Wholesale Price Determination Under the Threat of Demand-Side Substitution

Roman Inderst and Greg Shaffer

University of Frankfurt and Imperial College London, University of Rochester and University of East Anglia

October 2011
Wholesale Price Determination Under the Threat of Demand-Side Substitution

Roman Inderst* Greg Shaffer†

October 2011

Abstract

An upstream supplier that is constrained both by downstream competition and the threat of demand-side substitution faces a tradeoff between maximizing joint-profit and extracting surplus. Although joint-profit maximization calls for relatively high marginal wholesale prices in order to dampen intra-brand competition, surplus extraction will be higher when the supplier instead charges relatively low marginal wholesale prices. The reason is that by inducing more intra-brand competition through lower wholesale prices, the supplier makes it less attractive for downstream firms to switch to alternative sources of supply. We show how this can make it optimal for the supplier to disadvantage more efficient and thus ultimately larger buyers, thereby smoothing out differences in their market shares. We further show that despite the use of non-linear supply contracts, marginal wholesale prices and thus final goods’ prices will decrease when either downstream competition intensifies or the supplier becomes more constrained by the threat of demand-side substitution.

Keywords: Vertical control; Input markets; Price discrimination; Buyer power

*University of Frankfurt and Imperial College London; Correspondence: University of Frankfurt, Mer- tonstrasse 17, 60054 Frankfurt am Main, Germany; E-mail: inderst@finance.uni-frankfurt.de.
†University of Rochester and University of East Anglia; Correspondence: Simon School of Business, University of Rochester, Rochester NY 14627, USA; E-mail: shaffer@simon.rochester.edu.
1 Introduction

Suppliers with strong brands are often in a position to set the terms and conditions they offer to retailers or other downstream firms. Typically, however, their pricing power is constrained by their customers’ ability to switch to alternative inputs or products. This paper explores a simple conceptual framework that analyzes an upstream firm’s optimal pricing decisions when the downstream firms it sells to have access to such alternatives.

Our starting point is a setting with non-linear contracts. The use of non-linear contracts is important in what follows because it allows one to make a crucial distinction between a firm’s average wholesale price (which serves to divide the surplus) and its marginal wholesale price (which affects the firm’s competitiveness vis à vis its rivals). When contracts are observable, for example, it is well known that a supplier can dampen intra-brand competition among its retailers by choosing a sufficiently high marginal wholesale price for every downstream firm. And, in fact, an unconstrained upstream monopolist would optimally set the downstream firms’ average and marginal wholesale prices so as to induce thereby the monopoly outcome and extract all surplus. We show, however, that inducing the monopoly outcome would not be optimal for the upstream supplier when the downstream firms have the viable alternative of procuring the respective input or product elsewhere.

Our first insight is that such a supplier faces a trade-off between maximizing joint-profit and extracting surplus. Although joint-profit maximization calls for relatively high marginal wholesale prices for all downstream firms, surplus extraction will be higher when the supplier instead charges relatively low marginal wholesale prices. The reason is that by reducing the marginal wholesale price for one downstream firm, say downstream firm $i$, the value of the outside option of rival firms, say of downstream firm $j$, is reduced. This follows because even after switching to another source of supply, firm $j$ will still find itself in competition with firm $i$, which prices more competitively when it benefits from a lower marginal wholesale price. With the attractiveness of switching suppliers thus reduced for firm $j$, the supplier can extract more of their joint surplus. Because the same logic applies also with respect to the value of firm $i$’s outside option, the supplier thus has an incentive to stimulate competition by lowering the marginal wholesale prices of all downstream firms.

We further find that as the alternative options become closer substitutes, or as the downstream market becomes more competitive (e.g., the number of firms increases or they...
become less differentiated), the supplier will respond by cutting its marginal wholesale prices even further. This result is strikingly different from what an unconstrained upstream monopolist would do. An unconstrained supplier would optimally counteract an increase in the competitiveness of the downstream market by raising its marginal wholesale prices in order to maintain the monopoly outcome in the face of smaller downstream mark-ups.

The empirical implications in the two cases are thus very different. In the benchmark case of an unconstrained upstream monopolist, final goods’ prices are independent of the degree of downstream competition. In contrast, when the supplier is constrained by the threat of demand-side substitution to an alternative source of supply, wholesale prices and thus prices in the final goods’ market are lower when either the supplier or the downstream firms are more constrained by competition — despite the presence of non-linear contracts.

Our second insight concerns the unequal treatment of heterogeneous downstream firms. If all firms were offered the same marginal wholesale price, the more efficient downstream firms would set lower final goods’ prices and have larger market shares as well as procure larger amounts of the supplier’s input than the less efficient downstream firms. Given such heterogeneity, how should the supplier optimally adjust its marginal wholesale prices? Should it charge lower marginal prices to the more efficient firms or the less efficient firms?

We abstract in our model from well-known explanations in the literature for why some buyers may obtain (size-related) discounts, such as the ability of larger buyers to distribute a fixed cost over a larger volume, and instead identify two conflicting forces that determine the degree of wholesale price differentiation when contracts are non-linear. On the one hand, to maximize total industry surplus, we find that a supplier should handicap less efficient downstream firms by charging them higher marginal wholesale prices (cf. Inderst and Shaffer, 2009). On the other hand, we find that the adverse impact on a firm’s outside option when its rivals become more competitive is larger when the firm is more efficient and thus has, ceteris paribus, a larger market share. This latter effect works in the opposite direction of the former effect (i.e., towards a lower marginal wholesale price).

\footnote{Fixed costs may be incurred when switching suppliers (cf. Katz, 1987; and Inderst and Valletti, 2009a). Another view is that larger buyers are more powerful and can thus demand a discount, which enhances their competitiveness vis-à-vis rivals. However, with nonlinear contracts, it is not clear why such a discount would have to be given “on the margin,” thereby affecting intra-brand competition. The same applies when a buyer may be in a position to demand better terms because it has access to a more favorable outside option. Again, the supplier could concede more profit to the buyer by adjusting only “infra-marginal” prices, leaving constant the marginal wholesale price, and thereby, the buyer’s competitive position.}
price for the less efficient firms). With linear demands, it turns out that the latter effect always dominates. This implies that the supplier will optimally want to smooth out the competitive disadvantage of the less efficient firms — (handicap the more efficient firms).

The supplier offers the less efficient firms a discount on its marginal wholesale price because this enhances its bargaining position with respect to the more efficient firms.\(^2\) As a consequence, the supplier’s sales become more “balanced” across the different downstream firms (e.g., across the supplier’s different distribution channels, or across different outlets in the case of retailing). This is consistent with an observation made by the UK Competition Commission in their recent investigation into the grocery retail market, in which it was stated that seemingly less competitive retailers sometimes receive more advantageous terms of supply.\(^3\) It is also consistent with what one of the authors learned in the course of (unpublished) interviews that were conducted during this investigation. In these interviews, producers of branded products alleged that they often employ various strategies designed to strengthen particular distribution channels. In order to support smaller retailers that purchase through wholesalers, for instance, producers occasionally sell lots and package sizes at a discount that is exclusive to these smaller retailers. Using a simple static framework, our model provides a formal underpinning for such a strategy.

There exists by now a large literature on vertical contracting with both observable and unobservable contracts. Hart and Tirole (1990) and others have shown that when contracts are unobservable, an upstream supplier may have an incentive to engage in opportunistic behavior against its downstream firms, which can lead to extreme intra-brand competition. In these cases, the supplier becomes in essence its own worst competitor and - when the opportunism problem is sufficiently severe - wholesale prices no longer depend on either downstream or upstream competition.\(^4\) In contrast, when contracts are observable, and when they are nonlinear but there is no threat of demand-side substitution, Inderst and Shaffer (2009) have shown that an upstream supplier can achieve the vertically-integrated

\(^2\)Through a different mechanism, Raskovich (2003) has also found that less efficient and thus ultimately smaller buyers might receive greater marginal discounts than larger buyers. He finds that although large buyers can be pivotal to a supplier’s decision whether or not to produce, and although one might think that a buyer’s bargaining position would be improved by being pivotal, the opposite can hold because “If other buyers’ payments fall short of costs, a pivotal buyer must cover the shortfall or forfeit consumption.”


\(^4\)For example, under a passive-beliefs restriction (under downstream Cournot competition, the result is more general), or in a contract equilibrium (cf. O’Brien and Shaffer, 1992), the supplier sets its marginal wholesale prices equal to its marginal cost. See Rey and Vergé (2004) for a comprehensive treatment.
outcome by offering discounts to the more efficient downstream firms, thereby amplifying differences in market shares. An implication of this behavior — as we show below — is that more intense downstream competition then leads to higher marginal wholesale prices.

Neither of these predictions (i.e., that wholesale prices will be independent of downstream competition or that wholesale prices will be increasing in downstream competition), however, is particularly appealing. An implicit assumption in competition policy, for instance, is that a tightening of competitive constraints would depress prices, regardless of whether the tightening occurs in the upstream or downstream markets. To this end, we are the first to provide a single framework with nonlinear contracts (we assume they are observable) in which competitive constraints in both the upstream and downstream markets have a monotonic and intuitive impact on both wholesale and final goods’ prices.\(^5\)

Thus, in view of the literature, our contribution is the following. We first isolate a novel determinant of wholesale prices. Under the threat of demand-side substitution, we show that marginal wholesale prices affect not only final prices but also how much surplus a supplier can extract. Importantly, lower wholesale prices arise not out of an opportunism problem, but because they are strategically set lower than what would be necessary to eliminate intra-brand competition. Our second key observation is that a supplier who is constrained by demand-side substitution on the part of buyers but who can still choose wholesale prices strategically may disadvantage a more efficient buyer, thereby creating a more level playing field among competing downstream firms. Finally, we contribute a simple framework of wholesale price determination that, in our view, leads to predictions that are less extreme than those obtained in the existing literature. In our model, there is no opportunism problem, nor does the supplier fully eliminate intra-brand competition. As a consequence, we obtain monotonic predictions about how competitive constraints in the upstream and downstream markets affect wholesale prices as well as final goods’ prices.

The rest of the paper is organized as follows. Section 2 introduces the model. Section 3 shows how intra-brand competition persists under the supplier’s optimal contracts. Sections 4 and 5 conduct a comparative statics analysis both with respect to the downstream

\(^5\)A different strand of the literature constrains the supplier to choose only linear contracts, either without constraints (DeGraba, 1990; Yoshida, 2000) or under the threat of demand-side substitution (Inderst and Valletti, 2009b). The choice of a model with linear contracts may be motivated by the particular application at hand, given that contracts in an industry may be restricted in this particular way. The same holds for the aforementioned distinction between observable and unobservable contracting, where the degree of observability, in particular in longer-term relationships, may vary between industries.
firms’ cost characteristics and with respect to the level of competitive constraints faced in the upstream and downstream markets. Section 6 discusses various extensions and alternative applications of our modeling approach to vertical contracting. Section 7 offers some concluding remarks. The Appendix contains omitted proofs and additional calculations.

2 The Model

Consider a market with \( N \) competing downstream firms. Each firm \( n \in \{1, \ldots, N\} \) operates at constant marginal cost \( c_n \). Each firm also requires the use of an upstream input. Although our results extend to any technology that uses the input in fixed proportions, we assume for simplicity that one unit of the upstream good can be transformed into one unit of the downstream good. This would apply, in particular, to the case of retailing.

We follow the extant literature on wholesale price determination and stipulate that, at least in equilibrium, one (incumbent) supplier will serve all \( N \) downstream firms. We differ, however, from much of this literature in allowing the firms to turn to other sources of supply, as opposed to simply exiting the market, if they dislike the incumbent’s terms.\(^6\)

The incumbent supplier operates at constant marginal cost \( m \), which, to simplify expressions, we normalize to zero: \( m = 0 \). We assume, however, that downstream firms also have access to alternative sources of supply. For instance, retailers may list different products or manufacturers may procure the same or a similar input elsewhere. Abstracting from other (quality) differences between the different inputs, it is convenient to suppose that the incumbent has only a cost advantage. We thus suppose that downstream firms can alternatively procure the required input at constant marginal cost \( \hat{m} > m = 0 \).\(^7\) For instance, the alternative input might be procured competitively in the “world market.”\(^8\)

We assume the incumbent supplier can make take-it-or-leave-it offers. Each offer specifies a pair \( (t_n, w_n) \), which consists of a fixed transfer \( t_n \) together with a constant marginal

\(^6\)In that sense, we follow Katz (1987), O’Brien (1989, 2011), and Inderst and Valetti (2009a) in explicitly modeling the threat of demand-side substitution. However, in contrast to these other papers, which assume linear supply contracts, we allow for non-linear contracting between the upstream and downstream firms.

\(^7\)One can thus think of the difference \( \Delta m \equiv \hat{m} - m \) as a measure of the (market) power that the supplier has due to its superior product or technology. Intuitively, the smaller is this difference, provided it is not too large in the first place, the less the supplier will be able to extract from each of the downstream firms.

\(^8\)Accessing the alternative input might require that firms incur additional fixed costs, \( F \geq 0 \). As will become clear, however, in contrast to the models in Katz (1987) and Inderst and Valetti (2009a), which restrict attention to linear contracts, these additional fixed costs would not affect the marginal wholesale prices and thus final prices in the model. To simplify the expressions, we have thus implicitly set \( F = 0 \).
wholesale price \( w_n \). Let \( k_n = c_n + w_n \) denote downstream firm \( n \)'s marginal cost of operating if it accepts the offer, and let \( \hat{k}_n = c_n + \hat{m} \) denote downstream firm \( n \)'s marginal cost of operating if it rejects the offer and instead chooses the alternative supply option.

The offers, as well as the accept or reject decisions, are observed by all market participants. After the offers are made, and after they are accepted or rejected, the downstream firms compete by choosing their final prices \( p_n \) according to the demand each faces for its respective good. Sales are then made and payoffs are realized for all market participants. Note that our set-up thus does not allow for the possibility of renegotiation if an offer is rejected. However, as we argue in Section 6 below, our results would be unchanged if renegotiation were allowed and the supplier could make a new offer not only to those buyers who had accepted their offers but also to those buyers who had rejected their offers.

To summarize, the timing of moves is as follows. First, the incumbent supplier makes its offers \( (t_n, w_n) \). Second, offers are accepted or rejected. If a downstream firm rejects its offer, it can turn to its alternative supply option and procure the input at constant marginal cost \( \hat{m} > 0 \). If it accepts its offer, it can procure the input according to the terms of its accepted contract. Third, the downstream firms compete by setting final prices \( p_n \).

### 3 Persistence of Downstream Competition

Let \( \mathbf{k} \) denote the vector of downstream marginal costs when all offers are accepted, where the \( n \)th component of \( \mathbf{k} \) is given by \( k_n = c_n + w_n \). We assume for convenience that, for given \( \mathbf{k} \), there is a unique price equilibrium in pure strategies in the downstream market.

Denote the respective equilibrium final prices by \( p_n(\mathbf{k}) \), the vector of final prices by \( \mathbf{p}(\mathbf{k}) \), and the respective demands that follow from these prices by \( q_n(\mathbf{p}) \). Then, the respective downstream profits \textit{gross of the fixed transfers} \((t_1, \ldots, t_n)\) are given by

\[
\pi_n(\mathbf{k}) = q_n(\mathbf{p}) [p_n(\mathbf{k}) - k_n],
\]

and thus the sum of the upstream and downstream profits when all offers are accepted is

\[
\Omega(\mathbf{k}) = \sum_{n=1}^{N} [\pi_n(\mathbf{k}) + q_n(\mathbf{p}) w_n]
\]

\[
= \sum_{n=1}^{N} q_n(\mathbf{p}) [p_n(\mathbf{k}) - c_n].
\]

We will refer to \( \Omega(\mathbf{k}) \) as the total industry surplus, or equivalently, as the industry profit.
**Unconstrained Monopolist.** Suppose for a moment that the supplier is an *unconstrained* monopolist, so that downstream firms have no outside option (or, equivalently, have an outside option of value zero). Then, through the respective choice of the fixed parts $t_n$, the supplier’s take-it-or-leave-it offers would extract all industry profits, $\Omega(k)$. As a consequence, the supplier’s optimal choices of $w_1, \ldots, w_n$ would maximize $\Omega(k)$. That is, they would induce the downstream firms to fully internalize the effect of the final goods’ prices on $\Omega(k)$. The outcome in this case would thus be the same as what a vertically-integrated firm would obtain if it were to control both the input production and the sale of the final goods at all $N$ downstream firms (cf. Inderst and Shaffer (2009) for details).

**Optimal Wholesale Prices of a Constrained Supplier.** Now suppose that the downstream firms do in fact have a viable outside option. In this case, the incumbent supplier will be constrained by the threat of demand-side substitution, which in turn will constrain its optimal choice of contracts. In the event a single downstream firm $n$ were to reject the supplier’s offer and turn instead to its alternative supply option, downstream equilibrium prices and profits would no longer be determined using the marginal operating costs $k_n = c_n + w_n$ for firm $n$, but, instead, would be determined using $\hat{k}_n = c_n + \hat{m}$. In a slight abuse of notation, we let $\pi_n(\hat{k}_n)$ denote the (off-equilibrium) profits of firm $n$ when firm $n$ is the only firm that rejects the incumbent’s offer. Here, $\hat{k}_n$ is the relevant *vector* of downstream marginal costs, where $k_n' = c_n' + w_n'$ for all firms $n' \neq n$ and $\hat{k}_n$ for firm $n$.

To solve for the equilibrium contracts in this case, we proceed in two steps. We first derive the respective fixed parts $t_n$. We then turn to the derivation of the marginal wholesale prices $w_n$. The respective downstream firm accepts the supplier’s offer only if

$$\pi_n(k) - t_n \geq \pi_n(\hat{k}_n).$$

By optimality for the supplier, this participation constraint will be binding, which yields

$$t_n(k) = \pi_n(k) - \pi_n(\hat{k}_n).$$ (1)

Using the expression for $t_n(k)$ in (1), the supplier’s total profits are thus

$$\Pi = \sum_{n=1}^{N} [t_n(k) + q_n(p)w_n]$$

$$= \Omega(k) - \sum_{n=1}^{N} \pi_n(\hat{k}_n).$$

7
Differentiating $\Pi$ with respect to $w_n$, and using the fact that $w_n$ affects not only $\Omega(k)$ but also the outside options of all firms other than firm $n$, we thus have the constrained incumbent supplier’s respective first-order condition for the optimal wholesale price $w_n$:

$$
\frac{d}{dw_n} \Omega(k) = \frac{d}{dw_n} \sum_{n' \neq n} \pi_{n'}(\hat{k}_{n'}).
$$

(2)

It is convenient to suppose that the supplier’s maximization problem is strictly quasi concave and that, in equilibrium, all downstream firms are active. Then, the system of first-order conditions in (2) pins down a unique set of wholesale prices at which $q_n > 0$.

The choice of wholesale prices maximizes total industry surplus if and only if the right-hand side of (2) is zero. However, the right-hand side of (2) will not in general be zero when the downstream firms’ demands interact because in that case the marginal operating cost of firm $n$, i.e., $k_n$, would affect the (off-equilibrium) profits of all other firms $n'$. This follows because $w_n$, and hence $k_n$, affects the price $p_n$ set by firm $n$. In this paper, we focus on the case where the firms’ products are substitutes, so that for all $q_n > 0$ and $\pi_{n'} > 0$,  

$$
\frac{d\pi_{n'}}{dk_n} > 0 \text{ for all } n' \neq n.
$$

(3)

That is, all else equal, firms gain when their rivals have higher marginal operating costs.$^9$

In the presence of binding outside options for downstream firms, the supplier’s respective choice of each marginal wholesale price $w_n$ thus trades off two conflicting objectives: to maximize total industry surplus, given that the supplier becomes the “residual claimant” through his choice of $t_n$, and to minimize the value of each downstream firm’s outside option. These objectives correspond to the left and right-hand sides of (2), respectively.

If $w_n$ were chosen by the supplier so as to maximize total industry surplus, for a given choice of $w_{n'}$ for all other firms, a marginal change of $w_n$ would have a zero first-order effect on industry profits. On the other hand, from (3), a marginal decrease in $w_n$, and thus in $k_n = c_n + w_n$, would have a strictly negative first-order effect on the outside option of all other firms $n'$. Thus, it follows straightforwardly that, despite the presence of non-linear contracts, a constrained supplier will no longer set its marginal wholesale prices so as to fully monopolize the downstream market. We have thus arrived at the following result.

Proposition 1 When the supplier’s choice of nonlinear contracts is constrained by the outside options of the downstream firms, the optimal marginal wholesale prices $w_n$ satisfy

$^9$Under standard regularity conditions, this follows when the cross-price effect is positive (substitutes).
the first-order condition in (2). It follows that, for given \( w_{n'} \), the choice of \( w_n \) is strictly lower compared to the case in which the downstream firms have no valuable outside option.

Relative to the case of unconstrained monopoly, the supplier benefits from a reduction in \( w_n \) because this reduces the outside option value of all other downstream firms \( n' \neq n \).

In what follows, we will use the characterization of wholesale prices in (2) to obtain additional implications. In Section 4 we obtain comparative-statics results across downstream firms with different cost efficiencies and thus different sizes. In Section 5 we analyze how competition in the upstream and downstream markets affect wholesale and final prices.

4 Heterogeneous Downstream Firms

We now analyze how marginal wholesale prices are affected by cost differences among firms. Although our qualitative results extend readily to an arbitrary number of downstream firms, we facilitate exposition by setting \( N = 2 \) and stipulating that demand be symmetric (i.e., \( q_n = q(p_n, p_{n'}) \) for \( n' \neq n \)). Suppose then, without loss of generality, that \( c_1 < c_2 \).

To proceed, we begin by writing out the left-hand side of condition (2) more explicitly. Once again, it is convenient to consider first the case of an unconstrained monopolist.

**Unconstrained Monopolist.** Using the downstream firms’ first-order conditions with respect to \( p_n \), which imply \( \frac{\partial \pi_1}{\partial p_1} = \frac{\partial \pi_2}{\partial p_2} = 0 \), the condition \( d \Omega/d w_1 = 0 \) can be written as

\[
\frac{\partial \pi_1}{\partial p_2} \frac{dp_2}{dw_1} + \frac{\partial \pi_2}{\partial p_1} \frac{dp_1}{dw_1} + w_1 \left[ \frac{\partial q_1}{\partial p_1} \frac{dp_1}{dw_1} + \frac{\partial q_1}{\partial p_2} \frac{dp_2}{dw_1} \right] + w_2 \left[ \frac{\partial q_2}{\partial p_1} \frac{dp_1}{dw_1} + \frac{\partial q_2}{\partial p_2} \frac{dp_2}{dw_1} \right] = 0, \tag{4}
\]

where the first two terms in (4) correspond to the indirect effect on the downstream firms’ profits from an increase in \( w_1 \), and the last two terms correspond to the indirect effect on the supplier’s profit from an increase in \( w_1 \).\(^{10}\) The condition \( d \Omega/d w_2 = 0 \) is symmetric.

Simplifying (4) and combining the respective conditions for \( n = 1, 2 \) yields

\[
\begin{aligned}
\left[ \frac{dp_1}{dw_1} - \frac{dp_1}{dw_2} \right] \left[ (p_2 - c_2) \frac{\partial q_2}{\partial p_1} + w_1 \frac{\partial q_1}{\partial p_1} \right] &= \left[ \frac{dp_2}{dw_2} - \frac{dp_2}{dw_1} \right] \left[ (p_1 - c_1) \frac{\partial q_1}{\partial p_2} + w_2 \frac{\partial q_2}{\partial p_2} \right], \tag{5}
\end{aligned}
\]

\(^{10}\)The direct effect of an increase in \( w_1 \) is a pure transfer from downstream firm 1 to the supplier.
Condition (5) must hold if the supplier is choosing $w_1$ and $w_2$ to maximize total industry surplus. The first term in brackets on both sides of (5) captures the difference in the pass-through to a firm’s final price of a change in a firm’s own marginal wholesale price compared to a change in its rival’s marginal wholesale price. Although these terms would be the same under symmetry (i.e., when evaluated at the same quantities and final prices), with cost differences, they will generally not be the same. This follows because the quantities sold and the prices set may in general differ. As a result, without some restrictions on demand, it is not possible to determine \textit{a priori} for which firm the difference will be larger.

To obtain further results, and to isolate the new effects we wish to examine, we now turn our attention to the case of linear demands, where we can use the fact that the respective marginal effects on both sides of (5) are independent of the realized demand. This means that the pass through effects will be the same for both firms, regardless of the different quantities sold, as will the own-price and cross-price effects, which we denote by

$$\frac{\partial q_n}{\partial p_n} = \beta < 0 \quad \text{and} \quad \frac{\partial q_n}{\partial p_n'} = \gamma > 0.$$ 

With this simplification, condition (5) can thus be rewritten as

$$(w_1 - w_2)(\beta - \gamma) = \gamma [(p_1 - k_1) - (p_2 - k_2)],$$

(6)

which implies that the difference in wholesale prices must be inversely proportional to the difference in markups. It follows that the more efficient firm ($n = 1$), which will have the larger markup in equilibrium, must be given a strictly lower marginal wholesale price ($w_1 < w_2$) if total industry surplus is to be maximized (cf. also Inderst and Shaffer 2009).

\textbf{Reducing Outside Options.} When downstream firms have access to valuable outside options, the \textit{opposite} prediction holds. To see this, note that condition (2) implies that in addition to considering the marginal impact of $w_n$ on total industry surplus, the incumbent supplier must also consider the marginal impact of $w_n$ on the outside option of firm $n'$.

We have already shown that in order to reduce this outside option, the supplier has an incentive to lower $w_n$ below what it would charge if it were an unconstrained monopolist.

\footnote{To see this, suppose that $k_n < k_n'$, so that firm $n$’s operating cost is lower. Then, firm $n$ will have the lower price in equilibrium, $p_n < p_{n'}$, which implies that it will have the larger markup $p_n - k_n > p_{n'} - k_{n'}$ (this follows from the first-order condition $(p_n - k_n)\frac{\partial q_n}{\partial p_n} + q_n = 0$, where $\frac{\partial q_n}{\partial p_n} = \beta$). In particular, when $w_n = w$, so that, together with $c_1 < c_2$, this implies $k_1 < k_2$, the right-hand side of (6) would be strictly positive, while the left-hand side would be zero. To equalize the two sides, it must hold that $w_1 < w_2$.}
Given the cost differences downstream, however, there is no reason to believe that it would lower its marginal wholesale price equally for both firms. Thus, it is important to consider the effect a given wholesale price reduction would have on the outside options of the two firms, and to ask for which firm this effect would be stronger (note that firm 1’s outside option would decrease by more than firm 2’s outside option if $d\pi_1(\hat{k}_1)/dw_2 > d\pi_2(\hat{k}_2)/dw_1$).

Determining for which firm the “outside option effect” would be stronger turns out to be straightforward once it is recognized that the more efficient firm always has the lower marginal operating cost $k_n$ in equilibrium (i.e., $c_1 < c_2$ implies $k_1 < k_2$ in equilibrium). It follows from this that the effect will be stronger for the more efficient firm if and only if it is also stronger for the firm with the lower marginal operating cost, which holds when\footnote{Note that (7) is calculated at arbitrary values $k_n$ and $k_{n'}$, not just at equilibrium values.}

$$\frac{d^2\pi_n}{dk_n dk_{n'}} < 0 \text{ for } n' \neq n. \tag{7}$$

Condition (7) is commonly used in the literature and is satisfied by many functional specifications (cf. Athey and Schmutzler, 2001), including the case of linear demands. When it holds, a reduction of the rival firm’s marginal operating cost will have a larger negative effect on a firm’s profit when the firm itself has a lower marginal operating cost (i.e., when it holds, the reduction in the more efficient firm’s outside option will be greater).

Intuitively, when $k_n$ is smaller, making firm $n$ more competitive, firm $n$ sells at a higher per-unit markup (cf. our previous discussion). It follows that when the firm’s rival becomes more competitive following a reduction of its own marginal operating cost, the resulting decrease in firm $n$’s demand is more costly in terms of lost profits, compared to a situation in which firm $n$ has a higher marginal operating cost and thus a lower per-unit markup. This suggests that the lower wholesale price should go to the less efficient firm.

**Optimal Discriminatory Wholesale Prices.** We now explore the implications of the discussion above and consider which is stronger, the incentive that arises from (6) (which suggests that a lower wholesale price should be given to the more efficient firm) or the incentive that arises from (7) (which, assuming it holds, suggests that a lower wholesale price should be given to the less efficient firm). In general, it is not possible to say which one will dominate. However, it turns out that in the case of linear demands, the tradeoff is always resolved in favor of the latter: the optimal marginal wholesale price will be higher for the more efficient firm (i.e., $w_1 > w_2$) when the incumbent supplier is constrained.
Proposition 2  Consider the optimal non-linear contract of a supplier that is constrained by the downstream firms’ option of switching to a different source of supply. On the one hand, the supplier has an incentive to give the more efficient downstream firm a lower marginal wholesale price in order to increase total industry surplus, which can then be extracted via the fixed parts of its contract. On the other hand, when condition (7) is satisfied, the supplier also has an incentive to do the opposite, as giving an advantage to a less efficient firm has a larger negative effect on the outside option of its rival. In the case of linear demands, the latter incentive is sufficiently strong that marginal wholesale prices are always higher in equilibrium for the more efficient and thus larger downstream firm.

Proof. See Appendix.

With linear demands, the effect that we newly identify in this paper is sufficiently strong that it makes the supplier want to handicap the more efficient firm. This reduces total industry surplus. To see why, recall our observation that a firm with a lower marginal cost of operating, \( k_n \), will have a strictly higher markup, \( p_n - k_n \), in equilibrium. This means that the more efficient firm’s cost advantage will not adequately be reflected in a reduction in prices, and thus a shift in sales at the margin from the less efficient firm to the more efficient firm will increase industry surplus.\(^{13}\) It follows that, all else equal, industry surplus will be larger if the less efficient firm receives a higher marginal wholesale price. This is indeed the outcome when the supplier is unconstrained. Instead, by setting \( w_1 > w_2 \), the constrained supplier makes the allocative inefficiency worse. Although total industry surplus is thereby reduced, the supplier obtains a larger overall share of the profit.

Note that Proposition 2 focuses on the determination of the marginal wholesale price \( w_n \) because the choice of \( t_n \) is not relevant for downstream competition and prices. However, \( w_n \) should be distinguished from the average wholesale price \( W_n \), where \( W_n \) is given by

\[
W_n = w_n + \frac{t_n}{q_n}. \tag{8}
\]

Although the less efficient downstream firm obtains a lower marginal price when demand is linear, this does not imply that it will be able to purchase its input from the incumbent supplier at overall more favorable conditions (i.e., it should not be taken to suggest that

\(^{13}\)Note that we consider symmetric demand, which allows us to abstract from differences in demand elasticities when prices are symmetric.
\( W_2 < W_1 \) necessarily follows from \( w_2 < w_1 \) because this need not be true in general.\(^{14}\) Rather, as noted in the Introduction, a main contribution of our analysis is to show how a supplier that has some scope to optimally manage different distribution channels can instead gain by making smaller, less efficient buyers more competitive “at the margin”. We will return to a comparison of the average and marginal wholesale prices in Section 6.

## 5 Changing Competition

We now explore the model’s insights and predictions with respect to how varying the degree of competition in the downstream market affects equilibrium prices, and with respect to how these prices will be affected by the closeness of substitutes to the supplier’s input.

### 5.1 Hotelling Competition

Consider first a model of Hotelling competition with two firms. We assume demand is uniformly distributed over the unit interval with the two firms located at the extremes. Each consumer has valuation \( v > 0 \) and transportation costs are denoted by \( \tau > 0 \). The demand function when all consumers buy and both firms sell positive quantities is then

\[
q_n = \frac{1}{2} - \frac{1}{2\tau}(p_n - p_{n'}).
\]

Note that the properties of \( q_n \) are standard. It is decreasing in firm \( n \)’s own price, increasing in firm \( n' \)’s price, and equal to 1/2 when the final goods’ prices are the same (in that case, each firm has 50% of the market). Substituting this demand function into the first-order conditions from the proof of Proposition 2, we obtain the marginal wholesale prices\(^{15}\)

\[
w_n = \hat{m} + \frac{1}{2} \left[ 3\tau - \frac{1}{3}(c_n - c_{n'}) \right],
\]

and the resulting induced final prices\(^{16}\)

\[
p_n = \hat{m} + \frac{5}{2} \tau + c_n - \frac{7}{18}(c_n - c_m).
\]

\(^{14}\)It is also important to note that our set-up abstracts from other sources of size-related advantages in procuring input. For instance, Katz (1987) and Inderst and Valletti (2009a) show that lump-sum switching costs can give rise to size advantages because larger buyers can distribute these costs over a larger volume.

\(^{15}\)Note that for this characterization to apply, the consumers’ utility from purchasing the product must be sufficiently high that, under the characterized equilibrium prices, the market is indeed fully covered.

\(^{16}\)Details of the calculations for the Hotelling model can be found in Appendix B.
The characterization of optimal prices confirms our previous results. The minus sign in front of the term \( \frac{1}{3}(c_n - c_{n'}) \) in the expression for \( w_n \) implies that the more efficient downstream firm does indeed receive a higher marginal wholesale price from the supplier (c.f. Proposition 2). And it is straightforward to show that for all \( \hat{m} \) such that the market is covered and both firms have positive market shares, marginal wholesale prices are indeed lower than what an unconstrained monopolist supplier would charge (c.f. Proposition 1).

The expressions in (10) and (11) also relate the wholesale and final prices to the degree of competition \( \tau \) in the downstream market and the constraint \( \hat{m} \) that the supplier faces.

**Proposition 3** With Hotelling competition, when the market is covered and both firms sell positive quantities, the marginal wholesale price \( w_n \) and final price \( p_n \) are increasing in \( \tau \) and \( \hat{m} \). Thus, for example, if \( \tau \) decreases (i.e., if downstream competition were to become more intense), or if \( \hat{m} \) decreases (i.e., if the cost of obtaining the input elsewhere were to decrease), the supplier would respond by offering a lower marginal wholesale price.

The intuition for the result on \( \tau \) is straightforward. The supplier’s ability to reduce the value of each firm’s outside option becomes easier as the degree of downstream competition increases (note that in the absence of competition, it is not possible for \( w_n \) to affect the value of firm \( n' \)’s outside option), causing it to place more weight on this objective and less weight on maximizing industry profits when choosing its wholesale prices. As a result, the model predicts that wholesale prices would be expected to decrease in the face of increasing competition downstream. And final prices would also be expected to decrease in the face of increasing competition, as the decrease in wholesale prices would reinforce the direct increase in competitive pressure on final prices that would already result from a lower \( \tau \).

The intuition for the result on \( \hat{m} \) is perhaps less immediate. After all, it is not obvious why the supplier, when choosing its contracts \( (t_n, w_n) \), would be expected to place less weight on maximizing industry profits when the threat of demand-side substitution increases. Why would it not just reduce the fixed component \( t_n \) in its contracts, thereby paying the downstream firms directly for the increase in the value of their outside option?

The key to understanding it comes from condition (7), which holds with linear demands, and thus, in particular, for the Hotelling model. Note first that a downstream firm which operates under its outside option becomes more competitive when its respective marginal cost, \( \hat{m} + c \), decreases. From (7), such a deviating firm, which rejects the supplier’s offer,
is hurt more when the supplier lowers the marginal wholesale price, \( w_{n'} = w \), of the firm’s rival. Consequently, once \( \hat{m} \) decreases, the (negative) effect that a marginal decrease of \( w \) has on the outside option of all downstream firms is larger. Given the trade-off in the first-order condition (2) for \( w_n \), this induces the supplier to place more weight on the reduction of the firms’ outside options and less on the maximization of industry profits. As a result, the model predicts that the incumbent’s wholesale prices would be expected to decrease in the face of an increasingly attractive alternative source of supply—even though the outside option is not used in equilibrium and even though the incumbent uses nonlinear contracts.

Taken together, our setting with non-linear contracts and a constrained incumbent supplier, when applied to the workhorse model of Hotelling competition, thus has the potentially appealing feature that supply-side constraints (as expressed by a change in \( \hat{m} \)) as well as demand-side constraints (as expressed by a change in \( \tau \)) affect both wholesale prices and final prices. In particular, wholesale and final prices decrease as either downstream firms or the supplier become less differentiated compared to their respective competitors.

Finally, note that with symmetry, the expressions in Proposition 3 reduce to

\[
 w_n = w = \hat{m} + \frac{3}{2} \tau
\]

and

\[
 p_n = p = \hat{m} + \frac{5}{2} \tau + c.
\]

As is well known, the margin for each downstream firm, \( p - c - w \), is then simply \( \tau \). As competition increases, both the supplier’s and the downstream firms’ margins decrease.

**Comparison to Benchmarks.** To conclude our analysis of the Hotelling model, we compare the explicit characterization of \( w_n \) and \( p_n \) in Proposition 3 to the two benchmark cases that we discussed in the Introduction (recall that in the first benchmark, contracts are unobservable, whereas in the second benchmark, the incumbent supplier is unconstrained).

Regarding the first benchmark case, there can be an extreme opportunism problem when contracts are non-linear and unobservable. In the commonly assumed case of “passive beliefs” (cf. McAfee and Schwartz, 1994; Rey and Verge, 2004), for example, wholesale prices equal the supplier’s marginal cost, which in our setting implies \( w_n = 0 \).\(^{17}\) In this case, final goods’ prices strictly decrease when \( \tau \) decreases, while marginal wholesale prices are

\(^{17}\)Marginal-cost pricing also arises in the contract equilibrium setting in O’Brien and Shaffer (1992).
unaffected. Note that marginal wholesale prices are also unaffected when competition from another source of supply increases, as the incumbent supplier is its own worst competitor.

In the second benchmark case, $w_n$ is chosen to maximize total industry surplus $\Omega(w)$. In this case, if the market is fully covered, we obtain the (monopolizing) wholesale prices\textsuperscript{18}

$$w_n^{Mon} = v - \frac{3}{2} \tau - \frac{1}{4} (c_n + 3c_{n'})$$

and the resulting induced final prices

$$p_n^{Mon} = v - \frac{1}{2} \tau + \frac{1}{4} (c_n - c_{n'}).$$

In this case, the difference in the marginal wholesale prices between firms 1 and 2 is $(c_1 - c_2)/2$, which is strictly negative when $c_1 < c_2$ (firm 1 is more efficient), while the respective difference was strictly positive for the constrained supplier (cf. expression (10)).

Further, as transportation costs decrease, making the downstream firms closer substitutes, both wholesale and final prices increase. Again, this is the opposite of what we obtain for the constrained supplier (cf. Proposition 3). Admittedly, with an unconstrained supplier and Hotelling competition, the fact that $p_n^{Mon}$ is inversely related to $\tau$ is somewhat mechanic, provided that the market remains covered. Then, as $\tau$ increases, to ensure that all consumers will continue to buy the product, final prices must fall, which in turn requires that the firms’ marginal wholesale prices $w_n^{Mon}$ must fall as well. However, we show next that even when total demand is elastic, the case of an unconstrained supplier yields opposite predictions compared to those of our model where the supplier is constrained.

5.2 Elastic Total Demand

We next consider a linear demand system in which total demand is elastic (unlike in the Hotelling model where demand was assumed to be inelastic) and $N \geq 2$. In particular, using the symmetric, linear-quadratic specification of the representative consumer’s utility that is found in Shubik and Levitan (1980), we obtain for good $n$ the demand function

$$q_n = \frac{1}{N} \left[ 1 - p_n - \theta \left( p_n - \frac{\sum_{n' \in N} p_{n'}}{N} \right) \right]$$

\textsuperscript{18}To obtain these expressions, it is convenient to solve first for the retail prices that maximize $q_1(p_1 - c_1) + q_2(p_2 - c_2)$, where $q_n$ is given by (9). (The market is covered by at least one product when $p_1 + p_2 \leq 2v - \tau$.) This yields $p_n^{Mon}$. Requiring then that these prices are obtained from $p_n = \tau + (2k_n + k_{n'})/3$ yields $w_n^{Mon}$. 

16
and the indirect demand function

\[ p_n = 1 - \frac{N + \theta}{1 + \theta} q_n - \frac{\theta}{1 + \theta} \sum_{n' \neq n} q_{n'} \]

This demand system has the attractive property that when changing the degree of substitution, \( \theta \), the sum of the individual quantities does not change when firms charge symmetric prices.\(^{19}\) The same is true when the number of firms \( N \) increases. Hence, the degree of competition along both dimensions can be varied without affecting the size of the market.

In order to obtain explicit expressions when there are more than two downstream firms, we focus on the symmetric case: \( c_n = c \). And, to make production profitable, we stipulate that \( c < 1 \).\(^{20}\) Given these restrictions, we derive in the proof of Proposition 4 an explicit characterization of the equilibrium final prices and profits as a function of the marginal wholesale prices. It is useful to note that in the symmetric case with \( w_n = w \), we obtain

\[ p_n = p = \frac{1}{2 + \theta - \frac{\theta}{N}} \left[ 1 + (c + w) \left( 1 + \theta - \frac{1}{N} \theta \right) \right] \quad (13) \]

and

\[ \pi_n = \pi = \frac{1}{N} \left[ \frac{1 + \theta - \frac{\theta}{N}}{2 + \theta - \frac{1}{N} \theta} \right] (1 - c - w)^2. \quad (14) \]

For a given marginal wholesale price \( w \), it follows that both prices and profits strictly decrease as either \( N \) or \( \theta \) increase. Further, it follows that, \( p \to c + w \) as \( \theta \to \infty \). We analyze below how these prices change once we take into account the adjustment of \( w \).\(^{21}\)

**Benchmarks.** Before doing so, however, it is convenient to first set the stage by deriving the comparative statics for the two respective benchmarks. Regarding the first benchmark, the case of unobservable contracts, \( w_n = 0 \) once again when beliefs are passive, so that final prices strictly decrease when goods become less differentiated (higher \( \theta \)) or when the number of firms increases (higher \( N \)), while the marginal wholesale prices are not affected.

Regarding the second benchmark, an unconstrained supplier will optimally choose \( w_n = w \) to maximize total industry surplus. This yields the (monopolizing) wholesale prices

\[ w^{Mon} = \frac{1}{2} \theta \left( \frac{N - 1}{N} \right) \frac{1}{1 + \theta - \frac{1}{N} \theta} (1 + c) \]

\(^{19}\)Goods are independent if \( \theta = 0 \) and become increasingly substitutable as \( \theta \) increases.

\(^{20}\)This restriction on the downstream firms’ marginal costs is needed because the overall market size is equal to one and we have stipulated, for convenience only, that the supplier’s own marginal cost is zero.

\(^{21}\)Incidentally, with the demand system (12), for a given marginal wholesale price \( w \), the symmetric final price \( p \) does not converge to \( c + w \) as \( N \to \infty \), although clearly downstream profits satisfy \( \pi \to 0 \).
and the resulting induced final prices
\[ p^{Mon} = \frac{1}{2} (1 + c). \]

As can be seen from these expressions, the monopoly price \( p^{Mon} \) is independent of downstream competition (and thus does not vary with either \( \theta \) or \( N \)).\(^{22}\) This outcome holds because, as in the Hotelling case, the marginal wholesale price strictly increases when downstream competition intensifies (i.e., \( w^{Mon} \) strictly increases when either \( \theta \) or \( N \) increase). In the case of independent downstream firms (\( \theta = 0 \)), the supplier would choose \( w^{Mon} = 0 \). In the case of perfect competition downstream (\( \theta = \infty \)), the supplier would choose \( w^{Mon} = p^{Mon} \). For intermediate levels of competition, the supplier would choose \( w^{Mon} \) between 0 and \( p^{Mon} \), compensating for higher intra-brand competition in the downstream market by pushing up the marginal wholesale price in order to dampen competition.

**Characterization.** Making use of the first-order conditions in (2), the constrained supplier’s profit-maximization problem can be solved for the optimal marginal wholesale price \( w_n = w. \)^{23} This price can then be substituted into (13) to obtain the symmetric equilibrium final goods’ price \( p_n = p. \) Doing so yields the following comparative static results.

**Proposition 4** With Shubik-Levitan demands, when the downstream firms are equally efficient, \( c_n = c \), the marginal wholesale and final prices have the following properties:

- The marginal wholesale price \( w \) and final price \( p \) are increasing in \( \hat{m} \). If the cost of the alternative supply were to increase (decrease), the supplier would respond by raising (lowering) its wholesale price, thereby inducing a higher (lower) final price \( p \).

- The final price \( p \) is decreasing in \( \theta \) and \( N \). Final prices decrease (increase) as the downstream market becomes more (less) competitive, either because the final goods become less (more) differentiated or the number of competitors increases (decreases).

**Proof.** See Appendix.

The comparative static results in Proposition 4 follow from the same logic as in the previously analyzed model with Hotelling competition. The only difference is that here the

---

\(^{22}\)The calculations are straightforward and obtained from the expressions in the proof of Proposition 4.

\(^{23}\)We relegate an explicit characterization of the expression for \( w_n \) to the proof of Proposition 4.
market demand is elastic, whereas in the Hotelling model, the market demand is inelastic. Nevertheless, in both models, prices in the marketplace decrease both as the value of the outside option increases (the supplier becomes more constrained) or as the downstream market becomes more competitive (final goods become more substitutable or the number of firms increases). These results hold even though contracts are observable and nonlinear.

We can push the results in the Shubik-Levitan case further by noting that in the model \( \lim_{\theta \to \infty} p = \hat{m} + c \).

That is, as goods become essentially undifferentiated, final prices converge to the downstream firms’ marginal operating costs under the alternative supply option. Once again, we note that this result holds despite the ability of the supplier to use non-linear contracts. Moreover, as the supplier also becomes less differentiated (i.e., as his cost advantage \( \hat{m} > 0 \) erodes), the outcome becomes perfectly competitive, with \( p = c \).

As noted previously, our set-up allows one to capture constraints on both the upstream and downstream markets and offers intuitively appealing predictions. This has implications for applied work. In terms of competition policy, for example, a reduction of competitive pressure in our set-up, in either the upstream or the downstream market, gives rise to a change in both the input prices and the thereby induced final prices. For instance, a downstream merger would be expected to impact prices in both the upstream and the downstream markets. These predictions contrast with the predictions of the previously discussed benchmark of an unconstrained supplier, whether or not contracts are observable.

6 Discussion

Renegotiations. The core contribution of the paper is to develop a simple setting of vertical contracting in which a supplier uses its marginal wholesale prices strategically to extract a larger share of profits from competing downstream firms. Marginal wholesale prices affect industry profits both on equilibrium and off equilibrium. It is the latter effect which is novel in our contribution and which is behind our comparative static results.

Our model is kept purposely simple — the supplier makes take-it-or-leave-it offers to...
downstream firms — without the possibility of renegotiation. In particular, this embodies the following two assumptions off equilibrium. First, if a downstream firm rejects the supplier’s offer, there is no renegotiation of the contracts that were accepted by the other downstream firms. Second, there is also no attempt by the supplier to make a new offer to a downstream firm that rejected its offer. We now relax both of these assumptions.

When allowing for continuing negotiations after a rejection, one must specify what delay of an individual agreement would mean for production and sales at all firms. In this regard, it seems reasonable to assume that the other downstream firms and even the rejecting buyer would continue to be active in the market while the renegotiations proceed.

To this end, we extend our model by supposing that time runs discretely, \( \tau = 1, 2, \ldots \), and that all players discount future cash flows by the same discount factor \( \delta \). We further assume that firms produce and sell in each period (with per-period payoffs that are captured by our previous notation), and that the following (stage) game unfolds. First, the supplier makes offers \( (t_n, w_n) \) to each of the \( N \) buyers. Second, each buyer accepts or rejects its respective offer (rejection means it purchases from its alternative supplier). Third, each buyer competes (in the current period) under its respective supply arrangement. To rule out implicit collusion, we consider only outcomes which are equilibria of the stage game.

We now show that the possibility to renegotiate contracts does not affect our results. To show this, we proceed in two steps. First, suppose contracts only last for one period. Then, it is clear that, in a stationary equilibrium, our previous characterization fully applies to the outcome in each period. Next, suppose contracts last for multiple periods. Then, at least in principle, when facing a new offer, a buyer who had previously accepted a contract could alternatively compete under the old terms. But this would be inconsequential because, following an (off-equilibrium) rejection by some buyer \( n \), on the equilibrium path of the continuation game, all buyers would expect that buyer \( n \) would accept the same offer in the following round, thereby procuring from the alternative supplier only for a single period.

Note that the possibility to renegotiate contracts would affect our results if the supplier could not make a new offer to a buyer who rejected its contract — because then marginal wholesale prices would no longer serve the purpose of strategically affecting the outside option of other buyers. This can be seen most clearly for the case in which there are only two buyers. In this case, if one buyer were to reject its offer, and by assumption could no longer be approached, the other buyer’s contract would no longer serve the purpose of
reducing the first buyer’s outside option. The supplier and second buyer would instead want to renegotiate their contract so as to give the respective downstream firm a “first-mover” advantage vis-à-vis its competitor (a strategy known as “strategic delegation”). That said, we would expect that, in many instances, a rejecting buyer would in fact be approachable by the supplier even when it temporarily procures from an alternative source.

**Average vs. Marginal Wholesale Prices.** At the end of Section 4, we introduced the concept of a firm’s average wholesale price \( W_n \), as opposed to its marginal wholesale price \( w_n \). One can also distinguish between marginal operating costs \( k_n = c_n + w_n \) and average operating costs \( K_n = c_n + W_n \). With linear demands, we found that a firm with a lower \( c_n \) would pay a higher marginal wholesale price \( w_n \). To complete the analysis, for the Hotelling model (cf. Section 5.1), we now compare the firms’ respective average costs. Before doing so, however, recall that our analysis abstracts from various alternative sources of buyer power, which a downstream firm may lever into better purchasing conditions. Both on and off equilibrium, a more efficient downstream firm procures a larger volume, which may give rise to additional purchasing efficiencies. For instance, a buyer could invest, e.g., through more extended search, to make his outside option more attractive, which is more profitable to do so when it has a larger volume to procure. We abstract from these possibilities, which is why the preceding and subsequent comparative analysis of purchasing conditions is not intended to provide a full-fledged picture of buyer power.

In Appendix B, we show that the more efficient firm pays a higher average wholesale price. However, we also show that the more efficient firm still enjoys both a lower marginal cost of operation, \( k_n = c_n + w_n \), as well as a lower average cost of operation, \( K_n = c_n + W_n \).

**Other Applications.** We have seen how the incentives of a supplier to use marginal wholesale prices to decrease the outside option of competing downstream firms — even at the expense of sacrificing industry profit — distinguishes the case of an unconstrained supplier from that of a constrained supplier. And we have seen how these incentives can give rise to different comparative-static predictions. Moving beyond our focus on wholesale price determination, it would seem the framework we have set out could potentially be applied to other issues in vertical contracting as well. In what follows, we discuss how it might be applied, for example, to shed insight into the profitability of vertical mergers.
Recall the two benchmark cases. When an unconstrained supplier employs observable non-linear contracts, industry profits are maximized via arms-length contracting and there is no need or role for vertical integration. Indeed, partial vertical integration, where the supplier merges with some but not all of the downstream firms, would be counterproductive because transfer prices to the merged entities would then be at cost, which would induce the non-merged firms to be too aggressive in their pricing. In contrast, when contracts are unobservable, supplier opportunism prevents industry profits from being maximized via arms-length contracting. In this case, vertical integration — whether partial or full — can be profitable because it can mitigate the observability problem, resulting in higher prices. This is one of the insights from Hart and Tirole’s (1990) work, and it provides an explanation of how vertical mergers can be privately profitable and yet socially undesirable.

Things are different in our model with a constrained supplier. In the absence of vertical integration, wholesale prices, and the thereby induced final prices, fall short of maximizing industry profits for the reasons we have discussed. In contrast, with full vertical integration, there is no need to worry about the downstream firms’ outside options, and as a result, prices and industry profits will be higher. In this sense, the predictions of the model are similar to those from Hart and Tirole’s (1990) model. But, unlike in Hart and Tirole’s (1990) model, welfare need not be lower. In fact, when total demand is inelastic (cf. the Hotelling case in Section 5.1), welfare is actually higher under full vertical integration because this improves allocative efficiency relative to the case where the constrained supplier’s optimal wholesale prices would handicap the more efficient downstream firm.\(^{26}\)

The case of partial vertical integration is more nuanced. In this case, for example, when the supplier is only integrated with firm 1 but not with its duopolistic competitor, firm 2, the determination of the wholesale price \(w_2\) no longer serves the objective of decreasing firm 1’s outside option. All else being equal, this should push the choice of \(w_2\) up. Further, as is well known, the determination of \(w_2\) now serves the purpose of dampening competition through increasing the vertically integrated firm’s opportunity costs of subsequently choosing a low final price \(p_1\). This is so because firm 1 is now no longer an “agent” of the supplier who can be controlled by an observable contract. Instead, it is always supplied

\(^{26}\)In the Hotelling case, given a price difference \(p_2 - p_1\) and thus a ”critical consumer” location \(\hat{x} = q_1\), aggregate ”shoe leather costs” are given by \(\tau \left[ \int_0^{\hat{x}} x dx + \int_{\hat{x}}^1 (1 - x) dx \right]\), which have to be added to aggregate production costs \(q_1 c_1 + q_2 c_2\). Total costs are minimized when \(\hat{x} = q_1 = 1/2 + (c_2 - c_1)/(2 \tau)\), which is obtained exactly under an integrated monopoly (or, likewise, when \(w_n = w_{n^{Mon}}\), as derived in Section 5.1).
competitively. We leave the analysis of the impact of a partial vertical merger on consumer surplus and welfare, as well as that of the profitability of such a merger, to future work.

7 Concluding Remarks

This paper introduces a framework to model the joint determination of prices in an upstream and downstream market. We assume that an incumbent supplier is constrained by the ability of the downstream firms to switch to an alternative source of supply. Importantly, we allow the supplier to offer non-linear contracts. Because we also assume these contracts are observable, in principle the supplier could use marginal wholesale prices to dampen intra-brand competition so as to thereby maximize total industry surplus. However, our first result is that this is not optimal. By inducing more downstream competition, namely by reducing the marginal wholesale price for all supplied firms, the incumbent supplier can increase its profit by extracting a larger fraction of a smaller industry surplus. This follows because by making each downstream firm more competitive, the profits that would be obtained from switching suppliers are reduced for all other downstream firms.

The focus of our analysis lies in the comparative statics of the determined wholesale and final prices. In a first step, we analyze how our model gives rise to price discrimination in the upstream market. We identify two opposing forces. To generate a larger industry surplus, the supplier should offer more efficient and thus ultimately larger downstream firms a more competitive contract. However, the negative impact on a firm’s outside option when rivals are made more competitive is also larger when the respective firm is more efficient. This suggests that the supplier should offer more efficient firms a less competitive contract. We show that, with linear demands, the second effect dominates. The supplier can gain from partially ironing out the cost disadvantage of the less efficient firms. We discussed how our theory of wholesale price determination could be enriched by other theories of (size-related) buyer power, so as to obtain a richer picture of size-related discounts that is more commensurable with observed practice in different industries.

In an application to linear demand, we explicitly derive comparative statics results in terms of the constraints that firms face in the upstream and the downstream markets. We show how, in contrast to other benchmark models from the literature, our set-up with non-linear contracts and a binding constraint for the supplier generates monotonic comparative statics results that intuitively map such competitive constraints into a change in both
wholesale and final prices. We argue that, with respect to applications to competition policy, this set-up should enrich the modeling choices that are available to both theorists and practitioners. In this respect, we proposed questions for future research, such as the interaction of vertical mergers and the strategic use of wholesale prices to affect the value of downstream firms’ outside options and, thereby, the sharing of total industry surplus.

8 References


Proof of Proposition 2. In a slight abuse of notation, where the terms “with a hat” are understood to relate to prices and costs when \( n' \) rejects the supplier’s offer, we have

\[
\frac{d\pi'_n(\hat{k}'_n)}{dw_n} = -\frac{dp_n}{dw_n} (\hat{p}_n' - c_n' - \hat{m}) \frac{\partial q_n'}{\partial p_n}.
\]

With linear demand, we can express this more simply as

\[
\frac{d\pi'_n(\hat{k}'_n)}{dw_n} = -\frac{dp_n}{dk_n} (\hat{p}_n' - c_n' - \hat{m}) \gamma.
\]

(Note that it is more instructive to express the derivative of the downstream price with respect to our expression of overall marginal cost \( k_n = c_n + w_n \), instead of with respect to \( w_n \).) Together with the derivative of total industry surplus, this yields, from combining the two first-order conditions,

\[
\beta \left[ \frac{dp_n}{dk_n} - \frac{dp_n}{dk_n'} \right] (w_1 - w_2) = \gamma \left[ (p_1 - c_1) - (p_2 - c_2) \right] + \gamma \frac{dp_n}{dk_n} \gamma \left[ (\hat{p}_2 - c_2 - \hat{m}) - (\hat{p}_1 - c_1 - \hat{m}) \right].
\]

The right-hand side of this can be rearranged to read as

\[
\gamma \frac{dp_n}{dk_n'} \left[ (p_1 - c_1) - (p_2 - c_2) \right] + \gamma \frac{dp_n}{dk_n} \left[ (p_1 - \hat{p}_1) - (p_2 - \hat{p}_2) \right].
\]

For linear demand we can further use that

\[
p_n - \hat{p}_n = (w_n - \hat{m}) \frac{dp_n}{dk_n},
\]

so that

\[
(p_1 - \hat{p}_1) - (p_2 - \hat{p}_2) = \frac{dp_n}{dk_n} (w_1 - w_2)
\]

as well as

\[
(p_2 - p_1) = (k_2 - k_1) \left( \frac{dp_n}{dk_n} - \frac{dp_n}{dk_n'} \right).
\]

Then, substituting this into (15), we have the requirement that

\[
(w_1 - w_2) \left[ \beta + \gamma \frac{dp_n}{dk_n'} \left( \frac{dp_n}{dk_n} - \frac{dp_n}{dk_n'} \right) \right] - \gamma \left( \frac{dp_n}{dk_n} \right)^2 = \gamma \frac{dp_n}{dk_n'} (c_2 - c_1) \left[ \left( \frac{dp_n}{dk_n} - \frac{dp_n}{dk_n'} \right) - 1 \right].
\]
We now substitute for the derivatives \( \frac{dp_n}{dk_n} \) and \( \frac{dp_n}{dk_n'} \), which are readily obtained as

\[
\frac{dp_n}{dk_n} = \frac{2\beta^2}{4\beta^2 - \gamma^2} \quad \text{and} \quad \frac{dp_n}{dk_n'} = -\frac{\beta \gamma}{4\beta^2 - \gamma^2}.
\]

With this at hand, we finally obtain after some transformations that

\[
(w_1 - w_2) = (c_2 - c_1) \frac{\gamma^2}{2\beta} \frac{2\beta^2 - \gamma^2}{4\beta^3 - 2\beta^2 \gamma - \gamma^2},
\]

which, from \( \beta < 0 \) and \(-\beta \geq \gamma\), implies \( w_1 > w_2 \) as long as \( c_1 < c_2 \). Q.E.D.

**Proof of Proposition 4.** Denote \( p_\emptyset := \sum_n p_n/N \) and \( w_\emptyset := \sum_n w_n/N \). Recall that \( c_n = c \) in this case. To derive equilibrium final prices, we have to maximize for each \( n \) the respective profits

\[
\pi_n = \frac{1}{N} \left[ 1 - p_n - \theta \left( p_n - \sum_{n'\in N} p_{n'} \right) \right] (p_n - w_n - c).
\]

From the first-order condition we obtain, after making use of the average price \( p_\emptyset \),

\[
p_n = \frac{1 + \theta (p_\emptyset - p_n) + (c + w_n) \left( 1 + \theta - \frac{1}{N} \theta \right)}{2 + \theta - \frac{1}{N} \theta}.
\]

Aggregating over all \( n \) firms, this yields the system of optimal prices

\[
p_n(k) = \frac{1}{2 + \theta - \frac{1}{N} \theta} \left[ 1 + \left( c + w_n \right) \left( 1 + \theta - \frac{1}{N} \theta \right) + (w_\emptyset - w_n) \theta \frac{1 + \theta - \frac{1}{N} \theta}{2 + 2 \theta - \frac{1}{N} \theta} \right]. \tag{17}
\]

To calculate equilibrium profits, we use next from the first-order condition \( d\pi_n/dp_n = 0 \) that, after substituting for \( \frac{dp_n}{dp_n} \),

\[q_n = (p_n - w_n - c) \frac{1}{N} \left[ 1 + \theta \left( 1 - \frac{1}{N} \right) \right].\]

Substituting this back into \( \pi_n \) yields

\[
\pi_n(k) = \frac{1}{N} \frac{1 + \theta - \frac{\theta}{N}}{(2 + \theta - \frac{1}{N} \theta)^2} \left[ 1 - c - w_\emptyset - (\theta + 1) (w_n - w_\emptyset) \frac{2 + \theta - \frac{1}{N} \theta}{2 + 2 \theta - \frac{1}{N} \theta} \right]^2. \tag{18}
\]

In light of subsequent calculations, it is further useful to write this alternatively as

\[
\pi_n(k) = \frac{1}{N} \frac{1 + \theta - \frac{\theta}{N}}{(2 + \theta - \frac{1}{N} \theta)^2} \left[ 1 - c - \frac{w_n}{N} \left( \theta + 1 \right) \frac{(2 + \theta - \frac{1}{N} \theta)}{(2 + 2 \theta - \frac{1}{N} \theta)} (N - 1) + 1 \right] + \frac{\sum w_{n'} \theta \left( 1 + \theta - \frac{1}{N} \theta \right)}{(2 + 2 \theta - \frac{1}{N} \theta)^2}. \tag{19}
\]

\[27\]
We derive next total industry profits, \( Nq(p-c) \), from which we obtain
\[
\frac{d}{dw} \Omega(k) \bigg|_{w_n = w} = \frac{1 + \theta - \frac{1}{N} \theta}{2 + \theta - \frac{1}{N} \theta} (1 + c - 2p). \tag{19}
\]
We can substitute \( p \) from (13).

From (18) the outside option of some firm \( n \), provided that all rivals face \( w_n' = w \), is given, with a slight simplification of notation, by
\[
\hat{\pi}_n = \frac{1}{N} \left[ 1 - c - \frac{\hat{m}}{N} \left( \frac{\theta}{2 + \theta - \frac{1}{N} \theta} \right) \left( \frac{1 - c}{1 - \frac{\theta}{2 + \frac{1}{N} \theta}} \right) + w \frac{N - 1}{N} \frac{1 + \theta - \frac{1}{N} \theta}{2 + 2 \theta - \frac{1}{N} \theta} \right]. \tag{20}
\]
At an interior equilibrium for \( w \), and noting that the program is strictly concave, it holds that
\[
\frac{d}{dw} \Omega(k) \bigg|_{w_n = w} = N \frac{d\hat{\pi}}{dw},
\]
so that, together with (19), we have the requirement that
\[
\left( \frac{1 + \theta - \frac{1}{N} \theta}{2 + \theta - \frac{1}{N} \theta} \right) \left( \frac{\theta N - 1}{N} (1 - c) - 2w \left( \frac{1 + \theta - \frac{1}{N} \theta}{2 + \theta - \frac{1}{N} \theta} \right) \right) = 2 \left( \frac{N - 1}{N} \theta \right) \left( \frac{1 + \theta - \frac{1}{N} \theta}{2 + \theta - \frac{1}{N} \theta} \right)^2.
\]
This finally can be solved for \( w \) to obtain
\[
w = \frac{N - 1}{N^2} \left[ \theta (1 - c) - 2 \hat{m} \frac{1 + \theta - \frac{1}{N} \theta}{2 + 2 \theta - \frac{1}{N} \theta} \left( (\theta + 1) (N - 1) \left( 2 + \frac{N - 1}{N} \theta \right) + (2 + \theta + \frac{N - 1}{N} \theta) \right) \right] \left( \frac{N - 1}{N} \theta \right) \left( 1 + \frac{N - 1}{N} \theta \right) \left[ (2(1 + \theta) - \frac{1}{N} \theta) + \left( \frac{N - 1}{N} \theta \right)^2 \frac{1 + \theta - \frac{1}{N} \theta}{2(1 + \theta) - \frac{1}{N} \theta} \right]. \tag{21}
\]
While the comparative result in \( \hat{m} \) follows immediately, substituting \( w \) to obtain \( p \) and then differentiating with respect to \( \theta \) or \( N \) turns out to be too unwieldy. We, therefore, proceed somewhat indirectly.

For this we return to the supplier’s objective function, but suppose now that the control variable is the equilibrium final price \( p \). Note that any given \( p \) is supported by a wholesale price of
\[
w = \frac{p(2 + \theta - \frac{1}{N} \theta) - 1}{1 + \theta - \frac{1}{N} \theta} - c. \tag{22}
\]
Instead of (20), we thus consider now the first-order condition

\[
\frac{d}{dp} \Omega = N \frac{d\hat{\pi}}{dw} \frac{dw}{dp},
\]

(23)

where we have somewhat abbreviated the notation. Note now first that the derivative of industry profits on the left-hand side of (23) equals \(1 - 2p + c\), which is independent of \(\theta\). A sufficient condition for the assertion to hold is thus from (23) that for any \(p\) the right-hand side is strictly increasing in \(\theta\).

After substituting for \(w\) from (22), we obtain for the right-hand side of (23), after some further transformation, the expression

\[
\frac{2(N-1)}{N} \theta \left( \frac{1 + \theta - \frac{1}{N} \theta}{(2 + \theta - \frac{1}{N} \theta)(2 + 2\theta - \frac{1}{N} \theta)} \right).
\]

(24)

In the rest of the proof, we show that (24) is indeed strictly increasing in \(\theta\). For this we provide sufficient conditions by analyzing the different terms in (24) in turn. Note first that the term in rectangular brackets is strictly positive. To see next that the multiplier, which comprises the first line in (24), is increasing in \(\theta\), note that, using a further transformation,

\[
\frac{d}{d\theta} \left( \frac{1/\theta + 1 - \frac{1}{N}}{2/\theta + 1 - \frac{1}{N}} \right) > 0 \quad \text{and} \quad \frac{d}{d\theta} \left( \frac{1}{2/\theta + 2 - \frac{1}{N}} \right) > 0.
\]

We turn next to the term in rectangular brackets in (24). To show monotonicity in \(\theta\), we clearly have to deal only with the first line. This can be further transformed to

\[
c \left( \frac{\theta}{2 + 2\theta - \frac{1}{N} \theta} \right) - \frac{N - 1}{N} \hat{m} \left( \frac{2 + \theta - \frac{1}{N} \theta}{2 + 2\theta - \frac{1}{N} \theta} \right) + (p - \hat{m} - c) \frac{N - 1}{N} \theta \left( \frac{2 + \theta - \frac{1}{N} \theta}{2 + 2\theta - \frac{1}{N} \theta} \right).
\]

(25)

Take now the three terms in (25) in turn. First, we already know that the term multiplied by \(c\) is strictly increasing in \(\theta\). For the second term, which is multiplied by \(\hat{m}\), note that

\[
\frac{d}{d\theta} \left( \frac{2/\theta + 1 - \frac{1}{N}}{2/\theta + 2 - \frac{1}{N}} \right) < 0.
\]

Regarding the last term, note that \(p - \hat{m} - c > 0\), while it is immediate that the respective multiplier is strictly increasing in \(\theta\). This concludes the comparative analysis in \(\theta\).
With respect to a change in \(N\), we can now be brief by following the preceding argument. We thus need to show now that the right-hand side of (24) is strictly increasing in \(N\). To show this first for the multiplier, i.e., for the first line in (24), note that this transforms to
\[
2\theta \left( \frac{(1 + \theta)N - \theta}{(2 + \theta)N - \theta} \right) \left( \frac{N - 1}{2(1 + \theta)N - \theta} \right),
\]
where we can show for each of the two terms that they are indeed strictly increasing in \(N\). We next transform the term in rectangular brackets, which is strictly positive, to obtain
\[
1 - \hat{m} \left( \frac{(2 + \theta)N}{2(1 + \theta)N - \theta} \right) + (p - c) \frac{N - 1}{N} \left( \frac{(1 + \theta)N - \theta}{2(1 + \theta)N - \theta} \right) + p \left( \frac{N - 1}{2(1 + \theta)N - \theta} \right).
\]
Here, the term multiplied by \(-\hat{m}\) is strictly decreasing in \(N\). The term multiplied by \(p - c > 0\) is, next, strictly increasing in \(N\), as this holds for both parts. Finally, once again the term multiplied by \(p\) is strictly increasing in \(N\), as used also for (26). Q.E.D.

10 Appendix B: Calculations for the Hotelling Case

Solving the model of Hotelling competition, we obtain for the price of \(n = 1\)
\[
p_1 = \frac{2}{3}k_1 + \frac{1}{3}k_2 + \tau
\]
and for the firm’s downstream profits
\[
\pi_1 = \frac{1}{2\tau}(\tau + \frac{1}{3}(k_2 - k_1))^2,
\]
\[
\hat{\pi}_1 = \frac{1}{2\tau}(\tau + \frac{1}{3}(\hat{k}_2 - \hat{k}_1))^2.
\]
Expressions for \(n = 2\) are symmetric. (Again, in a slight abuse of notation we denote by \(\hat{\pi}_1\) the profits when only firm \(n = 1\) turns to the alternative source of supply.) Equilibrium marginal wholesale prices can be obtained immediately from combining the respective first-order conditions in (2). (Alternatively, we can use the expressions from Proposition 2, such as (16), after noting that, with Hotelling competition, \(\beta = 1\) and \(\gamma = 1\).) Q.E.D.

To support the discussion of average wholesale prices, \(W_n\), and average total costs, \(K_n\), in Section 6, we derive the respective expressions for the Hotelling case (cf. Section 5.1) under the assumed restriction that, both on and off equilibrium, the downstream firms are
active and the market is fully covered. On equilibrium, we thus require that $0 < q_1 < 1$, which after substituting for the equilibrium wholesale and final prices, holds if
\[ 0 < \frac{1}{\tau} \left( \frac{\tau}{2} + \frac{1}{9}(c_2 - c_1) \right) < 1. \]

Given that $c_1 \leq c_2$, this is satisfied if
\[ \tau > \frac{2}{9}(c_2 - c_1). \]  
(28)

Denote next the (off-equilibrium) quantity that prevails when firm $n = 1$ rejects the supplier’s offer by $\hat{q}_1$. We then have likewise that $0 < \hat{q}_1 < 1$ if
\[ \tau > \frac{5}{9}(c_2 - c_1). \]  
(29)

Finally, if firm $n = 2$ rejects, then off-equilibrium both firms will make positive sales if
\[ \tau > \frac{5}{27}(c_2 - c_1). \]  
(30)

As is easily seen, condition (30) is the strictest of these three conditions.

We ask next how $W_1$ and $W_2$ compare. Note first that, after some transformations,
\[ W_n = p_n - c_n - \frac{\hat{\pi}_n}{q_n}, \]
so that, after substituting for $p_n$,
\[ W_1 > W_2 \Leftrightarrow \frac{\hat{\pi}_1}{q_1} - \frac{\hat{\pi}_2}{q_2} < \frac{7}{9}(c_2 - c_1). \]  
(31)

Using next that
\[ \frac{\hat{\pi}_n}{q_n} = \left( \frac{3}{2} \tau - \frac{c_n - c_n'}{9} \right)^2, \]
we have that $W_1 > W_2$ holds if and only if
\[ \tau^2 > \frac{1}{27}(c_2 - c_1)^2. \]

This is implied by condition (30).

Finally, making use of condition (31), we have that
\[ K_1 < K_2 \Leftrightarrow \frac{\hat{\pi}_1^*}{q_1^*} - \frac{\hat{\pi}_2^*}{q_2^*} > -\frac{2}{9}(c_2 - c_1), \]
which transforms into condition
\[ \tau^2 > \frac{1}{72}(c_2 - c_1)^2. \]

This is again implied by condition (30).

\[ ^{27} \text{Note that we assume that consumers’ utility (gross of ”transportation cost”), } v, \text{ is sufficiently high.} \]