Imperfect collusion in an asymmetric duopoly*

Colusión imperfecta en un duopolio asimétrico

Marc Escrihuela-Villar**
Carlos Gutiérrez-Hita***

Abstract

Using the coefficient of cooperation, we analyse the effect of cost asymmetries on collusive agreements when firms are able to coordinate on distinct output levels than the unrestricted joint profit maximization outcome. In this context, we first investigate the extent to which collusive agreements are feasible. Secondly, we focus on collusion sustainability in an infinitely repeated game. We show that, regardless of the degree of cost asymmetry, at least some collusion is always sustainable. Finally, the degree of collusion is also endogeneised to show that cooperation has an upper bound determined by the most inefficient firm.

Key words: *Imperfect collusion, cost asymmetries, sustainability.*

JEL Classification: *L11, L13, L41, D43.*

Resumen

Usando el coeficiente de cooperación, analizamos el efecto de las asimetrías en costes en los acuerdos colusorios cuando las empresas son capaces de coordinarse en niveles de producción distintos de aquel que maximiza el beneficio conjunto. En este contexto, primero investigamos en qué medida son factibles los acuerdos colusorios. En segundo lugar, nos centramos en la sostenibilidad

* Authors are grateful to the EARIE and ASSET participants and to one anonymous referee for their helpful comments. Financial support by the "Ministerio de Economía y Competitividad" and the "Fondo Europeo de Desarrollo Regional" through their projects "Conflicto e influencia estratégica: Los efectos de la actividades de captación de rentas en las decisiones colectivas" (Ref: ECO2015-67901-P MINECO/FEDER) and "Strategic management on complex systems with multiplicity of agents" (Ref: MTM2014-54199-P) are gratefully acknowledged. The usual disclaimer applies.

** Corresponding author. Departamento de Economía Aplicada. Universitat de les Illes Balears. Edificio Jovellanos Ctra. Valldemossa km 7.5. 07122 Palma de Mallorca (Spain). Tel. +34 971173242. Fax: +34 971172389. E-mail: marc.escrihuela@uib.es

*** Departamento de Fundamentos del Análisis Económico. Universitat d'Alacant. Campus San Vicente del Raspeig, Alicante (Spain). Tel. +34 966658869. E-mail: carlosghita@ua.es

Received: March, 2017. Accepted: February, 2018.

de la colusión en un juego repetido infinitos periodos. Se demuestra que, independientemente del grado de asimetría en costes, al menos cierto nivel de colusión siempre es sostenible. Finalmente, también se obtiene el grado de colusión endógeno para demostrar que la cooperación tiene un límite superior determinado por la empresa más ineficiente.

Palabras clave: *Colusión, asimetrías en costes, sostenibilidad.*

Clasificación JEL: *L11, L13, L41, D43.*

1. Introduction

A standard assumption in cartel formation and collusion sustainability has been that all firms are identical in terms of their costs, even if there was a degree of differentiation amongst their products or in firm's timing decision. However, the assumption of cost symmetry is unrealistic and very restrictive when one looks to real markets. A primer approach to tackle cost asymmetries was inferred from an early paper by Patinkin (1947), where a cartel maximises total industry profits and therefore allocates output quotas so that the cartelised industry operates as if it was a multiplant monopolist allocating output between plants. Moreover, costs asymmetries are proved to play also an important role when firms attempt to reach collusion[1]. As Bain (1948) points out cost heterogeneity would mean that, in the absence of side payments between firms, such an allocation might not be viable as inefficient firms may obtain lower profits in the cartel than in the non-cooperative equilibrium. The intuition is that firms may find it difficult to agree on a common collusive policy because firms with a lower marginal cost will insist in lower prices than those firms with higher marginal cost would wish to sustain. More generally, the common wisdom is that the diversity of cost structures may rule out any possible agreement in pricing policies and so exacerbate coordination problems. In addition, technical efficiency would require allocating higher production quotas to low-cost firms, but this would clearly be difficult to sustain in the absence of explicit agreements or side transfers. Thus, it seems natural to characterise a class of agreements different to the most collusive outcome (i.e., the monopoly solution), namely imperfect collusion[2]. Such an agreement allows firms to achieve some degree of collusion and also to sustain that agreement over time.

Admittedly, the possibility of collusive firms operating with different cost functions has also received some attention in the modern Industrial Organisation literature. For instance, Osborne and Pitchik (1983) in a static non-cooperative model where firms are capacity-constrained allow for side payments and show

1 The literature on partial cartels with a dominant firm (namely, a cartel) facing a competitive fringe has also examined this issue. As an example, Donsimoni (1986) has shown that in this framework a degree of cost heterogeneity can be accommodated.

2 This concept is different to partial collusion which refers to a market where only a subset of firms colludes. See for instance Verboven (1997), Escrihuela-Villar (2008) or Mendi *et al.* (2011).

that the profit per unit of capacity of the small firm is higher than that of the large one. Schmalensee (1987) in a static game with linear costs in a Cournot setting characterises the set of profit vectors by applying a number of selection criteria such as the Nash bargaining solution. He finds that if a leading firm's cost advantage is substantial, its potential gains from collusion are relatively small. By their very nature, however, in a static model cartel members do not cheat on a cartel agreement since it is assumed that agreements are sustained through binding contracts. This may, therefore, be viewed as a model of explicit or binding collusion. These papers thus do not impose the incentive compatibility constraints of subgame perfection and the collusive outcome derived in their models may not be self-enforced. Looking at Friedman's (1971) supergame-theoretic approach to collusion a few papers have also considered the problem of enforcement of collusive behaviour with asymmetric firms. Rothschild (1999) shows that the stability of the cartel may depend crucially upon the relative efficiencies of the firms and that joint profit maximisation becomes less likely as cost functions differ across firms. Vasconcelos (2005) in a quantity setting oligopoly model assumes asymmetry by modelling that firms have different shares of a specific asset and shows that the sustainability of perfect collusion crucially depends on the most inefficient firm in the agreement, which represents the main obstacle to the enforcement of collusion. More recently, Miklós-Thal (2011) shows that in a Bertrand supergame some collusion is also sustainable under cost asymmetry whenever collusion is sustainable under cost symmetry and Contreras *et al.* (2008) have shown that with differentiated products a cartel may also be stable provided that returns to scale are high enough. Summarising, both the literature on static cartel stability and the dynamic models of tacit collusion suggest that collusion is unlikely to be observed in the presence of substantial competitive advantage, and therefore, a prior step before studying collusion sustainability when costs are heterogeneous and firms agree on output quotas is to consider whether collusion is viable. In line with these results, the analysis of the empirical literature also indicates that cost asymmetries hinder collusion (see for instance Levenstein and Suslow, 2006)[3].

In this paper, we investigate the extent to which imperfect collusion can be a way to sustain a collusive agreement in the presence of large cost asymmetries when firms produce a homogeneous product. First, we study whether cost heterogeneity is sufficient to make it impossible for firms to collude (and its sustainability over an infinite horizon) when coordination is not necessarily on the allocation that maximises total industry profits. In this sense, in their empirical studies Eckbo (1976) and Griffin (1989) provide an interesting motivation with this respect by finding that even though cartels that are made up of similar-sized firms are more able to raise prices, in some cases high-cost members of a cartel may produce at a cost larger than 50% above low-cost members[4]. Consequently, the question about why and to what extent should cost asymmetries be a re-

3 Experimental works provide equivalent results. For instance, Mason, Phillips and Nowell (1992) show that in an experimental duopoly game cooperation is also more likely when players face symmetric production costs.

4 In this line, Wirl (2015) empirically analyses OPEC behavior in a Cournot setting. It is assumed that OPEC adjustments affect country members in a different way depending

straint for collusion naturally arises. Secondly, we also study how the degree of collusion can be endogenously determined. Such degree of collusion may be interpreted as part of an explicit cartel agreement (and allocating thus different output quotas depending on the firm's relative efficiency) or it can be thought of as a degree of coordination when collusion is tacit.

Once the degree of collusion is characterised we investigate the sustainability of imperfect collusive agreements in a multi-period duopoly model. We use subgame perfect Nash equilibria –henceforth, SPNE– as solution concept. It is well known that this repeated game setting exhibits multiple equilibria. To select among those equilibria we adopt the particular criterion of restricting strategies to grim trigger strategies where firms adhere to the collusive agreement until there is a defection, in which case they revert forever to the static Cournot equilibrium. The key feature is the assumption that firms maximise the summation of its own profits plus a proportion of the profits of the other firms. As a consequence, this proportion may be considered as the degree of coordination under an (imperfect) collusive agreement. This approach has received growing attention of scholars (see for instance Symeonidis, 2008 or Matsumura *et al.*, 2013), it is also closely related to the coefficient of cooperation defined by Cyert and DeGroot (1973), and captures the relative performance approach that is evolutionary stable (Vega-Redondo, 1997). These objective functions are also in line with the growing and more recent behavioural economics literature, as well as with experimental games that test the extent to which subjects are concerned with reciprocity (see for instance Fehr and Schmidt, 1999 and Charness and Rabin, 2002 respectively).

Our main contribution is to show that in a quantity-setting model some degree of imperfect collusion can always be sustained regardless of the cost heterogeneity. The intuition is that even though cost asymmetry hinders collusion, for each possible level of their discount factor firms can always coordinate on an output level below the competitive one. Hence, one can expect collusion between firms to occur, at least to some extent, also with very asymmetric firms since these firms may still have an incentive to adhere to an imperfectly collusive agreement. Consequently, the antitrust authorities should also be cautious when firms in an industry have significantly different cost functions because firms' willingness to collude may be still present. It is also obtained that the endogenous degree of collusion to be sustained has an upper bound that can never be overcome. This boundary is determined by the most inefficient firm and depends negatively on the degree of cost asymmetry. In other words, although some degree of coordination (generally) might lead to higher profits overall, at the same time the aforementioned adjustments in the degree of collusion seem to be limited by the willingness of the most inefficient firms. Therefore, in a market with significant production cost asymmetries, a negative relationship

on its production capacity. His results suggest that the OPEC does not behave as a perfect cartel in the sense that country members should reduce outputs below its current level.

between firms' efficiency and the endogenous choice of the degree of sustainable collusion can be expected[5].

The remainder of the paper is organised as follows. In Section 2 the model is presented. In Sections 3 and 4 we analyse imperfect collusion and its sustainability. In Section 5 the degree of collusion is made endogenous. Section 6 presents some extensions of the model showing, for instance, that the effect of cost asymmetries on firms' collusion incentives is also robust to other ways to parameterise the product-market competition. Section 7 concludes. All proofs are grouped together in the Appendix.

2. Set Up of The Model

We consider an industry with two asymmetric firms indexed by $i = 1,2$ where each firm simultaneously produces a homogeneous product. Even though quantity competition is assumed, it is well-known from Kreps and Scheinkman (1983) that in a two-stage oligopoly game where, first, there is simultaneous production, and, second, after production levels are made public, there is price competition, the unique equilibrium outcome is the Cournot outcome. Consequently, the present model could also be interpreted as a market in which firstly capacities are determined and secondly price competition takes place. Regarding production costs, we assume a technology such that firms produce with a quadratic cost function $c_i(q_i) = \rho_i(c)q_i^2$ where q_i is the output produced by firm *i*. The function $\rho_i(c)$ accounts for the asymmetry between firms, taking values $1 + c$ and $1 - c$, respectively, with $0 < c < 1$. Without loss of generality, we assume that firm 1 is the inefficient and firm 2 is the efficient one[6]. Hence, as *c* approaches the unity firms become more asymmetric. Differences in the slope of marginal costs can be interpreted for instance as resulting from differences in capital stocks[7]. Some features of the cost functions should be emphasised at the outset. Firstly, following the reasoning provided by Rothschild (1999) and in order for the collusive outcome not to become trivial, the cost functions are not linear. If on the contrary, firms had different but constant marginal costs, it would clearly be practical for the firm with the lowest costs to produce the entire output. In the present model though, a switch of production exclusively to the most efficient firm would raise industry costs. The second feature is that fixed costs are taken

---

5 Our model presents also a similarity with partial ownership arrangements. Even though cross-ownership also takes into account rivals' profits in the static Cournot game, in our model, the above-mentioned reciprocity only implies cooperation among firms if collusion is sustainable whereas firms' reciprocal concern disappears in a competitive environment. In this sense, our approach differs from that in the tradition of Fershtman and Judd (1987) where the owners of each firm may write a contract in the first stage in which its managers are remunerated according to their performance compared to their competitors.

6 The results of the present paper also carry over to an oligopoly model with n efficient and n inefficient firms. We present the duopoly model for ease of exposition since comparative statics with respect to n just confirms the intuitions provided in the comprehensive surveys and the basic game-theoretical results on collusion exploring which factors facilitate or hinder collusion. Details from this extension are available from the authors upon request.

7 In Section 6 this idea is extended to discuss the effect of capacity constraints in the model.

to be zero. This simplifying assumption is common in the literature, partly for simplicity but also because provided that fixed costs are not so high as to force firms out of the market, the relative magnitudes of the payoffs to firms from different actions are unaffected by the omission of a fixed cost.

The industry inverse demand is given by the piecewise linear function $p(Q) = max(0,1 - Q)$, where p is the output price and $Q = q_1 + q_2$ is the industry output. In the absence of coordination, each firm plays a Cournot stage game. The profit function for firm i is given by:

$$\Pi_i\left(q_i, q_j\right) = p\left(q_i, q_j\right)q_i - \rho_i(c)q_i^2 \quad i, j = 1,2 \quad i \neq j \tag{1}$$

We characterise imperfect collusion in the Cournot-stage game considering a particular model where each firm maximises the sum of its own profit and a fraction of the profit of the other firm. Explicitly, each firm i maximises $\prod_i(q_i, q_j) + \alpha \prod_j(q_i, q_j)$ where $\alpha \in [0,1]$. We assume α to be constant and symmetric in such a way that, regardless of whether a firm is efficient or inefficient their degrees of reciprocal concern with the rival coincide. The parameter α thus can be interpreted as representing the degree of collusion[8]. Admittedly, although we build a direct link between α and the degree of collusion, the reader might feel more comfortable when such link is made explicit and based on a direct behavioural assumption like the output produced. Arguably, we could also interpret our model as one in which firms' strategy set is a quantity in the interval between the joint-profit maximising allocation and the asymmetric Cournot equilibrium where α merely parameterises the most collusive output achievable as a result of the efficiency differences between firms.

Definition 1*: Collusion is said to be imperfect if* $\alpha \in (0,1)$. *On the contrary, collusion is said to be perfect if* $\alpha = 1$.

An alternative consideration is that since we assume away side-payments, one could think that firms should bargain over possible outputs[9]. Our approach is somewhat different. We assume that firms coordinate to behave as in a model of symmetric cross-ownership but according to their different efficiency level, even though in the present model firms behave as Cournot competitors when they do not collude. In other words, we consider a profit-sharing rule where firms' profits are to some extent proportional to capital stocks. Bos and Harrington (2010) and the references cited therein provide abundant motivation about this rule often referred to as a proportional rule.

---

8 We note that Escrihuela-Villar (2015) shows that using conjectural variations and the coefficient of cooperation leads to equivalent closed-form solutions.

9 In this line, Schmalensee (1987) uses an axiomatic bargaining model to show that low-cost firms may have little to gain from collusion.

3. Characterisation of an Imperfect Collusive Agreement

Let q_i^c denote the quantity corresponding to the collusive output. Given a degree of asymmetry *c*, we obtain the following collusive equilibrium quantities and profits for firm *i*,

$$(2)\quad q_i^c(\alpha)=\frac{1-\alpha-2\rho_j(c)}{(3-\alpha)(5+\alpha)-4c^2},\quad \Pi_i^c(\alpha)=\left[q_i^c(\alpha)\right]^2\left[\frac{5+\rho_j(c)+\alpha\rho_i(c)-\alpha^2-2c^2}{1-\alpha+2\rho_j(c)}\right],$$

where throughout the paper and abusing notation, we assume as exogenously given *c*. It is straightforward to obtain the Nash-Cournot non-cooperative equilibrium and the associated level of profits for each firm in (2) for $\alpha = 0$. We denote them by q_i^*, and Π_i^* respectively.

We note that $\partial q_i^c(\alpha)/\partial\alpha < 0$ since if α increases firms' production tends to the one of a perfectly collusive market where outputs are reduced in order to increase the price. However, $q_i^c(\alpha)$ decreases (increases) with *c* for firm 1 (firm 2) since collusion requires an efficient reallocation of outputs concentrating production in the more efficient firm. Consequently, when c increases over a certain critical value, an inefficient firm may not be interested in being part of a perfectly collusive agreement because profits attained under such an agreement are lower than those obtained at the Cournot stage game. Intuitively, if collusion is perfect, in order to maximise joint profits the reduction needed in the quantity produced by the inefficient firm compared to the Cournot equilibrium is not compensated by the price increase. This argument also applies if α is large enough but lower than 1. However, if α is low enough imperfect collusion can still allow the inefficient firm to obtain larger profits than without collusion for any level of efficiency differences. On the contrary, it is straightforward to check that $\Pi_2^c(\alpha) > \Pi_2^*$ for any $\alpha \in (0,1]$. Since collusion implies that the cartel minimises the total costs of producing a given output level, allocating thus production amongst the firms in such a way that marginal costs are equalised, the efficient firm always benefits from collusion. In fact, this implies that a firm's share in the output of the cartel, and the profits which it obtains are larger if the firm is relatively more efficient[10]. The following lemma proves that at least some collusion is always feasible with an upper bound negatively associated with the cost asymmetry[11].

[10] We note that this is only one of several possible bases for allocating cartel output but it has the advantage of being simple to implement and intuitively derived from a cost minimization process. See Schmalensee (1987) for different plausible alternatives.

[11] We note thus that with increasing marginal costs and imperfect collusion the less efficient firm has still positive production levels. This is in line with the results of Schmalensee (1987) with bargaining and a linear cost function.

Lemma 1: *For any* $c \in (0,1)$ *there always exists* $\tilde{\alpha} \in (0,1)$ *such that* $\Pi_1^c(\alpha) > \Pi_1^*$ *if* $\alpha < \tilde{\alpha}$, *where* $\tilde{\alpha}$ *decreases with c. Conversely,* $\Pi_1^c(\alpha) < \Pi_1^*$ *if* $\alpha > \tilde{\alpha}$.

4. Sustainability of Imperfect Collusion

We assume in this section that firms compete repeatedly over an infinite horizon with complete information (i.e. both firms observe the whole history of actions) and discount the future according to a common discount factor $\delta \in (0,1)$. At any stage the profit function is given by (1). Time is discrete and dates are denoted by t = 1, 2,... In this framework, a pure strategy for firm i is an infinite sequence of functions $\{S_i^t\}_{t=1}^{\infty}$ with $S_i^t : \sum^{t-1} \rightarrow Q$ where $\sum^{t-1}$ is the set of all possible histories of actions (output choices) of each firm up to t – 1, with typical element σ_i^τ, $i = 1,2$, $\tau = 1,...,t\text{-}1$, and Q is the set of output choices available to each firm. We follow Friedman (1971) restricting our attention to the case where each firm is only allowed to follow grim trigger strategies such that firms adhere to a collusive agreement until there is a defection, in which case they revert forever to the static Nash-Cournot equilibrium. Hence, $\{S_i^t\}_{t=1}^{\infty}$ can be specified as follows. At t = 1, $S_i^1 = q_i^c$, while at t = 2,3,...

$$S_i^t(\sigma_j^\tau) = \begin{cases} q_i^c(\alpha) \text{ if } \sigma_j^\tau = q_j^c(\alpha) \text{ for all } j = 1,2,\ \tau = 1,...,\text{t-1} \\ q_i^* \qquad\qquad\qquad\qquad \text{otherwise,} \end{cases} \tag{3}$$

Although there is a multiplicity of equilibria since the above strategies sustain different collusive outputs, we focus on an equilibrium that depends on α and hence on the efficiency differences between firms[12]. Thus, firms producing $q_i^c(\alpha)$ in each period can be sustained as a SPNE of the repeated game with the strategy profile (3) if and only if the following conditions are satisfied

$$\delta \geq \left[\Pi_i^d(\alpha) - \Pi_i^c(\alpha)\right]\left[\Pi_i^d(\alpha) - \Pi_i^*\right]^{-1} \text{ for } i = 1,2 \tag{4}$$

where $\Pi_i^d(\alpha)$ denotes the profits obtained by firm i in an optimal deviation from the collusive output $q_i^c(\alpha)$. In other words, if δ exceeds a certain critical level the inequalities described in (4) are satisfied. We denote by $\tilde{\delta}_i(\alpha)$ this critical value of the discount factor where $q_i^c(\alpha)$ is a SPNE of the repeated game if $\delta \geq \max\left\{\tilde{\delta}_1(\alpha), \tilde{\delta}_2(\alpha)\right\}$. In order to characterise $\tilde{\delta}_i(\alpha)$, we need to define

---

[12] Note that we do not use α as a device to select among the multiple equilibria of a collusive dynamic game. As mentioned above, α is merely a parameter that characterizes and justifies the extent of imperfect collusion.

$\prod_i^d(\alpha)$ which is obtained by replacing $q_i^d(\alpha)$ in $\prod_i(q_i, q_j^c(\alpha))$ for $j \neq i$ where $q_i^d(\alpha) = \arg\ \max_{q_i} \prod_i(q_i, q_j^c(\alpha))$. Therefore, $\prod_i^d(\alpha) = \left[q_i^d(\alpha)\right]^2 (1+\rho_i(c))$, where

$$q_i^d(\alpha) = \frac{10 + 2\rho_j(c) + \alpha(5-3\alpha) - 4c^2}{2\left(1+\rho_i(c)\right)\left((3-\alpha)(7+3\alpha) - 4c^2\right)}, \text{ for } i, j = 1,2 \text{ and } i \neq j.$$

As we prove in the proof of Proposition 1 the condition on δ in (4) is always more easily satisfied for the efficient firm than for the inefficient one (namely, $\tilde{\delta}_1(\alpha) > \tilde{\delta}_2(\alpha)$)[13]. Therefore, we can define collusion sustainability as follows:

Definition 2: *Imperfect collusion is sustainable if* $\delta \geq \tilde{\delta}_1(\alpha)$.

The following proposition shows that (some) imperfect collusion is always sustainable regardless of the cost asymmetry of firms.

Proposition 1: *For any* $c \in (0,1)$ *there always exists* $\tilde{\alpha} \in (0,1)$ *such that no matter how small* δ *is, collusion on an output level below the competitive one is sustainable if* $\alpha < \tilde{\alpha}$ *with* $\tilde{\delta}_1(\alpha) < 1$. *Furthermore,* $\tilde{\delta}_1(\alpha)$ *increases with* α *and c.*

In other words, α is an upper bound on the degree of collusion that can be sustained in the infinitely repeated game. The above proposition thus extends Lemma 1 in the sense that collusion sustainability is only possible if firms agree on a lower degree of collusion[14]. The intuition is fairly simple. From Rothschild (1999) we know that if in our model $\alpha = 1$, perfect collusion becomes harder to sustain if firms' cost asymmetry increases in such a way that if c is large enough, perfect collusion is not sustainable[15]. However, if we allow firms to sustain less collusion, the agreement may be sustained. Consequently, there is always a small enough degree of collusion that can be sustained despite the difference between firms' costs. A numerical example may help clarify our result. For instance, perfect collusion with $c > 0.202$ yields to $\tilde{\delta}_1(\alpha) > 1$. Consequently, the standard joint profit maximisation allocation cannot be sustained. However, when

---

13 We note that Rothschild (1999) obtained the equivalent result for $\alpha = 1$.

14 We also note that, since (4) can also be written as $\delta \geq \left[\Pi_i^d(\alpha) - \Pi_i^c(\alpha)\right]\left[\Pi_i^d(\alpha) - \Pi_i^*\right]^{-1}$ and $\Pi_i^d(\alpha) - \Pi_i^c(\alpha)$ is always true regardless the value of α, the upper bound on α such that $\tilde{\delta}_1(\alpha) < 1$ coincides with the one obtained in Lemma 1 needed in order for $\Pi_1^d(\alpha) > \Pi_1^*$ to be satisfied.

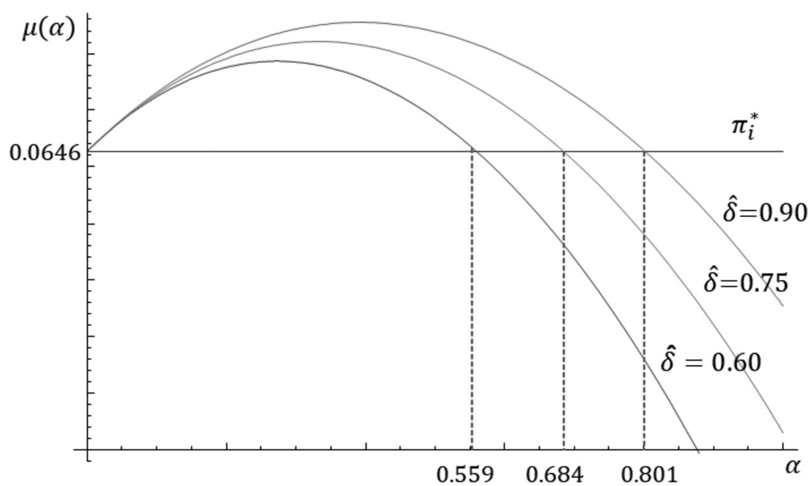
15 Additionally, Harrington (1991) uses the Nash bargaining solution concept to obtain that, in general, the larger the cost differences the higher the discount factor needed to sustain tacit collusion. More recently Ciarreta and Gutiérrez-Hita (2012) reach similar conclusions when firms compete in supply functions.

we consider imperfect collusion, $\alpha \leq 0.5$ can be sustained if $\delta \geq 0.467$ (since $\tilde{\delta}_1(0.5) = 0.467$). Lemma 1 and Proposition 1 show that if α is small enough, the standard Prisoner's Dilemma scheme of a quantity-setting collusive market is restored since $\prod_i^d(\alpha) > \prod_i^c(\alpha) > \prod_i^* > 0$ where (some) imperfect collusion is always sustainable. This result can also be interpreted in the following way. One can check that for a given value of δ and when α = 0,

1) The incentive constraint described in (4) is binding for the inefficient firm: $\prod_1^d(0) = \prod_1^c(0) = \prod_1^*$.
2) $\partial\prod_1^c(\alpha)/\partial\alpha > (1-\delta)(\partial\prod_1^d(\alpha)/\partial\alpha) > 0$ when α increases from zero.

Hence, there always exists an incentive to assume at least some coordination between firms because, in a Cournot market, collusion profits increase with the degree of collusion more than what deviation profits do. In fact, for a given $\delta \in (0,1)$ the left hand side of the inequality $\left[\prod_i^c(\alpha) - (1-\delta)\prod_i^d(\alpha)\right]/\delta \geq \prod_i^*$, which is obtained from (4), has an inverse U-shape function relationship with respect to α. Figure 1 shows this insight for $\delta = \{0.6, 0.75, 0.9\}$ and $c = 0.25$ where $\mu(\alpha) = \prod_1^c(\alpha) - (1-\delta)\prod_1^d(\alpha)$. In these cases, the maximum level of imperfect collusion that can be sustained as a SPNE of the infinitely repeated game is respectively $\alpha = \{0.559,\ 0.684,\ 0.801\}$.

FIGURE 1



5. Endogenous Imperfect Collusion

In this section, α is made endogenous by adding an initial stage in which firms choose the extent of imperfect collusion. We assume that firms firstly and simultaneously choose a degree of collusion, namely α_i, to afterwards continue with the infinitely repeated game described above. We characterise the problem that firms solve in the first stage denoting the profits that firms obtain in the initial stage by $\Pi_i^{\alpha}(\alpha_i,\alpha_j)$ for i and j = 1, 2 and $i \neq j$ where firm i maximises

$$(5) \quad \max_{\alpha_i} \Pi_i^{\alpha}(\alpha_i,\alpha_j) = \begin{cases} \Pi_i^{c}(\alpha_i,\alpha_j) \text{ if } \delta \geq \max\left\{\tilde{\delta}_i(\alpha_i,\alpha_j),\tilde{\delta}_j(\alpha_i,\alpha_j)\right\} \\ \Pi_i^{*} \quad \text{otherwise,} \end{cases}$$

and where $\Pi_i^{c}(\alpha_i,\alpha_j)$ is analogous to the collusive profit function defined in (2) but for the case of asymmetric α. Firm i chooses α_i anticipating that both firms will only maximise their (imperfectly) collusive profits whenever the degree of collusion chosen by firms is sustainable. Otherwise, firms' profits correspond with the non-cooperative Nash-Cournot equilibrium. In other words, once firms decide about α_i and α_j, and according to (3), imperfect collusion may be sustained as a SPNE of the repeated game only if $\delta \geq \max\left\{\tilde{\delta}_1(\alpha_1,\alpha_2),\tilde{\delta}_2(\alpha_1,\alpha_2)\right\}$. On the contrary, if $\delta < \max\left\{\tilde{\delta}_1(\alpha_1,\alpha_2),\tilde{\delta}_2(\alpha_1,\alpha_2)\right\}$ imperfect collusion cannot be sustained. Consequently, the solution to (5) gives rise to two different reaction functions $\alpha_i(\alpha_j)$ for i and j = 1, 2 and $i \neq j$ for each possible value of δ. Therefore, the intersection of these reaction functions leads us to a solution with different values for the endogenous degree of collusion for firms. We note that even though firms might deviate from the output agreed, we implicitly assume that firms cannot deviate from the degree of collusion decided in the first stage as long as α_i and α_j are non-cooperatively decided in the first stage[16].

We assume for simplicity though that firms agree on a common value (such that $\alpha_i = \alpha_j = \alpha$) that we denote by α^* and that we can interpret as an upper

[16] We believe that several reasons justify this assumption. Firstly, assuming cooperation on the initial degree of collusion is a considerably more complex model as long as possible deviations on α_i and α_j in every period should also be considered markedly extending the set of strategies available to firms. Secondly, in a non-cooperative game where the type of collusion considered is merely tacit one could also think that the extent of collusion to be sustained should be decided also non-cooperatively. Relaxing this assumption has been left for future research.

bound on the degree of collusion to be sustained[17]. In order to obtain α^*, we firstly analyse how the functions defined in (2) change with α[18].

Lemma 2: *The function $\prod_2^c(\alpha)$ defined in (2) always increases with α in the interval $\alpha \in (0,1)$. Conversely, $\prod_1^c(\alpha)$ increases with α up to $\hat{\alpha} \in (0,1)$ where it reaches a maximum value.*

The most efficient firm always prefers more collusion whereas for the inefficient firm, if $\alpha > \hat{\alpha}$ the rule for output allocation in order to (imperfectly) collude implies that its production is highly reduced and its profits are smaller than when $\alpha = \hat{\alpha}$. On the contrary, if $\alpha < \hat{\alpha}$ inefficient firm's profits increases with α and, therefore, this firm would rather choose the highest possible (namely, sustainable) value of α over the interval $(0,\hat{\alpha})$. Consequently, a potential mutual agreement on α to which both firms adhere may be obtained since it follows directly from Lemma 2 that the inefficient firm's decision on α is binding in order to sustain collusion.

Proposition 2: *Let's consider the function $\tilde{\delta}_1(\alpha)$. Then, if $\delta \geq \tilde{\delta}_1(\hat{\alpha})$ the endogenous degree of collusion is such that $\hat{\alpha} = \alpha^*$. On the contrary, for a given $\overline{\delta} < \tilde{\delta}_1(\hat{\alpha})$, the endogenous degree of collusion α^* is the α that solves the equation $\overline{\delta} = \tilde{\delta}_1(\alpha)$.*

Proposition 2 states that the endogenous degree of collusion is the one that maximises inefficient firm's profits whenever firms are patient enough to sustain it as a SPNE of the repeated game. In this case, the efficient firm would rather sustain perfect collusion. However, for the inefficient firm among all the possible sustainable degrees of collusion, $\hat{\alpha}$ is the one where its profits are maximised. Therefore, $\hat{\alpha}$ is the endogenous degree of collusion. Otherwise, that is if $\hat{\alpha}$ is not sustainable, the endogenous degree of collusion is the largest sustainable one by the inefficient firm for a given δ. In other words, since collusion can only be sustained if both firms agree, the inefficient firm is the one that imposes her will. For each possible level of the discount factor, the inefficient firm is always willing to sustain a lower degree of collusion than the efficient firm, either

---

17 When we allow for different levels of imperfect collusion, results do not qualitatively change. We obtain that both firms have incentives to choose the minimal possible degree of cooperation to the extent that collusion does not collapse turning profits to the ones of the Cournot allocation. Roughly speaking, free riding incentives are alleviated by the fact that if a firm does not cooperate enough with the rival, collusion collapses and both firms are worse off. Further details are available at https://goo.gl/V87MbS or from the authors upon request.

18 We note that since $\tilde{\delta}_1(0) = 0$ and $\tilde{\delta}_1(\alpha)$ increases with α then α^* might always be sustainable if $\delta > 0$. Moreover, we assume also that both firms correctly anticipate that $\tilde{\delta}_1(\alpha) > \tilde{\delta}_2(\alpha)$ for all α.

$\hat{\alpha} < 1$ if firms are patient enough or $\alpha < \hat{\alpha}$ otherwise. Therefore, even though imperfect collusion will always arise in equilibrium, the degree of collusion is limited by the degree of cost asymmetry between both firms. It is also natural to analyse thus how α^* varies with δ and c.

Proposition 3: *The endogenous degree of collusion decreases with c. Also, if* $\delta < \tilde{\delta}_1(\hat{\alpha})$, *the endogenous degree of collusion increases with* δ.

Intuitively, as the cost asymmetry increases, the inefficient firm is less willing to cooperate since collusion would imply a larger switch of production to the most efficient firm. Turning back to the numerical example provided above, we can better illustrate the results of the present section. Assume for instance $c = 0.202$. Then, the degree of collusion that maximises profits of the inefficient firm is $\hat{\alpha} = 0.486$ (note that if $c = 0$, obviously $\hat{\alpha} = 1$). Therefore, if for example $\delta = 3$, since $\tilde{\delta}_1(0.486)\big|_{c=0.202} = 0.45 > 0.3$, $\hat{\alpha}$ cannot be sustained. As a consequence, one can obtain (Proposition 2) the endogenous degree of collusion by solving the equation $\tilde{\delta}_1(\alpha) = 0.3$. The solution is $\alpha = 0.32$ which, in this case, is the endogenous symmetric degree of collusion.

6. Extensions

Although some fundamental issues have been raised in the present paper, some potentially important questions still need to be addressed. In this Section, we present an alternative approach to managing the degree of cooperation, the presence of capacity constraints, and finally, we discuss the case in which firms make their strategic choices sequentially.

Considering other business practices that may possibly have anticompetitive effects and enable firms to coordinate price increases provide also a rich area for future research. For instance, Holt and Scheffman (1987) provide some interesting examples like the use of best-price policies or the public advance notification of list-price increases. In the same line, García-Díaz, González and Kujal (2009) show that, in the standard Bertrand-Edgeworth duopoly model, the use of list pricing might be a possible collusion facilitating device.

6.1. Discount factor and the degree of cooperation

We test here whether Proposition 1 hinges on the assumption made in Section 2 regarding the way to measure the intensity of competition. In particular, we assume here that the degree of cooperation is captured by the discount factor. We consider an industry like the one described in our benchmark model but for the case where $\alpha = 1$. We assume also that firms play an infinitely repeated game at dates $t = 1,...,\infty$ with a common discount factor $\delta \in [0,1)$ and restricting attention to the well-known grim *trigger strategies.* Each firm producing a collusive output corresponds to a SPNE if and only if the following condition is satisfied for each firm:

$$\Pi_i^c \geq \Pi_i^d (1-\delta) + \delta \Pi_i^* \tag{6}$$

where Π_i^d denotes the one period profit from deviation and Π_i^c the profits obtained by each firm at the perfect collusive equilibrium. There are many SPNE collusive output vectors that satisfy the system of inequalities in condition (6) above. As in Verboven (1997) and Escrihuela-Villar (2008), we select an equilibrium from this large set assuming that if δ exceeds a certain critical level, the set of SPNE vectors is not a binding constraint, and the distribution of output is the symmetric distribution of the output under perfect collusion. We also note that this critical level is the $\tilde{\delta}_i(\alpha)$ previously defined evaluated at α = 1. Then, a perfectly collusive outcome is a SPNE of the repeated game if $\delta \geq \max\left\{\tilde{\delta}_1(1), \tilde{\delta}_2(1)\right\}$ where, as shown in the proof of Proposition 1, $\tilde{\delta}_1(1) \geq \tilde{\delta}_2(2)$ is satisfied. On the contrary, if δ is below that critical level, then the set of SPNE vectors is a binding constraint, and the distribution of output is the solution to the equality constraint in (6). Unfortunately, the underlying system of equations cannot be further simplified. However, Rothschild (1999) proved that with quadratic cost functions the incentive to deviate is increasing in the deviant's inefficiency and that, therefore, collusion is feasible whenever the most inefficient firm adheres to the agreement. Consequently, we make the simplifying assumption that we can just focus on firm 1. Let us denote by q_2 the collusive output produced by firm 2. Then, it can be easily checked that the quantity produced by firm 1 in the collusive equilibrium also depends on δ. We denote it by $q_1(q_2, c, \delta)$ and it is given by

$$q_1(q_2,c,\delta) = \begin{cases} \frac{1}{2}\left(\frac{1-q_2}{2+c} + \sqrt{\frac{\delta(27-2c+4c^2(q_2-2)-15q_2)(3+2c-15q_2+4c^2q_2)}{(2+c)^2(15-4c^2)^2}}\right) & \text{if } \delta < \tilde{\delta}_1(1) \\ \frac{1-q_2}{4+2c} & \text{otherwise,} \end{cases} \tag{7}$$

Notice that when $\delta = 0$ the Cournot outcome holds, whereas $\delta \to \tilde{\delta}_1(1)$ the perfectly collusive equilibrium is reached. Hence, as δ varies from zero to $\tilde{\delta}_1(1)$ the degree of collusion increases. As the collusive profits of firm 1, that we denote by $\Pi_1^c(q_2, \delta, c)$, depend on δ it can be proved that for all $q_2 \in (0, q_2^*)$ the following is true.

Proposition 4: *If c is small enough* $\Pi_1^c(q_2,\delta,c) > \Pi_1^*$, *and* $\Pi_1^c(q_2,\delta,c)$ *increases with δ. Conversely, if c is large enough* $\Pi_1^c(q_2,\delta,c) > \Pi_1^*$ *is only true if δ is low enough.*

The intuition is as follows. In a collusive equilibrium, firms are willing to cut production compared to the non-collusive equilibrium. This output reduction favours the inefficient firm as long as this firm is not "too inefficient". Therefore, the firm is not punished to further cut its production in order to maximise joint profits. On the contrary, if firm 1 is markedly inefficient compared to firm 2, further joint profit maximisation might imply a detrimental output contraction of firm 1. Hence, Proposition 4 shows that Proposition 1 also carries over to the case where δ captures the degree of collusion.

6.2. Capacity constraints

Capacity constraints also play a key role in the analysis of tacit collusion. In our model demand is constant, so capacity constraints unambiguously affect collusion and its sustainability[19]. It is well known that the level of firm's capital stock determines the maximum level of production capacity, i.e. the capacity constraint. Capacity constraints affect collusion sustainability in at least two ways: they reduce the incentives to deviate as well as the severity of retaliation. Many studies on this issue have focused on symmetric situations where all firms have the same capacity (see for instance Abreu, 1986). It seems plausible to think, however, that in a model with cost asymmetries, firms' production capacities are also asymmetric. Let us assume that firms may bear a common maximum level of cost $\bar{C}$, which is already determined by the access to capital market. Notice that there is no reason to assume that firms have different conditions to access the capital market. However, differences in efficiency may come from labour organisation and other internal production issues. In our model, this is captured by the parameter c, which in turn determines $\rho_i(c)$. Hence, we assume that a given level of $\tilde{c}$, thus $\rho_i(\tilde{c})$, it determines a maximum capacity level $\tilde{q}_i$. Therefore, for any level of production q, it is hold $C_1(\rho_1(\tilde{c}),q) > C_2(\rho_2(\tilde{c}),q)$. As q increases both firms approach $\bar{C}$ at a different path because marginal costs are higher for the inefficient firm. As a result, for a given level $\tilde{c}$ it can be determined a maximum level of capacity $\tilde{q}_i$,

$$C_1(\rho_1(\tilde{c}),\tilde{q}_1) = C_2(\rho_2(\tilde{c}),\tilde{q}_2) = \bar{C},$$

where $\tilde{q}_2 > \tilde{q}_1$. Thus, our model can be also interpreted as one where maximum capacity is determined by the level of c. Some studies suggest that the introduction of asymmetric capacities makes indeed collusion more difficult to sustain when the aggregate capacity is limited (see for instance Davidson and Deneckere, 1990 or Compte, Jenny and Rey, 2002). The intuition is that punishing a firm with a low capacity puts an upper bound on the punishment that the other firms may

[19] In fact, when either demand is subject to cyclical fluctuations (see for instance Fabra 2006), or it exists demand uncertainty (see for instance Staiger and Wolak 1992), the effect of capacity constraints in collusion and its sustainability need a deep analysis that go beyond our study.

inflict. Hence the other firms have to suffer the punishment they impose on the low capacity firm. Consequently, a firm with a large capacity might be reluctant to participate in such a punishment. In our model, this effect is relaxed by α. Indeed, the larger the difference in capacity, the lower α is needed to imperfectly collude. Thus, in the case of retaliation, the punishment is less severe. Hence, it seems that the introduction of asymmetric capacity constraints in our model would work in the same direction as cost asymmetries hurting also collusion sustainability. Presumably, then, a reduction in the degree of collusion could alleviate the effect of capacity constraints on collusion sustainability.

6.3. Simultaneous vs. sequential strategic choice

As Mouraviev and Rey (2011) show in a fairly general framework, such leadership is not effective in case of quantity competition since, following an aggressive deviation by the leader, the follower would rather limit its own output, making it more difficult to punish the deviation[20]. They also show that quantity leadership along the equilibrium path does not allow the firms to achieve a Pareto improvement. The intuition is that since quantities are strategic substitutes and firms should lower their outputs to increase their profits, if the leader reduces its output then the follower should increase its own quantity in order to sustain collusion. In fact, this can also be empirically observed. In their interesting survey of EC cartel decisions, Mouraviev and Rey show that while (production or distribution) capacity appears as the key strategic variable in some cases, leadership does not feature in any of these observations.

7. Concluding Comments

We have developed a theoretical framework to study how firms' cost asymmetry affects the possibility that a collusive agreement can be sustained over time. Contrary to the usual assumption made in many oligopoly models, we introduced that an imperfectly (tacit) collusive agreement can be sustained in the event that firms also care about the other firms' profits but just to some extent. Our main contribution is twofold. We show that even though cost asymmetry hinders collusion, imperfect collusion can always be sustained regardless of the cost asymmetry. Secondly, we also analyse the endogenous degree of collusion by assuming that firms agree on a common degree of cooperation. We show that there is a limit to the degree of sustainable collusion that depends on the most inefficient firm of the industry. Another interpretation of our results is also that cost asymmetry is not necessarily a restraint for collusion as long as firms are able to sustain the maximum degree of collusion contingent on their discount factor. In this sense, some evidence suggesting collusive agreements among firms of significantly different costs of production represents an empirical justification for our findings.

---

[20] Conversely, Deneckere and Kovenock (1992) show that when firms have capacity constraints, the outcome under price leadership is more collusive than the outcome under simultaneous price-setting.

The framework we have worked with is only a particular approach to a more general issue. To analyse real-world cartels, additional research is required, and for instance, a wider range of demand functions should also be considered. It would also be interesting to test if our results are robust to using an optimal punishment like the "stick-and-carrot strategies" proposed by Abreu (1986, 1988). We believe that those are subjects for future research.

Appendix

Proof of Lemma 1: From (2) when i = 1 and α = 0, Cournot profits are at $\Pi_1^* = \left[(3-2c)^2(2+c)\right]/(15-4c^2)^2$. Let us denote by $\Pi_1^c - \Pi_1^* \equiv f(c,\alpha)$ the function that captures the difference between collusive and Cournot profits. Obviously, it can be verified that $f(c, 0) = 0$.

Then, $f(c,1) = -\left[(3-2c)^2(2+c)\right]/(15-4c^2)^2 + (1-c)(12-4c^2) < 0$ only if c is large enough; it implies that the most inefficient firm is only worse off in the perfectly collusive agreement compared to the Cournot equilibrium if this firm is relatively inefficient enough compared to the efficient ones. Therefore, since Π_1^* does not depend on α and $\Pi_1^c(\alpha)$ increases with α when α is small enough (see the proof of Lemma 2), there exists $\tilde{\alpha}$ small enough such that if $\alpha < \tilde{\alpha}$, which yields $f(c,\alpha) > 0$. It implies that the inefficient firm obtains larger collusive profits than in the Cournot equilibrium. Since the upper bound on α such that $\tilde{\delta}_1(\alpha) < 1$ coincides with the one needed in order for $\Pi_1^c(\alpha) > \Pi_1^*$ to be true, then whenever $\tilde{\delta}_1(\alpha)$ increases with c, α decreases with c. Details have been omitted from the paper to save space, but are available in an additional Appendix[21]. □

Proof of Proposition 1: Since $\tilde{S}_1(\alpha) = \left[\Pi_1^d(\alpha) - \Pi_1^c(\alpha)\right]/[\Pi_1^d(\alpha) - \Pi_1^*]$, $\tilde{S}_1(\alpha) \in (0,1)$ if $\Pi_1^c(\alpha) > \Pi_1^*$ and $\Pi_1^d(\alpha) > \Pi_1^c(\alpha)$. The condition under which the first inequality holds is proved in Lemma 1, so in order to prove Proposition 1 it suffices to check that

$$\Pi_1^d(\alpha) - \Pi_1^c(\alpha) = \frac{\left[\alpha^2(\alpha-3-2c)^2\right]}{\left[4(2+c)\left(4c^2+(\alpha-3)(5+\alpha)\right)^2\right]} > 0.$$

We also prove that collusion is more easily sustained for the most efficient firm: $\tilde{S}_1(\alpha) \geq \tilde{S}_2(\alpha)$ for all $\alpha \in (0,1]$.. Although the expressions for these cutoffs cannot be easily simplified to be included in the paper, they can still be obtained from the values of the profit functions evaluated under Cournot competition,

---

21 We used the program Wolfram Mathematica 7.0. Further details for this and subsequent proofs are available at https://goo.gl/K4muLu or from the authors upon request.

imperfect collusion and an eventual deviation from it. They are reported in the main text for i = 1, 2. Obviously, both cutoffs coincide in the symmetric case $c=0$. Then, $\tilde{S}_1(\alpha)=\tilde{S}_2(\alpha)=25\alpha/40+9\alpha$.. Since $\tilde{S}_1$ and $\tilde{S}_2$ also depends on c, we can denote them by $\tilde{S}_i(\alpha,c)$. It can be easily checked using a mathematical software that the equation $\tilde{S}_1(\alpha,c)=\tilde{S}_2(\alpha,c)$ has no real root for $c\in(0,1)$. then, it is enough to check that for a given c, the inequality to be proved holds. We know from Lemma 1 that when c is large enough $\Pi_1^c<\Pi_1^*$ and therefore $\tilde{S}_1(\alpha,c)>1>\tilde{S}_2(\alpha,c)$. It suffices to check that in this case $1>\tilde{S}_2(\alpha,c)$. This is true as long as

$$\Pi_2^d(\alpha)-\Pi_2^c(\alpha)=\left[\alpha^2(2c+\alpha-3)^2\right]/\left[4(2-c)\left(4c^2+(\alpha-3)(5+\alpha)\right)^2\right]>0,$$

and $\Pi_2^c(\alpha)>\Pi_2^*$. The last inequality holds since $\Pi_2^c(0)=\Pi_2^*$, and Lemma 2 proves that $d\Pi_2^c(\alpha)/d\alpha>0$ while Π_2^* does not depend on α. Regarding the second part of the proposition, it is enough to show that

$$\partial\tilde{\delta}_1(\alpha,c)/\partial\alpha\Big|_{\alpha=0}=\frac{\left[\left(3-2c^2\right)\left(15-4c^2\right)\right]}{\left[4\left(c+2c^2-6\right)\left(4c(1+c)-9\right)\right]}>0;$$

i.e; the cutoff of the discount factor increases with α when α is very small. Then, an increase in α always increases $\tilde{\delta}_1(\alpha,c)$. Indeed, with mathematical software, we can check that the equation $\partial\tilde{\delta}_1(\alpha,c)/\partial\alpha=0$ has no real root for $\alpha\in(0,1)$; in other words, that the cutoff changes (increases) monotonically with α. Regarding the effect of c, we apply the same procedure. First,

$$\partial\tilde{\delta}_1(\alpha,c)/\partial c\Big|_{c=0}=\frac{\left[100c(5+\alpha)(35+4\alpha)^2\right]}{\left[3(3-\alpha)(40+9\alpha)^2\right]}>0,$$

and the equation $\dfrac{\partial\tilde{\delta}_1(\alpha,c)}{\partial c}=0$ has no real root for $c\in(0,1)$. □

Proof of Lemma 2: Let's consider the firms' collusive profits that we denoted by $\Pi_1^c(\alpha)$ and $\Pi_1^c(\alpha)$. It can be easily proved with the software Mathematica® that there is no real root for the equation $\dfrac{\partial\Pi_2^c(\alpha)}{\partial\alpha}=0$ if $\alpha\in(0,1)$, which implies that profits of the efficient firm change monotonically with α. Then, since profits for the efficient firm when α = 1 are larger than those when $\alpha=0$, $\Pi_2^c(1)-\Pi_2^c(0)=\left[3\left(4c^3-3-15c\right)\right]/\left[4\left(15-4c^2\right)^2\left(c^2-3\right)\right]>0$, the result holds. Besides, regarding the inefficient firm, $\dfrac{\partial\Pi_1^c(\alpha)}{\partial\alpha}=0$ has only one root in α and as long as profits increase with α for small enough α:

$$\partial\Pi_1^c(\alpha)/\partial\alpha\Big|_{\alpha=0}=\frac{\left[30c+4c^2-8c^3-27\right]}{\left(4c^2-15\right)^3}>0,$$

Inefficient firm's profits necessarily decreases with α when α is larger than the root mentioned before. □

Proof of Proposition 2: Let's define $\hat{\delta}\overset{\text{def}}{=}\tilde{\delta}_1(\hat{\alpha},c)$ as the minimum value of the discount factor such that $\hat{\alpha}$ can be sustained as a SPNE of the repeated game. We need thus to analyse to different cases. On one hand, $\delta>\hat{\delta}$. It is obvious from Lemma 2 that in this case inefficient firm's best response to any α_2 is always $\alpha_1=\hat{\alpha}$ if $\alpha_2\geq\hat{\alpha}$ because both a smaller and a larger degree of collusion than $\hat{\alpha}$ is profit-dominated by $\hat{\alpha}$ as long as $\hat{\alpha}$ is sustainable. On the other hand, $\alpha_2\leq\hat{\alpha}$ cannot be a Nash equilibrium since $\alpha_2<\alpha_1<\hat{\alpha}$ is dominated by $\alpha_2=\alpha_1<\hat{\alpha}$ (the efficient firm could push to sustain more collusion) and $\alpha_1<\alpha_2<\hat{\alpha}$ is also dominated by $\alpha_2=\alpha_1<\hat{\alpha}$ (the inefficient firm could still increase its collusive effort). Therefore, the only possible Nash equilibrium in the first case $(\delta>\hat{\delta})$ is thus $\alpha_1=\hat{\alpha}$ and $\alpha_2\geq\hat{\alpha}$. Regarding the second case where $\delta<\hat{\delta}$, let's denote by $\bar{\alpha}$ the maximum degree of collusion that can be sustained as SPNE for each possible $\bar{\delta}<\widehat{\delta}$, namely $\bar{\alpha}$ is the solution to the equation $\tilde{\delta}_1(\alpha,c)=\bar{\delta}$ for α. We note that we implicitly assume that firms may correctly anticipate that $\tilde{\delta}_1(\alpha,c)$ increases with α. The actions $\alpha_1>\alpha_2>\bar{\alpha}$ and $\alpha_2>\alpha_1>\bar{\alpha}$ do not constitute a Nash equilibrium. This is true since in both cases firms would obtain the non-collusive equilibrium profits (Π_i^*) as long as collusion would not be sustainable while each firm could optimally deviate by choosing $\bar{\alpha}$ where $\Pi_i^c(\bar{\alpha})>\Pi_i^*$. Moreover, $\alpha_1<\alpha_2\leq\bar{\alpha}$ and $\alpha_2<\alpha_1\leq\bar{\alpha}$ do not constitute an equilibrium either since inefficient and efficient firms can optimally deviate by choosing $\alpha_1=\alpha_2\leq\bar{\alpha}$, respectively. Analogously, $\alpha_1<\bar{\alpha}\leq\alpha_2$ and $2<\bar{\alpha}\leq\alpha_1$ do not constitute an equilibrium either since $\alpha_1=\bar{\alpha}$ and $\alpha_2=\bar{\alpha}$ represent profit enhancing unilateral deviations, respectively. Finally, we need to check that there is no optimal deviation from $\alpha_1=\alpha_2=\bar{\alpha}$. . Any lower α would decrease firms' profits (Lemma 2) and any unilateral deviation on α by one firm increasing α would not affect firm's profits. As a consequence, $\alpha_i=\bar{\alpha}\leq\alpha_j$ for $i,j=1,2$ and $i\neq j$ constitute a Nash equilibrium. □

Proof of Proposition 3: For the first part of the result, we have to check that $\partial\hat{\alpha}/\partial c<0$. As $\Pi_1^c(\alpha)$ does not depend on δ we can obtain $\hat{\alpha}$ like the argmax $\Pi_1^c(\alpha)$. Although the solution cannot be explicitly obtained, the equation $\partial\Pi_1^c(\alpha)/\partial\alpha=0$ can be solved for *c* giving

$$c(\alpha)=\frac{6+12\alpha^2-\alpha^3+\alpha\left(-13+\sqrt{(\alpha-3)^2\left(4+12\alpha+21\alpha^2-2\alpha^3+\alpha^4\right)/\alpha^2}\right)}{16\alpha}.$$

Then, $\hat{\alpha} = c^{-1}(c)$. Hence, one can easily check that $\partial c(\alpha)/\partial\alpha = 0$. Using Lagrange's notation, the derivative of the inverse function is given by $\partial\hat{\alpha}/\partial c = 1/(\partial c(\alpha)/\partial\alpha|_{c^{-1}(c)}) < 0$, and thus it is also decreasing its inverse function. Regarding the second part of the proposition, from the proof of Lemma 1 and Proposition 2 we know that $\partial\tilde{\delta}_1(\alpha,c)/\partial\alpha < 0$ and $\partial\tilde{\delta}_2(\alpha,c)/\partial\alpha > 0$. Finally, the result follows directly because the endogenous degree of collusion is obtained $\alpha = \tilde{\delta}_1^{-1}(\alpha,c)$. □

Proof of Proposition 4: From (7) profits for firm 1 can be easily obtained,

$$\Pi_1^c(q_2,\delta,c) = \frac{225 - 120c^2 + 16c^4 - 81\delta - 48c\delta + 28c^2\delta + 16c^3\delta + 2\left(15-4c^2\right)^2(\delta-1)q_2 - \left(15-4c^2\right)^2(\delta-1)q_2^2}{4(2+c)\left(15-4c^2\right)^2}.$$

Non-cooperative Nash equilibrium profits are given by $\Pi^* = \left[(3-2c)^2(2+c)\right]/\left(15-4c^2\right)^2$. Then, we simply have to check that for all $q_2 \leq q_2^* \overset{\text{def}}{=} (3+2c)/\left(15-4c^2\right)$ if $c < \frac{1}{4}\sqrt{\left(1-12q_2+60q_2^2\right)/q_2^2} - \frac{1}{4q_2}$,3 then $\frac{\partial\Pi_1^c(q_2,\delta,c)}{\partial\delta} > 0$ and $\Pi_1^c\left(q_2,\delta,c\right) > \Pi^*$. Conversely, if $c > \frac{1}{4}\sqrt{\left(1-12q_2+60q_2^2\right)/q_2^2} - \frac{1}{4q_2}$, we can easily check that $\Pi_1^c\left(q_2,1,c\right) < \Pi^*$ while $\Pi_1^c\left(q_2,0,c\right) = \Pi^*$ since, by definition, $\delta = 0$ corresponds to the non-cooperative Nash equilibrium when $q_2^* \overset{\text{def}}{=} \frac{3+2c}{15-4c^2}$. Then, since $\left.\frac{\partial\Pi_1^c(q_2,\delta,c)}{\partial\delta}\right|_{\delta} > 0$ by continuity, there exist $\delta \in (0,1)$ such that $\Pi_1^c\left(q_2,\delta,c\right) > \Pi^*$. □

References

Abreu, D.J. (1986). "Extremal Equilibria of Oligopolistic Supergames", *Journal of Economic Theory.* Vol. 39 (1); 191-225.

Abreu, D.J. (1988). "On the Theory of Infinitely Repeated Games with Discounting", *Econometrica.* Vol. 56; 383-396.

Bain, J. S. (1948). "Output quotas in imperfect cartels", *The Quarterly Journal of Economics*. Vol. 62; 617-622.

Bos, I. and Harrington, J. E. (2010). "Endogenous cartel formation with heterogeneous firms", *The RAND Journal of Economics*. Vol. 41 (1); 92-117.

Charness, G. and Rabin, M. (2002). "Understanding Social Preferences with Simple Tests", *Quarterly Journal of Economics*. Vol. 17 (3); 817-69.

Ciarreta, A. and Gutiérrez-Hita, C. (2012). "Collusive behaviour under costs asymmetries when firms compete in supply functions", *Journal of Economics.* Vol. 106 (3); 195-219.

Compte, O., Jenny, F. and Rey, P. (2002). "Capacity constraints, mergers and collusion", *European Economic Review*. Vol. 46 (1), 1-29.

Contreras, R.; Georgantzis, N. and Ginés, M. (2008). "Collusion in a Bertrand duopoly model with decreasing returns and product differentiation", *Estudios de Economía*. Vol. 35 (1); 19-31.

Cyert, R. M. and DeGroot, M. H. (1973). "An Analysis of Cooperation and Learning in a Duopoly Context", *American Economic Review*. Vol. 63 (1); 24-37.

Davidson, C. and Deneckere, R. J. (1990). "Excess capacity and collusion", *International Economic Review*. Vol. 31; 521-541.

Deneckere, R. J. and Kovenock, D. (1992). "Price leadership", *Review of Economic Studies*. Vol. 59 (1); 143-162.

Díaz, A. G., González, R. H. and Kujal, P. (2009). "List pricing and discounting in a Bertrand-Edgeworth duopoly", *International Journal of Industrial Organization*. Vol. 27 (6); 719-727.

Donsimoni, M. P.; Economides, N. S. and Polemarchakis, H. M. (1986). "Stable Cartels", *International Economic Review*. Vol. 27 (2); 317-327.

Eckbo, P. L. (1976). "The Future of World Oil", Cambridge: Ballinger Publishing.

Escrihuela-Villar, M. (2008). "Partial coordination and mergers among quantity-setting firms", *International Journal of Industrial Organization*. Vol. 26 (3); 803-810.

Escrihuela-Villar, M. (2015). "A note on the equivalence of the conjectural variations solution and the coefficient of cooperation", *The B.E. Journal of Theoretical Economics*. Vol. 15 (2); 473-480.

Fabra, N. (2006). "Collusion with capacity constraints over the business cycle", *International Journal of Industrial Organization*. Vol. 24; 69-81.

Fehr, E. and Schmidt, K. M. (1999). "A Theory of Fairness, Competition, and Cooperation", *Quarterly Journal of Economics*, Vol. 114 (3); 817-68.

Fershtman, C. and Judd, K. L. (1987). "Equilibrium incentives in oligopoly", *American Economic Review*. Vol. 77 (5); 827-840.

Friedman, J. W. (1971). "A non-cooperative equilibrium for supergames", *Review of Economic Studies*, Vol. 38 (1); 1-12.

García-Díaz, A.; Hernán-González, R. and Kujal, P. (2009). "List pricing and discounting in a Bertrand-Edgeworth duopoly", *International Journal of Industrial Organization.* Vol. 27; 719-727.

Griffin, J. M. (1989). "Previous cartel experience: Any lessons for OPEC?". *In Economics in theory and practice: An eclectic approach,* pp. 179-206, Springer, The Netherlands.

Harrington, J. E. (1991). "The determination of price and output quotas in a heterogeneous cartel", *International Economic Review*. Vol. 32 (4); 767-792.

Holt, C. A. and Scheffman, D. T. (1987). "Facilitating practices: the effects of advance notice and best-price policies", *RAND Journal of Economics*. Vol. 18 (2); 187-197.

Kreps, D. M. and Scheinkman, J. A. (1983). "Quantity precommitment and Bertrand competition yield Cournot outcomes", *Bell Journal of Economics*. Vol. 14 (2); 326-337.

Levenstein, M. C. and Suslow, V. Y. (2006). "What determines cartel success?", *Journal of Economic Literature*. Vol. 44 (1); 43-95.

Mason, C. F.; Phillips, O. R. and Nowell, C. (1992). "Duopoly behaviour in asymmetric markets: An experimental evaluation", *Review of Economics and Statistics.* Vol. 74 (4); 662-670.

Matsumura, T.; Matsushima, N. and Cato, S. (2013). "Competitiveness and R&D Competition Revisited", *Economic Modelling*. Vol. 31; 541-547.

Mendi, P.; Moner-Colonques, R. and Sempere-Monerris, J. J. (2011). "Vertical integration, collusion, and tariffs", *SERIEs*. Vol. 2 (3); 359-378.

Miklos-Thal, J. (2011). "Optimal collusion under cost asymmetry", *Economic Theory*. Vol. 46 (1); 99-125.

Mouraviev, I. and Rey, P. (2011). "Collusion and leadership", *International Journal of Industrial Organization*. Vol. 29 (6); 705-717.

Osborne, M. J. and Pitchik, C. (1983). "Profit-sharing in a collusive industry", *European Economic Review*. Vol. 22 (1); 59-74.

Patinkin, D. (1947). "Multiple-plant firms, cartels, and imperfect competition", *Quarterly Journal of Economics*. Vol. 61; 173-205.

Rothschild, R. (1999). "Cartel stability when costs are heterogeneous", *International Journal of Industrial Organization*. Vol. 17 (5); 717-734.

Schmalensee, R. (1987). "Competitive advantage and collusive optima", *International Journal of Industrial Organization*. Vol. 5 (4); 351-367.

Staiger, R. W. and Wolak, F. A. (1992). "Collusive pricing with capacity constraints in the presence of demand uncertainty", *RAND Journal of Economics*. Vol. 23 (2); 203-220.

Symeonidis, G. (2008). "Downstream Competition, Bargaining, and Welfare", *Journal of Economics and Management Strategy*. Vol. 17 (1); 247-70.

Vasconcelos, H. (2005). "Tacit collusion, cost asymmetries, and mergers", *RAND Journal of Economics*. Vol. 36; 39-62.

Vega-Redondo, F. (1997). "The evolution of Walrasian behaviour", *Econometrica*. Vol. 65; 375-384.

Verboven, F. (1997). "Collusive behaviour with heterogeneous firms", *Journal of Economic Behaviour & Organization.* Vol. 33 (1); 121-136.

Wirl, F. (2015). "Output adjusting cartels facing dynamic, convex demand under uncertainty: The case of OPEC", *Economic Modelling*. Vol. 44; 307-316.