uniqueness in the duopoly case. We first show that it is not an equilibrium outcome that both firms adopt separate sales. Consider the hypothetical equilibrium where both firms sell their products separately at price $p$. Now suppose firm $j$ unilaterally bundles. Then the situation is just like both firms bundling, and firm $j$ can at least earn the same profit as before by setting a bundle price $mp$. But it can do strictly better by adjusting prices as well. Suppose firm $j$ sets a bundle price $mp - m\varepsilon$, where $\varepsilon$ is a small positive number. The negative (first-order) effect of this deviation on firm $j$’s profits is $\frac{mp}{2}\varepsilon$. (Half of the consumers buy from firm $j$ when $\varepsilon = 0$, and now they pay $m\varepsilon$ less.) The demand for firm $j$’s bundle becomes

$$
\Pr(X_j + m\varepsilon > X_k) = \int G(x + \varepsilon) dG(x) ,
$$

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

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

(The equality is because of (5) for $n = 2$.) Therefore, the deviation is profitable if $P < mp$. Similarly, one can show that if $P > mp$, then charging a bundle price $mp + m\varepsilon$ will be a profitable deviation.

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

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

(ii) We have known (14) fails to hold for \( n = 2 \). On the other hand, we have \( \lim_{n \to \infty} p < \int_{1}^{\infty} F(x)dx \) as shown in the proof of Proposition 2 (even if the latter does not equal zero). Then (14) 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 (14) as
\[
\Delta(n) \equiv p(1 - \frac{1}{n}) - \int_{0}^{1} \frac{t - t^{n-1}}{l(t)} dt < 0.
\]
It suffices to show that \( \Delta(n) \) decreases in \( n \). This is not obvious given \( 1 - \frac{1}{n} \) is increasing in \( n \). (For example, when the match utility distribution is close to the exponential, \( p \) is almost constant in \( n \), and so \( p(1 - \frac{1}{n}) \) increases in \( n \).)

Let \( p_n \) denote the separate sales price when there are \( n \) firms. Then we have
\[
\Delta(n + 1) - \Delta(n) = p_{n+1} \frac{n}{n+1} - p_{n} \frac{n-1}{n} - \int_{0}^{1} \frac{t^{n-1}(1-t)}{l(t)} dt.
\]
From Lemma 1, we know that \( p_{n+1} < p_n \) when \( f \) is log-concave. So
\[
\frac{n}{n+1} - \frac{n-1}{n} < \frac{1}{n(n+1)} \frac{1}{\int_{0}^{1} l(t) t^{n-2} dt}.
\]
(The equality used \( \frac{1}{p_n} = n(n-1) \int_{0}^{1} l(t) t^{n-2} dt \).) Let \( \kappa(t) \equiv n(n+1)t^{n-1}(1-t) \). One can check that \( \kappa(t) \) is a pdf on \([0,1]\). Then we have
\[
\int_{0}^{1} \frac{t^{n-1}(1-t)}{l(t)} dt = \frac{1}{n(n+1)} \int_{0}^{1} \kappa(t) \frac{1}{l(t)} dt > \frac{1}{n(n+1)} \int_{0}^{1} l(t) l(t) \kappa(t) dt.
\]
(The inequality is from Jensen’s Inequality.)

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

**Proof of Lemma 4:** We only prove the results for \( p \). (The same logic works for \( P \) since the log-concavity of \( f \) implies the log-concavity of \( g \).) Notice that \( 1 - F \) is log-concave given \( f \) is log-concave. When \( p = 0 \), it is clear that the left-hand side of (17) is less than
the right-hand side. We can also show the opposite is true when \( p = p_M \). By using the second order statistic as in the proof of Lemma 1, the right-hand side of (17) equals
\[
\frac{1 - F(p)^n}{nF(p)^{n-1}f(p) + \int_p^\infty \frac{f(x)}{1-F(x)}dF(x)(x)} < \frac{1 - F(p)^n}{nF(p)^{n-1}f(p) + \frac{f(p)}{1-F(p)}(1-F(2)(p))} = \frac{1 - F(p)}{f(p)}.
\]
(The inequality is because \( f/(1 - F) \) is increasing, and the equality used the cdf of the second order statistic \( F(2)(p) = F(p)^n + nF(p)^{n-1}(1-F(p)) \).) Then the fact that \( p_M = \frac{1-F(p_M)}{f(p_M)} \) implies the result we want. This shows that (17) has a solution \( p \in (0, p_M) \).

To show the uniqueness, we prove that the right-hand side of (17) decreases with \( p \). One can verify that its derivative with respect to \( p \) is negative if and only if
\[
f'(p)(1 - F(p)^n) + nf(p) \left( F(p)^{n-1}f(p) + \int_p^\infty f(x)dF(x)^{n-1} \right) > 0.
\]
Using \( (1 - F)f' + f^2 > 0 \) (which is implied by the log-concavity of \( 1 - F \)), one can check that the above inequality holds if
\[
n \int_p^\infty f(x)dF(x)^{n-1} > (1 - F(p)^n) \frac{f(p)}{1-F(p)} - nf(p)F(p)^{n-1}.
\]
The left-hand side equals \( \int_p^\infty \frac{f(x)}{1-F(x)}dF(2)(x) \), and the right-hand side equals \( \frac{f(p)}{1-F(p)}(1 - F(2)(p)) \). So the inequality is implied by \( \frac{f(x)}{1-F(x)} > \frac{f(p)}{1-F(p)} \) for \( x > p \).

To prove the second comparative static result, let us first rewrite (17) as
\[
\frac{1}{p} = \frac{f(\overline{x}) - \int_p^\infty f'(x)F(x)^{n-1}dx}{(1 - F(p)^n)/n} = \frac{nf(\overline{x})}{1-F(p)^n} - \int_p^\infty f'(x)\frac{F(x)^n - F(p)^n}{1-F(p)^n}.
\]  
(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 \( \frac{f'}{f} \) 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 (34) increases with \( n \). So the unique solution \( p \) must decrease with \( n \).

**Elastic consumer demand.** We extend the baseline model by considering elastic consumer demand. Suppose each product is divisible and consumers can buy any quantity of a product. In other aspects, for convenience let us focus on the baseline model. As in the previous section we also assume that the \( m \) products are independent to each other. In this setting with elastic demand, if a firm adopts pure bundling strategy, it requires a consumer to buy all products from it or nothing at all.

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

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

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

Firm \( j \)'s profit from product \( i \) is then \(-v'(p')p'q(p')\). It is more convenient to work on indirect utility directly. We then look for a symmetric equilibrium where each firm offers indirect utility \( s \). Given \( v(p) \) is monotonic in \( p \), there is a one-to-one correspondence between \( p \) and \( s \). When a firm offers indirect utility \( s_i \) it must be charging a price \( v^{-1}(s) \) and the optimal quantity a consumer will buy is \(-v'(v^{-1}(s))\). Denote by \( r(s) \equiv v^{-1}(s)(-v'(v^{-1}(s))) \) the per-consumer profit when a firm offers indirect utility \( s \). If firm \( j \) deviates and offers \( s' \), then the measure of consumers who choose to buy from it is

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

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

\[
-\frac{r'(s)}{r(s)} = n \int f(x)dF(x)^{n-1} . \tag{35}
\]

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 when \( r(s) \) is log-concave (so \(-\frac{r(s)}{r(s)}\) is increasing in \( s \)). Once we solve \( s \), we can back out a unique equilibrium price \( v^{-1}(s) \).

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

We look for a symmetric equilibrium where each firm offers an indirect utility \( S \). Suppose firm \( j \) deviates to \( S' \). Then the measure of consumers who buy all products from it is

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

\(^{53}\)In the case with a linear demand function \(-v'(p) = 1 - p\), one can check that \( r(s) = \sqrt{2s} - 2s \) (which is actually concave) and \(-\frac{r(s)}{r(s)}\) increases from \(-\infty\) to \(\infty\) when \( s \) varies from 0 to \(\frac{1}{2} \).
Firm $j$’s profit is $mr\left(\frac{S_j}{m}\right)Q(S')$. So the first-order condition is

$$-\frac{r'(\frac{S_j}{m})}{r(\frac{S_j}{m})} = n \int g(x) dG(x)^{n-1}. \quad (36)$$

Therefore, as long as $-\frac{r'(s)}{r(s)}$ is increasing in $s$ (or $r(s)$ is log-concave), (35) and (36) are similar as the equilibrium price conditions (1) and (5) in the baseline model with unit demand. Then our price comparison results in Proposition 1 continues to hold.

Proof of Lemma 5: For convenience, let $I(y_i)$, $i = 1, 2$, be the identity of the firm where $y_i$ is drawn. The lower bound of $z$ is because the lowest possible bundle match utility at firm $I(y_i)$ is $y_i + \tilde{x}$. We now calculate the conditional probability of $\max_{k \neq j}\{x_1^k + x_2^k\} < z$. This event occurs if and only if all the following three conditions are satisfied: (i) $y_1 + x_2^I(y_1) < z$, (ii) $x_1^I(y_2) + y_2 < z$, and (iii) $x_1^k + x_2^k < z$ for all $k \neq j, I(y_1), I(y_2)$. Given $y_1$ and $y_2$, condition (i) holds with probability $F(z - y_1)/F(y_2)$, as the cdf of $x_2^I(y_1)$ conditional on $x_2^I(y_1) < y_2$ is $F(x)/F(y_2)$. Similarly, condition (ii) holds with probability $F(z - y_2)/F(y_1)$. One can also check (by resorting to a graph, for example) that the probability that $x_1^k + x_2^k < z$ holds for one firm other than $I(y_1)$ and $I(y_2)$ is

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

Conditional on $y_1$ and $y_2$, all the above three events are independent of each other. Therefore, the conditional probability of $\max_{k \neq j}\{x_1^k + x_2^k\} < z$ is the right-hand side of (21).

Proof of Proposition 6: We first prove the price comparison result in (ii). From (27) and (28), we know that the bundle price in the duopoly case is $2\rho - \delta = 1/[2(\alpha + \gamma)]$ (the variable $\delta$ in $\alpha(\delta)$ and $\gamma(\delta)$ has been suppressed). And the bundle price in the regime of separate sales is $1/h(0)$. The former is lower if

$$\alpha + \gamma = h(\delta)[1 - H(\delta)] + 2 \int_0^\delta h(t)^2 dt \geq \frac{h(0)}{2}. \quad (36)$$

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 $f$.

We now prove result (iii). The proof consists of a few steps.

Step 1: Approximate $\alpha(\delta)$, $\beta(\delta)$, $\gamma(\delta)$ and $\Omega_1(\delta)$ when $\delta$ is small.
Lemma 9 For a given $n$, if $\delta \approx 0$, we have

\[
\alpha(\delta) \approx \frac{h(0)}{n} - \left( \frac{h'(0)}{n} + \frac{h(0)^2}{n-1} \right) \delta , \\
\beta(\delta) \approx \left( 1 - \frac{1}{n} \right) h(0) + \left( \frac{h'(0)}{n} - h(0)^2 \right) \delta , \\
\gamma(\delta) \approx \frac{nh(0)^2}{n-1} \delta , \\
\Omega_1(\delta) \approx \frac{1}{n} \left( 1 - \frac{1}{n} \right) - \frac{2h(0)}{n} \delta ,
\]

where $h(0) = \int f(x) dF(x)^{n-1}$ and $h'(0) = \int f'(x) dF(x)^{n-1}$.\textsuperscript{54}

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 (20), 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 (21). By integration by parts and using $L(y_1 + y_2) = 1$, we can simplify this to

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

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

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

where

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

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

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

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

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

\textsuperscript{54}When the support of $x_i$ is finite and $f(\bar{x}) > 0$, the density of $x_i - y_i$ has a kink at zero such that $h'(0)$ is not well defined. However, one can check that $\lim_{t \to 0^-} h'(t) = \int f'(x) dF(x)^{n-1}$ (whenever $n \geq 3$) and $\lim_{t \to 0^+} h'(t) = \int f'(x) dF(x)^{n-1} - (n-1) f(\bar{x})^2$. We use $h'(0^-)$ in our approximations.
The integral term $\varphi(\delta)$ looks more complicated, but $\varphi(0) = 0$ and $\varphi'(0) = f(y_1)f(y_2)$ by noticing that $L(z)$ is independent of $\delta$ and $L(y_1 + y_2) = 1$. Hence,

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

Substituting these approximations into (38) and discarding all higher order terms yield the approximation for $\alpha(\delta)$ in (37). The other approximations can be derived similarly.

**Step 2:** When $n$ is large, the system of (25) and (26) 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 (25) and (26) has a solution with $\delta \in (0, \frac{1}{nh(0)})$.

**Proof.** Recall that (26) is

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

Denote the left-hand side by $\chi_L(\delta)$ and the right-hand side by $\chi_R(\delta)$. Notice that the assumption that $\frac{|f'(x)|}{f(x)}$ is bounded implies that $\frac{|h^0(\delta)|}{h(0)}$ is uniformly bounded for any $n$. (Suppose $\frac{|f'(x)|}{f(x)} < M$ for a constant $M < \infty$. Then $-Mf(x) < f'(x) < Mf(x)$, and so $-M \int f(x)dF(x)^{n-1} < \int f'(x)dF(x)^{n-1} < M \int f(x)dF(x)^{n-1}$ for any $n$. That is, $-Mh(0) < h'(0) < Mh(0)$ for any $n$, and so $\frac{|h^0(\delta)|}{h(0)}$ is uniformly bounded.)

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

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

Next, we show that $\chi_L(\delta) > \chi_R(\delta)$ at $\delta = \frac{1}{nh(0)}$ when $n$ is sufficiently large. The condition $\lim_{n \to \infty} p = 0$ implies that $\delta = \frac{1}{nh(0)} \approx 0$ when $n$ is large. Replacing $\delta$ in (37) by $\frac{1}{nh(0)}$, we have

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

Similarly,

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

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

55
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 when we replace \( \delta \) by \( \frac{1}{nh(0)} \) we did not do further approximations.)

Notice that \( \chi_L(\delta) > \chi_R(\delta) \) if and only if
\[
\left[ \frac{1}{n} + \delta(\alpha(\delta) + \gamma(\delta)) \right] [\beta(\delta) - \alpha(\delta)] > [\Omega_1(\delta) - \delta \alpha(\delta)] [\alpha(\delta) + \beta(\delta) + 2\gamma(\delta)] .
\]
Using the above approximations, we have
\[
\alpha(\delta) + \gamma(\delta) \approx \frac{2h(0)}{n} - \frac{h'(0)}{n^2h(0)} \quad \text{and} \quad \beta(\delta) - \alpha(\delta) \approx \left( 1 - \frac{3}{n} \right) h(0) + \frac{2h'(0)}{n^2h(0)} + \frac{h(0)}{n(n-1)} .
\]
Then the left-hand side of (39) equals
\[
\left[ \frac{1}{n} + \frac{1}{nh(0)} \left( \frac{2h(0)}{n} - \frac{h'(0)}{n^2h(0)} \right) \right] \times \left[ \left( 1 - \frac{3}{n} \right) h(0) + \frac{2h'(0)}{n^2h(0)} + \frac{h(0)}{n(n-1)} \right] \\
= \left[ \frac{1}{n} + \frac{2}{n^2} - \frac{1}{n^3 h(0)^2} \right] \times \left[ h(0) - \frac{3}{n} h(0) + \frac{2h'(0)}{n^2h(0)} + \frac{h(0)}{n(n-1)} \right] \\
\approx \left( 1 - \frac{1}{n^2} \right) h(0) .
\]
(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
\[
\Omega_1(\delta) - \delta \alpha(\delta) \approx \frac{1}{n} - \frac{4}{n^2} + \frac{1}{n^2(n-1)} + \frac{h'(0)}{n^3h(0)^2} ,
\]
and
\[
\alpha(\delta) + \beta(\delta) + 2\gamma(\delta) \approx \frac{h(0)}{n} - \frac{h'(0)}{n^2h(0)} - \frac{h(0)}{n(n-1)} + \left( 1 - \frac{2}{n} \right) h(0) + \frac{h'(0)}{n^2h(0)} + \frac{2h(0)}{n-1} \\
= \left( 1 - \frac{1}{n} \right) h(0) + \frac{h(0)}{n-1} \left( 2 - \frac{1}{n} \right) \\
= \frac{nh(0)}{n-1} \\
\]
Then the right-hand side of (39) equals
\[
\left[ \frac{1}{n} - \frac{4}{n^2} + \frac{1}{n^2(n-1)} + \frac{h'(0)}{n^3h(0)^2} \right] \times \frac{nh(0)}{n-1} \approx \left( 1 - \frac{3}{n(n-1)} \right) h(0) .
\]
(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 (25) and (26) when n is large.

Given the system has a solution with a small δ when n is large, we can approximate each side of (26) around δ ≈ 0 by using (37) and discarding all higher order terms. Then one can solve
\[
\rho \approx \frac{1}{nh(0)} \left( 1 + \frac{\delta h(0)}{n} \right); \quad \delta \approx \frac{2h'(0)}{n h(0)} + \frac{2n^2 - 3n + 2}{n^2 - n} \frac{nh(0)}{\rho}.
\]

(40)

Since n is large and \( \frac{h'(0)}{h(0)} \) is uniformly bounded for any n, this can be further approximated as
\[
\rho \approx \frac{1}{nh(0)}; \quad \delta \approx \frac{1}{2nh(0)}.
\]

References


