

ISSN 1936-5349 (print)
ISSN 1936-5357 (online)

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

BUNDLING AND FIRM REPUTATION

James D. Dana, Jr. & Kathryn E. Spier

Discussion Paper No. 649

09/2009

Harvard Law School
Cambridge, MA 02138

This paper can be downloaded without charge from:

The Harvard John M. Olin Discussion Paper Series:
http://www.law.harvard.edu/programs/olin_center/

The Social Science Research Network Electronic Paper Collection:
<http://ssrn.com/abstract=1349971>

Bundling and Firm Reputation*

James D. Dana, Jr.
Department of Economics
Northeastern University
360 Huntington Avenue
Boston, MA 02115
j.dana@neu.com

Kathryn E. Spier
Harvard Law School
1563 Massachusetts Avenue
Cambridge, MA 02138
kspier@law.harvard.edu

September 14, 2009

Abstract

By bundling experience goods, a manufacturer can more easily maintain a reputation for high quality over time. Formally, we extend Klein and Lefler's (1981) repeated moral hazard model of product quality to consider multi-product firms and imperfect private learning by consumers. When consumers are small, receive imperfect private signals of product quality, and have heterogeneous preferences over available products, then purchasing multiple products from the same firm makes consumers more effective monitors of the firm's behavior. These consumers observe more signals of firm behavior and detect shirking with a higher probability, which creates stronger incentives for the firm to produce high quality products. By constraining all of the firm's consumers to use more effective monitoring and punishment strategies, bundling creates an even stronger incentive for a multi-product firm to produce high quality products. The impact of bundling on incentives is even greater when consumers cannot identify which of the goods is responsible for poor overall product performance.

*We would like to thank Barry Nalebuff, Louis Kaplow, Howard Marvel, James Peck, and Luis Vasconcelos for helpful comments. Kathryn Spier acknowledges financial support from the John M. Olin Center for Law, Economics, and Business at the Harvard Law School.

1 Introduction

The latin expression *falsus in uno, falsus in omnibus*, meaning “false in one, false in everything,” aptly depicts consumers’ expectations about product quality. Consumers expect a firm’s entire product line to be of poor quality when they learn that one of a firm’s products is of poor quality. Similarly, positive experience with one of a firm’s products is a stimulus for consumers to purchase other products from the firm. Multi-product firms understand this dimension of consumer behavior and may even adopt a variety of strategies to exploit it. Product-line branding, commonly referred to as umbrella branding, is one strategy firms use to leverage an established reputation for high quality into additional lines of business. Some firms go one step further by tying or bundling their branded products together, requiring consumers who purchase one product to purchase other products as well.

Although tying and bundling have been viewed with suspicion by US courts – and are deemed per se illegal in some circumstances – it is common for quality assurance to be used as an explicit efficiency defense in antitrust cases.¹ In a landmark case from the 1930s, IBM unsuccessfully argued that tying the sale of paper tabulating cards to the lease of tabulating machines was a necessary and legitimate business practice. According to IBM, even small deviations in the size or thickness of the cards, or the presence of slime or carbon spots, “could cause inaccuracies in the function of the machine, serious in their consequences and difficult to trace to their source, with consequent injury to the reputation

¹See *Jefferson Parish Hop. Dist. No. 2 v. Hyde* (466 U.S. 2[1984]) and *Mozard C. v. Mercedes-Benz of Morth America Inc.* (833 F. 2d 1342, 1348 [9th Cir. 1987]).

of the machines and the good will of the lessors.”² A similar argument was successfully used by General Motors while defending their business practice of requiring their dealers to only use only General Motors parts in aftermarket service and repairs of their cars. “Defective parts, preventing efficient operations of cars, bring dissatisfaction with automobiles themselves. The material result is blame of the manufacturer and consequent loss of sales.”³

This paper analyzes the economic incentives of a manufacturer – the “branded firm” – to bundle experience goods. Formally, we extend Klein and Leffler’s (1981) repeated moral hazard model of reputations for product quality to consider multiple products. In our model, the branded firm competes for consumers with a fringe of competitive manufacturers who produce unbranded, low quality products. There are many long-lived consumers who have heterogeneous preferences for the branded products. In each period, the branded firm decides whether to invest in product quality, reducing the probability that consumers will experience product failures. Although consumers do not directly observe product quality at the time of sale, they infer it over time through the observation of imperfect private signals.

We begin by assuming that consumers correctly attribute product failures to the product that in fact caused the problem. In the absence of bundling, a significant proportion of the branded firm’s customer base choose to mix and match their purchases, consuming a combination of branded and unbranded products.

²*IBM v. United States* (298 U.S. 131 [1936]). While accepting the need for quality assurance the Supreme Court affirmed the lower court’s decision, observing that IBM could include contractual restrictions in its leases requiring lessees to only use cards that met the necessary specifications for accurate functioning.

³*Pick Mfg. Co. v. General Motors Corp. et al.* (80 F. 2d 641 [7th Cir. 1935]).

Because they purchase only one of the branded firm's products, these consumers detect quality deviations by the branded firm relatively slowly, increasing the incentive for the branded firm to shirk on quality. Bundling, by forcing consumers to experience both branded products, hastens detections of quality deviations and decreases the incentive of the branded firm to shirk on quality. Next we assume that consumers are unable to attribute their bad overall experience to particular products. In equilibrium, consumers who purchase only one of the branded firm's products are likely to place some or all of the blame for a bad experience on the low quality fringe firms. The *reputation externality* created by the consumers' propensity to blame the fringe exacerbates the branded firm's incentive problem as it implies that consumers who mix and match are even less effective at monitoring the branded firm's behavior.

In our model, the branded firm's private decision to bundle is aligned with the interests of society as a whole. Bundling creates social value in our model by eliminating a free-rider problem. Consumers who purchase multiple products from the branded firm are better able to monitor the branded firm's choice of product quality, since they receive multiple signals instead of just one. This is socially valuable, since these consumers can respond more quickly when the firm shirks on quality, strengthening the branded firm's incentive compatibility constraint. Absent bundling, consumers would pursue their individual interests without taking into account the impact that their reduced monitoring capability has on society more broadly. When many consumers mix and match their purchases, the branded firm faces a stronger incentive to cheat and reduce quality, since deviations are detected more slowly. In other words, product bundling increases the

branded firm’s incentive to provide high quality products by implementing more effective monitoring and punishment strategies.

Although there is a large literature on bundling and tying in the the legal and economic literatures,⁴ the issue of firm reputation and quality assurance has received relatively little attention. Bork (1978) and Posner and Easterbrook (1981) argued informally that when consumers use low quality goods along with high quality products, then low overall performance may be erroneously attributed to the producer of the high quality product.⁵ By tying or bundling the products together, the seller can protect its reputation. Iacobucci (2003) refined this argument further, highlighting the importance of attribution errors and the corresponding “confusion externalities.”⁶ Our work provides additional rigor and clarity to these arguments by providing an explicit treatment of the dynamic mechanisms by which consumers learn about product quality and by which firms maintain their reputations over time. In contrast to Iacobucci, we find that the consumers’ attribution problem strengthens the benefits of bundling but is not necessary for bundling to be privately or socially desirable.

Our model is closely related to the existing literature on umbrella branding. As in our model, much of that literature considers multiproduct versions of Klein and Leffler (1981). In a perfect public monitoring model related to Bernheim

⁴Adams and Yellen (1976) and McAfee, McMillan and Whinston (1989) have argued that bundling can be an effective form of price discrimination, allowing a monopolist to extract greater rents from consumers. See also Fang and Norman (2006). Others, including Whinston (1990) and Nalebuff (2004), have argued that bundling can be used as a competitive weapon by an incumbent to foreclose the market. Bundling may also reduce the costs of production and distribution of products. It is more efficient for automobile manufacturers, for example, to bundle tires with cars than for consumers to do it themselves. See Salinger (1995).

⁵For additional references and discussion of the legal literature, see Iacobucci (2003, at 437).

⁶See also Bar-Gill (2006).

and Whinston’s (1990) model of multimarket contact, Andersson (2002) shows that as long as the products are asymmetric and at least one of the incentive constraints for producing high quality binds, joint production is profitable for a monopolist.⁷ Cabral (2009) and Cai and Obara (2006) demonstrate the value of joint production when products are symmetric but there is imperfect public monitoring. Punishment occurs on the equilibrium path as the public learns about quality deviations, and umbrella branding reduces the prevalence of equilibrium path punishments.⁸

In the existing umbrella branding literature, information about product failures is public knowledge. All consumers have the same information, regardless of their individual purchase decisions and there is no reason for the branded firm to bundle its products together. In contrast, we focus on the role of private monitoring and free-riding by consumers.⁹ Consumers who mix and match their purchases learn about product quality deviations more slowly than consumers who buy multiple branded products. There are two reasons why these consumers learn more slowly. First, these consumers receive fewer imperfect private sig-

⁷See also Myatt and Rasmusen (2009).

⁸Hakenes and Peitz (2008) also consider imperfect public monitoring, but they assume (as we do) that there is a positive probability of accurately observing that the firm produced a low quality product. Montgomery and Wernerfelt (1992) consider a free-entry repeated moral hazard model in which some consumers observe quality. (In their model, umbrella branding is associated with lower variation in quality for the consumer and lower, not higher, average product quality.) Other papers in the umbrella branding literature, including Choi (1998), Wernerfelt (1998), and Cabral (2000), analyze umbrella branding as a signal of product quality. Still others consider demand-side and supply-side economies of scope to explain umbrella branding. For example, Pepall and Ricards (2002) consider the use of umbrella branding to create cross-product consumption externalities when consumers prefer to buy products carrying widely known brands.

⁹Imperfect private monitoring often implies that when firms use pure strategies, bad experience does not reveal anything about quality on the equilibrium path, so consumers’ behavior is insensitive to their experience. We assume that when a product fails, this reveals that the firm chose low quality, but a low quality product does not necessarily fail.

nals about the branded firm's quality choices. Second, there is a greater chance that these mix-and-match consumers will wrongly attribute failures caused by the branded product to the competitive fringe. Through bundling, the branded firm increases the speed and accuracy of consumers' collective monitoring capabilities and hence more easily sustains a reputation for producing high quality products.

The empirical literature on branding and firm reputation, while limited, is consistent with our theory. The marketing literature has consistently found that consumers expect new product introductions by established brands to be of higher quality when the brand has historically been associated with high quality (see for example Aaker and Keller, 1990 and Erdem, 1998). More recently, Jin and Leslie (2009) found that restaurants that are part of a restaurant chain are cleaner than those that are not affiliated with a chain. This suggests that restaurant chains, which are multi-product firms by virtue of their multiple restaurant locations, internalize the effect that a consumer's private experience from dining at one location has on the consumer's future demand for dining at the other locations. Our theory is also consistent with an early companion paper, Jin and Leslie (2003), which showed that making information about quality more public by requiring restaurants to post their cleanliness ratings in the window also increased quality.

The next section of the paper presents a simple stylized example that gives much of the intuition for our main results. Section 3 describes the model and characterizes the benchmark case in which product quality is observable at the time of sale. Section 4 characterizes the equilibrium price and quality when quality is unobservable for three cases: first, the case in which the two products are

sold by independent firms; second, the case in which one firm sells both products; and third, the case in which one firm sells both products as a bundle. This section also characterizes the conditions under which bundling is profitable (i.e., the conditions under which bundling is necessary to sustain high quality). Section 5 extends the model to analyze the attribution problem and shows that reputation externalities imply bundling is even more likely to be privately profitable and socially desirable. Section 6 concludes.

2 A Stylized Example

A monopolist sells a non-durable experience good to a group of repeat-purchase customers. The manufacturer and the consumers have a common discount factor δ . At the beginning of each period the manufacturer decides whether to produce a high quality good at unit cost 4 or a low quality good at unit cost 2. The consumers value the high quality good at 10 but place zero value on the low quality good. If quality were observed by the consumers before they made their purchase decisions, then the manufacturer would produce high-quality products and charge a price of 10, extracting all consumer surplus. This would be the first-best outcome. Although consumers do not directly observe product quality at the time of purchase, they do receive an imperfect private signals of quality immediately after purchase. If the quality of the unit is low, there is a 50% chance that the consumer will receive a negative signal alerting him to that fact. If the quality of the unit is high, then no signal is generated (so there are no false positives).

Through the repeated market interaction between the manufacturer and the consumers, it may be possible to sustain an equilibrium where the manufacturer produces high quality products over time. The logic, first outlined by Klein and Leffler (1981) in the context of product quality and later generalized in the large literature on repeated games, is by now familiar. If they expect quality to be high, consumers would be willing to pay a price of up to 10. As long as the consumer never receives a negative signal, he or she would be willing to continue to purchase the product at the premium price. If the monopolist ever cheats and reduces the quality to save on production costs, then a fraction of the consumers – 50% in our example – would receive the negative signal. Expecting that quality would be low in the future as well, those disenfranchised consumers would not be willing to pay a premium price going forward.¹⁰ When the discount factor is sufficiently high, the monopolist refrains from cutting quality because the short run gain is outweighed by the future loss of monopoly rents.

For example, suppose first that market consists of 150 separate consumers, each of whom demand at most one unit of the good. It is possible to sustain high quality over time when the following condition holds:

$$150 \sum_{t=1}^{\infty} \delta^{t-1} (10 - 4) \geq 150 \sum_{t=1}^{\infty} \delta^{t-1} (.5)^{t-1} (10 - 2).$$

The left-hand side is the present discounted value of monopoly profits from

¹⁰While many equilibria exist (the Folk Theorem), the equilibrium described here provides the strongest incentives for high quality. The price of 10 creates the greatest possible “punishment” for the monopolist once caught cheating, much in the same way that an efficiency wage provides strong incentives for workers. The beliefs of the consumers that future quality will be low following a negative signal creates the greatest future loss of market share for the monopolist, further strengthening his incentives not to cheat.

selling the high quality good at a price of 10 to all 150 customers. On the right-hand side is the profit that the monopolist would receive if it were to shirk on quality at $t = 1$ and in every period after that.¹¹ If quality were low at $t = 1$, then $(.5)^{t-1}$ or half the consumers would not detect the warning signal and would continue to purchase at the premium price in period $t = 2$, giving the monopolist a profit margin of $10 - 2$. In period $t = 3$, only a quarter of the original consumers would still be active in the market. Simple algebra reveals that high quality is sustainable in equilibrium when $\delta \geq .40$.

Now suppose instead that the market consists of 100 customers, 50 of whom purchase one unit of the good and 50 of whom purchase two units in each period. High quality is sustainable when

$$150 \sum_{t=1}^{\infty} \delta^{t-1} (10 - 4) \geq 50 \sum_{t=1}^{\infty} \delta^{t-1} (.5)^{t-1} (10 - 2) + 100 \sum_{t=1}^{\infty} \delta^{t-1} (.5)^{2(t-1)} (10 - 2).$$

The term on the left-hand side is exactly the same as before.¹² On the right-hand side, the first term is the present discounted value of the profits from selling a low quality product to the 50 consumers who demand just one unit each. The second term, which does not appear in the first condition, is the profits from the 50 consumers who purchase two units of the product each. These customers have twice the opportunity to detect cheating by the manufacturer in each period. And, when these consumers observe a negative signal associated with just one unit of the two units, they respond by not purchasing any units at all in the next

¹¹If it is profitable for the monopolist to shirk and reduce quality in period $t = 1$, then it would be profitable for him to shirk in period $t = 2$ and in every period after that.

¹²If the manufacturer chooses high quality in every period, it will sell 150 units in each period with a profit margin of $10 - 4$.

period and in every period thereafter.¹³ After t periods of shirking, only $(.5)^{2(t-1)}$ of these multiple-unit consumers will remain. Rearranging terms, we find that high quality is sustainable when $\delta > .34$, a larger range of parameter values than before. High quality is easier to sustain than before. Intuitively, consumers of multiple units can detect shirking more easily and retaliate more harshly than consumers of single units which strengthens the incentives of the monopolist to maintain high quality.

Finally, suppose that instead of simply charging 10 per unit, the monopolist bundles the product into pairs and charges 20 for the bundle. The 50 consumers who only want one unit of the product are priced out of the market, but the 50 consumers who want two units would be willing to pay up to 20 for the bundle (assuming they expect high quality). High quality is sustainable in equilibrium when

$$100 \sum_{t=1}^{\infty} \delta^{t-1} (10 - 4) \geq 100 \sum_{t=1}^{\infty} \delta^{t-1} (.5)^{2(t-1)} (10 - 2).$$

This condition differs from the previous condition in two ways. First, on the left-hand side, the volume of sales is smaller than before. Instead of selling 150 units to all consumers, the manufacturer is now limiting its sales to the 50 consumers who each demand two units each. The right-hand side is different as well, since there is no profit generated from cheating consumers who purchase one unit each. Rearranging terms, we find that high quality is sustainable when $\delta > .31$, which is an even greater range of parameter values. Without bundling the single-unit consumers were free-riding on the monitoring capabilities of the multiple-unit

¹³If the monopolist reduces quality in a given period, a multiple-unit consumer observes a negative signal with probability 75%.

consumers. Bundling screens out these free riders and allows the firm to more easily sustain high quality.

The advantages of bundling can be even stronger when consumers have difficulty attributing negative signals to the monopolist. To see why, imagine that the 50 consumers who only demanded a single unit of the premium product are, in fact, also consuming a second low-quality unit. This unit might be a home-grown unit produced by the consumer (for example, some consumers might be able to maintain and repair things themselves) or it might be produced at low cost by a competitive fringe. When these 50 consumers receive a negative signal they are unable to distinguish the source of that signal. Was the negative signal generated by the premium unit? Or was the negative signal coming from the home-grown unit?

This attribution problem makes it significantly more difficult for the manufacturer to maintain high quality without bundling. In an equilibrium in which the monopolist only produces high quality products, the single-unit consumers would observe negative signals but would rationally attribute them to the lower-quality home-grown unit, not to the premium unit supplied by the monopolist. Hence the single-unit consumers will continue to buy regardless of the negative signals they have received, and the monopolist has an even stronger incentive to shirk. The attribution problem exacerbates the cost of allowing single-unit consumers to free ride on the superior monitoring capabilities of the multiple-unit consumers. So preventing free riding is even more valuable and bundling is even more attractive.

3 The Model

A unit mass of infinitely-lived consumers demand at most a single unit of two goods, a and b , in each period of their lives. These non-durable goods are available from two different sources, a branded firm and a competitive fringe of unbranded firms. The fringe firms produce low quality, undifferentiated products at unit cost c_l . Each consumer's maximal willingness to pay for these low quality unbranded products is normalized to $V_0 > c_l$. The branded firm, on the other hand, manufactures products that are differentiated both horizontally and (potentially) vertically. If the branded firm produces a low quality product then the unit production cost is c_l , exactly the same as for the fringe firms. If the branded firm produces a high quality product, the unit cost is $c_h > c_l$. All players have a common discount factor, δ .

Consumer i 's willingness to pay for the branded product j is:

$$V_j^i = V_0 + z_j^i + v_j.$$

The parameter $v_j \in \{0, (1 - \pi)h\}$ is the same for all consumers and reflects the quality, or vertical differentiation, of the branded product. Low quality products, whether produced by a fringe firm or by the branded firm, fail with probability $1 - \pi$. The direct cost that a consumer suffers from a product failure is h . If the branded firm produces a low quality product then it is no more reliable than the fringe product, and the incremental value is $v_j = 0$. High quality products are strictly better for consumers because they never fail or malfunction. The incremental value of a high quality branded product is $v_j = (1 - \pi)h$, the expected

benefit of fewer product failures.¹⁴

The parameter $z_j^i \in \{-x, x\}$ reflects the horizontal differentiation of the branded product. While some consumers have a personal preference for the branded product j , namely those for whom $z_j^i = x$, others prefer the unbranded product (those for whom $z_j^i = -x$). Consumer i 's preferences for the two products, a and b , are drawn from a symmetric binary distribution at the beginning of the game and do not change over time. The unconditional probability that $z_j^i = x$ is equal to one half, as is the unconditional probability that $z_j^i = -x$; in other words, consumers are just as likely *ex ante* to prefer the branded product as not. We allow for correlation between an individual consumers' preferences for the two products and define $\rho = \Pr\{z_a^i = x | z_b^i = x\} = \Pr\{z_a^i = -x | z_b^i = -x\}$. Note that $\rho = 1$ corresponds to perfect positive correlation for an individual consumer, while $\rho = 0$ corresponds to perfect negative correlation. It follows that the probability that a consumer is of type (x, x) or, equivalently, of type $(-x, -x)$ is $\rho/2$. The probability that a consumer is type $(-x, x)$ or, equivalently, type $(x, -x)$ is $(1 - \rho)/2$. To simplify the notation going forward, we will suppress the superscript i referring to individual consumers.

Throughout the analysis, we will assume that $(1 - \pi)h > c_h - c_l$ so it is socially efficient for the branded firm to produce high quality, reliable products. More substantively, we also assume that $x \geq (1 - \pi)h$, so horizontal differentiation is sufficiently important relative to vertical differentiation. In other words, we

¹⁴Note that we are implicitly assuming that the incremental willingness to pay for a high quality product, $(1 - \pi)h$, is independent of which other products the consumer buys. This holds if the probability of failure is independently distributed and the cost of failure, h , is independent of whether or not the other product fails. In the next section we discuss alternative assumptions that yield the same implication.

assume that consumers care more about the horizontal differences between the branded firms and the fringe firms than any vertical differences.¹⁵ These assumptions imply that it is socially optimal for consumers with $z^j = x$ to purchase high quality branded products and for consumers with $z_j = -x$ to purchase low quality unbranded products, even though these unbranded products have a higher risk of failure.¹⁶

Timing. The timing of moves within each period is as follows. First, the branded firm decides whether to produce a low quality product at cost c_l or a high quality product at cost c_h . This decision is the private information of the branded firm. Next, the branded firm and the competitive fringe firms simultaneously choose prices for their products. Then, given these prices and their privately observed horizontal preferences z_j^i , consumers choose whether to purchase each product and from whom.¹⁷ Finally, after making their purchase decisions, consumers privately observe the performance, success or failure, of their own purchases.

Several remarks are in order. First, consumers do not observe the successes or failures of other consumers' purchases. They learn only through their own purchases. Moreover, if the branded firm produces a low quality product, each consumer who purchases the product learns that the firm has chosen low quality with probability $1 - \pi$ even if low quality is an off-the-equilibrium-path action.

¹⁵This assumption is stronger than we really need. The weaker conditions are that $x \geq (1 - \pi)h(1 - \rho)$ and $x \geq (1 - \pi)h/3$.

¹⁶It is efficient for consumers with $z_j = -x$ to purchase the low quality unbranded product when $V_0 - c_l > V_0 - x + (1 - \pi)h - c_h$, or $c_h - c_l > -x + (1 - \pi)h$. This last condition is true by the assumption that $x > (1 - \pi)h$.

¹⁷At the time of purchase, the consumers do not directly observe the vertical quality dimension, v_j , although they form inferences.

This technology is a blend of imperfect and perfect private monitoring. The signal of product quality is imperfect because consumers do not always observe low product quality when the firm shirks, however it is perfect in the sense that consumers know with certainty that the product was low quality conditional on receiving a signal. Finally, we assume that the product failures are independently distributed across consumers.

Demand. Competition among the fringe firms drives the price of the unbranded products down to their production cost, c_l . The branded firm can charge more than this, capturing the value created by horizontal and (possibly) vertical differentiation. We will now characterize the demand functions for unbundled and bundled products, assuming that the quality level is known to consumers.

The most that a consumer of type $z_j \in \{-x, x\}$ is willing to pay for a low quality product from the branded firm is the price at which his surplus from consuming the branded product equals his surplus from consuming the unbranded product. Formally, it is the price $\hat{p}_l(z)$ where $V_0 + z_j - \hat{p}_l(z_j) = V_0 - c_l$. Rearranging terms,

$$\hat{p}_l(z_j) = c_l + z_j.$$

The total demand for low quality products produced by the branded firm is therefore

$$d_l^j(p) = \begin{cases} 0 & \text{if } p > \hat{p}_l(x) \\ \frac{1}{2} & \text{if } p \in [\hat{p}_l(-x), \hat{p}_l(x)] \\ 1 & \text{if } p < \hat{p}_l(-x). \end{cases}$$

When $p > \hat{p}_l(x)$, all consumers prefer to purchase the unbranded product from the

fringe manufacturers for c_l . When $p \in [\hat{p}_l(-x), \hat{p}_l(x)]$, consumers with a positive horizontal preference for the branded product buy it, but those with a negative horizontal preference do not. When $p < \hat{p}_l(-x)$ then both types of consumer purchase the branded product.

Similarly, the most that a consumer of type $z_j \in \{-x, x\}$ is willing to pay for a high quality product from the branded firm, $\hat{p}_h(z)$, is given by $V_0 + z_j + (1 - \pi)h - \hat{p}_h(z_j) = V_0 - c_l$, or

$$\hat{p}_h(z_j) = c_l + z_j + (1 - \pi)h.$$

This price, which is higher than $\hat{p}_l(z_j)$, reflects the incremental value from enhanced product reliability, $(1 - \pi)h$. The total demand for high quality products is

$$d_h^j(p) = \begin{cases} 0 & \text{if } p > \hat{p}_h(x) \\ \frac{1}{2} & \text{if } p \in [\hat{p}_h(-x), \hat{p}_h(x)] \\ 1 & \text{if } p < \hat{p}_h(-x). \end{cases}$$

Finally, suppose that two high quality products are bundled together and sold as a package for price P .¹⁸ The demand for the bundle would be

$$d_{hh}(P) = \begin{cases} 0 & \text{if } P > 2\hat{p}_h(x) \\ \rho/2 & \text{if } P \in (\hat{p}_h(-x) + \hat{p}_h(x), 2\hat{p}_h(x)] \\ 1 - \rho/2 & \text{if } P \in (2\hat{p}_h(-x), \hat{p}_h(-x) + \hat{p}_h(x)] \\ 1 & \text{if } P \leq 2\hat{p}_h(-x). \end{cases}$$

¹⁸An analogous demand correspondence could be derived for bundles of low quality products, or bundles including one low quality item and one high quality item.

If $P \in (\hat{p}_h(-x) + \hat{p}_h(x), 2\hat{p}_h(x)]$ only consumers for whom $z_a = z_b = x$ would purchase the bundle, and the mass of consumers is $\rho/2$. If $P \in (2\hat{p}_h(-x), \hat{p}_h(-x) + \hat{p}_h(x)]$, then consumers with a horizontal preference for just one of the two products ($z_a = x$ and $z_b = -x$ or vice versa) would purchase the bundle as well. The mass of these consumers is $1 - \rho/2$.

Benchmark: Observable Quality. If quality were observable at the time of sale, or contractible, then the branded firm would not find it profitable to bundle the products. The best the branded firm could do is sell high quality versions of both a and b at the price $p_h^* = \hat{p}_h(x) = c_l + x + (1 - \pi)h$ for each product.¹⁹ Only consumers who have a positive horizontal preference for the branded products would purchase them. The branded firm extracts all of the incremental surplus from the type x consumers and does not sell to the type $-x$ consumers. Note that this implements the social welfare benchmark described earlier.²⁰

The strategy of selling individual units at a price of $p_h^* = \hat{p}_h(x) = c_l + x + (1 - \pi)h$ is clearly more profitable for the branded firm than bundling two high-quality products together and charging $2p_h^* = 2\hat{p}_h(x) = 2[c_l + x + (1 - \pi)h]$. Intuitively, bundling drives the consumers who only want one branded product out of the market, generating lower profits for the branded firm.²¹ The branded

¹⁹If quality were constrained to be low, the branded firm would sell both a and b at the price $p_l^* = \hat{p}_l(x)$ for each product.

²⁰This is more profitable than selling at a price $\hat{p}_h(-x) = c_l - x + (1 - \pi)h$ because our earlier assumption that $x \geq (1 - \pi)h$ implies the firm would prefer to earn a margin of $x + (1 - \pi)h$ selling an average of one product per customer than a margin of $-x + (1 - \pi)h$ selling two products to each customer. This is also clearly more profitable than bundling two high quality products together. The branded firm prefers to earn a profit margin on $x + (1 - \pi)h$ selling on average one unit per customer than selling the bundle at a price of $2x + 2(1 - \pi)h$ to $(1 - \rho/2)$ consumers.

²¹Without bundling, the branded firm sells on average one unit to each consumer. With bundling, the branded firm sells two units on average to $\rho/2 < 1/2$ consumers. This is clearly

firm's profits would be even lower if they were to charge $\hat{p}_h(x) + \hat{p}_l(-x) = x + (1 - \pi)h - x + (1 - \pi)h = 2(1 - \pi)h$ for the bundle, a low enough price to appeal to consumers who have preferences for just one branded good. Since $x > (1 - \pi)h$ by assumption, the branded firm prefers to earn $x + (1 - \pi)h$ selling an average of one product per customer than $(1 - \pi)h$ selling two products each to $1 - \rho/2$ customers.

The fact that bundling is not profitable here is driven by the restrictions on the parameter values. If the values were such that it were socially efficient for the consumers with $z_j = -x$ to purchase branded products as well, then bundling could certainly emerge as a profitable strategy for the branded firm. Most clearly, if consumers' preferences for a and b were perfectly negatively correlated, that is, $\rho = 0$, then the monopolist could extract the full consumer surplus through bundling but not through linear pricing. We restrict attention to the simpler case in which bundling is not profitable under when quality is observable in order to focus our attention on the additional private and social benefits that arise from bundling when quality is unobservable.

4 Equilibrium with Unobservable Quality

Suppose that the quality of the branded product, v_j , is not observable to consumers. It is clear that a perfect Bayesian equilibrium exists in which the branded firm produces low quality versions of its products and sells them at a price of $\hat{p}_l(x) = c_l + x$. In this equilibrium, consumers who have a positive horizontal

less profitable for the monopolist for all levels of correlation with $\rho < 1$. When $\rho = 1$ then the branded firm does just as well with bundling as without.

preference for the branded good expect low quality and are (just) willing to pay this price. Given the consumers' expectations that quality will be low in the future, the branded firm has no incentive to invest resources to improve the quality of its products. This is not necessarily the only equilibrium, however. Indeed, scholars have established folk-theorem type results in related contexts showing a range of achievable outcomes.

In this section, we find conditions under which it is possible to sustain equilibria in which the branded firm produces only high quality products. Specifically, we focus on symmetric, stationary perfect Bayesian equilibria in which consumers believe that the branded firm will continue produce high quality products as long as they have not experienced any product failures in the past, regardless of the prices charged.²² We also focus on equilibrium in which if a consumer ever observes a product failure – hard evidence that the branded firm deviated and produced a low quality product – the consumer believes that the quality of the branded products will be low in all future periods. These off-the-equilibrium-path beliefs convey the harshest punishment on the branded firm for shirking and thus provide the best incentives for the branded firm to invest in product quality. These equilibrium beliefs maximize the range of parameters for which the branded firm is able to maintain prices that extract all of the consumer surplus: p_h^* for an unbundled product and $2p_h^*$ for a bundled product.²³

Our analysis will proceed in several steps. First, we will consider a benchmark

²²In particular, we rule out equilibria in which consumers expect the branded firm to produce a low quality product simply because they deviate to a higher price.

²³We also focus only on equilibrium in which the branded firm charges its monopoly prices. Focusing on these prices further strengthens the brand firm's incentive not to cheat and cut quality.

case in which the two branded products are produced by two separate firms (or two independent divisions of the same firm). Second, we will consider the case in which the multiproduct firm sells unbundled branded products. Third, we will consider the case in which the multiproduct firm sells bundled branded products. Finally, we will compare these three cases in terms of their private and social desirability.

Single-Product Firms

Suppose the branded firm were split into two separate firms, one producing product a and the other producing product b . The managers of the two branded firms make their quality and pricing decisions independently of each other, only considering their own firm's profits.

In this section, we show that if the discount factor, δ , is sufficiently close to one, there exists a perfect Bayesian equilibrium in which the branded firms produce high quality products and charge $p_h^* = c_l + x + (1 - \pi)h$, earning profits of $(p_h^* - c_h)d_h^j(p_h^*)$ in each period. Consumers, believing that quality is high, are willing to pay this amount but no more. If a consumer ever observed a product failure, he or she would infer (correctly) that the firm shirked on quality and would then expect that the firm would shirk on quality in the future as well.

This outcome is sustainable as an equilibrium so long as there doesn't exist a profitable deviation for the branded firm. Given the assumed beliefs of consumers, the best possible deviation for a branded firm has the following form: the firm would shirk on quality while maintaining the high price of p_h^* for a fixed number of periods, after which it would continue to shirk on quality but drop the price

to p_l^* . This makes intuitive sense. When charging p_h^* but providing low quality, the firm reaps the benefits from tricking consumers in the short run, enjoying the lower unit costs and associated higher profit margins. The downside is that the firm loses a proportion $1 - \pi$ of their remaining customer base in each period, as these customers experience product failures and switch to the unbranded products. Over time, the mass of remaining consumers who have never experienced a product failure shrinks and eventually the firm would find it more profitable to lower its price to p_l^* and regain full market share.²⁴

Formally, incentive compatibility for a single-product branded firm requires

$$\sum_{t=1}^{\infty} \delta^{t-1} (p_h^* - c_h)(1/2) \geq \max_T \sum_{t=1}^T \delta^{t-1} \pi^{t-1} (p_h^* - c_l)(1/2) + \sum_{t=T+1}^{\infty} \delta^{t-1} (p_l^* - c_l)(1/2). \quad (1)$$

The left-hand side is the present discounted value of the branded firm's equilibrium profits from selling high quality products. The right-hand side is the firm's profit if it were to deviate and sell low quality products at $t = 1$ and every period after that. The firm will initially keep its prices high at p_h^* and earn a profit margin of $p_h^* - c_l$. The demand shrinks over time to a proportion π of lagged demand in each subsequent period. The right-hand side is monotonically decreasing in T and eventually the firm finds it profitable to lower its price to p_l^* . Let T_s be the last year in which the firm makes more money selling at a price p_h^* than selling at a price p_l^* , or, equivalently, let T_{s+1} be the first year where it becomes more

²⁴At a price of p_l^* , all consumers of type x will purchase the branded product, regardless of their past experience with product failures and their beliefs about future quality. Consumers of type $-x$ would not purchase the product, since $x > (1 - \pi)h$.

profitable for the firm to drop the price to p_l^* , That is, T_s is the unique value of T such that:

$$\pi^{T_s-1}(p_h^* - c_l) \geq p_l^* - c_l > \pi^{T_s}(p_h^* - c_l). \quad (2)$$

Expression (1) can be rewritten as

$$\left(\frac{p_h^* - c_h}{1 - \delta}\right) \geq \left(\frac{p_h^* - c_l}{1 - \delta\pi}\right) (1 - \delta^{T_s}\pi^{T_s}) + \delta^{T_s} \left(\frac{p_l^* - c_l}{1 - \delta}\right). \quad (3)$$

Comparing this incentive constraint to the standard incentive constraint with perfect public monitoring,

$$\left(\frac{p_h^* - c_h}{1 - \delta}\right) \geq p_h^* - c_l + \delta \left(\frac{p_l^* - c_l}{1 - \delta}\right),$$

it is obvious that when a consumer learns only from his or her own private experience, and not from other consumers' experiences, that the incentive constraint is more difficult to satisfy. So a policy change which improves consumer learning, such as the law requiring restaurants to post cleanliness ratings in their windows that was analyzed by Jin and Leslie (2003), can lead to an increase in firm quality.²⁵

²⁵Note that because we assume consumers are small and failures are independently distributed, assuming public monitoring makes learning perfect as well. More generally, it is straightforward to show that even with imperfect monitoring, if consumers' monitoring is semi-public so they observe not only from their private monitoring of their own product purchases, but also from public monitoring of a finite number of other consumers' product purchases, then the range of discount factors for which high quality can be sustained is monotonically increasing in the amount of public learning.

Multi-Product Firms

Now suppose instead that a single multi-product firm produces both branded products, but does not bundle them. As before, we will show that when the discount factor δ is sufficiently close to 1, then there exists a perfect Bayesian equilibrium where the firm produces high quality products and charges $p_h^* = c_l + x + (1 - \pi)h$ for each product. The consumer beliefs that support this equilibrium are analogous to those in the previous section. A consumer believes that quality of both products will be high unless the consumer witnesses a product failure. Specifically, if one or both products fail in a given period, the consumer believes that the quality of both products will be low in the future and is no longer willing to pay p_h^* .

Given the assumed consumer beliefs, the best possible deviation for the branded firm is to reduce the quality of both products while initially keeping prices high, reaping profit margins of $p_h^* - c_l$ from the remaining unsuspecting customers. As consumers experience product failures and the customer base shrinks below a certain level, the firm would choose to lower its price to p_l^* to regain its market share (albeit at a lower profit margin). Formally, incentive compatibility for the branded firm can be written as

$$\sum_{t=1}^{\infty} \delta^{t-1} (p_h^* - c_h) (1/2 + 1/2) \geq \max_T \left[\sum_{t=1}^T \delta^{t-1} (p_h^* - c_l) \pi^{t-1} (1 - \rho) + \sum_{t=1}^T \delta^{t-1} 2(p_h^* - c_l) \pi^{2(t-1)} (\rho/2) + \sum_{t=T+1}^{\infty} \delta^{t-1} (p_l^* - c_l) (1/2 + 1/2) \right]. \quad (4)$$

The left-hand side of this expression is the present discounted value of equilibrium profits. The right-hand side is the present discounted value from deviating. For the first T periods, the firm sells low quality products while keeping prices steady at p_h^* . In the first term, $(1 - \rho)$ is the mass of consumers who purchase exactly one unit of either good a or good b , but not both. These consumers experience product failures at a rate of $1 - \pi$ per period, so after t periods a proportion π^t of these consumers remain in the market. In the second term, $\rho/2$ is the mass of consumers who purchase one unit of good a and one unit of good b in each period. These consumers learn that the firm deviated more quickly, since they have two opportunities to observe a product failure in each period. After T periods, a proportion π^{2t} of these consumers remain in the market. Let T_m denote the last year in which the firm makes more money selling the low quality branded products at a price of p_h^* to the remaining customers than selling them at a price of p_l^* and capturing market share. Formally, T_m is the unique value of T such that:

$$\begin{aligned} \pi^{T_m-1}(p_h^* - c_l)(1 - \rho) + \pi^{2(T_m-1)}(p_h^* - c_l)\rho &\geq \\ p_l^* - c_l &> \pi^{T_m}(p_h^* - c_l)(1 - \rho) + \pi^{2T_m}(p_h^* - c_l)\rho. \end{aligned} \quad (5)$$

Expression (4) can be rewritten as

$$\begin{aligned} \left(\frac{p_h^* - c_h}{1 - \delta} \right) &\geq \left(\frac{p_h^* - c_l}{1 - \delta\pi} \right) (1 - \delta^{T_m}\pi^{T_m})(1 - \rho) \\ &+ \left(\frac{p_h^* - c_l}{1 - \delta\pi^2} \right) (1 - \delta^{T_m}\pi^{2T_m})\rho + \delta^{T_m} \left(\frac{p_l^* - c_l}{1 - \delta} \right) \end{aligned} \quad (6)$$

Note that equation (6) holds for a strictly greater range of discount factors than equation (3). That is, when consumers learn only from their own private experiences, then selling multiple products makes the firm's incentive constraint easier to satisfy. Under perfect public monitoring, however, selling multiple products has no effect on the firm's incentive constraint. Our analysis of imperfect private monitoring with multi-product firms is consistent with Jin and Leslie (2009), who find that restaurants that belong to chains have higher cleanliness ratings than restaurants that are not members of chains. According to our theory, consumers learn faster about the cleanliness of the the chain after dining in multiple locations, so high quality is easier to support in equilibrium.

Bundling

Next, suppose the branded firm chooses to bundle the two products together. Specifically, suppose the firm requires every consumer to buy both goods or neither. That is, we are considering a world of pure bundling, not mixed bundling.

Recall that if quality were observable to consumers, the branded firm would charge $2p_h^* = 2\hat{p}(x)$ for a bundle of high quality products, thereby limiting sales to consumers who have preference for both branded products ($z_a = z_b = x$). Our assumption that $x > (1 - \pi)h$ implies that this is a more profitable strategy for the branded firm than charging $\hat{p}_h(x) + \hat{p}_h(-x)$ for the bundle and broadening the market to include buyers who value just one of the branded products and not the other.

High quality can be supported as a perfect Bayesian equilibrium as long as the present discounted value of all future profits on the equilibrium path exceeds

the profit associated with shirking on both products simultaneously and selling a low quality bundle thereafter. The most profitable deviation the firm could make is to lower the quality but maintain a high price, p_h^* , for a certain number of periods. After sufficiently many consumers have detected low quality (and expect low quality in the future) the firm would lower its price to p_l^* . By analogy to the previous section, the incentive constraint can be written as

$$\sum_{t=1}^{\infty} \delta^{t-1} 2(p_h^* - c_h)(\rho/2) \geq \max_T \left[\sum_{t=1}^T \delta^{t-1} \pi^{2(t-1)} 2(p_h^* - c_l)(\rho/2) + \sum_{t=T+1}^{\infty} \delta^{t-1} 2(p_l^* - c_l)(\rho/2) \right]. \quad (7)$$

At a price of $2p_h^*$ the firm earns a profit of $2(p_h^* - c_h)$ on a mass of $\rho/2$ customers. Let T_b denote the last year in which the firm makes more money selling at a price p_h^* then selling at a price p_l^* . Formally, T_b is the unique value of T such that:

$$\pi^{2(T_b-1)}(p_h^* - c_l) \geq (p_l^* - c_l) > \pi^{2T_b}(p_h^* - c_l). \quad (8)$$

Note that with bundling, the probability that a consumer will detect that the firm shirked is $(1 - \pi)$ for each product consumed. This implies that only π^2 consumers fail to detect shirking for either product. Therefore the detection rate is higher for consumers that buy both products. In contrast to the case in which multi-product firms do not bundle, all customers here experience both products. It follows that the average detection rate is higher for a multi-product firm that bundles than one that does not.

The incentive compatibility constraint (7) can be rewritten as

$$\left(\frac{p_h^* - c_h}{1 - \delta}\right) \geq \left(\frac{p_h^* - c_l}{1 - \delta\pi^2}\right) (1 - \delta^{T_b}\pi^{2T_b}) + \delta^{T_b} \left(\frac{p_l^* - c_l}{1 - \delta}\right). \quad (9)$$

Again, notice that the bundling would have no effect on the incentive constraint under public monitoring, but makes the incentive constraint easier to satisfy under private monitoring.

Comparison

In all three cases considered above, the most profitable deviation for the branded firm (if one exists at all) involves the branded firm lowering quality while initially maintaining high prices. After sufficiently many consumers have detected the quality reduction and are no longer purchasing the product, the branded firm lowers its price to regain market share. The three cases differ, however, in how quickly the consumers, as a group, detect the reduction in quality and (consequently) in the length of time in which the branded firm benefits from tricking the unsuspecting consumers. This is very important, since the rate at which consumers learn is a disciplinary device in the market for experience goods.

Comparing (2), (5), and (8), it is clear that $T_s \geq T_m \geq T_b$ and that in the absence of integer constraints on T_s , T_m , and T_b , these inequalities would be strict. The time at which the branded firm lowers its price comes latest in the single-product case. This is due to the fact that the consumers, as a group, learn about reductions in quality the slowest when a branded firm produces a single product. The multi-product branded firm selling unbundled products is an intermediate

case. In that case, consumers learn about defections faster since a group of them experience more than one product from the same firm. When the multi-product firm bundles the products together, the average rate of learning is faster still. It follows that the firm lowers its price earliest when the products are bundled.

The fact that consumers, as a group, learn that the branded firm shirked and cut quality sooner when the products are bundled creates a stronger incentive for the branded firm to maintain high quality. Comparing the no-shirking constraints, i.e., (9) to (6) and (3), it is clear that the incentive to shirk is strictly greatest for the single-product firm and strictly weakest for the bundling firm. The long run profitability of a deviation is the same in each case, but the short run gains are highest when consumers learn the slowest, and are smallest when consumers learn the quickest. Hence, bundling makes it easier to support high quality.

More specifically, let δ_s be the value of δ for which (3) holds with equality, δ_m be the value of δ for which this (6) holds with equality, and δ_b be the value of δ for which (9) holds with equality, and note that lowering price increases each of these values, so we are not giving up anything by focusing on the most profitable price. Then $\delta_b < \delta_m < \delta_s$ as long as $\rho < 1$, that is, the minimum discount rate needed to sustain high quality is lowest when the firm bundles.²⁶

We can state this formally in the following proposition.

Proposition 1 *High quality is easiest to sustain by a firm that bundles its products, harder to sustain by a firm that sells multiple products, and hardest to sustain by a firm that sells only one product. More precisely, $\delta_b \leq \delta_m \leq \delta_s$.*

²⁶When $\rho = 1$ all consumers prefer to bundle their purchases so there is no free rider problem.

PROOF: From (2), (5), and (8) it follows that δ_s , δ_m , and δ_b are uniquely defined by

$$\sum_{t=1}^{\infty} \delta_s^{t-1} (p_h^* - c_h) = \sum_{t=1}^{T_s} \delta_s^{t-1} \pi^{t-1} (p_h^* - c_l) + \sum_{t=T_s+1}^{\infty} \delta_s^{t-1} (p_l^* - c_l), \quad (10)$$

$$\sum_{t=1}^{\infty} \delta_m^{t-1} (p_h^* - c_h) = \sum_{t=1}^{T_m} \delta_m^{t-1} (p_h^* - c_l) [\pi^{t-1} (1 - \rho) + \pi^{2(t-1)} \rho] + \sum_{t=T_m+1}^{\infty} \delta_m^{t-1} (p_l^* - c_l), \quad (11)$$

and

$$\sum_{t=1}^{\infty} \delta_b^{t-1} (p_h^* - c_h) = \sum_{t=1}^{T_b} \delta_b^{t-1} \pi^{2(t-1)} (p_h^* - c_l) + \sum_{t=T_b+1}^{\infty} \delta_b^{t-1} (p_l^* - c_l), \quad (12)$$

where T_s , T_m , and T_b are defined by (2), (5), and (8) respectively and, equivalently, maximize the right-hand side of the above equations. Ignoring integer constraints, we have that $T_s > T_m > T_b$. Note that the definitions of δ_s , δ_m , and δ_b are unique because the both the left and right-hand sides of each equation above are monotonically increasing in δ , but the right-hand sides all increase at a slower rate.

Suppose first that $\delta_s = \delta_m = \delta_b = \delta$. Since the left-hand sides of (10), (11), and (12) are equal the right-hand sides must be equal as well. However, we will now show that the right-hand sides cannot be equal, and that the differences between them establish the rankings of the threshold discount factors.

First compare the right-hand sides of the first two incentive compatibility constraints. Since T_s maximizes the right-hand side of (10), we know that the right-hand side of (10) is greater than the same expression evaluated at T_m , or

$$\sum_{t=1}^{T_m} \delta^{t-1} \pi^{t-1} (p_h^* - c_l) + \sum_{t=T_m+1}^{\infty} \delta^{t-1} (p_l^* - c_l),$$

and we now show that this is greater than or equal to than the right-hand side of (11). Note that last term of this expression is equal to the last term on the right-hand side of (11), so it is sufficient to show that

$$\sum_{t=1}^{T_m} \delta^{t-1} \pi^{t-1} > \sum_{t=1}^{T_m} \delta^{t-1} [\pi^{t-1} (1 - \rho) + \pi^{2(t-1)} \rho].$$

Rearranging terms, this becomes

$$\sum_{t=1}^{T_m} \delta^{t-1} \pi^{t-1} [1 - \pi^{t-1}] \rho > 0,$$

which is clearly true. Therefore the branded firm's payoff from lowering its quality is strictly larger if it is a single-product firm than if it is a multi-product firm. It follows that the threshold discount factor must be higher as well in order to maintain incentive compatibility, $\delta_s > \delta_m$.

Similarly, since T_m is chosen by the branded multi-product firm to maximize its deviation payoff, we know that the right-hand side of (11) is greater than or equal to its value evaluated at T_b ,

$$\sum_{t=1}^{T_b} \delta^{t-1} (p_h^* - c_l) [\pi^{t-1} (1 - \rho) + \pi^{2(t-1)} \rho] + \sum_{t=T_b+1}^{\infty} \delta^{t-1} (p_l^* - c_l),$$

and we will now show that this is greater than the right-hand side of (12). The last term of this expression is equal to the last term of (12), so it is sufficient to

show that

$$\sum_{t=1}^{T_b} \delta^{t-1} [\pi^{t-1}(1-\rho) + \pi^{2(t-1)}\rho] > \sum_{t=1}^{T_b} \delta^{t-1} \pi^{2(t-1)}.$$

Rearranging terms establishes that this is true,

$$\sum_{t=1}^{T_b} \delta^{t-1} \pi^{t-1} (1-\rho)(1-\pi^{t-1}) > 0.$$

So the deviation payoff for a multi-product branded firm's deviation payoff is higher when it does not bundle than when it does bundle. It follows that δ_m must be greater than δ_b as well.

Finally, since T_s , T_m , and T_b are constrained to be integers, they need not be strictly decreasing. And if they aren't strictly decreasing then the threshold discount rates will not be strictly increasing. So given the integer constraints, we can only conclude $\delta_s \geq \delta_m \geq \delta_b$. ■

The fact that high-quality equilibria are easier to sustain when the firm bundles its products together does not by itself imply that the firm will choose to bundle its products. First, the theory presented so far doesn't tell us which equilibrium will be selected among the many possibilities. Whether the firm bundles or not, there will always exist a perfect Bayesian equilibrium where consumers expect low quality branded products and the branded firm has no incentive to deliver anything better than that. If consumers expect low quality whether or not the firm bundles, then the firm would strictly prefer not to bundle because it restricts the size of the market.

Second, the branded firm might prefer an equilibrium with low quality unbundled products to an equilibrium with high quality bundled products. Although

the bundled products have higher profit margins, the market is smaller (since consumers who only prefer one of the branded firm's products to the products offered by the competitive fringe choose to purchase both goods from fringe). Formally, bundling increases the set of feasible profitable if and only if

$$\sum_{t=1}^{\infty} \delta^{t-1} 2(p_h^* - c_h) d_{hh}(2p_h^*) \geq \sum_{t=1}^{\infty} \delta^{t-1} (p_l^* - c_l) (d_l^a(p_l^*) + d_l^b(p_l^*)). \quad (13)$$

Since $d_{hh}(2p_h^*) = \rho/2$ and $d_l^a(p_l^*) + d_l^b(p_l^*) = 1/2 + 1/2 = 1$, we can rearrange the expression to establish the following result:

Proposition 2 *A necessary and sufficient condition for bundling to increase the highest profit sustainable by the branded firm in a PBE is $\delta_b \leq \delta < \delta_m$ and*

$$(p_h^* - c_h)\rho \geq (p_l^* - c_l). \quad (14)$$

Although high quality products are more efficient than low quality products, i.e., $(p_h^* - c_h) \geq (p_l^* - c_l)$, Proposition 2 shows that bundling won't be attractive unless the gains to quality are sufficient to offset the loss in demand, $(1 - \rho)$, associated with bundling. Moreover, it is not sufficient to let $\rho = 1$ since in that case bundling has no impact the incentive compatibility constraint, i.e. $\delta_b = \delta_m$.

5 Reputation Externalities

We now consider the case in which consumers cannot separately observe the quality of the two products. For example, when a consumer's internet service is disrupted, he or she may not know whether the modem or the service provider

is responsible. Similarly, if their printer doesn't work, the consumer may not know if the printer or the ink cartridge is to blame. The consequence of this assumption is that if consumers expect high quality from the branded firm, but low quality from the competitive fringe, then if they purchase only one product from the branded firm and consume the second product from the competitive fringe, the consumer will always attribute bad performance to the competitive fringe as opposed to the branded firm.²⁷ This problem, which we call a *reputation externality*, means that the consumers purchasing just one product are even less likely to punish the firm. In the absence of product bundling, the branded firm's incentive to provide high quality is reduced.

Under our stylized assumptions, poor quality is blamed on the rival's product, not the branded firm's product. This doesn't actually harm rival's reputation because the rival is already known to sell low quality products. In a more general model, the inability to attribute bad performance to products implies that shirking by one firm affects both its reputation and other firms' reputations. That is, each firm bears only part of the cost of shirking. Bundling will still be attractive because it internalizes this reputation externality.

The model in this section is the same as in the previous section, but with one change. Although the consumer knows when a product failure occurs and the magnitude of the harms that he or she has suffered, he or she cannot directly observe which of the two products has caused harm. Based on his prior beliefs and on the level of harm suffered, h or $2h$, the consumer forms Bayesian inferences about the products' quality. If the consumer suffers harm $2h$ he or she knows

²⁷This is because the competitive fringe produces a low quality product while on the equilibrium path the branded firm produces a high quality good.

that both products have failed, and will invariably deduce that both products are of low quality. But if the consumer suffers harm h , evidence that a single product failed, the consumer's prior beliefs will be critical in the assignment of blame. We will show that consumers, as a group, learn more slowly in this environment, making it more difficult to sustain equilibria with high quality products.

Single-Product Firms

Suppose that the branded firm produces one product only, and that the consumer purchases a second product from the competitive fringe. Imagine an equilibrium in which the branded firm produces high quality products. Product failures certainly occur on the equilibrium path, since the unbranded product is of low quality. If a consumer suffers harm h , they would rationally attribute the problem to the competitive fringe. Consumers would detect a deviation by the branded firm only when both products fail. In this case the harm is $2h$ which is indisputable evidence that the branded product is of low quality. Following a deviation by the branded firm, the proportion of remaining customers would be $\pi_0 = 1 - (1 - \pi)^2$ (i.e., the probability that there is at most one product failure, or equivalently, that both low quality products don't fail).

Formally, incentive compatibility for the branded firm requires

$$\sum_{t=1}^{\infty} \delta^{t-1} (p_h^* - c_h)(1/2) \geq \sum_{t=1}^{\hat{T}_s} \delta^{t-1} \pi_0^{t-1} (p_h^* - c_l)(1/2) + \sum_{t=\hat{T}_s+1}^{\infty} \delta^{t-1} (p_l^* - c_l)(1/2), \quad (15)$$

where \hat{T}_s is the last year in which the firm makes more money selling at a price p_h^* than selling at a price p_l^* , that is \hat{T}_s is the unique value of T such that

$$\pi_0^{\hat{T}_s-1}(p_h^* - c_l) \geq p_l^* - c_l > \pi_0^{\hat{T}_s}(p_h^* - c_l). \quad (16)$$

The main difference between (15) and our earlier incentive compatibility constraint (1) appears in the first term on the right-hand side, where π_0 appears instead of π . This term represents the branded firm's profits from tricking the remaining unsuspecting consumers, charging a high price for a low quality product. Before, a consumer would shift to the unbranded fringe product after observing the failure of the branded product, leaving a proportion π of consumers remaining. Now, the consumer switches to the unbranded fringe only after observing the failure of both products, leaving a proportion of $\pi_0 = 1 - (1 - \pi)^2 > \pi$.

Expression (15) can be rewritten as

$$\left(\frac{p_h^* - c_h}{1 - \delta} \right) \geq \left(\frac{p_h^* - c_l}{1 - \delta\pi_0} \right) (1 - \delta^{\hat{T}_s} \pi_0^{\hat{T}_s}) + \delta^{\hat{T}_s} \left(\frac{p_l^* - c_l}{1 - \delta} \right). \quad (17)$$

Multi-Product Firms

When the firm produces branded products, then of course consumers who choose to buy both products from the branded firm will attribute any harm suffered to the branded firm. In equilibrium, consumers who buy only one product from the branded firm will necessarily attribute a single-product failure to the unbranded product. As in the single product case, these consumer would have to experience two failures simultaneously for the consumer to know that a deviation occurred.

The firm's incentive constraint can be written

$$\begin{aligned}
\sum_{t=1}^{\infty} \delta^{t-1} (p_h^* - c_h) (1/2 + 1/2) \geq \\
\sum_{t=1}^{\hat{T}_m} \delta^{t-1} \pi_0^{t-1} (p_h^* - c_l) (1 - \rho) + \sum_{t=1}^{\hat{T}_m} \delta^{t-1} \pi^{2(t-1)} 2 (p_h^* - c_l) (\rho/2) \\
+ \sum_{t=\hat{T}_m+1}^{\infty} \delta^{t-1} (p_l^* - c_l) (1/2 + 1/2), \quad (18)
\end{aligned}$$

where \hat{T}_m is the last year in which the firm makes more money selling at a price p_h^* then selling at a price p_l^* , or equivalently, \hat{T}_m is the unique value of T such that

$$\begin{aligned}
\pi_0^{\hat{T}_m-1} (p_h^* - c_l) (1 - \rho) + \pi^{2(\hat{T}_m-1)} (p_h^* - c_l) \rho \geq \\
p_l^* - c_l > \pi_0^{\hat{T}_m} (p_h^* - c_l) (1 - \rho) + \pi^{2\hat{T}_m} (p_h^* - c_l) \rho. \quad (19)
\end{aligned}$$

The incentive compatibility constraint differs from the earlier constraint, (4), in the first term on the right-hand side. This term reflects continued sales following a deviation to consumers who purchase only one branded product. As just discussed, the proportion remaining in each round is simply a lagged proportion, $\pi_0 = 1 - (1 - \pi)^2$, of the previous demand. Equation (18) may be rewritten as

$$\begin{aligned}
\left(\frac{p_h^* - c_h}{1 - \delta} \right) \geq \left(\frac{p_h^* - c_l}{1 - \delta \pi_0} \right) (1 - \delta^{\hat{T}_m} \pi_0^{\hat{T}_m}) (1 - \rho) \\
+ \left(\frac{p_h^* - c_l}{1 - \delta \pi^2} \right) (1 - \delta^{\hat{T}_m} \pi^{2\hat{T}_m}) \rho + \delta^{\hat{T}_m} \left(\frac{p_l^* - c_l}{1 - \delta} \right). \quad (20)
\end{aligned}$$

Bundling

Finally, suppose the branded firm bundles its products, requiring every consumer to buy both goods or neither. In this case, there is no problem with attribution. Consumers who buy the bundle would know that the branded firm deviated after observing the failure of one or both products. In this case, the incentive constraint is exactly the same as before in expression (7), or

$$\sum_{t=1}^{\infty} \delta^{t-1} 2(p_h^* - c_h)(\rho/2) \geq \sum_{t=1}^{\hat{T}_b} \delta^{t-1} \pi^{2(t-1)} 2(p_h^* - c_l)(\rho/2) + \sum_{t=\hat{T}_b+1}^{\infty} \delta^{t-1} 2(p_l^* - c_l)(\rho/2), \quad (21)$$

where \hat{T}_b is the last year in which the firm makes more money selling at a price p_h^* then selling at a price p_l^* , or equivalently, \hat{T}_b the unique value of T such that

$$\pi^{2(\hat{T}_b-1)}(p_h^* - c_l) \geq (p_l^* - c_l) > \pi^{2\hat{T}_b}(p_h^* - c_l). \quad (22)$$

The incentive compatibility constraint, (21), can be rewritten as

$$\left(\frac{p_h^* - c_h}{1 - \delta} \right) \geq \left(\frac{p_h^* - c_l}{1 - \delta\pi^2} \right) (1 - \delta^{\hat{T}_b} \pi^2) + \delta^{\hat{T}_b} \left(\frac{p_l^* - c_l}{1 - \delta} \right). \quad (23)$$

Note that this is the same as (9). That is, bundling completely eliminates the imperfect attribution problem by eliminating the externality. Any product failure is blamed on the branded firm, regardless of whether the consumer knows which product failed.

Comparison

Comparing (16), (19), and (22), it is again clear that, in the absence of integer constraints, $\hat{T}_s > \hat{T}_m > \hat{T}_b$. Consumers as a group learn about product deviations most slowly when the branded firm sells a single product, more quickly when the branded firm sells multiple products, and fastest when the products are bundled. The attribution problem is the most severe with single-product firms, since consumers attribute single-product failures the unbranded product provided by the competitive fringe. When the products are bundled, there is no problem with attribution – when a consumer observes a failure of the bundle, he or she knows to attribute the harm to the branded firm.

Comparing the right-hand sides of the incentive compatibility constraints, (21), (18), and (15), the branded firm's constraint is weakest for the single-product firm and strongest for the bundling firm. A single-product firm can exploit customers more easily, especially given their tendency to attribute product failures to the fringe. With bundling, consumers learn about quality faster and so the payoff from cheating is lower. This provides greater discipline for the branded firm, making it easier to sustain high quality. More specifically, let $\hat{\delta}_s$ be the value of δ for which (15) holds with equality, $\hat{\delta}_m$ be the value of δ for which this (18) holds with equality, and $\hat{\delta}_b$ be the value of δ for which this (21) holds with equality. Then we have the following result.

Proposition 3 *In the presence of reputation externalities, high quality is easiest to sustain by a firm that bundles its products, harder to sustain by a firm that sells multiple products, and hardest to sustain by a firm that sells only one product. More precisely, $\hat{\delta}_b \leq \hat{\delta}_m \leq \hat{\delta}_s$ when $\rho < 1$.*

As before, the ranking of the discount rates is strict if $\hat{T}_s > \hat{T}_m > \hat{T}_b$, but given integer constraints, the switching times and discount rates may be equal.

The fact that bundling may be necessary to sustain high quality does not imply that the firm will necessarily bundle. This is because bundling has two effects. First, as we have just emphasized, bundling allows the firm to sustain high quality and earn a strictly higher margin. But, second, at the profit-maximizing price of $P = 2p_h^*$, bundling decreases the demand for the firms products by turning away consumers who would have bought just one of the firms products, so fewer consumers buy from the firm. A necessary condition is that it is not possible to sustain high quality otherwise and that multi-product firm would derive higher profits from a high-quality equilibrium with bundled products than a low-quality equilibrium with unbundled products. Formally,

$$\sum_{t=1}^{\infty} \delta^{t-1} 2(p_h^* - c_h) d_{hh}(2p_h^*) \geq \sum_{t=1}^{\infty} \delta^{t-1} (p_i^* - c_i) (d_i^a(p_i^*) + d_i^b(p_i^*)), \quad (24)$$

Using the fact that $d_{hh}(2p_h^*) = \rho/2$ and $d_i^a(p_i^*) + d_i^b(p_i^*) = 1/2 + 1/2 = 1$, we can state the following result.

Proposition 4 *In the presence of reputation externalities, a necessary and sufficient condition for bundling to be a profitable strategy for the branded firm is*

$$\hat{\delta}_b \leq \delta < \hat{\delta}_m \text{ and}$$

$$(p_h^* - c_h)\rho \geq p_i^* - c_i. \quad (25)$$

Finally, comparing the incentive compatibility constraints to those in the previous section, it is clear that imperfect attribution strictly increases the firm's incentive to bundle. Specifically,

Proposition 5 *Absent bundling, reputation externalities make it more difficult to sustain high quality. More precisely, $\delta_s > \hat{\delta}_s$ and $\delta_m > \hat{\delta}_m$.*

6 Conclusion

This paper extends Klein and Leffler's (1981) repeated moral hazard model of product quality to the case of multiproduct firms. In our model, consumers are small, heterogeneous, and receive imperfect private signals about the quality of their past purchases. Through bundling, the multiproduct firm constrains consumers to gather better information about the firm's product quality decisions. Using this information, consumers are collectively better able to punish the firm when it sells low quality products. This in turn gives the firm a stronger incentive to only produce high quality products. Intuitively, the multiproduct firm will rationally adopt a bundling strategy only when bundling enables the firm to offer high quality and the resulting loss of market share is offset by the higher margins that high quality products command.

The private and social benefits of bundling are even higher when consumers are unable to attribute a bad overall experience to the particular product that failed. In the absence of product bundling, many consumers purchase only one product from the multiproduct firm and don't hold that firm to blame when they have a bad experience. As a consequence of these free riders, and the associated reputation externality, shirking becomes even more attractive to the multiproduct firm. With bundling, although consumers cannot directly verify which product failed, they can nevertheless easily deduce that the branded firm is to blame.

The model and the results rely on several key assumptions. First, we assumed that consumers were small and had heterogeneous preferences. If there were a single consumer, that consumer would internalize the impact of his or her behavior on the firm's quality choice and bundling would be unnecessary. Second, we assumed that consumers received imperfect private signals of product quality. If consumers received perfect private signals instead, then the quality of monitoring would not depend on whether or not the consumer purchased both products.²⁸ Finally, if the signals were publicly observed instead of privately observed, then the consumers' information would be independent of their private purchase decisions and bundling would be ineffective.

While our model focused on the private and social desirability of pure bundling, the framework could be extended to consider the role of mixed bundling. Recall that quality-enhancing benefit of pure bundling in our model came at a cost. With bundling, consumers who placed high value on one branded product were inefficiently priced out of the market. The branded firm can potentially mitigate this cost through a mixed-bundling strategy, offering a limited number of individual units in addition to the bundled units. By rationing individual unit sales, the branded firm can keep the proportion of consumers buying multiple units just high enough to satisfy its incentive constraint. More generally, with a continuum of consumer types, the firm might offer a pure bundle discount in order to increase the proportion of consumers who purchase multiple products.²⁹

Our model had that feature that, absent quality assurance concerns, the

²⁸The conditions for sustaining high quality in equilibrium would be the same for a single-product firm, a multi-product firm, and a multi-product firm that bundles.

²⁹However in a model with such heterogeneity, the firm would also be able to use bundling to price discriminate, making the analysis much more complicated.

branded firm would never find it profitable to bundle its products together. Bundling is only privately and socially attractive when the branded firm cannot otherwise maintain a reputation for high quality. In reality, there are other benefits that that the branded firm can derive by bundling, such as price discrimination and market foreclosure. In a more general model, bundling might increase product quality but also deter entry by potential competitors.³⁰ The social costs of these anticompetitive effects would need to be weighed against the social benefits of quality improvements. Theoretical and empirical research combining these effects may be fruitful avenues for future research.³¹

³⁰More generally, one could explore how competition among branded firms changes the analysis. Intuitively, the mix-and-match behavior we analyze here should occur more easily when consumers are choosing among multiple high quality firms, however the analysis of the incentive constraints, particularly with reputation externalities, would be considerably more complicated.

³¹One implication is that empirically demonstrating bundling improves quality would be suggestive, but not conclusive, evidence the bundling is efficient.

References

- [1] Aaker, D. A., and K. L. Keller (1990) "Consumer Evaluations of Brand Extensions," *Journal of Marketing*, 54(1), 27-41.
- [2] Adams, W.J. and Yellen, J.L. (1976) "Commodity Bundling and the Burden of Monopoly," *Quarterly Journal of Economics*, 90, 475-498.
- [3] Andersson, F. (2002) "Pooling Reputations," *International Journal of Industrial Organization*, 20(5), 715-730.
- [4] Bar-Gill, Oren (2006) "Bundling and Consumer Misperception," *The University of Chicago Law Review*, 73, 33-61.
- [5] Bernheim, Douglas and Michael Whinston (1990) "Multimarket Contact and Collusive Behavior," *RAND Journal of Economics*, 21, 1-26.
- [6] Cabral, L. M. B. (2000) "Stretching Firm and Brand Reputation," *RAND Journal of Economics*, 31(4), 658-673.
- [7] Cabral, L. (2009) "Umbrella Branding with Imperfect Observability and Moral Hazard," *International Journal of Industrial Organization*, 27, 206-213.
- [8] Cai, H. and I. Obara (2006) "Firm Reputation and Horizontal Integration" UCLA working paper.
- [9] Choi, J. P. (1998) "Brand Extension as Information Leverage," *Review of Economic Studies*, pp. 655-670.
- [10] Erdem, T. (1998) "An Empirical Analysis of Umbrella Branding," *Journal of Marketing Research*, 35(3), 339-351.
- [11] Fang and Norman (2006) "To Bundle or Not to Bundle," *RAND Journal of Economics*, Winter.
- [12] Hakenes, H. and M. Peitz, (2008) "Umbrella Branding and the Provision of Quality," *International Journal of Industrial Organization*, 26: 546-556.
- [13] Iacobucci, Edward M. (2003) "Tying as Quality Control: A Legal and Economic Analysis," *Journal of Legal Studies*, 32, 435-464.

- [14] Jin, G. and P. Leslie (2003): “The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards,” *Quarterly Journal of Economics*, 118(2), 409-51.
- [15] Jin, Ginger Zhe, and Phillip Leslie. 2009. “Reputational Incentives for Restaurant Hygiene,” *American Economic Journal: Microeconomics*, 1(1): 237-67.
- [16] Klein, B., and K. Leffler (1981) “The Role of Market Forces in Assuring Contractual Performance,” *Journal of Political Economy*, 89(4), 615-641.
- [17] Milgrom, P. R., and J. Roberts (1986) “Price and Advertising Signals of Product Quality,” *Journal of Political Economy*, 94, 796-821.
- [18] Montgomery, C. A., and B. Wernerfelt (1992) “Risk Reduction and Umbrella Branding,” *Journal of Business*, 65(1), 31-50.
- [19] Myatt, D. P., and Eric B. Rasmusen (2009) “How Useful Is a Product Umbrella for Reputation?,” working paper.
- [20] Nalebuff, B. (2004) “Bundling as an Entry Barrier,” *Quarterly Journal of Economics*, 119(1), 159-187.
- [21] Pepall, L. M., and D. Richards (2002) “The Simple Economics of Brand Stretching,” *Journal of Business*, 75, 535-552.
- [22] Salinger, Michael A. (1995) “A Graphical Analysis of Bundling,” *Journal of Business*, 68, 85-98.
- [23] Wernerfelt, B. (1988) “Umbrella Branding as a Signal of New Product Quality,” *RAND Journal of Economics*, 19(3), 458-466.
- [24] Whinston, Michael (1990) “Tying, Foreclosure, and Exclusion,” *American Economic Review*, 80(4), 837-859.