

Chapter 40

QUANTITY FORCING AND EXCLUSION: BUNDLED DISCOUNTS AND NONLINEAR PRICING

Marius Schwartz and Daniel Vincent*

Quantity “forcing” refers to pricing schemes that reward a buyer for purchasing some threshold quantity from a firm. When there are significant scale economies and buyers are unable to coordinate, economic theory shows that a firm can profitably use quantity forcing to exclude rivals, reducing overall welfare and harming some buyers. Inducements to reach the quantity threshold may be provided through nonlinear pricing of the target product alone or through bundled discounts on that firm’s other “monopoly” product(s). Open questions remain about whether bundled discounts are the most effective way to achieve exclusion. Alternatively, bundled discounts can be used to extract rent from a monopoly market, but again, single-good nonlinear pricing schemes seem superior. Cost-based rules for detecting predation are problematic when applied to bundled discounts or to single good nonlinear pricing. A workable policy rule that also recognizes the efficiency potential of such pricing practices should combine structural screens with a more detailed conduct inquiry.

1. Introduction

Allegations of anticompetitive exclusion through quantity “forcing” have featured prominently in a line of recent cases, including *Ortho*, *Virgin-British Airways*, *Concord Boat*, and *LePage’s*.¹ These cases share some important features:

- Unlike explicit exclusive dealing, the alleged exclusion sometimes occurs not by barring purchases from rivals but by inducing (or “forcing”) customers to increase their purchases from the defendant—seemingly a procompetitive action.
- Unlike ordinary predatory pricing, the inducement is not via a simple cut in the per unit price of that product but essentially through a payment that is conditional on reaching a certain quantity threshold. With “market share

* Department of Economics, Georgetown University; Department of Economics, University of Maryland, College Park. We thank seminar participants at the U.S. Department of Justice and the Bates White 2006 Antitrust Conference, Roy Englert, and Don Russell for helpful discussions, and Ralph Winter for thorough comments.

1. *Ortho Diagnostic Sys. v. Abbott Labs.*, 920 F. Supp. 455 (S.D.N.Y. 1996); *Virgin Atl. Airways v. British Airways*, 69 F. Supp. 2d 571 (S.D.N.Y. 1999), *aff’d*, 257 F.3d 256 (2d Cir. 2001); *Concord Boat Corp. v. Brunswick Corp.*, 207 F.3d 1039 (8th Cir. 2000); *LePage’s Inc. v. 3M*, 324 F.3d 141 (3d Cir. 2003) (en banc). An earlier case raising similar issues is *SmithKline Corp. v. Eli Lilly & Co.*, 575 F.2d 1056 (3d Cir. 1978).

discounts,” the threshold is expressed as a percentage of the buyer’s total purchases of that product rather than an absolute quantity.

- At times, the inducement is offered in that market alone (e.g., *Concord Boat*), through various pricing schemes, such as all-units discounts.
- In other cases, the inducement is via discounts on different products sold by the alleged excluder over which it possesses significant market power (as in *SmithKline*, *Ortho*, *Virgin*, and *LePage’s*; the latter featured also outright payments by 3M to some stores for exclusivity). The “discounts” are off the prices of those other products if purchased alone. Because such multimarket discounts are conditional on reaching the target on the first product, they are sometimes called “bundled” or “tied” discounts.
- In *LePage’s*, the inducements were not offered marketwide but selectively to key accounts.

These practices raise important issues for economic analysis and antitrust policy:

- Can consumers be harmed by output-expanding initiatives that, unlike predatory pricing, need not be temporary?
- If so, how might one identify anticompetitive uses of these practices? For example, does the so-called *Ortho* test properly identify predatory pricing that may be implemented through bundled discounts?² Do standard price-cost tests for predatory pricing suffer from additional drawbacks when nonlinear pricing is used?
- Can tying be used, as some have claimed, to exclude competitors without incurring any profit sacrifice—i.e., “costless exclusion”? If so, the competitive threat from tying would be much greater than from other practices judged under the rule of reason. Moreover, a profit sacrifice test for exclusionary or predatory conduct³ would be vacuous in the case of exclusionary tying.

We find it important to distinguish between two potential roles of the practices in question: rent extraction and monopolization. The former refers to collecting more fully for any market power that the firm has legitimately acquired; the latter entails acquiring,

2. The test was proposed by the plaintiff’s economic expert, Janusz Ordover, but has come to be known more generally as the *Ortho* test. *Ortho*, 920 F. Supp. at 470. This test is used where goods are made available both as a bundle and on a stand-alone basis. It attributes as a cost of selling the tied product the sacrifice incurred by offering the tying product at a discount from its unbundled price. See *infra* Section 3.

3. The profit sacrifice test, or some variant, is advocated by ROBERT BORK, *THE ANTITRUST PARADOX: A POLICY AT WAR WITH ITSELF* (1978); Janusz Ordover & Robert Willig, *An Economic Definition of Predation: Pricing and Product Innovation*, 91 *YALE L.J.* 8 (1981); Douglas Melamed, *Exclusive Dealing Agreements and Other Exclusionary Conduct—Are There Unifying Principles?*, 73 *ANTITRUST L.J.* 375 (2006); and Gregory Werden, *Identifying Exclusionary Conduct Under Section 2: The “No Economic Sense” Test*, 73 *ANTITRUST L.J.* 413 (2006). Schwartz points out some limitations, while Salop rejects the approach, in favor of a consumer harm standard. See Marius Schwartz, *Investments in Oligopoly: Welfare Effects and Tests for Predation*, 41 *OXFORD ECON. PAPERS* 698 (1989); Steven Salop, *Exclusionary Conduct, Consumer Harm, and the Flawed Profit-Sacrifice Standard*, 73 *ANTITRUST L.J.* 311 (2006).

maintaining, or extending market power in ways that harm the firm's trading partners.⁴ At the normative level, one could take the position that only monopolization should be prohibited; if the market power was legitimately acquired, unfettered rent extraction should be allowed. Moreover, a rent extraction motive may be present also where structural conditions make monopolization implausible.

It will be useful at the outset to define some terms. Under standard "linear pricing," a buyer is allowed to buy *any* quantity the buyer desires at a constant, per unit price with no fixed fee.⁵ A "two-part tariff" simply adds a fixed fee to linear pricing. We use the term "quantity forcing" to denote the goal of inducing a downstream buyer to purchase (at least) a target quantity of product from an upstream supplier. This goal may be achieved through a variety of different "nonlinear contracts." *Within* the same market, quantity forcing may be achieved, for example, through (1) "all-or-nothing" contracts that allow the buyer only a single quantity choice, the target quantity, and either make other quantities unavailable or available at a prohibitively high price; or (2) "all units discounts," where if the quantity purchased exceeds a specified threshold, the buyer obtains a discount on all units. Alternatively, quantity-forcing contracts could involve products *across* markets. A "tie-out" (or "requirements tying") requires the buyer to commit to purchase all (or a very high share) of its requirements in the tied market from the supplier in return, typically, for a price break on a product in a separate, tying market, usually one in which the supplier has significant market power. A "tie-in" requires the purchase of only a modest amount of the tied good from that supplier.

We begin this chapter with tying. Section 2 presents a simple model where a monopolist over some product 1 also sells an unrelated product 2 and faces competition there. If the monopolist is only able to use linear pricing for each product, tying can permit fuller rent extraction than available when offering the monopoly good independently.⁶ Tying then serves as an imperfect substitute for richer pricing instruments such as two-part tariffs. We identify the effects of tying on consumers and overall welfare in two analytically useful benchmark cases: (1) when the monopolist can commit to offer product 1 only tied with 2, known as pure tying, or (2) can offer a tie but must also continue to offer product 1 unbundled at its simple monopoly price, a practice that Nalebuff calls mixed tying.⁷

Section 3 illustrates shortcomings of the *Ortho* test for predatory pricing when the defendant firm is offering tied discounts. If the tied products are used in fixed proportions, as implicitly assumed when comparing prices for unbundled and tied offers,

4. For a recent reminder of this distinction, see Ken Heyer, *A World of Uncertainty: Economics and the Globalization of Antitrust*, 72 ANTITRUST L.J. 375 (2005). We shall use the term "monopoly" not literally but as shorthand for substantial market power, and similarly for "monopolization."

5. Hence the term "linear," since the total amount paid is linear in the quantity purchased. As is well known, a monopolist restricted to linear pricing typically must cede some consumer surplus to its buyers.

6. See Burstein and Mathewson and Winter, who extend his analysis. Meyer Burstein, *The Economics of Tie-in Sales*, 42 REV. ECON. & STAT. 68 (1960); Frank Mathewson & Ralph Winter, *Tying as a Response to Demand Uncertainty*, 28 RAND J. ECON. 566 (1997).

7. See Barry Nalebuff, *Bundling as a Way to Leverage Monopoly* (Yale School of Management, Working Paper ES-36, 2004).

the firm can fail the test even when its conduct is motivated by price discrimination rather than predation and rivals remain in the tied market. When the products are used in variable proportions, comparing prices is not a meaningful indicator of potential sacrifice, and we present an alternative “revenue” version of the test. Using the analysis of Section 2, we show that the firm can fail the test even when tying is not predatory, does not harm consumers, and increases overall welfare. The test overstates the profit sacrifice, in large part because it ignores any output expansion for the tying product.

Section 4 questions the assumption maintained thus far that a monopolist can implement tying but is otherwise limited to charging per unit pricing. Suppose a seller can require a buyer to purchase a second good only from that seller (a tie-out). The same seller should also be able to collect a fixed fee, by requiring the buyer—as a condition for obtaining the monopoly good—to purchase *even a small quantity* of the tied product at a suitably inflated price. Moreover, with identical buyers and full information, charging a fixed fee through such a “minimal tie-in” is superior to a tie-out if the purpose is rent extraction. Mathewson and Winter show that a rent extraction role for tying can resurface if the monopolist lacks full information on consumers’ demands and if demands for the two products are affiliated (loosely, are positively correlated in all cases).⁸ We review their argument, note some limitations, and find significant open questions about the optimality of using tie-outs purely for rent extraction. Monopolization of the tied market, on the other hand, can offer a clear motive for inducing buyers to purchase a large share of their requirements from the would be excluder—quantity forcing—when the target product exhibits significant economies of scale. Quantity forcing then can deny scale economies to competitors and threaten their vigor or even viability. One way to implement quantity forcing in the target market is by offering a tied discount on a second product. However, providing the incentives through tied discounts is generally more costly than utilizing nonlinear pricing of the target product alone. Thus, the exclusionary motive for quantity forcing also fails to explain the use of tie-outs instead of reliance on single product nonlinear pricing.

Section 5 illustrates how exclusionary quantity forcing can be profitably implemented through nonlinear pricing of the target product alone. We review briefly the theory of “naked exclusion,” that is, why monopolization through exclusive dealing contracts can be profitable given scale economies and lack of buyer coordination.⁹ In most of the antitrust cases mentioned, the contested practices did not include explicit exclusive dealing.¹⁰ However, drawing on Bernheim and Whinston, we show that the basic divide-and-conquer logic can allow profitable exclusion (though less frequently) also through quantity-forcing contracts.¹¹ While reliance on quantity forcing makes the

8. Mathewson & Winter, *supra* note 6.

9. Eric Rasmusen, J. Mark Ramseyer & John Shepard Wiley, Jr., *Naked Exclusion*, 81 AM. ECON. REV. 1137 (1991); *see also* Ilya Segal & Michael Whinston, *Naked Exclusion: Comment*, 90 AM. ECON. REV. 296 (2000); Robert Innes & Richard Sexton, *Strategic Buyers and Exclusionary Contracts*, 84 AM. ECON. REV. 566 (1994).

10. An exception is *LePage’s*, where allegations included payments for exclusivity. *See* Brief for Respondents in Opposition, 3M v. LePage’s, 2003 WL 22428377, at *1 (2003) (No. 02-1865).

11. B. Douglas Bernheim & Michael Whinston, *Exclusive Dealing*, 106 J. POL. ECON. 64 (1998).

firm's output higher than under naked exclusion, consumer welfare and overall welfare can still be lower than in the nonexclusionary benchmark.

Section 6 reviews some recent nonexclusionary explanations for quantity-forcing contracts, including price discrimination when buyers have private information about their demands, and inducement of effort by dealers.

Section 7 presents concluding remarks. We see no clean test to distinguish benign from exclusionary motives for quantity-forcing practices. However, our analysis suggests that the profit sacrifice approach has some bite, at least conceptually—exclusion typically requires behavior different from that which maximizes rent extraction from preexisting market power. A workable policy approach is likely to combine this principle with structural screens, such as the degree of scale economies in the target market and whether competitors' viability is seriously at risk.

2. Tying for rent extraction under only linear pricing

For simplicity, and in keeping with the main articles discussed below, we focus on tied products that are unrelated in both cost and demand. With complementarities, it is well known that tying can arise for efficiency reasons, for price discrimination, or for exclusion. However, the argument that tying can induce costless exclusion has been made recently even for such unrelated products. In our view, this argument confounds exclusion with the use of tying for fuller rent extraction by a monopolist who can only employ linear pricing. The rent extraction argument, due to Burstein, and Mathewson and Winter, can be illustrated in the following simple environment.¹²

Suppose firm A is a secure monopolist over product 1. In product 2, it faces an actual or potential rival, firm B (or a competitive fringe), capable of offering an identical product. Thus, suppliers of product 2 offer perfect substitutes (homogeneous goods). Demand for each product depends only on its own price. Firm A has constant marginal and average costs for both products, (c^1, c^2) .¹³ For now, assume that firm B also has marginal cost c^2 and no fixed cost. These conditions imply that firm B is equally efficient as A , and they rule out tying to monopolize market 2 since firm B can operate efficiently at any small scale.¹⁴ We will later relax both conditions. A typical buyer's demands for products 1 and 2 are denoted $q^1 = D^1(p^1)$, $q^2 = D^2(p^2)$ and are shown in Figure 1 as linear purely for simplicity.

We note two alternative interpretations of these demand curves for welfare analysis. A buyer may be a final consumer as above. The height of a demand curve then represents the buyer's *marginal* willingness to pay (in dollars per unit) for that product; the area under the demand curve from 0 to any quantity q shows the buyer's *total* willingness to pay (in dollars) for that entire quantity q when the alternative is to get

12. Burstein, *supra* note 6; Mathewson & Winter, *supra* note 6.

13. We will generally use superscripts to denote the product or market.

14. Monopolization of market 2 through tying would be implausible also if there were modest economies of scale but a large share of all customers who desire product 2 were not interested in product 1; the stand-alone demand for product 2 is large enough to support numerous single product competitors.

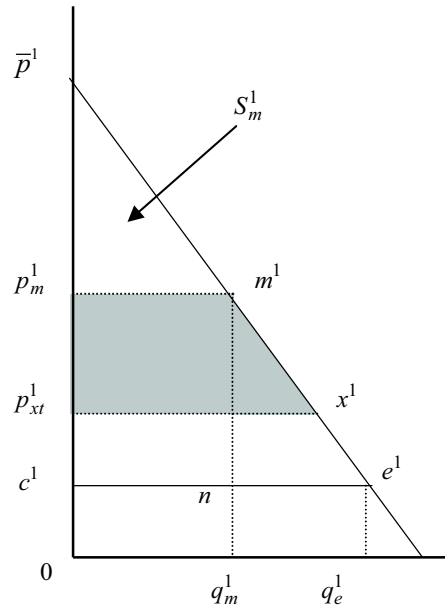


Figure 1a.

Typical buyer's demand for product 1.

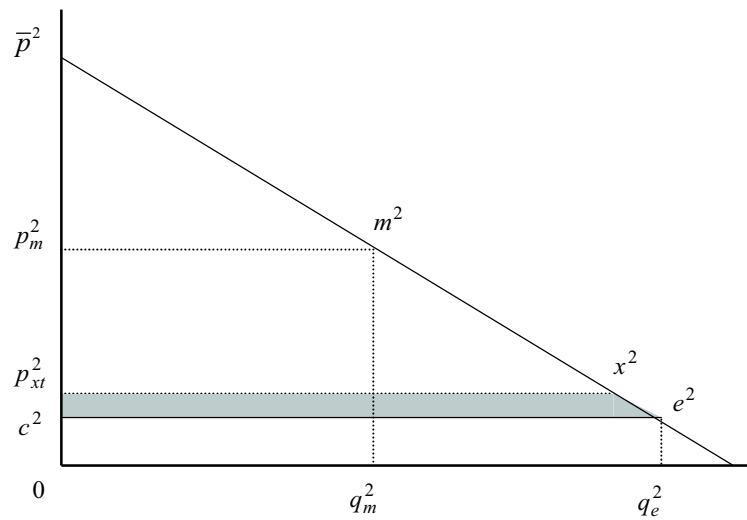


Figure 1b.

Typical buyer's demand for product 2.

none of this product.¹⁵ This area minus the actual payment to the supplier—the triangular area under the demand curve but above a given price—is consumer surplus, the buyer’s net gain from purchasing the particular quantity at the given price. Alternatively, and more reflective of the antitrust tying cases, the buyer may be an intermediary, say a retailer. Following much of the literature, suppose that the retailer is a local monopolist in selling to final consumers. If the retailer can only charge per unit prices, then the height of its demand curve for each product equals its marginal revenue from the sale of that product, net of any other costs; the area under its demand curve up to any quantity q reflects the *retailer’s* maximum willingness to pay for that quantity, but this amount now understates the full social value of that quantity by leaving out consumer surplus of end users.¹⁶ For brevity, we shall use “consumer surplus” to denote the buyer’s surplus also in this case, and “total surplus” or “welfare” to denote the joint surplus of the sellers and the immediate buyer. Let W denote total welfare—consumer surplus plus profit.

Figure 1 shows that simple monopoly pricing fails to maximize profit compared to richer pricing schemes for two quite different reasons: it leaves consumer surplus, and causes a welfare loss from underproduction. The simple monopoly price-quantity point is m^1 in market 1 and m^2 in market 2. These are the price-quantity pairs firm A would choose if it were an unconstrained monopolist in both markets. If firm A faces perfect competition in 2 and cannot tie its products, it stays at m^1 but prices 2 at marginal cost, leading to the socially efficient point in market 2, denoted e^2 . In market 1, the monopoly price leaves consumer surplus equal to the triangle S_m^1 , and a “deadweight loss” equal to the triangle m^1e^1n —the loss in welfare from choosing the monopoly quantity instead of expanding to the efficient point e^1 where demand intersects marginal cost. If firm A could charge a fixed fee as well as a per unit price (or use other forms of nonlinear pricing), it would capture both triangles as increased profit: it would cut price to marginal cost c^1 , expanding consumption and eliminating the deadweight loss, and charge a fixed fee equal to the consumer surplus realized at point e^1 .¹⁷ However, given the assumed restriction to only per unit prices, there is a rent extraction role for tying.

Section 2.1 analyzes a case that Nalebuff calls mixed tying.¹⁸ While the term commonly denotes the sale of goods both independently and jointly, Nalebuff uses the

15. The above interpretation holds under the assumption (which we will maintain throughout) that the consumer’s utility function is quasilinear with respect to money. Independent demands for 1 and 2 emerge if the utility function is also additively separable in products 1 and 2. See Mathewson & Winter, *supra* note 6.

16. If the retailer instead is a perfectly price-discriminating monopolist, then its demand fully reflects the product’s value to final consumers and the welfare analysis proceeds as if the supplier sold directly to them.

17. Outside our simplified full information setting, a monopolist may prefer to set the per-unit price above marginal cost and extract part of its rent through a margin on sales rather than solely through the fixed fee. See JEAN TIROLE, *THE THEORY OF INDUSTRIAL ORGANIZATION* (1988). However, claims of costless exclusion through tying have been made even for this simple setting.

18. Nalebuff, *supra* note 7. Greenlee, Reitman, and Sibley also analyze this case. Patrick Greenlee, David Reitman & David Sibley, *An Antitrust Analysis of Bundled Loyalty Discounts* (U.S. Department of Justice, Discussion Paper 04-13, 2004).

term in a specific way: firm A may offer a tied package but must also continue to offer product 1 independently at the simple monopoly price. Since the old option remains available, the firm will set tied prices that yield the same consumer surplus as without tying; if the tied offer were inferior, it would be rejected. Thus, the firm's gain from mixed tying comes entirely because overall welfare increases compared to independent monopoly pricing, due to expanded consumption of the monopoly good. In contrast, Section 2.2 assumes that the firm can tie and raise the stand-alone price of the monopoly good to a prohibitive level, which amounts to pure tying. Tying now unambiguously harms consumers, since their alternative, stand-alone option has been degraded. Overall welfare, however, can still rise due to the beneficial expansion of the monopoly good, the same force as under mixed tying. (Whether overall welfare rises or falls will depend on the relative decrease in the price of good 1 versus the increase in the price of 2, which in turn depends on factors such as the relative sizes of these two markets.)

Mixed tying and pure tying are useful benchmarks and help identify the welfare effects of tying in the remaining situations. If the stand-alone price of the monopoly good falls (instead of staying the same as assumed under mixed tying), then consumer surplus increases (as does overall welfare in our model). If the stand-alone price rises but not to a prohibitive level, then consumer surplus decreases, but by less than under pure tying.¹⁹

2.1. *Mixed tying: Monopoly product remains available independently at the monopoly price*

Mixed tying is a menu of options presented to the consumer: a tied option and an unbundled one. The unbundled option is assumed to be the combination of point m^1 in market 1, yielding consumer surplus S_m^1 , and point e^2 in market 2, that yields an additional S_e^2 (the triangle $\bar{p}^2 e^2 c^2$), for a total of $S_U = S_m^1 + S_e^2$. To be accepted, any *tied* offer must yield consumer surplus of at least S_U . Thus, firm A 's gain from adding a tied option cannot come from harming the buyer but from mitigating the pricing distortion on good 1, as explained shortly.

Before proceeding, we note two possible justifications for the above assumption that firm A must continue to offer the monopoly good unbundled at its simple monopoly

19. The latter is illustrated in *SmithKline*: Lilly sought to induce hospitals to buy its drug Kefzol instead of SmithKline's competing drug Ancef by offering a "bonus rebate" plan. *SmithKline Corp. v. Eli Lilly & Co.*, 427 F. Supp. 1089, 1106 (E.D. Pa. 1976). It (1) reduced by approximately 3% the preexisting rebate on the "total number of grams of Lilly cephalosporins purchased by the hospital" and simultaneously (2) provided an additional 3% bonus rebate on the hospital's total purchases of cephalosporin "if the hospital buys established minimum quantities (separately established for each hospital) of each of any three of Lilly's five cephalosporin products." *Id.* at 1105. Step (2) reflects a bundled discount, while step (1) reflects an increase in stand-alone prices. By contrast, in *LePage's*, 3M claimed that "every product [was] here available independently on commercially reasonable terms," which might be construed as a claim that no stand-alone price was increased, corresponding to our mixed tying. Petition for Writ of Certiorari at 18, *3M v. LePage's*, 2003 WL 22428375, at *1 (2003) (No. 02-1865).

price, p_m^1 . First, the firm may fear that raising the unbundled price could raise antitrust charges of pure tying. A second possibility emerges if we modify our model slightly. Suppose there are three groups of consumers, with identical members within each: in group 1-and-2, each consumer has demands D^1 and D^2 ; in group 1, each consumer has demand D^1 ; in group 2, each consumer has demand D^2 . If it could price discriminate, firm A would charge group 1 the simple monopoly price p_m^1 and present group 1-and-2 only a tied offer.²⁰ Assuming that price discrimination is not feasible, the profit-maximizing tied prices will be constrained by the unbundled price chosen for good 1. To relax this constraint, firm A will raise this unbundled price somewhat above p_m^1 ; however, as group 1 becomes large relative to combined group 1-and-2, the profit-maximizing unbundled price will approach p_m^1 . Thus, our mixed tying regime is an approximation to the firm's unconstrained optimal pricing when group 1 is relatively large.

We now show that mixed tying can benefit firm A purely by reducing the deadweight loss on product 1. This is done by structuring the tied option as follows: cut the price of good 1 somewhat below p_m^1 and raise the price of 2 somewhat above marginal cost. The deadweight loss will be reduced because starting at the efficient point e^2 in market 2, a small rise in price and fall in quantity reduces welfare W only negligibly (since the consumer's marginal value approximately equals marginal cost), but the quantity expansion in 1 increases W significantly (since marginal value starting at the monopoly quantity is well above marginal cost). Since total welfare can be increased, there exist many alternative price pairs that leave the same consumer surplus S_U as the unbundled option but yield higher profit.²¹ From this set, firm A chooses the pair that maximizes its profit. These *mixed tying prices*, (p_{xt}^1, p_{xt}^2) , lie below the simple monopoly levels and above marginal costs $c^1 < p_{xt}^1 < p_m^1$ and $c^2 < p_{xt}^2 < p_m^2$, henceforth: $(c^1, c^2) < (p_{xt}^1, p_{xt}^2) < (p_m^1, p_m^2)$.

Geometrically, to maintain S_U , the mixed tying prices (p_{xt}^1, p_{xt}^2) must be such that the consumer's gain from the price cut on product 1 (trapezoid $p_m^1 m^1 x^1 p_{xt}^1$ in Figure 1) equals the loss from the price rise on product 2 (trapezoid $p_{xt}^2 x^2 e^2 c^2$). We return to this property of the mixed tying prices when discussing the *Ortho* test. Observe that this pricing pattern still leaves both quantities below their efficient levels and, hence, is still inferior to charging a fixed fee in market 1 and setting both prices at marginal costs.²²

20. See *infra* Section 2.2.

21. The explanation for the existence of such welfare increasing prices is sometimes couched differently: that a small price cut from the monopoly level in market 1 has a negligible (zero to the first order) effect on profit while a small increase starting from marginal cost pricing in market 2 will significantly raise profit. However, this argument does not establish that consumer surplus has not fallen, which is needed for the tie to be accepted.

22. This is why, in the absence of fixed fees, the monopolist would gain by tying to the monopoly good 1 numerous competitive products if possible, and setting lower markups on each of the tied goods so as to reduce the overall consumption distortion. See Burstein, *supra* note 6. Interestingly, 3M tied many different office products to purchases of its generic tape.

2.2. *Pure tying: Monopoly product is only offered tied*

Now consider an alternative case.²³ Firm *A* is not constrained—by fear of antitrust or otherwise—to continue offering product 1 at the monopoly price; it may offer product 1 only tied with 2, or also unbundled but at a price of its choosing. By raising the unbundled price of 1 to a prohibitive level (\bar{p}^1 or higher), firm *A* can de facto adopt pure tying and reduce the consumer surplus under the unbundled option from $S_U = S_m^1 + S_e^2$ to S_e^2 , the level available by buying only product 2 competitively. Since consumers' unbundled option is less attractive under pure tying than under mixed tying, firm 1 sets the pure tying prices (p_1^1, p_1^2) above the mixed tying levels (p_{xt}^1, p_{xt}^2) .

2.2.1. *Effect of pure tying on consumers and overall welfare*

Pure tying captures part or all (cases (1) and (2) below, respectively) of the consumer surplus S_m^1 that was available from buying product 1 unbundled at its monopoly price. The pure tying prices are determined by the following alternative conditions:

- (1) *Unconstrained monopoly*: If the simple monopoly prices (p_m^1, p_m^2) yield consumer surplus of at least S_e^2 , what the consumer would get by rejecting the tie and buying only product 2 competitively, then firm *A* will set its tied prices at the simple monopoly levels in both markets: $(p_1^1, p_1^2) = (p_m^1, p_m^2)$.
- (2) *Constrained monopoly*: If consumer surplus at (p_m^1, p_m^2) is less than S_e^2 , then to yield consumer surplus of S_e^2 firm *A* will set its pure tying prices below the monopoly levels, but still higher than under mixed tying: $(p_{xt}^1, p_{xt}^2) < (p_1^1, p_1^2) < (p_m^1, p_m^2)$.

Case (1) arises, for example, if market 2 is small relative to 1, while (2) arises if market 2 is relatively large.

The possible welfare effects of pure tying are as follows:

- *Consumer surplus under pure tying is lower than under no tying or mixed tying.* This result follows immediately because pure tying raises the prices of both goods relative to mixed tying which, in turn, yields the same consumer surplus as no tying ($S_U = S_m^1 + S_e^2$).
- *Welfare under pure tying is lower than under mixed tying.* This, too, follows because pure tying raises prices above the mixed tying levels, which already exceed marginal costs. Thus, consumption of both goods is reduced further below the efficient levels.
- *Welfare under pure tying can be lower or higher than under no tying.* Welfare is higher under pure tying if the prices are close to marginal costs, as will occur if the tied market 2 is large relative to 1. Tying then lets firm *A* exploit its monopoly power relatively efficiently: it accepts a large price reduction in the monopoly market in exchange for a positive but small margin in the larger market 2. Figure 1 illustrates such a case. Welfare is lower under pure tying if market 2 is relatively small. The pure tying prices then will be near, or equal to,

23. Mathewson & Winter, *supra* note 6; Greenlee et al, *supra* note 18; Joseph Farrell *Deconstructing Chicago on Exclusive Dealing*, 50 ANTITRUST BULL. 465 (2005); Nalebuff, *supra* note 7.

the monopoly levels, so tying reduces price only slightly if at all in market 1 but raises price in market 2 substantially above its no tying level of marginal cost.

2.2.2. *Costless exclusion?*

Case (1) above, unconstrained monopoly, seemingly illustrates costless exclusion through tying.²⁴ By offering the monopoly product 1 only in a tie, firm *A* is able to set the monopoly price also for product 2, while denying sales to single-product competitors there. However, the proper characterization of such an outcome in our view is fuller rent extraction rather than exclusion. By assumption, market 2 here is not susceptible to monopolization. Furthermore, the motive for tying here would vanish if full rent extraction were achievable in market 1: pure tying, even if it permits the monopoly prices in both markets, is only an inferior substitute to charging a fixed fee in market 1. Tying is inferior because it distorts consumption levels away from the first best, whereas a profit-maximizing two-part tariff for product 1 would avoid inefficiency by setting the per unit price at marginal cost and collect the maximal profit through the fixed fee.

2.2.3. *The Greenlee, Reitman, and Sibley test*

Greenlee, Reitman, and Sibley propose a simple test to judge how tying affects consumer surplus: if the unbundled price of the monopoly good rises (perhaps to a prohibitive level) after the introduction of tying, then consumer surplus must fall.²⁵ As explained above, the firm will set its tied prices to leave consumer surplus equal to (or negligibly higher than) the maximal level attainable by buying, instead, at the new unbundled prices. Since the unbundled option has been degraded—the unbundled price of the monopoly product has risen by hypothesis, while the price of the competitive product is unchanged (at marginal cost)—consumer surplus at the tied prices also will be lower. Conversely, if the unbundled price of the monopoly product falls or remains unchanged, then the introduction of tying cannot harm consumers.²⁶

3. The *Ortho* test for predatory tying

The *Ortho* test²⁷ is proposed when a firm offers its products both as a bundle and individually. The test compares the prices to determine whether the discount in the bundle entails pricing below cost and is presumptively predatory. In Section 3.1 we describe the test and show that it performs well in a particular setting: when the goods are used in fixed proportions and when the tying firm and its rivals offer homogeneous

24. See Nalebuff, *supra* note 7, at 326-27, 343. Note that his markets *A* and *B* correspond to our 1 and 2 (we adopt the alternative notation to conform with Bernheim and Whinston, whose work we discuss later). Nalebuff writes: “With tying, the firm . . . can achieve the exclusion at no cost . . . by threatening to raise the à la carte price of *A* while maintaining the optimal price if the customer accepts the tied sale.” *Id.* at 327. He is describing pure tying.

25. Greenlee et al., *supra* note 18.

26. Note that, strictly speaking, this case is inconsistent with Greenlee, Reitman and Sibley’s formal analysis, where a monopolist would always raise its unbundled price after introducing a tied offer. However, a straightforward extension of their model would admit the possibility of mixed tying.

27. See *supra* note 2.

versions of the competitive good. We then consider two departures from this setting and show that in each case a firm can fail the test when tying is not driven by predation.

Section 3.2 maintains fixed proportions but assumes that firms offer differentiated versions of the competitive good 2. The good 1 monopolist can then employ tying essentially as a way to implement (third degree) price discrimination between consumers that choose its version of the competitive good and those who choose the rivals' version, unrelated to predation against the rivals. The test wrongly finds a profit sacrifice from tying, inferring predation, by assuming that each additional unit sold by the monopoly firm in the tied market fully displaces a unit of output from rivals.

Section 3.3 returns to homogeneous products in the tied market but assumes that the tying and tied goods are used in variable proportions, so that comparing prices across goods is not meaningful. We adapt the *Ortho* test to compare, instead, the profit from the tied good with measures of profit sacrifice on the tying good. A fundamental problem with the test is that it overstates the sacrifice by failing to include the profit from expanded sales of the tying good resulting from the tied discount. This overstatement can cause the firm to fail the test even when engaging in mixed tying, which, as we have shown, raises welfare without harming consumers. We illustrate additional reasons why the sacrifice may be overstated. In fact, if demands are linear, a firm engaging in pure tying will always fail the test, even though pure tying in our examples is motivated solely by rent extraction and, a fortiori, is profitable.

3.1. Fixed proportions and homogeneous products in the tied market

Consider a firm that offers its products both independently and at lower tied prices available only if all the products are bought from that firm. Do the tied prices reflect predatory pricing in the target market 2? The *Ortho* test presumes a violation if

$$[\text{sum of all tied prices} - \text{sum of unbundled prices except 2}] < \text{unit cost of 2} \quad (1)$$

The term in brackets is the *implicit price* of good 2 when purchased through the bundled offer. Predation is presumed if this implicit price is below some measure of unit cost (typically, average variable cost or marginal cost). In our two-goods setting, if firm *A* offers tied/bundled prices (p_b^1, p_b^2) and unbundled prices (p_u^1, p_u^2) , the test reads:

$$[p_b^2 - (p_u^1 - p_b^1)] < c^2 \quad (2)$$

The term $(p_u^1 - p_b^1)$ is the implicit sacrifice attributed to good 1 when sold in a tie with good 2 instead of unbundled. The test subtracts this imputed sacrifice from the posted tied price of good 2 to arrive at the effective price for good 2 when sold in the bundle.

The test performs well in the following setting. Assume: (A1) goods 1 and 2 must be consumed in fixed proportions, for simplicity one-to-one, and (A2) firm *A* and its rivals *B* offer homogeneous products in market 2. Rivals' unit cost is c_B^2 and $\Delta \equiv c^2 - c_B^2$ denotes their cost advantage; $\Delta > 0$ means that rivals' cost is lower. Suppose firm *A* offered the same price for good 1 regardless of whether good 2 is bought from it or from rivals and sought to drive out rivals by setting its good 2 unbundled price, p_u^2 , below their cost. This requires $p_u^2 < c_B^2 = c^2 - \Delta$ or $p_u^2 - c^2 < -\Delta$. Thus, to exclude more efficient rivals ($\Delta \geq 0$) without tying, firm *A* would have to set its price below cost.

Now consider how bundled pricing can disguise such below-cost pricing. Let firm A offer both unbundled and bundled prices: (p_b^1, p_b^2) , (p_u^1, p_u^2) . Since goods 1 and 2 are consumed in 1:1 proportions, and since the two versions of good 2 are homogeneous, consumers will prefer to buy both goods from firm A instead of good 2 from rivals at their price p_B^2 and good 1 from firm A at its unbundled price p_u^1 if and only if the sum of the prices is lower:

$$(p_b^1 + p_b^2) < (p_u^1 + p_B^2) \tag{3}$$

Assume that rivals price their good at cost, $p_B^2 = c_B^2 = c^2 - \Delta$. Substituting in Equation (3) and rearranging yields the condition for exclusion to be

$$p_b^2 - c^2 < (p_u^1 - p_b^1) - \Delta \tag{4}$$

Suppose firm A sets its bundled price for good 2 at cost, $p_b^2 = c^2$, passing a standard predation test. It can satisfy Equation (4), thereby displacing in market 2 more efficient rivals who price at cost, by offering a bundled discount on good 1 that exceeds rivals' cost advantage on good 2: $(p_u^1 - p_b^1) > \Delta$. Instead of providing an explicit unbundled discount on good 2 through below-cost pricing, $p_u^2 - c^2 < -\Delta$, firm A now provides a bundled discount on good 1 of the same magnitude that achieves the same exclusionary effect.

Finally, observe that satisfying Equation (4) when $\Delta > 0$ (rivals have lower cost) requires $(p_b^2 - c^2) < (p_u^1 - p_b^1)$, hence failing the *Ortho* test in Equation (2). Moreover, failing the test implies that the firm is indeed pricing below economic cost. Under assumptions (A1) and (A2), firm A incurs an equivalent profit sacrifice by granting (1) an explicit discount on good 2 of size δ , or (2) a bundled discount on good 1 of size δ . (To exclude rivals requires $\delta \geq \Delta$.) The former is an explicit cost; the latter is an opportunity cost—foregone profit on good 1 from selling its own version of good 2 instead of rivals' version along with a higher price on good 1. Jointly, assumptions (A1) and (A2) imply that a consumer chooses between the two versions of good 2 solely based on the total purchase cost of good 2 plus good 1, hence granting a discount on firm A 's version of good 2 is equivalent to granting the same discount by accepting a lower tied price on good 1. In this setting, therefore, the *Ortho* test performs well.

We now show how the *Ortho* test can condemn nonpredatory behavior when assumption (A2) is relaxed. The next section shows a similar result by relaxing (A1).

3.2. Fixed proportions but differentiated products in the tied market

Suppose firms A and B offer differentiated products in market 2, $A2$ and $B2$, with equal unit costs c^2 . Each product is consumed in proportion 1:1 with the monopoly good 1 so, as before, only the sum of the prices for good 1 and good 2 is relevant. This time, however, $A2$ and $B2$ are differentiated, so both can be sold even if their total prices differ. Moreover, since product differentiation is valuable to consumers, firm A has an incentive to let rivals operate in market 2.

Consider first the hypothetical fully integrated benchmark where firm A controls both $A2$ and $B2$, as well as good 1. Let p_A^f denote the total profit-maximizing price to a fully integrated firm for a bundle comprising of goods 1 and $A2$; define p_B^f analogously for a

bundle comprising of goods 1 and $B2$. Suppose demand for $B2$ is “stronger” than for $A2$, in the sense that these bundle prices satisfy $p_b^1 > p_A^1$. Now consider the actual case where firm A is not integrated into $B2$. Suppose for simplicity that $B2$ is competitive (but differentiated from $A2$) hence priced at marginal cost, $p_B^2 = c^2$. Firm A can duplicate the two integrated total prices by appropriately setting three prices as follows: a bundled price p_b^2 for its good $A2$, a bundled price p_b^1 for good 1 purchased with its good $A2$, and an unbundled price p_u^1 available with $B2$, such that

$$(p_b^1 + p_b^2) = p_A^1 < p_B^1 = (p_u^1 + c^2) \quad \text{or} \quad (5)$$

$$(p_b^2 - c^2) < (p_u^1 - p_b^1) \quad (6)$$

Equation (6) implies that the *Ortho* test in Equation (2) above is failed.

The ostensible logic is that firm A incurs a sacrifice by adopting such pricing: $(p_b^2 - c^2)$ is its margin on good $A2$ while $(p_u^1 - p_b^1)$ is the sacrifice it incurs on good 1 if a unit of $A2$ displaces a unit of $B2$. However, with product differentiation (imperfect substitution) between $A2$ and $B2$ there is no longer one-to-one displacement. The pricing in Equation (6) is motivated purely by differences in demand for $A2$ and $B2$, essentially a form of third degree price discrimination in the sale of the monopoly good, not by predation—indeed, in our example, rivals B will survive in the market (albeit still pricing at cost).²⁸ Thus, the *Ortho* test can convict the innocent.

Moreover, aside from not being predatory, the pricing in Equation (6) can increase both overall welfare and consumer welfare. The argument is the same as one commonly used in examples of price discrimination. If a standard predation rule forces firm A to price $A2$ above cost $(p_b^2 - c^2) \geq 0$, and the firm seeks to price as in Equation (6), it will fail *Ortho* if $(p_u^1 - p_b^1) > 0$, that is, if it essentially price discriminates against the independent product ($B2$) in access to the monopoly good. But if demand for $B2$ is sufficiently large relative to $A2$, or if the difference in the desired prices $(p_B^1 - p_A^1)$ is sufficiently large, then firm A will respond to an *Ortho* rule (in this setting, a ban on price discrimination) by dropping its own “weak” product $A2$ and only selling good 1 at the optimal price desired with $B2$. Compared to allowing price discrimination, the firm and consumers are harmed.²⁹

28. The above point tracks a critique of the Efficient Component Pricing Rule in a regulation context where a monopolist over an access good (our good 1) also sells an adjacent good (our $A2$) bundled with access via an affiliate, in competition with differentiated rivals (our $B2$) that also require access. The rule states that the affiliate’s total price to consumers minus its resource cost on the adjacent segment (in our notation, $p_b^1 + p_b^2 - c^2$) should be no lower than the price of access charged to rivals in good 2 (our p_u^1). However, under asymmetric product differentiation where demand for product $B2$ is stronger than for $A2$ in the earlier sense, second best efficient pricing violates the rule. See JEAN-JACQUES LAFFONT & JEAN TIROLE, *COMPETITION IN TELECOMMUNICATIONS* (2000).

29. A slight variation on the example allows the opposite error: that the *Ortho* test would acquit the guilty. Suppose that product $A2$ is stronger than $B2$ in a vertical differentiation sense: at equal prices, all consumers would prefer $A2$. If firm A wishes to engage in predatory pricing, it can price $A2$ at cost and good 1 uniformly (regardless of whether it is used with $A2$ or $B2$), thus passing *Ortho*. Yet $B2$ would be driven from the market. Moreover, with perfectly competitive sellers of $B2$ as assumed, firm A ’s profit-maximizing choice would typically allow them to operate, so as to tap the pricing benefits from vertical differentiation. Thus, profit-sacrificing and welfare-reducing predation can pass *Ortho*.

3.3. Variable proportions and a revenue version of the *Ortho* test

When the products are demanded in variable rather than fixed proportions, the *Ortho* test suffers from a deeper problem. Because there is no fixed amount of good 1 associated with a unit of good 2, comparing prices will not give an unambiguous measure of the sacrifice per unit of good 2 incurred by offering a bundled discount on good 1. For example, the test in Equation (2) finds a violation if $(p_b^2 - c^2) < (p_u^1 - p_b^1)$. This comparison will depend on the units that are being bundled. For example, suppose that instead of measuring good 2 in pounds we measure it in tons; then the term $(p^2 - c^2)$ will become 2,000 times larger. This is innocuous under fixed proportions, because the right side will also have to be 2,000 times larger, but is problematic with variable proportions.

An obvious alternative is to compare, instead of prices, the relevant revenues and variable costs. In the two-good case, this involves comparing the total revenue and total variable cost of good 2 when sold in a tie with good 1 against the revenue sacrifice on good 1 implied by the tied discount. Thus, we replace condition (2) with

$$[\text{revenue from 2} - \text{total variable cost of 2}] < [\text{revenue sacrifice on 1}] \quad (2')$$

Observe that, like the price version of the test, this version overstates firm *A*'s actual profit sacrifice from offering the tied discount in market 1 by ignoring its gain from the output expansion induced by the discount.³⁰

3.3.1. Mixed tying

We first examine the properties of the *Ortho* test when the firm engages in mixed tying, which benefits the firm without harming consumers (see Section 2.2). Firm *A* is initially an unconstrained monopolist in both markets, charging the monopoly prices. Now a fringe of competitors enter into market 2 and set price equal to marginal cost, c^2 . Suppose that firm *A* must continue to make available on a stand-alone basis its good 1 at the monopoly price. Referring to Figure 1, consumers' unbundled option consists of the monopoly point m^1 in market 1 and the efficient point e^2 in market 2, yielding a total consumer surplus of $S_U = S_m^1 + S_e^2$. We saw that firm *A* can increase its profit by adding a tied option that yields the same consumer surplus S_U (hence increases overall welfare) and whose mixed-tying prices lie below the simple monopoly levels and above marginal costs: $(c^1, c^2) < (p_{xt}^1, p_{xt}^2) < (p_m^1, p_m^2)$.

Will a firm engaging in the above mixed tying pass the *Ortho* test in Equation (2')? The left term in Equation (2'), firm *A*'s revenue minus variable cost in market 2, equals

However, this result does not hinge on the bundling aspect; it merely arises because price-cost tests for predation are inherently imperfect when the firms offer differentiated products.

30. This point was recognized by the district court in the *Ortho* case:

Ordovery's testimony assumes that the effect of the package pricing was simply to reduce the revenues generated by sales of HTLV and HIV-12 by an amount equal to the difference in the bundled and unbundled prices of those assays. . . . [But] Abbott's CCBC [package] pricing reasonably may be expected to have increased its unit sales of all of the assays and thereby to have generated added profits offsetting, at least in part, the revenue loss attributable to the price cuts. Ordovery's deposition testimony does not confront this matter with any specificity.

Ortho Diagnostic Sys. v. Abbott Labs., 920 F. Supp. 455, 470 (S.D.N.Y. 1996).

$(p_{xt}^2 - c^2)q_{xt}^2$, where q_{xt}^2 is the mixed tying quantity (not drawn in Figure 1, but corresponding to the point x^2), that lies above the monopoly quantity q_m^2 but below the efficient quantity q_e^2 . The right term in Equation (2'), the imputed revenue sacrifice in market 1 from granting the tied discount, is computed as $(p_m^1 - p_{xt}^1)q^1$, where q^1 is firm A 's output level in market 1 used to assess the sacrifice, discussed momentarily. Thus, the *Ortho* test in Equation (2') is failed if

$$(p_{xt}^2 - c^2)q_{xt}^2 < (p_m^1 - p_{xt}^1)q^1 \quad (7)$$

There is ambiguity about which level of q^1 should be used. We analyze two alternatives, q^1 is taken to be the old monopoly output q_m^1 or the new higher output q_{xt}^1 .

Firm A sets the tied discount on good 1, $(p_m^1 - p_{xt}^1)$, such that the gain in consumer surplus, denoted ΔS^1 , equals the amount lost in consumer surplus from the price increase in market 2 (relative to the unbundled option of buying from rivals at cost), denoted ΔS^2 . Referring to Figure 1, ΔS^1 equals the area of the trapezoid $p_m^1 m^1 x^1 p_{xt}^1$ and ΔS^2 is the trapezoid $p_{xt}^2 x^2 e^2 c^2$. Decompose each trapezoid into a "rectangle" and "triangle" and use the equality of ΔS^1 and ΔS^2 to yield

$$\Delta S^1 = (p_m^1 - p_{xt}^1)q_m^1 + t^1 = \Delta S^2 = (p_{xt}^2 - c^2)q_{xt}^2 + t^2 \quad (8)$$

where t^1 is the consumer surplus "triangle" gained in market 1 solely from the quantity expansion and t^2 is the consumer surplus "triangle" lost in market 2 solely from the quantity reduction, both measured in absolute value. Rearranging Equation (8) gives

$$(p_{xt}^2 - c^2)q_{xt}^2 = (p_m^1 - p_{xt}^1)q_m^1 + (t^1 - t^2) < (p_m^1 - p_{xt}^1)q_{xt}^1 \quad (9)$$

where the inequality follows since $q_m^1 < q_{xt}^1$, hence the trapezoid $(p_m^1 - p_{xt}^1)q_m^1 + t^1$ is contained in the rectangle $(p_m^1 - p_{xt}^1)q_{xt}^1$. To determine the test outcome, we use Equation (9) to help us apply Equation (7), with the output q^1 taken as either the new level q_{xt}^1 or the old level q_m^1 .

Old output: $q^1 = q_m^1$. The equality in Equation (9) implies that

$$(p_{xt}^2 - c^2)q_{xt}^2 < (p_m^1 - p_{xt}^1)q_m^1 \text{ if and only if } t^1 < t^2 \text{ (fail Ortho)}$$

Thus, the firm will fail if $t^1 < t^2$ and pass if $t^1 \geq t^2$. In general, either outcome is possible. For example, if demands are linear and have equal vertical intercepts but different slopes, the test is failed if demand in market 2 is sufficiently steeper than in market 1 (i.e., demand in 2 is sufficiently smaller) and is passed otherwise.³¹

New output: $q^1 = q_{xt}^1$. The inequality in Equation (9) shows that the *Ortho* test is unambiguously failed. Besides ignoring the gain $(q_{xt}^1 - q_m^1)(p_{xt}^1 - c^1)$ to firm A from the expansion in sales of good 1 resulting from mixed tying, the test now overstates the revenue sacrifice on good 1: the price cut $(p_m^1 - p_{xt}^1)$ is attributed also on the output

31. The geometric reason is that when market 2 is small relative to 1, maintaining consumer indifference requires a relatively small price cut in market 1 compared to the rise in market 2, and each triangle is proportional to the square of the size of the price change. (This effect outweighs the fact that demand in market 1 is larger, i.e., flatter demand.)

expansion $(q_{xt}^1 - q_m^1)$, an expansion that is made possible only because of the tied discount. This added bias causes the firm to fail the test.

These findings are summarized in the following proposition.

Proposition 1. Suppose the firm engages in mixed tying.

1. If the revenue sacrifice on the tying good is assessed at the old, pretying output, $q^1 = q_m^1$, the revenue version of the *Ortho* test, Equation (7), is passed if and only if $t^1 < t^2$.
2. If the revenue sacrifice is assessed at the new, posttying output, $q^1 = q_{xt}^1$, the test in Equation (7) is failed always.

Intuitively, a fundamental problem with the test is that when assessing the “sacrifice” on product 1 embodied in tying it with 2, the test ignores the profit from increased sales of product 1. In addition, if the sacrifice is assessed on the posttying output level (case 2) in the proposition), the test overstates the revenue sacrifice from the price cut as well. As a result, the test can condemn tying that is both profitable for the firm and innocuous for consumers.

3.3.2. Pure tying

Instead of mixed tying, suppose firm 1 engages in pure tying, as discussed in Section 2.2. The unbundled price of good 1, p_u^1 , is set above the monopoly level in order to reduce the attractiveness of the unbundled option, allowing the pure tying prices to be raised above the mixed tying levels: $(p_t^1, p_t^2) > (p_{xt}^1, p_{xt}^2)$. The pure tying prices will be either (1) at the monopoly levels, $(p_t^1, p_t^2) = (p_m^1, p_m^2)$, if market 1 is large relative to 2; or (2) below them otherwise, $(p_t^1, p_t^2) < (p_m^1, p_m^2)$. In either case, consumer surplus under pure tying is lower than under no tying (the latter yields the same consumer surplus as mixed tying). However, total surplus, W , can be higher under pure tying than under no tying in case (2), provided market 2 is sufficiently larger than 1. We show that, at least with linear demands, the firm using pure tying will always fail the *Ortho* test in case (2)—even though welfare rises in some such cases (false positive).

In case (2), the unbundled option constrains firm 1’s pure tying prices, hence firm A will set its unbundled price for good 1 no lower than the choke level, $p_u^1 \geq \bar{p}^1$. Consider $p_u^1 = \bar{p}^1$; any higher price makes it easier to fail the test (since a higher unbundled price is taken to show a larger sacrifice from tying). The revenue version of the test becomes Equation (7’):

$$(p_t^2 - c^2)q_t^2 < (\bar{p}^1 - p_t^1)q^1 \tag{7’}$$

This is analogous to Equation (7) except for replacing the mixed tying variables by their pure tying counterparts, and the unbundled price p_m^1 by \bar{p}^1 . The consumer indifference condition used to determine the pure tying prices in case (2) is

$$\Delta S^2 = \Delta S^1 \Leftrightarrow (p_t^2 - c^2)q_t^2 + t^2 = S_m^1 + (p_m^1 - p_t^1)q_m^1 + t^1 \tag{8’}$$

where ΔS^1 and ΔS^2 are the changes in consumer surplus (in absolute value) moving from the unbundled option to pure tying; the gain in market 1 is decomposed as the consumer surplus attained at the monopoly price, S_m^1 , plus the gain from cutting price to

$p_t^1 < p_m^1$, which is the sum of a rectangle and the “triangle” t^1 . Using Equation (8') in Equation (7') shows that the test is failed under pure tying if

$$S_m^1 + (p_m^1 - p_t^1)q_m^1 + (t^1 - t^2) < (\bar{p}^1 - p_t^1)q^1 \quad (9')$$

where q^1 is the output level in market 1 used as the base for computing the sacrifice from the tied discount $(\bar{p}^1 - p_t^1)$. Suppose the old (pretying) output is used: $q^1 = q_m^1$. Writing $(\bar{p}^1 - p_t^1)q^1 = (\bar{p}^1 - p_m^1)q_m^1 + (p_m^1 - p_t^1)q_m^1$, Equation (9') can be expressed as

$$S_m^1 + (t^1 - t^2) < (\bar{p}^1 - p_m^1)q_m^1,$$

which is met with linear demand in market 1 since $(\bar{p}^1 - p_m^1)q_m^1 = 2S_m^1$ and $t^1 < S_m^1$ given $p_t^1 > c$. Thus, the test is failed. A fortiori, the test also will be failed if the market 1 sacrifice is assessed using the new (tying level) output, $q^1 = q_t^1 > q_m^1$, since the right term in Equation (9') will be larger.

These findings are summarized in the following proposition.

Proposition 2. Suppose the firm engages in pure tying and its tied prices are below the simple monopoly levels (the constrained monopoly case noted in Section 2.3.1). If demands for the products are linear, then the revenue version of the *Ortho* test, Equation (7), is failed always, even though welfare sometimes increases.

4. Why tying? A closer look

Although tying (in the sense of a tie-out) can extract more rent than linear stand-alone pricing, in environments where tying is feasible there often exist alternative pricing instruments that would yield greater rent. This observation might be thought to suggest ulterior motives for tying; however, similar arguments indicate that exclusionary outcomes may also be more effectively achieved with instruments other than tying. Thus the motivations for tie-outs remain a significant open question.

4.1. Is tying useful for rent extraction when fixed fees are feasible?

Our analysis thus far has assumed that the tying firm can only charge per unit prices, it is unable to charge also a fixed fee, or to use other forms of nonlinear pricing. While there may exist circumstances where the only feasible contract in the monopoly market is a per unit price, Mathewson and Winter point out that such conditions are inconsistent with the premise that tying is feasible.³² If tying can be imposed, then a fixed fee can be charged in the tied market simply through a price-quantity contract: by making good 1 available only if the customer also buys *any* specified quantity of good 2 at a suitably inflated price, the monopolist can essentially levy a fixed fee for the right to purchase good 1. A fixed fee can be charged by specifying even a minor quantity; we shall call such a contract a minimal tie-in to distinguish it from a tie-out, whereby the customer is required to buy all (or a very high share of) its good 2 purchases from this firm. A minimal tie-in contract seems no harder to enforce than a tie-out.

Furthermore, with identical buyers and complete information, a fixed fee implemented through a minimal tie-in avoids the inefficiencies created by a tie-out, and

32. Mathewson & Winter, *supra* note 6.

therefore allows the firm to extract the maximal rent for its good 1 monopoly. A minimal tie-in leaves consumers' quantity choices in the other market undistorted at the margin, while a tie-out distorts their choices, which reduces the total rent that can be extracted. Thus, Mathewson and Winter conclude that tying—in the sense of a tie-out—would not be chosen for rent extraction purposes in the case of identical buyers.³³ Whenever such a contract is feasible, it is dominated by a simple price-quantity contract in the tied market.

The same conclusion holds with heterogeneous buyers if the seller is informed about their individual demands and can offer them customized contracts with differing fixed fees (or other customized nonlinear contracts). Indeed, in some of the antitrust cases noted earlier, customized nonlinear contracts were offered to different buyers (e.g., *SmithKline, LePage's v. 3M*). Full tying or approximations thereof, whereby discounts were conditioned on the buyer sourcing from that firm a very high share of its total purchases, were added—according to plaintiffs—only in response to entry (e.g., *SmithKline*) or to an increase in the competitive threat posed by a growing rival (e.g., *LePage's v. 3M*).

Mathewson and Winter argue that a rent extraction role for tying can resurface even if the firm can offer a two-part tariff to all customers, provided there is asymmetric information about demands. Specifically, suppose only the buyer knows his level of demand for good 1 (the buyer's "private information"), but the monopolist knows that demands for goods 1 and 2 are affiliated (loosely, a customer with relatively high demand for one good is also more likely to have relatively high demand for the other). Tying the competitive product and pricing it above marginal cost then can extract additional rent (and at least some customer types are better off with tying). Note the importance of the assumption that the two demands are statistically linked (i.e., affiliated). If they are statistically independent, Mathewson and Winter show (for a particular example) that tying for this purpose is ineffective.

Affiliated demands across otherwise unrelated products therefore expands the universe where tying may be used for rent extraction beyond the traditional case of direct metering, where a fixed good is consumed together with variable quantities of a complementary good, such as a camera and film. Observe, however, that when tying is used to exploit affiliation in demands, such a motive is more accurately portrayed as rent extraction rather than exclusion. Mathewson and Winter's framework, like ours, precludes monopolization of market 2 by assumption. The purpose for tying is to more fully extract the rent available due to the monopoly over good 1 and this motive is present independent of whether conditions in market 2 are conducive to foreclosure.

Observe also that the affiliated demands explanation for tying fits less well in environments where share contracts are used, as was true in some of the pertinent antitrust cases. In order to verify that the buyer is purchasing from firm *A* the specified share of its total purchases of the tied product, firm *A* must also observe the buyer's purchases from other firms. But in such a case, firm *A* may be able to use total purchases of the tied product as a signal about the buyer's demand for the monopoly good and adjust the total payment accordingly; it may not have to specify that a high

33. *Id.*

share in the tied market be bought from it, or require a complete tie-out. In short, outside of metering contexts, if the goal is purely rent extraction from the monopoly good, then achieving it does not require—and may conflict with—forcing a buyer to obtain a large percentage of its total requirements of the competitive good from that same firm.

4.2. *Tying for monopolization*

Section 5 demonstrates that requiring customers to buy a large share of their demand for the competitive good from one firm—quantity forcing—is consistent with a monopolization motive if rivals face economies of scale instead of constant unit costs as assumed thus far. If explicit exclusive dealing requirements are not feasible, denying scale economies to rivals may have to be done indirectly, by requiring consumers to purchase a large share of their requirements from the would-be excluder.

While monopolization through quantity forcing must be taken seriously in the presence of large scale economies, it is worth stepping back and asking whether quantity forcing in the target market is aided by tying another product and, if so, how. The economic literature suggests reasons why a firm's presence in one market (our market 1) can increase its gain from monopolizing another market (market 2) or its ability to do so.³⁴ These reasons often hinge on the firm being a threatened monopolist in one market and seeking to impede entry there by monopolizing an adjacent product that is strongly complementary and therefore needed by competitors (or their customers) to contest the monopoly.³⁵ In such cases, tying can be analyzed as an attempt to drive out rivals in the complementary product through a refusal to deal in the monopoly good rather than through quantity forcing in the target market.

Many of the articles on exclusion through tied discounts, however, as well as some of the cited cases, do not involve strongly complementary goods.³⁶ We therefore abstract from these scenarios by assuming that the products are unrelated in both costs and demands. For independent goods, we ask: why might the firm prefer to implement

34. For a comprehensive survey, see the report for the European Commission's Directorate-General for Competition authored by JEFFREY CHURCH, *THE IMPACT OF VERTICAL AND CONGLOMERATE MERGERS ON COMPETITION* (2004).

35. Recent contributions include Jeffrey Church & Neil Gandal, *Systems Competition, Vertical Merger and Foreclosure*, 9 J. ECON. & MGMT. STRATEGY 25 (2000); Jay Pil Choi & Christodoulos Stefanides, *Tying, Investment and the Dynamic Leverage Theory*, 32 RAND. J. ECON. 52 (2001); Dennis Carlton & Michael Waldman, *The Strategic Use of Tying to Preserve and Create Market Power in Evolving Industries*, 33 RAND. J. ECON. 194 (2002); Patrick Rey & Jean Tirole, *A Primer on Foreclosure* (Institut d'Economie Industrielle, Working Paper, 2005). Nalebuff presents a model where tying deters entry when demand for the two products are independent. Nalebuff, *supra* note 7. For a critique, see Timothy Brennan, *Competition as an Entry Barrier? Consumer and Total Welfare Benefits of Bundling* (AEI-Brookings Joint Center for Regulatory Studies, Related Publication 05-08, June 2005).

36. *Ortho* is a case where strong complements were tied and, indeed, among the plaintiff's claims was an assertion that Abbott anticompetitively attempted to maintain its monopoly position in some of the blood tests. However, the bulk of the opinion focused on foreclosure in the nonmonopolized market. See *Ortho Diagnostic Sys. v. Abbott Labs.*, 920 F. Supp. 455 (S.D.N.Y. 1996).

quantity forcing in the target market by granting discounts off other tied products rather than providing inducements solely in the target market?

Whinston proposes one explanation: by making its monopoly product available only bundled with the competitive one, the firm is essentially committing itself to charge a lower price in the second market, a commitment which is profitable if and only if it induces the rival to exit or refrain from entering.³⁷

Another possible explanation is that providing the discounts in other markets may make the inducement offered for expanding purchases in the target market less transparent to antitrust enforcers, especially if the discount is spread over multiple markets. This, in turn, may help escape condemnation of a discount as predatory.

Aside from these two scenarios, however, it is not evident that tying provides a cheaper way to implement quantity forcing than through direct incentives in the target market. The next section considers such quantity forcing in a single product setting.

5. Quantity forcing and exclusion via nonlinear pricing of a single product

The basic idea for how quantity forcing can be used to exclude rivals in the presence of scale economies can be traced back to the limit output model of Bain and Sylos-Labini, as explicated by Modigliani.³⁸ There, an incumbent monopolist deters entry by committing to an output level that depresses the residual demand left for entrants just below their average cost curve. That analysis, however, did not explicitly analyze the contracting process, in particular, why far-sighted buyers would willingly participate in a scheme that deters competition.

Subsequent authors showed that exclusion could succeed despite the harm to at least some buyers, due to lack of buyer coordination. We first review this logic when a firm can selectively offer contracts that require customers not to buy from other suppliers. This practice was alleged in *LePage's*, where 3M was described as selectively making “outright payments to distributors—or offered special prices on Scotch tape—in exchange for exclusivity.”³⁹ We then illustrate Bernheim and Whinston’s argument that the divide-and-conquer logic can make exclusion profitable also when, instead of

37. Michael Whinston, *Tying, Foreclosure and Exclusion*, 80 AM. ECON. REV. 837 (1990). In Whinston’s example, tying cannot be for rent extraction, since in his monopoly market consumers have identical unit demand and linear pricing then suffices to extract all consumer surplus. Tying indirectly commits the good 1 monopolist to charge a lower price for good 2 because when sales of 1 are made only in a bundle with 2, losing a sale of 2 means also losing the monopoly margin on 1. To prevent this, the firm sets the implicit price of good 2 (in the bundle) lower than the explicit price it would charge for 2 if selling the goods separately. As Whinston points out, this argument requires that the firm be able to commit to sell the products only as a bundle. Given such commitment ability, tying may enable costless exclusion because the tie may not harm the firm’s profit when it is the sole seller of both goods, in the special case where all consumers demand both goods. However, tying would still fail the “no economic sense” test for exclusionary conduct advanced by Melamed, Werden, and others because tying would reduce the firm’s profit if the rival were not excluded. See Melamed, *supra* note 3; Werden, *supra* note 3.

38. Franco Modigliani, *New Developments on the Oligopoly Front*, 66 J. POL. ECON. 215 (1958).

39. See Brief for Respondents in Opposition at 4, *3M v. LePage's*, 2003 WL 22428377, at *1 (2003) (No. 02-1865).

explicit exclusivity, a firm uses quantity forcing—offers nonlinear contracts that induce high purchases from it.⁴⁰ Such practices also were alleged in *LePage's*: 3M was said to have targeted its bundled discounts to “certain” customers⁴¹ and set “targets individually for each of the distributors that were essential to LePage’s and calculated the targets to make it impractical for the distributor to meet while retaining LePage’s as a supplier.”⁴²

5.1. Scale economies and divide-and-conquer via exclusive dealing contracts

Rasmusen, Ramseyer, and Wiley showed how a firm could profitably foreclose competitors by offering to some buyers a lump sum payment in exchange for a commitment to buy only from it—naked exclusion.⁴³ In their model, there is an initial monopolist, firm M , but in the second period a second firm, E , can enter with an identical product. Each firm’s marginal cost is constant at a level c for outputs beyond some minimum level Q^* but infinite at lower outputs.⁴⁴ Thus, Q^* is the minimum viable scale (MVS). Market demand at price c exceeds $2Q^*$, so the market can support two competing firms with equal shares (and, because of the discontinuous nature of scale economies, both firms would still operate at minimum unit cost). If entry occurs, the market is split evenly and homogeneous products Bertrand competition drives both prices to c . If firm M remains a monopolist, it can only charge linear pricing and sets the simple monopoly price m , depressing consumption and yielding the familiar monopoly deadweight loss.

Suppose firm M can offer any buyer a lump sum payment for committing to buy only from it in the second period. Let X^* denote a buyer’s loss of consumer surplus due to a price increase from the competitive level c to the monopoly level m (geometrically, X^* is a trapezoid). If a single buyer’s demand at price c , denoted $d(c)$, exceeds MVS, $d(c) \geq Q^*$, then exclusion is unprofitable. Any individual buyer can then support entry, so exclusion would require signing up all buyers and paying each X^* ; but since monopoly pricing reduces overall welfare, the per buyer monopoly profit π falls short of X^* . However, if two or more buyers are needed to support entry, then divide-and-conquer could make exclusion profitable. If scale economies are sufficiently large, the excluder can afford to lose money on each buyer to whom it pays X^* because it will extract the monopoly profit π from all buyers, including those who are not compensated.

The role of scale economies can be seen as follows. Let N denote the number of independent, identical buyers. The minimal number of buyers firm M must sign up to

40. Bernheim & Whinston, *supra* note 11.

41. Brief of Amici Curiae Morgan Stanley, Procter & Gamble Company, and Coca-Cola Company in support of Petition for Certiorari at 2, 3M v. LePage’s, 2003 WL 22428378, at *1 (2004) (No. 02-1865).

42. See Brief for Respondents in Opposition at 5, 3M v. LePage’s, 2003 WL 22428377, at *1 (2003) (No. 02-1865).

43. Rasmusen et al., *supra* note 9. Segal and Whinston refined the analysis, and showed that the ability to offer discriminatory contracts across buyers is necessary for profitable exclusion. Segal & Whinston, *supra* note 9; see also Innes & Sexton, *supra* note 9.

44. Thus, the average and marginal cost is L-shaped; this is an extreme representation of scale economies, but the results do not hinge on this simplifying assumption.

foreclose entry, N^* , is determined by the condition $(N - N^*)d(c)/2 = Q^*$; total demand by “free” buyers, $(N - N^*)d(c)$, of which 1/2 would go to the entrant, equals MVS. Therefore $N^* = N - 2Q^*/d(c)$. Firm M can sign up N^* buyers to exclusivity if it pays them a total of N^*X^* . Its second period profit from monopoly pricing to *all* buyers is $N\pi$. Thus, exclusion is profitable if $N^*X^* \leq N\pi$, or $N^*/N \leq \pi/X^*$, the fraction of all buyers that must be signed is less than the ratio of monopoly profit per buyer to a buyer’s loss in consumer surplus from monopoly pricing. Substituting for N^* in $N^*/N \leq \pi/X^*$ yields, after some algebra, that exclusion is profitable provided MVS as a fraction of market demand at the competitive price exceeds a given threshold:

$$\frac{Q^*}{Nd(c)} \geq \frac{1}{2} \left[1 - \frac{\pi}{X^*} \right]$$

For example, if each buyer’s demand $d(p)$ is linear, then $\pi/X^* = 2/3$, so exclusion is profitable provided MVS relative to the competitive quantity exceeds 1/6.⁴⁵

Segal and Whinston and Innes and Sexton stress that profitable exclusion hinges on the ability to make discriminatory offers across buyers.⁴⁶ Even when all buyers accept exclusivity, this outcome is driven by buyers’ recognition that if they rejected exclusivity, the firm could and would find it profitable to pay others for exclusivity. The role of discriminatory offers in enabling divide-and-conquer recurs in the analysis of Bernheim and Whinston, discussed below.

Finally, in Rasmusen, Ramseyer, and Wiley’s analysis the excluding firm enjoys a first mover advantage in contracting: it can approach all buyers before the potential entrant arrives on the scene. Spector allows the entrant to compete for contracts.⁴⁷ He finds that if the firms are sufficiently asymmetric, exclusion can still be profitable and the outcome is ex post monopoly pricing. However, the rents from monopolization are captured not solely by the excluder but also by a subset of the buyers who are signed to exclusivity, a finding reminiscent of Bernheim and Whinston’s analysis as well as that of O’Brien and Shaffer.⁴⁸

5.2. Exclusion via quantity forcing

Rasmusen, Ramseyer, and Wiley’s analysis is relevant because it highlights the role of divide-and-conquer. But it must be adapted in important ways to address the antitrust cases discussed in Section 1. There, the contested practices did not involve explicit

45. The above condition was derived assuming that the excluder must pay X^* to each of N^* buyers. Rasmusen, Ramseyer, and Wiley show that there exists equilibria where firm M excludes entry while paying virtually nothing, because each buyer believes he is not pivotal to the entrant’s success and therefore accepts exclusivity with M for the slightest payment. Segal and Whinston argue (for the case where contracts are offered to buyers simultaneously rather than sequentially) that these costless exclusion equilibria fail to satisfy a commonly applied refinement (coalition proofness) that is satisfied by the costly (but profitable) equilibrium.

46. Segal & Whinston, *supra* note 9; Innes & Sexton, *supra* note 9.

47. David Spector, *Exclusive Contracts and Demand Foreclosure* (Paris Sciences Economiques, Working Paper, 2006).

48. Daniel O’Brien & Greg Shaffer, *Nonlinear Supply Contracts, Exclusive Dealing, and Equilibrium Market Foreclosure*, 6 J. ECON. & MGMT. STRATEGY 755 (1997).

exclusive dealing, i.e., prohibitions on also buying from rivals. Rather, demand for rivals' products was depressed *indirectly* through contracts that provided strong incentives for some buyers to expand purchases from the alleged excluder: quantity forcing. Furthermore, as noted above, the framework of Rasmusen, Ramseyer, and Wiley grants the excluder a strong first mover advantage, at odds with cases discussed earlier, where the rival producers were already present and could offer their own quantity-forcing contracts. Finally, note that the use of quantity-forcing contracts implies that firms in such cases are, in practice, not limited to only linear pricing, as assumed by Rasmusen, Ramseyer, and Wiley. To analyze exclusion through quantity forcing, therefore, requires a framework where rival firms compete by offering nonlinear pricing contracts.

5.2.1. *Nonlinear pricing to a single buyer*

O'Brien and Shaffer and Bernheim and Whinston analyze the following situation. Firms A and B compete with differentiated products: their products are imperfect substitutes. Suppose for now that they sell to a single buyer. The buyer may be a local monopolist dealer; it demands both products, reflecting the diverse preferences of its final customers. For simplicity, assume that both firms have constant (but possibly different) marginal costs. Both firms (indexed below by $J = A, B$) are able to offer general nonlinear contracts.

Let $T_J(q_J)$ the total payment to firm J as a function of the quantity purchased from it, q_J , and q_J^t denote a specified target quantity. Let p_J denote the per unit price charged by firm J and F_J denote a fixed fee payment to firm J . We can focus on two types of contracts:

$$\text{Two-part tariff: } T_J(q_J) = F_J + p_J q_J$$

$$\text{All-or-nothing: } T_J(q_J) = F_J^t \text{ if } q_J = q_J^t \text{ and } \infty \text{ for any } q_J \neq q_J^t$$

A two-part tariff gives the buyer the flexibility to select any quantity while paying a constant per unit price and a fixed fee.⁴⁹ The particular forcing contract above allows a target quantity q_J^t for a fixed payment F_J^t and any other quantity is refused (i.e., requires an infinite payment). Less extreme contracts can also implement quantity forcing. For example, if the buyer chooses any quantity below q_J^t the per unit price is p ; but if it meets or exceeds q_J^t , the price drops to $p - d$ on *all purchases*, the so-called all units discounts. This amounts to cutting the marginal price to $p - d$ once the target is reached (as with ordinary sliding scale pricing) *and*, in addition, refunding the buyer a lump sum equal to $q_J^t d$ for reaching the target.

When such contracts are available, the standard monopoly deadweight loss arising under linear pricing will be eliminated; a seller and a buyer will choose the quantity that maximizes their joint surplus and use a fixed fee (or other nonlinear pricing) to divide

49. With constant marginal costs, two-part tariffs that set $p_J = c_J$ suffice to achieve the efficient quantities. If marginal costs are not constant, the natural extension is a contract $F_J + C_J(q_J)$ where $C_J(q_J)$ is the firm's total variable cost of supplying the associated output. Bernheim and Whinston call these "sellout contracts."

this surplus.⁵⁰ Thus, potential efficiency losses from exclusion in this setting arise not from monopolistic per unit pricing (as in Rasmusen, Ramseyer, and Wiley) but from diminished product variety. Denote by q_A^s the quantity of product A that would maximize the joint surplus of the buyer and firm A when its product is the single one available: q_A^s is the quantity where firm A 's marginal cost intersects the buyer's demand curve for A derived conditional on buying none of B . Let (q_A^c, q_B^c) denote the quantities that would maximize the common surplus of the buyer and *both firms* (equivalently, of the buyer and an integrated firm that controlled both products) if efficiency requires carrying *both* products. If, instead, scale economies (e.g., due to fixed costs) make it jointly efficient to carry only one product, let it be product A , and the efficient quantity q_A^s defined earlier.⁵¹ Since the products are substitutes, $q_A^s > q_A^c$. Two-part tariffs suffice to implement the efficient solution in either case: if both firms offer per unit prices equal to marginal costs and fixed fees needed to cover at least their fixed costs, the buyer would choose (q_A^c, q_B^c) or $(q_A^s, 0)$, depending on which option yields higher joint surplus.

Foreclosure is said to occur when the jointly efficient outcome involves sale of both products, (q_A^c, q_B^c) , but firm A instead induces exclusion of product B . In the case of a *single* buyer, O'Brien and Shaffer show, by example, that foreclosure through quantity forcing is possible sometimes—if the incremental contribution of product B to the joint surplus of firm B and the buyer is not positive, given that the buyer is purchasing q_A^s .⁵² Foreclosure in this case is still inefficient overall, taking into account the loss to the excluded rival, because the incremental value of product B to the buyer is artificially depressed by firm A “forcing” an excessive quantity of its product, q_A^s instead of q_A^c . This foreclosure outcome can be implemented, for example, using the forcing contract described earlier, by specifying the target quantity as q_A^s .

O'Brien and Shaffer and Bernheim and Whinston do not analyze market share contracts. However, such contracts also can potentially achieve foreclosure. Let product A 's market share in the nonforeclosure equilibrium be $s = q_A^c / (q_A^c + q_B^c)$. Suppose firm A offers the buyer an inducement to buy at least a higher share than s from firm A . To meet this higher share target, the buyer must either raise its purchases from A , which indirectly reduces its demand for B , or directly reduce its purchases from B . Both effects reduce the sales of product B , possibly rendering it nonviable.

Even though foreclosure can be feasible, both O'Brien and Shaffer and Bernheim and Whinston show that foreclosure will not be profitable to firm A in the single buyer setting. Whenever it is jointly efficient for the buyer and the two firms to have both

50. Recall from Section 2 that in the case of an intermediate buyer rather than final consumers, we use total surplus to denote the joint surplus of the sellers and the immediate buyer.

51. The identity of the stronger product will depend on both demand and cost differences. There is, of course, no loss of generality in assuming that product A is the stronger one—this is just a matter of labeling.

52. Observe that, if exclusive dealing is feasible, foreclosure can always occur: firm A could stipulate exclusive dealing, and offer the bilaterally efficient output for it and the buyer conditional on product B being out, coupled with a suitable fixed fee (the highest such fee leaves the buyer just indifferent to dropping product A and purchasing the optimal quantity from firm B at its cost). However, for reasons discussed next, foreclosure will not be profitable in this single buyer setting.

firms supply, firm *A*'s profit is reduced if it induces exclusion of firm *B*. The intuition is that when nonlinear contracts are available, firm *A* can extract profit not solely through a per unit price on its own sales but also through a fixed fee whose level can capture some of the surplus that the buyer earns from the other product. Thus, if it is better for the three parties collectively to have both products supplied, then firm *A* can realize higher profit by accepting such an outcome and setting a higher fixed fee.⁵³

5.2.2. *Nonlinear pricing to multiple buyers: Profitable but costly exclusion and predation at the margin*

Intuition suggests that (inefficient) exclusion can become profitable once we move from a single buyer environment to multiple (noncompeting) buyers, due to the divide-and-conquer logic. Bernheim and Whinston confirm this intuition. We review their analysis in more detail in the appendix and summarize it here.

Bernheim and Whinston develop a two-buyer example where both firms simultaneously offer nonlinear contracts to each buyer in sequence. That is, when contracting with the first buyer, it is not possible to also specify the terms to the second, for example, because that second market has not yet developed (what they call noncoincident markets). Suppose that (1) firm *B* must incur a fixed cost if it wishes to produce and makes this decision following the outcome of negotiations with the first buyer, (2) this fixed cost is high enough that firm *B* can cover it only if it sells to both buyers (i.e., there are important scale economies), and (3) firm *A* has already incurred this fixed cost. Bernheim and Whinston show that even when it is socially efficient to have both firms sell to both buyers, foreclosure of firm *B* can be profitable to firm *A* and can arise in equilibrium.

The key is that foreclosure allows firm *A* to extract as monopoly profit the second buyer's entire value for firm *A*'s product, whereas competition between *A* and *B* would leave surplus to the second buyer. If the loss from reduced variety to both buyers is small relative to the gain from extracting monopoly rents for product *A* from the second buyer, then foreclosure can arise. The first buyer is aware that its high purchases from firm *A* will induce exclusion of firm *B* and deprive the first buyer of valuable variety, but firm *A* is able to "bribe" the buyer to partake in foreclosing *B* by drawing on the anticipated monopoly profits from the second buyer. The analysis yields several interesting conclusions.

Foreclosure does not harm the pivotal buyer. In the models analyzed by Bernheim and Whinston and O'Brien and Shaffer, the first buyer always gains from exclusive dealing because this forces firms to engage in tougher all-or-nothing competition for its business.⁵⁴ The harmed parties are the excluded firm and the second buyer, who cannot participate in the initial contracting with the first buyer. This echoes a common finding in the economics literature: profitable contracting among a subset of agents can inflict

53. If, instead, firms are limited to linear prices, then the firm with the product differentiation advantage may find exclusive dealing profitable because it derives profit only on a per unit sales basis and exclusive dealing increases those sales. See Frank Mathewson & Ralph Winter, *The Competitive Effects of Vertical Agreement*, 77 AM. ECON. REV. 1057 (1987).

54. Because of the all-or-nothing form of competition, firm *A* can profit from exclusion only if it has *some* advantage over *B*, in this case, by having already sunk its fixed cost.

harm on parties absent from the contracting table. The antitrust lesson is that inefficient exclusion may not be detected by asking if the participating buyers have been “coerced”—the harm is likely to fall on others.

Profit sacrifice on the first buyer. Compared to the profits firm *A* would earn if it accepted firm *B* (by refraining from quantity forcing and instead employing a two-part tariff), under foreclosure firm *A* earns lower profit on the first buyer, to whom it must offer a lower fee for distorting the quantity choices by reducing variety. Since foreclosure is profitable on balance, an antitrust test that asked if firm *A* was sacrificing profit overall by excluding would obviously acquit. However, if the test was applied to the first buyer alone, it would reveal that exclusion entailed a profit sacrifice.

Predation at the margin. If firm *B*'s fixed cost is sufficiently high, firm *A* may be able to exclude *B* by “forcing” the first buyer to purchase from it their bilaterally efficient quantity conditional on firm *B* being out, q_A^s . If firm *B*'s fixed cost is lower, however, foreclosure may require inducing the first buyer to expand its purchases of *A* even beyond q_A^s , so as to further depress that buyer's residual demand for the substitute product *B*. In such a case, there is predation at the margin: the quantity exceeds the level that would result if firm *A* set price equal to marginal cost and let the buyer select the quantity. The implicit price for the marginal units is thus below marginal cost. Observe that a test that compared average revenue with average cost would acquit; a test that compared *incremental* revenue with incremental cost would find predation, provided it adopted the correct quantity increment. In practice, of course, this is a daunting task. We return to these issues in Section 7.

6. Nonexclusionary motives for quantity forcing

6.1. Heterogeneous consumers and indirect price discrimination

With identical customers, simple two-part tariffs suffice for rent extraction, so quantity-forcing contracts discussed above may raise eyebrows. However, two-part tariffs are no longer sufficient for maximal rent extraction when a monopolist faces heterogeneous buyers and cannot directly price discriminate, e.g., because he cannot identify the particular demand curve of any given buyer (the buyer's “type”), though he knows the distribution of demands.⁵⁵ In such cases, the monopolist's profit-maximizing indirect (second degree) price discrimination entails offering the same menu of nonlinear contracts to all potential buyers, with the menu designed to induce self selection—different buyer types (indexed by $i = 1, \dots, n$) select different contracts. Kolay, Shaffer, and Ordovery remind us that optimal indirect discrimination cannot be implemented with menus of two-part tariffs and show it can be implemented through contracts such as menus of all units discounts.⁵⁶ Note that all unit discount contracts all have the “suspicious” property that total payments fall over some range as purchases rise

55. Direct discrimination may also be infeasible for legal reasons, e.g., in some contexts, any offer made to one customer must be made available to all.

56. Sreya Kolay, Greg Shaffer & Janusz Ordovery, *All-Units Discounts in Retail Contracts*, 13 J. ECON. & MGMT. STRATEGY 429 (2004).

past the threshold. The message of Kolay, Shaffer, and Ordovery is that such contracts need not be motivated by exclusion.

However, if the sole purpose were monopoly price discrimination, this could also be achieved by offering a menu of contracts $\{q_i, T_i\}$ where q_i is the quantity intended for (and, in equilibrium, chosen by) a particular type i , and T_i is the total payment for that total quantity.⁵⁷ This optimal menu has the natural property that total payment demanded always rises with quantity (T_i increases with q_i).⁵⁸ These more natural contracts would seem simpler to implement.

6.1.1. Market share discounts

With competing firms instead of a monopolist, specifying the minimal purchases needed to qualify for a discount not as absolute quantities but as *percentages* of the buyer's total purchases—share contracts—raises a perception that the purpose may be to deny business to competitors rather than increase it for the firm. An alternative explanation put forth for share discounts is that buyers differ *exogenously* in their sizes—e.g., small versus large retail stores—and hence in their ability to attain specified volume levels; basing the discounts on the *share* of each buyer's total purchases made from that manufacturer may provide incentives for all buyers to increase their purchases.⁵⁹

One rationale for the above behavior may be a blend of third degree price discrimination (there are exogenous differences between certain observable buyer classes such as large versus small stores, to motivate reliance on share of purchases rather than absolute volumes) and second degree price discrimination (within a given group there is buyer heterogeneity that is unobserved by the seller, such as high or low preference for its product). At an informal level, optimal screening within the class of small buyers and optimal screening within the class of large buyers takes the same relative form: some types choose, in equilibrium, smaller quantities than other types within the same class. A menu of contracts that is intended to serve as a screening contract for all sizes of buyers might then conceivably take the form of a market share contract since such contracts would have this same qualitative feature.

However, two issues arise. First, it appears that only highly particular environments would yield that the optimal (size specific) menu of contracts implies precisely the same market share thresholds for both smaller and larger buyers. Conditions under which share discounts indeed implement the profit-maximizing form of price discrimination given the information constraints faced by the seller remain to be explored. Second, and more importantly, a contract that makes outcomes contingent on market share gives the privately informed buyer an extra way to impede price discrimination by distorting the information available to the seller. A buyer could, in some circumstances, meet a market share threshold partly by reducing its purchases from other sellers, as an alternative to expanding its purchases from the first seller.

57. For example, wireless phone companies offer buckets with differing number of minutes for different fixed monthly fees.

58. See TIROLE, *supra* note 17.

59. This argument is noted in the *Concord Boat* case and attributed to Brunswick's economic expert, Dr. Frederick Warren-Boulton. *Concord Boat Corp. v. Brunswick Corp.*, 207 F.3d 1039 (8th Cir. 2000).

Both issues highlight that if the goal is solely price discrimination, market share discounts are *prima facie* suboptimal mechanisms. One could imagine reasons why such contracts nevertheless would be chosen for price discrimination—perhaps it is simply too costly to tailor each contract for each individual buyer, or perhaps because of equity or legal concerns⁶⁰ a seller is unwilling to present significantly different menus to different customers. But a defendant making such a claim should be prepared to offer some evidence.

6.2. *Market share discounts to induce dealer effort*

Mills offers an efficiency explanation for share discounts, suggesting they may be a means to induce retailers to engage in otherwise unenforceable merchandising activity.⁶¹ He develops a model where retailers who are heterogeneous in size and skill can take an action that increases demand for the product of a monopolist supplier. The retailer's consumers purchase either the monopolist's product or an inferior (in the vertical differentiation sense) rival product. If the monopolist were able to monitor the retailer's activities then he would simply pay the retailer explicitly for such activities whenever it is profitable (and, therefore, in this model, socially desirable). In the absence of a monitoring capability, the monopolist could conceivably reward the retailer for achieving a target sales level, but if the retailers differ exogenously in size and these differences are unobservable by the supplier, then the appropriate target level for a given retailer may not be known and therefore may not be contractible. Mills provides an example where merchandising activities affect sales of the supplier's product proportionately to the retailer's size. In such circumstances, a market share target level may be sufficient to induce the retailer to take the appropriate actions. This efficiency-based motive for market share discounts relies heavily on the requirement that, at the contracting stage, the supplier be unable to recognize whether the retailer is large or small relative to the market and that the retailer's optimal activities be directly proportionate to its size.

7. Conclusions

We began with the case where quantity forcing in the target market is implemented via a tied discount off a second, monopoly product. The economic literature shows that, compared to stand-alone per unit pricing of the monopoly good, tying can extract additional rent. This role for tying is present under the assumed pricing conditions whether or not foreclosure is possible in the tied market. Moreover, tying in this setting can improve economic efficiency by inducing an output expansion for the monopoly good. Building on this analysis, we identified why an *Ortho*-type test could condemn welfare increasing (and consumer neutral) tying as predatory.

60. For example, the Robinson-Patman Act may prohibit a seller from offering different buyers different menus of contracts.

61. David Mills, *Market Share Discounts* (University of Virginia, Working Paper, 2004).

However, there is a serious difficulty with the above rent extraction argument for tying. The underlying premise that the monopolist can charge only linear pricing is at odds with the ability to impose tying. A tie-out (or a tie that requires significant purchases of the tied product) requires greater contracting power than a requirement also to purchase some small amount of the tied product from that firm. A minimal tie-in, coupled with a suitably high price for the tied good, would suffice to levy a fixed fee for the right to buy the monopoly good. Furthermore, if buyers' demands are known, the ability to charge a fixed fee—or differing fixed fees if buyers are heterogeneous—avoids the inefficiency of a tie-out, and therefore can extract more rents. If the monopolist is imperfectly informed about demands, tying can help with rent extraction in certain settings: the traditional one where a tied complementary product serves to meter the intensity of demand for the tying good, and perhaps also where the products are unrelated in consumption but are statistically dependent (so a buyer's demand for one good provides a signal about his level of demand for the other). These explanations, however, do not apply when demands for the goods are unrelated and independent nor when the goods are consumed in fixed proportions. In both cases, the demand for one good then conveys no added information about the demand for another.

Thus, in many situations tying is likely to conflict with optimal rent extraction because the pricing contracts that are plausibly available whenever tying is feasible are likely to induce smaller consumption distortions. By contrast, inducing a customer to buy all or a high share of its tied good requirements from the tying good monopolist is consistent with a desire to exclude rivals from the tied market when economies of scale are significant. But why would the excluder prefer to implement quantity forcing by providing the inducement via tied discounts rather than directly via nonlinear pricing of the tied good? If the monopolist has used its contracting instruments optimally to extract rents from the other markets, there should not remain "cheap money" in the form of unexploited rents available to "bribe" the buyer into accepting quantity forcing in the target market.

A possible reason for preferring tied discounts is to make the inducements less transparent to antitrust enforcers, especially if the discounts are provided on numerous tied products. However, the antitrust concern should center on the quantity forcing aspect and whether it poses a serious risk of exclusion, not on *how* quantity forcing is implemented, whether through tying or nonlinear pricing in the target market alone.

Turning to nonlinear pricing in the target market, such contracts make exclusion less costly to the perpetrator, in the presence of scale economies, than having to rely on uniform price cuts. This fact alone, however, clearly does not justify condemnation or even greater scrutiny of nonlinear pricing than of per unit pricing. Nonlinear pricing makes it cheaper for the firm to expand output whatever the purpose, exclusion *or* legitimate. Unfortunately, our analysis suggests there is no simple test that will identify the main purpose and likely effect of all nonlinear pricing schemes.

Indeed, formulating an economically justified safe harbor is likely to be even harder than for simple per unit pricing. There, setting price above marginal cost might be taken as a proxy for lack of profit sacrifice and an absence of predatory intent. With nonlinear pricing, implementing such an incremental revenue/incremental cost test poses added difficulties:

- The incremental price or revenue will often not be directly observable because it must be computed net of the discounts, which themselves may be somewhat opaque to outsiders.
- The calculations will depend on the size of the output increment considered; as noted earlier, all units discounts inherently imply a negative marginal revenue for small output increments that take the buyer just past a discount threshold.
- A predatory profit sacrifice from expanding output beyond the marginal-cost-pricing level will be easier to detect when there is an explicit per unit price than when the seller offers, for example, an all-or-nothing quantity contract.
- Finally, with nonlinear pricing, there can be a profit sacrifice even if the output is at a marginal-cost pricing level. For instance, in Bernheim and Whinston's analysis, inefficient foreclosure through quantity forcing can entail choosing the output level that is jointly efficient for the firm and buyers given that the rival is excluded, so at the margin, price is set equal to marginal cost. Yet there is profit sacrifice, taking the form of collecting a smaller fixed fee relative to the unobservable non-exclusion benchmark.

In the current state of knowledge, we are not hopeful about the prospects of crafting an economically grounded safe harbor for nonlinear pricing, at least not one that would cover all such contracts. The prudent course at this stage, we believe, is to conduct the analysis in two broad steps: (1) use structural screens to gauge if the risk of exclusion is significant, and (2) if and only if the answer is yes, proceed to a more detailed inquiry to determine whether the only or the predominant motive for the practice was exclusion.⁶² A compelling reason for focusing on intent, once the above structural conditions have been met, is that it provides a strong signal about likely effects.⁶³ As a market insider, the firm is likely to possess superior knowledge than an outsider about the likely effects of its action, so an ex post inquiry into the firm's rationale for adopting a practice can shed light also on the likely prospects for success, i.e., does the practice pose a serious risk to competition?

Structural screens. Given the strong efficiency potential of nonlinear pricing, we would set the bar quite high before entertaining allegations of exclusion. To be useful, screens must be based on information that is fairly readily available. Thus, we would confine the inquiry to questions such as:

- *Concentration:* Does the alleged excluder possess a high market share? Is the plaintiff one of only very few strong competitors, such that excluding it would have a significant adverse effect on competition?
- *Scale economies and volume denied:* Are there important scale economies so that a competitor incurs a significant cost penalty per unit if forced to operate at a low scale? If so, has the incumbent, through its forcing contracts, locked up

62. Spector also offers a structured rule of reason but differs somewhat from ours in the order of, and relative emphasis on, various steps. David Spector, *Loyalty Rebates: An Assessment of Competition Concern and a Proposed Structured Rule of Reason*, 1 COMPETITION POL'Y INT'L 89 (2005).

63. Thus intent conveys relevant information even when the sole issue is whether to enjoin the practice going forward. Intent is obviously necessary when the issue is whether to penalize the firm for its past willful actions.

enough of the demand that the competitor's ability to realize scale economies is seriously jeopardized?⁶⁴ These conditions are likely to be less observable than concentration, so we would not demand detailed information or set the bar too high for the plaintiff at this screening stage. However, we would require some showing that these conditions are met. For example, the scale economy hurdle is failed if one observes other firms operating successfully at fairly small scale.

Conduct inquiry. If the structural screens are passed, we turn to a more detailed and nuanced analysis of the particular conduct and the specific market context. Price-cost tests can be informative but are subject to added difficulties in the case of nonlinear pricing. Recall, for example, that exclusion through quantity forcing can arise even when the firm's marginal price equals its marginal cost. The form of the particular contracts and the circumstances of their adoption are likely to be just as revealing.

- *Quantity forcing versus more flexible contracts:* Contracts that induce a customer to purchase a very high share of its total requirements from the defendant have greater exclusionary potential than, for example, two-part tariffs. In Bernheim and Whinston's analysis, two-part tariffs offered by both firms will suffice to implement the efficient solution—whether it entails both being active or only one of them—but inefficient exclusion requires use of quantity forcing.⁶⁵
- *Divide-and-conquer:* Profitable exclusion through denying scale economies to a rival is facilitated by offering quantity-forcing contracts to only a subset of customers. Thus, concern is heightened if the contracts are not offered uniformly to all comparably situated customers but only to some (as occurred in *LePage's*). Of course, the term “comparably situated” is key, because with customer heterogeneity, there are nonexclusionary reasons for differential offers. Transaction costs could explain, for example, why a manufacturer would only engage in complex contracts with larger customers and sell on simple terms to small ones. We would allow defenses based on transaction costs or on price discrimination that reflects demonstrable differences across customers in their willingness to pay. However, suspicion is raised if discrimination seems driven by an intent to divide and conquer.
- *Calibrating the quantity forcing:* Suspicion is further heightened if there is evidence that the defendant chose its quantity-forcing targets with an eye to denying the rival scale economies rather than for its own internal reasons. The timing of when such contracts were adopted can also be informative: did the disputed contracts appear to have been adopted in response to an increased in the perceived entry threat?
- *Efficiency motives:* Our starting position is that firms should not be required to explain the business reasons for their actions; those should be presumed

64. In *Concord Boat*, the court of appeals reversed the original judgment for the plaintiff, finding that the plaintiffs had “failed to . . . demonstrate that Brunswick had foreclosed a substantial share” of the market and “did not show that significant barriers to entry existed in the stern drive engine market.” *Concord Boat*, 207 F.3d at 1059.

65. Bernheim & Whinston, *supra* note 11.

legitimate barring evidence to the contrary. However, as evidence of exclusionary motive and likely effect accumulates, the burden increases on the defendant to show an alternative reason. We recognize that there could be nonexclusionary purposes for market-share contracts and the like. But at this stage of the inquiry, we would require more than superficial assertions about such purposes.

In sum, quantity-forcing requirements accounting for a large share of purchases, in a concentrated market, and where operating on a small scale seriously threatens the rival as a robust competitor will raise a yellow flag. The flag *may* turn red if the quantity targets and the number of affected customers are calibrated with an eye on denying the rival scale economies and there is other evidence of a concerted plan to exclude. We recognize that this is not a simple inquiry, and we would therefore impose a fairly stringent threshold at the screening stage. However, such an inquiry seems unavoidable in monopolization cases, whether the affected conduct includes price or nonprice terms.

Appendix:

Competition with Nonlinear Pricing

This appendix draws on the analysis of Bernheim and Whinston (BW) and O’Brien and Shaffer (OS).

Single Buyer

Assume initially that firms *A* and *B* sell differentiated products to a single monopolist retailer, buyer 1. Let $R(q_A, q_B)$ denote the retailer’s revenue net of any variable retailing costs as a function of the pair of outputs sold, and $C_J(q_J)$ the cost function of firm $J (J = A, B)$, with $C_J(0) = 0$. Let (q_A^c, q_B^c) denote the quantities that maximize the common surplus of both firms and the buyer if efficiency requires selling *both* products, and Π^c the associated total surplus, $\Pi^c = R(q_A^c, q_B^c) - C_A(q_A^c) - C_B(q_B^c)$. Let q_J^s denote the quantity of product J that maximizes the joint surplus of the buyer and firm J when only product J is sold, and Π^J the associated total surplus, $\Pi^J = R(q_J^s) - C_J(q_J^s)$. Assume

$$\partial^2 R(\cdot) / \partial q_A \partial q_B < 0 \quad (\text{implying } q_J^c < q_J^s) \tag{A1}$$

the marginal profitability of each product is reduced as the quantity of the other is increased, a condition typical of substitutes. This condition implies

$$\Pi^c < \Pi^A + \Pi^B \tag{A2}$$

A’s incremental contribution to surplus is smaller when *B* is also sold than when *A* alone is sold, $\Pi^c - \Pi^B < \Pi^A$, and similarly for *B*. Finally, suppose each product is viable, one product is “stronger” (without loss of generality, let it be *A*), and efficiency requires carrying both products:

$$\Pi^c > \Pi^A > \Pi^B > 0 \tag{A3}$$

This is the interesting case for studying foreclosure incentives: one firm has an advantage if competing for exclusivity ($\Pi^A > \Pi^B$), but foreclosure is inefficient overall ($\Pi^c > \Pi^A$).

The contracting takes place as follows. (1) Both firms simultaneously present *contract offers*, a pair $(T_J^e(q_J), T_J^c(q_J))$, where T_J is the total payment demanded by firm J as a function of J 's output, for exclusive (e) or common (c) representation (exclusivity is enforced, for example, by specifying $T_J^c = \infty$). (2) Next, the retailer chooses whether to represent one firm or both (if the latter is an option). (3) The retailer chooses its output(s), realizes sales, and makes payments to the chosen firm(s). Given the simultaneous choices in stage (1), exclusivity can occur because of coordination failures—if one firm demands exclusivity, the other is de facto forced to do so as well; to rule this out, BW focus on equilibria that are undominated for firms A and B . They show:

Proposition 1 (BW): In any undominated equilibrium (for the manufacturers):

- (1) the retailer chooses the quantities that maximize overall surplus;
- (2) the payoffs to firms A and B are: $v_A^c = \Pi^c - \Pi^B$, $v_B^c = \Pi^c - \Pi^A$ (i.e., $v_J^c = \Pi^c - \Pi^{-J}$, each firm earns its incremental contribution to overall surplus);
- (3) the retailer's payoff is $v_1^c \equiv \Pi^c - v_A^c - v_B^c = \Pi^A + \Pi^B - \Pi^c (= \Pi^J - (\Pi^c - \Pi^{-J}))$, i.e., the retailer earns the difference between the stand-alone surplus generated by either firm J and that firm's incremental contribution to total surplus); and
- (4) there is always a common equilibrium yielding this undominated outcome.

Part (1) is intuitive: if total surplus were not maximized, there would be an alternative outcome yielding at least the same payoff to the retailer and more to the firms, thus violating the restriction to undominated equilibria.

Part (2) can be understood as follows. The argument is identical for both firms, so consider one case, why $v_A^c = \Pi^c - \Pi^B$ or, equivalently, $\Pi^c - v_A^c = \Pi^B$. The left side is the joint surplus of the retailer and firm B in the common dealing equilibrium; the right side is their joint surplus under exclusivity with B . If $\Pi^c - v_A^c < \Pi^B$ (firm A demands more than its incremental contribution to total surplus), then the joint surplus of firm B and the retailer is lower under common dealing than under exclusivity with B . Firm B can then offer a contract that both induces the retailer to switch to exclusivity and increases B 's profit. If $\Pi^c - v_A^c > \Pi^B$, firm A can profitably raise the payoff it demands for common dealing above v_A^c until the joint surplus of the retailer and firm B equals Π^B .

[Details: Let (P_J^c, P_J^e) denote the net payoffs demanded by firm J under common dealing and exclusive dealing, respectively. For example, under sellout contracts these demanded payoffs are simply the fixed fees. In a common dealing equilibrium each firm's actual payoff v_J^c equals P_J^c , and P_J^e must prevent any profitable deviation. Suppose $v_A^c > (\Pi^c - \Pi^B)$. The buyer's surplus under common dealing is $v_1^c = \Pi^c - v_A^c - v_B^c$. For this to be an equilibrium, the buyer must weakly prefer it over exclusivity with B : $v_1^c \geq \Pi^B - P_B^e$, implying $P_B^e \geq \Pi^B - v_1^c$. Let firm B instead offer $\hat{P}_B^e = \Pi^B - v_1^c - \varepsilon$ where $0 < \varepsilon < v_A^c - (\Pi^c - \Pi^B)$. (1) The buyer will deviate to exclusivity with B , since its payoff rises to $\Pi^B - \hat{P}_B^e = v_1^c + \varepsilon$, and (2) firm B 's profit also

risers, by $(\hat{P}_B^e - v_B^c) = (\Pi^B - v_1^c - \varepsilon) - (\Pi^c - v_1^c - v_A^c) = [v_A^c - (\Pi^c - \Pi^B) - \varepsilon] > 0$, breaking the hypothesized equilibrium. Next, consider $v_A^c < (\Pi^c - \Pi^B)$ so $\Pi^c - v_A^c > \Pi^B$. This too cannot be an equilibrium since firm A can profitably raise its demanded payoff from v_A^c to $\hat{P}_A^c = \Pi^c - \Pi^B$, without pushing the buyer and firm B away from common dealing.]

Finally, the above outcome can be supported by, for example, sellout contracts: $T_J^c(q_J) = F_J^c + C_J(q_J)$, $T_J^e(q_J) = F_J^e + C_J(q_J)$, $J=A, B$, with $F_J^e = F_J^c = v_J^c = \Pi^c - \Pi^{-J}$. The latter says that in equilibrium the payoff demanded by either firm for dealing exclusively is equal to the actual payoff earned by that firm under common dealing.

This property follows because the buyer must also be indifferent between these regimes, as explained next.

Buyer indifference: $v_1^c \equiv \Pi^c - F_A^c - F_B^c = \Pi^A - F_A^e = \Pi^B - F_B^e$. (BW expression (3), OS Lemma 1.) If $\Pi^c - F_A^c - F_B^c > \Pi^J - F_J^e$, the other firm I could profitably demand a higher payoff for common dealing (F_I^c) without pushing the buyer to exclusivity with J . If $\Pi^c - F_A^c - F_B^c < \Pi^J - F_J^e$, the buyer would prefer exclusivity with J over common dealing. Intuitively, equilibrium contracts require both that (1) the buyer and each firm J are jointly indifferent between exclusivity with J and common dealing, and (2) the buyer individually is indifferent. Thus, firm J 's demanded payoffs must also make J indifferent.

Observe that firm A could foreclose B while earning positive profit. Under exclusivity, firm A earns $\Pi^A - \Pi^B$ (to obtain exclusivity, firm A must match the maximal surplus the retailer would get from B). If exclusivity contracts are not feasible, firm A still may be able to foreclose through quantity forcing; foreclosure is feasible if the output that maximizes surplus when only A is sold, q_A^s , leaves firm B with nonpositive profit. (Recall from Equation (A1) that $q_A^s > q_A^c$. OS provide an example where q_A^s indeed renders firm B not viable; see OS Proposition 4. Any output above q_A^s cannot be an equilibrium, as it fails to maximize the joint surplus of the retailer and firm A .) In that case, firm A can foreclose B by offering the all-or-nothing contract $T_A^c(q_A^s) = C_A(q_A^s) + (\Pi^A - \Pi^B)$ (and $T_A^c = \infty$ for any other output). However, with a single buyer, foreclosure entails a profit sacrifice for firm A relative to common dealing since $\Pi^A - \Pi^B < \Pi^c - \Pi^B$ by Equation (A3).

Intuitively, under common dealing firm B earns only its incremental contribution to total surplus, $\Pi^c - \Pi^A$, leaving firm A and the retailer with $\Pi^c - (\Pi^c - \Pi^A) = \Pi^A$, the same surplus they could generate under exclusivity with A . Of this, firm A captures more under common dealing than under exclusivity, $(\Pi^c - \Pi^B) - (\Pi^A - \Pi^B) = \Pi^c - \Pi^A > 0$, and the retailer captures correspondingly less, $(\Pi^A + \Pi^B - \Pi^c) - \Pi^B = -(\Pi^c - \Pi^A)$. (For this reason, the retailer would prefer firms A and B to compete for exclusivity.)

Two buyers: Noncoincident markets

Now modify the above scenario as follows. Firms A and B can contract (simultaneously as before) also with a second monopolist retailer in a different market, buyer 2, but only after contracting with the buyer 1. (For example, the second market will develop only later, but its emergence is foreseen at the outset.) Unlike previously,

assume firm B must incur a fixed cost K to be active in either market, in addition to its variable cost $C_B(q_B)$. For simplicity, assume constant variable costs (so the operating profits across markets are independent), let the two markets be otherwise identical, and ignore discounting. Firm B can decide whether to incur K after contracting with the first buyer. Assume also that

$$(\Pi^c - \Pi^A) < K < 2(\Pi^c - \Pi^A) \quad (\text{A4})$$

Thus, firm B can make enough profit under common dealing to cover the entry cost if serving both markets (hence its entry is efficient overall) but not if serving only one.

Contractual exclusive dealing

We first analyze explicit contractual exclusivity, then exclusion through quantity forcing. If there is common dealing with buyer 1, this will also be the equilibrium outcome with buyer 2. By Proposition 1, gross profits from each market will be $v_A^c = \Pi^c - \Pi^B$ and $v_B^c = \Pi^c - \Pi^A$ (in addition, firm B incurs K). If firm A obtains exclusivity with buyer 1, firm B is foreclosed also from buyer 2. Thus, buyer 1 is pivotal. To obtain exclusivity, firm A must match the maximum surplus that firm B can bid to buyer 1 for exclusivity: $b = \Pi^B + (\Pi^c - \Pi^A) - K$, the surplus in market 1 under exclusivity plus B 's net profit in market 2 under common dealing. (In market 2, the outcome would be common dealing as described in Proposition 1 since firm A cannot be profitably foreclosed.) Thus, under exclusivity firm A 's net profit is $\Pi^A - b$ from buyer 1 plus its unconstrained monopoly profit Π^A from buyer 2 (having foreclosed firm B). Firm A 's net overall surplus under exclusivity with 1 (S_A^e) and under common dealing (S_A^c) are summarized below:

$$S_A^e = \begin{array}{l} \text{from buyer 2} \\ \Pi^A \end{array} + \begin{array}{l} \text{from buyer 1} \\ \Pi^A - b \end{array} \quad (b = \Pi^B + (\Pi^c - \Pi^A) - K) \quad (\text{A5})$$

$$S_A^c = (\Pi^c - \Pi^B) + (\Pi^c - \Pi^B) \quad (\text{A6})$$

Firm A chooses foreclosure through contractual exclusive dealing if $S_A^e > S_A^c$ or

$$\Pi^A - (\Pi^c - \Pi^B) > 2(\Pi^c - \Pi^A) - K \quad (\text{A7})$$

(This is a version of BW's condition (C3), assuming symmetric markets.)

The left term in Equation (A7) is positive by Equation (A2) and represents the increased rent firm A extracts from buyer 2 by excluding firm B . It equals (1) buyer 2's surplus under common dealing (Proposition 1.3) and, hence, (2) buyer 2's loss from exclusion since firm A then extracts 2's entire surplus. The right term is also positive, by Equation (A4), and represents the loss in overall welfare across both markets from foreclosure. (We explain shortly why this term affects firm A 's profitability from exclusion.) It equals (3) firm B 's net profit under common dealing in both markets, and, hence, (4) firm B 's loss from exclusion. The distributional effects are summarized as follows.

Gainers and losers. Denote the change in surplus to party J ($J = A, B, 1, 2$) moving from common to exclusive dealing by $\Delta S_J \equiv S_J^e - S_J^c$. Then

- firm B loses its entire common dealing surplus: $\Delta S_B = -[2(\Pi^c - \Pi^A) - K] < 0$;
- buyer 2 loses its entire common dealing surplus: $\Delta S_2 = -[\Pi^A - (\Pi^c - \Pi^B)] < 0$;
- firm A gains on balance given Equation (A5): $\Delta S_A = -\Delta S_2 + \Delta S_B > 0$; and
- buyer 1 gains an amount equal to firm B 's loss which, in turn, equals the loss in total surplus: $\Delta TS \equiv [\Delta S_A + \Delta S_B + \Delta S_2 + \Delta S_1] = \Delta S_B < 0$ using $\Delta S_1 = -\Delta S_B$.

Profit sacrifice on pivotal buyer. The right term in Equation (A7) is both firm A 's profit sacrifice on buyer 1 to obtain exclusivity, $2(\Pi^c - \Pi^A) - K = (\Pi^c - \Pi^B) - (\Pi^A - b)$, and the loss of total surplus across both markets. The entire welfare loss falls on firm A as a profit sacrifice for exclusivity with buyer 1 because firm A must match the surplus that firm B bids for exclusivity, $b = \Pi^B + (\Pi^c - \Pi^A) - K$, the entire gross surplus in market 1 under exclusivity with B plus firm B 's net profit in market 2 under common dealing. The difference in firm A 's profit from buyer 1 under common dealing versus exclusivity, $(\Pi^c - \Pi^B) - (\Pi^A - b)$, can be decomposed as $[(\Pi^c - \Pi^B) - (\Pi^A - \Pi^B)] - (\Pi^c - \Pi^A - K)$.

Under exclusivity, firm A incurs two sources of sacrifice. First, firm A collects from buyer 1 the difference in market 1 surplus with product A alone versus with B alone, $(\Pi^A - \Pi^B)$, instead of the larger difference in surplus under common dealing versus with B alone, $(\Pi^c - \Pi^B)$. This inefficiency from lost variety in market 1 reduces the payment firm A can earn by the amount (1) $(\Pi^c - \Pi^A)$. Second, firm A must pay buyer 1 the net profit that firm B would have earned in the second market absent foreclosure, term (2) $(\Pi^c - \Pi^A - K)$. Terms (1) and (2) equal the total welfare loss.

Finally, this sum equals buyer 1's gain ($\Delta S_1 = -\Delta TS$). To see why, write buyer 1's surplus under common dealing as $\Pi^B - (\Pi^c - \Pi^A)$, B 's incremental contribution to market surplus when only B is offered versus when also A is offered. Under exclusivity, buyer 1 gets $b = \Pi^B + [(\Pi^c - \Pi^A) - K]$, which can be interpreted as B 's contribution to market 1 surplus when only B is offered, plus firm B 's net profit from market 2. Thus, exclusivity raises buyer 1's surplus by $2(\Pi^c - \Pi^A) - K = -\Delta TS$. Exclusivity intensifies the competition, forcing firm B to bid its entire two-market profit in an attempt to prevent foreclosure, and firm A to match this payoff to the pivotal buyer. By contrast, under common dealing competition is not all-or-none so the buyer attains less surplus.

Exploiting the future buyer. Adding $\Pi^A + (\Pi^c - \Pi^B)$ to both sides in Equation (A7) yields an alternative version of the condition for profitable foreclosure:

$$2\Pi^A > \Pi^A + (\Pi^c - \Pi^B) + 2(\Pi^c - \Pi^A) - K \quad \text{or}$$

$$2\Pi^A > \Pi^c + [(\Pi^c - \Pi^B) + (\Pi^c - \Pi^A) - K] \quad (\text{A7}')$$

The left term is the joint surplus of firms A, B and buyer 1 if firm A obtains exclusivity with buyer 1. The right term is their joint surplus under common dealing: Π^c in market 1 plus the profits of firm A and firm B from market 2 (attributing for accounting purposes the cost K to that market). Although firm B loses from foreclosure, foreclosure is only profitable if the gains to firm A and the buyer exceed this loss because firm B offers the pivotal buyer its entire prospective surplus to prevent being

foreclosed. The only party not adequately represented is buyer 2, because it is absent from the initial contracting.

Exclusivity through quantity forcing

If contractual exclusivity is prohibited, can firm A profitably foreclose B using quantity forcing? For any quantity q_A accepted by buyer 1, define $q_B^*(q_A)$ as B 's best response output—it maximizes market 1 surplus given q_A . Denote the resulting surplus $\hat{\Pi}^c(q_A) \equiv R(q_A, q_B^*(q_A)) - c_A q_A - c_B q_B^*(q_A)$, where c_A, c_B are the constant marginal costs. If B stays out given q_A , the market 1 surplus is denoted $\hat{\Pi}^A(q_A) \equiv R(q_A, 0) - c_A q_A$. A deterring output q_A^d is one that leaves firm B with nonpositive net profit overall:

$$[\hat{\Pi}^c(q_A^d) - \hat{\Pi}^A(q_A^d)] + (\Pi^c - \Pi^A - K) \leq 0 \quad (\text{A8})$$

The term in square brackets is B 's profit in market 1, B 's incremental contribution to total surplus conditional on q_A^d ; the other term is B 's net profit in market 2 under the efficient common dealing outputs (since outputs there are determined by spot competition). By Equation (A4), firm A 's optimal common dealing output q_A^c would not deter B ; a higher output is required. Consider two of the cases analyzed by BW (Proposition 5).

Case 2—Foreclosure at firm A 's optimal stand-alone output: q_A^s suffices to deter and foreclosure under contractual exclusivity would be profitable (Equation (A7') holds and Equation (A8) is satisfied at $q_A^d = q_A^s$). In this case, firm A can duplicate through quantity forcing the contractual exclusivity outcome. To foreclose, firm A must give buyer 1 at least the maximum surplus firm B offers for de facto exclusivity, $b = \Pi^B + (\Pi^c - \Pi^A) - K$ from Equation (A5). (Firm B can offer this surplus through a contract specifying the output q_B^s , which generates joint surplus Π^B in market 1 if only B is active, and a suitable fixed fee.) Suppose firm A offers the quantity-forcing contract $T_A(q_A^s) = \Pi^A - b$, $T_A(q) = \infty$ for any $q \neq q_A^s$. If buyer 1 accepts, firm B is foreclosed (since q_A^s satisfies Equation (A8)) and the buyer's payoff is $\Pi^A - T_A(q_A^s) = b$, so the buyer will accept. Firm A 's profit across both markets is $2\Pi^A - b$, which exceeds its common dealing profit $2(\Pi^c - \Pi^B)$ given Equation (A7').

Case 1a—Foreclosure with predation at the margin: q_A^s does not deter and Equation (A9) below holds (BW's C4). Let $\tilde{q}_A^d (> q_A^s)$ denote firm A 's optimal deterring output: it maximizes market 1 surplus with product A alone $[R(q_A, 0) - c_A q_A]$ subject to satisfying the deterrence constraint Equation (A8). To deter, firm A must still grant buyer 1 net surplus b . But its gross profit in market 1 is now $\tilde{\Pi}^A \equiv [R(\tilde{q}_A^d, 0) - c_A \tilde{q}_A^d] < \Pi^A$. Thus, instead of $2\Pi^A > b + 2(\Pi^c - \Pi^B)$, Equation (A7'), profitable deterrence now requires the stronger condition $\tilde{\Pi}^A + \Pi^A > b + 2(\Pi^c - \Pi^B)$ or substituting for $\tilde{\Pi}^A$ and b :

$$[R(\tilde{q}_A^d, 0) - c_A \tilde{q}_A^d] + \Pi^A > \Pi^c + [(\Pi^c - \Pi^B) + (\Pi^c - \Pi^A - K)] \quad (\text{A9})$$

If this condition is met, firm A will foreclose B by offering a quantity-forcing contract that induces the pivotal buyer to choose a larger quantity than the one maximizing their joint surplus (q_A^s). Thus, firm A is engaging in predation at the margin, in that the incremental value of the output increment $\tilde{q}_A^d - q_A^s$ is below its incremental cost.