Vertical limit pricing*

Aggey Semenov† and Julian Wright‡

Abstract

A new theory of limit pricing is provided which works through the vertical contract signed between an incumbent manufacturer and a retailer. We establish conditions under which the incumbent can obtain full monopoly profits, even if the potential entrant is more efficient. A key feature of the optimal vertical contract we describe is quantity discounting, typically involving three-part incremental-units or all-units tariffs, with a marginal wholesale price that is below the incumbent’s marginal cost for sufficiently large quantities.

1 Introduction

Existing theories of limit pricing and predation treat buyers as final consumers, focusing on the price charged by an incumbent firm to consumers, and whether this price is set low enough to keep out (or drive out) a rival. In contrast, in many important cases of limit pricing and predation, the incumbent is actually a manufacturer that offers a non-linear wholesale price schedule (i.e. a non-linear tariff) to downstream buyers. For example, in the ongoing dispute between Intel and AMD, the firms manufacture microprocessors that they sell to competing computer-makers like Dell and Hewlett-Packard. One of AMD’s complaints is that Intel offers computer-makers substantial discounts for large purchases through an all-units quantity discounting scheme, so that in some cases the price of incremental purchases to computer-makers is below Intel’s own marginal cost (sometimes being zero or even negative) and that “Intel’s practices exacerbate normal impediments

---

*We are grateful to Mark Armstrong, Jay Pil Choi, Patrick Rey and Jean Tirole for helpful comments, and to Yongchao Zhang for excellent research assistance. Any errors are our own.

†Department of Economics, National University of Singapore, AS2 Level 6, 1 Arts Link, Singapore 117570. Email: ccssa@nus.edu.sg.

‡Corresponding author. Department of Economics, National University of Singapore, AS2 Level 6, 1 Arts Link, Singapore 117570. Email jwright@nus.edu.sg.
to entry and expansion”.¹ Similarly, in the landmark predatory pricing case *Brooke Group Ltd. v. Brown & Williamson Tobacco Corp* (509 U.S. 209: 1993), the firms in question were manufacturers of cigarettes and the alleged predatory pricing related to volume discounts given for large wholesale purchases of generic cigarettes by distributors.² Such cases naturally raise the question of whether limit pricing and predation can still work in a vertical setting, and if so, how the mechanism behind limit pricing and predation might differ from that traditionally studied?

This paper offers an answer to these questions by providing a new theory of limit pricing, one which works through the vertical contract signed between an incumbent manufacturer and a retailer (i.e. its distributor). We call this vertical limit pricing.³ Unlike standard theories of limit pricing and predation, such as those based on signaling (as introduced by Milgrom and Roberts, 1982), the theory we propose does not rely on any asymmetric information between the incumbent and entrant. Despite this, in our theory the incumbent sets a low (wholesale) price, necessarily below its own marginal cost over some range of output. The result is that efficient entry is deterred.

Our theory is related to a substantial body of work that studies the commitment benefits of vertical contracts. A standard result in this literature is that manufacturers can soften price competition if they can commit to contracts with retailers in which wholesale prices are inflated above cost. Examples of papers in this line include Bonanno and Vickers (1988), Irmen (1998), Kühn (1997), Rey and Stiglitz (1988, 1995), Sklivas (1987), and Vickers (1985). We explore a previously overlooked implication of the commitment effects of vertical contracts, which is to deter entry. An incumbent manufacturer writes a vertical contract with a retailer in which it commits to provide goods at a low cost, in fact, below its marginal cost for purchases beyond a certain level. In equilibrium, entry is deterred. Profit is recovered from the retailer either through a fixed fee or high initial wholesale prices. We establish conditions under which the incumbent can obtain full monopoly profits, even if the rival is more efficient, even if entry costs are trivial, both in cases in which the rival’s product is identical and where it is somewhat differentiated, and even if the incumbent’s vertical contract is only observed with a small probability.

³The theory can also be applied to the case of predation, if the incumbent manufacturer uses a vertical contract to induce the exit of a rival where the rival needs to incur additional fixed costs to stay in the market. This might be called vertical predation.
A key feature of the optimal vertical contract we describe is quantity discounting or declining *marginal* wholesale prices. We establish that both of the common forms of quantity discounting used in practice, incremental-units or all-units quantity discounting (Munson and Rosenblatt, 1998) are optimal. For low levels of purchases, the retailer purchases at a wholesale price set above the incumbent’s marginal cost, thereby providing a way for the manufacturer to extract the retailer’s profit (alternatively, a fixed fee can be used for this purpose). For purchases in some intermediate range, the retailer purchases at a wholesale price set equal to the incumbent’s marginal cost, thereby ensuring the retailer sets the correct monopoly price when it is indeed a monopolist. For purchases beyond some yet higher level, the retailer purchases at a wholesale price set below the incumbent’s marginal cost, thereby ensuring that in the face of competition, the retailer will want to compete aggressively, in such a way that the rival will not want to enter. Three-part tariffs are therefore the simplest optimal tariffs.

We consider both the case in which firms are homogenous price competitors and the case in which the goods are imperfect substitutes (competing in prices or quantities). In both cases, there are situations where to deter entry the incumbent must give its retailer a wholesale schedule so that the retailer is willing to price at a point where marginal revenue is negative. In the absence of entry, the incumbent’s retailer could take advantage of this fact by disposing of some units to move back up its monopoly revenue function. To prevent this, the incumbent must leave the retailer with some rent in equilibrium. We call this the retailer’s “disposal-rent”. In case goods are imperfect substitutes, a different type of rent may also be required in order that the incumbent’s retailer is willing to choose the monopoly quantity in equilibrium rather than the out-of-equilibrium entry deterring quantity, which we call the retailer’s “incentive-rent”. Furthermore, if the incumbent can also make use of an upfront fee, it can extract these rents when the retailer signs the contract, thereby still obtaining its full monopoly profit.

Quantity discounting naturally arises with vertical limit pricing, reflecting the underlying concavity of a monopolist’s revenue function. As such, our theory is complementary to the theory of Kühn (1997), who provides a related explanation for such quantity discounting. His theory is also based on commitment effects through vertical contracts, but in a context where retailers compete in quantities and there is demand uncertainty. In his symmetric environment, concave tariffs are used by each manufacturer to commit their retailers to be more aggressive competitors, since retailers will face lower marginal costs if they sell a lot. Without the possibility of entry deterrence, quantity discounting ends up hurting manufacturers in equilib-
rium. However, in the face of price competition, he shows the opposite result holds. To soften competition, manufacturers will offer convex contracts with increasing marginal wholesale prices. Thus, in contrast to Kühn’s results, our results can explain quantity discounting even if firms compete in prices, and moreover, even if interbrand competition takes the homogenous Bertrand form (in which no such softening of competition is possible).

Vertical limit pricing relates to the literature on exclusive dealing. When the incumbent can commit to a wholesale pricing schedule as part of its initial exclusive deal and buyers are downstream competitors, the setting is quite similar to that in our paper. This has been considered by Simpson and Wickelgren (2001), Stefanadis (1998), Yong (1999) and Appendix B of Fumagalli and Motta (2006). For instance, Fumagalli and Motta show the incumbent manufacturer will commit to a low wholesale price (to deter entry), extracting the surplus enjoyed by retail buyers paying this low wholesale price through an upfront fee which it receives when the exclusive deal is signed. This enables the incumbent to deter entry. Our results imply the incumbent can do better, often obtaining the full monopoly profit, with a more sophisticated contract involving quantity discounting but which often does not require an upfront fee or exclusionary terms.

Our study of vertical limit pricing is somewhat less related to a recent line of literature exploring the ability of bundled rebates, market-share discounts and all-units discounts to have exclusionary effects. As Tom et al. (2000, p.615) note “The traditional analysis governing exclusive dealing arrangements has focused on a manufacturer’s requirement that its distributors deal exclusively with it. In recent years, however, some manufacturers have begun to use subtler arrangements in which incentives replace requirements ...”. Our paper differs in two respects. First, much of this literature has focused on contracts with end-users (or a retailer representing end-users’ interests) which is not the case in our theory. Second, our paper provides a predatory-type purpose for quantity discounting. Whether or not volume discounts also include exclusivity provisions, their purpose in our theory is to commit the incumbent to price below cost where this is necessary to drive out the (potential) rival. They are not simply replicating exclusive deals that are designed to prevent distributors sourcing inputs from competing manufacturers.

From a policy viewpoint, our theory provides a particular setting which supports the use of a predatory pricing standard for dealing with wholesale price discounts in

---

4In other theories in which the incumbent uses exclusive (or partially exclusive) contracts as a barrier to entry (e.g. Aghion and Bolton, 1987, Rasmusen et al., 1991, and Segal and Whinston, 2000), contracts are signed directly with final consumers and the mechanisms at work do not apply in the setting we consider.
single-product cases. In our theory, marginal wholesale prices must fall below a firm’s own marginal cost for sufficiently large quantities in order to deter entry. Where there are no efficiency justifications for below-cost wholesale prices, such contracts are therefore anticompetitive. Forcing the incumbent to raise its marginal wholesale price to be no less than marginal cost will encourage efficient entry and increase welfare in our setting. More generally, vertical limit pricing provides a rationale for competition authorities to be concerned about vertical contracts which involve declining marginal wholesale prices which become very low for high quantities.

The rest of the paper proceeds as follows. Section 2 presents a benchmark model in which firms are homogeneous price competitors. Our main results are presented in Section 3. Section 4 discusses a variety of extensions including to simultaneous or unobservable contract offers, the role of upfront fees, some implications of allowing limited renegotiation possibilities, and to the case with imperfect substitutes (including quantity competition). Finally, section 5 concludes with some directions for future research.

2 Benchmark model

Market demand is denoted $Q(P)$ where $P$ is the market price. $Q(P)$ is assumed to be non-negative, continuously differentiable and strictly decreasing in price over the relevant price interval (i.e. the interval of prices from the lowest non-negative price at which demand is defined up to the point where demand just becomes zero, if indeed such a maximum price exists) and is zero thereafter. The inverse demand function is denoted $P(Q)$ which is therefore non-negative, continuously differentiable and strictly decreasing in quantity over the corresponding interval of quantities. We restrict to this interval of quantities throughout the following analysis.

Define the revenue function $R(Q) = P(Q)Q$. We assume that $R(Q)$ is strictly concave in $Q$. To allow for the possibility that revenue is decreasing in output (as is the case with linear demand, for instance), we also define the following modified revenue function

$$\tilde{R}(Q) = \max_{0 \leq q \leq Q} R(q),$$

which is non-decreasing in $Q$. The monopoly price given any constant marginal cost $w$ (assumed to be less than $P(0)$) is denoted

$$P_M(w) = \arg \max_P (P - w)Q(P).$$

For notational convenience, define $Q_M(w) = Q(P_M(w))$. 
To start with we focus on a model in which firms sell an identical good and set prices (i.e. homogenous Bertrand competition). The extension to imperfect competition will be studied in Section 4. There are two upstream firms, i.e. manufacturers. There is an incumbent upstream firm, which we will denote as $U_1$, which faces constant marginal costs of $c_1$. The incumbent’s monopoly price and quantity are defined as $P_M = P_M(c_1)$ and $Q_M = Q_M(c_1)$, with corresponding monopoly profit $\Pi_M = (P_M - c_1)Q_M$. Assume $P(0) > c_1$ which ensures that if the incumbent is a monopolist it will produce a positive output (and so can obtain a positive profit). A potential upstream entrant, denoted $U_2$, faces lower marginal costs of $c_2 < c_1$ but some fixed cost of entry $F$ which satisfies

$$0 \leq F < (c_1 - c_2)Q(c_1).$$

$U_2$ is assumed to only enter if it makes positive profit (which is the reason we can include the case in which fixed costs are assumed to be zero). With these assumptions it will always be profitable for $U_2$ to enter if it competes directly with $U_1$. Define the break-even price for $U_2$ to be (the lowest value of) $p_2$ that solves

$$(p_2 - c_2)Q(p_2) = F,$$

which we denote as $P$. This $P$ exists and satisfies $c_2 \leq P < c_1$ given the strict concavity of the revenue function and since (2) implies $(P - c_2)Q(P) > F$ when $P = c_1$ and $(P - c_2)Q(P) \leq F$ when $P = c_2 < c_1$.

Initially, we assume $R'(Q(P)) \geq 0$ so that the market revenue function is non-decreasing at $U_2$’s break-even price. This condition holds for standard demand specifications such as constant elasticity and logit demand where the monopolist’s revenue function $R(Q)$ is always increasing in $Q$. For demand specifications such as linear and exponential, where the revenue function can decrease, the condition requires the price elasticity of the market demand $Q(P)$ is greater than unity (in magnitude) at $Q(P)$. We will subsequently discuss what happens without this condition.

To proceed, we employ a standard vertical chains structure, following Rey and Stiglitz (1988). Each upstream firm (or manufacturer) sells through a downstream firm (or retailer) denoted respectively $D_1$ and $D_2$. Retailers set final prices. They face no costs other than those determined by the manufacturers’ wholesale tariffs. Whichever retailer sets the lower price obtains the entire market demand at that

---

5This particular vertical structure is not crucial for our results. The analysis would still apply if more than one retailer competed to be a manufacturer’s distributor or if retailers could sign with both manufacturers, provided $U_1$ can still sign a binding contract with at least one of the retailers before $U_2$ has to decide whether to incur its fixed costs of entry.
price. Tie-breaking rules are that if retailers set the same price, one of the retailers will obtain the entire market, and this retailer will be determined so as to avoid open set problems in defining equilibria (i.e. in standard cases, the one facing a lower marginal cost).

The timing is that in stage 1, $U_1$ offers a contract to $D_1$, who accepts or not. If $D_1$ rejects the offer both, $U_1$ and $D_1$ earn nothing.\footnote{Equivalently, $U_1$ could supply the market directly in the case its contract is rejected. Given its higher costs, $U_1$ would not be able to compete and $U_1$ and $D_1$ will not obtain any profit.} Then in stage 2, after observing $U_1$’s contract and $D_1$’s decision, $U_2$ can decide whether to enter the market (incurring the cost $F$), and if it does it offers a contract to $D_2$ who accepts or not. If $D_2$ rejects the offer, $U_2$ and $D_2$ earn nothing.\footnote{Alternatively, we can allow that $U_2$ can supply the market directly in this case, in which case $U_2$ is no better off than had it been able to sell through $D_2$.} In stage 3, active retailers then set prices and pay wholesalers based on the amount they order from them. Upstream firms can only observe how much their downstream firm buys from them, but not anything else. In particular, the contract cannot depend on the amount the downstream firm sells or the price it sets (as in revenue or profit sharing contracts), or on the rival’s decisions or outcomes. This setup also assumes downstream firms only observe a rival’s contract after it has been signed. These informational assumptions are standard in the literature (see Kühn, 1997 and Rey and Stiglitz, 1995).

This informational structure implies $U_i$’s tariff $T_i(q_i)$ to $D_i$ depends only on $q_i$, the amount purchased by downstream firm $D_i$ ($i = 1, 2$). For the most part we restrict contracts further by considering tariff schedules without upfront fees, either a fee paid prior to stage 3 (e.g. a franchise fee) or a fee paid in stage 3 even if $D_1$ does not purchase any units. Such upfront fees make it substantially easier for an incumbent to obtain monopoly profits through vertical limit pricing since they provide a further first-mover advantage to the incumbent. Large upfront fees may also not be feasible or efficient in practice, for reasons outside of this model. Taken literally, they may imply a very large upfront payment so that the manufacturer can extract the ongoing monopoly profits of the retailer. Such payments may be infeasible in practice due to credit constraints or may be inefficient if they put too much risk on the retailer. Later, we will explain how our results change allowing for such upfront fees.\footnote{Note the absence of a fixed fee in stage 3 is implied if $D_1$ can walk away from any contract which it finds unprofitable ex-post (i.e. after observing firm 2’s contract) by not buying anything from $U_1$ and not paying anything to $U_1$. In other words, while we assume that the upstream firm can commit to its contract, unlike Kühn we do not require the downstream firm can do the same.}

Without loss of generality, tariffs in the allowed class are defined as $T_i(q_i) = T_i I_{q_i > 0} + W_i(q_i)$ on $[0, \infty)$, where $T_i$ is a possible “fixed fee” that is only paid if $D_i$
makes a positive purchase from $U_i$ (i.e. $I_{q_i > 0}$ is an indicator variable which is 0 if $q_i = 0$ and 1 if $q_i > 0$) and $W_i(q_i)$ corresponds to some wholesale price schedule with $W_i(0) \leq 0$. Note $D_i$ is allowed to buy more units than it sells and freely dispose of the extra units, to take advantage of any wholesale schedule with decreasing tariffs. This constrains the type of contracts that can be profitably offered. Define the class of such tariffs as $\mathcal{T}$.

Even tariffs within the class $\mathcal{T}$ may be far more complicated than can reasonably be used in practice. An important focus of our analysis will be to see whether the incumbent’s optimal contract can be implemented with plausible or realistic tariff schedules. For instance, one possible concern arises if the optimal contract involves very high tariffs for certain quantities, implying the incumbent is effectively engaged in quantity forcing, which could be illegal. (Note we have already ruled out direct quantity forcing in which the upstream firms restrict the range of quantities that can be chosen by downstream firms). Another concern might arise if the contract is required to be very complicated. In practice, tariff schedules are not presented as non-linear functions, but rather schedules of wholesale prices that apply for different quantities purchased, possibly with a fixed fee.\footnote{Munson and Rosenblatt (1998) survey contracts offered by manufacturers or received by wholesalers and retailers and find (p.364) “None of the participants have seen continuous schedules in practice.” Simplicity of the contract was a key concern expressed by interviewees.} This is the class of tariffs that are piece-wise linear.

The literature on operations management (see Munson and Rosenblatt, 1998) identifies two types of piece-wise linear tariffs that are studied by researchers and widely used by industry to quantity discount.\footnote{According to the survey of Munson and Rosenblatt, incremental-units quantity discounting was used by 37% of their sampled firms. Even more commonly used was all-units quantity discounting (used by 95% of sampled firms). In addition to the use of fixed fees by 29% of the interviewees, these were the only forms of quantity discounting identified by the authors. Kolay et al. (2005) also note all-units quantity discounting is widely used in intermediate-goods markets, with the list of companies known to have used them including Coca-Cola, British Airways, and Michelin.} The first type of tariff is associated with incremental-units quantity discounting, which is a block declining tariff, in which the marginal wholesale prices declines at each increment. More generally, allowing for the possibility of a fixed fee $T_i$ and not requiring quantity discounting, this is the case where the wholesale schedule $W_i(q_i)$ is a continuous piece-wise linear function. To be more precise, define $W_i^{(n)}(q) = W_i^{(n)}(q; w_1, ..., w_{n-1}, S_1, ..., S_{n-2})$ as

$$W_i^{(n)}(q) = \begin{cases} W_i(0) & \text{if } q = 0 \\ W_i(q) & \text{if } q > 0 \end{cases}$$
Then an $n$-part incremental-units tariff is either $T_I^{(n)} (q) = T_I I_{q > 0} + W_I^{(n-1)} (q)$ in the case a positive fixed fee $T_I > 0$ is used or $T_I^{(n)} (q) = W_I^{(n)} (q)$ in the case of no fixed fee (i.e. $T_I = 0$). The class of $n$-part incremental-units tariff is denoted $T_I^{(n)}$. In general, we will refer to the class of contracts which belong to some $T_I^{(n)}$ as incremental-units contracts, denoted $T_I$. Then $T_I = \cup_{n=1}^{\infty} T_I^{(n)}$.

The second type of tariff is associated with all-units quantity discounting, in which wholesale prices decline at each increment, but the lower wholesale price applies to all units purchased rather than just marginal units. More generally, again allowing for the possibility of a fixed fee and no longer requiring quantity discounting, this is the case where the wholesale schedule is a special type of piecewise linear function with discontinuities associated with the different price-breaks. More precisely, define $W_A^{(n)} (q) = W_A^{(n)} (q; w_1, \ldots, w_{n-1}, S_1, \ldots, S_{n-2})$ as

$$
W_A^{(n)} (q) = \begin{cases}
    w_1 q & \text{if } 0 < q \leq S_1, \\
    w_2 q & \text{if } S_1 < q \leq S_2, \\
    \vdots & \vdots \\
    w_n q & \text{if } S_{n-1} < q.
\end{cases}
$$

Then an $n$-part all-units tariff is either $T_A^{(n)} (q) = T_A I_{q > 0} + W_A^{(n-1)} (q)$ in the case a fixed fee $T_A$ is used or $T_A^{(n)} (q) = W_A^{(n)} (q)$ in the case it is not. The class of $n$-part all-units tariff is denoted $T_A^{(n)}$. In general, we will refer to the class of contracts which belong to some $T_A^{(n)}$ as all-units contracts, denoted $T_A$. Then $T_A = \cup_{n=1}^{\infty} T_A^{(n)}$.

In what follows we will drop the subscript which characterizes the type of the tariff. Instead, we will indicate the set to which this tariff belongs. For example, the notation $T_1 (q) = T_1 I_{q > 0} + W_1 (q) \in T_1^{(n)}$ means that the tariff $T_1 (q)$ is a $n$-part incremental-units tariff.

In the next section, in addition to characterizing optimal contracts, we will be interested in whether incremental-units tariffs and/or all-units tariffs can achieve optimality, and if they can, the properties of the simplest of such tariffs (those where $n$, the number of parts, is lowest). These questions will be addressed both with and without the use of a fixed fee.
3 Results

In this section we characterize optimal contracts. Different restrictions on the upstream firms’ tariffs will be considered. Actually, these restrictions are only required on the incumbent’s tariff. The results would remain the same if the rival was allowed to use any tariff from the general class $T$ or sell directly to consumers. However, to avoid giving the false impression that the results depend on the rival having access to a wider tariff class, all our results will be stated assuming both upstream firms have the same restrictions on the tariffs they can write.

Before considering non-linear tariffs, it is useful to point out that by restricting to linear tariffs (which in the terminology of Section 2 is just a one-part tariff without a fixed fee), the incumbent manufacturer cannot prevent entry. Trivially, we have that:

**Proposition 1** If upstream firms are restricted to using linear tariffs, then the incumbent never makes any sales or profit. The potential entrant always enters.

**Proof.** So as to cover its costs for any level of sales, $U_1$ must set its wholesale price at or above $c_1$. $U_2$ can always undercut with a slightly lower wholesale price (if necessary), so that given (2), $D_2$ will profitably take the whole market. $U_1$ ends up with no sales or profit. ■

This result motivates the use of non-linear tariffs. Naturally, one wonders whether simple two-part tariffs can do better. Here we consider two-part tariffs belonging to the classes $T_I$ or $T_A$. This includes a standard two-part tariff in which $T_i > 0$ and $W_i(q_i)$ is linear, or tariffs in which there is no fixed fee but two different regions with different wholesale prices applying. As the next proposition establishes, such tariffs, while sometimes enabling entry deterrence, are never optimal.

**Proposition 2** If upstream firms are restricted to using a two-part tariff belonging to the class $T_I$ or $T_A$, then the incumbent can profitably take the whole market by setting a wholesale price below its own cost (so that the potential entrant stays out) but its profit is strictly less than monopoly profit.

**Proof.** We consider each of the three possible forms of tariffs in turn.

(i) Consider first a two-part tariff belonging to the classes $T_I$ or $T_A$ in which a fixed fee is used. This is a standard two-part tariff. The best $U_1$ can do by offering a two-part tariff of the form $T_1(q_1) = T_1I_{q_1 > 0} + w_1q_1$ is to set $T_1 = (P - w_1)Q(P)$ and $w_1 < P < c_1$. With this offer, $D_1$ will be willing to price down to $P$ to take the whole market, since by doing so it will sell $Q(P)$ units and obtain a profit of
 Observing this offer, $U_2$ will not enter since it cannot make a positive profit given $D_2$ can only sell units at a price below $P$, its break-even price. Given $U_2$ will not enter, $U_1'$s profit is

$$
\Pi_1 = (P - w_1) Q(P) - (c_1 - w_1) Q_M(w_1)
$$

and $D_1'$s profit is

$$
\pi_1 = (P_M(w_1) - w_1) Q_M(w_1) - (P - w_1) Q(P).
$$

Clearly, $\pi_1 \geq 0$ since by definition $P_M(w_1)$ maximizes $(p - w_1) Q(p)$. For $U_1$ to deter entry requires there exists $0 \leq w_1 < P$ with $\Pi_1 > 0$. If such a $w_1$ exists, then we have $\Pi_1 \leq \Pi_1 + \pi_1 = (P_M(w_1) - c_1) Q_M(w_1) < (P_M - c_1) Q_M$ since $P_M$ maximizes $(p - c_1) Q(p)$. That is, $U_1'$s profit is strictly less than monopoly profit. (An example of this case is $Q(p) = 1 - p$, $c_1 = 12/20$, $P = 11/20$, $w_1 = 7/20$. Figure 1(a) illustrates the corresponding tariff). If such a $w_1$ does not exist, then $U_1$ cannot deter entry with such a two-part tariff. (For instance, when $c_1$ is increased to 13/20 in the previous example.)

(ii) Consider a two-part tariff belonging to the class $\mathcal{T}_i$ without a fixed fee. Denote the first wholesale price $w_1$ which applies to purchases within $0 < q_1 < S_1$ and the second wholesale price $w_2$ which applies to marginal purchases when $q_1 \geq S_1$. First, consider the contract designed so $D_1$ chooses some $q_1 < S_1$ in equilibrium. Then the best $U_1$ can do is to set $w_1 = \arg \max_w (w - c_1) Q_M(w) > c_1$. To deter entry, set $w_2 = R'(Q(P))$ and $S_1$ to solve $w_1 S_1 + w_2 (Q(P) - S_1) = PQ(P)$. This ensures that $D_1$ is willing to price down to $P$ if required. As a result, $D_1$ will be a monopolist. Since $w_1$ is chosen so that $Q_M(w_1) > 0$, $D_1$ can always choose a price sufficiently close to $w_1$ such that $R(q_1) > W_1(q_1)$ for some positive quantity. This ensures $D_1$ will choose to sell a positive quantity. Moreover, $D_1'$s choice of $q_1$ is less than $S_1$ since $W_1(q_1) \geq R(q_1)$ for $q_1 \geq S_1$ (this follows given $W_1(q_1) = R(q_1)$ at $Q(P)$, $W_1(q_1)$ is linear for $q_1 \geq S_1$, and $R(q_1)$ is strictly concave). Since $D_1$ chooses $q_1 < S_1$, this means $U_1$ obtains the positive profit $\max_w (w - c_1) Q_M(w)$, which is less than the full monopoly profit $\max_w (w - c_1) Q(w)$. Figure 1(b) illustrates.

Similarly, $U_1$ will also not be able to obtain monopoly profit if it tries to deter entry with a contract designed such that as a monopolist $D_1$ chooses $q_1 \geq S_1$. To deter entry we require $w_1 S_1 + w_2 (Q(P) - S_1) \leq PQ(P)$. This implies $U_1'$s profit

$$
(w_1 - c_1) S_1 + (w_2 - c_1) (Q_M(w_2) - S_1) \leq (P - w_2) Q(P) + (w_2 - c_1) Q_M(w_2).
$$

If $U_1$ tries to induce $D_2$ to choose the monopoly price, it must set $w_2 = c_1$, which given $P < c_1$ implies it will make a loss.
(iii) Consider a two-part tariff belonging to the class $T_A$ without a fixed fee. Denote the first wholesale price $w_1$ which applies when $0 < q_1 < S_1$ and the second wholesale price $w_2$ which applies to all units when $q_1 \geq S_1$. $U_1$ can always profitably deter entry obtaining the same profit as in (ii). Set $S_1 = Q(P)$, $w_1 = \arg \max_w (w - c_1) Q_M(w) > c_1$ and $w_2 = P$. Then by construction, $D_1$ will be willing to price down to $P$ to take the whole market. As a result, $D_1$ will be a monopolist and will choose the same price as in (ii). Figure 1(c) illustrates. Note this contract is the best $U_1$ can do. Any contract designed such that as a monopolist $D_1$ chooses $q_1 \geq S_1$ can be ruled out since to deter entry it would require $w_2 \leq P$ which would imply a loss for $U_1$.

Proposition 2 shows $U_1$ may be able to use a two-part tariff contract with $D_1$ to prevent a more efficient firm entering. Such a contract involves using a wholesale price that is lower than $U_1$’s marginal cost. This illustrates, in the simplest possible setting, the ability of non-linear vertical contracts to be used by an incumbent to prevent entry, which we refer to as vertical limit pricing. This relies on the commitment of the incumbent to make its downstream firm a tough competitor in the face of competition, which it does through a vertical contract. Pricing below cost at the wholesale level can be profitable since (i) by so doing, its retailer can deter entry, thereby remaining a monopolist so it can sustain a high retail price and (ii) some of the profits associated with this high retail price can be recovered through a fixed fee. Thus, both instruments of the two-part tariff (the low wholesale price and the positive fixed fee) are needed for vertical limit pricing to work.

As proposition 2 established, the incumbent has to give up some part of its profit.
in order to deter entry. For instance, with linear demand, it can easily be established that an upper bound on the incumbent’s profit from using vertical limit pricing with two-part tariffs from the classes $T_I$ or $T_A$ is one half of its normal monopoly profit. This suggests the incumbent has strong incentives to make use of more sophisticated tariffs to deter entry, which is the case we consider next.

Once we allow for more complicated tariffs, even say the class of piece-wise linear tariffs, the dimensionality of the contract space becomes large relative to the problem that the incumbent solves, so not surprisingly the optimal contract is not uniquely determined (see Kühn, 1997 for a more general discussion of this type of problem). However, it turns out we can still say something useful about the incumbent’s optimal tariff. We start by showing that the optimal tariff within the general class $T$ will always allow the incumbent to obtain its full monopoly profit.

**Proposition 3** When upstream firms can set general non-linear tariff functions within the class $T$, the incumbent will obtain full monopoly profits, deterring entry in the process.

**Proof.** Consider $U_1$’s tariff $T_1(q_1)$ from the contract space $T$. First note that $T_1(q_1) \geq 0$ for all $q_1$ since otherwise there will be some output at which $D_1$ obtains positive profit through a subsidy from $U_1$ which it will strictly prefer to the equilibrium, given in equilibrium $U_1$ has to extract the full monopoly profit from $D_1$. This implies $T_1(0) = 0$ (since $T_1(0) > 0$ has already been ruled out).

To allow $U_1$ to obtain monopoly profits (i.e. its optimal outcome) we require that if $D_1$ does not face competition from $D_2$, then $D_1$ should choose to price at $P_M$ and buy the monopoly output $Q_M$ from $U_1$. This incentive constraint requires $D_1$ be better off (or no worse off) buying the monopoly output compared to any other positive output level. Formally, we require

$$ R(Q_M) - T_1 - W_1(Q_M) \geq \tilde{R}(q_1) - T_1 - W_1(q_1), \quad (6) $$

for all $q_1 > 0$.\footnote{The use of the modified revenue function $\tilde{R}(q)$ arises since we allow for the possibility of free disposal.} Second, $U_1$ should recover the full monopoly profits in equilibrium but leave $D_1$ willing to participate (since it can always choose to buy nothing and pay nothing). This participation constraint requires

$$ T_1 = R(Q_M) - W_1(Q_M). \quad (7) $$

Third, we require that if $D_2$ enters and sets a price of $P$ (so $U_2$ can just cover its costs)
costs of entry), $D_1$ will be willing to price so as to take the whole market. By offering such a contract, $U_1$ ensures that $U_2$ will never want to enter in stage 2. Suppose $D_1$ faces a rival pricing at $P$. If $D_1$ sets a higher price than $P$, it will not sell anything, obtaining a payoff of zero. Alternatively, $D_1$ can set a price which is the same or lower than $D_2$’s price and take the whole market. This will imply $D_1$ sells $q_1 \geq Q(P)$ units. Given it chooses $q_1$ optimally (through its choice of $p_1$), $D_1$ will get a profit of $R(q_1) - T_1 - W(q)$ for any $q \geq q_1$, where the number of units it offers for sale $q_1$ may be less than the number it purchases $q$ due to free disposal. This “minimal deterrence” constraint requires

$$\max_{q \geq Q(P)} \left( R(q_1) - T_1 - \min_{q \geq q_1} W(q) \right) \geq 0,$$

which ensures $D_1$ does at least as well with this option as it can setting a higher price in which it sells nothing.

Suppose $U_1$ offers a tariff with the following features: (i) it fixes any arbitrary $T_1 \geq 0$; (ii) it sets $W_1(Q_M) = R(Q_M) - T_1$; (iii) it sets $W_1(Q(P)) = R(Q(P)) - T_1$; and (iv) it sets $W_1(q_1)$ sufficiently high for any other $q_1$ (for example, it is sufficient to set $W_1(q_1) > \bar{R}(q_1) - T_1$ for any other $q_1$, which prevents $D_1$ from doing better with a different output when it is a monopolist). Substituting these conditions into the above constraints, it is clear (6)-(8) hold given the revenue function is assumed to be non-decreasing at $Q(P)$. Two examples of such tariffs are illustrated in figure 2.

Given the construction of any such tariff, $U_2$ cannot profitably enter. For $U_2$ to want to enter it must expect $D_2$ to price above $P$ in stage 3 (so it can expect

\footnote{In fact, even if it attracts some demand at this price, this will be sufficient to deliver the result.}
to make a profit from entry). However, if this were the case then given the above tariff, \( D_1 \) would profitably undercut to take the whole market. For this reason, \( U_2 \) will stay out. Facing the contract, \( D_1 \) cannot do better than set the monopoly price \( P_M \), generating monopoly profits of \( (P_M - c_1) Q_M \), which \( U_1 \) extracts through its wholesale schedule and/or initial fixed fee. Thus, we have that the incumbent can deter entry and obtain monopoly profit.

The optimal tariff in Proposition 3 can work regardless of whether the fixed fee is positive or zero. What is important is that as a monopolist, \( D_1 \) chooses the correct monopoly price from \( U_1 \)'s perspective (the optimal pricing constraint), and that facing a rival, it will be willing to undercut whenever the rival prices at a level which ensures \( U_2 \) can cover its costs (the minimal deterrence constraint). Even with these constraints, there is considerable redundancy in any optimal tariff within the class \( T \), since there are some quantities for which the tariff's only purpose is to avoid \( D_1 \) wanting to deviate from one of these two situations. This can always be achieved with any sufficiently high tariffs. Figure 2 illustrates two somewhat generic optimal tariffs. As illustrated in figure 2, the key properties for optimality are \( T_1 (0) = R (0) \), \( T_1 (Q_M) = R (Q_M) \) and \( T_1 (q_1) = R (q_1) \) for some \( q_1 \geq Q (P) \), with \( T_1 (q_1) \geq \tilde{R} (q_1) \) elsewhere.\(^{13} \)

Next we wish to see whether tariffs within the more reasonable classes \( T_I \) and \( T_A \) can achieve optimality for \( U_1 \), and if so, what are the properties of the simplest such tariffs. First, considering the class \( T_I \), we find optimality can indeed be achieved, and the simplest such tariff is a three-part block declining tariff, in which the marginal wholesale prices declines at each increment.\(^{14} \) The tariff is concave, exhibiting “incremental-units” quantity discounting.

**Proposition 4** The incumbent can obtain full monopoly profit, deterring entry in the process, by using a three-part block declining tariff in which \( W_1 (q_1) \) is continuous and concave. This is the simplest tariff within the class \( T_I \) that allows the incumbent to extract the full monopoly profit. The lowest wholesale price in the tariff is below the incumbent’s marginal cost.

**Proof.** The proof is by construction. We will start with the case in which a fixed fee is used, so that \( T_1 > 0 \). It is helpful to start with this case since it is the

\(^{13} \)To give \( D_1 \) a strict incentive to choose the monopoly output when there is no entry, or to choose a high output in the face of entry, the optimal tariffs characterized in Proposition 3 can always be approximated arbitrarily closely by slightly lower tariffs which leave \( D_1 \) with a positive profit in each case, but higher profit at \( Q_M \) than at \( Q (P) \) when \( D_1 \) is a monopolist. Similar approximations can be applied to the results throughout the paper.

\(^{14} \)In the following two propositions, \( U_2 \) is still assumed to be able to choose tariffs from the more general class \( T \). The same results also apply if it is restricted to the same (more restrictive) class of tariffs as \( U_1 \).
limiting case of the equivalent tariff in which no fixed fee is used, which we will introduce in the proof below. Also a fixed fee is necessary if the demand function is such that $P(0)$ is undefined, as it is with logit or constant elasticity demand. In such cases, $R'(0)$ is arbitrarily high, so it is not possible to simultaneously satisfy $T_1(0) = 0$ and $T_1(q_1) > R(q_1)$ as $q_1 \to 0$ with a piece-wise linear tariff that does not have a fixed fee.

Suppose $U_1$ offers the three-part tariff, as illustrated in figure 3(a), in which $T_1 = (P_M - c_1)Q_M$, $W_1(q_1) = c_1q_1$ if $0 \leq q_1 < S_1$ and $W(q_1) = c_1S_1 + w_1(q_1 - S_1)$ if $q_1 \geq S_1$ where $Q_M < S_1 < Q(P)$. Set $w_1 = R'(Q(P))$, in which case $0 \leq w_1$ by assumption and $w_1 = R'(Q(P)) < R(Q(P))/Q(P) = P$ from the concavity of $R(Q)$. Then set $S_1$ so that $T_1(Q(P)) = R(Q(P))$; i.e.

$$S_1 = \frac{(P - R'(Q(P)))Q(P) - (P_M - c_1)Q_M}{c_1 - R'(Q(P))}.$$ 

The inequality $S_1 < Q(P)$ holds given $P_M > c_1$ and $c_1 > P$. The inequality $Q_M < S_1$ holds given

$$(P - R'(Q(P)))Q(P) > (P_M - R'(Q(P)))Q_M,$$

which reflects that $Q(P) = \arg\max_{q_1} (R(q_1) - R'(Q(P))q_1)$ and that $R(q_1)$ is strictly concave. This establishes any such contract satisfies incremental-unit discounting.

![Figure 3: Optimal incremental-units tariffs](image)

It remains to check the conditions in proposition 3 hold for optimality. Recall a sufficient condition is $T_1(0) = 0$, $T_1(Q_M) = R(Q_M)$ and $T_1(Q(P)) = R(Q(P))$, with $T_1(q_1) > \bar{R}(q_1)$ everywhere else. The equality conditions hold
by construction. The inequality condition holds for $0 < q_1 < S_1$ given that $Q_M = \arg \max_{q_1} (R(q_1) - c_1 q_1)$ and it holds for $q_1 > S_1$ reflecting that the strictly concave function $R(q_1)$ lies below the tariff line $T_1(q_1)$ for $q_1 > S_1$ which is tangent to $R(q_1)$ at $q_1 = Q(P)$.

The same conditions for optimality can be maintained by shifting the tangency point to the right of $Q(P)$, so that $w_1$ is lower and $S_1$ is higher than those specified above. Specifically, for any $Q \geq Q(P)$ such that $R'(Q) \geq 0$, then the proposition continues to hold with $w_1 = R'(Q)$ and $S_1 = ((P(Q) - R'(Q)) Q - (P_M - c_1) Q_M) / (c_1 - R'(Q))$.

Now we show that the same conditions for optimality hold without the use of a fixed fee, provided market demand is such that $P(0)$ is finite. $U_1$ offers a tariff where the fixed fee is replaced with an extra step in the tariff schedule, as illustrated in figure 3(b), with a high initial wholesale price. In particular, the following tariff satisfies all the conditions for optimality: $T_1 = 0$, $W(q_1) = R'(0) q_1$ for $0 \leq q_1 < S_1 = (P_M - c_1) Q_M / (R'(0) - c_1)$ and $W_1(q_1) = R'(0) S_1 + c_1 (q_1 - S_1)$ for $S_1 \leq q_1 < S_2$ and $W_1(q_1) = R'(0) S_1 + c_1 (S_2 - S_1) + w_2 (q_1 - S_2)$ for $q_1 \geq S_2$ where $w_2$ and $S_2$ are equal to $w_1$ and $S_1$ defined above for the tariff with a fixed fee. This works given the strict concavity of $R(Q)$ provided $R'(0) = P(0)$ is finite. This ensures the tariff is everywhere above the revenue function for positive quantities up to the point where this tariff schedule intersects with the tariff $T_1(q_1)$ at $S_1$, with the rest of the proof as before.

Clearly the tariffs here are the simplest tariff within the class $T_I$, since they involve just two price breaks (three parts). Proposition 2 established that a tariff in $T_I$ with just one price break (two parts) could not achieve optimality. ■

This proposition shows that quantity discounting, defined in a very standard way, can be used by the incumbent to deter entry and obtain its full monopoly profits. Incremental-units quantity discounting involves the downstream firm enjoying a progressively lower wholesale price as it purchases more. The optimal contract can thus be expressed in the form that the downstream firm pays a fixed fee to buy its initial units, paying $c_1$ per unit for purchases up to $S_1$ units, and thereafter $w_1 < c_1$ per unit for additional units. Figure 3(a) illustrates the tariff. No instrument in the tariff is redundant. Clearly the two different wholesale prices serve different purposes. The wholesale price of $c_1$ ensures that $D_1$ chooses the monopoly price, so that the monopoly level of quantity will be purchased. The wholesale price of $w_1$ that applies if at least $S_1$ units are purchased ensures that $U_2$ does not find entry profitable (since $D_1$ would be willing to price down to $P$ to serve the market). The constraint that at least $S_1$ units be purchased for $D_1$ to enjoy the lower wholesale
price ensures that \( D_1 \) does not want to move onto this lower wholesale price schedule in the absence of competition. The fixed fee is chosen to extract \( D'_1 \)'s monopoly profit in equilibrium.\(^{15}\) Proposition 4 also establishes that (except for demand specifications where it takes an arbitrarily high price to drive demand to zero), the same outcome can be achieved by replacing the fixed fee with an additional step in the existing tariff schedule (so it remains a three-part tariff), but with a sufficiently high initial wholesale price. Figure 3(b) illustrates. The purpose of this very high initial wholesale price is exactly the same as the fixed fee, to extract \( D'_1 \)'s monopoly profit in equilibrium.

A similar result is obtained when looking at tariffs within the class \( T_A \). Again, the simplest such tariff that is optimal exhibits quantity discounting, in this case all-units quantity discounting.

**Proposition 5** *The incumbent can obtain full monopoly profit, deterring entry in the process, by using a three-part tariff from the class \( T_A \) which exhibits all-units quantity discounting. This is the simplest tariff in the class \( T_A \) that allows the incumbent to extract the full monopoly profit. The lowest wholesale price is below the incumbent’s marginal cost.*

**Proof.** Suppose \( U_1 \) offers the tariff in which \( D_1 \) pays \( R'(0) \) per unit if \( q_1 < Q_M \), \( P_M \) per unit if \( Q_M \leq q_1 < Q(P) \) and \( P \) per unit if \( q_1 \geq Q(P) \). Such a contract is illustrated in figure 4(b). All-units discounting follows from the strict concavity of \( R(Q) \) which implies \( R'(0) > P_M > c_1 > P \). By construction, \( T_1(0) = R(0), T_1(Q_M) = R(Q_M) \) and \( T_1(Q(P)) = R(Q(P)) \). Given the strict concavity of \( R(q_1) \) we also have that \( T_1(q_1) > \tilde{R}(q_1) \) everywhere except at \( q_1 = 0, q_1 = Q_M \) and \( Q(P) \), so that the conditions in proposition 3 hold for optimality provided \( R'(0) = P(0) \) is finite. Figure 4(b) makes it clear why this tariff is the simplest tariff within \( T_A \) that can satisfy the conditions for optimality without the use of a fixed fee.

Exactly as with optimal incremental-units tariffs, the optimal contract can also be achieved with a fixed fee \( T_1 = (P_M - c_1)Q_M \) replacing the first step in the tariff function, as illustrated in figure 4(a). The use of this fixed fee will in fact be necessary if \( P(0) \) is not defined. Then \( W_1(q_1) = c_1q_1 \) if \( 0 \leq q_1 < Q(P) \) and \( W_1(q_1) = w_1q_1 \) with \( w_1 = P - T_1/Q(P) < c \) if \( q_1 \geq Q(P) \). Note \( R'(Q(P)) \geq 0 \) implies \( PQ(P) > P_MQ_M \) from the concavity of the revenue function, so \( w_1 > (Q_M/Q(P))c_1 > 0 \) ensuring the wholesale price still remains positive. \( \blacksquare \)

\(^{15}\)This optimal three-part tariff is just the lower envelope of two two-part tariffs. One two-part tariff would be the tariff offered by a monopolist, which applies in equilibrium, and the other two-part tariff is designed to deter entry. \( D_1 \) self selects the appropriate two-part tariff, depending on whether it faces entry or not.
Figure 4: Optimal all-units tariffs

Again, no instrument in the tariff is redundant. Here the high initial wholesale price of $R'(0)$ for output levels below monopoly ensures $D_1$ does not want to move into this region when it is indeed a monopolist. The wholesale price of $P_M$ for intermediate levels of output starting from the monopoly output ensures the monopoly quantity level will be purchased and extracts the corresponding profit. The lower wholesale price of $P$ that applies if at least $Q(P)$ units are purchased ensures that $U_2$ does not find entry profitable. The constraint that at least $Q(P)$ units be purchased to enjoy the lower wholesale price ensures that $D_1$ does not want to move onto this lower wholesale price schedule in the absence of competition.

4 Extensions

In this section, we consider some important extensions of the above benchmark model. Section 4.1 shows that our results may continue to hold even if the incumbent and entrant make their contract offers simultaneously rather than sequentially, including if the incumbent’s contract offer is only observed with a small probability. Section 4.2 explains how the ability to use upfront fees only make it easier for an incumbent to deter entry. In particular, we show how an upfront fee can be used to extract the disposal-rent which must be left to the incumbent’s retailer when the revenue function is decreasing at $Q(P)$. Section 4.3 discusses the fundamental commitment problem that arises from the incentive to renegotiate the vertical contract in the case entry does arise (i.e. out-of-equilibrium) arguing this suggests some types of vertical contracts may be preferred to others. Section 4.4 explores an implication of this point, where a constraint on vertical contracts imposed by the in-
centive to renegotiate allows us to explain how limit pricing and quantity discounting can arise at the actual equilibrium quantity purchased by the incumbent’s retailer (rather than just as part of the incumbent’s offer). Finally, section 4.5 explains how the analysis changes when firms offer imperfect substitutes.

4.1 Simultaneous and unobservable offers

While our benchmark results were derived in a framework in which the incumbent has a first-mover advantage (since it makes its observable contract offer first), the equilibria we obtained remain equilibria even when both manufacturers make their offers simultaneously, as in the existing literature on the commitment effects of vertical contracts. Given the incumbent’s quantity discounting contract, the rival manufacturer would prefer to remain inactive thereby avoiding any (possibly trivial) fixed costs.

Specifically, suppose that $U_1'$s offer to $D_1$ is made at the same time that $U_2$ makes its offer to $D_2$ (i.e. in stage 1). $U_2$ decides whether to enter in stage 2 (incurring some, possibly trivial, fixed cost). In stage 3, active retailers compete. Suppose, first, that $U_2$ observes $U_1'$s contract before deciding whether to enter in stage 2. Then our previous analysis and results continue to apply, even though contracts are offered simultaneously.

A more interesting case arises when $U_2$ does not always observe $U_1'$s contract before deciding whether to enter in stage 2. Suppose $U_2$ observes $U_1'$s contract only with some positive probability and consider $U_1$ making an optimal contract offer as in propositions 3-5. If $U_2$ expects $U_1$ to make such an offer, it will prefer to stay out even if the offer is unobserved. Moreover, it would be pointless for $U_2$ to instead enter and try to counter with some other contract, since any such deviation will not be observed by $U_1$ and $D_1$ (at least until after their contract has been signed). Since $U_1$ cannot obtain more than full monopoly profits, our optimality results continue to apply even if offers are made simultaneously and only observed with some small probability.

This finding is in contrast to the existing literature on competing vertical chains in which the commitment benefits of delegation rely on the observability of vertical contracts. Given contracts are not observed (or observed with an arbitrarily small probability), a manufacturer always does best selling to its retailer at its true cost and recovering the maximum profit possible through its fixed fee (Katz, 1991). Note, though, for our results to hold in a robust way, it is necessary that there be at least some small probability of $U_1'$s contract being observed. If not, then the vertical limit pricing strategy would be weakly dominated by a simple two-part tariff in which the
wholesale price equals $c_1$ and the fixed fee extracts monopoly profit (which gives the same profit in equilibrium but would avoid the possibility that $U_1$ has to supply units below cost in case of entry).

Simultaneous and partially unobserved offers do, however, open up the possibility of other equilibria arising. Entry deterrence is no longer the unique equilibrium outcome. For instance, as the probability of observing $U'_1$'s contract becomes small, another equilibrium arises in which $U_2$ chooses to be active and offers a tariff with a linear wholesale price at (or if necessary) just below $c_1$. Since $c_1 > P$ this ensures that $U_2$ more than covers its fixed costs if it takes the market. Given this offer is expected by $U_1$, it is not profitable for $U_1$ to offer a contract which induces $D_1$ to undercut. It cannot do better than to set a wholesale price of $c_1$. Thus, in cases in which the incumbent does not have a first-mover advantage in determining its contract and in which contracts are most likely to be unobserved, entry deterrence is not inevitable, although it is still possible.

### 4.2 Upfront fees and the “disposal-rent”

In our analysis up to this point we have restricted attention to a certain class of contracts in which upfront fees (fixed fees paid at the time the retailers accept their respective contracts) are not allowed but only fixed fees that apply when some positive quantity is purchased by the retailer are considered. We argued in section 2, large upfront fees (e.g. equal to the ongoing monopoly profit of the industry) may not be feasible or efficient in practice. However, even if such upfront fees are possible, the existing contracts characterized by propositions 3-5 remain optimal. In other words, $U_2$ can still not profitably enter even if it can use upfront fees (while $U_1$ does not). Given the contracts characterized by propositions 3-5, for $D_2$ to capture any share of the market it must price at or below $P$, meaning there is no way for $U_2$ to make a profit. Given $U_1$ already obtains the monopoly profit, it has no reason to use upfront fees. Nevertheless, the ability to use upfront fees as opposed to the fixed fees analyzed up until now can make it easier for $U_1$ to deter entry. As we will show, they are in fact necessary to ensure optimality if we allow the possibility that $R'(Q(P)) < 0$.

Upfront fees can make it easier for an incumbent to obtain monopoly profits through vertical limit pricing since they provide a further first-mover advantage to the incumbent, whose offer is accepted first. In equilibrium, $D_1$ is a monopolist and $U_1$ extracts the expected monopoly profit in equilibrium through its upfront fee. This means if $D_1$ does face competition, this upfront fee is a sunk cost for $D_1$, allowing $U_1$ to collect more in total from $D_1$ while still ensuring it will undercut $U_2$. 

21
or $D_2$ as is required to prevent entry. This also means, with upfront fees, $D_1$ may regret signing its contract with $U_1$, in the case there is entry. Despite this difference, the existing optimal contracts continue to work as in propositions 3-5 if the fixed fee $T_1$ is simply replaced by an upfront fee of equal magnitude instead. Moreover, even if $U_1$ uses the upfront fee to recover the maximum amount possible (so $D_1$ will regret signing its contract if $D_2$ also enters), the same types of contracts as analyzed above will still exhibit quantity discounting since they will still have slope $c_1$ for intermediate quantity levels and a lower slope (less than $P$) for sufficiently high quantities (in order that the wholesale tariff passes through $R(Q)$ at $Q(P)$). Figure 5(a) illustrates the point that with this upfront fee, the wholesale price that applies at the margin for large purchases can be increased, although it must still remain less than $P$.

![Figure 5: Optimal tariffs with upfront fees.](image)

An upfront fee becomes necessary to achieve optimality when the assumption $R'(Q(P)) \geq 0$ does not hold. Without upfront fees, $U_1$ cannot do better than to offer a three-part all-units tariff. However, unlike the case before in which $D_1$ is left with no profit in equilibrium, here $U_1$ must leave some rent for $D_1$. The reason is that if there is no entry (as will be the case in equilibrium), $D_1$ can always make a positive profit buying $Q(P)$ units for $T(Q(P))$ but then selling fewer units so as to obtain a higher revenue given $T(Q(P)) = R(Q(P))$ and $R'(Q(P)) < 0$, freely disposing of the additional units.\[16\] In fact, $D_1$ will optimally sell only the revenue

\[\text{If instead the marginal cost of disposing of units is at least } |R'(Q(P))|, \text{ then no such rent is needed when } R'(Q(P)) < 0.\]
maximizing number of units; i.e. it will sell \( Q_R = \arg \max_Q R(Q) \) units. To prevent this situation, \( U_1 \) will offer \( D_1 \) a rent \( r_D \) for selling the monopoly output level \( Q_M \) which must equal \( R(Q_R) - R(Q(P)) = \tilde{R}(Q(P)) - R(Q(P)) \). We call this rent \( D'_1 \)'s “disposal-rent”, the rent \( D_1 \) can obtain in equilibrium given it can freely dispose of the good. Thus, the incumbent may still deter entry, but its monopoly profit will be reduced by the size of this rent. If an upfront fee is possible, it will then allow \( U_1 \) to extract this rent. The resulting optimal tariff involves an upfront fee plus a three-part all-units tariff (in other words, it has four parts). Figure 5(b) illustrates.

### 4.3 Commitment and renegotiation

All our results to this point have illustrated that quantity discounting is a somewhat generic property of optimality contracts, reflecting that the optimal contract is pinned down by the underlying concave revenue function. Apart from counter-examples relying on unnecessarily complicated tariffs (e.g. as illustrated in figure 2(b)), there is one simple counter-example to quantity discounting that arises within the class of piece-wise linear tariffs, but only when an upfront fee can be used. It is the special case of an incremental-units tariff in which the price break is set at \( Q_M \), such that the wholesale price is less than \( c_1 \) up to \( Q_M \) and above \( c_1 \) for quantities exceeding \( Q_M \). As figure 5(c) illustrates, this also satisfies the key properties of optimality, but only because the fixed fee is upfront. As a monopolist, \( D_1 \) will buy \( Q_M \) units, given additional units cost more than \( c_1 \). The tariff works rather like quantity forcing in equilibrium. Facing competition for \( D_2 \) and with its upfront fee sunk, \( D_1 \) will be willing to price down to \( P \) given the wholesale tariff passes through \( R(Q) \) at \( Q(P) \). Since the initial wholesale price (up to \( Q_M \)) is lower than \( c_1 \), this contract requires an upfront fee which exceeds the full monopoly profit (previously, we argued collecting the full monopoly profit through an upfront fee may be unreasonable). For this reason, \( U_1 \) would like to renege on the contract after collecting the upfront fee, in equilibrium as well as out-of-equilibrium (i.e. regardless of whether the rival enters or not).

As the previous contract illustrated, our results rely heavily on the assumption that the incumbent can commit to its vertical contract. (Note we do not require downstream firms to commit to the contract.) Some contracts require more commitment power than others in that the incentives for \( U_1 \) to renege or renegotiate are stronger. The previous contract is one such case. More generally, the commitment problem arises since if the rival does enter, the incumbent would prefer not to provide goods according to its original tariff schedule given this would involve supplying its retailer below its own marginal cost. Post entry, if the incumbent and
its retailer could freely renegotiate, they would want to do so, thereby making such entry profitable.

To avoid any such commitment problem, we have implicitly treated the vertical contract as a commercial contract which is enforceable by law. This is, of course, also a common requirement in the existing literature on vertical contracts. Allowing for costly renegotiation may still enable vertical contracts to work in similar ways, as would the case in which the breach of contract is possible but with sufficient penalties. Alternatively, if a third party is available which can help enforce the contract, then upfront fees may be used to avoid the commitment problem. The upfront fee could be paid to the third party upon signing the contract. The fee could then only be passed onto the incumbent if it honors its initial contract. This can make it profitable for the incumbent to supply its retailer even in the face of entry, since the amount it receives in total (with the upfront fee) is $PQ(P) + (P_M - c_1)Q_M$ which can exceed its costs $c_1Q(P)$ of supplying this many units even though $P < c_1$.

4.4 Equilibrium limit pricing and renegotiation

For the most part, our theory of limit pricing is one in which limit pricing would not actually be observed in equilibrium. Although $D_1$ is offered a tariff which involves units that can be purchased at below $U_1's$ marginal cost over some range, in equilibrium $D_1$ will not actually purchase any units at a price below $U_1's$ marginal cost. As figure 1(a) illustrates, limit pricing can arise in equilibrium when $U_1$ is constrained in the tariffs it can offer, such as when it can only offer two-part tariffs. To deter entry requires $U_1$ offers a contract which in equilibrium will leave $D_1$ with some rent. This means $U_1$ can no longer extract $D_1's$ monopoly profit in equilibrium. $U_1$ will then choose a wholesale price below marginal cost which optimally trades-off the cost of increasing its quantity above its preferred monopoly level $Q_M$ with the reduction in rent it must leave $D_1$.

Other types of constraints may lead to similar results. One constraint on tariffs may arise from the incentive for $U_1$ to renegotiate tariffs in the face of entry. To illustrate the point, consider an optimal tariff such as the three-part incremental-units tariff characterized in Proposition 4. Suppose $U_1's$ perceived cost of reneging on the contract is $\kappa$, so that $U_1$ will renge on a contract if it expects a loss of more than $\kappa$ from continuing to supply according to the contract at any point. Given the information structure we have assumed, up until the equilibrium quantity $Q_M$, $U_1$ will expect to sell $Q_M$ units. Beyond this, it will face a loss equal to $(R(Q_M) - c_1Q_M) - (R(Q(P)) - c_1Q(P))$. If this is greater than $\kappa$, $U_1$ will have to reduce this loss in order to avoid the ex-post incentive to renge in the case of entry.
It can do so by shifting some of the loss to the left of the equilibrium point. Since $U_1$ only expects to sell the equilibrium quantity, this loss reduces the profit $U_1$ expects in equilibrium but increases the incremental profit out-of-equilibrium. Shifting the loss to the left of the equilibrium point may be achieved by reducing the marginal wholesale price to be below cost before the equilibrium point is reached, giving rise to limit pricing in equilibrium.

To see why limit pricing may arise in equilibrium, note that $\kappa$ together with $Q(P)$ determine the loci of equilibrium choices for different tariffs. These various equilibrium choices lie on a line with slope $c_1$ that is below $R(Q_M)$ at $Q_M$. The amount by which the line is below $R(Q_M)$ just depends on $\kappa$. By construction, the equilibrium choices all give $U_1$ the same equilibrium profit and the same loss from continuing to supply beyond the equilibrium point (equal to $\kappa$). One such tariff is the optimal three-part tariff with a fixed fee from Figure 3a with the wholesale price $w_1 = c_1$ but with the switching point $S_1$ shifted to the right so that $(c_1 - w_2)(Q(P) - S_1) = \kappa$, where $w_2 = R'(Q(P)) < c_1$. The corresponding profit of $U_1$ for this tariff is

$$\Pi_1 = R(Q(P)) - c_1 S_1 - w_2(Q(P) - S_1) = -(c_1 - P) Q(P) + \kappa. \quad (9)$$

Obviously, we require $\kappa > (c_1 - P) Q(P)$ to ensure that $U_1$ can make a profit in equilibrium.

Given the cost of reneging, no other tariff can do better. Other three-part tariffs (with different $w_1$) which imply the same loss from reneging also imply the same equilibrium profits as in (9) but can involve $w_1 < c$ so that equilibrium limit pricing arises. This would also be true even without a fixed fee. Moreover, the simplest optimal tariff given $\kappa > 0$ is actually a two-part tariff which exhibits equilibrium limit pricing given $w_1 < c_1$ must hold so that the equilibrium point lies on the loci described above. Provided the cost of reneging is neither too high or too small, such a two-part tariff does just as well as the optimal three-part tariff described above.\footnote{If the cost of reneging is too large, then $U_1$ will do better offering a three-part tariff given two-part tariffs constrain the maximum profit that can be extracted by $U_1$ (e.g. in the case of linear demand, to half the monopoly profit). Nor can the cost of reneging be too small, since it must exceed $(c_1 - P) Q(P)$ for $U_1$ to make a profit.}

4.5 Imperfect substitutes

In this extension, we show that tariffs with quantity discounting deter entry and may still lead to full monopoly profit even when goods are imperfect substitutes.\footnote{A supplementary appendix contains more detailed analysis including formal proofs and specific examples.} The analysis in the case with imperfect substitutes introduces some new elements.
Obtaining the monopoly profit while blocking entry may no longer be possible since for some demand specifications such as logit, neither retailer can be driven from the market. In this case, if $U_1$ tried to extract the monopoly revenue then the rival could always profitably enter knowing that $D_1$ would want to exit given it cannot recover the monopoly revenue facing such competition. Even though the incumbent makes a take-it-or-leave-it offer to its retailer and even though this contract deters entry, the incumbent has to leave some profit with its retailer to satisfy ex-post participation by $D_1$. Furthermore, since in this case $U_1$ can no longer obtain monopoly profit, sometimes it may prefer not to deter entry, preferring instead to share the market.

To proceed, we retain the timing of the game and the structure of tariffs proposed for homogenous goods. We focus first on price competition. At the end of this section, we note how our analysis extends to the case of quantity competition. To keep matters as simple as possible, in what follows we focus on all-units tariffs. This is sufficient for our purposes since the outcome for any entry deterring tariff can be replicated by a tariff from $T_A$. Specifically, we consider tariffs of the form
\[ T_1(q_1) = W_A^{(3)}(q_1; w_1, w_2, w_3, S_1, S_2), \]
where $w_i$ are marginal wholesale prices and $S_i$ are switch points in the tariff. The demands $q_i(p_1, p_2)$ are downward sloping and goods are imperfect substitutes: $\partial_i q_i(p_1, p_2) < 0$ and $\partial_j q_i(p_1, p_2) > 0$ for $i = 1, 2$, where $p_1$ and $p_2$ are the prices set by $D_1$ and $D_2$ respectively. The Jacobian of the demand system is assumed to be negative definite, so the system of demand functions can be inverted to obtain inverse demands $p_i(q_1, q_2)$ for $i = 1, 2$. The monopoly demand for $D_1$ is defined as $q_1(p_1, p_2)$, where $p_2 = p_2(q_1, 0)$. It is denoted by $Q_1(p_1)$ with $p_1(Q_1)$ being the inverse monopoly demand. We assume that the residual revenue function, defined as $R_i(q_i, p_j) = p_i(q_i, p_j) q_i$, is strictly concave in $q_i$, twice differentiable and
\[ \partial_{12} R_1(q_1, p_2) \geq 0. \]

The inequality in (10) holds for a wide range of demand functions such as linear and logit.

The fixed cost of entry $F$ is non-negative and less than $\overline{F}$, where $\overline{F}$ is the level of $U_2$'s fixed cost such that it would (just) not want to enter even if it competed directly with $U_1$. This upper limit $\overline{F}$ is equal to the maximum profit $U_2$ can obtain if $U_1$ sets its price equal to its marginal cost $c_1$. Given the fixed cost of entry $F$, the break-even price $\underline{p}$ is defined by (3). Note that when goods are homogenous, then $\overline{F} = (c_1 - c_2) Q(c_1)$.

The objective of $U_1$, assuming it prefers to deter entry, is to choose a tariff from the above class which most profitably deters entry. The analysis differs depending on the shape of the residual revenue function. Figure 6 illustrates the optimal entry
deterring tariffs for three different shapes when the maximum of the residual revenue function is finite.

(a) Equivalent to homogenous case

(b) Disposal and deterrence rents

(c) Equilibrium quantity greater than monopoly

Figure 6: All-units quantity discounting (imperfect substitutes)

Figure 6(a) considers the case the residual revenue $R_1(q, P)$ intersects the monopoly revenue function $R(q)$ where it is non-decreasing. The logic of vertical limit pricing in this case is exactly as in the homogenous setting. Under competition, $D_1$ faces the residual revenue function and is willing to price to sell $S_2$ units so that it takes the whole market even if $D_2$ prices at $P$. Given this, $U_2$ will not enter. As a monopolist, $D_1$ is willing to buy exactly $Q_M$ units as is required for optimality. Full monopoly profit is possible.\(^{19}\)

In the case discussed above, $D_1$ would prefer to take the whole market if $D_2$ enters and sets any price at or above $P$. This is because the residual revenue function is non-decreasing at the point it intersects the monopoly revenue function. Figure 6(b) considers the case the residual revenue function is decreasing at the point it intersects the monopoly revenue function. This introduces a new rent, denoted $r_I$, which we call the incentive-rent. This rent arises since the tariff required to deter entry is below the monopoly revenue function at $S_2$. This means, when $D_1$ does not face entry, it can earn a profit equal to the difference between the monopoly revenue function and

\(^{19}\)The case when the residual revenue $R_1(q, P)$ intersects the monopoly revenue function $R(q)$ where it is decreasing is also possible. As with homogenous goods, $U_1$ must leave $D_1$ with the disposal-rent $r_D$. This rent prevents $D_1$ preferring to buy $S_2$ units as a monopolist (to get the low wholesale price) and disposing of the additional units to move to the peak of the monopoly revenue function.
the residual revenue function at $S_2$, which is the amount $r_I = R_1(S_2) - R_1(S_2, P)$ in figure 6(b). To get $D_1$ to choose the monopoly quantity $Q_M$ in equilibrium, $U_1$ must therefore leave $D_1$ at least $r_I$. In addition, $U_1$ must also leave $D_1$ with some disposal-rent $r_D = \bar{R}_1(S_2) - R_1(S_2)$ due to the fact that the monopoly revenue function is also decreasing at $S_2$, which brings the total rent which must be left up to $r = r_D + r_I = \bar{R}_1(S_2) - R_1(S_2, P)$ in figure 6(b).

In the cases illustrated in figure 6(a) and 6(b), $U_1$ can obtain full monopoly profit if, in addition to the given tariff, it uses an upfront fee to extract ex-ante the rent that $D_1$ must be left ex-post. However, in the case described in the figure 6(c), where the residual revenue function is closer to the monopoly revenue function at $Q_M$ than at $S_2$, it is no longer optimal to induce $D_1$ to choose the monopoly quantity in equilibrium (i.e. when entry is deterred). For this reason, even if $U_1$ can use an upfront fee, it will obtain less than the full monopoly profit. As a result, when goods are not close substitutes, $U_1$ may be better off offering a contract which accommodates entry even though entry deterrence is possible. With price competition, this contract will involve softening competition as in the standard literature.\footnote{The remaining case in which the maximum of the residual revenue function is finite is where the maximum of the residual revenue function lies to the left of the maximum of the monopoly revenue function. This case, which is more involved, is considered in the supplementary appendix. The supplementary appendix also considers three other cases where the maximum of the residual revenue function is no longer finite but in these cases entry deterrence will generally not be achieved.}

A key property of each of the optimal tariffs illustrated in figure 6 is they involve quantity discounting, with the wholesale price declining at each switch-point in the tariff, and such that the lowest wholesale price is strictly less than $U_1$'s marginal cost $c_1$. This result generalizes. Allowing for more general forms of the residual revenue function satisfying our general assumptions (e.g. we can allow for a residual revenue function which is always increasing, as is the case with logit, or with a maximum of the residual revenue function to the left of the maximum of the monopoly revenue function), it can be established (a supplementary appendix contains the formal details):

**Proposition 6** The optimal entry deterring tariff in the class of all-units three-part tariffs exhibits quantity discounting, with the wholesale price declining at each switch-point in the tariff, and such that the lowest wholesale price is (weakly) less than $U_1$'s marginal cost $c_1$.

The analysis above can be easily extended to quantity competition. Instead of assuming $\partial_{12} R_1(q_1, p_2) \geq 0$, we require $\partial_{12} \bar{R}_1(q_1, q_2) \leq 0$, where $\bar{R}_1(q_1, q_2) =$
\( p_1(q_1, q_2) q_1 \) is the residual revenue function under quantity competition. This assumption is satisfied for a wide variety of demands including linear demands. The construction of the optimal entry deterring tariff is essentially the same as for the case of price competition. However, instead of the break-even price \( P \), we now consider the break-even quantity \( Q = Q_1(P) \). Since firms are less competitive under quantity competition than price competition, to deter entry under quantity competition, \( D_1 \) has to be induced to be even more competitive than it otherwise would be. This makes entry deterrence somewhat more difficult. On the other hand, renegotiation issues may be less severe under quantity competition since even when accommodating entry, \( U_1 \) can benefit by committing to make its downstream firm more aggressive. That is, under quantity competition a “top-dog” strategy as in Fudenberg and Tirole (1984) is optimal for both entry deterrence and entry accommodation.

5 Conclusions

The key new idea developed in this paper is that commonly used forms of wholesale contracts involving quantity discounting can have entry deterring effects. An upstream incumbent can use such contracts to commit its downstream distributor or retailer to be more aggressive in the face of competition. In a benchmark setting, with homogenous price-setting firms, the simplest optimal contract is a three-part tariff. For low levels of purchases, the retailer purchases at a wholesale price set above the incumbent’s marginal cost, thereby providing a way for the manufacturer to extract the retailer’s monopoly profit (alternatively, a fixed fee can be used for this purpose). For purchases in some intermediate range, the retailer purchases at a wholesale price set equal to the incumbent’s marginal cost, thereby ensuring the retailer sets the correct monopoly price when it indeed does not face competition. For purchases beyond some yet higher level, the retailer purchases at a wholesale price set below the incumbent’s marginal cost, thereby ensuring that in the face of competition, the retailer will want to compete aggressively, in such a way that the rival will not want to enter. Thus, we provide a new explanation of limit pricing, one which does not depend on asymmetric information.

The benchmark model we have provided can be extended in numerous directions. Several natural extensions have been analyzed in this paper, most significantly to the case with imperfect substitutes. In this and some other cases, conditions are provided under which the incumbent still deters entry using a vertical contract although often this requires the incumbent to leave its retailer with some rent. We
also explored what happens under alternative timing and informational assumptions. The key assumptions necessary to establish our results are that the incumbent can commit, at least partially, to a wholesale contract and that this contract is observed, at least some fraction of the time, prior to a potential entrant making a decision about incurring fixed costs to enter the market (or in the case of predation, prior to an existing rival making a decision about incurring additional fixed costs to stay in the market).

An interesting direction for future research would be to explore whether a similar theory can be constructed when there is more than one incumbent firm, in which case the design of the optimal tariff is likely to be more complicated given the tension between deterring entry and softening competition between incumbents in the case of price competition (this tension may be less of a constraint under quantity competition). Also of interest is to consider a dynamic version of the vertical limit pricing story, in which downstream firms make a sequence of purchase decisions. We discussed such a possibility informally when analyzing optimal contracts in the face of renegotiation. A dynamic version of our vertical limit pricing story should be able to formally explain the use of rebates and loyalty programs to deter entry or drive existing rivals out.

Finally, we note, that a very natural extension of the established limit pricing literature to a vertical setting, would be to modify the standard signaling story of limit pricing in which the incumbent’s cost is unobservable so as to incorporate the fact that the incumbent sells to retailers rather than final consumers. In such a theory, a low wholesale price might signal that the incumbent has low cost and therefore would deter entry. However, a low wholesale price can also have a direct entry deterring effect, in addition to its signaling effect, along the lines considered in this paper. In other words, the analysis of signaling through limit pricing in a vertical setting is likely to make for an interesting extension of the existing theories of limit pricing and predation.

6 References


Fudenberg, Drew and Jean Tirole (1984) “The Fat-Cat Effect, the Puppy-


