Working Paper Series ES
Economics

“Bundling as a Way to Leverage Monopoly”

Barry Nalebuff, Yale School of Management, National Bureau of Economic Research

Working Paper # 36

This paper can be downloaded without charge from the Social Science Research Network Electronic Paper Collection:
http://ssrn.com/abstract=586648
BUNDLING AS A WAY TO LEVERAGE MONOPOLY*

by Barry Nalebuff

ABSTRACT: This paper shows how a monopolist generally can increase its profits by offering a discount on its monopolized product if the customer agrees to buy a competitively supplied good from it at a price premium.

The use of bundling to leverage market power has a long (and checkered) history in law and economics. The Chicago School seemed to end the debate with their result that there is only one monopoly profit and thus there is no gain from bundling. This “folk theorem” relies on some special assumptions, most importantly that the goods are consumed in fixed proportions. Once we allow for continuous consumption levels, then it is generally the case that a firm can extend a monopoly from A into a competitive B market.

While it is well understood how a monopolist can use tying to extract more consumer surplus or to engage in price discrimination, this paper pursues a different motivation. The emphasis of this paper is on optional, as opposed to forced, tied sales. The firm offers to scale back its monopoly price in return for getting a price premium in a second market. The reduction in monopoly price causes no first-order loss to the firm, while providing a first-order incentive for customers to voluntarily accept the deal.

The ability of a monopolist to extend its influence to adjacent markets is a challenge both to the competitors in those markets and to economists looking to understand the antitrust implications of bundling.

* I am very grateful for the feedback given by Ben Polak. Contact information: Barry Nalebuff, Yale School of Management, Box 208200, New Haven, CT 06520-8200, barry.nalebuff@yale.edu.
BUNDLING AS A WAY TO LEVERAGE MONOPOLY

BARRY NALEBUFF

I. INTRODUCTION

To the layperson, it would seem to make perfect sense that a firm with a monopoly in good A could use that power to extend its market power into another market, B. The law might prevent such an extension of market power, but absent such a restraint, a monopolist would have an incentive to do so.

Thus, it was surprising when the Chicago School provided an argument that seemed to show that leveraging monopoly into a competitive market provided no advantage.¹ As Bork (1995) famously wrote: “There is only one monopoly profit.”

The point of this paper is to show that the layperson is mostly right, after all. While there are some special cases in which leverage does not lead to higher profits, in the general case, a monopolist can earn higher profits by leveraging its power into a competitive market.

The idea that a monopolist might increase its profits by selling its monopoly good A only as part of a package with a competitive good B is well known (Burstein (1960)). The monopolist uses a markup on the competitive good to recover some of the consumer surplus left on the table at the monopoly price. But, as tied sales by a monopolist are a violation of §1 of the Sherman Act and §3 of the Clayton Act, such naked ties are rare.²

¹ In a nutshell, the Chicago School argument (Director and Levi (1956)) goes as follows: Imagine that the monopoly price of good A on its own is m, and the competitive price of good B is c. If the monopolist were to earn higher profits at price b for a bundle of A and B, then consider the implied monopoly price \( m' = b - c \). Since good B is available at c, anyone who buys the bundle is willing to pay an incremental price of \( b - c \) for A. Were the monopolist to charge \( b - c \) for A alone and eliminate the bundle, its demand and, hence, its profits would be at least as large (as there may be some consumers who do not value good B even at its cost c). This suggests that bundling does not lead to higher profits.

² As the Supreme Court held in International Salt Co. v. United States, 322 U.S. 392, 396 (1947): “It is violative per se of §1 of the Sherman Act and §3 of the Clayton Act for a corporation engaged in interstate commerce in salt, of which it is the country's largest producer for industrial uses, and which also owns
This paper is concerned with a more subtle form of tying or bundled sales. Here, the monopolist offers a tied sale as an option. Of course, if the untied good is priced arbitrarily high, then for practical purposes, there is only one option, namely the tied sale. Thus, we require the monopolist to continue to offer its good A at the optimal à la carte price, assuming no bundling is possible. Even with the à la carte prices available, a monopolist will generally be able to increase its profits via an optional tied offering.

Imagine that the monopoly price of good A is m, and the competitive price of good B is c. The monopolist offers its customers the following deal: I’ll sell you A at a discount, say m-ε, in return for which you agree to buy all of your good B from me at a premium, say c+δ. Or, you can still buy good A from me at m.

The basic intuition for the advantage of bundling comes from the envelope theorem. From the monopolist’s perspective, changing the price of m, either way, has no first-order effect. The monopolist doesn’t really care if it charges m or m-ε. In contrast, consumers care a lot. They have a first-order gain from paying less than m for the monopoly good and are willing to reward the monopolist by buying its B good, even at slight market premium. Thus, the change in the price of the A good is a first-order wash, while the change in the B good pricing leads to a first-order gain.

As this bundling arrangement will generally result in higher profits, it must depart from the framework implied by the Chicago School argument (see Director and Levi (1956)). Unlike the Chicago School, I do not assume that the consumer buys only one unit of A or B. Nor do I require that the bundle be sold in a fixed proportion: one unit of A with one

---

3 There is a semantic question of whether this strategy should be called bundling or tying. With mixed bundling, the monopolist offers A and B at stand-alone prices and a bundled price for A and B together. Under tying, customers can buy B on its own, but they cannot buy good A without also buying the firm’s good B. Here we introduce “mixed tying,” where A is available at an à la carte price and at a lower price when bought with the firm’s good B.
unit of B. Different customers may want to consume A and B in different ratios. As long as they buy all of their B from the monopolist, they can get the discount on A.

If all of the customer’s demand for good B must be purchased from the monopolist, this requires that the purchase of B can be monitored. For most business-to-business transactions, this is a reasonable assumption. In the recent LePage’s decision, 3M was found guilty of excluding its rival LePage’s from selling generic tape by offering a discount to superstores based on total volume of 3M purchases.\(^4\) If Staples contractually agrees to buy all of its transparent tape from 3M, Staples cannot surreptitiously buy generic tape from LePage’s. Business-to-business interactions is the arena where we are most concerned about the antitrust implications of bundling. In the Microsoft case, Apple was forced to pick Explorer as its default navigator as a condition for Microsoft upgrading Office. Mastercard and Visa required their member banks to issue only Mastercard and Visa and not American Express.

I will be more formal shortly, but let me first distinguish between two cases: (1) what I call the *forced* bundle and (II) the *optional* bundle. The results of this paper emphasize optional bundling. I make this distinction because it shows that the source of the gain is neither the extraction of consumer surplus nor price discrimination. If the monopolist can force the bundle on the consumer as the only way to buy A, then it can use the increase in the price of B as something akin to a two-part tariff. But if the monopolist simply makes the bundled pricing pair as another option, this limits its ability to extract surplus from the consumers. With optional bundling where the original à la carte prices remain available, any change is a Pareto improvement (at least in the short run).

In the forced bundle case, the monopolist replaces its offer of \((m, c)\) with \((m, c+\delta)\). Thus, a customer who wants to buy A has no other choice but to buy the monopolist’s overpriced B good. The argument for forced bundling was originally presented by Burstein (1960), who recognized that the consumer typically has some surplus when buying the monopoly good at \(m\). The monopolist can extract some of that surplus by

\(^4\) LePage’s Inc. v. 3M (Minnesota Mining and Manufacturing Company), 2003, 324 F.3d 141.
requiring the consumer to buy good B at a price of c+\delta along with its monopolized good A. Provided that the consumer is better off paying an extra \delta for B than not consuming any A at all, the monopolist will be able to increase its profits.

This point can be illustrated in a one-consumer economy. Let demand for good A be Q(p) = 1 – p. If the cost of A is zero, the monopoly price of A is \frac{1}{2}. This consumer has surplus of \frac{1}{4} \times \frac{1}{2} = 1/8. Thus, if the consumer has exogenous demand of B of 1, then the monopolist can tie the two goods together and charge up to \frac{1}{8} over the competitive price of B and still gain the customer’s purchase of A. (More generally, the optimal pricing of A and B are Ramsey prices, where the markup is inversely proportional to the demand elasticity; see Katz (1989) and Mathewson and Winter (1997).)

As Bowman (1957) and Friedman (1976) recognized, a forced tied sale can also be used as a price discrimination device. If A and B are used together, and those with a high value of A (printer) buy a greater amount of B (toner cartridges), then the firm can effectively engage in price discrimination via metering of the tied good.\textsuperscript{5} Bowman argued that metering via bundling is not really leveraging monopoly in that the same result could also be achieved via a per-use fee on good A directly. However, it is often impractical to measure output directly and, thus, the only way to meter is via a tied good. As a result, rival B firms are foreclosed from the market, and this may have a long-run effect on the competitive nature of that market.

The extraction of consumer surplus generally holds even with just one consumer type, while the advantages of metering is based on the heterogeneity of preferences. While forced bundling typically leads to higher profits, with identical consumers, the monopolist would do even better with the use of a two-part tariff. Mathewson and Winter (1997) show that even with a two-part tariff in place, the monopolist might also want to

\textsuperscript{5} Kaplow (1995) observes that firms charged with tying rarely attempt to defend this practice by appealing to price discrimination via metering. As he further points out, looking at the specific facts of the case may suggest that metering is not the explanation. For example, in the landmark International Salt case, the tying provision had an escape clause—buyers were not required to purchase from Internal Salt if they could find cheaper salt elsewhere (see 322 U.S at 394, n.5). This would have made it difficult to charge a premium price on salt and thereby extract more surplus from high-volume buyers.
employ a tied sales contract. Tying increases profits provided that “inframarginal consumers have a relatively more intense preference for [the tied] good than do marginal customers.” In particular, it then follows that tying does not increase profits when demand for the two goods is independent (and two-part pricing is possible).

A central difference between this paper and the earlier literature is that I require the monopolist to leave its profit-maximizing contract as an option for consumers. The case in which the monopolist effectively forces its customers to buy B at an inflated price is not the source of its increased profits in the present results. I demonstrate the advantage of selling an \((m-\epsilon, c+\delta)\) package \textit{even when \((m, c)\) is left on the table}. Since the original prices are still an option for all consumers, there is no sense in which they are forced into this deal. Consumers who choose the tying contract must be better off. Indeed, if it is profitable for the monopolist, then we have a Pareto-improving bundled price offer. The gain from tying comes from mitigating the inefficiency of monopoly pricing.

This is the main result: that there generally exists a Pareto-improving price offer for a monopolist in A facing competition in B. That Pareto-improving offer involves requiring customers who take this offer to buy all of their B from the monopolist at an inflated price.

The intuition for the result is similar to the approach taken by McAfee, McMillan, and Whinston (1989). They establish a general result on the optimality of bundled contracts when the demand for A and the demand for B are independent. Yet their result does not apply to the present case. Assume that good A is optimally priced at \(m\). Then the monopolist would generally increase its profits by offering an A-B bundle at a discount. To the first order, there is no change in profits in selling A at a price of \(m-\epsilon\) to all customers or just to customers of good B. In the latter case, the firm gives up \(\epsilon\) to customers who were already buying A and B, but that is just offset by extra sales to customers of B who were on the margin of buying A. The gain from bundling arises as the same discount also motivates A customers who were on the margin of buying B. While it is still true that additional B sales are created, these sales come at a price of \(p_b\) -
ε. Thus, if \( p_b = c \), the incremental sales would be at a loss and there is no advantage to bundling. McAfee \textit{et al.} shows the general advantage of bundled pricing for all cases except where \( B \) is sold in a competitive market.

There is also a strand of literature that shows the advantage of bundled sales in protecting market power. For example, Whinston (1990) shows how a commitment to bundling can deter entrants into a \( B \) market. The reason is that if the \( A \) monopolist can sell \( A \) only along with \( B \), then it will have to be very aggressive in discounting the \( A-B \) bundle to fend off attacks from equal or even superior \( B \) products. But if entry occurs, then the incumbent is worse off by bundling. Thus, if a competitive market for \( B \) already exists, then there is no advantage to bundling. Although the motivations for bundling are different in Carlton and Waldman (2002) and Nalebuff (2004a), there is no gain from bundling in the event that there is a competitive market for \( B \). No entry deterrence can be achieved, nor can the cost of entry be mitigated.

The literature on bundling as an entry deterrent emphasizes how bundling changes the incentive of entrants and/or of the incumbent’s response to entry. That is not a factor once there is a competitive market in \( B \) and hence changing incentives is not relevant to this paper. In that sense, this paper focuses on a different question: does a firm with a monopoly in good \( A \) have an incentive to use that monopoly to leverage its way into a competitive \( B \) market? The example below demonstrates the basic intuition for the result using a simple linear demand specification. The more general theorem follows.

II. \textbf{Example of Optional Bundling Leading to Higher Profits}

In the initial results below, I assume that a competitive market is characterized by perfect substitutes and production at constant marginal cost, with no fixed costs. Following the proofs, I come back to reexamine this definition of perfect competition. If perfect competition means no product differentiation and free entry, then there may be further advantages of leveraging a monopoly into a competitive market.
I first illustrate the leverage result with a simple example. Customer of type $\alpha$ has demand for A equal to:

$$D(\alpha, p) = \alpha - p.$$  

For simplicity, we assume that $\alpha$ is distributed uniformly over $[0, 1]$. Total demand at price $p$ is

$$D(p) = \int_p^1 (\alpha - p) d\alpha = (1 - p)^2 / 2.$$  

We further simplify the problem by assuming that the production cost of A and B are both zero. The monopoly price for A maximizes $p^*(1-p)^2/2$:

$$m = 1/3.$$  

Thus, two-thirds of the customers buy A, and among those who buy, their average demand is one-third.

As for good B, the competitive price is 0. The good B produced by the monopolist and its rivals are perfect substitutes for each other. For simplicity, we assume that the demand for good B is exogenous and equal to one-third.

Now we need to show that there is some combination of $(\varepsilon, \delta)$ such that the resulting demand among the consumers who prefer $(m-\varepsilon, c+\delta)$ to $(m, c)$ leads to higher profits.

If all consumers took the $(m-\varepsilon, c+\delta)$ deal, we’d be done. The extra price of $\delta$ would increase the firm’s profits in the B market by $\delta/3$, while the price cut of $\varepsilon$ in the A market would wash out with the increased demand it stimulates.

The problem is that the selection of customers who take this deal works against the firm. For small $\varepsilon$, a customer would take up the offer if and only if
Based on his purchases at prices \((m, c)\), the customer saves more money on A than it pays extra on B. It follows that if the customer is paying less, then the firm must be getting less. The source of the gain that offsets this loss is that the customer also buys a bit more A, and this extra profit more than covers the reduction in price on the existing bundle.

It is clear that those who buy the most A have the greatest incentive to take up the offer. Thus, the firm will be giving up more than its average revenue while only getting the average incremental sales (as the marginal increase in demand is the same for all customers). Of course, the degree of adverse selection depends on the relative size of \(\varepsilon\) and \(\delta\).

In the present case, consider the further simplification of \(\delta = \varepsilon\). The offer will be accepted by those whose demand for A exceeds 1/3—namely, customers with \(\alpha \geq 2/3\). On average, these customers buy 1/2 (the average of 2/3 for \(\alpha=1\) and 1/3 for \(\alpha = 2/3\)), and so the cost of the bundle deal is

\[
(1/3)*[-1/2 \varepsilon + 1/3 \varepsilon + m * \varepsilon] = (1/3)*(-1/6 + 1/3) \varepsilon = \varepsilon/18 > 0.
\]

One-third of the customers take up the deal. These customers save \(\varepsilon/6\), on average, based on the price reduction on A net of the offset on B. But their demand for A increases by \(\varepsilon\). With \(p=1/3\), that gain of \(\varepsilon/3\) more than covers the monopolist’s sacrifice in revenue for those who take up the offer.

This example illustrates the basic intuition. The firm has a way of rewarding customers very cheaply. By lowering the price of A slightly below the monopoly level, it can make customers better off at almost no cost. Customers who take this deal are required to give something back to the firm in the form of paying an artificially high price for the

\[\text{D}(\alpha, p) \times \varepsilon - (1/3) \times \delta \geq 0.\]

---

\[^{6}\text{Generally, the consumer will prefer } (m-\varepsilon, c+\delta) \text{ to } (m, c) \text{ if and only if } \varepsilon^2 + D(\alpha, p) \times \varepsilon \geq (1/3) \times \delta. \text{ For small } \varepsilon, \text{ this reduces to } D(\alpha, p) \times \varepsilon \geq (1/3) \times \delta.\]
competitive good. Provided the premium in the B good isn’t too high, enough of the customers will take up the deal, and the gain in B will be additional profits to the seller. In the proof, we show that if \( \delta \) is a small enough multiple of \( \epsilon \), then (almost) everyone will accept the offer and profits will be increasing in \( \epsilon \).

Note that, in our example, if everyone were forced to accept the \((m-\epsilon, c+\epsilon)\) deal, then the monopolist would gain by an amount \( \epsilon/3 \). Adverse selection cuts into but does not eliminate the advantage of leveraging into the competitive market. As the consumption of good B is fixed, the deal is equivalent to the offer of a discount on A in return for the payment of a lump-sum tax. More generally, the existence of a lump-sum tax would not eliminate the advantage of increasing the price of B for reasons similar to those in Mathewson and Winter (1977).

III. Why Doesn’t the Chicago School Argument Work Here?

Given the general advantage of leveraging monopoly, the more surprising result is why the argument doesn’t extend to the case considered by the Chicago School.

In the Chicago model, consumers buy either 0 or 1 units of each of goods A and B. Thus, lowering the price of A below \( m \) won’t have the advantage of encouraging existing customers to buy more A. But it will still expand demand by bringing some additional customers on board. The extra revenue from those customers will just offset (to the first order) the loss in revenue from the discount on A.

Thus, it might seem that the intuition from the example would still carry through. The problem is that if the proportions of A and B are fixed (say at one to one), then an increase in the price of B is really just like an increase in the price of A. If we move from \((m, c)\) to \((m-\epsilon, c+\epsilon)\), then the marginal consumer for A has no reason to buy discounted A, as he has to give all the discount back via the overpriced B.
If demand doesn’t expand, then the monopolist doesn’t get a reward for lowering price. There is no cost, either, as the price of B has risen just enough to offset the fall in the price of A. The result is a wash.

Another way of putting it is the following: consider a customer who is already buying A and B, and, just by chance, in the proportion of one to one. In response to a change in prices from \((m, c)\) to \((m-\epsilon, c+\epsilon)\), the customer can be no worse off. If he continues to buy the same amounts, the cost will be the same and, thus, his previous consumption level is an option. But the customer can shift his consumption around in response to these price changes. The expected response will be to increase the purchase of A and reduce the purchase of B. Cutting back the purchase of B is of no (first-order) loss to the firm, as the profits on B are marginal. Increasing the purchases of A leads to a first-order gain for the monopolist, and that is the source of the leverage gain.

When customers are buying only one or none of A and B, then any potential gain has to come from marginal customers making a big shift in their consumption, rather than from existing customers making small shifts. But the marginal customer is indifferent before and is indifferent after and, thus, isn’t led to make a purchase.

What happened to Judge Bork’s “one monopoly profit” story? If it is possible to leverage monopoly from A to B, then how is the monopolist making more than its monopoly profit in A alone?\(^7\)

The answer is that the monopoly in A is inefficient. If it is possible to reduce the inefficiency (which is the result of lowering the monopoly price below \(m\)), then the total pie is bigger. If the monopolist can find a way to get some of that extra surplus (say by

---

\(^7\) Kaplow (1995) takes a different tack in his critique of the one monopoly profit argument. He notes: “Consider the case of a terrorist with one stick of dynamite. The fixed sum thesis posits that since the power is fixed—that is, the terrorist has one and only one dynamite stick—we should be indifferent to where the dynamite is placed.” Our example provides an additional reason for concern as to how the power is exercised. By applying the dynamite across two markets, the monopolist can do a better job exercising its power.
raising the price in good B), it can make more than the regular monopoly profits. Thus, while there may be only one monopoly profit, the size of that profit may vary.

When customers buy one or nothing, the only way to reduce the inefficiency is to get more customers to buy A. But to get them into the market, the total price of A and B has to be lower, and there’s the rub—there’s no way for the monopolist to get any of the extra surplus it has created.

One can debate whether the single unit purchase or the variable quantity purchase is the more broadly applicable case. In the case of a software product such as Word, it is fair to say that the end consumer is interested in one or none and that there is no natural variable quantity. But in the case of bundling long distance and wireless calls, the relative quantities of the two products are almost continuously variable and, thus, leveraging monopoly does make sense. Similarly, when Staples decides how much Scotch tape it will buy along with generic transparent tape, the two quantities are not fixed. Even in the case of SmithKline versus Lilly, there is some substitution between the different cephalosporins, and so again leveraging monopoly is possible.

IV. GENERAL PROOF OF LEVERAGING MONOPOLY RESULT

Consumer of type $\alpha$ has demand for products A and B, $D_a(\alpha, p_a, p_b, p_b^*)$ and $D_b(\alpha, p_a, p_b, p_b^*)$, where $p_b$ is the monopolist’s price of good B and $p_b^*$ are the other prices for good B in the market. Note that $D_b(\alpha, p_a, p_b, p_b^*)$ is demand for the monopolist’s version of good B. Define the cumulative density for type $\alpha$ by $F(\alpha)$ and let the density function be denoted by $f(\alpha)$.

By assumption, B is supplied into a competitive market, so that the rivals’ prices for B all equal marginal cost: $p_b^* = c$. Under the assumption that the B goods are perfect substitutes, then, absent any bundling contracts, $D_b(\alpha, p_a, p_b, c) = 0$ for $p_b > c$. Thus, the monopolist earns zero profits on B in any case and we can set $p_b = c$ without loss of
generality. For ease of notation, we write $D_a(\alpha, p_a, c)$ as the demand for good A when all firms charge c for good B.

The monopolist takes the price of good B as fixed at $c$ and chooses $p_a$ to maximize its profit. For simplicity, we assume that good A is produced at zero cost. The monopolist’s profits are:

$$\Pi(p_a,c) = \int_\alpha [p_a D_a(\alpha, p_a, c)] f(\alpha) d\alpha. \quad (5)$$

We denote the profit-maximizing price for good A by $m$.

The main result is that unforced bundling can lead to higher profits than $\Pi(m,c)$. This means that, even with $(m, c)$ available to consumers, the monopolist can find another offer $(m', c')$ that leads to higher profits. In establishing this result, the relationship between $D_a$ and $D_b$ can be quite general. The goods can be substitutes or complements, or they can be independent.

The result relies only on the assumption that the monopolist would like to charge more than c for the competitive good. It is possible that goods A and B have sufficient complementarity that lowering the price of B would increase demand of good A by more than enough to compensate for the subsidy of good B. If that were true, then there would be nothing preventing the monopolist in A from carrying this out. Even without a bundling strategy, it could offer good B at a discount. Our assumption is that selling B below cost is not desirable. Specifically, if the monopolist in A also had pricing power in good B, it would choose to raise the price of B above $c$.

Define the profits of a monopolist over goods A and B by $\Pi^*(p_a, p_b)$. Here, $D_b(\alpha, p_a, p_b)$ is the demand by type $\alpha$ for good B, assuming that he is paying the monopolist’s price of $p_b$.

$$\Pi^*(p_a, p_b) = \int_\alpha [p_a D_a(\alpha, p_a, p_b) + (p_b - c)(D_b(\alpha, p_a, p_b))] f(\alpha) d\alpha. \quad (6)$$
**Assumption 1:** At \((p_a, p_b) = (m, c)\), \(d\Pi^* (p_a, p_b) / dp_b > 0\).

**Theorem 1:** Under Assumption 1, there always exists an \(\varepsilon\) and \(\lambda\) such that \((m-\varepsilon, c+\lambda\varepsilon)\) leads to higher profits than \((m, c)\) when the firm offers consumers the choice of buying A and B on an à la carte basis at prices of \((m, c)\) or buying *all* of their B from the monopolist under prices of \((m-\varepsilon, c+\lambda\varepsilon)\).

**Proof of Theorem 1:** In appendix

The general intuition for this result is the envelope theorem. Starting at the optimal price \(m\), when the firm cuts the price of A, it loses money from the discount but gets back an equivalent amount from increased demand. The result is a wash. The customer, however, has a first-order welfare gain. Thus, the customer is willing to give back some (or even all) of that gain via paying a higher price for the B good. This giveback leads to higher profits for the monopolist.

The only problem with this intuition is its implicit assumption that all customers accept the bundle deal. If the giveback is small enough \((\lambda\) is near zero\), then nearly all of the customers will accept the bundled offer and, thus, the argument goes through.

In general, customers would like to reward the firm for lowering the price of A. But they don’t have a way to do this just on good A. The bundle with good B provides this opportunity.

From a consumer welfare perspective, one can still be worried about the consequences of bundling. While Theorem 1 shows the existence of a bundled offer that results in a Pareto improvement, the firm may not choose to let \((m, c)\) remain as an outstanding offer in the market. Or, anticipating that it will also be offering a bundled contract, the firm might not have chosen \((m, c)\) to begin with. Thus, it is less clear that the presence of bundling will
always results in a welfare gain to consumers, even when they have the choice of whether or not to take the bundled contract.

Another cause for concern is the long-run effect. Essentially all consumers are attracted to this bundle and, thus, all the other producers of good B will be foreclosed from the market. If reentry is difficult, then the monopolist may not be constrained by the existence of a competitive market in B in the future.

Given the generality of this result, we should go back and consider where the argument breaks down under the Chicago School framework. Assume that all customers consume A and B in a constant ratio of one to r. Then, customers care only about the bundle price of \( p_a + r p_b \). At \( (p_a, p_b) = (m, c) \),

\[
\frac{d\Pi^*(p_a, p_b)}{dp_b} = r \frac{d\Pi^*(p_a, p_b)}{dp_a} = 0,
\]

violating Assumption 1. Even if the firm could eliminate competitors and thereby raise the price of \( p_b \) above \( c \), profits would not go up. The increase in the price of B leads to a reduction in the demand for A that just offsets the increased revenue on B.\(^8\) Thus, Assumption 1 is not innocuous in that it rules out Leontief preferences.

Theorem 1 is based on an argument that considers a bundled deal that (nearly) everyone will accept. The proof considers a bundled offering with \( \lambda \) near zero. We also can show that bundling is attractive by looking at a case in which \( \lambda \) is sufficiently large that a vanishingly small set of customers accepts the bundle offer. The fact that unforced bundling is attractive both for small and large \( \lambda \) suggests that the gain from bundling will be large.

The intuition for the result with a large \( \lambda \) is different. Recall that customers can choose between \( (m, c) \) and \( (m-\varepsilon, c+\lambda \varepsilon) \). As \( \lambda \) increases, fewer customers will find the bundle

\(^8\) Note also that the monopolist would not like to lower the price of \( c \), either. The fixed ratio consumption is the knife-edge case in which \( (m, c) \) is an optimal price pair for a firm with a monopoly in both A and B.
offering attractive. Imagine that \( \lambda \) is so large that only one customer type is still willing to take the bundle. For that customer, the change in prices is a wash at his original consumption. The decrease in the price of good A is exactly offset by the increase in the price of good B. Thus, if the consumer maintains his original consumption, the firm’s profits would also be unchanged. But note that good A has become cheaper and good B more expensive, and if that change in relative prices leads the consumer to buy more good A, then profits will rise. With those two offsetting price changes, the firm has no direct revenue loss, but still gains from the increased consumption of A. The offsetting reduction in the consumption of B is irrelevant, as the firm earns zero profits on B (prior to the price change).

The second result on the desirability of unforced bundling relies on Assumptions A2—A5. I assume that the distribution of types has strictly positive density over its support. This assumption will be satisfied by any of the common demand assumptions, such as uniform or truncated normal.

**Assumption 2:** For \( \alpha \) such that \( 0 < F(\alpha) < 1 \), \( f(\alpha) \geq f^* > 0 \).

My next assumption is that preferences are smooth. Note that this rules out the case in which preferences are Leontief, and so consumption is always in a constant ratio.

**Assumption 3:** For all \( \alpha \), preferences for goods A and B are smooth, quasi-concave, and satisfy local non-satiation.

I assume that consumers for the monopoly product also be interested in good B when it is supplied at the competitive price.

**Assumption 4:** For all \( \alpha \), there exists a \( \lambda^* \) such that \( D_a(\alpha, m, c)/D_b(\alpha, m, c) \leq \lambda^* \).

The last assumption requires that A and B are the only two goods in the market. This is stronger than required and is used to abstract away issues of how the consumers
substitute between A and B and all other goods. Combined with Assumption 3, the implication of this assumption is that consumers spend all of their income on goods A and B. For simplicity, I call this Assumption 5.

\textit{Assumption 5:} \( p_a D_a(\alpha, m, c) + p_b D_b(\alpha, m, c) = Y \).

Under Assumptions 2-5, a firm with a monopoly in product A and facing competition in good B can earn higher profits by leveraging its monopoly in A. Consumers are offered a discount on A if they buy all of their B demand from the firm, where B is sold at above the competitive price. The monopolist can earn higher profits through this offer even when consumers retain the option of buying A and B at the pre-bundling prices of \((m, c)\).

\textbf{Theorem 2:} Under Assumptions 2-5, there always exists an \( \varepsilon \) and \( \lambda \) such that \((m-\varepsilon, c+\lambda \varepsilon)\) leads to higher profits than \((m, c)\) when the firm offers consumers the choice of buying A and B on an \( \text{à la carte} \) basis at prices of \((m, c)\) or buying \textit{all} of their B from the monopolist under prices of \((m-\varepsilon, c+\lambda \varepsilon)\).

\textit{Proof of Theorem 2:} In Appendix.

Theorem 2 is suggestive of how bundling continues to be effective even as \( \lambda \) increases. In the proof, the monopolist chooses a large enough \( \lambda \) so that only the consumers with the largest ratio of \( D_a \) to \( D_b \) will accept the offer. The monopolist takes back (almost) all of its discount in A through a price increase in B. The net savings to customers will be close to a wash, as will be the cost of the discount. So long as the price tilt in favor of good A results in increased A consumption, this will lead to higher profits.

As the prior consumption is still affordable, and good A is cheaper while B is more expensive, we expect consumption of A to rise and B to fall. With only two goods in the market, this won’t happen if (contrary to Assumption 3) customers have Leontief utility. In that case, there will be a kink at the tangency to the original budget line, and the optimal consumption will remain unchanged at the new prices. Profits also would be
unchanged if the potential increase in demand comes entirely from new consumers who buy A and B only in discrete fixed proportions. For these types of consumers, there is no ability to adjust A on the margin. Given the fixed proportions, the offsetting price changes exactly cancel out; thus, they see no price fall and, hence, have no incentive to increase demand. Again, this violates Assumption 3.

It is possible that with more than two goods, the increase in the price of B results in a decreased level of consumption of both A and B, which more than offsets the increase in consumption of A from its reduced price. While this seems to be the less likely scenario, it is not impossible, and, hence, Theorem 2 is less general than Theorem 1.

V. A BROADER VIEW OF BUNDLING INTO COMPETITIVE MARKETS

Having established the advantage of leveraging monopoly, I want to return to the Chicago School argument and reevaluate it on its own terms. It is true that there is no gain from leveraging monopoly when consumption of A and B are either 0 or 1 and B is supplied by a competitive market with constant marginal cost and zero fixed costs.

But the definition of a competitive market extends beyond production at constant marginal costs and no fixed costs. Indeed, it is hard to think of any industry with constant returns to scale and zero fixed costs.

A broader definition of competitive markets is that there is free entry and that the marginal firm makes zero profits. In that setting, there can be an advantage to leveraging monopoly.

The layperson intuition is again useful. The typical firm seeks additional customers at the market price. Indeed, as Hal Varian has pointed out, firms spend a good deal of money on advertising and marketing to attract additional customers. This suggests that price is
typically above marginal cost. The profits cover the fixed costs of doing business, so that in the end, firms do not earn above competitive rates of return.

Consider what would happen if a monopolist charging \((m, c)\) tells its customers that they have to buy all of their B demand from it. As the competing B goods are perfect substitutes, consumers are no worse off.

This is the flip side of the Chicago argument. While the firm doesn’t gain anything from foreclosure, it doesn’t cost anything, either. If most consumers of good B also buy A, then at no cost, the monopolist can exclude all or almost all of its rivals in the B market.\(^9\)

It hasn’t made any more money in a world where price equals marginal cost. But our layperson would argue that a firm that could costlessly increase its market share, from \(1/n\) to 1, would stand to earn money in the B market.\(^{10}\)

The A monopolist might make money in the future by achieving a second monopoly in B. It might make money by preventing others from using the B market as a launching pad into A and thereby preserve its A monopoly. (Denying rivals scale in a complementary B market is the point emphasized by Carlton and Waldman (2002).) But our layperson would suspect that the firm would make money even in the present through its greatly increased B sales and market share. Of course, our layperson isn’t always right, and we need to make this argument more formal.

Since price is above cost, increased B sales are profitable. That leads to the question of why the existing B firms don’t cut price to capture the entire market and thereby increase their profits. Our answer is that demand is not perfectly elastic. Even if the goods are perfect substitutes, customers may not be aware of this.

---

\(^9\) This point is emphasized by Kaplow (1985) and is the subject of a companion paper, Nalebuff (2004b).

\(^{10}\) The monopolist can go one step further and even offer an inducement to buy its B good. It can offer to lower the cost of \(m\) to \(m - \epsilon\) if the customer buys all of its good B demand from the monopolist, and that would give consumers a strict incentive to buy its B good. This ability to exclude even more efficient rivals would come at a cost, but only a second-order cost, to the monopolist.
Firms can increase their market share in several ways: increasing product quality, spending more money on marketing, and cutting price. Marketing dollars are more like a fixed cost, while increased quality may have both fixed and variable cost components. The monopoly in A can give the firm an advantage in each of these spheres. There may be complementarity in production between A and B, leading to lower costs or higher quality. The good A monopoly also can make marketing or price cuts more cost-effective.

Even if firms have the same variable cost production technology for B, the existence of market power may influence the market shares. Requiring customers to buy the firm’s B along with its A is an inexpensive form of advertising. The customer might have been indifferent about which B to buy. The firm uses the fact that it has the customer’s attention when purchasing A to capture that customer’s demand for B, as well. In this case, the A monopolist is more efficient than its rivals in marketing. Its good A monopoly gives it an advantage in getting access to the customer. Microsoft, for example, might include its Windows Media Player along with the operating system. While customers might otherwise have split 50-50 between Real Media player and Windows Media player, Microsoft’s monopoly in A gives it preferential access to 100% of the potential B customers.

The result of Theorem 1 suggests another way in which the monopolist can leverage its power. The monopolist uses the discount on A as a way to almost costlessly increase its share of good B. The customer who is buying both goods doesn’t care if A or B is discounted. But discounting A is cheaper for the monopolist as it expands demand. Rival firms who sell only good B don’t have the same option.

This effect can be illustrated in a simple model. Consider a market where demand for good A is independent of consumer type. For all consumers, demand for good A is \( D_a(p) = 1 - p_a \). Assuming zero cost, the monopoly price of A is \( \frac{1}{2} \).
Each consumer’s demand for Good B is inelastic at 1. However, due to imperfect search or idiosyncratic valuation for the product of different B firms, customers will not necessarily purchase B from the lowest-priced firm. For this example, we employ the random utility model where the error term has a multivariate (Type I) extreme value distribution. Given a price vector \( \mathbf{p} \), and \( n \) competing firms, each customer’s chance of buying from firm \( i \) is

\[
D_b(p) = \frac{e^{-p_i}}{\sum_{j=1}^{n} e^{-p_j}}. \tag{7}
\]

Assume that B is produced with a zero marginal cost and a fixed cost of 0.1.\(^{11}\) With these parameters, the profit-maximizing price satisfies the following first-order condition:

\[
D_b^* (1 - p_i + p_i D_b) = 0. \tag{8}
\]

In a symmetric equilibrium, \( D_b = 1/n \), so that \( p_i = n/(n-1) \) and \( \pi_i = 1/(n-1) - F \). Thus, with \( F=0.1 \), the equilibrium number of firms will be 11: \( p_i = 1.1 \), and \( \pi_i = 0.1 - 0.1 = 0 \), so each B firm will earn zero profits. We assume that the firm with the A monopoly is one of these 11 participants in the B market.

Now that we have the baseline competitive market, we are ready to show how bundling leads to increased profits. Consider the result if the monopolist in A offers good A at a price of 0.4 if the customer also buys its version of good B.

To measure the impact of this offer, we can divide the customers into two groups, the

\( 1/11^{th} \) of the market already buying B from the monopolist and the \( 10/11^{th} \) buying good B from other firms.

From the customers in the first group, the monopolist sees demand of good A expand to 0.6, and it earns 0.24 rather than 0.25 (a second-order loss in profits). The aggregate cost

\[^{11}\) Here, the fixed cost is related to production, not marketing, and thus does not influence demand.
is .01/11 approx 0.001. There is a second-order loss from the fact that 0.4 is slightly below the optimal price of 0.5.

For the other 10/11ths of the customers, the discount on A will be available only if they switch over their B purchase to the monopolist. On their existing purchase of 0.5 units of A, the monopolist’s offer will save them 0.05.\textsuperscript{12} From the customer’s perspective, this is just like a price reduction in good B of 0.05.

As a result, the A monopolist’s share of the B market will increase from $1/11$ by approximately

$$D_b(p)\Delta p = (-D_b(p) + D_b(p)^2) \times (-0.05) = \frac{10}{121} \times \frac{0.05}{121} = \frac{0.5}{121}$$

This gain in market share is almost pure incremental profits. The reason is that the customers who accept the offer also increase their demand for A. Thus, the cost of the discount is only the difference between 0.5*0.5 and 0.4*0.6, or 0.01. Each new B customer brings in incremental revenue of 1 (from B) offset only by a profit loss of 0.01 in A. The net gain to the monopolist from the switchers is

$$0.5/121 \times (1 - 0.01) > 0.001.$$  

The intuition for this result is related to the general argument for bundling provided by McAfee, McMillan, and Whinston (1989). The primary difference is that we do not assume fixed demand, and so bundling is even more attractive.\textsuperscript{13} The gain comes from

\begin{itemize}
  \item \textsuperscript{12} In fact, they will gain slightly more as a result of their increased consumption of A. But, to be conservative, we assume that the consumer considers the price savings on his existing demand only when considering switching B providers.
  \item \textsuperscript{13} If demand for A were 0 or 1, the result would also follow. Let the value of good A to type $\alpha$ be $\alpha$. With a uniform density over [0,1], the demand for A will be 1-$p$ and the optimal price will be $\frac{1}{2}$. If the monopolist offers A at 0.4 to customers who buy B from it as well (or an A-B bundle at a price of 1.4), then this will cost the firm 0.01 on the 1/11\textsuperscript{th} of the customers who are already buying B from the firm. As before, the firm will gain an additional 10/121*(1)*(1/2) in market share—here, the 0.1 represents the cost savings and the $\frac{1}{2}$ represents the fraction of the market buying A. The increased profit from the incremental share is 0.9, as the firm is effectively selling B at a price of 0.9. The point is that the firm gets the increase in B market share at almost no cost, as the discount given to its existing B customers is almost entirely offset by
\end{itemize}
the fact that the discount for good A does double duty. It expands the demand for good A among existing B customers, and since A was optimally priced, this is close to a wash. The discount also brings in new B customers. But, as the cost of the discount is almost nothing (as these customers buy more A), the incremental B customers are essentially pure profits.

To say this differently, the monopolist in A has an advantage in going after B customers. Its rivals in the B market have to cut the price of B to gain market share. The A monopoly can offer B customers a reduced price in A if they switch over their B demand. While the customer considers this almost the same as reducing the price of A, the result is much better for the monopolist. It gets incremental B customers essentially for free, as the price reduction is almost entirely offset by their increased consumption of A.

The firm’s monopoly in A allows it to earn a profit in what would otherwise be a zero-profit competitive B market.

VI. CONCLUSION

The results in this paper show how a monopolist can use its market power in good A to achieve an advantage in a competitive B market and increased profits. It is, perhaps, surprising that these gains can arise even without coercion in that the original unbundled prices all remain in the market.¹⁴

The source of the gain is inefficiency of monopoly pricing and the envelope theorem. The monopolist is almost indifferent about charging a slightly lower price, while consumers are strictly better off. The monopolist can recapture some of that gain either through a

---

¹⁴ The leverage could work via a threat instead of a promise. Instead of offering a discount if good B is purchased, the monopolist could threaten a price increase if B is not purchased. This would be similarly effective, although more coercive.
premium price in the B market or via increased share (in the case where price exceeds marginal cost).

The result of the bundled offer is a Pareto improvement, so there is clearly no immediate harm to consumers. However, there is clear harm to competitors, many or even all of whom may be foreclosed from the B market. Whether foreclosure will allow the A monopolist to later dominate the B market and what the impact that might have on consumers requires a dynamic model of competition, something beyond the scope of this paper. Even in the short run, the fact that a Pareto improvement is possible doesn’t mean that it will be offered. The monopolist can make even more profits if coercion or a forced bundle is allowed.

What we have shown is the why and how of bundling or tied sales. Except in the special case where goods A and B are consumed in a fixed ratio, we can expect firms with market power to extend their influence beyond their original monopoly, even without coercion. While it is true that this exercise of the one monopoly power across multiple markets leads to less distortion, it is also the case that this leads to an expansion of the monopolist’s sphere of influence. Are two half wrongs better than one? The ability of a monopolist to extend its influence to adjacent markets is a challenge both to the competitors in those markets and to economists looking to understand the antitrust implications of bundling.
REFERENCES


Theorem 1: Under Assumption 1, there always exists an \( \varepsilon \) and \( \lambda \) such that \( (m-\varepsilon, c+\lambda \varepsilon) \) leads to higher profits than \( (m, c) \) when the firm offers consumers the choice of buying A and B on an à la carte basis at prices of \( (m, c) \) or buying all of their B from the monopolist under prices of \( (m-\varepsilon, c+\lambda \varepsilon) \).

Proof of Theorem 1: Let the monopolist’s profits from the combined offers \( (m, c) \) and \( (m-\varepsilon, c+\lambda \varepsilon) \) be denoted by \( \Pi(m, c, \varepsilon, \lambda) \). For small \( \varepsilon \), the customers who prefer \( (m-\varepsilon, c+\lambda \varepsilon) \) to \( (m, c) \) are those for whom their demand at \( (m, c) \) is still affordable at \( (m-\varepsilon, c+\lambda \varepsilon) \).

\[
\Pi(m,c,\varepsilon,\lambda) = \int_{\alpha D_a(\alpha,m,c) \geq \lambda D_b(\alpha,m,c)} [(m-\varepsilon)D_u(\alpha, m-\varepsilon, c + \lambda \varepsilon) + \lambda \varepsilon D_b(\alpha, m-\varepsilon, c + \lambda \varepsilon)] f(\alpha) d\alpha 
+ \int_{\alpha D_a(\alpha,m,c) < \lambda D_b(\alpha,m,c)} [(m)D_u(\alpha, m, c)] f(\alpha) d\alpha
\]

The argument relies on the envelope theorem. Define \( V(m, c, \lambda) = \frac{\partial \Pi(m, c, \varepsilon, \lambda)}{\partial \varepsilon} \bigg|_{\varepsilon=0} \).

\[
V(m,c,\lambda) = \int_{\alpha D_a(\alpha,m,c) \geq \lambda D_b(\alpha,m,c)} [m \frac{dD_u(\alpha, m-\varepsilon, c + \lambda \varepsilon)}{d\varepsilon} - D_u(\alpha, m, c) + \lambda D_b(\alpha, m, c)] f(\alpha) d\alpha 
= \int_{\alpha D_a(\alpha,m,c) \geq \lambda D_b(\alpha,m,c)} [m (-\frac{\partial D_u(\alpha, m, c)}{\partial p_a} + \lambda \frac{\partial D_u(\alpha, m, c)}{\partial p_b}) - D_u(\alpha, m, c) + \lambda D_b(\alpha, m, c)] f(\alpha) d\alpha.
\]

At \( \lambda=0 \), all customers accept the bundled offer since it offers them a discount on A without having to pay more for B. Further note that with \( \lambda=0 \), a change in \( \varepsilon \) is just like a change in \( p_a \), as none of the discount is taken back. Thus, \( V(m, c, 0) = 0 \), by the optimal choice of \( p_a \). (This reflects the fact that, as all consumers accept the deal, the price reduction in A is exactly offset by the increase demand.)

The result will follow if we can demonstrate that \( \frac{\partial V(m, c, \lambda)}{\partial \lambda} \bigg|_{\lambda=0} > 0 \). That would imply \( \frac{\partial \Pi(m, c, \varepsilon, \lambda)}{\partial \varepsilon} \bigg|_{\varepsilon=0} > 0 \) for small \( \lambda \). If the monopolist offers a discount on good A with the condition that customers buying good B give back a small fraction of the
saving, almost all consumers will accept this offer, and the net effect on profits will be positive.

\[ \frac{dV(m,c,\lambda)}{d\lambda} \bigg|_{\epsilon=0} = \int [m \frac{\partial D_a(\alpha,m,c)}{\partial p_b} + D_b(\alpha,m,c)] f(\alpha) d\alpha \]

This last condition is positive by Assumption 1,

\[ \frac{d\Pi^*(m,c)}{dp_b} = \int [m \frac{\partial D_a(\alpha,m,c)}{\partial p_b} + D_b(\alpha,m,c)] f(\alpha) d\alpha > 0. \]

Q.E.D.

**Theorem 2:** Under Assumptions 2-5, there always exists an \( \epsilon \) and \( \lambda \) such that \((m-\epsilon, c+\lambda\epsilon)\) leads to higher profits than \((m, c)\) when the firm offers consumers the choice of buying A and B on an à la carte basis at prices of \((m, c)\) or buying *all* of their B from the monopolist under prices of \((m-\epsilon, c+\lambda\epsilon)\).

*Proof of Theorem 2:* Let the monopolist’s profits from the combined offers \((m, c)\) and \((m-\epsilon, c+\lambda\epsilon)\) be denoted by \(\Pi(m, c, \epsilon, \lambda)\). Recall from the proof of Theorem 1 that \(V(\lambda) = \frac{\partial \Pi(m, c, \epsilon, \lambda)}{\partial \epsilon} \bigg|_{\epsilon=0}\) where

\[ V(m,c,\lambda) = \int [m \frac{\partial D_a(\alpha,m,c)}{\partial p_b} + D_b(\alpha,m,c)] f(\alpha) d\alpha \]

Profits change only to the extent that customers accept the \((m-\epsilon, c+\lambda\epsilon)\) offer. As \(\epsilon \rightarrow 0\), the set of customers who prefer the bundled offer are those who save money at their existing levels of demand: \(D_a(\alpha, m, c) - \lambda D_b(\alpha, m, c) \geq 0\).

As \(\lambda\) increases, it follows that the range of the integral shrinks. The reason is that there are fewer consumers for whom saving \(\epsilon\) on \(D_a\) more than offsets paying an extra \(\lambda\epsilon\) on \(D_b\).
Define $\lambda^*$ as the minimum value of $\lambda$ that satisfies Assumption 4. Thus $V(\lambda) = 0$ for all $\lambda \geq \lambda^*$. At $\lambda^*$, the measure of consumers for whom $D_a(\alpha, m, c) - \lambda D_b(\alpha, m, c) \geq 0$ is zero (although there could be several different $\alpha$ for which this equality holds).

$$V(m,c,\lambda^*) = \int_{\alpha:D_a(\alpha,m,c)\geq \lambda^*D_b(\alpha,m,c)} [mdD_a(\alpha,m-\varepsilon,c+\lambda^*\varepsilon)/d\varepsilon + 0] f(\alpha) d\alpha.$$  

Consider the (downward derivative) value of $V'(\lambda)$ evaluated at $\lambda = \lambda^*$. The second and third terms of the integral $V(m, c, \lambda)$ cancel out as $-D_a(\alpha, m, c) + \lambda^*D_b(\alpha, m, c) = 0$ for the consumers just on the margin. Thus,

$$dV(m,c,\lambda^*)/d\lambda |_{\lambda^*} = -D_b(\alpha,m,c) \sum_{\alpha:D_a(\alpha,m,c)\geq \lambda^*D_b(\alpha,m,c)} m[dD_a(\alpha,m-\varepsilon,c+\lambda^*\varepsilon)/d\varepsilon] f(\alpha).$$  

Given Assumption 2 that $f(\alpha) > 0$, it follows that $V'(\lambda^*) < 0$ provided that $dD_a/d\varepsilon$ is positive. By the definition of $\lambda^*$, the original consumption vector, $(D_a(\alpha, m, c), D_b(\alpha, m, c))$ is still just affordable at the new prices $(m-\varepsilon, c+\lambda^*\varepsilon)$ for the $\alpha$ type(s) who choose the new prices. Thus, the new budget line runs through the original consumption point. But the new budget line has a steeper slope. Given smooth preferences (Assumption 3) and non-satiation, this will lead the consumer to buy more of good A and less of good B.

This conclusion follows from WARP and Assumption 5. Given that the consumer will be on a budget line through $(D_a(\alpha, m, c), D_b(\alpha, m, c))$, he can’t consume either more of both goods or less of both goods. From WARP it then follows that consumption of A rises and B falls.

This last step can be seen graphically by looking at the preferences for any of the customers for whom $D_a(\alpha, m, c) = \lambda^*D_b(\alpha, m, c)$. Note that at prices $(m, c)$, the firm doesn’t make any profits on B, and so its isoprofit lines are horizontal.
In response to the new prices, those who accept the bundle adjust their consumption in a manner that leads to more $D_a(\alpha, p_a, p_b)$ and less $D_b(\alpha, p_a, p_b)$. The increase in $D_a(\alpha, p_a, p_b)$ raises profits for the monopolist.

As $V(m, c, \lambda^*) = 0$ and $dV(m, c, \lambda)/d\lambda < 0$, for $\lambda$ slightly lower than $\lambda^*$ it follows that a small increase in $\varepsilon$ will lead to higher profits. The change in expenditure at existing consumption is near zero, while the gain from increased consumption of $D_a(\alpha, m, c)$ is strictly positive.

Q.E.D.