Price Discrimination in Cartels

Heiko Gerlach

School of Economics, University of Queensland

St Lucia, QLD 4069, Australia, email: h.gerlach@uq.edu.au

January 2020

Abstract

This paper analyzes price discrimination of an upstream cartel in the presence of a dominant firm at the retail level. Charging different wholesale prices creates a bond between the upstream cartel and the favored downstream firm. This bond reduces or eliminates this firm's incentives to accept deviation offers from upstream cartel members. When a cartel price discriminates in favor of the dominant downstream firm, it is able to implement prices above cost for any strictly positive value of the discount factor. This conclusion is robust with respect to the type of downstream competition, the form of contracts and the observability of offers. I also discuss the effect of vertical mergers on cartel sustainability in this setting.

JEL classification: D43, L41, L44.

Keywords: Cartels, Price Discrimination, Vertical Mergers, Antitrust Policy.

1 Introduction

Many of the uncovered cartels over the last two decades involved companies selling to intermediate good producers not final consumers.[1] In some of these cases, the cartel members practised price discrimination typically favoring the larger downstream firms. For example, in the *Bitumen* cartel in the Netherlands, the European Commission fined the five largest downstream customers for collusion with the upstream Bitumen cartel.[2] The upstream manufacturers had agreed to limit price rebates to smaller downstream competitors:

> For a period lasting at least between 1 April 1994 and 15 April 2002, collusion existed between and within a group of bitumen suppliers, consisting of [..], and a group of large Dutch road builders, consisting of [..], to regularly fix for sales and purchases of road pavement bitumen in the Netherlands as to the following: (1) the gross price; (2) a uniform (minimum) rebate on the gross price for that group of road builders; (3) a smaller (maximum) rebate on the gross price for other road builders.[3]

In this case, the larger road builders received a favorable treatment from the upstream cartel. This suggests that large downstream firms might tolerate higher input prices imposed by an upstream cartel if favorable price discrimination improves their market position relative to rival retailers. Or, as Levenstein and Suslow (2006) put it in their empirical cartel analysis, "they [large firms] would rather face an upstream cartel than downstream competition" (p. 64, para 1). In this paper, I formalize this idea and analyze the sustainability of upstream cartels in the presence of dominant retailers and price discrimination.

I consider a repeated game set-up with a vertically related industry where an upstream cartel sells to two retailers who are differentiated with respect to their retail

[1]For instance, 52 out of the 84 (61.9%) cartel cases decided by the European Commission between 2005 and 2018 involved vertically related industries.

[2]Other examples include the European Commission (EC) cases in Graphite Electrodes, Electrical and Mechanical Carbon Products and Sorbates.

[3]See European Commission (EC) Case 38.456, Bitumen NL, 21 December 2005, para 48.

cost. By price discriminating, the upstream cartel can create a bond with the favored retailer. This bond is the difference between the favored retailer's profits with an operating cartel and with competing upstream firms. A positive bond with the upstream cartel makes it more costly for the retailer to accept an offer from a defecting cartel member if acceptance implies that the cartel breaks down subsequently. For this to work, the upstream cartel must adopt a bond trigger strategy in which the upstream cartel reverts to competitive pricing if and only if the favored retailer accepts an offer from a defecting upstream firm. This introduces two types of deviations. Bilateral deviations attract both retailers and are punished with reversion to competition. Bypass deviations only target the non-favored retailer and are not punished with future price wars.

The main result of the paper is that, for any strictly positive value of the discount factor, the upstream cartel can sustain collusion with strictly positive profits by price discriminating in favor of the dominant retailer. In order to attract the favored downstream firm, a deviating upstream firm needs to offer significant price cuts to compensate for the loss of the cartel bond. For low discount factors, bilateral deviations can thus be prevented by lowering the cartel's wholesale price for the dominant retailer which, in turn, decreases the required deviation offer. At the same time, by-passing the dominant retailer and deviating with the less efficient retailer only gives limited downstream market access and is not profitable if the cartel price for the dominant retailer is sufficiently low. Hence, the presence of a dominant retailer and price discrimination are highly conducive to successful upstream collusion. When downstream firms compete in prices, this result holds for any cost asymmetry between the retailers. With quantity competition, the dominant firm needs to be sufficiently more efficient than its rival to achieve collusion for all strictly positive values of the discount factor.

The analysis is extended into several directions. It is demonstrated that the results are robust with respect to secret offers and wary beliefs on the part of the retailers. I also show that the use of two-part tariffs might make the cartel harder to sustain compared to linear wholesale prices. Finally, the effect of vertical integration on the sustainability of collusion with price discrimination depends on the identity of the merging downstream firm. When the dominant retailers pairs up with a manufacturer in this set-up, collusion becomes easier to sustain in the industry. By contrast, if the less efficient retailer is vertically related with an upstream firm, collusion becomes harder to sustain.

This paper contributes to two strands of the literature. Building on the work of Robinson (1933), there is an extant literature on price discrimination with respect to different types of final consumers identifying conditions under which aggregate output and welfare increase relative to uniform pricing.[4] There is also a string of papers that considers price discrimination by an upstream monopolist selling to intermediate good producers (Katz, 1987; DeGraba, 1990; Yoshida, 2000). They show that, in a static model, an upstream monopolist has an incentive to price discriminate against the more efficient downstream Cournot competitor.[5] This introduces an inefficient allocation of output among the downstream firms and price discrimination can either increase or decrease both total output and welfare. Here I show that price discrimination does not only affect downstream allocations but it can change the mode of upstream competition and allow for extensive collusion. Furthermore, the pattern of price discrimination depends on the discount factor of the upstream firms. A cartel with sufficiently patient firms can implement its first-best pricing and price discriminate against the more efficient retailer. For lower values of the discount factor, the cartel uses price discrimination in favor of the dominant retailer to implement the bond trigger strategy and sustain upstream collusion.

There is also a small but growing literature on cartels in vertically related markets.[6] Schinkel et al. (2008) show that an upstream cartel can shield itself from private damage claims of intermediate good producers by reducing supply to relax downstream competi-

[4]Schmalensee (1981) and Varian (1985) were among the first to generalize Robinson's results. Recent contributions include Armstrong and Vickers (2001) and Aguirre et al. (2010). For a good survey, see Stole (2007).

[5]Inderst and Valletti (2009) show that if there is the possibility of demand side substitution, this result can be reversed.

[6]Jullien and Rey (2007) consider the effect or resale price maintenance on the sustainability of an upstream cartel when otherwise symmetric retailers have private information about their local demand level. Piccolo and Miklos-Thal (2012) show that a downstream cartel can limit to incentives to under cut by paying slotting fees. Gu et al. (2019) considers optimal collusion in a successive oligopolies model with discount factors close to one.

tion and share part of the cartel profits. By contrast, this paper focuses on the upstream cartel's potential to sustain collusion through price discrimination. Nocke and White (2007) analyze the effect of vertical mergers on the sustainability of upstream cartels. Vertical integration implies that a defecting manufacturer has less downstream firms to sell to and and thus less access to consumers. At the same time, punishment is less severe for integrated firms. The outlets effect, however, dominates the punishment effect and vertical mergers facilitate collusion. Nocke and White (2010) extend this result to the case where downstream firms differ in their production capacities. However, their analysis focuses on the case where the upstream cartel extracts all industry rents and standard grim trigger strategies with reversion to the stage game equilibrium. Both features preclude their analysis from the effects discussed in this paper.

The plan of the paper is as follows. In Section 2, I set out the benchmark model. Section 3 briefly discusses optimal, monopolistic price discrimination. I derive equilibrium cartel formation with downstream price and quantity competition in Section 4. In Section 5, I extend the analysis to the case of unobservable contract offers. Section 6 discusses two scenario of vertical mergers. Section 7 analyzes collusion with two-part tariffs and the last section concludes. All proofs are relegated to the appendix.

2 The Model

Consider a vertical market structure with an upstream segment consisting of $n \geq 2$ identical manufacturers $i \in \{1, .., n\}$ producing a homogenous good. Variable upstream production costs are normalized to zero and any fixed costs are sunk. The output of the upstream industry is an input into the production of the downstream segment. One unit of the final good requires one unit of input from the upstream sector.

In the downstream segment, two retailers, $j \in \{1, 2\}$, compete in prices (or quantities) and sell their homogenous product to consumers. The downstream firms are differentiated with respect to their retail cost c_j. Retailer 1 (R1) is the dominant firm and has no retail cost, that is $c_1 = 0$. Retailer 2 (R2), by contrast, incurs an additional retail cost of $c_2 = c > 0$ per unit of the final good. Moreover, both retailers need to purchase their input from the upstream producers. Let a_j be the accepted wholesale price and $m_j = c_j + a_j$ the total per unit cost of retailer j. Downstream demand at price p is given by $D(p)$ which is continuous and strictly decreasing for any price such that $D(p) > 0$. The corresponding inverse demand is given by $P(Q)$. The monopoly profit at cost level m, $\pi^m(m) = (p - m)D(p)$, is strictly concave and maximized at the monopoly price $p^m(m)$. Suppose that the retail cost asymmetry between the downstream firms is not prohibitive, that is, $0 < c \leq p^m(0)$.

The analysis considers both price and quantity competition among retailers. Under price competition, the entire market demand goes to the lowest price retailer. In case of a price tie at the lowest price, total demand is evenly shared between the retailers. As Cournot competitors, each retailer chooses its output to maximize profits given its rival's output. Assume that $P'(Q) + QP''(Q) < 0$ which ensures that the profit functions are quasi-concave and the quantity game exhibits strategic substitutability (Vives, 1999). Further, let $q^c(m_j, m_{-j})$ and $\pi^c(m_j, m_{-j})$ be the Cournot equilibrium output and profit for retailer j with marginal cost m_j facing a rival with marginal cost m_{-j}. For all cost levels such that $q^c(m_j, m_{-j}) > 0$, it holds that

$$\frac{\partial q^c(m_j, m_{-j})}{\partial m_j} < 0 < \frac{\partial q^c(m_j, m_{-j})}{\partial m_{-j}}, \frac{\partial q^c(m_j, m_{-j})}{\partial m_j} < -\frac{\partial q^c(m_j, m_{-j})}{\partial m_{-j}} \tag{A1}$$

$$\frac{\partial \pi(m_j, m_{-j})}{\partial m_j} < 0 < \frac{\partial \pi(m_j, m_{-j})}{\partial m_{-j}} \tag{A2}$$

Both sets of assumptions hold in standard models of quantity competition. A retailer's equilibrium quantity and profit decreases in its own cost and increases in its rival's cost level. In absolute terms, the own-cost effect on equilibrium production is stronger than the cross-cost effect.

Firms play a repeated game in discrete time with an infinite horizon and a common discount factor δ for all upstream and downstream firms. As a benchmark, the following extensive form stage game is considered:

1. All upstream firms i simultaneously make a take-it-or-leave offer w_{ij} to each downstream firm j.

2. All offers are observed and retailers simultaneously decide whether and which offer to accept.

3. The accepted offers a_j are observed and the downstream firms compete in prices or quantities with a total marginal cost $m_j = c_j + a_j$.

The equilibrium concept for the benchmark model is subgame-perfect Nash equilibrium. Each manufacturer maximizes the discounted sum of its profits and conditions the continuation play on the history of all offers made to the retailers. In Section 5 below, the observability assumption is relaxed and I consider the effect of secret offers. As the main argument of the paper rests on price discrimination that affects the relative competitive position of the retailers, I focus on simple linear wholesale prices in the benchmark model. Section 7 below considers collusion with two-part tariffs.

It is easy to verify that the stage game has a unique subgame-perfect equilibrium in which all manufacturers offer their input at marginal cost to all retailers, that is, $w_{ij} = 0$, for all i and j, and the retailers compete downstream incurring only their retail costs. In this competitive equilibrium, each upstream firm makes zero profits. Notice that in repeated normal-form games, it is sufficient to focus on simple penal codes such that any deviation is punished with the same continuation path (Abreu, 1986). By contrast, as highlighted in Mailath, Nocke, and White (2017), optimal punishment in repeated extensive form games might involve more complex strategies which depend on the nature of the deviation.[7] In particular, in this setting, continuation play can be conditioned on upstream deviations and downstream acceptance of such deviations as well as on the identity of the deviator. I discuss the details of the punishment strategy in Section 4. Finally, for notational convenience, let $q(m_j, m_{-j})$ and $\pi(m_j, m_{-j})$ be retailer j's equilibrium quantity and profit independent of the mode of competition.

[7]Hatfield et al. (2019) is another example of a repeated extensive form game in which players use sophisticated strategies that make use of within-period punishment. In their model, cartel member bid for contracts and then form a joint production syndicate in the same period with the bid winner as leader. They show that cartels can sustain collusive pricing despite low levels of market concentration.

3 Monopolistic price discrimination

Before proceeding to the analysis of cartel formation, let us characterize the unconstrained optimal wholesale price pair (w_1^*, w_2^*) that maximizes upstream industry profits in the presence of asymmetric retailers. Let $R(m_j, m_{-j}) = w_j q(m_j, m_{-j})$ be the upstream revenue with retailer j for given total marginal costs, respectively. The profit-maximizing wholesale prices then solve

$$(w_1^*, w_2^*) = \arg \max_{(w_1, w_2)} M(w_1, w_2) \equiv R(c + w_1, w_2) + R(w_2, c + w_1).$$

Let us solve this problem for price and quantity competition, respectively.[8] When downstream firms compete in prices, the retailer with the lowest total cost serves the market at a price equal to the total cost level of its competitor. For a given cost level of the retailer that is serving the market, an upstream monopolist maximizes downstream demand by setting the rival's wholesale price to the lowest level such that it cannot profitably serve demand. Moreover, the upstream monopolist is always able to charge a higher wholesale price to the more efficient firm as this firm is able to raise its price higher without losing sales to the competitor. As a consequence, the optimal wholesale price w_1 is just below $c + w_2$ such that retailer 1 serves the downstream market. An upstream monopolist thus maximizes $w_1 D(w_1)$ and charges

$$w_1^* = p^m(0) > w_2^* = p^m(0) - c.$$

Now suppose the downstream firms are Cournot competitors. With quantity competition, total downstream production is a function of the sum of the total costs of both retailers and we can define the total downstream production as[9]

$$Q^c(c + w_1 + w_2) = q^c(w_1, c + w_2) + q^c(c + w_2, w_1).$$

[8] See the proof of Lemma 1 in the appendix for a formal analysis.

[9] This is demonstrated in the proof of the following lemma in the Appendix.

Using this with the two first-order conditions for the optimal wholesale prices yields

$$(w_1 - w_2)[\frac{\partial q^c(w_1, c + w_2)}{\partial w_1} - \frac{\partial q^c(w_1, c + w_2)}{\partial w_2}] + 2[q^c(w_1, c + w_2) - \frac{Q^c(c + w_1 + w_2)}{2}] = 0.$$

For any given total downstream cost (and production level), the above equation characterizes the optimal choice of wholesale prices. The first term is the effect of an increase in w_1 (and a simultaneous reduction in w_2 to keep overall cost constant) on revenues at the intensive margin. The squared bracket gives the net amount of consumers retailer 1 loses to retailer 2. This is multiplied by the difference in wholesale prices to give the marginal loss in revenues from raising the price for retailer 1. The first term is thus negative if and only if $w_1 > w_2$. The second term is the effect of raising w_1 on inframarginal consumers. This effect is positive if and only if retailer 1 has a higher market share than its rival, that is, if $w_1 < c + w_2$. Hence, in an optimal interior solution, an upstream monopolist gives a lower wholesale price to retailer 2 but allows a total cost advantage and a higher market share for retailer 1. We can thus summarize as follows.

Lemma 1. *A monopolist manufacturer optimally price discriminates against the more efficient firm such that*

$$w_2^* < w_1^* \leq w_2^* + c.$$

The optimal unconstrained wholesale prices of an upstream cartel imply a higher wholesale price but weakly lower total costs for the more efficient firm. The cartel price discriminates against the more efficient firm and puts downstream competitors on a more even playing field. The degree of price discrimination depends on the mode of competition. Price competition downstream leads to more price discrimination than quantity competition.

4 Collusion in the Benchmark Model

4.1 General framework

Suppose the upstream firms form a cartel that aims to sustain uniform wholesale prices for each of the two downstream firms, that is, $w_{ij} = w_j$. First, consider cartel trigger strategies that punish any deviation with eternal reversion to the competitive equilibrium of the stage game. This punishment involves zero continuation profits for the upstream firms and is the maximum punishment. On the equilibrium path, downstream firm j faces the same offer w_j from all manufacturers and randomly picks one supplier as long as the resulting profits are non-negative. After any deviation, retailer j anticipates the upstream punishment and maximizes its current period profits by accepting the lowest priced contract offer. This implies that the optimal deviation for any $w_j \leq w_j^*$ is to shave the cartel price for each of the two retailers and serve the entire downstream market. The cartel is then sustainable if and only if a manufacturer's discounted, future share in cartel profits exceeds the current period market share gain when deviating,

$$\frac{\delta}{1-\delta}\frac{M(w_1,w_2)}{n} \geq \frac{n-1}{n}M(w_1,w_2) \quad \text{or} \quad \delta \geq \frac{n-1}{n}. \tag{1}$$

Hence, if the manufacturers are sufficiently patient, they can sustain the optimal price levels (w_1^*, w_2^*). For lower values of the discount factor, no collusion is sustainable using grim trigger strategies after a deviation. In fact, the cartel constraint is the same as in the case where the manufacturers do not use intermediaries and compete in prices for consumers. In repeated normal form games, reverting to the deviator's minmax value of the stage game represents the maximum punishment. However, as mentioned above, reverting to the competitive equilibrium after a deviation is not necessarily the best continuation plan to sustain collusive outcomes in this repeated extensive form game. In what follows, I focus on discount factors that do not satisfy condition (1) and show that collusion is sustainable if manufacturers adjust their strategies to reward a retailer for not accepting deviation offers.

To fix ideas, let $m_j = c_j + w_j$ be the total cost of retailer j at the wholesale price fixed by the upstream cartel. It is useful to take the difference in profits for retailer j along the equilibrium and punishment path of the cartel,

$$B(m_j, m_{-j}) = \pi(m_j, m_{-j}) - \pi(c_j, c_{-j}).$$

This is the per period profit that retailer j gains or loses due to the existence of the upstream cartel. I refer to this difference as the bond of retailer j with the upstream cartel.

Lemma 2. *If $B(m_j, m_{-j}) > 0$, then $w_j < w_{-j}$ and $B(m_{-j}, m_j) < 0$.*

In order to provide a positive bond with a downstream firm, the upstream cartel needs to implement a sufficient degree of price discrimination favoring this retailer. This implies that if the upstream cartel forms a positive bond with one retailer, there must be a negative bond with the other retailer. Vice versa, if prices are similar, both retailers may form a negative bond and would be better off without upstream collusion.

Assume that the cartel implements wholesale prices such that retailer j forms a positive bond with the cartel and consider the following punishment strategy:

> The upstream firms trigger punishment and revert to the competitive equilibrium if and only if manufacturer i deviates with at least one offer ($w_{i1} \neq w_1$ and/or $w_{i2} \neq w_2$) and retailer j accepts its offer from this manufacturer ($a_j = w_{ij}$).

This strategy prescribes maximum punishment if and only if the favored retailer *accepts* the offer from a deviating manufacturer. If this retailer rejects, then collusion continues, independent of whether rival retailer $-j$ had a deviating offer, and whether this offer was accepted or not. In other words, after a deviation, the retailer with the positive bond becomes pivotal for the continuation of the cartel. The strategy also implies that if a manufacturer only offers a deviating price for the disadvantaged retailer $-j$, punishment sets in if the favored retailer accepts this deviating manufacturer's offer.

Now consider the retailers' incentives to accept deviation offers. Suppose a manufacturer offers contracts $(\hat{w}_j, \hat{w}_{-j})$ such that at least one of the offers differs from the equilibrium values (w_j, w_{-j}). Denote the corresponding total input costs as $(\hat{m}_j, \hat{m}_{-j})$. At the contract acceptance stage, both retailers simultaneously decide which of the offers to accept. Given the above bond trigger strategy, independent of what retailer j does, if retailer $-j$ has a deviating offer, its choice does not change the continuation value for either of the retailers. This means that retailer $-j$ maximizes current period profits and accepts any $\hat{w}_{-j} < w_{-j}$. By contrast, consider the retailer with the positive cartel bond. After observing a deviating offer, retailer j anticipates that accepting the deviator's offer triggers cartel punishment while, otherwise, upstream collusion continues. For a given acceptance of its rival, retailer j is better off accepting the deviator's offer if and only if

$$\pi(\hat{m}_j, \hat{m}_{-j}) - \pi(m_j, \hat{m}_{-j}) \geq \frac{\delta}{1-\delta}[\pi(m_j, m_{-j}) - \pi(c_j, c_{-j})], \tag{2}$$

The LHS is the short-run profit from accepting the deviation which increases in the difference between equilibrium and deviating offer as well as in the deviating offer to retailer $-j$. The RHS is the future, discounted value of the bond between retailer j and the cartel. In order to attract the retailer with a positive cartel bond, a deviating manufacturer has to compensate for the lost future profits that retailer j makes on the equilibrium path. This requires a price cut strictly below the cartel price level. The necessary price cut is higher, the stronger the bond, the lower the offer to the rival retailer and the higher the discount factor of the retailer. We thus get the following result.[10]

Lemma 3. *Consider the contract acceptance stage when the cartel forms a positive bond with retailer j. Retailer $-j$ accepts any offer $\hat{w}_{-j} < w_{-j}$. There exists a $\omega_j < w_j$ such that retailer j accepts the deviating offer if and only if $\hat{w}_j \leq \omega_j$.*

This implies that there are two types of continuation equilibria at the contract acceptance stage. If the price cut for the retailer with the positive bond is sufficiently strong, both retailers accept the offer of the deviating manufacturer. Otherwise, a deviating manufacturer is only able to attract the disadvantaged retailer at any price below its

[10] For notational convenience, I suppress that the threshold value ω_j is a function of $(w_j, w_{-j}, \hat{w}_{-j})$.

equilibrium cartel price. The favored retailer accepts a cartel price offer from a non-deviating manufacturer to avoid triggering punishment. I refer to deviations leading to these two types of equilibria as *bilateral* and *bypass deviations*, respectively.

We are now in a position to analyze upstream incentives when firms form a cartel agreement (w_1, w_2) involving a bond $B(m_j, m_{-j}) > 0$ with downstream firm j. Take a bilateral deviation $(\hat{w}_j, \hat{w}_{-j})$ satisfying $\hat{w}_{-j} < w_{-j}$ and condition (2). Any such deviation is accepted by both retailers and triggers cartel punishment from the next period onwards. A manufacturer has no incentive to use a bilateral deviation if and only if

$$\frac{1}{1-\delta}(R(m_j, m_{-j}) + R(m_{-j}, m_j))\frac{1}{n} \geq R(\hat{m}_j, \hat{m}_{-j}) + R(\hat{m}_{-j}, \hat{m}_j). \tag{3}$$

Bilateral deviations allow the deviator to capture the entire downstream market by selling through both retailers. However, attracting retailer j requires a significant price cut and the deviator loses future cartel profits.

A cartel agreement also has to resist the bypass deviation constraint in which $(\hat{w}_j, \hat{w}_{-j})$ do not satisfy condition (2) but the retailer with a negative bond accepts $\hat{w}_{-j} < w_{-j}$. This deviation does not trigger cartel punishment as the pivotal retailer rejects the deviator's offer. A bypass deviation is not profitable for a manufacturer if and only if

$$[R(m_j, m_{-j}) + R(m_{-j}, m_j]\frac{1}{n} \geq R(\hat{m}_{-j}, m_j). \tag{4}$$

Bypass deviation avoid the trigger of future punishment but the deviator only sells through retailer $-j$. A successful upstream cartel thus has to ensure that conditions (3) and (4) are satisfied for all admissible deviation offers. The qualitative nature of these conditions depends on the mode of downstream competition. In the next subsection, I discuss cartel sustainability with price and quantity competition between retailers, respectively.

4.2 Downstream Price competition

Suppose the downstream firms compete in prices and the upstream cartel forms a positive bond with retailer 1. First consider the bypass deviation, in which a deviating manufacturer sets wholesale prices such that retailer 2 accepts while retailer 1 rejects its offer (and accepts a price w_1 with a rival manufacturer). In order to attract any demand through retailer 2, the deviating manufacturer needs to set $\hat{w}_2 < w_1 - c$ and the optimal deviation is to give R2 marginally lower total costs. FIGURE 1 below depicts the two types of deviations in a $\hat{w}_1 - \hat{w}_2$ diagram. The equilibrium cartel prices (w_1, w_2) are marked at point A. The bypass deviation is to point B in this graph. Maximum deviation profit are then $(w_1 - c)D(w_1)$ and the bypass deviation constraint is

$$\frac{w_1 D(c + w_2)}{n} \geq (w_1 - c)D(w_1). \tag{4-P}$$

Bypassing retailer 1 with a strictly positive demand is only feasible if the cartel price for R1 is larger than the retail cost of R2. In fact, there exists a threshold value for w_1 strictly larger than c for any $c > 0$ such that the bypass deviation is never profitable for any w_1 less than this threshold.

Lemma 4. *Consider downstream price competition and a cartel bond with retailer 1, $B(w_1, c + w_2) > 0$. There exists a unique threshold $\omega_1' \in [c, p^m(0)]$ such that the bypass constraint is satisfied if and only if $w_1 \leq \omega_1'$. The threshold ω_1' decreases in w_2 and increases in c.*

The bypass constraint requires a cartel price for retailer 1 below the first-best level $w_1^* = p^m(0)$. Moreover, this constraint is easier to satisfy, the higher the cost difference c between the retailers. A higher retail cost requires larger price cuts to attract consumers through retailer 2 and makes deviations less attractive for a manufacturer. At the same time, a higher cost difference increases the equilibrium price and reduces upstream profits on the cartel equilibrium path. However, this second effect is dominated for any cartel prices that create a positive bond with retailer 1.

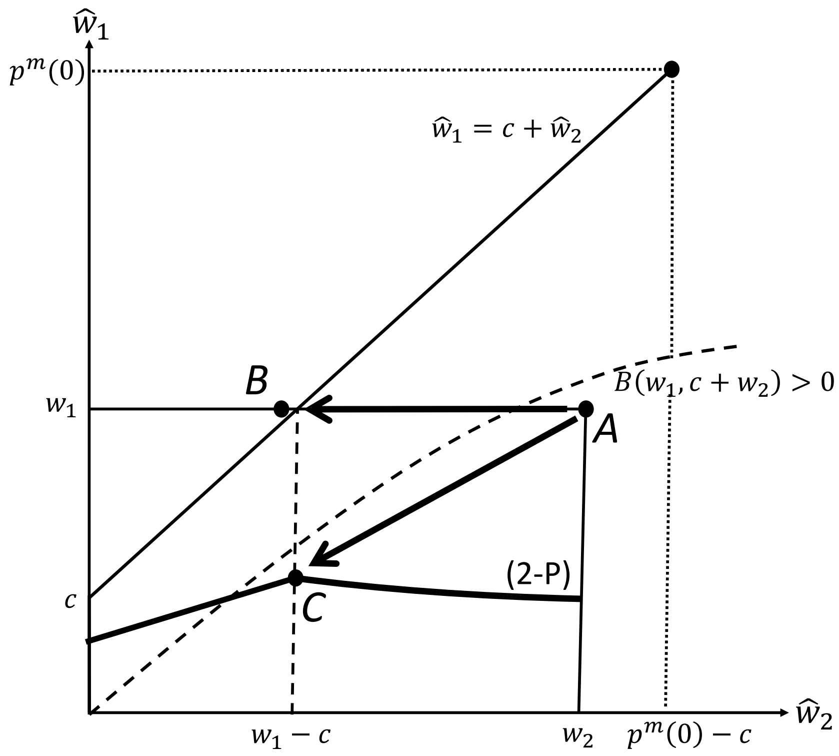

FIGURE 1: *Bilateral and bypass deviation with price competition*

Next consider the bilateral deviation constraint when both retailers accept the offers of a deviating manufacturer. With downstream price competition, the retailer with the lowest marginal cost serves the downstream market at a price equal to the cost of its rival. Retailer 1 accepts the offer from a deviating manufacturer if its short-run gain with the deviation offer outweighs the future value of the bond with the upstream cartel. That is, condition (2) is given by

$$(c + \hat{w}_2 - \hat{w}_1)D(c + \hat{w}_2) - (\max\{c+\hat{w}_2 - w_1, 0\})D(c + \hat{w}_2) \geq \frac{\delta}{1-\delta}[(c + w_2 - w_1)D(c + w_2) - cD(c)]. \quad \text{(2-P)}$$

In order to attract retailer 1, a deviating manufacturer has to cut the price to a level $\hat{w}_1 \leq \omega_1 < c + \hat{w}_2$. Note that the deviation offer to retailer 2 also affects the profits of retailer 1 if it rejects the deviator's offer and accepts the cartel price from another manufacturer. Two cases can arise. If $\hat{w}_2 < w_1 - c$, retailer 1 is priced out of the market in the current period after accepting the cartel price. By contrast, for higher deviating prices $\hat{w}_2$, R1 serves the market independent of which firm's contract it accepts. The next lemma further characterizes the minimum wholesale price offer ω_1.

Lemma 5. *With downstream price competition, it holds that*

$$\frac{\partial \omega_1}{\partial \hat{w}_2} \gtreqless 0 \text{ if } \hat{w}_2 \lesseqgtr w_1 - c, \quad \frac{\partial \omega_1}{\partial w_1} > 0 \text{ and } \frac{\partial \omega_1}{\partial w_2} < 0.$$

Condition (2-P) is depicted in FIGURE 1. The upper bound ω_1 has a kink at $\hat{w}_2 = w_1 - c$ and is increasing (decreasing) for lower (higher) values of $\hat{w}_2$. For lower values, the deviation price $\hat{w}_2$ only affects the favored retailer's profits when accepting the deviator's offer. A higher $\hat{w}_2$ thus makes the condition easier to satisfy. By contrast, for higher values, the deviation revenues $(c + \hat{w}_2)D(c + \hat{w}_2)$ cancel out on the LHS. The current period gains are equal to the cost savings from deviation. A higher $\hat{w}_2$ lowers the demand level over which cost savings are applied and reduces the incentives to accept the deviator's offer. This yields the comparative statics in the lemma with respect to $\hat{w}_2$. Furthermore, note that retailer 1 is less likely to accept the deviation, the lower the equilibrium cartel price w_1. A low cartel price raises current period profits when rejecting the deviator and increases the bond with retailer 1.

We are now in a position to characterize the optimal bilateral deviation for a manufacturer. Since an accepted bilateral deviation triggers cartel punishment, a deviating manufacturer maximizes current period profits $\hat{w}_1 D(c + \hat{w}_2)$ subject to (2-P).

Lemma 6. *Consider downstream price competition and $w_1 \leq p^m(0)$. The optimal bilateral deviation for a manufacturer is characterized by $\hat{w}_2 = \max\{w_1 - c, 0\}$ and*

$$\hat{w}_1 = w_1 - \frac{\delta}{1-\delta} \frac{B(w_1, c + w_2)}{D(max\{w_1, c\})}.$$

If the cartel price w_1 is sufficiently small, then there exists no bilateral deviation yielding positive profits for a manufacturer.

The lemma states that, for any cartel price $w_1 \leq p^m(0)$, the optimal deviation is to the highest possible price $\hat{w}_1$ satisfying condition (2-P) for any $\hat{w}_2 < w_2$. More specifically, if w_1 is larger than the cost difference c, a deviator optimally chooses the kink point on ω_1 (point C in FIGURE 1). If w_1 is less than c, then the kink does not exist. In this case, the optimal deviation maximizes downstream demand by setting $\hat{w}_2$ to zero and choosing the highest price $\hat{w}_1$ satisfying (2-P). This allows us to write the bilateral deviation constraint as follows

$$\frac{w_1 D(c+w_2)}{n(1-\delta)} \geq w_1 D(max\{w_1, c\}) - \frac{\delta}{1-\delta} B(w_1, c+w_2). \tag{3-P}$$

A manufacturer weighs the cartel profits on the equilibrium path against short run gains from deviation net of the discounted, future cartel bond which retailer 1 demands as compensation for accepting the deviation offer. As the equilibrium price becomes smaller, the short run gains go towards zero while the value of the bond increases. Hence, for sufficiently small equilibrium prices w_1, the value of the RHS is negative and the bilateral deviation constraint is always satisfied.

From our analysis so far, the first part of the next proposition follows immediately.

Proposition 1. *Suppose retailers compete in prices. For any value of $\delta > 0$ and $c > 0$, there exists a subgame-perfect equilibrium in which the upstream cartel can sustain strictly positive profits. In such an equilibrium, the cartel members price discriminate against the less efficient retailer.*

With downstream price competition, an upstream cartel is able to sustain partial collusion for any positive value of the discount factor. By skewing the relative wholesale prices in favour of the dominant retailer, the cartel transfers collusive rents and forms a bond with this retailer. This weakens R1's incentives to accept deviating offers. In turn, upstream deviations become more costly, or completely unprofitable when the cartel price w_1 is relatively low. Bypassing the dominant retailer, and deviating through R2 only gives limited and costly access to the downstream market as this retailer is less efficient. In particular, for low prices for retailer 1, a bypass deviation is not profitable for a deviating manufacturer.

Furthermore, as the second part of the proposition suggests, collusion with a positive bond between the cartel and the less efficient retailer is not sustainable. To see this, it suffices to consider the bypass constraint of the cartel. A positive bond with R2 means that cartel prices satisfy $w_1 > c + w_2$. At these prices, a manufacturer makes per-period cartel profits of $w_2 D(w_1)/n$. A deviation to $\hat{w}_1 = c + w_2$ with retailer 1 would allow the manufacturer to serve the downstream market and achieve higher upstream industry profits.[11] Hence, the upstream cartel is only able to sustain a positive bond and collusion for any $\delta > 0$ if it price discriminates in favor of the dominant retailer.

Two more observations follow. First, the sustainability of the cartel hinges on the existence of a dominant retailer, that is $c > 0$, in order to relax the bypass deviation constraint. With symmetric firms at the retail level, collusion would only be sustainable for high discount factors that satisfy condition (1). Second, as seen in Section 3, an upstream monopolist optimally favors the less efficient retailer. By contrast, this result suggests that if a cartel is unable to sustain this first-best solution, then it is best to reverse price discrimination and favor the dominant retailer.

4.3 Downstream Quantity Competition

Consider now the case of downstream quantity competition between the retailers. Suppose again the cartel forms a positive bond with dominant retailer 1. With quantity competition downstream, condition (2) holds if the difference in current period Cournot profits exceeds the value of the future discounted bond with the cartel or

$$\pi^c(\hat{w}_1, c + \hat{w}_2) - \pi^c(w_1, c + \hat{w}_2) \geq \frac{\delta}{1-\delta}[\pi^c(w_1, c + w_2) - \pi^c(0, c)]. \tag{2-Q}$$

The price threshold ω_1 below which retailer 1 accepts the deviating offer has similar properties to the case with price competition. It increases in the cartel equilibrium price for retailer 1 and decreases in the price for retailer 2. Moreover, the higher the deviation

[11]Check that since $w_1 > c + w_2$ it holds that

$$(c + w_2)D(c + w_2) > w_2 D(c + w_2) > w_2 D(w_1) > w_2 D(w_1)/n.$$

price $\hat{w}_2$ for its rival, the more valuable are price cuts for retailer 1. This implies that the price threshold increases in the deviation offer to the other retailer.

Lemma 7. *With quantity competition, the minimum price offer ω_1 increases in the deviation price $\hat{w}_2$ for retailer 2. If w_1 is sufficiently small, there exists no profitable bilateral deviation for a manufacturer.*

The analysis of the bilateral deviation constraint is similar to the case with price competition. In particular, the fact that the bilateral deviation constraint can be satisfied with strictly positive wholesale prices for any $\delta > 0$ also holds in this case. If the equilibrium cartel price for R1 is sufficiently low, then there is no positive deviation price that could induce the dominant retailer to accept a deviation.

In order to analyze the bypass deviation constraint with quantity competition, let $w_2^r(w_1)$ be the wholesale price that maximizes revenues from retailer 2 for a given price w_1 for R1. In order to attract retailer 2, the deviation price needs to be less than the equilibrium price w_2 offered by rival manufacturers. Hence, the bypass deviation constraint is given by

$$\frac{1}{n}[w_1 q^c(w_1, c + w_2) + w_2 q^c(c + w_2, w_1)] \geq \hat{w}_2 q^c(c + \hat{w}_2, w_1) \tag{4-Q}$$

where $\hat{w}_2 = \min\{w_2 - \epsilon, w_2^r(w_1)\}$. This condition is easier to satisfy when the cartel price w_1 and the cost difference c are high. At $w_1 = 0$, there is no revenue from retailer 1 and this condition is never satisfied. However, as w_1 increases, cartel profits rise faster than deviation profits and there exists a threshold value above which the constraint is satisfied. This minimum price level depends on the cost of the bypass retailer 2. As c increases, both sides of condition (4-Q) decrease but the deviation profits go down faster since a higher c raises the revenues from retailer 1. Let $\hat{c} = \{c | q^c(c, 0) = 0\}$ be the prohibitive cost level for R2 and the following result obtains.

Lemma 8. *Consider the bypass deviation constraint with quantity competition. For any $w_1 > 0$, there exists a threshold value $\hat{c} \in (0, \bar{c})$ such that if $c \geq \hat{c}$, the bypass deviation can be satisfied.*

Any strictly positive price w_1 can be satisfied if the cost asymmetry between R1 and R2 is sufficiently high. As w_1 approaches zero, the required cost level is the prohibitive level $\bar{c}$. The main result with quantity competition follows.

Proposition 2. *Suppose retailers compete in quantities. For any $\delta > 0$ and $c \geq \hat{c}$, there exists a subgame-perfect equilibrium in which the manufacturers can sustain strictly positive profits and price discriminate against retailer 2.*

It is harder to sustain strictly positive profits for all levels of the discount factor when retailers compete in quantities rather than in prices. The reason for this is that the bypass constraint is more restrictive with quantity competition. A wholesale price reduction from a deviating manufacturer makes retailer 2 more aggressive in downstream competition. With price competition and strategic complements, this is countered aggressively by retailer 1 with lower prices. By contrast, with quantity competition and strategic substitutes, a lower wholesale price for R2 leads to a higher market share for this retailer. Hence, bypass deviations are more attractive for manufacturers when retailers compete in quantities.

5 Collusion with Secret Offers

Suppose retailers receive private offers from the manufacturers, that is, a retailer cannot observe its rival's offers before deciding which contract to accept. It is well known that if a monopolistic manufacturer can act opportunistically with an individual retailer, the manufacturer might lose market power with respect to all retailers.[12] In this section, I analyze the effects of secret offers on the sustainability of an upstream cartel.

To fix ideas, consider the following change to the stage game. First, all manufacturers offer contracts to both retailers. Retailer j observes his contract offers w_{ij} and decides which one to accept. After the acceptance decisions, the retailers observe the accepted contract of their rival and compete in the downstream market. Finally, manufacturers

[12] This issue has been first discussed by Hart and Tirole (1990) and O'Brien and Shaffer (1992). See Rey and Tirole (2007) for a survey of this literature.

observe all offers before the game continues in the next period.[13] I look for Perfect Public Equilibria of the repeated game in which manufacturers condition the continuation play on the publicly available information of the upstream cartel, that is, on actual and accepted offers.

The upstream cartel is trying to implement a uniform pricing scheme for each retailer such that retailer j only receives offers w_j. With secret offers, a retailer's willingness to accept a deviation depends on his belief regarding the offer that the rival retailer has received from the deviating firm. Hence, the out-of-equilibrium belief formation affects the incentive to accept deviations and, ultimately, the incentive to deviate from the upstream cartel. A retailer receiving a deviation offer should realize that this offer might affect the upstream firm's contract proposal to the other retailer. In other words, the retailer's beliefs should be consistent with the incentives of the deviating manufacturer. Following the notion of *wary beliefs* introduced by McAfee and Schwartz (1994), I assume that when a given retailer receives a deviation offer $\hat{w}_j$, it believes that the deviating manufacturer offers a contract to the rival retailer that maximizes the deviator's continuation value given $\hat{w}_j$.

More formally, assume that $u(a_j, a_{-j})$ and $v_j(a_j, a_{-j})$ are the continuation values for a manufacturer and retailer j, respectively. These values are determined by the strategies used by the upstream cartel. Furthermore, let $R_j(a_j, a_{-j})$ be the manufacturer's revenue with retailer j. Note that $R_j(a_j, a_{-j}) = R(a_j, a_{-j})$ if $a_j = \hat{w}_j$. By contrast, if $\hat{w}_j = w_j$, then $R_j(a_j, a_{-j}) = R(a_j, a_{-j})/n$. When being offered a contract $\hat{w}_j$, retailer j expects the deviating manufacturer to offer a contract $W_{-j}(\hat{w}_j)$ given by

$$\begin{aligned} W_{-j}(\hat{w}_j) = \arg\max_{w} \quad & R_j(a_j(\hat{w}_j, w), a_{-j}(w, W_j(w))) + R_{-j}(a_j(\hat{w}_j, w), a_{-j}(w, W_j(w))) \\ & + \delta u(a_j(\hat{w}_j, w), a_{-j}(w, W_j(w))) \end{aligned}$$

Using this wary belief, retailer j then decides whether to accept or reject a deviating

[13] This assumption is stronger than what is needed for the analysis below. All what is required is that after retailer 2 accepts a deviating offer, the manufacturers learn the deviator's offer to retailer 1.

offer according to

$$a_j(\hat{w}_j, W_{-j}(\hat{w}_j)) = \arg \max_{a \in \{\hat{w}_j, w_j\}} \quad \pi(a, a_{-j}(W_{-j}(\hat{w}_j), W_j(W_{-j}(\hat{w}_j)))) + \delta v_j(a, a_{-j}(W_{-j}(\hat{w}_j), W_j(W_{-j}(\hat{w}_j)))).$$

Suppose the manufacturers follow a similar bond trigger strategy as in the benchmark model, with one modification. With observable offers, retailer 1 receiving on-schedule cartel offers knows when a manufacturer only deviates with retailer 2. Retailer 1 can then avoid the break-up of the cartel by not signing with the deviator and accept the offer of another cartel member. With secret offers this is not possible. In a bypass deviation, a manufacturer could deviate with retailer 2 only and still have its offer accepted by retailer 1. Such a deviation would always be profitable with the bond trigger strategy used in the benchmark model. In order to prevent this type of deviations, the cartel has to punish a deviating manufacturer by reverting to the competitive equilibrium of the stage game. Hence, cartel strategy with secret offers triggers punishment by reversion to the competitive equilibrium if either *(i)* R1 accepts an offer from a deviating manufacturer or, *(ii)* a manufacturer deviates with R2 only.

Let us now discuss cartel sustainability with secret offers and quantity competition. The analysis with price competition is qualitatively similar and relegated to the appendix. First consider bilateral deviations aiming to attract both retailers. Let $w_2^R(w_1)$ be the wholesale price that maximizes upstream revenues from both retailers for a given price w_1. Again, retailer 2 accepts any deviation $\hat{w}_2 < w_2$ independent of the offer made to the rival retailer. With wary beliefs, retailer 1 accepts a deviation $\hat{w}_1$ if and only if the *expected* current deviation gains exceed the discounted, future bond with the cartel or

$$\pi^c(\hat{w}_1, c + W_2(\hat{w}_1)) - \pi^c(w_1, c + W_2(\hat{w}_1)) \geq \frac{\delta}{1-\delta}[\pi^c(w_1, c + w_2) - \pi^c(0, c)]. \quad \text{(2-S)}$$

where $W_2(\hat{w}_1) = \min\{w_2^R(\hat{w}_1), w_2 - \epsilon\}$. What is the effect of secret offers on retailer 1's decision to accept or reject deviating offers? Consider FIGURE 2 below which depicts retailer 1's highest acceptable price ω_1 for a given deviation offer $\hat{w}_2$ to retailer 2, that is condition (2-Q) from the case with observable offers. As discussed above, under quantity competition, the highest acceptable price increases in the price for the rival retailer. Moreover, a deviator's optimal price for R2, $w_2^R(\hat{w}_1)$, increases in the offer $\hat{w}_1$ for retailer 1. The lower the offer R1 receives, the lower is the expected offer for retailer 2. Hence, the optimal deviation with secret offers is point A on $w_2^R(\hat{w}_1)$ that just satisfies (2-Q). This deviation achieves a profit level of M_0. With observable offers, the prices only need to satisfy (2-Q). Since along w_2^R the iso-profit curves have a slope of zero, there must exist a more profitable bilateral deviation ($M_1 > M_0$) to point B. With observable offers, the deviator can commit to a higher price for retailer 2 which allows to attract the dominant retailer at a higher wholesale price. Secret offers thus make it easier for the cartel to prevent bilateral deviations. Moreover, as ω_1 decreases in the cartel price w_1, it is also clear that there exist cartel prices such that no profitable bilateral deviation exists with secret offers while there would be a strictly positive profit for a deviator with observable offers.

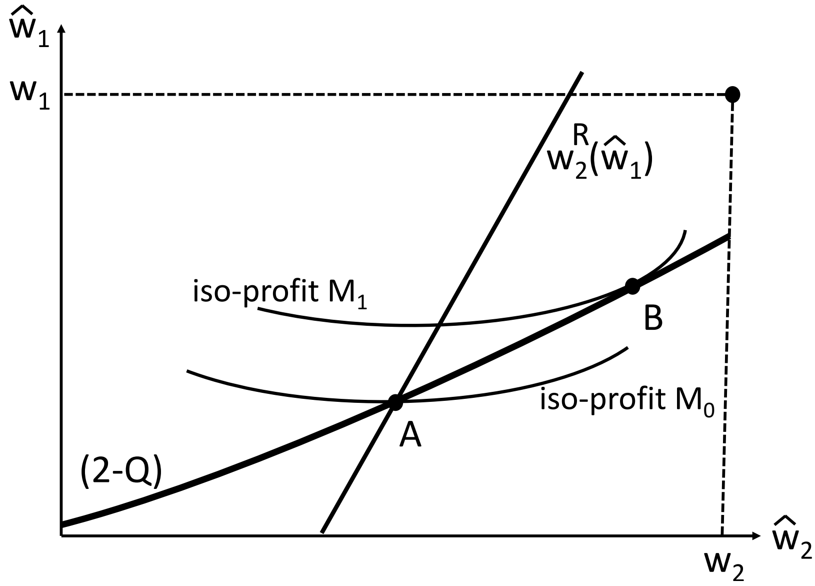

FIGURE 2: *Bilateral deviation with secret offers*

Now consider bypass deviations to prices that do not satisfy (2-S). There are two possible deviations. First, a manufacturer could set a price $\hat{w}_1 \neq w_1$. In this case, R1 rejects and no punishment is triggered. The optimal deviation is to $\hat{w}_1 = w_1 - \epsilon$

and $\hat{w}_2 = w_2^r(\hat{w}_1)$. This yields the same condition as the bypass constraint (4-Q) in the benchmark model. The second possibility is an on-schedule deviation $\hat{w}_1 = w_1$ with retailer 1 and $\hat{w}_2 < w_2$. This deviation is not observed by R1 but triggers punishment by the upstream cartel. Let $w_2^s(w_1)$ be the price that maximizes a manufacturer's current period revenues with this deviation.[14] The additional bypass constraint is then given by

$$\frac{w_1 q^c(w_1, c + w_2) + w_2 q^c(c + w_2, w_1)}{n(1-\delta)} \geq w_1 q^c(w_1, c + \hat{w}_2, w_1)\frac{1}{n} + \hat{w}_2 q^c(c + \hat{w}_2, w_1), \quad \text{(4-S)}$$

with $\hat{w}_2 = \min\{w_2 - \epsilon, w_2^s(w_1)\}$. It is clear that if the discount factor is sufficiently small, then this condition is harder to satisfy than condition (4-Q). In this case, bypass deviations are strictly more profitable with secret offers.

The possibility of secret offers in the presence of wary beliefs has two opposing effects on deviation incentives for the upstream cartel. It invites opportunistic behaviour from deviating manufacturers. This makes retailer 1 more reluctant to accept deviations which in turn implies that bilateral deviations are (weakly) less profitable. At the same time, undetected on-schedule deviations with retailer 1 allow manufacturers to gain current period market share with the dominant retailer. This makes bypass deviations (weakly) more profitable. With respect to the overall sustainability of an upstream cartel, similar qualitative results as in the benchmark model obtain.

Proposition 3. *Consider the possibility of secret offers when retailers hold wary beliefs. With price competition the upstream cartel is able to sustain strictly positive profits for any $\delta > 0$. With quantity competition, manufacturers can sustain strictly positive profits for any $\delta > 0$ and $c \geq \hat{\hat{c}}$ where $\hat{\hat{c}} \geq \hat{c}$.*

Like in the benchmark model, bilateral deviations can be prevented by the cartel for very low discount factors by reducing the equilibrium price for retailer 1. By contrast, for sufficiently low discount factors, the bypass constraint is harder to satisfy with secret price offers. In the case of retail price competition, the bypass constraint is easier to satisfy and an upstream cartel is still able to make strictly positive profits for any

[14]Note that this value is different from $w_2^r(w_1)$ as the manufacturer might also attract demand from retailer 1.

strictly positive discount factor and any level of cost asymmetry $c > 0$. With quantity competition, the bypass constraint limits cartel sustainability for discount factors close to zero if retailers are not sufficiently differentiated. As a result, with secret offer, retailer 1 needs to be relatively more efficient relative to the benchmark model to allow collusion for low discount factors.

6 Vertical Integration

The results so far raise interesting questions for the assessment of vertical mergers in this setting. In what follows, I briefly discuss the effects of vertical integration on cartel stability when a manufacturer merges with the dominant and the less efficient retailer, respectively.

First consider the case where the less efficient retailer is vertically integrated with a manufacturer while the dominant retailer is independent. Suppose there is price competition downstream and firms try to implement a collusion scheme (w_1, p_1, p_2) where w_1 is the cartel's wholesale price to R1 and (p_1, p_2), with $w_1 < p_1 < p_2$, the downstream prices charged by R1 and the integrated downstream unit, respectively. The cartel provides a positive bond to the dominant retailer, that is, it holds that $(p_1 - w_1)D(p_1) > cD(c)$. Additionally, on the equilibrium path, each of the upstream firms, including the integrated upstream unit, supplies an equal share of the dominant retailer's input.[15]

What is the effect of vertical integration in this setting? First, there is what Nocke and White (2007) call the "outlet effect". After the vertical merger, there are no more upstream defections that sell through the downstream unit of the integrated firm. Hence, there are no more bypass deviations, which removes one of the cartel's constraints. Next, upstream deviations of non-integrated manufacturers using R1 can be punished by the downstream unit of the integrated firm in the same period as they occur. Accordingly, any deviation price $\hat{w}_1 \geq c$ is unprofitable as R1 is priced out of the market in the same period. Moreover, for any $\hat{w}_1 < c$, retailer 1 has no incentive to accept the deviation

[15] This assumption simplifies the exposition without affecting the qualitative results of the analysis.

since

$$(c - \hat{w}_1)D(c) - (p_1 - w_1)D(p_1) < \frac{\delta}{1-\delta}[(p_1 - w_1)D(p_1) - cD(c)]$$

for any $\hat{w}_1 \geq 0$.[16] Finally, consider the incentives of the integrated firm to deviate in the downstream market. The integrated downstream unit artificially inflates its price p_2 above its marginal cost c to allow a positive bond for the dominant retailer. The optimal deviation at the downstream pricing stage is to undercut the dominant retailer at $p_1 - \epsilon$ and serving the entire downstream market using the supply of the integrated upstream unit. In fact, such a deviation is profitable for the integrated firm since

$$(p_1 - c)D(p_1) > p_1 D(p_1) - cD(c) > w_1 D(p_1) > \frac{w_1 D(p_1)}{n(1-\delta)},$$

where the second inequality holds due to the positive bond for R1 and the third inequality is true for discount factors not satisfying (1). Hence, vertical integration involving the less efficient retailer 2 destroys the sustainability of the cartel for low discount factors as cartel deviation opportunities reach through to the downstream segment.

Now consider a merger between a manufacturer and the dominant retailer 1. Firms try to sustain a collusive scheme in which each of the upstream units (including the integrated firm) supplies an equal share of the integrated dominant retailer's input requirement on the equilibrium path at a price w_1. Moreover, the firms agree to offer a wholesale price w_2 to independent retailer 2 which implies that the integrated firm serves the downstream market at $p_1 = c + w_2 - \epsilon$. Again, there is the outlet effect preventing any deviation with the downstream unit of the integrated firm. This means there is no bilateral deviation constraint for the cartel. In addition, a non-integrated manufacturer's deviation with R2 is unprofitable as the integrated firm can punish in the same period and serve the entire downstream market. Finally, the integrated firm could deviate and supply its downstream unit exclusively through its integrated manufacturer. This deviation is not profitable if the discounted stream of profits from its upstream and downstream unit is at least as large as the current period profits from being the only

[16]Check that the condition is hardest to satisfy for $\hat{w}_1 = 0$ such that the LHS is negative for any positive bond value on the RHS.

supplier plus the future competitive profits in the downstream sector, that is,

$$\frac{w_1 D(c+w_2)}{n(1-\delta)} + \frac{(c+w_2-w_1)D(c+w_2)}{1-\delta} \geq (c+w_2)D(c+w_2) + \frac{\delta c D(c)}{1-\delta},$$

which can be rewritten as

$$\frac{w_1 D(c+w_2)}{n(1-\delta)} \geq w_1 D(c+w_2) - \frac{\delta}{1-\delta} B(w_1, c+w_2).$$

Since $D(c+w_2) < D(max\{w_1, c\})$, this condition is easier to satisfy than condition (3-P) with independent firms in the benchmark model. With independent firms, a deviating manufacturer who wants to attract the dominant retailer cuts the wholesale price for the rival retailer to put more pressure on R1 to lower its own cost by accepting the deviation offer. This destroys product market rents for the dominant retailer. By contrast, an integrated firm internalizes this competition effect through its downstream unit and does not offer a lower wholesale price for retailer 2. Hence, deviations are less profitable and a vertical merger involving the dominant firm increases cartel stability. We can thus summarize as follows.

Proposition 4. *Consider downstream price competition. A vertical merger between a manufacturer and the dominant retailer increases the stability of the upstream cartel. Vertical integration between a manufacturer and the less efficient retailer decreases the cartel's sustainability.*

The US Non-Horizontal merger guidelines (1984, section 4.222) state that a vertical merger involving a "sufficiently important" downstream firm may facilitate collusion in the upstream market. The reasoning is that with an independent dominant retailer, upstream firms are tempted to disrupt the cartel agreement in an effort to secure business with this retailer. In the argument in this paper, the dominant retailer is implicitly part of the upstream cartel. Nevertheless, there is a channel through which vertical integration involving this retailer reduces the scope for upstream collusion. An integrated manufacturer-dominant retailer pair stands to gain less from deviating relative to independent manufacturers who can steal upstream market share and appropriate rents from the downstream segment. Additionally, as shown, a vertical merger with the smaller downstream firm eliminates the potential for collusion (for low discount factors) and is pro-competitive as cartel deviations can reach through to the downstream market. Thus, although for different reasons, the results in this paper warrant the caution in the merger guidelines with respect to vertical merger involving dominant retailers.

7 Collusion with Two-Part Tariffs

In the baseline model, the upstream cartel creates the bond with the dominant retailer using linear wholesale prices. In the following analysis, I introduce two-part tariffs and show that while fixed fees can extract more surplus from the downstream segment, they might also make deviations more profitable and collusion harder to sustain.

Suppose firms use two-part tariffs such that retailer j is offered (w_j, F_j) where $F_j \geq 0$ is a fixed fee independent of the amount of supplied input. It is clear that fixed fees are competed away in the competitive stage equilibrium, in which all manufacturer charge marginal cost and make zero profits. Assume that retailers compete in prices and that the cartel forms a positive bond with the dominant retailer where

$$B = (c + w_2 - w_1)D(c + w_2) - F_1 - cD(c).$$

Any fixed fee charged along the equilibrium path reduces the value of the bond for R1. Note that, with price competition, the manufacturers can only charge a strictly positive fixed fee to the dominant retailer serving the market. Hence, the equilibrium fee for retailer 2 is $F_2 = 0$.

Consider a bilateral deviation by an upstream firm offering $(\hat{w}_j, \hat{F}_j)$ to retailer j. Using the same argument as in the benchmark model, retailer 1 accepts a bilateral deviation if and only if

$$\hat{F}_1 + \hat{w}_1 D(c + \hat{w}_2) \leq F_1 + (min\{c + \hat{w}_2, w_1\})D(c + \hat{w}_2) - \frac{\delta}{1-\delta}B. \qquad (5)$$

The LHS gives the retailer's total cost when accepting the deviation offer. The RHS is the total input cost when rejecting minus the value of the bond with the upstream cartel. A deviating upstream firm chooses its deviation offers to maximize the current period profits, which are equivalent to the LHS of (5), subject to satisfying this condition. This implies that the optimal deviation uses $(\hat{w}_1, \hat{F}_1)$ to satisfy (5) with equality and a linear fee $\hat{w}_2 = \max\{w_1 - c, 0\}$ to maximize the RHS. This yields the following bilateral deviation constraint for the cartel,

$$\frac{w_1 D(c + w_2) + F_1}{n(1 - \delta)} \geq w_1 D(\max\{w_1, c\}) + F_1 - \frac{\delta}{1 - \delta} B.$$

What is the effect of the fixed fee F_1 on the sustainability of this condition? A fixed fee allows the cartel to extract rents from retailer 1 which increases the cartel profits on the LHS. At the same time, it increases the current period deviation profits as the cartel's fixed fee makes it easier for a deviating manufacturer to attract the dominant retailer. This second effect outweighs the positive effect for low discount factors, for which condition (1) is not satisfied. Hence, fixed fees make the bilateral deviation constraint harder to sustain.

Now consider the bypass deviation. Retailer 2 accepts the deviation offer $(\hat{w}_2, \hat{F}_2)$ if

$$(w_1 - c - \hat{w}_2) D(w_1) - \hat{F}_2 \geq 0.$$

The optimal deviation for the upstream firm is to charge $\hat{w}_2 = 0$ and a fixed fee $\hat{F}_2$ that extracts the downstream deviation profits of retailer 2. This gives the following bypass deviation constraint for the cartel,

$$\frac{w_1 D(c + w_2) + F_1}{n} \geq (w_1 - c) D(w_1).$$

In this case, the fixed fee F_1 increases the cartel's equilibrium path profits without affecting the deviation profits with retailer 2. The bypass deviation constraint is thus easier to sustain and we get the following result.

Proposition 5. *Consider a cartel using two-part tariffs when retailers compete in prices. Charging retailers a fixed fee may increase or decrease the sustainability of the cartel.*

While fixed fees allow the cartel to extract more rents, they reduce the bond with the retailer and can make bilateral deviations more profitable. Whether a fixed fee then facilitates collusion or not depends on the parameter values and on which deviation constraint is binding. In particular, if the discount factor is low and the cartel price satisfies $w_1 \leq c$, the bilateral deviation constraint is binding and the cartel is easier to sustain without fixed fee. This means that the same qualitative results hold as in the benchmark model.

8 Conclusions

In this paper I show that the presence of a dominant retailer is highly conducive to successful upstream collusion. By price discriminating in favor of the dominant retailer, an upstream cartel is able to make strictly positive profits for any strictly positive value of the discount factor. Price discrimination creates a bond with the favored retailer which reduces or eliminates this firm's incentive to accept deviation offers from upstream cartel members. At the same time, using the less efficient rival retailer in a deviation is not profitable as it gives insufficient access to the downstream market. As a consequence, price discrimination increases cartel stability.

This result is robust with respect to the type of downstream competition, the possibility of secret offers and the type of contracts used by the cartel. When retailers are unable to observe their rival's offers and hold wary beliefs, there are two opposing effects on cartel stability. Secret offers invite opportunistic behaviour from the deviating upstream firm which makes the dominant retailer more reluctant to accept deviations. This effect strengthens the cartel. At the same time, unobservable offers introduce undetectable, on-schedule deviation which make deviations using the less efficient retailer more profitable and the cartel harder to sustain. Similarly, when the cartel charges two-part tariffs, it can use fixed fees to extract more rents from downstream firms. However, fixed fees also make it easier for a deviating upstream firm to attract the dominant retailer. In both cases, the qualitative results of the analysis are unchanged.

The long-standing discussion of the competitive effects of price discrimination has focused on static competition and unilateral conduct. In the context of coordinated price setting, this paper highlights an unambiguously negative effect of price discrimination. The analysis has thus potentially important implications for antitrust authorities. In vertically related industries, evaluating the potential of upstream collusion requires careful consideration of downstream market concentration. This is particularly relevant for structural cartel screening as well as for the assessment of horizontal (downstream) and vertical mergers. Furthermore, in this industry setting, upstream competition destroys any potential for price discrimination. Hence, price discrimination in itself is a marker for collusive upstream conduct. This holds, *a fortiori*, for low values of the discount factor, when price discrimination (in favor of the dominant retailer) is a necessary prerequisite for successful collusion.

References

Abreu, D. (1986). Extremal equilibria of oligopolistic supergames. *Journal of Economic Theory 39*(1), 191–225.

Aguirre, I., S. Cowan, and J. Vickers (2010). Monopoly price discrimination and demand curvature. *American Economic Review 100*(4), 1601–1615.

Armstrong, M. and J. Vickers (2001). Competitive price discrimination. *RAND Journal of Economics 32*(4), 579–605.

DeGraba, P. (1990). Input market price discrimination and the choice of technology. *American Economic Review 80*(5), 1246–1253.

Gu, D., Z. Yao, W. Zhou, and R. Bai (2019). When is upstream collusion profitable? *RAND Journal of Economics 50*(2), 326–341.

Hart, O. and J. Tirole (1990). Vertical integration and market foreclosure. *Brookings Papers on Economic Activity. Microeconomics. 1990*, 205–276.

Hatfield, J. W., S. D. Kominers, R. Lowery, and J. M. Barry (2019). Collusion in markets with syndication. Harvard Business School Entrepreneurial Management Working Paper No. 18-009. Available at SSRN: https://ssrn.com/abstract=3007819.

Inderst, R. and T. Valletti (2009). Price discrimination in input markets. *RAND Journal of Economics 40*(1), 1–19.

Jullien, B. and P. Rey (2007). Resale price maintenance and collusion. *RAND Journal of Economics 38*(4), 983–1001.

Katz, M. L. (1987). The welfare effects of third-degree price discrimination in intermediate good markets. *The American Economic Review 77*(1), 154–167.

Levenstein, M. C. and V. Y. Suslow (2006). What determines cartel success? *Journal of Economic Literature 44*, 43–95.

Mailath, G. J., V. Nocke, and L. White (2017). When and how the punishment must fit the crime. *International Economic Review 58*(2), 315–330.

McAfee, R. P. and M. Schwartz (1994). Opportunism in multilateral vertical contracting: Nondiscrimination, exclusivity and uniformity. *American Economic Review 84*(1), 210–230.

Nocke, V. and L. White (2007). Do vertical mergers facilitate upstream collusion. *American Economic Review 97*(4), 1321–1339.

Nocke, V. and L. White (2010). Vertical merger, collusion, and disruptive buyers. *International Journal of Industrial Organization 28*, 350–354.

O'Brien, D. and G. Shaffer (1992). Vertical control with bilateral contracts. *RAND Journal of Economics 23*, 299–308.

Piccolo, S. and J. Miklos-Thal (2012). Colluding through suppliers. *RAND Journal of Economics 43*(3), 492–513.

Rey, P. and J. Tirole (2007). A primer on foreclosure. In M. Armstrong and R. H. Porter (Eds.), *Handbook of Industrial Organisation*, Volume 3, Chapter 33, pp. 2147–2215. Elsevier BV.

Robinson, J. (1933). *The Economics of Imperfect Competition.* London: Macmillan.

Schinkel, M. P., J. Tuinstra, and J. Rueggeberg (2008). Illinois walls: how barring indirect purchaser suits facilitates collusion. *RAND Journal of Economics 39*(3), 683–698.

Schmalensee, R. (1981). Output and welfare implications of monopolistic third-degree price discrimination. *American Economic Review 71*(1), 242–247.

Stole, L. A. (2007). Price discrimination in competitive envrionments. In M. Armstrong and R. H. Porter (Eds.), *Handbook of Industrial Organisation*, Volume 3, Chapter 34, pp. 2221–2292. Elsevier BV.

Varian, H. (1985). Price discrimination and social welfare. *American Economic Review 75*(4), 870–875.

Vives, X. (1999). *Oligopoly Pricing: Old Ideas and New Tools.* Cambridge, Maa.: MIT Press.

Yoshida, Y. (2000). Third-degree price discrimination in input markets: Output and welfare. *American Economic Review 90*(1), 240–246.

Appendix

Proof of Lemma 1

First consider price competition downstream. For $w_1 \leq c + w_2 \leq p^m(w_1)$, the upstream profits are $M = w_1 D(c + w_2)$ which for given w_2 are maximized at $w_1 = c + w_2$.

For $c + w_2 < w_1 \leq p^m(c + w_2)$, profits are $M = w_2 D(w_1)$ and the optimal price for a given w_2 is $w_1 = c + w_2 + \epsilon$, where $\epsilon > 0$ is the small currency unit. Since $w_2 D(c + w_2 + \epsilon) < (c + w_2)D(c + w_2)$ for ϵ sufficiently small, it is always optimal to decrease w_1 from $c+w_2+\epsilon$ to $c+w_2$. It follows that the optimal prices for $c+w_2 \leq p^m(w_1)$ and $w_1 \leq p^m(c + w_2)$ satisfy $w_1 = c + w_2$. Using this and maximizing $w_1 D(w_1)$ yields $w_1^* = p^m(0)$ and $w_2^* = p^m(0) - c$. Finally, for $c + w_2 > p^m(w_1)$, the highest achievable upstream profits are $\pi^m(0) = p^m(0)D(p^m(0))$ which, by continuity of demand, are always dominated by setting $w_1 = p^m(0)$ and any $w_2 \in [p^m(0) - c, p^m(0))$. Similarly, for $w_1 > p^m(c + w_2)$, upstream profits are maximized at $w_2' = \arg\max_{w_2} w_2 D(c + w_2)$. This solution is dominated by any $w_1 \in (c + w_2, p^m(c + w_2')]$ and w_2'.

The Cournot equilibrium quantities solve for each firm j,

$$q^c(m_j, m_{-j}) = \arg\max_q [P(q + q^c(m_{-j}, m_j)) - m_j]q$$

which yields as first-order condition

$$P(Q^c) + P'(Q^c)q^c(m_j, m_{-j}) - m_j = 0,$$

where $Q^c = q^c(m_j, m_{-j}) + q^c(m_{-j}, m_j)$. Adding the first-order conditions of the firms yields

$$P(Q^c) + P'(Q^c)Q^c = m_1 + m_2$$

and the statement in the main text follows. The first-order conditions of the cartel's problem are

$$\frac{\partial M}{\partial w_1} = q^c(w_1, c + w_2) + w_1 \frac{\partial q^c(w_1, c + w_2)}{\partial w_1} + w_2 \frac{\partial q^c(c + w_2, w_1)}{\partial w_1} = 0$$
$$\frac{\partial M}{\partial w_2} = q^c(c + w_2, w_1) + w_2 \frac{\partial q^c(c + w_2, w_1)}{\partial w_2} + w_1 \frac{\partial q^c(w_1, c + w_2)}{\partial w_2 x} = 0$$

Substituting $q^c(c + w_2, w_1) = Q^c - q^c(w_1, c + w_2)$ in the second condition and equating the LHS of both conditions gives the equation in the main text. The lemma follows. QED.

Proof of Lemma 2

First consider price competition. Note that firm 1 makes $cD(c) > 0$ on the punishment path. If $w_1 \geq c + w_2$, firm is unable to attract demand on the equilibrium path and the bond is always negative. For $w_1 < c + w_2 \leq p^m(w_1)$, the bond of firm 1 is positive if and only if

$$B(w_1, c + w_2) = (c + w_2 - w_1)D(c + w_2) - cD(c) \geq 0$$

or

$$w_1 \leq c + w_2 - \frac{cD(c)}{D(c + w_2)} < w_2.$$

This threshold equals zero at $w_2 = 0$ and increases in w_2 with a slope

$$\begin{aligned}\frac{dw_1}{dw_2} &= 1 + \frac{cD(c)D'(c + w_2)}{D(c + w_2)^2} \\ &= \frac{1}{D(c + w_2)^2}[(c + w_2 - w_1)D'(c + w_2) + D(c + w_2)] \geq 0\end{aligned}$$

for any $c + w_2 \leq p^m(w_1)$. The slope is equal to zero at $c + w_2 = p^m(w_1)$. For higher values, $c + w_2 > p^m(w_1)$, the bond of firm 1 is independent of w_2 and positive if and only if $\pi^m(w_1) - cD(c) \geq 0$. By contrast, firm 2 makes zero profits on the punishment path under price competition. Hence, $B(c + w_2, w_1) > 0$ if and only if $w_2 < w_1 - c$. It will further be useful to show that

$$\frac{\partial B(w_1, c + w_2)}{\partial c} = D(c + w_2) + (c + w_2)D'(c + w_2) - D(c) - cD'(c) - w_1 D'(c + w_2) < 0$$

for all $c + w_2 < p^m(w_1)$. This condition is harder to satisfy for the highest feasible value $w_1 = w_2$ and we get

$$D(c + w_2) + cD'(c + w_2) - D(c) - cD'(c) < 0.$$

The LHS equals 0 at $w_2 = 0$ and decreases in w_2 since $D'(c + w_2) + cD''(c + w_2) < 0$ for all $c + w_2 < p^m(w_2) = p^m(w_1)$.

Next consider quantity competition and check that along $B(m_j, m_{-j}) = 0$ we get

$$\frac{dw_j}{dw_{-j}} = -\frac{\partial\pi/\partial w_{-j}}{\partial\pi/\partial w_j} = \frac{P'\frac{\partial Q^c}{\partial m_j}q^c(m_j, m_{-j}) + (P - m_j)\frac{\partial q^c(m_j,m_{-j})}{\partial m_{-j}}}{P'\frac{\partial Q^c}{\partial m_j}q^c(m_j, m_{-j}) + (P - m_j)\frac{\partial q^c(m_j,m_{-j})}{\partial m_j} - q^c(m_j, m_{-j})} < 1$$

if and only if

$$(P - m_j)(\frac{\partial q^c(m_j, m_{-j})}{\partial m_{-j}} - \frac{\partial q^c(m_j, m_{-j})}{\partial m_j}) + q^c(m_j, m_{-j}) > 0$$

which is always satisfied due to assumption (A1). Since at $(w_1, w_2) = (0, 0)$ we have $B(m_j, m_{-j}) = 0$ and the slope of the cross-derivative is less than 1, it has to hold that if $B(m_j, m_{-j}) \geq 0$, then $w_j < w_{-j}$. The lemma follows. QED.

Proof of Lemma 4

The LHS of (4-P) takes value 0 at $w_1 = 0$ and linearly increases in w_1. The RHS is strictly concave, maximized at $w_1 = p^m(c)$ and takes value 0 at $w_1 = c$ and $w_1 = \bar{p}$. Either the LHS is higher for any w_1 (which is not the case for $B(w_1, c+w_2) \geq 0$ as shown below) or solving condition (4-P) as an equation yields two roots. Note that the LHS decreases in w_2 and equals zero for $w_2 = \bar{p}$. Hence, the smaller root decreases in w_2 and satisfies $w_1 = c$ at $w_2 = \bar{p}$ while the larger root increases in w_2 and takes value $w_1 = \bar{p}$ at $w_2 = \bar{p}$. A positive bond $B(w_1, c+w_2) \geq 0$ implies $w_1 \leq w_2$. Hence, if the larger root is greater than w_2 for $w_2 = w_1$, then only the smaller root can satisfy $w_1 \leq w_2$. To see that this is indeed the case, check that at $w_2 = w_1$ the LHS is strictly concave in w_1 and takes value 0 at $w_1 = 0$ and at $w_1 = \bar{p} - c < \bar{p}$. Hence, LHS and RHS intersect exactly one and there exists a unique root ω_1' for $w_1 \leq w_2$ such that if $w_1 \leq \omega_1'$, then condition (4-P) is satisfied.

For the comparative statics in the lemma, rewrite the constraint as equality such that

$$\psi = w_1 D(c + w_2) - n(w_1 - c)D(w_1) = 0.$$

Check that

$$\frac{\partial \psi}{\partial w_1} = D(c + w_2) - (w_1 - c)nD(w_1) - nD(w_1) < 0$$

since at the smaller root of (4-P) the slope of the RHS is larger than the slope of the LHS. Moreover, verify that $\partial\psi/\partial w_2 = w_1 D'(c + w_2) < 0$ and

$$\frac{\partial \psi}{\partial c} = w_1 D'(c + w_2) + nD(w_1) = \frac{w_1}{w_1 - c}[(w_1 - c)D'(c + w_2) + D(c + w_2)] > 0.$$

The inequality holds since the term in the squared bracket decreases in w_1 and a positive bond implies $w_1 \leq c + w_2$. Thus, at $w_1 = c + w_2$, the squared bracket is minimized. Substituting $p = c + w_2$, the squared bracket can be written as

$$(p - c)D'(p) + D(p) > 0$$

for all $p = c + w_2 < p^m(c) < p^m(w_1)$. The comparative statics results are therefore

$$\frac{dw_1}{dw_2} = -\frac{\partial\psi/\partial w_2}{\partial\psi/\partial w_1} < 0, \frac{dw_1}{dc} = -\frac{\partial\psi/\partial w_2}{\partial\psi/\partial w_1} > 0.$$

As last step, I show that $\omega_1' \leq p^m(0)$ for $B(w_1, c + w_2) \geq 0$. From Lemma 2, we know that $dw_1/dw_2 \geq 0$ along $B(w_1, c + w_2) = 0$ and $p^m(w_1) \geq c + w_2$. At the same time, it holds that $dw_1/dw_2 < 0$ for the bypass constraint. The highest admissible w_1 is thus either at the point where both conditions hold with equality or at the highest value such that only $B(w_1, c + w_2) = 0$ is binding while (4-P) is slack. The highest value w_1 that satisfies $B(w_1, c + w_2) = 0$ is implicitly defined by $p^m(w_1) = c + w_2$ (see Lemma 2). Let $(\bar{w}_1', \bar{w}_2')$ be the intersection where both constraints hold jointly. Further let $(\bar{w}_1'', \bar{w}_2'')$ be the intersection of the bypass constraint and $p^m(w_1) = c + w_2$. Verify the following four properties: *(i)* $\bar{w}_2'(c = 0) = 0$. *(ii)* $\partial\bar{w}_2'/\partial c > 0$, since $dw_1/dc < 0$ along $B(w_1, c + w_2) = 0$ (see proof of Lemma 2) and $dw_1/dc > 0$ when (4-P) holds with equality. *(iii)* $\bar{w}_2''(c = 0) = p^m(0)$ since $w_1 = 0$ and $c + w_2 = p^m(0)$. *(iv)* $\partial\bar{w}_2''/\partial c < 0$ since

$$\frac{\partial\bar{w}_2''}{\partial c} = \frac{\partial p^m(\bar{w}_1'')}{\partial\bar{w}_1''}\frac{\partial\bar{w}_1''}{\partial c} - 1 < 0.$$

To see this, note that $\partial p^m(w)/\partial w < 1$. Moreover, from

$$\frac{\partial \bar{w}_1''}{\partial c} = -\frac{D(w_1)}{[D(p^m(w_1)) + w_1 D'(p^m(w_1)\partial p^m/\partial w_1)]/n - (w_1 - c)D'(w_1) - D(w_1)},$$
$$\frac{\partial^2 \bar{w}_1''}{(\partial c)^2} = \frac{D(w_1)D'(w_1)}{[[D(p^m(w_1)) + w_1 D'(p^m(w_1)\partial p^m/\partial w_1)]/n - (w_1 - c)D'(w_1) - D(w_1)]^2} < 0,$$

follows that the highest slope is at $c = 0$ where $w_1 = 0$ and

$$\frac{\partial \bar{w}_1''}{\partial c}|_{c=0} = \frac{D(0)}{D(0) - D(p^m(0))/n} < 1.$$

From the four properties above follows that there exists a $\bar{c}$ such that if and only if $c \leq \bar{c}$, then $\bar{w}_2' \leq \bar{w}_2''$. Note that if $\bar{w}_2' \leq \bar{w}_2''$, then the highest value of w_1 is given by $\bar{w}_1'$; otherwise, the bypass constraint is holding if $B(w_1, c + w_2) \geq 0$. The former case arises if $c \leq \bar{c}$ and it has to hold that

$$\bar{w}_1' \leq \bar{w}_2' \leq \bar{w}_2'' < p^m(0).$$

For $c > \bar{c}$, the highest value $w_1 = \bar{w}_1'''$ satisfies $B(w_1, c + w_2) = 0$ and $p^m(w_1) = c + w_2$ which yields

$$(p^m(\bar{w}_1''') - \bar{w}_1''')D(p^m(\bar{w}_1''')) = cD(c).$$

It follows that $\partial \bar{w}_1'''/\partial c < 0$ for all $c \leq p^m(0)$. Hence, it has to hold that

$$\bar{w}_1'''(c) \leq \bar{w}_1'''(c = \bar{c}) = \bar{w}_2'' < p^m(0).$$

This means $\omega_1' \leq p^m(0)$ and the lemma follows. QED.

Proof of Lemma 5

First consider deviations such that $\hat{w}_2 \leq w_1 - c$. Since $\hat{w}_2 < p^m(\hat{w}_1) - c$, total differentiation yields

$$\frac{d\hat{w}_1}{d\hat{w}_2} = \frac{D(c+\hat{w}_2) + (c+\hat{w}_2-\hat{w}_1)D'(c+\hat{w}_2)}{D(c+\hat{w}_2)} > 0, \qquad \frac{d\hat{w}_1}{dw_1} = \frac{\delta}{1-\delta} > 0.$$

For deviations to $\hat{w}_2 > w_1 - c$, the comparative statics are

$$\frac{d\hat{w}_1}{d\hat{w}_2} = \frac{(w_1-\hat{w}_1)D'(c+\hat{w}_2)}{D(c+\hat{w}_2)} < 0, \qquad \frac{d\hat{w}_1}{dw_1} = \frac{1}{1-\delta} > 0$$

In both cases, total differentiation yields

$$\frac{d\hat{w}_1}{dw_2} = -\frac{\delta[D(c+w_2) + (c+w_2-w_1)D'(c+w_2)]/(1-\delta)}{D(c+\hat{w}_2)} < 0. \quad \text{QED.}$$

Proof of Lemma 6

The optimal deviation is at the kink of the threshold function ω_1 if the slope of the iso-profit curve of the deviating manufacturer is less than the slope of ω_1 for any $\hat{w}_2 \leq w_1 - c$. Totally differentiating the deviation profits of a manufacturer $M = \hat{w}_1 D(c+\hat{w}_2)$ yields

$$\frac{d\hat{w}_2}{d\hat{w}_1} = -\frac{\hat{w}_1 D'(c+\hat{w}_2)}{D(c+\hat{w}_2)}.$$

This is less than the slope of ω_1 if and only if

$$-\frac{\hat{w}_1 D'(c+\hat{w}_2)}{D(c+\hat{w}_2)} < \frac{\partial \omega_1}{\partial \hat{w}_2} = 1 + (c+\hat{w}_2-\hat{w}_1)\frac{D'(c+\hat{w}_2)}{D(c+\hat{w}_2)} \quad \text{or}$$

$$1 + (c+\hat{w}_2)\frac{D'(c+\hat{w}_2)}{D(c+\hat{w}_2)} > 0 \quad \text{or} \quad D(c+\hat{w}_2) + (c+\hat{w}_2)D'(c+\hat{w}_2) > 0$$

which holds for $\hat{w}_2 \leq p^m(0) - c$ and, since $\hat{w}_2 \leq w_1 - c$, for $w_1 \leq p^m(0)$.

The kink does not exist for $w_1 \leq c$. In this case the optimal deviation is at $\hat{w}_2 = 0$ and

$$\hat{w}_1 = w_1 - \frac{\delta}{1-\delta}\frac{B(w_1, c+w_2)}{D(c)} < 0$$

if and only if

$$w_1 < \frac{\delta[(c+w_2)D(c+w_2) - cD(c)]}{(1-\delta)D(c) + \delta D(c+w_2)}.$$

The RHS is strictly positive for any $c > 0$, $w_2 > 0$ and $\delta > 0$. Hence, strictly positive cartel prices exist such that a bilateral deviation cannot yield strictly positive profits. QED.

Proof of Lemma 7

The threshold price ω_1 is defined as

$$\pi^c(\omega_1, c+\hat{w}_2) - \pi^c(w_1, c+\hat{w}_2) = \frac{\delta}{1-\delta}[\pi^c(w_1, c+w_2) - \pi^c(0,c)].$$

Check that

$$\frac{d\omega_1}{d\hat{w}_2} = -\frac{\partial\pi^c(\omega_1, \hat{m}_2)/\partial\hat{m}_2 - \partial\pi^c(w_1, \hat{m}_2)/\partial\hat{m}_2}{\partial\pi^c(\omega_1, \hat{m}_2)/\partial\omega_1} > 0$$

since the denominator is negative, and the numerator is positive due to $\omega_1 < w_1$ and

$$\frac{\partial^2\pi^c(m_j, m_{-j})}{\partial m_j \partial m_{-j}} = -\frac{\partial q^c(m_j, m_{-j})}{\partial m_{-j}} < 0.$$

Consider a bilateral deviation to the highest acceptable offer to retailer 2, which just shaves the equilibrium price w_2. Let ω_1' be the maximum price for retailer 1 at this deviation, that is,

$$\pi^c(\omega_1', c+w_2) - \pi^c(w_1, c+w_2) = \frac{\delta}{1-\delta}[\pi^c(w_1, c+w_2) - \pi^c(0,c)].$$

Notice that

$$\frac{d\omega_1'}{dw_1} = -\frac{-(1/(1-\delta))\partial\pi^c(w_1, m_2)/\partial w_1}{\partial\pi^c(\hat{m}_1, m_2)/\partial\hat{m}_1} > 0$$

as the numerator is positive and the denominator is negative. Furthermore, as w_1 approaches zero, the RHS of the definition of ω_1' is strictly positive for $w_2 > 0$, which means that $\omega_1' < 0$ at $w_1 = 0$. It follows that there must exist a value $w_1^o > 0$ such that if $w_1 < w_1^o$, then there exists no positive deviation offer that retailer 1 would accept. Moreover, since $d\omega_1/d\hat{w}_2 > 0$, no other deviation $\hat{w}_2$ allows for a positive deviation offer that retailer 1 accepts. Furthermore, check that w_1^o is strictly decreasing in δ and takes value zero at $\delta = 0$. Hence, the bilateral deviation constraint can be satisfied with strictly positive prices for any $\delta > 0$. QED.

Proof of Lemma 8

Let

$$w_2^r(w_1) = \arg\max_{w_2} \quad w_2 q^c(c + w_2, w_1)$$

and

$$w_2^R(w_1) = \arg\max_{w_2} \quad w_1 q^c(w_1, c + w_2) + w_2 q^c(c + w_2, w_1)$$

where $w_2^r(w_1) < w_2^R(w_1)$ for all $w_1 > 0$ and $w_2^r(0) = w_2^R(0)$. Two cases arise. First, suppose $w_2 \geq w_2^r(w_1)$. In this case the optimal deviation is to $w_2^r(w_1)$ and - for a given w_1 - the bypass constraint is sustainable if

$$\frac{1}{n}M(w_1, w_2^R(w_1)) \geq w_2^r(w_1) q^c(c + w_2^r(w_1), w_1).$$

Let $\hat{c}(w_1)$ be the value such that this condition holds with equality. Check that as w_1 approaches zero, this constraint can only be satisfied if $c = \bar{c}$ where $\bar{c} = \{c | q^c(c, 0) = 0\}$. Hence, $\hat{c}(w_1 = 0) = \bar{c}$. Furthermore, let Ψ denote the difference between LHS and RHS of the bypass condition. Check that

$$\frac{\partial\Psi}{\partial w_1} = \frac{1}{n}[w_1 \frac{\partial q^c(w_1, m_2)}{\partial w_1} + q^c(w_1, m_2) + w_2^R(w_1)\frac{\partial q^c(m_2, w_1)}{\partial w_1}] - w_2^r(w_1)\frac{\partial q^c(c + w_2^r(w_1), w_1)}{\partial w_1}$$

with $m_2 = c + w_2^R(w_1)$. At $w_1 = 0$ it holds that $c = \bar{c}$ and, hence, $w_2^r(w_1) = w_2^R(w_1) = 0$. This implies $\partial\Psi/\partial w_1 = q^c(0, \bar{c}) > 0$. Due to the continuity of all functions, the derivative is also positive for any w_1 sufficiently small. Similarly,

$$\frac{\partial \Psi}{\partial c} = \frac{1}{n}[w_1 \frac{\partial q^c(w_1, m_2)}{\partial m_2} + w_2^R(w_1) \frac{\partial q^c(m_2, w_1)}{\partial m_2}] - w_2^r(w_1) \frac{\partial q^c(c + w_2^r(w_1), w_1)}{\partial c}$$

At $w_1 = 0$ it holds that $\partial\Psi/\partial c = 0$. For w_1 sufficiently small, the derivative is positive since $w_2^R \approx w_2^r$ and

$$0 > \frac{1}{n} \frac{\partial M(w_1, w_2)}{\partial c} > \frac{\partial M(w_1, w_2)}{\partial c} > w_2 \frac{\partial q^c(m_2, w_1)}{\partial c}.$$

From this follows that the total differential $dc/dw_1 = -(\partial\Psi/\partial w_1)/(\partial\Psi/\partial c) < 0$ for w_1 sufficiently small and $w_2 = w_2^R(w_1)$, Hence, there must exist values for c, with $0 < \hat{c}(w_1) \leq c < \bar{c}$, such that this constraint is sustainable.

If $w_2 < w_2^r(w_1)$, then the optimal deviation is to slightly shave the cartel offer for retailer 2. In this case, the bypass condition is given by

$$\frac{1}{n} w_1 q^c(w_1, c + w_2) \geq \frac{n-1}{n} w_2 q^c(c + w_2, w_1).$$

Consider $n = 2$ which makes it easiest to satisfy this constraint. Consider $c = 0$. In this case, the LHS is equal to the RHS if $w_2 = w_1$. Moreover, as w_1 goes to zero, the RHS approaches zero while the LHS is strictly positive for all $w_2 > 0$. This implies that for w_2 in $(w_1, w_2^r(w_1)]$ and any $B(w_1, w_2) > 0$ the RHS is larger than the LHS. Furthermore, $B(w_1, w_2) > 0$ requires that $w_2 > w_1$ for any $w_1 > 0$. For given continuous functions, this means that if the bypass constraint can be satisfied with $w_2 < w_2^r(w_1)$, then it must be for values of c such that $c \geq c' > 0$. QED.

Proof of Proposition 3

1. Downstream Price Competition. Given the trigger strategy, retailer 2's choice does not affect the continuation value. For prices such that $B_1(w_1, w_2) > 0$, retailer 2 is weakly better off with any deviation $\hat{w}_2 < w_2$. Hence, from the point of view of retailer 1, we have

$$a_2(\hat{w}_2, W_1(\hat{w}_2)) = \hat{w}_2 = W_2(\hat{w}_1)$$

while the optimal price for retailer 2 as a function of the offer for retailer 1 is simply

$$W_2(\hat{w}_1) = \hat{w}_2 = \max\{0, \hat{w}_1 - c\}.$$

This implies no deviation $\hat{w}_1 \geq c$ yields positive profits for retailer 1. For $\hat{w}_1 < c$, retailer 1 accepts if

$$(\min\{w_1, c\} - \hat{w}_1)D(c) \geq \frac{\delta}{1-\delta}B(w_1, w_2).$$

This yields the following bilateral deviation constraint

$$\frac{w_1 D(c + w_2)}{n(1-\delta)} \geq \min\{w_1, c\}D(c) - \frac{\delta}{1-\delta}B(w_1, c + w_2). \tag{3-S}$$

which is strictly easier to satisfy for $w_1 > c$ and equivalent to condition (3-P) in the benchmark model, otherwise.

Now consider a bypass deviation with $\hat{w}_1 \neq w_1$ which does not satisfy condition (3-S). This deviation does not trigger punishment and retailer 1 does not accept the deviator's offer. The manufacturer's current deviation profits are maximized at ($\hat{w}_1 = w_1 - \epsilon, \hat{w}_2 = \hat{w}_1 - c - \epsilon$) which yields the same condition as the bypass condition (4-P) in the benchmark model. Now consider a bypass deviation with $\hat{w}_1 = w_1$. This induces punishment from next period onwards. A deviation to $\hat{w}_2 = w_1 - c - \epsilon$ yields $(w_1 - c)D(w_1)$ and is not optimal if the above condition (4-P) holds. A deviation to $\hat{w}_2 = \max\{w_1 - c, 0\}$ ensures that retailer 1 wins the market at the lowest possible downstream price. The deviator gets $w_1 D(\max\{w_1, c\})/n$. If $w_1 \leq c$, then a deviation is not profitable if and only if

$$D(c + w_2) \geq (1-\delta)D(c). \tag{6}$$

For any $\delta > 0$, there exists a $w_2 > 0$ such that this condition holds with equality. This completes the proof of Proposition 3 for price competition.

2. Downstream Quantity Competition. Suppose $w_2 = w_2^R(w_1) \geq w_2^s(w_1)$. Note that $w_2^r(w_1) < w_2^s(w_1) < w_2^R(w_1)$ for $w_1 > 0$ and $w_2^r(0) = w_2^s(0) = w_2^R(0)$. Let $\hat{\hat{c}}(w_1)$ be the value such that condition (4-S) holds with equality. As w_1 approaches zero, condition (4-S) can only be satisfied if $c = \bar{c}$, that is, $\hat{\hat{c}}(0) = \bar{c}$. Furthermore, let Ψ denote the difference between LHS and RHS of condition (4-S). Check that

$$\begin{aligned}\frac{\partial \Psi}{\partial w_1} =& \frac{1}{n(1-\delta)}[w_1 \frac{\partial q^c(w_1, m_2)}{\partial w_1} + q^c(w_1, m_2) + w_2^R(w_1) \frac{\partial q^c(m_2, w_1)}{\partial w_1}] \\ &- \frac{1}{n}[q^c(w_1, c + w_2^s(w_1)) - w_1 \frac{\partial q^c(w_1, c + w_2^s(w_1))}{\partial w_1}] - w_2^s(w_1) \frac{\partial q^c(c + w_2^s(w_1), w_1)}{\partial w_1}\end{aligned}$$

with $m_2 = c + w_2^R(w_1)$. At $w_1 = 0$, it holds that $c = \bar{c}$ and $w_2^R(w_1) = w_2^s(w_1) = 0$. This implies $\partial \Psi / \partial w_1 = q^c(0, \bar{c}) > 0$ and, since all functions are continuous, the derivative must be positive for values of w_1 sufficiently close to zero. Further check that

$$\begin{aligned}\frac{\partial \Psi}{\partial c} =& \frac{1}{n(1-\delta)}[w_1 \frac{\partial q^c(w_1, m_2)}{\partial m_2} + w_2^R(w_1) \frac{\partial q^c(m_2, w_1)}{\partial m_2}] \\ &- \frac{1}{n} w_1 \frac{\partial q^c(w_1, c + w_2^s(w_1))}{\partial c} - w_2^s(w_1) \frac{\partial q^c(c + w_2^s(w_1), w_1)}{\partial c}\end{aligned}$$

At $w_1 = 0$ it holds that $\partial \Psi / \partial c = 0$. For w_1 sufficiently small, the derivative is positive since $w_2^R(w_1) \approx w_2^s(w_1)$ and

$$0 > \frac{1}{n(1-\delta)} \frac{\partial M(w_1, w_2)}{\partial c} > \frac{\partial M(w_1, w_2)}{\partial c} > \frac{1}{n} w_1 \frac{\partial q^c(w_1, m_2)}{\partial c} + w_2 \frac{\partial q^c(m_2, w_1)}{\partial c}.$$

From this follows that the total differential $dc/dw_1 = -(\partial \Psi / \partial w_1)/(\partial \Psi / \partial c) < 0$ for w_1 sufficiently small and $w_2 = w_2^R(w_1)$, Hence, there must exist values for c, with $0 < \hat{\hat{c}}(w_1) \leq c < \bar{c}$, such that the bypass constraint is sustainable. QED.