Exclusive dealing with distortionary pricing*

Giacomo Calzolari*, Vincenzo Denicolò*
and Piercarlo Zanchettin+

* University of Bologna and CEPR; + University of Leicester

September 30, 2017

Abstract

It is now generally recognized that exclusive contracts may benefit a dominant firm indirectly in various ways, but can they be directly profitable? The answer would be negative if firms priced at marginal cost and extracted buyers' rent by means of fixed fees only. However, various factors may impede efficient rent extraction, leading firms to use distortionary prices. We show that in this case exclusive contracts are profitable and anticompetitive if the dominant firm enjoys a large competitive advantage over its rivals. If instead the competitive advantage is small, exclusive contracts tend to be procompetitive. These effects appear as soon as marginal prices are distorted upwards and irrespective of the specific reason why they are.

Keywords: Exclusive dealing; Price distortions; Rent extraction; Two-part tariffs; Antitrust; Dominant firm.

J.E.L. numbers: D42, D82, L42

1 Introduction

Exclusive dealing agreements prohibit a buyer who purchases a firm's product from buying the products of the firm's competitors. These agreements are generally regarded with suspicion by antitrust authorities. However, policy and theory are still both unsettled. In this paper, we propose a new, robust explanation for this practice. In most of the existing theories, exclusive dealing entails a sacrifice of profits that pays off in other stages of the game, and any

*We are grateful to seminar participants at Bern, Mannheim, Paris Dauphine, Shanghai School of Economics and Finance, Jiao Tong University, Luiss, Bologna, EUI, Catholic University (Milan), the 5th IO Eief-Unibo-Igier workshop, BECCLE Competition Policy Conference in Bergen, the CRESS conference in Rhodes and the European Commission for useful comments and suggestions. E-mail addresses: giacomo.calzolari@unibo.it, vincenzo.denicolo@unibo.it, pz11@le.ac.uk.

possible anticompetitive effects materialize only in the recoupment phase.[1] In contrast, we show that exclusive dealing may be *immediately* profitable and anticompetitive. The difference is critical for antitrust policy.

Our explanation relies on the presence of price distortions. As noted by Bernheim and Whinston (1998) in response to Mathewson and Winter (1987), such distortions cannot be lightly hypothesized in models of exclusive dealing. The reason for this is that exclusive dealing arrangements can be enforced only if trade is non-anonymous, but if this is so then firms could in principle extract their profits efficiently by means of non-distortionary, lump-sum fees.

However, we believe that in practice firms rarely rely exclusively on fixed fees to extract profits. More often, lump-sum payments are supplemented with marginal prices in excess of marginal costs. This pattern of pricing may indeed be optimal if fixed fees are an imperfect means of rent extraction. We analyze two possible reasons why this may be so, namely, moral hazard and adverse selection. In addition, we propose a reduced-form model that can capture other possible sources of price distortions. The reduced-form model is general and simple, and it may be of independent interest.

We show that as soon as marginal prices are optimally distorted upwards, even by a tiny margin, firms may have unilateral incentives to enter into exclusive dealing arrangements. For competitors of comparable strength, this leads to a prisoners' dilemma: prices are reduced, and all firms are eventually worse off. Things are different, however, when a dominant firm enjoys a strong enough competitive advantage over its rivals, in terms of higher quality or lower costs. In this case, the dominant firm will gain from exclusive dealing whereas its rivals lose. We prove these results both in fully fledged models that explicitly account for the imperfections mentioned above, and in the reduced-form model. We do this to clarify that the conclusion does not depend on the specific reason why rents cannot be extracted efficiently: what drives the result is the upwards distortion in marginal prices.

We further show that not only the profitability, but also the welfare effects of exclusive dealing depend on the size of the dominant firm's competitive advantage. That is, exclusive dealing is anticompetitive when the competitive advantage is large, procompetitive when it is small. Again, our analysis clarifies that these effects appear as soon as marginal prices are distorted upwards, and irrespective of the exact reason why they are.

The importance of these results for competition policy lies precisely in their generality. The effects that we uncover are robust and do not depend on the fine details of specific cases. Furthermore, our theory identifies one key factor that determines the competitive impact of exclusive dealing, namely, the size of the dominant firm's competitive advantage. This can be assessed quantitatively, as it correlates with variables, such as for instance the dominant firm's market

[1] In some special cases the short-run sacrifice may vanish, but still there is no immediate gain. The future gain may take a variety of forms, such as entry deterrence (e.g. Rasmusen et al., 1991), the exploitation of a future entrant (e.g. Aghion and Bolton, 1987), the protection of non-contractible investments (e.g. Marvel, 1982), and so on. See Whinston (2008) and Calcagno, Fumagalli and Motta (2016) for excellent surveys of the literature.

share, that are observable. This means that the theory is applicable broadly, and that its implications for policy are practical.

We shall return to these implications in the concluding section. At this point, it may be useful to discuss informally the economic mechanism underlying our main results. Abstracting from any intertemporal trade off, consider the competition among two or more firms that supply substitute products. In this setting, the upside of exclusive dealing arrangements is that they increase the demand for the firm's product. The downside is whatever price reduction may be necessary to compensate the buyer for the loss of the option of buying other products. This creates a price-volume trade off, but one of a special nature.

The optimal resolution to this trade off would be obvious if firms extracted buyers' surplus efficiently by means of fixed fees only. In this case, a firm engaging in exclusive dealing would not profit from the increase in volume, as its marginal price would optimally be set at cost. But the firm might have to reduce its fixed fee in order to compensate the buyers for the loss in variety. As a result, exclusive dealing would never be directly profitable: see O'Brien and Shaffer (1997) and Bernheim and Whinston (1998) for a rigorous proof.

With distortionary pricing, however, a firm that engages in exclusive dealing does benefit from the increase in volumes. Moreover, the price reduction that is needed to entice the buyer to sign up an exclusive contract may be small: in a duopoly, it is in fact nominal. This follows from the so-called principle of *individual excludability* that must hold whenever firms can rely on lump-sum payments. Obviously, each firm sets its fixed fee at the level where the buyer is just indifferent between accepting the contract or not. As a result, in the equilibrium that would prevail in the absence of exclusive contracts, the buyer would be indifferent between dealing with both firms or with only either one: that is, the retailer could exclude either supplier at no cost. But with exclusive contracts each firm can unilaterally force the buyer to deal with either supplier, but not both, and then a minimal discount suffices to break the remaining indifference in favour of the deviating firm. Since the deviating firm's volumes increase by a discrete amount, with upwards distorted marginal prices the move is definitely profitable.[2]

The above arguments confirm that the profitability of exclusive contracts crucially rests upon the impossibility of extracting buyers' surplus efficiently. Taking one step backwards, then, the question naturally arises of what may impede perfect rent extraction if fixed fees are, indeed, feasible. One possible answer is risk aversion, as in the moral hazard model originally proposed by Bernheim and Whinston (1998, sect. V) and further analyzed in this paper. In this model, buyers are risk-averse retailers and demand is uncertain. Here, the cost of fixed fees is that they expose retailers to the risk of making large fixed

[2] This argument demonstrates the unilateral incentive to offer exclusive dealing contracts, but does not imply that such contracts are profitable when the rival's reaction is taken into account. As we shall see, this requires that the deviating firm possesses a sufficiently large competitive advantage over the rival. Likewise, with more than two competitors even the existence of a unilateral incentive requires that the dominant firm is sufficiently stronger than its rivals.

payments even if demand turns out to be low. To reduce the risk, upstream firms lower their fixed fees and distort marginal prices upwards.

Another answer is that firms may not exactly know the buyers' willingness to pay for their products, as in the adverse selection model of Calzolari and Denicolò (2013, 2015), a variant of which is further explored here. In this case, fixed fees may create a distortion at the extensive margin by excluding some low-demand buyers. Balancing distortions at the extensive and intensive margins, firms again set marginal prices above marginal costs.

Other factors may likewise impede efficient rent extraction. Rather than considering each separately, we propose a reduced-form model that may encompass many of them. The reduced-form model simply assumes that it is costly to extract buyers' rent by means of fixed fees, without specifying the nature of the cost. We show that this model produces qualitatively similar results to the more highly structured ones, being exactly equivalent in some cases, and providing a close approximation in others.

The rest of the paper proceeds as follows. In section 2, we present a model of duopolistic competition in two-part tariffs with asymmetric information. In section 3, we show that in this framework firms generally have incentives to distort marginal prices, and we present a reduced-form model that can capture these incentives in a stylized way. Section 4 compares the equilibria that prevail when exclusive contracts are prohibited or permitted and derives our main results. Section 5 analyzes whether firms can coordinate on a better equilibrium by offering contracts that are destined not to be accepted. Section 6 discusses the implications for competition policy and concludes the paper. All proofs are in the Appendixes.

2 The model

We start from a fully fledged model of price competition with uncertain demand that follows closely Bernheim and Whinston (1998, sect. V). One can further specify the model as one of moral hazard (as Bernheim and Whinston do) or adverse selection, depending on the informational assumptions made.

To eschew inter-temporal trade offs, the model is one stage. There are two substitute goods, denoted by $i = 1, 2$, which are produced by upstream firm 1 and 2, respectively. Marginal costs c_i are constant, and we abstract from fixed costs. Upstream firms sell to retailers or, more generally, downstream firms. Retailers do not interact strategically with each other, so we can focus on the firms' relationships with a single retailer.

The gross profit that the retailer can make with Q_i units of good i and Q_j units of good j is denoted by $V(Q_i, Q_j, \theta)$. The variable θ represents the state of demand and is stochastic. It is distributed according to a distribution function $G(\theta)$ with positive, finite density $g(\theta) > 0$ over the support $(\underline{\theta}, \overline{\theta})$. Following Bernheim and Whinston, we assume that uncertainty is multiplicative. This requires that $V(Q_i, Q_j, \theta)$ is homogeneous of degree one, i.e. $V_{Q_i}(Q_i, Q_j, \theta) =$

$V_{Q_i}\left(\frac{Q_i}{\theta}, \frac{Q_j}{\theta}, 1\right)$, so that for any given marginal prices the demand for each product is proportional to θ. We can then write $V(Q_i, Q_j, \theta) = \theta v(q_i, q_j)$, where $q_i \equiv \frac{Q_i}{\theta}$ and $v(q_i, q_j) \equiv V\left(\frac{Q_i}{\theta}, \frac{Q_j}{\theta}, 1\right)$. With no further loss of generality, the average value of θ is normalized to 1: $\int_{\underline{\theta}}^{\bar{\theta}} \theta g(\theta) d\theta = 1$.

We make standard regularity conditions on the payoff function $v(q_i, q_j)$. In particular, we assume that it is at least twice continuously differentiable in q_i and q_j, and that the goods are imperfect substitutes:

$$v_{q_i q_i}(q_i, q_j) \leq v_{q_i q_j}(q_i, q_j) \leq 0.$$

(Subscripts denote partial derivatives, and the above inequalities are assumed to be strict when quantities are strictly positive.) Furthermore, the function $v(q_i, q_j)$ is assumed to: (i) vanish when $q_1 = q_2 = 0$; (ii) exhibit finite satiation points $\bar{q}_i$ implicitly defined by $v_{q_i}(\bar{q}_i, 0) = 0$; and (iii) exhibit finite choke prices $\bar{p}_i = v_{q_i}(0, 0)$.

Firms may be asymmetric; in particular, upstream firm 1 (which we shall refer to as the dominant firm) may have a competitive advantage over its rival. The competitive advantage may be due to lower costs, higher quality, or a combination of the two. To fix ideas, however, we shall assume that demand is symmetric (which is equivalent to the symmetry of $v(q_i, q_j)$) and focus on cost asymmetries. In particular, we shall take the cost gap $c \equiv c_2 - c_1 \geq 0$ as a measure of the dominant firm's competitive advantage.[3] Normalizing good 1's unit production cost to zero, c then is the unit cost of producing good 2. To keep the analysis interesting, we assume that c is not so large that firm 2 is automatically foreclosed: the condition is $c < c_{DRAS}$, where c_{DRAS} will be defined more precisely below.

Following again Bernheim and Whinston, we assume that upstream firms compete in two-part tariffs $P_i = F_i + p_i Q_i$, where F_i is the fixed fee and p_i is the marginal price. A tariff is denoted by (p_i, F_i). Note that with constant marginal costs, two-part tariffs could in principle allow efficient profit extraction. The price distortions identified below are therefore due to market imperfections, not to restrictions on the firms' strategy space made for the sake of tractability.

We distinguish between two modes of competition, depending on whether firms may or may not use exclusive contracts. When they may, upstream firm i may offer either an exclusive tariff (denoted by superscript E), which requires $Q_j = 0$, or a non-exclusive tariff (denoted by superscript NE), which allows for $Q_j > 0$. When exclusive contracts are prohibited, in contrast, upstream firms may offer only one tariff, which applies irrespective of whether the rival's sales are positive or nil.[4]

[3] The model is analytically equivalent to one in which c is an index of vertical differentiation, with product 1 being of greater quality, and hence in greater demand, than product 2. In this interpretation, the retailer's payoff function would be $v(q_1, q_2) - cq_2$, with symmetric production costs.

[4] The assumption that each firm can offer only one tariff may be restrictive, especially in the adverse selection framework described below, but is not essential for our results. In section 5 we shall allow each firm to offer two tariffs, an exclusive and a non-exclusive one.

Upstream firms choose their tariffs independently and simultaneously. The retailer then chooses which contracts to sign, and the quantities to buy.

We shall refer to the case in which the retailer elects to buy from one firm only as *exclusive representation*, and the case where he buys from both as *common representation*. Either firm can unilaterally enforce exclusive representation, whereas common representation may prevail only if both firms offer non-exclusive contracts.

2.1 Moral hazard

With these common assumptions, one can specify a model of moral hazard or adverse selection by introducing different forms of asymmetric information. For example, Bernheim and Whinston (1998) posit that the state of demand θ is unknown to all players at the contracting stage but is revealed to the retailer before actual quantities are chosen. The retailer therefore chooses which contracts to sign before observing θ, and the quantities Q_1 and Q_2 after. The retailer is risk averse and θ is not contractible (only quantities are), so we have a problem of moral hazard.

Bernheim and Whinston propose this model in order to provide rigorous theoretical foundations for the view that exclusive dealing may be an efficient contractual device. They assume a Constant Absolute Risk Aversion (CARA) utility function, which implies that it may be desirable to provide insurance by means of negative fixed fees (i.e., lump-sum subsidies). In the absence of exclusivity clauses, however, lump-sum subsidies cannot be offered when the goods are sufficiently close substitutes, as the retailer could then pocket the subsidy offered by one firm and purchase the product from the other. Exclusive dealing prevents this opportunistic behaviour, and thus can improve efficiency.

However, risk aversion also implies that upstream firms have an incentive to distort marginal prices upwards. We contend that this creates other, possibly less benign reasons for offering exclusive contracts. To abstract from the pro-efficiency effect that Bernheim and Whinston focus on and highlight the consequences of the upwards distortion in marginal prices, we choose a different specification of risk aversion. That is, we replace the CARA von Neumann-Morgenstern utility function with a piecewise linear one, with a kink at the origin. In other words, the retailer is risk neutral in the region of monetary gains and in the region of monetary losses, but dislikes losses more than he likes gains. This assumption rules out the possibility that lump-sum subsidies may be optimal,[5] eliminating the rationale for exclusive dealing discussed by Bernheim and Whinston (1998).

To proceed, normalize to one the slope of the utility function in the region of gains, and denote its slope in the region of losses by $1+\lambda$, so that the parameter

[5] Risk aversion still creates a demand for insurance that upstream firms meet by decreasing their fixed fees. However, as soon as the fixed fees vanish the retailer is guaranteed to stay in the region of gains, where he is effectively risk neutral and thus demands no more insurance. (Fixed costs could be easily accommodated by suitably adjusting the position of the kink in the utility function.)

$\lambda \geq 0$ measures the degree of risk (loss) aversion. The retailer's *ex ante* expected utility then is

$$U = (1+\lambda) \int_{\underline{\theta}}^{\hat{\theta}} \Pi_R(\theta) g(\theta) d\theta + \int_{\hat{\theta}}^{\bar{\theta}} \Pi_R(\theta) g(\theta) d\theta, \tag{1}$$

where $\Pi_R(\theta)$ denotes the retailer's *ex post* payoff, net of any payment to the firms, as a function of the state of demand θ. The cut-off $\hat{\theta}$ corresponds to $\Pi_R(\hat{\theta}) = 0$; if $\Pi_R(\theta)$ is always positive, or negative, then (1) applies with $\hat{\theta} = \underline{\theta}$, or $\hat{\theta} = \bar{\theta}$. Equation (1) may have a behavioural flavour, but is in fact fully consistent with expected utility theory.

Upstream firms are risk neutral and maximize expected profits

$$\Pi_i = \int_{\underline{\theta}}^{\bar{\theta}} \theta q_i (p_i - c_i) g(\theta) d\theta + F_i. \tag{2}$$

2.2 Adverse selection

In the adverse selection specification of the model, the retailer knows the state of demand θ at the contracting stage whereas upstream firms do not. The retailer then chooses both which contracts to sign and the volumes to purchase conditional on θ, and thus he maximizes $\Pi_R(\theta)$ pointwise. As a result, his attitude towards risk is now irrelevant. The upstream firms' payoffs are still given by (2).

The resulting model is similar to that analyzed by Calzolari and Denicolò (2013, 2015). The main difference is that they allow firms to offer *menus* of two-part tariffs, which gives plenty of scope for price discrimination. Their detailed analysis of the optimal screening of buyers may suggest that exclusive dealing should indeed be viewed as a means to better price discriminate. But in fact the effects of exclusive contracts that they uncover are driven to a large extent simply by the fact that marginal prices are distorted upwards, rather than by price discrimination in itself. To highlight this point, here we restrict firms to offer only one two-part tariff, reducing the scope for price discrimination as far as this is possible in this adverse selection framework. This clarifies that the reason why firms use exclusive contracts is not necessarily to better screen the buyers, but rather to increase the demand for their products.[6]

[6] Another difference is that Calzolari and Denicolò focus on the case of an uncovered market, which requires that $\underline{\theta}$ is sufficiently small. Here, in contrast, our main focus is on the case in which $\underline{\theta}$ is close to 1 so that the market is covered.

3 Distortionary pricing

In this section, we analyze price distortions under moral hazard and adverse selection, and how these distortions can be captured by a simple reduced-form model. The analysis focuses on a firm's best response to its rival's strategy; the characterization of the equilibrium is postponed to the next sections.

Generally speaking, with two-part tariffs a firm can extract its profits in two ways: by charging a positive fixed fee, or by charging a positive price-cost margin. To understand how profits are extracted optimally, we decompose the upstream firms' profit into two parts, which loosely speaking correspond to what can be obtained by each of those tools.

Let us consider firm i. For ease of notation, suppose that its rival, firm j, offers a competitive non-exclusive tariff $(c_j, 0)$; the logic of other cases is similar. Denote by $\theta\upsilon(q_i)$ the *indirect payoff function*, i.e. the highest payoff that the retailer can obtain by purchasing q_i units of good i and then trading optimally with firm j. Under exclusive dealing, the indirect payoff function is simply $\upsilon^E(q_i) = v(q_i, 0)$; under common representation, it is $\upsilon^{NE}(q_i) = \max_{q_j \geq 0}[v(q_i, q_j) - c_j q_j]$. Both functions $\upsilon^E(q_i)$ and $\upsilon^{NE}(q_i)$ are continuous, increasing up to the satiation point, and concave. They are also smooth almost everywhere; however, $\upsilon^{NE}(q_i)$ exhibits a kink where $\arg\max_{q_j \geq 0}[v(q_i, q_j) - c_j q_j]$ vanishes.[7]

While the difference between $\upsilon^E(q_i)$ and $\upsilon^{NE}(q_i)$ is crucial for the profitability of exclusive dealing arrangements, as we shall see below, it is largely irrelevant for the analysis of price distortions. Therefore, in this section we shall refer to a generic indirect payoff function $\upsilon(q_i)$ The inverse demand for product i (i.e., the residual demand) is found by maximizing $\upsilon(q_i) - p_i q_i$ and is therefore $p_i = \upsilon_{q_i}(q_i)$. Its inverse, the direct demand function, is denoted by $q_i^{\mathrm{d}}(p_i)$. Since there is a one-to-one correspondence between p_i and q_i, one can write profit and expected utility as functions of p_i or, equivalently, of q_i.

3.1 Moral hazard

Consider first the moral hazard model. To avoid uninteresting cases in which the retailer always ends up in the region of gains, and hence is effectively risk neutral, in the analysis of this case we shall henceforth set $\underline{\theta} = 0$.

[7] For $\upsilon^E(q_i)$, these properties follow immediately from the model's assumption. For $\upsilon^{NE}(q_i)$, denoting $q_j^*(q_i) = \arg\max_{q_j \geq 0}[v(q_i, q_j) - c_j q_j]$, by the envelope theorem we have

$$\upsilon_{q_i}^{NE}(q_i) = v_{q_i}(q_i, q_j^*(q_i)).$$

As long as $q_j^*(q_i) > 0$, we also have

$$\upsilon_{q_i q_i}^{NE}(q_i) = v_{q_i q_i}(q_i, q_j^*(q_i)) - \frac{\left[v_{q_i q_j}(q_i, q_j^*(q_i))\right]^2}{v_{q_j q_j}(q_i, q_j^*(q_i))} < 0.$$

When instead $q_j^*(q_i) = 0$, we have $\upsilon^{NE}(q_i) = \upsilon^E(q_i)$ and hence $\upsilon_{q_i q_i}^{NE}(q_i) = v_{q_i q_i}(q_i, 0)$. It is easy to verify that concavity is preserved at the kink.

To guarantee participation, firm i must meet the following constraint:

$$(1+\lambda)\int_{0}^{\hat{\theta}_i} \{\theta\left[\upsilon(q_i) - q_i\upsilon_{q_i}(q_i)\right] - F_i\} g(\theta)d\theta +$$
$$+\int_{\hat{\theta}_i}^{\bar{\theta}} \{\theta\left[\upsilon(q_i) - q_i\upsilon_{q_i}(q_i)\right] - F_i\} g(\theta)d\theta \geq \upsilon^{NE}(0), \qquad (3)$$

where $\upsilon^{NE}(0)$ is the expected utility that the retailer can obtain by dealing with firm j only,[8] and the cut-off $\hat{\theta}_i$ is

$$\hat{\theta}_i = \frac{F_i}{\upsilon(q_i) - q_i\upsilon_{q_i}(q_i)}$$

if $\frac{F_i}{\upsilon(q_i)-q_i\upsilon_{q_i}(q_i)} \in (0,\bar{\theta})$; otherwise, $\hat{\theta}_i$ is either 0 or $\bar{\theta}$.

Clearly, the participation constraint must bind in equilibrium. Solving for F_i, substituting back into (3) and rearranging one gets

$$\Pi_i(q_i) = \alpha_i\left[\upsilon(q_i) - c_iq_i - (1-\xi_i)\upsilon^{NE}(0)\right] + (1-\alpha_i)\pi_i(q_i), \qquad (4)$$

where $\pi_i(q_i) = q_i\left[\upsilon_{q_i}(q_i) - c_i\right]$, $\alpha_i = \frac{1+\Psi(\hat{\theta}_i)}{1+\lambda G(\hat{\theta}_i)}$, and $\xi_i = \frac{\Psi(\hat{\theta}_i)}{1+\Psi(\hat{\theta}_i)}$, with $\Psi(\hat{\theta}_i) \equiv \lambda\int_{0}^{\hat{\theta}_i}\theta g(\theta)d\theta$. Both coefficients α_i and ξ_i range in between 0 and 1.

Firm i's pricing strategy is implicitly determined by the choice of the output level q_i that maximizes $\Pi_i(q_i)$: the marginal price is $p_i = \upsilon_{q_i}(q_i)$, and the fixed fee is then pinned down by the participation constraint (3).

Expression (4) shows that a firm's profit may be written as a weighted average of its *marginal contribution* $\upsilon(q_i)-c_iq_i-(1-\xi_i)\upsilon^{NE}(0)$ and the *linear-pricing profit* $\pi_i(q_i)$, with weights α_i and $(1-\alpha_i)$, respectively. The marginal contribution is the bilateral surplus due to the fact that the retailer trades with firm i. Under certainty, this would be $\upsilon(q_i) - c_iq_i - \upsilon^{NE}(0)$, but in expression (4) the outside option $\upsilon^{NE}(0)$ is adjusted to account for risk aversion. The adjustment factor, $1-\xi_i$, collapses to one when $\lambda = 0$ (the retailer is risk neutral) or $\hat{\theta}_i = 0$ (the retailer is always in the region of gains): in these cases, no adjustment is needed.

Loosely speaking, the first term in expression (4) corresponds to the profit that can be extracted by means of the fixed fee, the second to the profit that can be extracted by charging a marginal price that exceeds the marginal cost. The weights α_i and $(1-\alpha_i)$ determine how much of the profit is extracted by means of each of these tools.

[8] Since $F_j = 0$, a retailer who buys from firm j only is guaranteed to always be in the region of gains. Therefore, his expected utility equals his expected payoff. With multiplicative uncertainty, the expected payoff is simply the payoff obtained when $\theta = 1$.

Consider, for instance, the limiting case $\alpha_i = 1$ that is obtained when $\lambda = 0$, i.e. under risk neutrality. In this case, the profit function reduces to the marginal contribution, $\upsilon(q_i) - c_i q_i - \upsilon^{NE}(0)$. The optimal marginal price p_i then equals the marginal cost c_i, and the profit is extracted by means of fixed fees only.

The opposite case, $\alpha_i = 0$, is obtained in the limit as $\lambda \to \infty$, i.e., with infinite risk aversion. In this case, the fixed fee vanishes, and the profit is extracted by means of positive price-cost margins only (see the Appendix for a formal proof).

For intermediate degrees of risk aversion, the weights are endogenous. However, it can be shown that firms always complement fixed fees with positive price-cost margins as means of profit extraction.

Proposition 1 *For any finite $\lambda > 0$, both the optimal fixed fee and the optimal price-cost margin are positive: $F_i > 0$ and $p_i > c_i$.*

A similar result can be found in Png and Wang (2010).[9] The intuition is that relying exclusively on the fixed fee as a means of rent extraction exposes the retailer to the risk of making large fixed payments even if demand turns out to be low. To reduce the risk, upstream firms lower the fixed fee and distort the marginal price upwards.

3.2 Adverse selection

Consider next the adverse selection model. To begin with, suppose that $\underline{\theta}$ is sufficiently close to one that the market is covered, meaning that it is optimal to make the retailer sign the contract in all states of demand. In this case, the participation constraint is:

$$\underline{\theta}\left[\upsilon(q_i) - \upsilon_{q_i}(q_i)q_i\right] - F_i = \underline{\theta}\upsilon^{NE}(0). \tag{5}$$

Solving for F_i and substituting into (2) one gets

$$\Pi_i(q_i) = \underline{\theta}\left[\upsilon(q_i) - c_i q_i - \upsilon^{NE}(0)\right] + (1 - \underline{\theta})\pi_i(q_i). \tag{6}$$

The profit is again a weighted average of the marginal contribution and the linear-pricing profit, as in the moral hazard model. However, the weights are now exogenous, which simplifies the analysis considerably.

With no uncertainty, i.e. $\underline{\theta} = 1 = \overline{\theta}$, the profit reduces to the marginal contribution, so the rent is extracted by means of the fixed fee only. But as soon as there is some uncertainty, so that $\underline{\theta} < 1$, the weight of the last term of (6) becomes positive and thus marginal prices are distorted upwards.

As $\underline{\theta}$ gets still lower, however, firm i may optimally choose to exclude some low-demand retailers, which implies that the market becomes uncovered. To account for this possibility, denote now by $\hat{\theta}_i$ the "marginal retailer", i.e. the

[9] To be precise, Png and Wang show that $p > c$ holds as long as the total and the marginal payoffs are positively correlated. The assumption of multiplicative uncertainty guarantees that this is always true in our model.

lowest state of demand for which the retailer purchases a positive amount of good i; it is implicitly defined by the condition:

$$\hat{\theta}_i \left[\upsilon(q_i) - \upsilon_{q_i}(q_i)q_i\right] - F_i = \hat{\theta}_i \upsilon^{NE}(0).$$

Solving for F_i and substituting into (2) one now gets

$$\Pi_i(q_i) = \beta_i \left[\upsilon(q_i) - c_i q_i - \upsilon^{NE}(0)\right] + \gamma_i \pi_i(q_i), \tag{7}$$

where $\beta_i = \hat{\theta}_i \left[1 - G(\hat{\theta}_i)\right]$ and $\gamma_i = \int_{\hat{\theta}_i}^{\bar{\theta}} \theta g(\theta) d\theta - \beta_i$. If $\hat{\theta}_i = \underline{\theta}$, we re-obtain formula (6). When instead the market is uncovered (i.e., if $\hat{\theta}_i > \underline{\theta}$) the weights β_i and γ_i add up to less than one. This reflects the profit lost because the retailer is inefficiently excluded in low states of demand. Furthermore, with an uncovered market the weights are again endogenous, as in the moral hazard model.

Irrespective of whether the market is covered or uncovered, as soon as $\underline{\theta} < \bar{\theta}$ the marginal price is distorted upwards, and the fixed fee is correspondingly reduced.

Proposition 2 *For any $\underline{\theta} < \bar{\theta}$, the optimal fixed fee is non-negative and the optimal price-cost margin is positive: $F_i \geq 0$ and $p_i > c_i$.*

The distortion in the marginal price is a familiar property of models of optimal screening. Intuitively, a firm that relies exclusively on the fixed fee maximizes the surplus extracted from the retailer when $\theta = \hat{\theta}_i$, but leaves him too much rent in higher demand states. Distorting prices upwards (and hence quantities downwards) reduces the rent that is left to infra-marginal retailers.

3.3 Reduced form

The analogy between the moral hazard and the adverse selection model suggests that a similar pattern may emerge whenever it is costly to extract the retailer' rent by means of fixed fees, and irrespective of the precise reason why this is so. Thus, let us directly assume that fixed fees are costly. We take the cost to be proportional to the size of the fixed fee; in other words, we assume that by charging a fixed fee $F_i > 0$ the firm gains F_i but the retailer loses $(1+\mu)F_i$, with $\mu \geq 0$.

The parameter μ may capture various imperfections that impede rent extraction by means of fixed fees. For example, buyers may underestimate future demand, as in Della Vigna and Malmendier (2004) and a rapidly growing body of related literature. In this case, sellers cannot fully extract the true expected gains from trade using their fixed fees because the perceived gains are lower than the true ones. Other imperfections like this may make fixed fees costly to use. The reduced-form model is agnostic about the exact source of the inefficiency, and focuses on the consequences of imperfect rent extraction rather than on its causes.

For consistency, we assume that the cost appears only when $F_i > 0$, so that lump-sum subsidies do not entail any gain. This guarantees that lump-sum subsidies cannot be optimal, as is true in the more highly structured models considered above. Analytically, in the reduced-form model the profit function will have a kink in correspondence of $F = 0$.

With these assumptions, the participation constraint is (demand uncertainty is no longer relevant here, so we simply set $\theta = 1$):

$$\upsilon(q_i) - \upsilon_{q_i}(q_i)q_i - (1+\mu)F_i = \upsilon^{NE}(0). \tag{8}$$

The upstream firm's profit Π_i then becomes:

$$\Pi_i(q_i) = \frac{1}{1+\mu}\left[\upsilon(q_i) - c_i q_i - \upsilon^{NE}(0)\right] + \frac{\mu}{1+\mu}\pi_i(q_i) \tag{9}$$

as long as $F_i > 0$; otherwise, $\Pi_i(q_i) = \pi_i(q_i)$.

The parameter μ plays the same role as λ or $\underline{\theta}$ in the structural models; that is, it determines what part of the firm's total profit is extracted by means of distorted marginal prices and what by means of fixed fees. In particular, expression (9) immediately implies a positive wedge between $p_i = \upsilon_{q_i}(q_i)$ and c_i as soon as $\mu > 0$; the magnitude of the distortion increases with μ. Notice that for $\mu = \frac{1-\underline{\theta}}{\underline{\theta}}$ one re-obtains exactly the case of adverse selection with a covered market.

* * *

This section has argued that fixed fees are generally imperfect means of rent extraction. The imperfection may be due to risk aversion, inefficient exclusion, or other similar factors. Whatever the source of the imperfection, as soon as *fixed fees are costly* (an expression that we use as a shorthand for the assumption that $\lambda > 0$ in the moral hazard model, $\underline{\theta} < \bar{\theta}$ in the adverse selection model, or $\mu > 0$ in the reduced form model) firms optimally distort marginal prices upwards. In the remainder of the paper, we shall analyze the implications of these distortions for exclusive dealing.

4 Equilibrium

In this section, we compare the equilibrium when exclusive dealing is permitted and when it is prohibited. As we proceed, it will become clear that our results are driven by the upwards distortion in marginal prices and hence hold in all models considered here, as soon as fixed fees are costly. To avoid repetitions, however, we present only the formulas for the reduced-form model, which are simplest. We note in footnotes any modifications applying to the moral hazard and adverse selection cases.

4.1 Exclusive dealing prohibited

Our benchmark is the case in which exclusive contracts are prohibited. In this case, both firms must offer a non-exclusive two-part tariff. To stress strategic interactions, the retailer's indirect payoff function will henceforth be denoted by $\upsilon^{NE}(q_i, p_j) = \max_{q_j \geq 0} [v(q_i, q_j) - p_j q_j]$, his reservation payoff (gross of the fixed fee) by $\upsilon^{NE}(0, p_j)$, and so on.

For any given tariff (p_j, F_j) offered by its rival, proceeding as in Section 3 we can write firm i's profit $\Pi_i(q_i, p_j)$ as a linear combination of the marginal contribution and the linear pricing profit $\pi_i^{NE}(q_i, p_j) = \left[\upsilon_{q_i}^{NE}(q_i, p_j) - c_i\right] q_i$. In the reduced-form model, the profit function is

$$\Pi_i(q_i, p_j) = \frac{1}{1+\mu}\left[\upsilon^{NE}(q_i, p_j) - c_i q_i - \upsilon^{NE}(0, p_j)\right] + \frac{\mu}{1+\mu}\pi_i^{NE}(q_i, p_j).^{10} \tag{10}$$

Firm i's best response to (p_j, F_j) is obtained by maximizing $\Pi_i(q_i, p_j)$ with respect to q_i. With no loss of generality, q_i may be restricted to range in the interval $[0, \bar{q}_i]$. Thus, an optimal choice q_i^* exists. The corresponding marginal price is $\upsilon_{q_i}^{NE}(q_i^*, p_j)$, and the fixed fee is pinned down by the participation constraint. In this way, q_i^* implicitly defines firm i's best response. The equilibrium of the pricing game then is determined as the fixed point of the best response functions. Existence of an equilibrium may require further regularity conditions (such as for instance that $\Pi_i(q_i, p_j)$ is quasi-concave in q_i) that we shall take for granted in what follows. Uniqueness is not really necessary for our results, but for ease of exposition we shall proceed as if the equilibrium were unique. We denote the equilibrium tariffs by (p_i^{NE}, F_i^{NE}), and the corresponding quantities by q_i^{NE}.

Our analysis relies on two general properties of this benchmark equilibrium. The first is that marginal prices are distorted upwards as soon as fixed fees are costly. As we have seen in the preceding section, this property is exhibited by any best response, and thus it must hold in any equilibrium. The second property is what in the common agency jargon is known as the principle of individual excludability. The idea is very simple: since in equilibrium each firm sets its fixed fee at the level where the retailer is just indifferent between accepting the contract or not, the retailer must be indifferent between dealing with both firms or with only either one (that is, he can exclude any one supplier at no cost). This implies that even if exclusive dealing is prohibited, the retailer is just indifferent between dealing only with firm 1 or only with firm 2 .

4.2 Exclusive dealing permitted

Now suppose that exclusive contracts are permitted, so that each firm can choose whether to offer a non-exclusive or an exclusive tariff. The main result of this

[10] Expression (9) implicity assumes that the buyer buys also from firm j. If he does not, the profit function includes also a constant term F_j. Being constant, however, this additional term does not affect the optimal quantity.

paper is that as soon as fixed fees are costly, in equilibrium exclusive representation always prevails. In particular, firm 2 makes no sales but stands ready to supply its product at cost; the dominant firm offers an exclusive tariff and wins the competition for exclusives by slightly undercutting its rival.

To state this result formally, let us consider the dominant firm's optimal exclusive tariff when its rival prices at cost, i.e. $p_2 = c$ and $F_2 = 0$. It is irrelevant whether that tariff is offered on an exclusive or non-exclusive basis. In any case, the retailer's reservation payoff is $\upsilon^{NE}(0, c)$, so the dominant firm maximizes

$$\Pi_1^E(q_1) = \frac{1}{1+\mu}\left[\upsilon^E(q_1) - \upsilon^{NE}(0,c)\right] + \frac{\mu}{1+\mu} q_1 \upsilon^E_{q_1}(q_1). \tag{11}$$

We denote the solution to this problem by q_1^E, and the corresponding tariff by (p_1^E, F_1^E).[11]

Proposition 3 *If exclusive contracts are permitted, as soon as fixed fees are costly there is a unique Nash equilibrium. In this equilibrium, exclusive representation prevails; firm 2 offers the tariff* $(c, 0)$ *(either with or without an exclusivity clause) but makes no sales; and firm 1 wins the competition for exclusives by offering the exclusive tariff of* (p_1^E, F_1^E).[12]

We present an informal sketch of the proof, leaving the details to the reader. To begin with, we show by contradiction that there exists no common representation equilibrium. Plainly, any such equilibrium must coincide with the equilibrium arising when exclusive contracts are prohibited. We have just noted that in such an equilibrium the retailer is indifferent between buying only from firm 1 or from firm 2. This implies that each firm can deviate by replacing the non-exclusive tariff with an exclusive one, at slightly reduced prices. Faced with the choice of dealing with either firm, but not both, the retailer would naturally opt for the deviating firm, which offers the better deal. The deviating firm's volumes would then increase by a discrete amount, as the products are substitutes. But this would increase its profits when price-cost margins are strictly positive. This means that starting from a putative common representation equilibrium, each firm has a profitable deviation to exclusivity.[13]

[11] By construction, the retailer is just indifferent between the dominant firm's tariff (p_1^E, F_1^E) and firm 2's tariff $(c, 0)$. To avoid issues of equilibrium existence, we assume that the tie is broken in favour of the dominant firm.

[12] In the adverse selection model, this result holds only as long as the market is covered. With an uncovered market, there might emerge equilibrium patterns in which high-demand types buy exclusively from the dominant firm while some low-demand types buy exclusively from the competitor.

[13] Several remarks are in order. Firstly, in the adverse selection model with an uncovered market, the argument would apply with reference to the lowest type θ who purchases from both firms. The firm that sells to the high types only will then definitely have an incentive to deviate.

Secondly, the argument does not depend on marginal prices being constant. This means that even if one allows firms to offer non-linear tariffs (or menus of two-part tariffs), the conclusion would not change.

Any equilibrium must therefore involve exclusive representation. But under exclusive representation firms compete in *utility space*, where their products are effectively homogeneous. The standard Bertrand logic then implies that the dominant firm wins the competition for exclusives by slightly undercutting its rival. The weaker firm, which is foreclosed, must stand ready to supply its product at competitive terms. It will therefore price at marginal cost, without charging any fixed fee.[14]

Proposition 3 obviously implies that the dominant firm's rival cannot gain from exclusive dealing, which invariably leads to its foreclosure. Perhaps less obviously, the retailer never loses from exclusive dealing. This follows from the principle of individual excludability, which implies that the retailer always obtains exactly the same payoff as if he dealt only with firm 2, and the fact that firm's 2 tariff is lower under exclusive dealing. However, the impact of exclusive dealing on the dominant firm and social welfare is generally ambiguous. We now identify the sources of such ambiguity and show that its resolution turns on the size of the dominant firm's competitive advantage.

4.3 Comparison

Generally speaking, there is a profound difference in the nature of competition with and without exclusive contracts. In the former case, firms compete for the entire volume demanded by a buyer; in the latter, they compete for each marginal unit. This has three important consequences. Firstly, exclusive dealing intensifies competition. Secondly, it reduces efficiency, depriving the buyer of product variety. Finally, exclusive dealing increases the demand for the product of the firm whose exclusive contract is accepted.

The competition-enhancing effect of exclusive dealing is apparent from the fact that in a common representation equilibrium firm 2 may exploit the market power that it possesses thanks to product differentiation, whereas under exclusive dealing it must price competitively. To the extent that prices are strategic complements, the dominant firm will also reduce its tariff. Intuitively, when the products are imperfect substitutes the competition for marginal units is attenuated by product differentiation; the competition for exclusives, in contrast, is not, as it takes place in utility space, where the products are effectively homogeneous.

The second and third effect can be disentangled from the first one by imagining that firm 2 always sticks to the competitive tariff $(c, 0)$. In this case, the retailer's indirect payoff function is $v^E(q_1)$ under exclusive dealing and $v^{NE}(q_1, c)$

Finally, the result would extend to the case of n firms as follows: every firm always has a unilateral incentive to offer a contract that excludes any one of its competitors. However, only a firm that enjoys a sufficiently large competitive advantage may have a unilateral incentive to exclude all of its competitors simultaneously.

[14] If fixed fees did not involve any cost, the deviation described above would not be profitable and so the common representation equilibrium would survive. The equilibrium described in Proposition 3 would still exist, but it would be Pareto dominated, as shown by O'Brien and Shaffer (1997) and Bernheim and Whinston (1998).

under common representation. A comparison of these functions immediately reveals that $v^{NE}(q_1, c) \geq v^E(q_1)$ (implying that the dominant firm's marginal contribution is greater under common representation) and $v_{q_1}^{NE}(q_1, c) \leq v_{q_1}^E(q_1)$ (meaning that the demand for the dominant firm's product is higher under exclusive representation). The first inequality holds by construction; the second follows from from the envelope theorem, which implies $v_{q_1}^{NE}(q_1, c) = v_{q_1}(q_1, q_2)$, and the assumption that $v_{q_1 q_2}(q_1, q_2) \leq 0$. Both inequalities are strict as long as $q_2 > 0$. Loosely speaking, these inequalities imply that exclusive dealing decreases the profits that can be extracted by means of the fixed fee, but increases those extracted by means of a positive price-cost margin.

Clearly, the first two effects tend to make exclusive dealing unprofitable. These effects were highlighted by the so called Chicago school and are the only ones at work in models where pricing is efficient. Therefore, in these models exclusive dealing entails an immediate sacrifice of profit and may become profitable only to the extent that it confers strategic advantages in other stages of the game. With distortionary pricing, however, the third effect also appears, and it tends to make exclusive dealing profitable. We have already seen that this creates unilateral incentives to offer exclusive dealing contracts. The problem is whether the dominant firm can ever eventually gain after taking into account equilibrium responses. We now show that it can, provided that its competitive advantage is sufficiently large.

4.3.1 Large competitive advantage

To prove the claim, we show that if the dominant firm's competitive advantage is sufficiently large, the first two effects discussed above vanish. Only the third effect, which is good for the dominant firm's profit, then remains. To understand how this may happen, one must distinguish between two possible equilibrium patterns that may emerge when exclusive dealing is prohibited.

The first possible outcome is a common representation one, in which both firms sell positive quantities. In this case, for each firm the profit is maximized at a point where the indirect payoff function is smooth. Thus, the following first-order conditions hold:

$$v_{q_i}^{NE}(q_i, p_j^{NE}) + \frac{\mu}{1+\mu} v_{q_i q_i}^{NE}(q_i, p_j^{NE}) q_i = c_i.^{15} \tag{12}$$

The other possibility is a limit pricing equilibrium where exclusive representation prevails even if the dominant firm does not impose any explicit exclusivity clause. At this equilibrium, the dominant firm sets its marginal price at $p_1^{\lim} = v_{q_1}^{NE}(q_1^{\lim}, c)$, where the limit quantity $q_1^{\lim}$ is implicitly defined by the condition that $v_{q_2}(q_1^{\lim}, 0) = c$. Firm 2 is foreclosed and hence prices at cost, setting $p_2^{NE} = c$ and $F_2^{NE} = 0$.

The limit pricing equilibrium arises when the dominant firm's competitive advantage is greater than a critical threshold, $c_{\lim}$, implicitly defined as the solution to

$$\upsilon_{q_1}^{NE}(q_1^{\lim}, c) + \frac{\mu}{1+\mu}\upsilon_{q_1 q_1}^{NE}(q_1^{\lim}, c) q_1^{\lim} = 0.^{16,17}$$

It is important to note that exclusive contracts are not necessarily irrelevant for $c \geq c_{\lim}$. In fact, the condition for exclusive contracts to be irrelevant is $c \geq c_{DRAS} \equiv v_{q_2}(q_1^E, 0)$. This guarantees that the competitive pressure exerted by the dominant firm's rival is so weak that foreclosure does not require any strategic expansion of the output (analytically, we have $q_1^E \geq q_1^{\lim}$). To make the analysis interesting, we assume that $c < c_{DRAS}$. But it is immediate to see that $c_{\lim} < c_{DRAS}$. Therefore, when exclusive dealing is banned we have a common representation equilibrium for $c < c_{\lim}$ and a limit pricing equilibrium for $c_{\lim} \leq c < c_{DRAS}$.[18]

In this latter case, it is evident that the dominant firm gains from exclusive dealing. This follows immediately from the fact that both with and without exclusive dealing its profit is $\Pi_1^E(q_1)$. However, in the absence of exclusive dealing the dominant firm must set $q_1 \geq q_1^{\lim}$; with exclusive dealing, in contrast, it can maximize $\Pi_1^E(q_1)$ freely.

It is also easy to show that when $c_{\lim} \leq c < c_{DRAS}$ exclusive dealing reduces the social surplus $S = v(q_1, q_2) - cq_2$. This follows immediately from the fact that $q_1^E < q_1^{NE}$ and $q_2^E = q_2^{NE} = 0$.[19] We can therefore conclude that when the benchmark is a limit pricing equilibrium, exclusive dealing is both profitable and anticompetitive.

By continuity, this conclusion will continue to hold even if c is slightly lower than $c_{\lim}$. But in fact the result remains true even if the competitive advantage is significantly lower than $c_{\lim}$. To show this, define the *positive primary outputs* threshold, c_{PPO}, as the critical level of c below which the efficient quantity

[16] At $c = c_{\lim}$, when $p_2 = c$ the dominant firm's first-order condition delivers an optimal quantity of product 1 exactly equal to $q_1^{\lim}$. This implies that the residual demand for product 2 lies entirely below the marginal cost curve. It follows that $q_2 = 0$: firm 2 is foreclosed, so to speak, inadvertently. For $c > c_{\lim}$, in contrast, the dominant firm prices precisely with the objective of foreclosing the rival. Analytically, the profit function $\Pi_1(q_1, c)$ is increasing to the left of $q_1^{\lim}$, whereas to the right of $q_1^{\lim}$ it coincides with $\Pi_1^E(q_1)$ and hence is decreasing provided that $q_1^E < q_1^{\lim}$. Therefore, the maximum is reached exactly at $q_1 = q_1^{\lim}$, where the profit function exhibits a kink.

[17] Existence of $c_{\lim}$ is guaranteed by our regularity conditions (in particular, the assumption of a finite choke price). Uniqueness of $c_{\lim}$ requires that q_1^* is non-decreasing in p_2. This sounds like a natural property but it can be guaranteed only by restrictions involving third derivatives. The property always hold when the third derivatives vanish, as in the linear-quadratic example presented below. For simplicity, the discussion in the main text assumes uniqueness of $c_{\lim}$.

[18] Notice that $c_{\lim}$, c_{DRAS} and q_1^E depend on μ; $q_1^{\lim}$ and $p_1^{\lim}$, in contrast, do not.

[19] A similar logic implies that final consumers lose from exclusive dealing when the benchmark is a limit pricing equilibrium. The reason for this is simple. Since $q_2^E = q_2^{NE} = 0$, final consumers are deprived of product variety in any case. However, inequality $q_1^E < q_1^{NE} = q_1^{\lim}$ implies that the wholesale price of the only product they purchase is higher under exclusive dealing. This generally implies that the final price will also be higher, and thus that final consumers are worse off.

of product 2 is strictly positive.[20] When the competitive advantage is close to c_{PPO}, even a minimal departure from marginal cost pricing makes exclusive dealing arrangements welfare decreasing and directly profitable for the dominant firm.

Proposition 4 *For any arbitrarily small $\mu > 0$, there exists a neighborhood of c_{PPO} in which exclusive dealing is profitable for the dominant firm and reduces social welfare.*

4.3.2 Small competitive advantage

When the dominant firm's competitive advantage is small, the effects of exclusive dealing are reversed: exclusive dealing is unprofitable and procompetitive.

To show this, take the limiting case in which firms are symmetric, i.e. $c = 0$. In this case, the dominant firm's profit under exclusive dealing vanishes. This follows immediately from the fact that that when $c = 0$, under exclusive dealing both the dominant firm's marginal contribution and the linear pricing profit vanish. Under common representation, in contrast, the dominant firm can take advantage of product differentiation to obtain a positive profit even if $c = 0$. By continuity, exclusive contracts will continue to be unprofitable as long as c is sufficiently small.

This means that in this case firms are trapped in a prisoners' dilemma: both have a unilateral incentive to enter into exclusive dealing arrangements, but both are eventually harmed by such a move. The intuition is as follows. As we have already noted, exclusive dealing intensifies competition by wiping out product differentiation. Who gains and who loses from such more intense competition depends on who is, effectively, protected by product differentiation. When the dominant firm's competitive advantage is large, product differentiation protects mainly the weaker firm, allowing it to maintain a market niche even in the face of a substantially more efficient competitor. In this case, wiping out product differentiation allows the dominant firm to seize the rival's market share, increasing its profit. When the competitive advantage is small, in contrast, product differentiation shields both firms from disruptive competition from the rival. Exclusive dealing arrangements destroy the shield and are therefore unprofitable for both firms.

When the dominant firm's competitive advantage is small, exclusive dealing may increase social welfare and benefit final consumers. This possibility arises

[20] Efficient quantities are given by the conditions

$$\begin{cases} v_{q_1}(\tilde{q}_1, \tilde{q}_2) = 0 \\ v_{q_2}(\tilde{q}_1, \tilde{q}_2) = c, \end{cases}$$

and are always well defined in view of our assumptions. Then, the threshold c_{PPO} is implicitly defined by the condition

$$v_{q_2}(\tilde{q}_1, 0) = c_{PPO}.$$

See Singh and Vives (1984) and Amir and Jin (2001). Therefore, at $c = c_{PPO}$ firm 2 is foreclosed when firm 1 prices efficiently. But in equilibrium foreclosure is less likely, as firm 1 distorts its marginal price upwards. This implies that $c_{\lim} > c_{PPO}$.

because the competition for exclusives lowers equilibrium prices. Of course exclusive dealing entails a loss in terms of reduced variety, but this may be more than offset by the lower prices.

To demonstrate this possibility, we assume that the welfare of final consumers can be proxied by the retailer's payoff *gross of the fixed fees,* $v(q_1) - p_1 q_1$. With this approximation, consider again the case of symmetric firms, $c = 0$, and take the limiting case of linear pricing, $\mu \to \infty$. The result then follows from Beard and Stern (2008), who show that the consumers surplus when only one product is available but is priced competitively is greater than the surplus when both products are monopolized.[21] By continuity, the result will continue to hold for c positive but small, provided that price distortions are sufficiently large.

4.4 Linear demand

Our analysis has shown that when the competitive advantage is large, the dominant firms gains from exclusive dealing while its rival loses; when instead the competitive advantage is small, exclusive dealing harms both firms. To illustrate these results, we work out the equilibrium for the simple case of a quadratic payoff function:[22]

$$v(q_1, q_2) = (q_1 + q_2) - \frac{1}{2}\left(q_1^2 + q_2^2\right) - \gamma q_1 q_2, \tag{13}$$

which implies linear demand functions. With no further loss of generality, we have normalized the coefficients of the payoff function in such a way that both the intercept and the slope of the exclusive demand curve are equal to one. The remaining parameter, γ, captures the degree of substitutability between the goods. It ranges from 0 (independent goods) to 1 (perfect substitutes).

With the payoff function (13), the demand for product 1 is

$$q_1^E = 1 - p_1^E$$

under exclusive dealing, and

$$q_1^{NE} = \frac{1 - p_1^{NE} - \gamma(1 - p_2^{NE})}{1 - \gamma^2}$$

under common representation. For any given price, demand is both lower and more elastic under common representation than under exclusive dealing.

[21] Beard and Stern (2008) actually prove the result for the case of independent products, but their technique can be easily extended to the case of two substitutes.

[22] With uncertain demand, the corresponding payoff function is

$$V(Q_1, Q_2, \theta) = (Q_1 + Q_2) - \frac{1}{2\theta}\left(Q_1^2 + Q_2^2\right) - \frac{\gamma}{\theta} Q_1 Q_2.$$

This can be rewritten as $\theta\left[(q_1 + q_2) - \frac{1}{2}\left(q_1^2 + q_2^2\right) - \gamma q_1 q_2\right]$, which reduces to (14) for $\theta = 1$. In the calculations reported in the Online Appendix, for the moral hazard model, θ has been taken to be uniformly distributed over the interval $(0, 2]$. For the adverse selection model, we have considered a narrower support so as to ensure that the market is covered.

The best response curves are linearly increasing, with a slope lower than one. This guarantees the existence of a unique common representation equilibrium, which can be calculated as:

$$\begin{aligned} p_1^{NE} &= \frac{\mu(1-\gamma)}{1+\mu(2-\gamma)} + c\frac{\gamma\mu(1+\mu)}{(1+2\mu)^2-\mu^2\gamma^2} \\ p_2^{NE} &= \frac{\mu(1-\gamma)}{1+\mu(2-\gamma)} + c\frac{(1+\mu)(1+2\mu)}{(1+2\mu)^2-\mu^2\gamma^2}. \end{aligned}$$

The corresponding equilibrium fixed fees are:

$$\begin{aligned} F_1^{NE} &= \frac{\left(1-\gamma-p_1^{NE}+\gamma p_2^{NE}\right)^2}{2(1-\gamma^2)(1+\mu)} \\ F_2^{NE} &= \frac{\left(1-\gamma-p_2^{NE}+\gamma p_1^{NE}\right)^2}{2(1-\gamma^2)(1+\mu)}. \end{aligned}$$

When $\mu = 0$, marginal prices equal marginal costs, and the surplus is extracted efficiently by means of fixed fees only. In this case, the equilibrium reproduces the truthful equilibrium of a standard common agency game. When instead $\mu \to \infty$, the fixed fees vanish, and the marginal prices converge to the Bertrand prices.[23]

The above tariffs apply as long as

$$c < c_{\lim}(\mu) = \frac{(1-\gamma)[1+\mu(2+\gamma)]}{1+\mu(2-\gamma^2)},$$

a condition that guarantees that the associated quantities are strictly positive. When $c \geq c_{\lim}(\mu)$, we have a limit pricing equilibrium, where

$$p_1 = p_1^{\lim} = 1 - \frac{1-c}{\gamma}. \tag{14}$$

and $p_2 = c$.

With exclusive contracts, in contrast, the optimal marginal price is

$$p_1^E = \min\left[c, \frac{\mu}{1+2\mu}\right].$$

This vanishes when $\mu = 0$ and converges to $\min[c, \frac{1}{2}]$ when $\mu \to \infty$. The corresponding fixed fee is

$$F_1^E = \max\left[0, \frac{\left(c - \frac{\mu}{1+2\mu}\right)\left(2 - c - \frac{\mu}{1+2\mu}\right)}{2(1+\mu)}\right].$$

Comparing profits and social welfare under exclusive dealing and common representation is now just a matter of calculation. We have:

[23] This is the linear pricing case that Mathewson and Winter (1987) focus on.

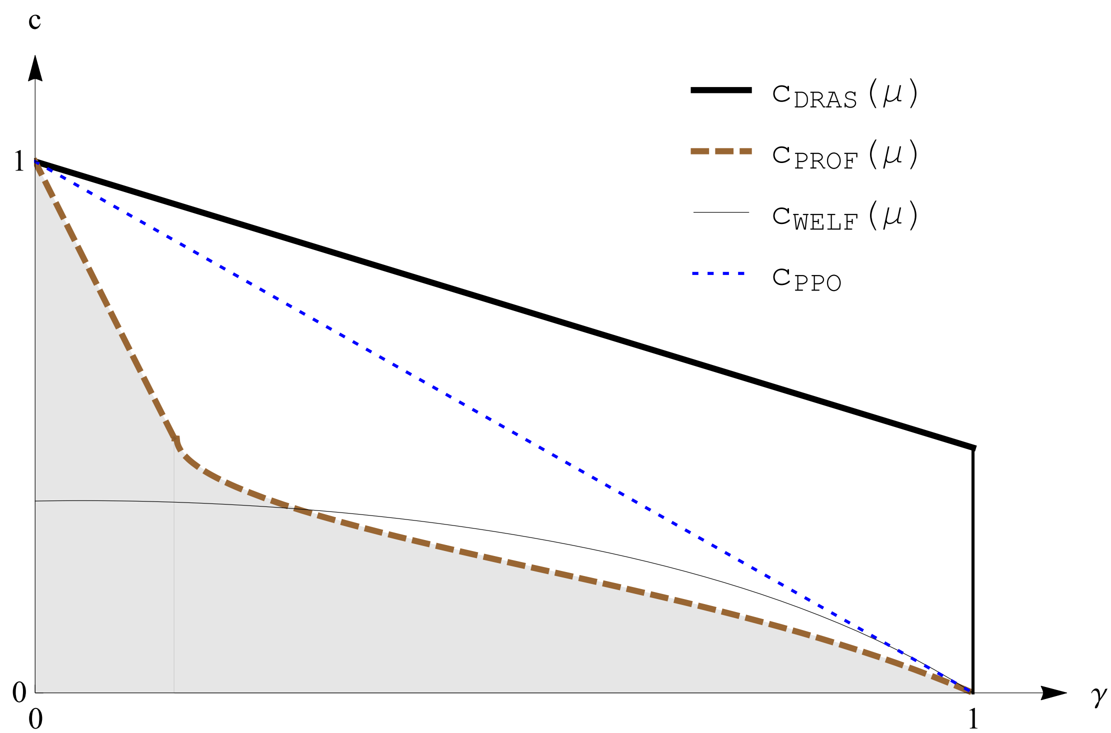

Figure 1: Profitability. Exclusive dealing is profitable in the grey region and reduces social surplus above the threshold $c_{WELF}(\mu)$ ($\mu = 6$).

Proposition 5 *With the quadratic payoff function (13), for any value of μ there exist a lower and an upper threshold, $c_{PROF}(\mu)$ and $c_{DRAS}(\mu)$, such that exclusive dealing is unprofitable for $c \leq c_{PROF}(\mu)$, profitable for $c_{PROF}(\mu) < c < c_{DRAS}(\mu)$, and irrelevant for $c \geq c_{DRAS}(\mu)$. Both $c_{DRAS}(\mu)$ and $c_{PROF}(\mu)$ converge to $c_{PPO} = 1 - \gamma$ as $\mu \to 0$.*

Proof. See the Online Appendix. ■

The curves $c_{PROF}(\mu)$ and $c_{DRAS}(\mu)$ are depicted in Figure 1, together with the frontier $c_{WELF}(\mu)$ between the cases in which exclusive dealing increases or decreases social surplus S. The explicit formulas are reported in the Online Appendix. The Appendix also provides the calculations for the moral hazard and adverse selection models, which lead to similar results.

An increase in μ, which widens price distortions, shifts the curves $c_{PROF}(\mu)$ and $c_{WELF}(\mu)$ down, and the curve $c_{DRAS}(\mu)$ up. Therefore, the stronger are the price distortions, the larger is region where exclusive dealing is profitable and anticompetitive.

A close look at Figure 1 shows that in fact the two thresholds $c_{PROF}(\mu)$ and $c_{WELF}(\mu)$ do not exactly coincide. Focusing on the case in which the products are fairly close substitutes (arguably, the only one relevant for the analysis of exclusive dealing) one notes that when μ is sufficiently large there exists a sub-region where exclusive dealing is both profitable and welfare improving.[24] However, this sub-region vanishes when μ is small. And, in any case, in most of the region where exclusive dealing is welfare improving, it is not profitable for the dominant firm. This means that the procompetitive effects of exclusive dealing are essentially due to a lack of coordination among the firms.

5 Coordination

The analysis above has shown that exclusive dealing is anticompetitive when the dominant firm's competitive advantage is large, procompetitive when it is small. However, these different effects may not seem both equally plausible.

The anticompetitive effects arise when exclusive dealing is profitable for the dominant firm and weakens its competitor. In this case, competition produces winners and losers, and winners have little incentives to alter the outcome. The procompetitive effects, in contrast, arise because of a lack of coordination among the firms: both lose from the cut-throat competition engendered by exclusive contracts. A skeptic might argue that such disruptive competition must in time tend to correct itself. For example, Mathewson and Winter (1987) posit that firms can commit, in a first stage of the game, to a type of contract. With this assumption, exclusive dealing would be observed if and only if it is profitable for the dominant firm, and hence, essentially, only if it is anticompetitive.

However, coordination problems may not be that easy to solve. Here we take the view that some coordination may be possible, but we rule out commitments. Rather, we explore the consequences of relaxing the assumption that each firm can offer only one type of contract – either exclusive or non exclusive. In contrast, we now allow both firms to offer two contracts, one of each type, letting the retailer choose which one to sign. This is the same assumption as made by Bernheim and Whinston (1998).[25] Since the retailer will sign at most one contract with each firm, the contracts that are destined not to be accepted may act as a coordination device.

This section analyzes the equilibria that emerge under this new hypothesis. To preview the results, there are no substantial changes when exclusive dealing is profitable and anticompetitive. The procompetitive effects of exclusive dealing

[24] This possibility was already noted by Mathewson and Winter (1987), who looked at the limiting case $\mu \to \infty$.

[25] This assumption also captures the common practice of "exclusivity discounts," which could not be observed if firms offered one tariff only. An exclusivity discount consists of a combination of two tariffs: the reference (non-exclusive) tariff and a tariff with reduced prices which applies in case the buyer does not purchase from competitors.

are less pronounced than in the baseline model but do not disappear, especially if the degree of product substitutability is significant.

5.1 Scope and limits of coordination

To begin with, note that the equilibrium of Proposition 3 continues to exist also under our new assumption, and remains the only exclusive representation equilibrium. Now, however, this equilibrium may not be unique: there might exist other equilibria in which common representation prevails.

Such common representation equilibria may exist because they can now be supported by exclusive contracts suitably designed in order to counter the unilateral incentives to switch to exclusive representation identified above. Therefore, in these equilibria each firm must offer both a non-exclusive and an exclusive contract. The non-exclusive contracts are the ones destined to be accepted, the exclusive contracts serve to prevent deviations from the equilibrium. In particular, the exclusive contract offered by firm i must make it unprofitable for firm j to switch to exclusive representation, and vice versa.

Reaching an equilibrium of this sort clearly requires a delicate coordination among the firms. Scope for coordination exists because the single sourcing outcome destroys the value of product variety. Firms might therefore gain by providing variety and then extracting the value that it has for the retailer.

If this strategy does work, exclusive contracts will not be signed by the retailer in equilibrium. This opens the way for another form of coordination: firms could also, to some extent, raise their exclusive tariffs in a coordinated fashion. This move cannot be profitable directly, as exclusive contracts are not signed, but can benefit the firms indirectly, allowing them to increase their non-exclusive prices as well.

However, in a non-cooperative equilibrium there are limits to the possibility of coordinating price strategies in this way. Firstly, the coordination mechanism is delicate in that the non-exclusive tariffs must perform two distinct functions: they must induce the retailer to opt for common representation, and split the extra surplus between the firms in such a way that both are better off than under exclusive dealing. It is not obvious that both goals can be achieved simultaneously. Secondly, raising the exclusive tariffs increases the temptation to deviate to exclusive representation. This puts an upper bound on the degree of price coordination that can be sustained.

Before proceeding to the formal analysis, a few preliminaries are in order. Firstly, recall that direct side payments between the firms, which would make the coordination problem trivial, are ruled out in our model. With two-part tariffs, however, side payments can be implemented indirectly, *via* the fixed fees: one firm may lower its fixed fee, thereby allowing the rival to increase its own by the same amount. We limit the possibility of transferring profits between the firms in this fashion by assuming that fixed fees cannot be negative: $F_i \geq 0$.

Secondly, to abstract from irrelevant details, we define any two equilibria as *outcome equivalent* if the retailer ends up choosing the same contracts in both.

All outcome equivalent equilibria are practically identical; as such, they will be regarded as one and the same.

Finally, for ease of exposition we exploit a convenient property of the reduced-form model. That is, the optimal choice q_i^* does not depend on the rival's fixed fee F_j; as is apparent from (9), it depends only on p_j. Therefore, firm i's best response can be written as $p_i = \upsilon_{q_i}^{NE}(q_i^*(p_j), p_j) \equiv BR_i(p_j)$.[26]

5.2 Characterization

We now provide a characterization of the common representation equilibria. As we have argued above, any common representation equilibrium must consist of two tariffs for each firm, a non-exclusive and an exclusive one. The exclusive contracts are not accepted, but serve to prevent deviations to exclusive representation. To begin with, we show that one can restrict attention to exclusive contracts that are efficient, in the sense that they maximize the firm's profit for any given payoff left to the retailer.

Formally, given a retailer's payoff of κ, the efficient exclusive contract for firm i corresponds to the solution to the following program:

$$\max \Pi_i^E(q_i) = \frac{1}{1+\mu}\left[\upsilon^E(q_i) - c_i q_i - \kappa\right] + \frac{\mu}{1+\mu} q_i \upsilon_{q_i}^E(q_i).$$

We denote the solution to this problem by $q_i^E(\kappa)$, the corresponding tariff by $(p_i^E(\kappa), F_i^E(\kappa))$, and the profit by $\Pi_i^E(\kappa)$. We have:[27]

Lemma 1 *For any common representation equilibrium in which the retailer obtains a payoff of κ, there is an outcome-equivalent equilibrium with the same non-exclusive tariffs, in which the exclusive tariffs are given by $(p_i^E(\kappa), F_i^E(\kappa))$.*

In view of the lemma, a common representation equilibrium is fully identified by the non exclusive tariffs (p_i, F_i). Given these, one can calculate the retailer's payoff

$$\kappa = \max_{q_1,q_2}\left[v(q_1,q_2) - p_1 q_1 - p_2 q_2 - (1+\mu)(F_1+F_2)\right], \tag{15}$$

and the corresponding exclusive tariffs $(p_i^E(\kappa), F_i^E(\kappa))$.

Next, let us denote by $\mathcal{E} \subset R_+^4$ the set of four-tuples (p_1, F_1, p_2, F_2) that satisfy the following conditions:

$$\begin{array}{ll} \pi_i^{NE}(p_1,p_2) + F_i \geq \Pi_i^E(\kappa) & (E_i) \\ BR_i(p_j) \geq p_i & (NE_i) \\ \left[BR_i(p_j) - p_i\right] F_i = 0, & (CS_i) \end{array} \tag{16}$$

where, with a slight abuse of notation, the linear-pricing profit π_i is now written as a function of marginal prices rather than quantities. We then have:

[26] The independence of F_j does not extend to those cases in which the weights in the profit functions are endogenous and depend on the fixed fees, as for instance in the moral hazard model. This property is not essential for our results, but simplifies the analysis and the notation.

[27] Our tie-breaking rule here is that a retailer who is indifferent between common and exclusive representation opts for common representation.

Proposition 6 *Any element of $\mathcal{E}$ is a common representation equilibrium; conversely, any common representation equilibrium is outcome equivalent to an element of $\mathcal{E}$.*

Proposition 6 clarifies the scope and limits to coordination informally discussed above. Inequalities (E) say that each firm must prefer its equilibrium non-exclusive contract to the best exclusive one which would be accepted by the retailer. Crucially, since both firms offer an exclusive contract which is worth κ to the retailer, a firm that considers deviating to exclusivity must guarantee to the retailer a payoff of at least κ. If exclusive contracts were not offered, in contrast, a firm could deviate simply by guaranteeing to the retailer what he could obtain by trading only with the rival *at the non-exclusive terms*. This makes deviations to exclusivity less attractive than in the baseline model and explains why common representation equilibria may now be sustained.

However, preventing deviations to exclusivity is not the only condition for the sustainability of common representation equilibria: deviations to a different common representation outcome must be prevented as well. This is the role of inequalities (NE) and the corresponding "complementary slackness" conditions (CS). These conditions say that each firm's marginal price must lie on the corresponding best response curve if the firm's fixed fee is positive, or below the best response curve if the fixed fee vanishes. The conditions imply that if firms try to redistribute the profit by means of the fixed fees, they must stick to the equilibrium marginal prices p_i^{NE}; if instead they try to change (reduce) the marginal prices, they must set the fixed fees at zero.

Because of these constraints, coordination cannot completely undo the effects of exclusive contracts. In other words, the equilibrium that prevails when exclusive contracts are prohibited can never be sustained when such contracts are permitted.

Proposition 7 *If exclusive contracts are permitted, as soon as fixed fees are costly there exists no equilibrium that is outcome equivalent to the equilibrium that arises when exclusive contracts are prohibited.*

What coordination might undo are the effects of exclusive contracts on marginal prices. That is, there may exist common representation equilibria in which the marginal prices are exactly p_i^{NE}, and only the fixed fees are different. In this case, exclusive contracts are *quasi-neutral* in that they do not affect the final consumers, at least to the extent that final prices depends only on wholesale marginal prices. The linear demand example shows that this possibility arises in particular when the products are poor substitutes, which makes intuitive sense.

When the products are closer substitutes, however, the effects of exclusive contracts are more pervasive. It might be impossible to sustain any common representation equilibrium at all, and in any case the highest marginal prices that can be sustained are necessarily lower than p_i^{NE}.

To understand why coordination may fail altogether, it is useful to note that coordination is easiest when $F_i = 0$. The reason for this is twofold. Firstly, setting $F_i > 0$ forces p_i to lay exactly on the best response $BR_i(p_j)$, eliminating one degree of freedom. Secondly, raising F_i by ΔF_i increases the left-hand side of (E_i) by the same amount. However, κ decreases by $(1+\mu)\Delta F_i$, and hence the right-hand side increases by at least ΔF_i. This means that when F_i increases conditions (E_i) cannot become looser and may in fact become tighter.

When $F_i = 0$, however, firms cannot redistribute the profits by manipulating the fixed fees. Therefore, marginal prices must play a double role: they must create an efficiency gain with respect to the single sourcing outcome, and at the same time they must distribute the surplus between the firms in an incentive-compatible way. When the dominant firm's competitive advantage is small, firms are almost symmetric. The two goals do not conflict with each other and hence can be attained simultaneously. But when the competitive advantage is larger, the goals become conflicting and cannot be both achieved.

5.2.1 Linear demand

For the case of linear demand – i.e., the quadratic payoff function (13) – it is possible to identify exactly the regions where each of the cases described above arises. We have:

Proposition 8 *With the quadratic payoff function (13), for any value of μ there exist two thresholds, $\gamma(\mu)$ and $c_{ED}(\mu)$. Common representation equilibria exist for $c \leq c_{ED}(\mu)$, whereas only the exclusive representation equilibrium exists when $c > c_{ED}(\mu)$. When common representation equilibria exist, if the products are poor substitutes ($\gamma \leq \gamma(\mu)$) exclusive contracts are quasi-neutral; if instead the products are close substitutes ($\gamma > \gamma(\mu)$), the marginal prices are lower than p_1^{NE} and p_2^{NE}. As μ tends to zero, both $c_{ED}(\mu)$ and $\gamma(\mu)$ collapse to c_{PPO}.*

The expressions for $c_{ED}(\mu)$ and $\gamma(\mu)$ are reported in the Online Appendix and are depicted in Figure 2.

The general picture that emerges is quite clear. First of all, exclusive contracts are of scarce relevance when the products are poor substitutes. In this case, the incentives to offer exclusive contracts are small, and in any case firms may coordinate on an equilibrium in which exclusive contracts do not affect marginal prices. This is consistent with the fact that in reality exclusivity clauses generally involve products that are fairly close substitutes.

When the degree of substitutability is sufficiently high, the crucial parameter is once again the dominant firm's competitive advantage. When it is small, firms may succeed in coordinating on a common representation equilibrium. The profit loss is reduced, but the possibility of offering exclusive contracts still lowers marginal prices and equilibrium profits. Therefore, exclusive contracts are still procompetitive. When the competitive advantage is large, in contrast, coordination fails and exclusive representation prevails. In this case, therefore, exclusive dealing has exactly the same effects as in the baseline model: it is profitable for the dominant firm and anticompetitive.[28]

[28] In fact, there exists a small sub-region where coordination fails even if exclusive dealing is marginally unprofitable for the dominant firm.

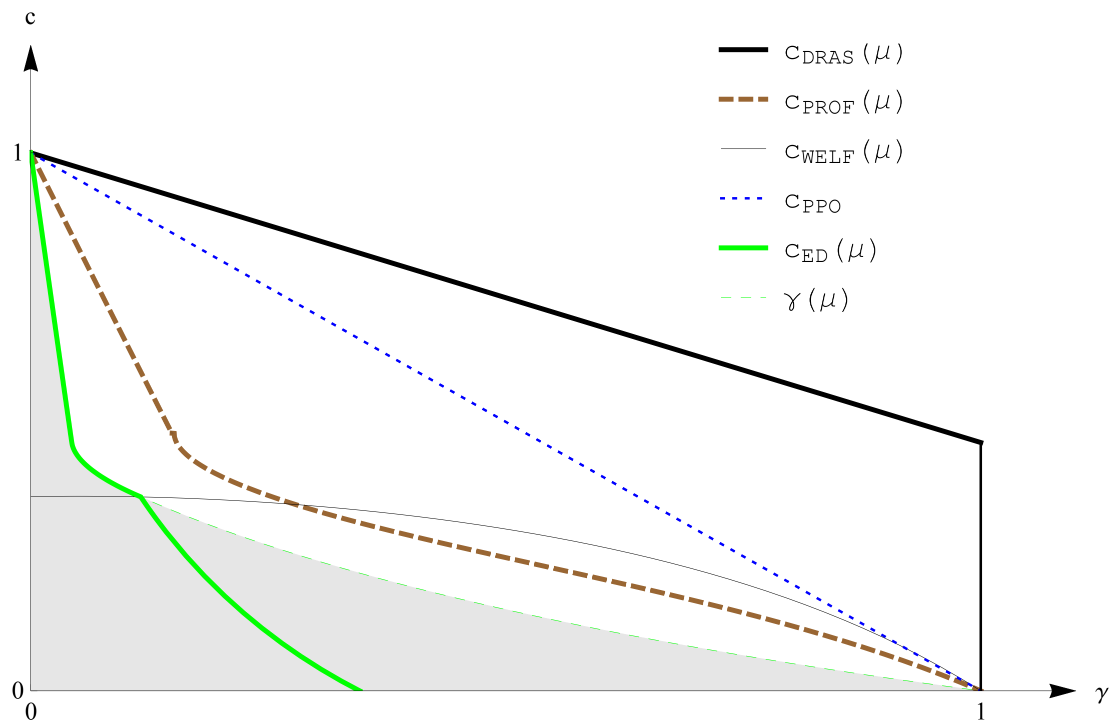

Figure 2: Coordination. Common representation still prevails in the gray region, although with reduced marginal prices between the lines $c_{ED}(\mu)$ and $\gamma(\mu)$ ($\mu = 6$).

Thus, the main difference with respect to the baseline model is that the pro-competitive effects of exclusive contracts are less pronounced, especially when price distortions are small. In the baseline model, even a tiny distortion in the marginal prices destroys the common representation equilibrium. Now, in contrast, when μ is small the region where coordination can undo the effects of exclusive dealing on marginal prices may be large. In the limit, as $\mu \to 0$ exclusive dealing affects marginal prices only in a neighborhood of c_{PPO}.

6 Conclusion

In this paper, we have analyzed exclusive dealing when firms compete in two-part tariffs. We have focused on product market competition, abstracting from any possible effects on entry, exit, investments and so on. Crucially, we have allowed for market imperfections that make it optimal for firms to distort marginal prices upwards.

In this framework, the effects of exclusive dealing depend on the size of the dominant firm's competitive advantage. When it is small, exclusive dealing reduces profits and increases social surplus; when it is large, exclusive dealing is profitable for the dominant firm and anticompetitive.

The analysis has both policy and methodological implications. In terms of policy, it provides a new theory of harm that antitrust authorities can apply in the analysis of exclusive dealing cases. The mechanism through which exclusive dealing produces anticompetitive effects is simply that it increases the demand for the dominant firm's product and decreases the demand for the rivals' products. These consequences of exclusive dealing are automatic, and hence need not be proved. The currently standard theory, in contrast, is based on the notion that exclusive dealing can be profitable only indirectly, by weakening rivals and allowing the dominant firm to gain in the future, or in adjacent markets. It follows the same profit sacrifice/recoupment logic as is commonly adopted in cases of predation. Hence, it implies that policy should weight the immediate social benefits from the allegedly anticompetitive practice against the future costs. These, however, may be difficult to assess as they may not have materialized yet.

Our theory of harm does not require speculating on future social costs. However, three conditions must be met. First, products must be close substitutes. Second, price-cost margins must be non negligible. We believe that these conditions are often satisfied in real world cases. The third and critical condition is that the dominant firm must have a sufficiently large competitive advantage over its rivals. One may wonder whether this last condition can be assessed in practice. We believe that it can, as the competitive advantage correlates with variables, such as the dominant firm's market share, which are readily observable.

It may be tempting to conclude that our findings are supportive of a policy of near *per se* illegality, subject only to possible efficiency defenses. One could indeed argue that the fact that exclusive dealing may procompetitive when the competitive advantage is small need not be a matter of concern. A *per se* prohibition would not apply to cases in which the competitive advantage is small, because in those cases exclusive contracts would hardly be observed, and the most efficient firm would not hold a dominant position (in the American jargon, a "monopoly") in the meaning of antitrust policy.

However, we believe that policy cannot be so simple. The problem is, a dominant position may be found even if the level of dominance is not large enough for exclusive dealing to be directly anticompetitive. For example, with the quadratic payoff function (13) exclusive dealing can have anticompetitive effects only when the dominant firm's market share is at least 65-70%, a value much higher than is often required for finding a dominant position.

A more nuanced approach is therefore necessary. One must distinguish between cases of strong dominance, where exclusive dealing may be presumed to be illegal (subject to possible efficiency defenses), and weak dominance, where the direct effects of exclusive dealing are procompetitive. In these latter cases, a rule of reason approach seems more appropriate. Plaintiffs could still mount a case based on the sacrifice/recoupment logic of the traditional theories, but defendants should have an easier time than in cases of strong dominance.

In terms of methodology, the paper provides tools for the analysis of optimal pricing when trade is non anonymous, as is normally the case in vertical relations. In principle, with non-anonymous trade sellers who possess market power could extract their profit efficiency, by means of fixed fees only. But many problems that are commonly perceived as real, such as for instance double marginalization, or royalty stacking in the licensing of complementary patents, would then disappear. To analyze those problems, economists often restrict firms to linear pricing. However, this restriction is clearly *ad hoc.* Getting rid of it, without throwing the baby out with the bathwater, requires modeling market imperfections of the sort discussed in the first part of this paper. Our treatment of such imperfections, and in particular our reduced-form model, shows that the problem is analytically more tractable than one might have thought. The approach proposed in this paper may therefore be applied to the analysis of other problems in the economics of vertical relations.

References

[1] Aghion, P. and P. Bolton (1987), "Contracts as a barrier to entry," *American Economic Review* 77, 388-401.

[2] Amir, R. and J.Y. Jin, (2001), "Cournot and Bertrand Equilibria Compared: Substitutability, Complementarity and Concavity," *International Journal of Industrial Organization*, 19, 303–317.

[3] Bernheim, D. and M. Whinston (1998), "Exclusive dealing," *Journal of Political Economy,* 106, 64-103.

[4] Calcagno C., C. Fumagalli and M. Motta (2016), *Monopolization: A Theory of Exclusionary Practices* Cambridge University Press, forthcoming.

[5] Calzolari, G. and V. Denicolò (2013), "Competition with exclusive contracts and market-share discounts," *American Economic Review* 103, 2384-2411.

[6] Calzolari, G. and V. Denicolò (2015), "Exclusive dealing and market dominance," *American Economic Review*, 105, 11, 3321-51.

[7] Della Vigna S. and U. Malmendier (2004), "Contract design and self-control: Theory and evidence," *Quarterly Journal of Economics*, 119, 353-402

[8] Hart O. and J. Tirole (1990), "Vertical integration and market foreclosure," Brooking papers: Microeconomics, 205-286.

[9] Matthewson, F. and R. Winter (1987), "The competitive effect of vertical agreements: Comment," *American Economic Review* 77, 1057-1062.

[10] O'Brien, D. and G. Shaffer (1997), "Nonlinear supply contracts, exclusive dealing, and equilibrium market foreclosure," *Journal of Economics and Management Strategy* 6, 755-785.

[11] Png I. and H. Wang (2010) "Buyer uncertainty and two-part pricing: Theory and applications," *Management Science*, 56, 334-342.

[12] Singh N. and X. Vives (1984), "Price and quantity competition in a differentiated duopoly," *RAND Journal of Economics*, 15, 546-554.

[13] Whinston, M. (2008), *Lectures on antitrust economics*, Cambridge, MA: MIT Press.

APPENDIX

The limiting case $\lambda \to \infty$. First of all, we prove the claim made in the text that with infinite risk aversion we get $\alpha_i = 0$, so that the fixed fee vanishes, and the profit is extracted by means of positive price-cost margins only. The first step in the proof is to show that

$$\lim_{\lambda \to \infty} \hat{\theta}_i = 0.$$

This follows directly from the participation constraint (3) by noting that if $\lim_{\lambda \to \infty} \hat{\theta}_i > 0$ the first term on the left-hand side would diverge to minus infinity. But this is impossible, as the remaining terms in (3) are finite.

Since $\upsilon(q_i) - q_i \upsilon_{q_i}(q_i) > 0$ as soon as $q_i > 0$, by the definition of $\hat{\theta}_i$ this immediately implies that

$$\lim_{\lambda \to \infty} F_i = 0.$$

In fact, since $\upsilon(q_i) - c_i q_i - (1 - \xi_i)\upsilon^{NE}(0) \geq 0$, F_1 vanishes only in the limit. We now show that this implies:

$$\lim_{\lambda \to \infty} \lambda \int_0^{\hat{\theta}_i} g(\theta) d\theta = \infty.$$

The proof is by contradiction. Suppose to the contrary that $\lim_{\lambda \to \infty} \lambda \int_0^{\hat{\theta}_i} g(\theta) d\theta$ is finite. Since

$$\int_0^{\hat{\theta}_i} \theta g(\theta) d\theta \leq \int_0^{\hat{\theta}_i} g(\theta) d\theta, \tag{A.3}$$

(this follows directly from our normalization $\int_0^{\bar{\theta}} \theta g(\theta) d\theta = 1$), this would imply that $\lim_{\lambda \to \infty} \lambda \int_0^{\hat{\theta}_i} \theta g(\theta) d\theta$ is also finite. But then it would follow from the participation constraint that $\lim_{\lambda \to \infty} F_i > 0$,a contradiction.

To complete the proof for this case, it remains to show that $\lim_{\lambda \to \infty} \alpha = 0$, i.e.

$$\lim_{\lambda \to \infty} \frac{1 + \lambda \int_0^{\hat{\theta}_i} \theta g(\theta) d\theta}{1 + \lambda \int_0^{\hat{\theta}_i} g(\theta) d\theta} = 0.$$

We have just shown that the denominator tends to infinity. As for the numerator, two cases are possible: either it converges to a finite limit, in which case the result is obvious, or

$$\lim_{\lambda\to\infty} \lambda \int_0^{\hat{\theta}} \theta g(\theta) d\theta = \infty,$$

in which case a straightforward application of l'Hopital rule and the fact that $\lim_{\lambda\to\infty} \hat{\theta}_i = 0$ yield

$$\lim_{\lambda\to\infty} \frac{1 + \lambda \int_0^{\hat{\theta}_i} \theta g(\theta) d\theta}{1 + \lambda \int_0^{\hat{\theta}_i} g(\theta) d\theta} = \lim_{\lambda\to\infty} \frac{\int_0^{\hat{\theta}_i} \theta g(\theta) d\theta}{\int_0^{\hat{\theta}_i} g(\theta) d\theta} = 0.$$

The remainder of this Appendix contains proofs of formal results omitted in the text.

Proof of Proposition 1. To show that $F_i > 0$, it suffices to note that $\upsilon(q_i) - q_i \upsilon_{q_i}(q_i)$ is always positive by the concavity of $\upsilon(q_i)$. Therefore, $F_i \leq 0$ would imply $\hat{\theta}_i = 0$. But then the profit would become $\Pi_i(q_i) = \upsilon(q_i) - c_i q_i - \upsilon^{NE}(0)$, which is maximized at $p_i = c_i$. Clearly, though, $p_i = c_i$ and $F_i \leq 0$ cannot be the optimal tariff.

To show that $p_i > c_i$, suppose to the contrary that $p_i = c_i$. Consider then a small increase $dp_i > 0$ in p_i and a corresponding decrease $q_i^{\mathrm{d}}(c_i) \times dp_i$ in F_i. In other words, the fixed payment F_i decreases by the same amount by which the average variable payment $p_i q_i$ increases. With this change in the price schedule, the firm's average profit by construction does not change. Since the total surplus $\upsilon(q_i) - c_i q_i$ is maximized at $p_i = c_i$, as small change in p_i has a second order effect. Therefore, the retailer's average profit, which is the difference between the average total surplus and the average profit of firm i, does not change. However, the retailer's profit has become less uncertain, so the participation constraint is now slack. This means that the fixed fee may actually be reduced by less than $q_i^{\mathrm{d}}(c_i) \times dp_i$, which makes the increase in the marginal price profitable. ■

Proof of Proposition 2. To show that $F_i \geq 0$, suppose to the contrary that $F_i < 0$. In this case, participation is guaranteed for all types θ. Now, consider type $\bar{\theta}$. Take a small decrease $dp_i < 0$ in p_i and a corresponding increase in F_i that leaves the profit extracted from retailer $\bar{\theta}$ unaffected. Since the profit extracted *via* the price cost margin is lower for types $\theta < \bar{\theta}$ than for type $\bar{\theta}$, this change increases the firm's total profit. This shows that $F_i < 0$ cannot be optimal.

To prove the second part of the proposition, we proceed as in the proof of Proposition 1. Thus, suppose to the contrary that $p_i = c_i$ and consider a small increase $dp_i > 0$ in p_i and a corresponding decrease in F_i equal to $dF_i = -\hat{\theta}_i \times q_i^{\mathrm{d}}(c_i) \times dp_i$. By construction, this change does not affect the marginal retailer $\hat{\theta}_i$, and hence the weights β_i and γ_i in the profit function (7). Furthermore, the change has a second-order impact on the first term of the profit function, but the impact on the second term (which is positive) is first order. This means that the increase in the marginal price is profitable. ■

Proof of Proposition 4. We start from the profitability of exclusive dealing. First of all, rewrite the profit as

$$\Pi_1 = \frac{1}{1+\mu}\left[\upsilon^M(q_1) - \upsilon^{NE}(0,p_2)\right] + \frac{\mu}{1+\mu} q_1 \upsilon^M_{q_1}(q_1)$$

with $M = E, NE$. Note that the solution is interior for $\mu = 0$; by continuity, this will be true also when μ is positive but small enough. The optimal quantity then is given by the first-order condition

$$\upsilon^M_{q_1}(q_1) + \frac{\mu}{1+\mu} q_1 \upsilon^M_{q_1 q_1}(q_1) = 0.$$

Under exclusive dealing, this formula immediately gives the equilibrium quantity q_1^E. Under common representation, the equilibrium is obtained in the way described in the preceding subsection. In any case, denote the equilibrium quantity by $q_1^M(\mu)$. The equilibrium profit is a function of μ

$$\Pi_1^M(\mu) = \frac{1}{1+\mu}\left\{\upsilon^M\left[q_1^M(\mu)\right] - \upsilon^{NE}(0,p_2^M)\right\} + \frac{\mu}{1+\mu} q_1^M(\mu)\upsilon^M_{q_1}\left[q_1^M(\mu)\right]$$

where p_2^M is the rival's equilibrium price. The formula holds both with and without exclusive contracts. Using Maclaurin's approximation, we get

$$\Pi_1^M(\mu) = \Pi_1^M(0) + \frac{d\Pi_1^M}{d\mu}\mu + \frac{1}{2}\frac{d^2\Pi_1^M}{d\mu^2}\mu^2 + o^3,$$

where o^3 denotes terms of order three or higher.

Now, at $\mu = 0$ and $c = c_{PPO}$ we have $q_2^M(\mu) = 0$ and $p_2^M = c$, both with and without exclusive contracts, and hence $\upsilon^E(q_1) = \upsilon^{NE}(q_1)$. It follows that $q_1^E(0) = q_1^{NE}(0)$, implying that

$$\Pi_1^{NE}(0) = \Pi_1^E(0).$$

Next, using the envelope theorem we get

$$\frac{d\Pi_1^M}{d\mu} = \frac{1}{(1+\mu)^2}\left\{q_1^M(\mu)\upsilon_{q_1}\left[q_1^M(\mu)\right] - \upsilon^M\left[q_1^M(\mu)\right] + \upsilon^{NE}(0,p_2^M) - (1+\mu)\, q_2^M(\mu)\frac{\partial p_2^M}{\partial \mu}\right\}$$

where $\frac{\partial p_2^M}{\partial \mu} = -\frac{1}{(1+\mu)^2} q_2^M(\mu) \upsilon_{q_2 q_2}\left[q_2^M(\mu)\right]$. Evaluating the derivative at $\mu = 0$ and keeping in mind that at $c = c_{PPO}$ we have $q_1^E(0) = q_1^{NE}(0)$, $p_2^{NE} = p_2^E = c$, and $q_2^M(\mu) = 0$, it follows that

$$\left.\frac{d\Pi_1^{NE}}{d\mu}\right|_{\mu=0} = \left.\frac{d\Pi_1^E}{d\mu}\right|_{\mu=0}.$$

We are thus left with the comparison of the second order effects. A glance at the above derivative confirms that the last two terms inside curly brackets are third order. To a second-order approximation, therefore, we have

$$\left.\frac{d^2\Pi_1^M}{d\mu^2}\right|_{\mu=0} = q_1^M(\mu)\upsilon_{q_1 q_1}^M\left[q_1^M(\mu)\right] \left.\frac{dq_1^M}{d\mu}\right|_{\mu=0}.$$

It follows that

$$\left.\frac{d^2\Pi_1^{NE}}{d\mu^2}\right|_{\mu=0} - \left.\frac{d^2\Pi_1^E}{d\mu^2}\right|_{\mu=0} = \left[\upsilon_{q_1 q_1}^E(q_1^E(0)) \left.\frac{dq_1^E}{d\mu}\right|_{\mu=0} - \upsilon_{q_1 q_1}^{NE}(q_1^{NE}(0), c) \left.\frac{dq_1^{NE}}{d\mu}\right|_{\mu=0}\right] q_1^M(0).$$

Next, notice that $\left.\frac{dq_1^M}{d\mu}\right|_{\mu=0}$ is the same both with and without exclusive contracts, as the first order-condition is the same at $\mu = 0$ and $c = c_{PPO}$. It is easy to see that $\left.\frac{dq_1^M}{d\mu}\right|_{\mu=0} < 0$. Therefore,

$$\left.\frac{d^2\Pi_1^{NE}}{d\mu^2}\right|_{\mu=0} - \left.\frac{d^2\Pi_1^E}{d\mu^2}\right|_{\mu=0}$$

has the same sign as

$$\upsilon_{q_1 q_1}^E(q_1^E(0)) - \upsilon_{q_1 q_1}^{NE}(q_1^{NE}(0), c) = \frac{\left[\upsilon_{q_1 q_2}(q_1^M(0), 0)\right]^2}{\upsilon_{q_2 q_2}(q_1^M(0), 0)} < 0.$$

It follows that

$$\left.\frac{d^2\Pi_1^E}{d\mu^2}\right|_{\mu=0} > \left.\frac{d^2\Pi_1^{NE}}{d\mu^2}\right|_{\mu=0}$$

and hence that $\Pi_1^E(\mu) > \Pi_1^{NE}(\mu)$ for μ arbitrarily small when $c = c_{PPO}$.

For social welfare, the logic of the proof is similar. Social welfare is simply

$$S^M = \upsilon^M(q_1),$$

so

$$S^M(\mu) = S^M(0) + \frac{dS^M}{d\mu}\mu + \frac{1}{2}\frac{d^2S^M}{d\mu^2}\mu^2 + o^3.$$

Clearly,

$$\frac{dS^M}{d\mu} = \upsilon_{q_1}^M(q_1^M)\frac{dq_1^M}{d\mu}$$

Proceeding as before, it is easy to confirm that at $c = c_{PPO}$ one has $S^E(0) = S^{NE}(0)$,

$$\left.\frac{dS^{NE}}{d\mu}\right|_{\mu=0} = \left.\frac{dS^{E}}{d\mu}\right|_{\mu=0},$$

and

$$\left.\frac{d^2S^{M}}{d\mu^2}\right|_{\mu=0} = \upsilon^M_{q_1q_1}\left[q_1^M(\mu)\right]\left(\left.\frac{dq_1^M}{d\mu}\right|_{\mu=0}\right)^2.$$

It follows that

$$\left.\frac{d^2S^{NE}}{d\mu^2}\right|_{\mu=0} - \left.\frac{d^2S^{E}}{d\mu^2}\right|_{\mu=0}$$

has the same sign as

$$\upsilon^{NE}_{q_1q_1}(q_1^{NE}(0)) - \upsilon^{E}_{q_1q_1}(q_1^{E}(0), c) = -\frac{\left[v_{q_1q_2}(q_1^M(0), 0)\right]^2}{v_{q_2q_2}(q_1^M(0), 0)} > 0,$$

meaning that exclusive dealing reduces social welfare. ■

Proof of Proposition 5. See the Online Appendix 1 at [address]. ■

Proof of Lemma 1. It is immediate to verify that any profitable deviation from the proposed candidate equilibrium would also be a profitable deviation from the original equilibrium. By definition, however, no such deviation exist, implying that the candidate equilibrium is, indeed, an equilibrium. ■

Proof of Proposition 6. *Necessity.* We must show that for any common representation equilibrium the non-exclusive tariffs (p_1, p_2, F_1, F_2) satisfy conditions (E_i), (NE_i) and (CS_i). The exclusive tariffs may be different from $\left(p_i^E(\kappa), F_i^E(\kappa)\right)$, but since these contracts are not signed in equilibrium, this does not affect the outcome.

If condition (E_1) does not hold, the dominant firm can profitably deviate from the original equilibrium by changing its non-exclusive tariff so as to make it less attractive for the retailer and offering the exclusive contract $\left(p_1^E(\kappa), F_1^E(\kappa)\right)$. The retailer would then sign the exclusive contract $\left(p_1^E(\kappa), F_1^E(\kappa)\right)$ since the rival's exclusive contract entail an expected utility not greater than κ (otherwise, the retailer would not have signed the non-exclusive contracts in the original equilibrium). But the exclusive contract $\left(p_1^E(\kappa), F_1^E(\kappa)\right)$ is more profitable than the equilibrium non-exclusive contract if condition (E_1) fails. The argument for (E_2) is similar, except that firm 2's deviation may have to slightly undercut the exclusive contract $\left(p_2^E(\kappa), F_2^E(\kappa)\right)$ to break a possible indifference.

If condition (NE_i) does not hold, a firm could profitably deviate by decreasing its marginal price. Finally, if condition (CS_i) does not hold, a firm could profitably deviate by increasing its marginal price. This completes the proof of necessity.

Sufficiency. We must show that any vector (p_1, p_2, F_1, F_2) that satisfies conditions (E_i), (NE_i) and (CS_i) is a common representation equilibrium. This requires that the retailer signs the non-exclusive contracts, and that no firm has any profitable deviation.

That the retailer signs the non-exclusive contracts is guaranteed by condition (16) and our tie-breaking rule.

There are three types of possible deviations: a firm could change its exclusive contract, its non-exclusive contract, or both. Consider first a deviation to a different exclusive contract. Since the rival's exclusive contract guarantees to the retailer an expected utility of κ, no exclusive contract could be signed if it does not entail at least the same expected utility. However, condition (E_i) says that the most profitable among the exclusive contracts that could be signed by the retailer is less profitable than the equilibrium non-exclusive contract. Thus, there is no profitable deviation to a different exclusive contract.

Second, a firm could deviate to a different non-exclusive contract. In particular, it could either increase or decrease its marginal price. However, if the deviating firm increases its marginal price, it must reduce the fixed fee; otherwise, the retailer would sign an exclusive contract. Therefore, an increase in the marginal price is feasible only if the original fixed fee is positive. But in this case condition (CS_i) says that the marginal price lies on the firm's best response curve. On the other hand, condition (NE_i) says that a decrease in the marginal price is unprofitable. Therefore, no firm can profitably deviate to a different non-exclusive contracts.

Finally, consider a deviation in which a firm changes simultaneously both the exclusive and the non-exclusive tariff. If the retailer still signs the non-exclusive contracts, the deviation is equivalent to a deviation in which only the non-exclusive contract is changed. The reason for this is that the retailer can still choose the rival's exclusive contract, and thus his reservation utility is at least κ. If instead the retailer chooses the new exclusive contract, this contract must give the retailer at least the same expected utility as the rival's exclusive contract, which has not changed. But then by condition (E_i) the new exclusive contract cannot be more profitable than the equilibrium non-exclusive contract. ■

Proof of Proposition 7. The proof is by contradiction. Suppose to the contrary that there exists a common representation equilibrium in which $(p_i, F_i) = (p_i^{NE}, F_i^{NE})$ for $i = 1, 2$. When $c \geq c_{\text{lim}}$, we already know that exclusive dealing is profitable for the dominant firm, which implies that (E_1) fails. Suppose then that $c < c_{\text{lim}}$. By the property of individual excludability, we have

$$\kappa = \max_{q_i \geq 0} \left[v(q_i, 0) - p_i^{NE} q_i - F_i^{NE} \right]$$

for $i = 1, 2$. But now suppose that the dominant firm offers the exclusive contract

$$p_1^E = p_1^{NE} \qquad \text{and} \qquad F_1^E = F_1^{NE},$$

and at the same time changes its non-exclusive tariff so as to make it less attractive for the retailer. The retailer will then sign the dominant firm's exclusive contract. But this is more profitable than the non exclusive contract if $c < c_{\text{lim}}$

(which implies that q_2^{NE} is positive, and hence $q_1^E > q_1^{NE}$). This means that condition (E_1) must be violated even if $c < c_{\lim}$. ■

Proof of Proposition 8. See the Online Appendix 2 at [address]. ■