Partial Vertical Ownership and Tacit Collusion

by

Miguel Trigo Franco Frazão

Dissertation submitted in partial fulfillment of the requirements for the degree of

Master of Science in Economics

at the

Universidade Católica Portuguesa

Thesis Supervisor: Professor Duarte Brito

June 2017

[Page intentionally left blank]

Partial Vertical Ownership and Tacit Collusion

Miguel Frazão

Supervisor: Professor Duarte Brito

June 2, 2017

Abstract

A partial ownership held by a downstream firm creates a perceived cost asymmetry towards its competitors. In this article, it is shown that this will have a negative impact on the sustainability of a collusive scenario. This is a similar result to natural differences in production costs of firms. However, this participation makes it so it is more likely to be the most eficient firm to deviate, and not the least one, as in natural assymetry. The existance of participation never makes collusion easier to sustain than its absence. This also creates a tool for the upstream firm to break, or incentivate, joint downstream decision-making, as it may also be used to increase its directed demand. Similarly, this tool can be used by a regulator to increase welfare by avoiding market concentration.

Resumo

Uma participação parcial detida por uma empresa a jusante cria uma assimetria de custos perante os seus concorrentes. Neste artigo mostra-se que isto terá um impacto negativo na sustentabilidade de um cenário de conluio. Isto é um resultado semelhante a uma diferença natural nos custos de produção das empresas. Contudo, esta participação faz com que seja a empresa mais eficiente a mais provável de desviar da situação, em vez da menos eficiente, como em assimetria natural. A existência da participação nunca torna o conluio mais sustentável do que a sua ausência. Isto também cria uma ferramenta à empresa a montante para quebrar, ou incentivar, a decisão conjunta a jusante. Do mesmo modo, esta ferramenta pode ser usada por um regulador para aumentar o bem estar, quebrando o conluio.

[Page intentionally left blank]

Acknowledgements

First, I would like to thank my supervisor, Professor Duarte Brito, whose guidance and dedication were fundamental for the conclusion of this thesis. Without his patience and incentive, I wouldn't have been able to achieve this goal. I am also grateful to Professor Catarina Reis, Professor Pedro Teles and Professor Fernando Branco, who have inspired me and grown my passion for economics.

I would also like to thank my friends and colleagues, from many walks of life, for their support and advice (sometimes even technical), specially as they helped in providing meaning to my work.

I am specially grateful to my family. My cousins, who I've grown with and learned a lot from. My sister, who I've always looked up to. My parents, who have taught and showed me what really matters in life and have always demanded the best of me.

Finally, and above all, I'd like to thank God for gracing my life with opportunities that I never deserved.

[Page intentionally left blank]

Contents

1 Introduction

In several vertical industries, downstream firms partially own an upstream input supplier. This partial backward integration may be a mere financial interest or may involve some degree of control. Greenlee and Raskovich (2006) have focused on passive partial vertical ownership (PVO), that is characterized by the downstream firms having claims on upstream profits that involve no degree of control with respect to the upstream firm's decisions. In this setting, a downstream firm is, in fact, paying part of the input price to herself. This will have some impact on the equilibrium of the downstream competition stage. Greenlee and Raskovich (2006) analyze how the downstream firm's input/output choice is affected by different ownership patterns and establish that, under uniform ownership, PVO would have no effects.

Similarly to what happens in a horizontal merger framework, it is likely that PVO has other effects in addition to these unilateral ones. In particular, it is likely that there are also some coordinated effects involved. These effects refer to the possibility that downstream collusion, explicit or tacit, may be more or less easy to sustain in the presence of PVO. The purpose of this paper is to analyze how PVO without control affects downstream collusion and to incorporate these effects in the upstream decisions.

We model the industry downstream as an homogeneous product duopoly, that needs to purchase an input from an upstream monopolist, which may be regulated or not. One of the downstream firms is assumed to own a percentage of the upstream producer and therefore receive a percentage of its profit. These assumptions resemble closely the Portuguese electricity industry where REN, Redes Energéticas Nacionais, is an energy transmission operator with two major business areas: the transmission of electricity at very high voltage and the transport of high-pressure natural gas. One of the five larger shareholders of REN is the EDP group, a vertically integrated generator, distributor and supplier of electricity in Portugal who is also one of its clients. The downstream industry is relatively concentrated and the access conditions to the transmission network are regulated.

This paper is related to several strands of literature. Firstly, it is related to the literature on PVO, with or without control, that includes the above mentioned Greenlee and Raskovich (2006). Chen and Ross (2003) and Rossini and Vergari (2011) analyze the case of Input Production Joint Ventures owned in equal parts by duopolists where some degree of control is present. The first find that, under some assumptions, this joint venture may lead to the same outcomes as a full merger between the downstream firms. Rossini and Vergari (2011) analyze an oligopoly with differentiated goods and find conditions for these joint ventures to exist and find that firms' incentives may reduce welfare.

Secondly, it shares some modelling assumptions with the literature that analyzes the unilateral effects of partial interests in competitors or in production joint ventures, that is, partial horizontal ownership. Flath (1992), using Cournot industries where competitors own participations on each other, finds that "effects of horizontal shareholding interlocks are greater if firms are mindful of indirect shareholding links than if only attentive to direct links". Bresnahan and Salop (1986), who analyze different types of arrangements for joint ventures, conclude that the incentives of the competing firms depend on the type of financial interest and control arrangements. They propose a Modified Herfindahl-Hirshman Index to quantify these incentives. O'Brien and Salop (2000) state that it might happen that partial investments raise even larger concerns than full control. Similar to the previous article quoted, they consider several arrangements of financial interest and of control.

Both strands of literature above do not consider the effects of partial ownership on collusion. Gilo et al. (2006) discuss such effects, but for the case of horizontal ownership. They find necessary and sufficient conditions for which these arrangements facilitate collusion. Using a Bertrand oligopoly with n firms, they are able to prove that an increase in the participation between rivals never hinders collusion and that it may facilitate it. Foros et al. (2010) use a three firms Cournot game to find that a partial cross ownership with control might lead to higher joint profits than full merger. Malueg (1992) shows that "if firms interact repeatedly, then increasing cross ownership may reduce the likelihood of collusion. A high level of cross ownership may even entail a lower likelihood of collusion than would no cross ownership."

The main conclusions of this article are the following. Similarly to asymmetry, PVO makes collusion less likely to be sustained. However, contrary to the case of natural asymmetry between firms, it will be the most efficient firm the one with higher incentives to deviate. Just like under natural asymmetry, PVO imposes an upper bound on the price of the intermediate good for collusion to be sustainable. On the contrary, if there are efficiency gains from distributing production between the downstream firms, which can be caused by factors such as increasing marginal costs or differentiated goods, collusion might be more sustainable under PVO for an additional set of higher prices of the intermediate good. It is also concluded that the upstream firm might have an incentive to charge a different price to its buyers (the downstream firms) in order to prevent collusion. However, if there are gains from efficiency in having production distributed between firms, it may happen the opposite: that the firm upstream would choose a different price to encourage collusion. Finally, it could be shown that, from a social welfare perspective, it might be optimal to choose a price for the intermediate good above marginal costs as to break incentives for concertation.

The paper is organized as follows. Section 2 presents the model and the basic assumptions. Section 3 studies the impacts on the sustainability of collusion downstream for a given price of the intermediate good, while comparing this scenario with a natural costs asymmetry one. Section 4 presents two ways of having the price of the intermediate good endogenously determined: the cases of a natural monopolist firm upstream and a regulated one.

2 The model

Consider a vertical industry with two stages. An upstream monopolist, firm U, produces an input that is sold to two downstream firms, 1 and 2. The price for the input is assumed to be uniform (no price discrimination is allowed) and fixed for the long run and firm U is assumed to have no costs. There is either a long term contract that stipulates the unit price k or the price is fixed by a regulator for a long period of time. Downstream firms interact frequently and play an infinitely repeated Cournot game. That is, for a given upstream input price k, downstream firms may be able to sustain a collusive outcome. As usual in this type of literature, we will focus on the standard trigger strategies, which can be defined as "firm i chooses the collusive output in each period if no outcome other than that one has happened before, and reverts to the equilibrium of the one shot game otherwise". By collusive output we refer to the output that maximizes the joint profits of the downstream firms.

At the downstream level, we consider an homogeneous-product industry with two firms competing *a la Cournot*. The market demand is assumed to be linear, and given by $X^D = A - bP$, where X^D denotes the total quantity demanded, P denotes the unit price and A and b are two positive parameters. Each unit of output produced requires one unit of input. In addition to the unit cost of the input, k, the downstream firms face some additional quadratic costs. The total cost of production of output x_i is then

$$TC_i(x_i) = kx_i + gx_i^2,\ i = 1, 2$$

where g is a positive parameter that is related to the slope of the marginal cost curve.

Downstream firm 1 owns a percentage s_1 of the supplier. This ownership relationship does not imply any degree of control but it entitles firm 1 to a percentage s_1 of the upstream firm's profit[1].

[1] The main results of this article hold for any $s_1 \in (0, 1)$. However, since it is assumed that this participation provides no control it would be reasonable to limit $s_1 \in (0, 0.5)$.

The profit functions are given by:

$$\begin{aligned}
\Pi_U &= k(x_1 + x_2) \\
\Pi_1 &= Px_1 - TC_1(x_1) + s_1\Pi_U = (A - b(x_1 + x_2))x_1 - kx_1 - gx_1^2 + s_1k(x_1 + x_2) \\
\Pi_2 &= Px_2 - TC_2(x_2) = (A - b(x_1 + x_2))x_2 - kx_2 - gx_2^2
\end{aligned}$$

As can easily be seen, the existence of a positive s_1 has two effects. First, it lowers the effective cost of firm 1, because this firm receives part of its input cost when it receives its share of the upstream firm's profit. Effectively, firms decide outputs as if the cost functions were $TC_i = k(1 - s_i)x_i + gx_i^2$, with $s_2 = 0$[2]. Second, firm 1 is having some additional revenue for each unit that firm 2 sells. The first effect introduces some asymmetry between the two downstream firms that has been considered in the literature. Rotschild (1999) studies how firms' relative efficiencies affect the stability of a cartel, using a very similar setting: Cournot competition and linear demand, with firm-specific quadratic costs of the type $c_iq_i + \gamma_iq_i^2$. In section 3.1, the case of equal γ_i's across firms is considered, and for $n = 3$ it is concluded that "at least for some values of the parameters, the incentive to deviate is increasing as the prospective deviant is more inefficient." As Collie (2004) points out, "the explanation is that the inefficient firm has more incentive to defect from the cartel as reversion to the Cournot equilibrium outputs following a defection from the cartel has less effect on its profits than on the profits of the efficient firm. This is because the market share of the inefficient firm in the Cournot equilibrium is larger than its market share in the cartel." To compare the PVO scenario with the one where the differences in costs are natural, a parameter a is introduced and the payoff functions are rewritten as

$$\begin{aligned}
\Pi_U &= k(x_1 + x_2) \\
\Pi_1 &= (A - b(x_1 + x_2))x_1 - k(1 - s_1)x_1 - gx_1^2 + as_1kx_2 \\
\Pi_2 &= (A - b(x_1 + x_2))x_2 - kx_2 - gx_2^2
\end{aligned}$$

If $a = 0$, a positive s_1 can be interpreted as firm 1 being more efficient than firm 2 (in the sense that it has naturally lower marginal costs) when the downstream stage is solved. If $a = 1$, the cost asymmetry is recognized as arising from PVO and not from a natural cost advantage.

[2] It is straightforward to show that the sustainability of collusion does not depend on the degree of PVO or on the input cost when PVO does not lead to some firm asymmetry. Please see the appendix.

It can be shown that the objective function of downstream firm i is equivalent to

$$\pi_i = (1 - q_i - q_j - c(1 - s_i))q_i - \gamma q_i^2 + a s_i c q_j,\ i = 1, 2$$

with $\gamma = \frac{g}{b}$, $c = \frac{k}{A}$, $q_i = \frac{g}{A\gamma} x_i$, $\pi_i = \frac{g\Pi_i}{A^2\gamma}$.

From now on, we look as c to be the price of the intermediate good, q_i the quantity for each firm downstream, and π_i the profits. In this context, γ is a parameter of the increasing marginal costs. We can, however, rewrite the previous expression as

$$\pi_i = (1 - (1 + \gamma)q_i - q_j - c(1 - s_i))q_i + a s_i c q_j$$

In this case, γ is a parameter of the own price elasticity, which also measures the degree of differentiation between the products of the two firms. When providing intuitions regarding γ, we will be using the increasing marginal costs interpretation, but they can easily be seen from this second point of view.

In order to establish under which circumstances the trigger strategies defined above constitute an equilibrium we need to obtain the payoffs under three different possibilities: when both firms play the collusive output, when one of them deviates, and when both are playing the Cournot equilibrium, which corresponds to the payoff under the punishment stage of the trigger strategies. This is presented in the following section.

3 Downstream stage

3.1 Punishment

In this section, we examine the standard Cournot equilibrium for the single period game, when firms act independently. This corresponds to the punishment period payoff and therefore we denote this case with a superscript P.

Taking c as given, the downstream firms' reaction functions are:

$$\begin{aligned} q_1 &= RF_1(q_2) = \frac{1 - c(1 - s_1) - q_2}{2(\gamma + 1)} \\ q_2 &= RF_2(q_1) = \frac{1 - c - q_1}{2(\gamma + 1)} \end{aligned}$$

Lemma 1: *i) For* $c < c_m^P \equiv \frac{2\gamma+1}{2\gamma+1+s_1}$*, the Cournot equilibrium outputs are:*

$$
\begin{aligned}
q_1^P &= \frac{1+2\gamma(1-c)+c(2s_1(\gamma+1)-1)}{(2\gamma+1)(2\gamma+3)} \\
q_2^P &= \frac{1+2\gamma(1-c)-c(1+s_1)}{(2\gamma+1)(2\gamma+3)}
\end{aligned}
$$

ii) For $c > c_m^P$*, the Cournot equilibrium outputs are:*

$$
\begin{aligned}
q_1^P &= \frac{1-c(1-s_1)}{2(\gamma+1)} \\
q_2^P &= 0
\end{aligned}
$$

By computing $q_1^P - q_2^P = \frac{cs_1}{2\gamma+1} > 0$, we can see that, for $0 < c < c_m^P$ and $s_1 > 0$, firm 1 has lower perceived costs and, thus, produces more than firm 2.

The profits become

$$
\begin{aligned}
\pi_1^P &= (\gamma+1)\left(q_1^P\right)^2 + acs_1q_2^P \\
\pi_2^P &= (\gamma+1)(q_2^P)^2
\end{aligned}
$$

Since quantities are independent of a (because the output decision is simultaneous and hence firm 1 considers q_2 as a constant), the profit of firm 1 is higher when the difference in perceived costs is due to the partial ownership, i.e., when $a = 1$. All else constant, this would make firm 1 more likely to deviate from the collusive output when the asymmetry is due to PVO than when it arises from higher efficiency. Firm 2 obtains the same payoff for $a = 0$ or $a = 1$.

3.2 Collusion

In this case, downstream firms will be choosing the output level that maximizes their joint profits. A superscript C will denote this case. The objective function is:

$$
\pi^C = \pi_1 + \pi_2 = (1-q_1-q_2-c(1-s_1))q_1 + (1-q_1-q_2-c(1-as_1))q_2 - \gamma q_1^2 - \gamma q_2^2
$$

The following Lemma presents the collusive output.

Lemma 2: *The equilibrium outputs that maximize joint profits are:*

$$\begin{aligned} q_1^C &= \frac{1-c\left(1-s_1\right)}{2\left(\gamma+2\right)}+\frac{\left(1-a\right)cs_1}{2\left(\gamma+2\right)\gamma} \\ q_2^C &= \frac{1-c\left(1-s_1\right)}{2\left(\gamma+2\right)}-\left(\gamma+1\right)\frac{\left(1-a\right)cs_1}{2\left(\gamma+2\right)\gamma} \end{aligned}$$

Note that, for $a=1$, both downstream firms are perceived as equally efficient (regardless of which downstream firm is producing, part of the input cost is being paid to firm 1 and hence it is considered in the joint profit) and, thus, quantities are the same:

$$q_i^C\left(a=1\right)=q^C=\frac{1-c\left(1-s_1\right)}{2\left(\gamma+2\right)}$$

In this case, we can show that the collusive output for firm 1 is always lower than that under competition, but for firm 2 that is not always the case. The intuition is as follows. When maximizing joint profits, firms internalize the negative horizontal externality they imposed on each other. This would lead to lower outputs. Additionally, the fact that they both are perceived as equally efficient leads to a transfer of output from firm 1 to firm 2, to minimize the quadratic part of the total cost. The aggregate effect on the output of firm 1 is negative, whereas each effect may dominate the other one for firm 2. We may also have that the total quantity increases in collusion, as this second effect, of optimization of production distribution, has a positive impact on total quantity produced and may dominate for $c>\frac{1}{s_1(\gamma+1)+1}$. This effect is larger when γ is high.

If $a=0$, we would have firm 1 being more efficient than firm 2, therefore producing more when joint profits are maximized: $q_1^C>q^C>q_2^C$.

The corresponding profits are

$$\begin{aligned} \pi_1^C &= (\gamma+2)q_1^C q^C+acs_1q_2^C \\ \pi_2^C &= (\gamma+2)q_2^C q^C-\frac{a+1}{2}cs_1q_2^C \end{aligned}$$

It is not direct whether firm 1's collusive profits are higher when this firm is naturally more efficient or when the cost asymmetries are just the result of PVO. If $a=0$, firm 1 is more efficient and is producing more in the collusive outcome, $q_1^C>q^C$, but it is not receiving any profit from the input purchases by firm 2. It is easy to show that, $\pi_1^C(a=1)>\pi_1^C(a=0)$ for $\gamma>\frac{1}{2}$. Interestingly, s_1 does not affect the sign of the difference but it does affect its intensity. This means that, all else constant, the profits obtained by firm 1 under collusion are higher when the asymmetries are due to PVO only when $\gamma > \frac{1}{2}$.

It is also relevant to verify under which conditions firm 2's profits are positive, which, for $a = 1$, happens when $c < \frac{1}{1+s_1} \equiv c_m^C < c_m^P$.[3] For $a = 0$, this limit is $c < \frac{\gamma}{\gamma+s_1} \equiv c_m^{a=0}$.

3.3 Deviation

The collusive scenario is, by definition, not an equilibrium of the one shot game. We can see the optimal unilateral deviation by replacing firm j's quantity in collusion, q_j^C, in firm i's reaction function.

The following Lemma presents each firm's optimal deviation.

Lemma 3: *The optimal deviation by each firm is:*

$$\begin{aligned} q_1^D &= RF_1(q_2 = q_2^C) = \frac{2\gamma+3}{2(\gamma+1)}q^C + \frac{s_1 c(1-a)}{4(\gamma+2)\gamma} \\ q_2^D &= RF_2(q_1 = q_1^C) = \frac{2\gamma+3}{2(\gamma+1)}q^C - \frac{s_1 c}{2(\gamma+1)} - \frac{s_1 c(1-a)}{4(\gamma+2)\gamma(\gamma+1)} \end{aligned}$$

Both firms' deviation is a linear function of q^C. However, if for firm 1 deviation always happens through an increase in its quantity, for firm 2 that is only the case, for $a = 1$, if $q^C > cs_1$, or $c < \frac{1}{s_1(2\gamma+3)+1}$, which is smaller than c_m^C. The possibility that firm 2's best deviation is to lower output arises from the fact that its marginal costs are higher than as perceived by the joint profits maximizer. It may be optimal to deviate by decreasing quantities to reduce marginal costs and increase marginal revenue.

The profits, for the firm that unilaterally deviates are:

$$\begin{aligned} \pi_1^D &= (\gamma+1)(q_1^D)^2 + acs_1 q_2^C \\ \pi_2^D &= (\gamma+1)(q_2^D)^2 \end{aligned}$$

In terms of downstream profits alone, firm 1 makes a higher profit for $a = 0$, since its output is larger. However, if $a = 1$, firm 1 makes an extra revenue from the other firm's input purchases. As a result, one can either have $\pi_1^D(a = 1) > \pi_1^D(a = 0)$ or the opposite.

It is worth analyzing, in the traditional reactions functions graph, the possibilities for the three scenarios, with $a = 1$.

[3] In this section, since c is exogenous, we shall analyze only values of $c < c_m^C$. Otherwise, collusion is obviously never sustainable

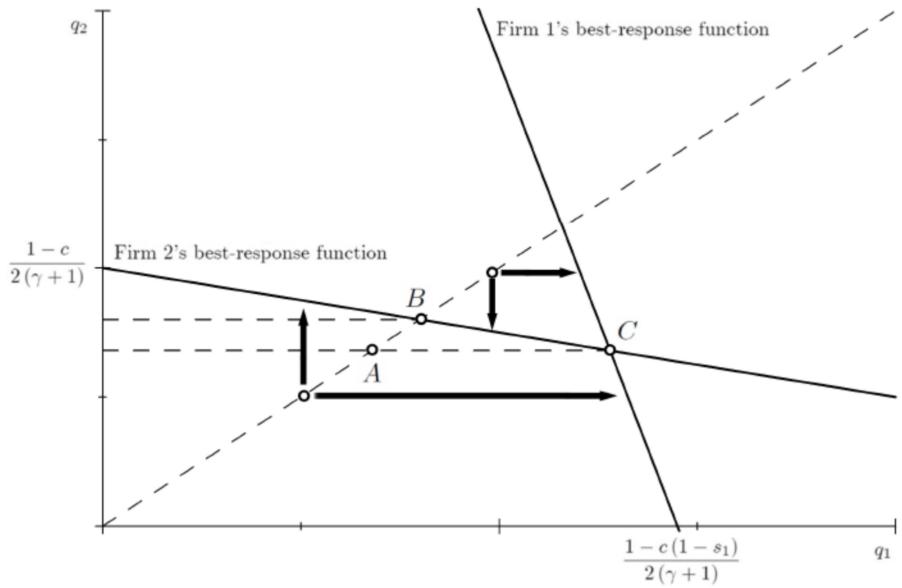

Figure 1: Best Response Functions and Optimal Deviations

In Figure 1 we can see that the Cournot equilibrium happens to the right of the 45^o line ($q_2 = q_1$), with firm 1 producing more than half of the output due to the lower perceived cost. As for the collusive outputs, they are somewhere on that line. We know that the deviation of firm 1 is to increase quantities, which is always a movement to the right, from the 45^o line to the reaction function. It can happen that such deviation leads to the equilibrium outputs. This occurs at point A, where $q_2^C = q_2^P$, with $c = \frac{2\gamma+1}{2\gamma+1+s_1(10\gamma+4\gamma^2+7)}$. Note that, at this point, firm 1's profit from collusion is lower than at the Cournot equilibrium, which, in turn, must be the same as that of deviation.

Regarding firm 2, as we saw, we can have both a movement to increase or to decrease output, which are, respectively, represented by moving upwards or downwards, from the 45^o line to firm 2's own reaction function. Notice that this implies the existence of one situation where firm 2 doesn't have incentives to deviate: $q_2^C = q_2^D$, at $c = \frac{1}{s_1(2\gamma+3)+1}$, as illustrated when the collusive output lies at point B.

3.4 Sustaining Collusion

Having set up the three previous scenarios, one can now investigate the sustainability of the collusive outcome in an infinitely repeated game downstream. Firms will be using the standard trigger strategies, which can be defined as "firm i chooses $q_i = q_i^C$ in each period if no outcome other than $\left(q_1^C, q_2^C\right)$ has happened before, and chooses $q_i = q_i^P$ otherwise". Collusion is sustained by the trigger strategies if

$$\pi_i^C \frac{1}{1-\delta} \geq \pi_i^D + \frac{\delta}{1-\delta}\pi_i^P \Leftrightarrow \pi_i^D - \pi_i^C \leq \delta\left(\pi_i^D - \pi_i^P\right)$$

If $\pi_i^D - \pi_i^P < 0$ for any firm, it is impossible to sustain collusion. Otherwise, the sustainability will depend on the discount factor $\delta_i \in (0,1)$ that firms apply to future incomes.

We can compute a critical discount factor $\delta_i^C = \frac{\pi_i^D - \pi_i^C}{\pi_i^D - \pi_i^P}$ such that, for $\delta_i \geq \delta_i^C$, $i = 1,2$ the trigger strategies that sustain the collusive scenario are an equilibrium. These conditions, $\delta_i \geq \delta_i^C$, will be later referred to as incentive compatibility constraints. Using the expression above, the critical factors are:

$$\begin{aligned} \delta_1^C &= \frac{\pi_1^D - \pi_1^C}{\pi_1^D - \pi_1^P} = \frac{(\gamma+1)(q_1^D)^2 - (\gamma+2)q_1^C q^C}{(\gamma+1)((q_1^D)^2 - (q_1^P)^2) + as_1c(q_2^C - q_2^P)} \\ \delta_2^C &= \frac{\pi_2^D - \pi_2^C}{\pi_2^D - \pi_2^P} = \frac{(\gamma+1)(q_2^D)^2 - (\gamma+2)q_2^C q^C + \frac{a+1}{2}cs_1 q_2^C}{(\gamma+1)((q_2^D)^2 - (q_2^P)^2)} \end{aligned}$$

Sustaining collusion is only possible if these critical discount factors are smaller than 1. That happens when $\pi_i^D - \pi_i^C < \pi_i^D - \pi_i^P \Leftrightarrow \pi_i^C > \pi_i^P$. This is a necessary and sufficient condition for $\delta_i^C \in (0,1)$ as, by definition, $\pi_i^D \geq \pi_i^C$. It is easy to see that this condition may not hold for firm 1. As mentioned before, it is possible to have $q_1^D = q_1^P$, which happens precisely when $q_2^C = q_2^P$, meaning that $\pi_1^D = \pi_1^P > \pi_1^C$. We, thus, check this condition for both firms. For firm 1:

$$\pi_1^C > \pi_1^P \Leftrightarrow (\gamma+2)q_1^C q^C + acs_1 q_2^C - \left((\gamma+1)\left(q_1^P\right)^2 + acs_1 q_2^P\right) > 0$$

For $a = 1$, this is a quadratic function on c, and the inequality is verified for

$$c < c_1^- \text{ or } c > c_1^+$$

It can be shown that $0 < c_1^- < c_1^+ < c_m^C$.[4] This means that for any $c \in (c_1^-, c_1^+)$ collusion is never sustainable. Outside this interval there is a critical value for the discount factor that must be compared to the effective discount factor in order to assess the sustainability of collusion. Interestingly, it should be noted that, at $\gamma = 0$, $c_1^+ = c_m^C$. This means that, if there are no efficiency gains in distributing production when colluding, there will not exist the possibility of sustaining collusion with $c > c_1^+$.

[4] The expressions for c_1^- and c_1^+, as well as many other below, are provided in the mathematical appendix

For firm 2, the condition becomes

$$\pi_2^C > \pi_2^P \Leftrightarrow (\gamma+2)\left(q^C\right)^2 - cs_1 q^C - (\gamma+1)\left(q_2^P\right) > 0$$

which, for $a = 1$, is satisfied for:

$$c < c_2^-$$

We will have that $c_2^- < c_m^C$. It is also possible to show that $c_1^+ < c_2^-$.

Lemma 4 *For $c \in (0, c_1^-) \cup (c_1^+, c_2^-)$, both firms' critical discount factors belong to the interval $(0,1)$, meaning that it might be possible to sustain collusion, if the actual values of δ_i are sufficiently high.*

Studying the behavior of the critical discount factors, when they are in the interval $(0,1)$, it is possible to show that δ_1^C is always increasing in c to the left of c_1^- and decreasing to the right of c_1^+. It can be shown that δ_2^C is decreasing for $c < \frac{1}{s_1(2\gamma+3)+1} \equiv c_2^{\min}$, and increasing otherwise. Note that at $c = c_2^{\min}$, $\delta_2^C = 0$. This is the case where firm 2 has no incentive to deviate: $q_2^C = q_2^D$ and, thus, collusion is always sustainable as far as firm 2 is concerned. To illustrate the results, one particular set of the critical factor functions and c_m^P are illustrated in Figure 2.

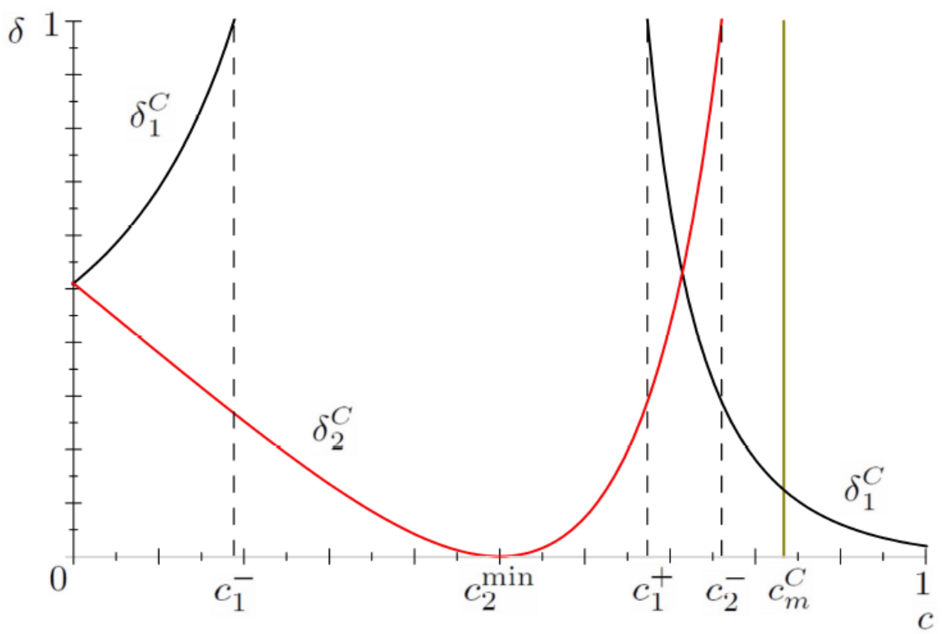

Figure 2: Critical discount factors as functions of c. $\gamma = 1$ and $s_1 = 0.2$

It is now relevant to compare this scenario with the one for $a = 0$. In that scenario, $\delta_1^C \in (0, 1)$ for any value of c belonging to

$$c < c_1^{a=0}$$

Since it can be shown that $c_1^{a=0} > c_m^{a=0}$, it can be concluded that the critical discount factor for firm 1 under the natural asymmetry scenario always falls in the interval[5].

For firm 2, the restriction will be such that there exists a non-empty interval if

$$c < c_2^{a=0}$$

Since $c_2^{a=0} < c_m^{a=0}$, it is interesting to note one immediate difference between these two origins of asymmetry. For small values of c, the firm that would more likely have incentives to deviate under PVO would be firm 1, the most efficient. Under natural cost differences, however, that firm would be firm 2, which is the one with higher costs.

Figure 3 illustrates these critical factor functions, for the same values of the parameters as before

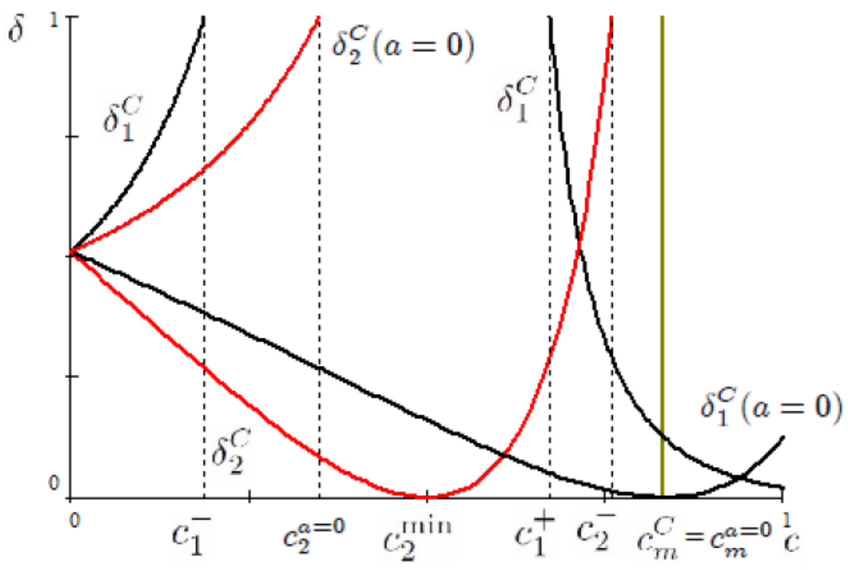

Figure 3: Critical discount factors for both values of a

[5]For the sake of comparison between the cases of $a = 0$ and $a = 1$, it will be assumed that $\gamma \geq 1$, which is when $c_m^{a=0} \geq c_m^C$. This allows us to keep considering c in the same interval. There is no loss of generality since $c_2^{a=0} < c_m^{a=0}$, which means that collusion would not be sustainable for c close to $c_m^{a=0}$.

Note that, in this case, $c_m^{a=0}$ is located between c_1^- and $c_2^{\min}$, but as stated before, that does not have to be the case, for different values of the parameters.

To be able to take conclusions on what are the different impacts that the two types of asymmetry have on the critical discount factor, we need to realize that there are two factors that change firm 1's incentives to sustain collusion when the ownership and the natural efficiency differences scenarios are compared. The first one happens in the retail profit and the second one takes place at the wholesale level. For this purpose it is convenient to separate firm 1's profit in two terms: the retail profit, π_{1r}^j, and the wholesale profits from firm 2's input expenditures, π_{1w}^j. So, $\pi_1^j = \pi_{1r}^j + \pi_{1w}^j$ with $j = C, P, D$. Firm 1's critical discount factor is then

$$\delta_1^C = \frac{\pi_{1r}^D + \pi_{1w}^D - \pi_{1r}^C - \pi_{1w}^C}{\pi_{1r}^D + \pi_{1w}^D - \pi_{1r}^P - \pi_{1w}^P} = \frac{\pi_{1r}^D - \pi_{1r}^C}{\pi_{1r}^D - \pi_{1r}^P + \pi_{1w}^D - \pi_{1w}^P}$$

The competitive equilibrium leads to the same quantities and retail profit, π_{1r}^P, regardless of $a = 0$ or $a = 1$. In both cases firm 1 has a higher market share than firm 2. Under collusion, however, both firms produce the same output under PVO but not if the cost differences were natural. This leads to lower retail profits under collusion for firm 1 in the PVO case, $\pi_{1r}^C(a = 1) < \pi_{1r}^C(a = 0)$. Even though the profits from deviating are also lower, $\pi_{1r}^D(a = 1) < \pi_{1r}^D(a = 0)$, the relative distances, $\frac{\pi_{1r}^D - \pi_{1r}^C}{\pi_{1r}^D - \pi_{1r}^P}$ will increase. From this effect, firm 1 would have a higher incentive to deviate under the partial ownership scenario.

The second effect happens at the wholesale profit and results from the fact that firm 1 receives part of the costs paid by firm 2. This one only exists under PVO. This revenue is given by

$$\pi_{1w}^D - \pi_{1w}^P = as_1c(q_2^C - q_2^P)$$

Thus, if firm 2 produces more in the collusive scenario in the presence of PVO, firm 1 might be better off under collusion when $a = 1$. Note that the sign of this effect depends on the sign of $q_2^C - q_2^P$. Thus, this effect might strengthen the previous one, making collusion more difficult to sustain under PVO if $q_2^C < q_2^P$. However, it might also have the opposite effect, for $q_2^C > q_2^P$, and eventually be strong enough to bring the critical discount factor back to possible values to sustain collusion, which happens for high values of c and s_1. This is the origin of c_1^+, which doesn't exist for the $a = 0$ case. Inspection of the function $\delta_1^C(a = 1)$ reveals that it has two vertical asymptotes. The first happens in the already mentioned scenario of $q_1^D = q_1^P$. The second one appears, when this "new" effect is strong enough to compensate for the loss in the retailer side of the profits, which creates another possibility of $\pi_1^D = \pi_1^P$. For the admissible values of c, it can be shown that $\delta_1^C(a = 0)$ is decreasing in c, having a minimum at $c = c_m^{a=0}$, where $\delta_1^C(a=0)=0$ as illustrated in Figure 3. Remembering that $\delta_1^C(a=1)$ is increasing in c for $c<c_1^-$, it will follow that, at least until some value of c, firm 1 would have higher incentives to deviate under PVO than under natural asymmetry.[6]

For firm 2 the only differences are through the retail profit. In general, it is easier to sustain collusion under PVO than under natural asymmetry. This happens because the distribution of production in the collusive scenario under PVO is such that $q_1^C=q_2^C$, which doesn't happen under natural asymmetry, where firm 2 would be producing less than half of the total output. It can be shown that $\delta_2^C(a=0)$ is increasing for $c\in(0,c_2^{a=0})$, while, as stated before, $\delta_2^C(a=1)$ is initially decreasing. This proves that, for at least some values of c, natural asymmetry is more destructive to collusion, through firm 2, than PVO. It might happen, for very high values of γ and low values of s_1, that $c_2^{a=0}>c_2^-$, meaning that such is not always the case. This possibility exists because, for such values: (1) there isn't much asymmetry between firms and, thus, colluding as in PVO wouldn't be that much better than doing it as in natural asymmetry and (2) firm 2 is not forced to sell as much above marginal costs under natural asymmetry than under PVO.

In conclusion, it is not clear whether it is PVO or natural asymmetry that makes collusion more difficult to sustain. It can be seen, however, that under PVO firm 1 would be the one more likely to deviate, while for the natural asymmetry case, that would be firm 2. Having understood this, we will now only use $a=1$.

4 Upstream decisions

So far, we have considered an exogenously defined c. At this point we are ready to look at alternative ways of choosing this price. We will study two scenarios. The first corresponds to an upstream monopolist that is able to set c to maximize its profit. The second is the case of a regulated upstream firm, with a regulator setting c to maximize welfare. For this section it will be assumed that, if the trigger strategies are an equilibrium of the infinitely repeated game, that will be the realized equilibrium.

4.1 Monopolist Wholesaler

The upstream firm will choose c to maximize its own profits. It is indifferent whether we consider that it maximizes all its profits, π_W, or just the part that remains after the subtraction

[6] It can easily be shown that a sufficient condition for this to happen for any $c\in(0,c_m^C)$ is $\gamma>1$.

of downstream's ownership, $(1-s_1)\pi_W$, as this is a monotonic increasing transformation of the objective function. By choosing c, besides affecting the demand by the usual effects, firm U might also influence whether there will be or not downstream collusion. For some values of c, collusion will be sustainable and for others it will not. Thus, the upstream monopolist may choose a different value of c to induce collusion or to break it up. We start by presenting the optimal c if there is always downstream collusion or if there is always competition between the downstream firms.

Under collusion, the wholesaler's per period payoff is:

$$\pi_U^C = c * 2q^C = c\left(\frac{1-c(1-s_1)}{(\gamma+2)}\right) = c\left(\frac{1}{\gamma+2} - \frac{1-s_1}{\gamma+2}c\right)$$

Lemma 5 *If downstream collusion is sustainable, the optimal* c *chosen by firm* U *is*

$$c^C = \frac{1}{2(1-s_1)}$$

Which will lead to profits of

$$\pi_U^C = \frac{1}{4(\gamma+2)(1-s_1)}$$

Under competition, the per period profit function becomes:

$$\pi_U^P = c * (q_1^P + q_2^P) = c(\frac{cs_1 + 2(1-c)}{2\gamma+3}) = c(\frac{2}{2\gamma+3} - \frac{2-s_1}{2\gamma+3}c)$$

Lemma 6 *If downstream collusion is not sustainable, the optimal* c *chosen by firm* U *is*

$$c^P = \frac{1}{2-s_1}$$

Leading to profits of

$$\pi_U^P = \frac{1}{(2-s_1)(2\gamma+3)}$$

There could be the possibility that firm U could prefer to foreclose firm 2, by setting $c \geq c_m^P$. It can be shown that such case will never occur.

It can easily be seen that $\pi_U^P > \pi_U^C$ for $s_1 < \frac{2}{2\gamma+5} \equiv s_1^F$. To understand this limit, one should realize that the type of downstream outcome (competition or collusion) affects the upstream firm through its demand function. At the outset, it is unclear which outcome leads to a higher demand for the upstream firm. To understand this, it is useful to think of collusion (joint profit maximization) as equivalent to a horizontal merger. On the one hand, collusion leads to the internalization of the horizontal price externality that downstream firms impose on one another, which reduces total output and, consequently, the demand for the intermediate product. On the other hand, collusion leads to lower marginal costs, which raises the directed demand for the upstream firm. There are two reasons for this marginal cost reduction. First, the efficient use of two "plants" and the possibility of reshuffling output between them allows for a lower marginal cost provided that $\gamma > 0$ (due to the quadratic term in the cost function). Second, from a joint profit maximization perspective, it is irrelevant which downstream firm purchases the input from the upstream firm: each unit purchased results in an additional revenue of cs_1 for the downstream firms. Therefore, production by firm 2 now also occurs with a lower perceived marginal cost. If s_1 is large enough, the demand faced by the upstream firm may be larger under collusion than under competition, because of the last reason mentioned above. On top of this, the demand is also less elastic: because both firms act as if they are paying part of the marginal cost to themselves, they will be less sensitive to price. As a result, a high s_1 may make collusion more profitable than competition for the upstream firm. We can also see that s_1^F is negatively dependent on γ, coming from the fact that the more efficiency gains there are in collusion, the more the upstream firm benefits from it. One can see that there might not exist an area where collusion is sustainable at $c = c^C$ as $c^C < c_m^C$ only if $s_1 < \frac{1}{3}$. It can also be shown that there are values of the parameters s_1 and γ that allow for each c^C and c^P to fall in all sides of our "non-empty interval" for the critical discount factor. It is simple to see that for very small values of s_1, both c^C and c^P are close to $\frac{1}{2}$, while we could have c_1^- bigger than that. It is also easy to notice that for some s_1 that makes c^P close to c_m^C, a positive γ is enough to create the other extreme possibility ($c_2^- < c^P$).

We now consider the possibility that the upstream firm can choose some c different from the optimal for the unconstrained scenarios if it prefers to change the existence of collusion. We can split this analysis into two scenarios:

The first happens for $s_1 < s_1^F$, which means that unconstrained competition is better than unconstrained collusion.[7] Thus, the problem only arises if collusion is sustainable for $c = c^P$. In this case firm U might want to break collusion, which implies to make either of the incentive compatibility constraints bind, which means, either $\delta_1 = \delta_1^C(c, s_1, \gamma)$ or $\delta_2 = \delta_2^C(c, s_1, \gamma)$. As we saw in the previous section, this scenario is only possible in two areas, $c^P < c_1^-$ and $c^P > c_1^+$, and with δ_i sufficiently high. In the first area, it must be that c must be increased, to make firm 1's constraint binding, while for the second area it would be the opposite. Collusion may

[7] By "unconstrained collusion", we refer to the case where collusion is sustainable at $c = c^C$, whereas "constrained collusion" means that a different c must be set for collusion to be sustainable. Likewise for the case of competition.

also be made unsustainable through firm 2. From the shape of δ_2^C described in the previous section, it is possible to understand that the cheapest way to do it is by the opposite movement than the one necessary for firm 1. This means that, for $c^P > c_1^+$, firm U can decrease c to give incentives for firm 1 to deviate, or increase it so that it is firm 2 the one not willing to stay in collusion. For $c^P < c_1^+$ it would only be optimal to bind firm 2's constraint if its discount factor is sufficiently lower than that of firm 1 (for a common value of $\delta = \delta_1 = \delta_2$ it would never compensate to change c as to do this, as it would always be cheaper to break collusion through firm 1).

Having found the cheapest way to break collusion downstream, firm U will compare it with the collusive scenario. It might not be possible to achieve the unconstrained collusion case, since $c^C > c^P$, in which case it would be needed to find the best constrained collusion, to be used for comparison. Finding constrained collusion follows a similar method. The main difference is that instead of breaking/violating one constraint, firm U will have to ensure that both firms want to collude. To illustrate this case, one numerical example is provided:

Let $\gamma = 1$, $s_1 = 0.03$ and $\delta = \delta_1 = \delta_2 = 0.9$. In this scenario, firm U would prefer unconstrained competition over unconstrained collusion, but the first one cannot be achieved, as $\delta > \delta_1^C(c = c^P) > \delta_2^C(c = c^P)$. It, thus, faces two options. It can choose a c to the right of of c^P such that the $\delta = \delta_1^C$ (in this case, it is clear that breaking through firm 2 would be more expensive), or settle for unconstrained collusion, with $c = c^C$ and the downstream firms colluding. The first option, which is achieved by setting $c = 0.57$ yields a profit of $\pi_U = 0.1$, while choosing $c = c^C = 0.515$ leads to profits of only $\pi_U^C = 0.086$. Under these conditions, firm U is better off by choosing an non-optimal c to force competition, than accepting downstream collusion, even if that would be unconstrained. Figure 4 illustrates this.

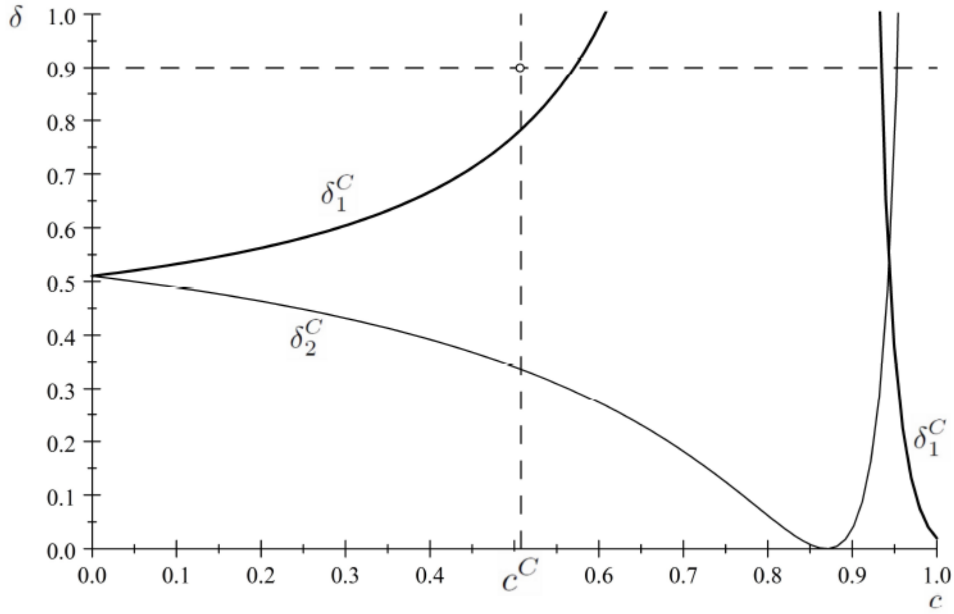

Figure 4: Example of wholesaler incentives to c when unconstrained competition would be preferable

The second scenario is for $s_1 > s_1^F$. Since, under these circumstances, firm U prefers unconstrained downstream collusion, there is only a decision to be analyzed if the trigger strategies are not an equilibrium for $c = c^C$. The choices are similar to the previous scenario. Firm U might choose c such that both firms have their constraints satisfied. It then compares the best way to implement constrained collusion, with a competitive scenario. This scenario will be, if possible, the unconstrained competition one. If not possible, firm U has the best constrained competitive case to compare. Thus, we find a very interesting result: that, if there are enough efficiency gains, it might be better for an upstream firm to facilitate collusion downstream, even if unconstrained competition is an available option. To illustrate, another numerical example is provided.

Let $s_1 = 0.1$ and $\gamma = 5$. For a common value of the critical factor $\delta = \delta_1 = \delta_2$, we can define a function that characterizes the optimal $c(\delta)$. For $\delta > 0.825$ unconstrained collusion is feasible, while for a lower value firm 2 would want to deviate. We can find that, for $0.505 < \delta < 0.825$ it is optimal for firm U to bind the restriction of firm 2 and have constrained collusion. For lower values of the discount factor, firm U would settle for unconstrained competition. For these values of the parameters, firm 1's constraint will never be binding. Figure 5 presents this example.

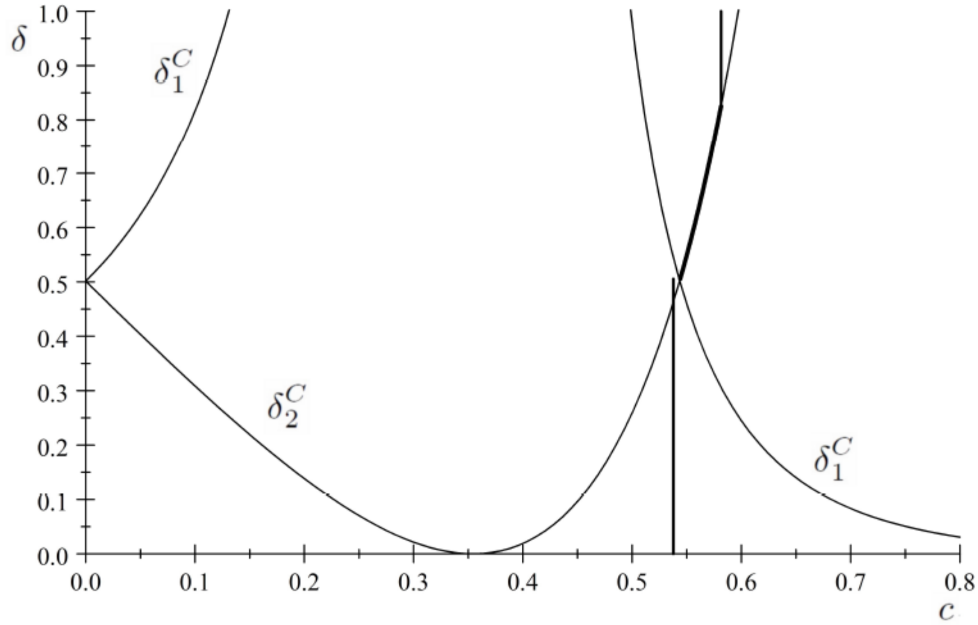

Figure 5: Example of best c for any common value of δ if unconstrained collusion would be optimal

4.2 Welfare

A different setup is now considered, where the upstream firm is regulated. The regulator's goal is to maximize total welfare, subject to $c \geq 0$, i.e., firm U cannot sell below marginal costs. Consumer gross surplus is

$$\int_0^Q (1-x)dx = Q - \frac{1}{2}Q^2 = (q_1 + q_2) - \frac{(q_1+q_2)^2}{2}$$

Total welfare is given by

$$W = (q_1 + q_2) - \frac{(q_1+q_2)^2}{2} - \gamma(q_1^2 + q_2^2)$$

It is easy to show that, regardless of the collusive scenario, the optimal c should be negative, to compensate for the market power that downstream firms possess.

Lemma 7 *Assume setting $c < 0$ is impossible. If the collusive scenario is taken as given,* *the value that maximizes welfare is $c = 0$, for both scenarios*

Under competition total welfare is

$$W_0^P = 2\frac{\gamma + 2}{(2\gamma + 3)^2}$$

Under downstream collusion, total welfare is given by

$$W_0^C = \frac{1}{2}\frac{\gamma + 3}{(\gamma + 2)^2}$$

As expected, $W^P > W^C$. The question that is now imposed, is whether it will be optimal for the regulator to choose $c > 0$, increasing the distance between the price and the marginal cost, in order to break the agreement between the downstream firms. This is, obviously, only an issue if the incentive compatibility constraints are verified, meaning $\delta_1 > \delta_1^C(c = 0)$ and $\delta_2 > \delta_2^C(c = 0)$. Remember that $\delta_1^C(c = 0) = \delta_2^C(c = 0) = \frac{(2\gamma+3)^2}{24\gamma+8\gamma^2+17}$ since with $c = 0$ there is no asymmetry between firms, with the critical values being the same for both firms and independent of s_1.

As seen in the previous sections, there are many ways available to make collusion unsustainable. It is easy to see, however, that the one with the least impact on welfare is by increasing c enough to give firm 1 incentives to deviate (assuming $\delta_1 = \delta_2$). The regulator only needs to compare the welfare of unconstrained collusion at $c = 0$, with the welfare of competition, by choosing $c = c^*$ such that $\delta_1 = \delta_1^C(c^*)$. It can be shown that it is possible to have $W^P(c = c_1^-) > W_0^C$, meaning that, in such circumstances, it would be optimal for firm U to to break downstream collusion, regardless of the value of $\dot{\delta}_1$. There are, however, for $\delta_1 < 1$, less inefficient ways of reaching such results. By finding the value c_x such that $W^P(c = c_x) = W_0^C$, it will compensate for firm U to set $c = c^*$ if and only if $c^* < c_x$. With this, it follows that:

Proposition 1 *For $\delta_1 \leq \delta_1^C(c_x)$ the regulator is willing to set $c = c^* > 0$ as to break collusion and improve social welfare*

Thus, the function $\delta_1^C(c_x(s_1, \gamma), s_1, \gamma)$ specifies the maximum value that δ_1 can take, for which the regulator is willing to set $c = c^*$. The behavior of this function is such that

1. $\delta_1^X(\gamma, 0) = \delta_1^C(c = 0)$. Since, for $s_1 = 0$, changing the value of c does not change the incentives to collude, the regulator has no tool to break it.

2. It is increasing in s_1. The higher the asymmetry between firms, the easier it is to break collusion.

3. Decreasing in γ.[8] Here two effects exist. For a higher γ, collusion is easier to break. However, it also increases the inefficiency of competition with positive c, as the distribution of production will not be optimal. The second effect will dominate.

This function is represented in Figure 6.

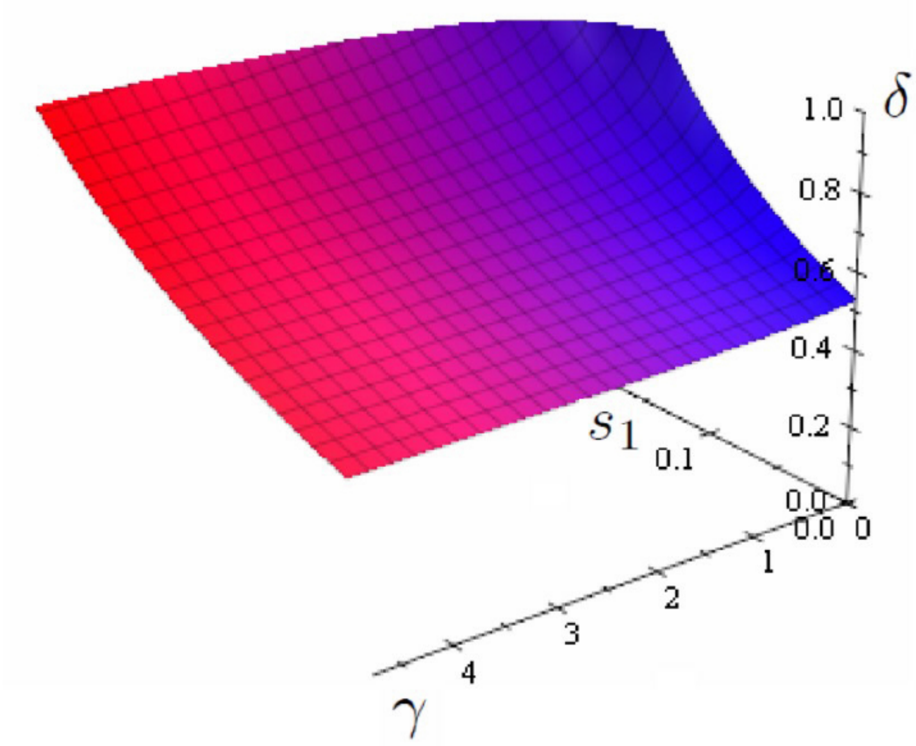

Figure 6: Maximum value of δ for which the regulator is willing to set $c > 0$

If δ_1 is located above the represented area, the regulator prefers to settle for collusion, while if it is below the area, the better option is to choose $c = c^*$ and break it.

5 Conclusion

Partial vertical ownership is common practice in modern markets. In the literature, it has been seen as a way to Pareto improve welfare in a market, as it decreases double marginalization, improving both consumers' surplus and producers' profit. Greenlee and Raskovich (2006), for example, find that for homogenous goods such will be the case, even though the biggest effect is on producer's profits and not consumers' surplus. They do argue, however, that for differentiated goods total welfare might actually be reduced. In this paper, a framework for

[8]This could not be proven formally. All numeric simulations and graphical analysis confirm it

the analysis of the incentives for downstream collusion under PVO is provided. The first result is that PVO creates asymmetry between firms, as the ones that have a participation in the supplier receive part of their costs back. This asymmetry makes collusion more unlikely to be sustained. This result is similar to that obtained for the case of natural differences between firm's costs structures, but it changes which firm has the most incentives to deviate. Under PVO, this will be the most efficient firm, while under natural cost asymmetry, it would be the least efficient one. A more efficient firm has a higher incentive to deviate from a collusive scenario since its efficiency is not totally reflected when firms are maximizing joint profits. There is, however, another major difference between natural asymmetry and the one caused by PVO. If there are total efficiency gains from colluding, a firm that owns a part of the upstream firm may indirectly benefit in this scenario from such gains, thus creating a new possibility for sustainable collusion.

The existence of this ownership creates a new tool for the upstream firm. Besides impacting its profits by the normal effects on demand, the price that this firm sets may also change whether downstream firms will collude or not. This tool can be used to break downstream collusion as that scenario provides, in general, a lower demand for the intermediate product that the upstream firm is selling. The tool, however, may also be used to create such joint profit maximization, which might be the optimal choice if there are sufficient efficiency gains from the distribution of production between downstream firms.

Finally, this tool may also be beneficial in the case of a regulated market, such as, typically, energy grids or telecommunications. A regulator may use the price of the intermediate good differently from what is usually considered the best option in a static environment (the marginal cost). It may be better to have a sufficiently high price of such good that avoids collusion, even if it increases the double marginalization problem.

This analysis abstracted from some important issues of PVO. The first is the inexistence of control. In many cases, this integration structure may provide partial or total control to the downstream firm. A second is to make the participation share endogenous. In this framework, the ownership was not chosen by the downstream firms. New issues would be raised, were it an endogenous decision, such as whether it would compensate to change the participation just to make collusion sustainable. Finally, differentiated prices for the intermediate product were ignored. This would provide a more complex tool form the upstream firm to change the collusive scenario.

References

[1] Bresnahan, Timothy, and Steven Salop (1986), "Quantifying the Competitive Effects of Production Joint Ventures", International Journal of Industrial Organization 4, 155–175.

[2] Chen, Z. and T. W. Ross, 2003, "Cooperating upstream while competing downstream: a theory of input joint ventures", International Journal of Industrial Organization, 21(3), 381-397.

[3] Flath, David (1992), "Horizontal Shareholding Interlocks", Managerial and Decision Economics 13, 75–77.

[4] Foros, O., H.J. Kind and G. Shaffer (2010), "Mergers and Partial Ownership", The Bradley Policy Research Center, Financial Research and Policy Working Paper No. FR 10-11.

[5] Gilo, D., Y. Moshe and Y. Spiegel (2006), "Partial Cross Ownership and Tacit Collusion", RAND Journal of Economics 37, 81–99.

[6] Greenlee, P. and A. Raskovich, 2006, "Partial vertical ownership", European Economic Review, 50(4), 1017-1041

[7] Malueg, D. (1992), "Collusive Behavior and Partial Ownership of Rivals", International Journal of Industrial Organization 10, 27–34.

[8] O'Brien, Daniel P. and Salop, Steven C. (2000), "Competitive Effects of Partial Ownership: Financial Interest and Corporate Control", Antitrust Law Journal 67, 559–614.

[9] Reitman, D. (1994), "Partial Ownership Arrangements and the Potential for Collusion", The Journal of Industrial Economics 42, 313–322.

[10] Reynolds, Robert J, and Bruce R. Snapp (1986), "The Competitive Effects of Partial Equity Interests and Joint Ventures", International Journal of Industrial Organization 4, 141–153.

[11] Rossini, G. and Vergari, C. (2011), "Input Production Joint Venture", The B.E. Journal of Theoretical Economics 11(1), Article 5.

[12] Compte, O. et al. (2002), "Capacity constraints, Mergers and Collusion", European Economic Review, 46, 1-29

[13] Collie, D. (2004), "Collusion and the elasticity of demand", Economics Bulletin, 12, 1-6

6 Appendix

The following is a list of the expressions that were too long to be presented in the article

$$c_1^- = \frac{(2\gamma+1)\left(2\gamma+10s_1+18\gamma s_1+8\gamma^2 s_1+1\right) - s_1(2\gamma+3)(2\gamma+1)\sqrt{9+16\gamma+8\gamma^2}}{(2\gamma+1)^2+s_1\left(76\gamma+19s_1+108\gamma s_1+88\gamma^2+32\gamma^3+184\gamma^2 s_1+128\gamma^3 s_1+32\gamma^4 s_1+20\right)}$$

$$c_1^+ = \frac{(2\gamma+1)\left(2\gamma+10s_1+18\gamma s_1+8\gamma^2 s_1+1\right) + s_1(2\gamma+3)(2\gamma+1)\sqrt{9+16\gamma+8\gamma^2}}{(2\gamma+1)^2+s_1\left(76\gamma+19s_1+108\gamma s_1+88\gamma^2+32\gamma^3+184\gamma^2 s_1+128\gamma^3 s_1+32\gamma^4 s_1+20\right)}$$

$$c_1^{a=0} = \frac{(2\gamma+1)\left(2\gamma-9s_1-16\gamma s_1+4\gamma^2-8\gamma^2 s_1\right) - (2\gamma+3)(2\gamma+1)s_1\sqrt{16\gamma+8\gamma^2+9}}{2\left(\gamma-9s_1-34\gamma s_1+4\gamma^2+4\gamma^3+9s_1^2+25\gamma s_1^2-40\gamma^2 s_1-16\gamma^3 s_1+24\gamma^2 s_1^2+8\gamma^3 s_1^2\right)}$$

$$c_2^- = \frac{(2\gamma+1)\left(8s_1-2\gamma+12\gamma s_1+4\gamma^2 s_1-1\right) + s_1(2\gamma+3)(2\gamma+1)\sqrt{(16\gamma+8\gamma^2+9)}}{16s_1+56\gamma s_1+17s_1^2+60\gamma s_1^2+56\gamma^2 s_1+16\gamma^3 s_1+92\gamma^2 s_1^2+64\gamma^3 s_1^2+16\gamma^4 s_1^2-(2\gamma+1)^2}$$

$$c_2^{a=0} = \frac{(2\gamma+1)\left(2\gamma+9s_1+14\gamma s_1+4\gamma^2+4\gamma^2 s_1\right) + s_1(2\gamma+3)(2\gamma+1)\sqrt{(16\gamma+8\gamma^2+9)}}{2\gamma+18s_1+64\gamma s_1+8\gamma^2+8\gamma^3-16\gamma s_1^2+64\gamma^2 s_1+16\gamma^3 s_1-24\gamma^2 s_1^2-8\gamma^3 s_1^2}$$

$$c_x = \frac{(2\gamma+3)(2\gamma+1)\sqrt{4(2\gamma+1)^2(1-s_1)+s_1^2\left(9\gamma+7\gamma^2+1\right)} - (\gamma+2)(2\gamma+1)^2(2-s_1)}{(\gamma+2)\left(4(\gamma+1)(2\gamma+1)^2(1-s_1)+s_1^2+14\gamma s_1^2+4s_1^2\gamma^2(2\gamma+5)\right)}$$

Proofs

Here are presented the formal proofs of the statements in the article. If any are left out, it's because they were trivial.

For most of the following proofs, computer software[9] was used for simplification of expressions as well as to solve some equations and inequations. For this reason, many steps of simplification are left out. Also, replicating many of them would be nearly impossible without similar computational aid. Each time a new proof is being presented, it is denoted by *(P)*

Proofs of Section 2

(P) Proving that PVO only affects the sustainability of collusion if it creates asymmetry. This proof uses notation that is presented during section 2, but hasn't been introduced before footnote 2. It is advisable to finish reading the section before analyzing this proof.

If one introduces $s_2 = s_1 = s$ and uses the critical discount factor method explained in section 3, it will follow:

Under competition:

$$q_i^P = \frac{1-c(1-s)}{2\gamma+3}$$
$$\pi_i^P = \frac{(1-c+cs)(1-c+4cs+\gamma(3cs+1-c))}{(2\gamma+3)^2}$$

Under collusion:

$$q_i^C = \frac{1-c+2cs}{2\gamma+4}$$
$$\pi_i^C = \frac{(2cs+1-c)^2}{4(\gamma+2)}$$

Under deviation by firm i:

[9] The program was Scientific Workplace v5.0

$$q_i^D = \frac{2cs + 3(1-c) + 2\gamma(cs + 1 - c)}{12\gamma + 4\gamma^2 + 8}$$

$$\pi_i^D = \frac{\gamma^2 4\left(cs(-4c + 5cs + 4) + (c-1)^2\right) + \gamma 4\left(cs(-11c + 14cs + 11) + 3(c-1)^2\right)}{16(\gamma+1)(\gamma+2)^2} + \frac{+9(c-1)^2 + 4cs(-7c + 9cs + 7)}{16(\gamma+1)(\gamma+2)^2}$$

The critical discount factor becomes:

$$\delta_i^C = \frac{\pi_i^D - \pi_i^C}{\pi_i^D - \pi_i^P} = \frac{(2\gamma+3)^2}{24\gamma + 8\gamma^2 + 17}$$

which does not depend on s or c.

Proofs of Section 3

(P) The proofs of Lemmas 1 to 3 are trivial.

Lemma 1: The reaction functions are just a maximization of the objective function, taking the other firm's quantity as given. The equilibrium quantities come from the intersection of the reaction functions.

Lemma 2: Simply maximize the joint profits with respect to q_1 and q_2

Lemma 3: Replace q_j^C in firm i reaction function.

All the second order conditions are also trivially satisfied. To get the expressions of the profits only a rearrangement of terms is required.

(P) Proof that $\pi_1^C(a=1) > \pi_1^C(a=0)$ for $\gamma > \frac{1}{2}$

Computing the difference

$$\pi_1^C(a=1) - \pi_1^C(a=0)$$

which is

$$(\gamma+2)\left(\frac{cs_1}{2(\gamma+2)\gamma} + \frac{1-c(1-s_1)}{2(\gamma+2)}\right)\left(\frac{1-c(1-s_1)}{2(\gamma+2)}\right) - \\ -(\gamma+2)\left(\frac{1-c(1-s_1)}{2(\gamma+2)}\right)^2 - cs_1\left(\frac{1-c(1-s_1)}{2(\gamma+2)}\right)$$

is equivalent to

$$\frac{(1-c(1-s_1))(1-2\gamma)cs_1}{4(\gamma+2)\gamma}$$

Which is positive iff $\gamma > \frac{1}{2}$

(P) Profits of firm 2 positive under collusion. $\pi_2^C > 0$

For $a = 1$ is verified for

$$(\gamma+2)\left(\frac{1-c(1-s_1)}{2(\gamma+2)}\right)^2 - cs_1\frac{1-c(1-s_1)}{2(\gamma+2)} > 0$$

which happens if

$$c < \frac{1}{1+s_1}$$

For $a = 0$ is verified for

$$\frac{(\gamma-c(\gamma+s_1))(1-c)}{4(\gamma+2)\gamma} > 0$$

which happens if

$$c < \frac{\gamma}{\gamma+s_1}$$

(P) When is firm 2 increasing or decreasing in deviation for $a = 1$, i.e., $q_2^D > q^C$

Computing the difference

$$q_2^D - q^C > 0$$

which is

$$\frac{2\gamma+3}{2(\gamma+1)}q^C - \frac{s_1c}{2(\gamma+1)} - q^C > 0$$

is verified for

$$q^C > s_1c$$

Replacing $q^C = \frac{1-c(1-s_1)}{2(\gamma+2)}$ and solving for c

$$c < \frac{1}{s_1(2\gamma+3)+1}$$

(P) When is $\pi_1^D(a=1) > \pi_1^D(a=0)$

Computing the difference

$$\pi_1^D(a=1) - \pi_1^D(a=0) > 0$$

which is

$$(\gamma+1)(\frac{2\gamma+3}{2(\gamma+1)}\frac{1-c(1-s_1)}{2(\gamma+2)})^2 + s_1 c\left(\frac{1-c(1-s_1)}{2(\gamma+2)}\right) - \\ -(\gamma+1)(\frac{2\gamma+3}{2(\gamma+1)}\frac{1-c(1-s_1)}{2(\gamma+2)} + \frac{s_1 c}{4(\gamma+2)\gamma})^2 > 0$$

equivalent to

$$c\left((1-2\gamma)\left(8\gamma+4\gamma^2+1\right)(1-s_1) - (\gamma s_1+1)\right) + 2\gamma\left(6\gamma+4\gamma^2-3\right) > 0$$

It may or not be verified, depending on the value of the parameters.

(P) Intersection between q_2^P and q^C. $q_2^P = q^C$

$$\frac{1+2\gamma(1-c) - c(1+s_1)}{(2\gamma+1)(2\gamma+3)} = \frac{1-c(1-s_1)}{2(\gamma+2)}$$

verified for

$$c = \frac{2\gamma+1}{2\gamma+1+s_1(10\gamma+4\gamma^2+7)}$$

(P) Finding the conditions for a non empty interval of the discount factor.

Firm 1: $\delta_1^C \in (0,1)$

As stated, a necessary and sufficient condition is $\pi_1^D \geq \pi_1^C$. For $a=1$, the resulting expression is

$$c^2\left(4\gamma+20s_1+76\gamma s_1+4\gamma^2+19s_1^2+108\gamma s_1^2+88\gamma^2 s_1+32\gamma^3 s_1+184\gamma^2 s_1^2+128\gamma^3 s_1^2+32\gamma^4 s_1^2+1\right) \\ -2c(2\gamma+1)\left(2\gamma+10s_1+18\gamma s_1+8\gamma^2 s_1+1\right) + (2\gamma+1)^2 > 0$$

The roots are

$$c_1^- = \frac{(2\gamma+1)\left(2\gamma+10s_1+18\gamma s_1+8\gamma^2 s_1+1\right)-s_1(2\gamma+3)(2\gamma+1)\sqrt{9+16\gamma+8\gamma^2}}{(2\gamma+1)^2+s_1\left(76\gamma+19s_1+108\gamma s_1+88\gamma^2+32\gamma^3+184\gamma^2 s_1+128\gamma^3 s_1+32\gamma^4 s_1+20\right)}$$

$$c_1^+ = \frac{(2\gamma+1)\left(2\gamma+10s_1+18\gamma s_1+8\gamma^2 s_1+1\right)+s_1(2\gamma+3)(2\gamma+1)\sqrt{9+16\gamma+8\gamma^2}}{(2\gamma+1)^2+s_1\left(76\gamma+19s_1+108\gamma s_1+88\gamma^2+32\gamma^3+184\gamma^2 s_1+128\gamma^3 s_1+32\gamma^4 s_1+20\right)}$$

To be sure these fall in the desired intervals: $c_1^- \in (0, c_m^C)$ and $c_1^+ \in (0, c_m^C)$

Since $c_1^- < c_1^+$, it is sufficient to check whether (1) $c_1^- > 0$ and (2) $c_1^+ < c_m^C$.

(1) happens if

$$\left(2\gamma+10s_1+18\gamma s_1+8\gamma^2 s_1+1\right) > s_1(2\gamma+3)\sqrt{9+16\gamma+8\gamma^2}$$

which is equivalent to

$$s_1^2\left(108\gamma+184\gamma^2+128\gamma^3+32\gamma^4+19\right)+4s_1(2\gamma+1)\left(9\gamma-18\gamma^2+44\gamma^3+5\right)+4\gamma+1>0$$

which is verified for

$$\begin{aligned} s_1 &> \frac{\sqrt{576\gamma+828\gamma^2-544\gamma^3+4192\gamma^4+13\,248\gamma^5+256\gamma^6+5632\gamma^7+30\,976\gamma^8+81}}{108\gamma+184\gamma^2+128\gamma^3+32\gamma^4+19} - \\ &\quad \frac{-2(2\gamma+1)\left(9\gamma-18\gamma^2+44\gamma^3+5\right)}{108\gamma+184\gamma^2+128\gamma^3+32\gamma^4+19} \\ &\text{or} \\ s_1 &< \frac{-\sqrt{576\gamma+828\gamma^2-544\gamma^3+4192\gamma^4+13\,248\gamma^5+256\gamma^6+5632\gamma^7+30\,976\gamma^8+81}}{108\gamma+184\gamma^2+128\gamma^3+32\gamma^4+19} - \\ &\quad \frac{-2(2\gamma+1)\left(9\gamma-18\gamma^2+44\gamma^3+5\right)}{108\gamma+184\gamma^2+128\gamma^3+32\gamma^4+19} \end{aligned}$$

the second root is always smaller than 0 and, thus, is never verified. The first is also always lower than 0 if

$$-(4\gamma+1)\left(108\gamma+184\gamma^2+128\gamma^3+32\gamma^4+19\right)<0$$

which is always verified. (1) is proved

(2) happens if

$$\frac{(2\gamma+1)(2\gamma+10s_1+18\gamma s_1+8\gamma^2 s_1+1)+s_1(2\gamma+3)(2\gamma+1)\sqrt{9+16\gamma+8\gamma^2}}{s_1(76\gamma+19s_1+108\gamma s_1+88\gamma^2+32\gamma^3+184\gamma^2 s_1+128\gamma^3 s_1+32\gamma^4 s_1+20)+(2\gamma+1)^2} < \frac{1}{1+s_1}$$

which is equivalent to

$$-4s_1^2\gamma(\gamma+1)\left(16\gamma+8\gamma^2+9\right)(4\gamma+20s_1+76\gamma s_1+4\gamma^2++19s_1^2$$
$$+108\gamma s_1^2+88\gamma^2 s_1+32\gamma^3 s_1+184\gamma^2 s_1^2+128\gamma^3 s_1^2+32\gamma^4 s_1^2+1)<0$$

This expression is always true. (2) is proved

(P) Non-empty interval condition form firm 2. We need that $\pi_2^D \geq \pi_2^C$, which happens for

$$c^2\left(4\gamma-16s_1-56\gamma s_1+4\gamma^2-17s_1^2-60\gamma s_1^2-56\gamma^2 s_1-16\gamma^3 s_1-92\gamma^2 s_1^2-64\gamma^3 s_1^2-16\gamma^4 s_1^2+1\right)+$$
$$+2c(2\gamma+1)\left(8s_1-2\gamma+12\gamma s_1+4\gamma^2 s_1-1\right)+(2\gamma+1)^2>0$$

The roots of the function are

$$c^A = \frac{(2\gamma+1)(8s_1-2\gamma+12\gamma s_1+4\gamma^2 s_1-1)-s_1(2\gamma+3)(2\gamma+1)\sqrt{(16\gamma+8\gamma^2+9)}}{s_1(56\gamma+17s_1+60\gamma s_1+56\gamma^2+16\gamma^3+92\gamma^2 s_1+64\gamma^3 s_1+16\gamma^4 s_1+16)-(2\gamma+1)^2}$$

$$c_2^- = \frac{(2\gamma+1)(8s_1-2\gamma+12\gamma s_1+4\gamma^2 s_1-1)+s_1(2\gamma+3)(2\gamma+1)\sqrt{(16\gamma+8\gamma^2+9)}}{s_1(56\gamma+17s_1+60\gamma s_1+56\gamma^2+16\gamma^3+92\gamma^2 s_1+64\gamma^3 s_1+16\gamma^4 s_1+16)-(2\gamma+1)^2}$$

We also need to analyze the concavity of the initial expression. It is convex if:

$$-s_1^2\left(60\gamma+92\gamma^2+64\gamma^3+16\gamma^4+17\right)-8s_1(2\gamma+1)(\gamma+2)(\gamma+1)+(2\gamma+1)^2>0$$

which is verified for

$$s_1 < \frac{(2\gamma+3)(2\gamma+1)\sqrt{16\gamma+8\gamma^2+9}-4(2\gamma+1)(\gamma+2)(\gamma+1)}{60\gamma+92\gamma^2+64\gamma^3+16\gamma^4+17} \equiv s_1^D$$

and

$$s_1 > \frac{-4(2\gamma+1)(\gamma+2)(\gamma+1)-(2\gamma+3)(2\gamma+1)\sqrt{16\gamma+8\gamma^2+9}}{60\gamma+92\gamma^2+64\gamma^3+16\gamma^4+17} \equiv s_1^E$$

It is easy to see that the second root is always smaller than 0, meaning that condition is always verified.

The first root will always be between 0 and 1. To prove it , $s_1^D \in (0,1)$ happens if

$$\left(60\gamma+92\gamma^2+64\gamma^3+16\gamma^4+17\right)(2\gamma+1)^2 > 0$$
$$-16(\gamma+2)\left(60\gamma+92\gamma^2+64\gamma^3+16\gamma^4+17\right)(\gamma+1)^3 < 0$$

Both conditions are always true. This means that, for $s_1 < s_1^D$ concavity is positive, being negative otherwise.

It is easy to show that, for $s_1 < s_1^D$, $c_2^- < c^A$, and the opposite for $s_1 > s_1^D$. This means that, if $s_1 < s_1^D$ the restriction is $c < c_2^-$, but if $s_1 > s_1^D$, it is also needed that $c > c^A$. Finally, we need to show that $c^A < 0$ for $s_1 > s_1^D$, meaning that it is never a restriction on our interval. Since for those values of s_1 the denominator is positive, we only need to show that the numerator is always negative. This happens if

$$-\left(60\gamma+92\gamma^2+64\gamma^3+16\gamma^4+17\right)s_1^2-8(\gamma+1)(\gamma+2)(2\gamma+1)s_1+(2\gamma+1)^2 < 0$$

Which is verified for

$$s_1 > \frac{(2\gamma+3)(2\gamma+1)\sqrt{(16\gamma+8\gamma^2+9)}-4(2\gamma+1)(\gamma+2)(\gamma+1)}{60\gamma+92\gamma^2+64\gamma^3+16\gamma^4+17}$$

or

$$s_1 < \frac{4(2\gamma+1)(\gamma+2)(\gamma+1)+(2\gamma+3)(2\gamma+1)\sqrt{(16\gamma+8\gamma^2+9)}}{-60\gamma-92\gamma^2-64\gamma^3-16\gamma^4-17}$$

The first root is always verified since it is the same as s_1^D. It is shown that the condition for the existence of a non-empty interval for the critical value for firm 2 is $c < c_2^-$.

(P) To show that $c_2^- > c_1^+$. It is equivalent to have

$$16s_1 - 4\gamma + 56\gamma s_1 - 4\gamma^2 + 17s_1^2 + 60\gamma s_1^2 + 56\gamma^2 s_1 + \\ +16\gamma^3 s_1 + 92\gamma^2 s_1^2 + 64\gamma^3 s_1^2 + 16\gamma^4 s_1^2 - 1 < 0, \text{ for } s_1 > s_1^D$$

$$16s_1 - 4\gamma + 56\gamma s_1 - 4\gamma^2 + 17s_1^2 + 60\gamma s_1^2 + 56\gamma^2 s_1 + \\ +16\gamma^3 s_1 + 92\gamma^2 s_1^2 + 64\gamma^3 s_1^2 + 16\gamma^4 s_1^2 - 1 > 0, \text{ for } s_1 < s_1^D$$

Each of these conditions is verified precisely by the condition on its respective branch, i.e., the first is verified if $s_1 > s_1^D$, while the second is verified if $s_1 < s_1^D$.

(P) Behavior of $\delta_1^C(a = 1)$. It is increasing in c if

$$\frac{\partial \delta_1^C}{\partial c} > 0$$

which happens if

$$11c - 42\gamma + 42c\gamma + 29cs_1 + 146c\gamma s_1 - 48\gamma^2 - 16\gamma^3 + \\ +48c\gamma^2 + 16c\gamma^3 + 228c\gamma^2 s_1 + 144c\gamma^3 s_1 + 32c\gamma^4 s_1 - 11 < 0$$

which is verified for

$$c < \frac{(2\gamma + 1)(20\gamma + 8\gamma^2 + 11)}{42\gamma + 29s_1 + 146\gamma s_1 + 48\gamma^2 + 16\gamma^3 + 228\gamma^2 s_1 + 144\gamma^3 s_1 + 32\gamma^4 s_1 + 11} \equiv c_1^0$$

Showing that $c_1^0 \in (c_1^-, c_1^+)$

1.$c_1^0 > c_1^-$ is verified if

$$\left(42\gamma + 29s_1 + 146\gamma s_1 + 48\gamma^2 + 16\gamma^3 + 228\gamma^2 s_1 + 144\gamma^3 s_1 + 32\gamma^4 s_1 + 11\right) \\ > -\sqrt{16\gamma + 8\gamma^2 + 9}\,(2\gamma + 3)(2\gamma + 1)(1 - s_1)$$

which is always true

2.$c_1^0 < c_1^-$ is verified if

$$0 < 4s_1^2(\gamma+2)(\gamma+1)\left(16\gamma+8\gamma^2+9\right)\left(16\gamma+8\gamma^2+5\right)(2\gamma+3)^2(2\gamma+1)^2(4\gamma+20s_1+$$
$$+76\gamma s_1+4\gamma^2+19s_1^2+108\gamma s_1^2+88\gamma^2 s_1+32\gamma^3 s_1+184\gamma^2 s_1^2+128\gamma^3 s_1^2+32\gamma^4 s_1^2+1)$$

which is always true

(P) Behavior of $\delta_2^C(a=1)$. Taking

$$\frac{\partial \delta_2^C}{\partial c} < 0$$

is verified for

$$c+3cs_1+2c\gamma s_1-1<0$$

Thus it is decreasing for

$$c < \frac{1}{3s_1+2\gamma s_1+1}$$

And increasing otherwise.

(P) Finding the conditions for the non-empty interval for firm 1, when $a=0$. Again, we need: $\pi_1^D \geq \pi_1^C$, which is now equivalent to:

$$c^2\left(\gamma-9s_1-34\gamma s_1+4\gamma^2+4\gamma^3+9s_1^2+25\gamma s_1^2-40\gamma^2 s_1-16\gamma^3 s_1+24\gamma^2 s_1^2+8\gamma^3 s_1^2\right)+$$
$$+c(2\gamma+1)\left(9s_1-2\gamma+16\gamma s_1-4\gamma^2+8\gamma^2 s_1\right)+\gamma(2\gamma+1)^2>0$$

the roots of the function are

$$c_1^{a=0} = \frac{(2\gamma+1)(2\gamma-9s_1-16\gamma s_1+4\gamma^2-8\gamma^2 s_1)-(2\gamma+3)(2\gamma+1)s_1\sqrt{16\gamma+8\gamma^2+9}}{2(\gamma-9s_1-34\gamma s_1+4\gamma^2+4\gamma^3+9s_1^2+25\gamma s_1^2-40\gamma^2 s_1-16\gamma^3 s_1+24\gamma^2 s_1^2+8\gamma^3 s_1^2)}$$

$$c^B = \frac{(2\gamma+1)(2\gamma-9s_1-16\gamma s_1+4\gamma^2-8\gamma^2 s_1)+(2\gamma+3)(2\gamma+1)s_1\sqrt{16\gamma+8\gamma^2+9}}{2(\gamma-9s_1-34\gamma s_1+4\gamma^2+4\gamma^3+9s_1^2+25\gamma s_1^2-40\gamma^2 s_1-16\gamma^3 s_1+24\gamma^2 s_1^2+8\gamma^3 s_1^2)}$$

We also need to analyze the concavity of the initial expression. It is convex if:

$$s_1^2(\gamma+1)\left(16\gamma+8\gamma^2+9\right)-s_1\left(16\gamma+8\gamma^2+9\right)(2\gamma+1)+\gamma(2\gamma+1)^2>0$$

which is verified for

$$
\begin{aligned}
s_1^A &< \frac{(2\gamma+1)(16\gamma+8\gamma^2+9)-(2\gamma+3)(2\gamma+1)\sqrt{(16\gamma+8\gamma^2+9)}}{2(\gamma+1)(16\gamma+8\gamma^2+9)} \\
&\text{or} \\
s_1^B &> \frac{(2\gamma+1)(16\gamma+8\gamma^2+9)+(2\gamma+3)(2\gamma+1)\sqrt{(16\gamma+8\gamma^2+9)}}{2(\gamma+1)(16\gamma+8\gamma^2+9)}
\end{aligned}
$$

The second root is always larger than one. To prove it, $s_1^B > 1$ is equivalent to

$$16\gamma(\gamma+2)\left(16\gamma+8\gamma^2+9\right)(\gamma+1)^2 > 0$$

This expression is always true.

The first root will always be between 0 and 1. To prove it , $s_1^A \in (0,1)$ happens if

$$4\gamma(\gamma+1)\left(16\gamma+8\gamma^2+9\right)(2\gamma+1)^2 > 0$$

$$(2\gamma+3)(2\gamma+1)\sqrt{(16\gamma+8\gamma^2+9)}+\left(16\gamma+8\gamma^2+9\right) > 0$$

Both conditions are always true. This means that, for $s_1 < s_1^A$ concavity is positive, being negative otherwise.

It is easy to show that, for $s_1 < s_1^A$, $c_1^{a=0} < c^B$, and the opposite for $s_1 > s_1^A$. This means that, if $s_1 < s_1^A$ the restriction is $c < c_1^{a=0}$, but if $s_1 > s_1^A$, it is also needed that $c > c^B$. Finally, we need to show that $c^B < 0$ for $s_1 > s_1^A$, meaning that it is never a restriction on our interval. Since for those values of s_1 the denominator is negative, we only need to show that the numerator is always positive. This happens if

$$s_1 < \frac{2\gamma(2\gamma+1)}{(16\gamma+8\gamma^2+9)-(2\gamma+3)\sqrt{16\gamma+8\gamma^2+9}}$$

This condition is always verified for $s_1 < 1$. We have shown that the condition for the existence of a non-empty interval for the critical value for firm 1 when $a = 0$ is $c < c_1^{a=0}$. This value would never happen, however, since $c_1^{a=0} > c_m^{a=0}$.

(P) Conditions for firm 2 if $a = 0$

$$\pi_2^C > \pi_2^P$$

The roots of the inequality are

$$
\begin{aligned}
c^{D} &= \frac{(2\gamma+1)\left(2\gamma+9s_1+14\gamma s_1+4\gamma^2+4\gamma^2 s_1\right)+s_1\left(2\gamma+3\right)\left(2\gamma+1\right)\sqrt{(16\gamma+8\gamma^2+9)}}{2\gamma+18s_1+64\gamma s_1+8\gamma^2+8\gamma^3-16\gamma s_1^2+64\gamma^2 s_1+16\gamma^3 s_1-24\gamma^2 s_1^2-8\gamma^3 s_1^2} \\
c_2^{a=0} &= \frac{(2\gamma+1)\left(2\gamma+9s_1+14\gamma s_1+4\gamma^2+4\gamma^2 s_1\right)-s_1\left(2\gamma+3\right)\left(2\gamma+1\right)\sqrt{(16\gamma+8\gamma^2+9)}}{2\gamma+18s_1+64\gamma s_1+8\gamma^2+8\gamma^3-16\gamma s_1^2+64\gamma^2 s_1+16\gamma^3 s_1-24\gamma^2 s_1^2-8\gamma^3 s_1^2}
\end{aligned}
$$

Concavity

$$
s_1^2\left(8\gamma+12\gamma^2+4\gamma^3\right)-\left(14\gamma+4\gamma^2+9\right)\left(2\gamma+1\right)s_1-\gamma\left(2\gamma+1\right)^2>0
$$

$$
\begin{aligned}
s_1 &< \frac{(2\gamma+1)\left(14\gamma+4\gamma^2+9\right)-(2\gamma+3)\left(2\gamma+1\right)\sqrt{(16\gamma+8\gamma^2+9)}}{16\gamma+24\gamma^2+8\gamma^3}<0 \\
&\text{or} \\
s_1 &> \frac{32\gamma+32\gamma^2+8\gamma^3+9+(2\gamma+3)\left(2\gamma+1\right)\sqrt{(16\gamma+8\gamma^2+9)}}{16\gamma+24\gamma^2+8\gamma^3}>1
\end{aligned}
$$

Concavity is always negative.

One can show that $c^D > c_m^C$. The condition for non-empty interval becomes $c < c_2^{a=0}$

(P) $\delta_1^C(a=0)$ decreasing until $c=\frac{\gamma}{\gamma+s_1}$. Increasing otherwise - this proof is provided before being mentioned in the article, since it is necessary for the following proof. The inequality is

$$
\frac{\partial\delta_1^C(a=0)}{\partial c}<0
$$

which is verified for

$$
c<\frac{\gamma}{\gamma+s_1}
$$

(P) To show that the ratio $\frac{\pi_{1r}^D-\pi_{1r}^C}{\pi_{1r}^D-\pi_{1r}^P}$ is always increasing for $a=1$:

$$
\frac{\partial\left(\frac{\pi_{1r}^D-\pi_{1r}^C}{\pi_{1r}^D-\pi_{1r}^P}\right)}{\partial c}>0
$$

Is the same as

$$
\frac{32}{\left(c-2\gamma+2c\gamma+7cs_1+10c\gamma s_1+4c\gamma^2 s_1-1\right)^2}>0
$$

which is always true.

For $a = 0$ that ratio is always equal to $\delta_1^C(a=0)$, which, as seen before, is decreasing for the relevant values of c.

(P) Showing that $\delta_2^C(a=0)$ is increasing for $c \in (0, c_2^{a=0})$. The condition is

$$\frac{\partial \delta_2^C(a=0)}{\partial c} > 0$$

which is verified for

$$\left(\gamma(2\gamma+1)\left(20\gamma+8\gamma^2+11\right)(c-1)+9cs_1+41c\gamma s_1+48c\gamma^2 s_1+16c\gamma^3 s_1\right)(\gamma-c\gamma+cs_1+c\gamma s_1)<0$$

The left term: $(\gamma(2\gamma+1)(20\gamma+8\gamma^2+11)(c-1)+9cs_1+41c\gamma s_1+48c\gamma^2 s_1+16c\gamma^3 s_1) < 0$ for

$$c < \frac{\gamma(2\gamma+1)(20\gamma+8\gamma^2+11)}{9s_1+41\gamma s_1+48\gamma^2 s_1+16\gamma^3 s_1+\gamma(2\gamma+1)(20\gamma+8\gamma^2+11)}$$

The right term: $\gamma + c(s_1 - \gamma(1-s_1)) > 0$ for

$$\begin{aligned} c &> \frac{-\gamma}{s_1-\gamma(1-s_1)}, \text{ if } s_1-\gamma(1-s_1)>0 \\ c &< \frac{-\gamma}{s_1-\gamma(1-s_1)}, \text{ if } s_1-\gamma(1-s_1)<0 \end{aligned}$$

The first inequality is always true. If the second case happens we have

$$\frac{\gamma(2\gamma+1)(20\gamma+8\gamma^2+11)}{9s_1+41\gamma s_1+48\gamma^2 s_1+16\gamma^3 s_1+\gamma(2\gamma+1)(20\gamma+8\gamma^2+11)}-\frac{-\gamma}{s_1-\gamma(1-s_1)}= \\ \frac{2(16\gamma+8\gamma^2+5)(\gamma+1)(\gamma+2)\gamma s_1}{(11\gamma+9s_1+41\gamma s_1+42\gamma^2+48\gamma^3+16\gamma^4+48\gamma^2 s_1+16\gamma^3 s_1)(s_1-\gamma(1-s_1))}$$

which is always negative. Thus, the derivative is positive if

$$c < \frac{\gamma(2\gamma+1)(20\gamma+8\gamma^2+11)}{9s_1+41\gamma s_1+48\gamma^2 s_1+16\gamma^3 s_1+\gamma(2\gamma+1)(20\gamma+8\gamma^2+11)}$$

Finally, it is easy to show that this threshold for c is higher than c^D, which is higher than $c_2^{a=0}$. This means that, for $c \in (0, c_2^{a=0})$, $\delta_2^C(a=0)$ is increasing.

Proofs of Section 4

(P) Showing that it would never be optimal for firm U to foreclose firm 2

If there is a downstream monopoly by firm 1, firm U would choose $c = \frac{1}{2(1-s_1)}$.

Comparing with the competition scenario, firm U's profits would be higher if

$$\frac{1}{8(1-s_1)(2+\gamma)} > \frac{1}{(2-s_1)(2\gamma+3)}$$

Which happens for

$$s_1 > \frac{4\gamma+10}{6\gamma+13}$$

Noting that this value of s_1 is higher than s_1^F, one can see that if indeed it happens, firm U would be even better in unconstrained collusion than under firm 1's monopoly, since

$$\frac{1}{8(1-s_1)(2+\gamma)} > \frac{1}{4(\gamma+2)(1-s_1)}$$

This inequality is verified for any value of $s_1 \in (0,1)$.

It might happen that unconstrained collusion could not be achieved for such s_1. In which case foreclosing firm 2 would be optimal. We can see, however, that this limit on s_1 is quite higher than 0.5, meaning it is not consistent with our non-control assumption. This is the only case where imposing the upper bound of s_1 to be 0.5 and not 1 could change a result. This result, however, is irrelevant for the rest of the article and, thus, none of the main results would be affected by changing the bound.

(P) Computing c_x. This value is implicitly defined by $W^P(c=c_x) = W_0^C$. This means

$$\frac{-\left(4(\gamma+1)(2\gamma+1)^2(1-s_1) + s_1^2\left(14\gamma+20\gamma^2+8\gamma^3+1\right)\right)c^2}{2(2\gamma+3)^2(2\gamma+1)^2} + \\ + \frac{-2(2\gamma+1)^2(2-s_1)c + 4(\gamma+2)(2\gamma+1)^2}{2(2\gamma+3)^2(2\gamma+1)^2} - \frac{1}{2}\frac{\gamma+3}{(\gamma+2)^2} = 0$$

This expression will have two roots:

$$c_x = \frac{(2\gamma+3)(2\gamma+1)\sqrt{4(2\gamma+1)^2(1-s_1)+s_1^2(9\gamma+7\gamma^2+1)}-(\gamma+2)(2\gamma+1)^2(2-s_1)}{(\gamma+2)\left(4(\gamma+1)(2\gamma+1)^2(1-s_1)+s_1^2+14\gamma s_1^2+4s_1^2\gamma^2(2\gamma+5)\right)}$$

$$c = \frac{-(2\gamma+3)(2\gamma+1)\sqrt{4(2\gamma+1)^2(1-s_1)+s_1^2(9\gamma+7\gamma^2+1)}-(\gamma+2)(2\gamma+1)^2(2-s_1)}{(\gamma+2)\left(4(\gamma+1)(2\gamma+1)^2(1-s_1)+s_1^2+14\gamma s_1^2+4s_1^2\gamma^2(2\gamma+5)\right)}$$

It is clear to see that the second root is always negative, which means the solution is found in the first one. To guarantee that $c_x > 0$ we need

$$(3\gamma+5)(2\gamma+1)^2\left(4(\gamma+1)(2\gamma+1)^2(1-s_1)+14\gamma s_1^2+s_1^2+20\gamma^2 s_1^2+8\gamma^3 s_1^2\right)>0$$

this is always true.

(P) Proving that δ_1^X is increasing in s_1.

The function δ_1^X can be achieved by replacing $\delta_1^C(c=c_x)$. It has been shown before that δ_1^C is increasing in c. Now it only remains to be shown that c_x is increasing in s_1.

Since c_x is implicitly defined by $W^P(c=c_x)=W_0^C$, and W_o^C does not depend on either c or s_1 it must happen that

$$\frac{\partial c_x}{\partial s_1} = -\frac{\frac{\partial W^P}{\partial s_1}}{\frac{\partial W^P}{\partial c}}$$

Since $\frac{\partial W}{\partial c}<0$, c_x will be increasing in s_1 if W^P is increasing in s_1. This happens if

$$2c+4\gamma+10c\gamma-cs_1-14c\gamma s_1+4\gamma^2+16c\gamma^2+8c\gamma^3-20c\gamma^2 s_1-8c\gamma^3 s_1+1>0$$

which is verified for

$$s_1 < \frac{(2\gamma+1)^2(2c+2c\gamma+1)}{(14\gamma+20\gamma^2+8\gamma^3+1)c}$$

For $s_1<1$ this is always true. It is proven that c_x is increasing in s_1.