

# A note on cartel stability and endogenous sequencing with tacit collusion.

Marc Escrihuela-Villar\*

Universitat de les Illes Balears

October, 2008

## Abstract

We use the concept of cartel stability defined by d'Aspremont et al. (1983) to obtain that the sequence of play between the cartel and the fringe affects cartel stability in a quantity-competition setting where firms tacitly collude. We also prove that an endogenous sequence of play between a cartel and a fringe depends on the discount factor. If the discount factor is large enough, the cartel and the fringe simultaneously choose quantities since the stable cartel may contain more firms under simultaneous play than under cartel leadership. This is due to the fact that under simultaneous play cartel firms have incentives to participate in the cartel because otherwise no collusion is possible.

*JEL Classification:* L11; L41; D43.

*Keywords:* Collusion; Cartel stability; Stackelberg leadership

---

\*Mailing address: Departamento Economía Aplicada. Edificio Jovellanos Ctra. Valldemossa km.7.5. 07122 Palma de Mallorca Baleares - España. Tel. +34 971173242. Fax: +34 971172389. Email: marc.escrihuela@uib.es

# 1 Introduction

The concept of a stable cartel defined in d’Aspremont et al. (1983) is one of the first contributions in the analysis of cartel formation in oligopoly markets. A cartel is stable if no firm inside the cartel finds it desirable to exit and no firm outside the cartel finds it desirable to enter. More recent work follows this path and explores conditions under which the stable cartels exist. For instance, Donsimoni et al. (1986) and Shaffer (1995) characterize the set of stable cartels in a model of price and quantity competition respectively. In these papers it has been shown that cartels containing approximately just over half the firms in the industry are stable. In such models, and to the best of our knowledge in all the cartel and fringe literature (for instance Selten (1973), Donsimoni (1985), Thoron (1998) or Lofaro (1999)), it is assumed without further analysis that the cartel behaves as a leader with respect to the fringe. By their very nature, these models may be viewed as models of binding collusion since they also assume the existence of an enforcement mechanism for collusive behavior within the cartel such that once a firm decides to join the cartel, the possibility of cheating on the agreement is ruled out. Another strand in the literature on collusion, the supergame-theoretic approach also termed tacit or implicit collusion, has focused on the problem of enforcement of collusive behavior. This approach analyzes the symmetric subgame perfect Nash equilibrium —henceforth, SPNE— that maximizes industry profits (see for example the seminal paper by Friedman (1971), Davidson and Deneckere (1984), Martin (1993) or Rothschild (1999)). However, this approach focuses on firms’ “incentive constraints” and leaves out firms’ “participation constraints”,<sup>1</sup> and therefore the suitability of the leadership assumption is not considered either. In this paper, we address the question of whether cartels have an incentive to behave as a Stackelberg leader by analyzing the cartel-stability problem in an infinitely repeated setting using SPNE as solution concept. We develop a multi-period oligopoly model with homogeneous, quantity-setting firms, a subset of which are assumed to collude

---

<sup>1</sup>“Participation constraints” determine the incentives that firms have to join a cartel. However, “incentive constraints” determine the incentives of cartel firms to deviate from the collusive agreement.

while the remaining (fringe) firms choose their output levels noncooperatively.<sup>2</sup> Then, using the concept proposed by d'Aspremont et al. (1983) for cartel stability, we endogenize cartel formation by analyzing the number of firms that are willing to participate in a cartel when firms have the possibility to tacitly collude. Subsequently, we examine and discuss the equilibrium sequence of moves between the cartel and the fringe.

We show that when the cartel does not behave as a leader, cartel firms may be willing to participate in the smallest cartel among those which can be sustained as a SPNE since although firms might have incentives to exit the cartel, the threat of no collusion at equilibrium induces them to collude. On the other hand, when the cartel is allowed to lead, the previous mentioned results on cartel stability with binding collusion extend to the case of tacit collusion. This leads us to obtain that when the discount factor is large enough, the stable cartel may contain more firms under simultaneous play than under cartel leadership. Consequently, although for a given cartel size cartel firms obtain larger profits with the leadership, when the cartel is allowed to impose its most preferred production timing, the sequence of play between a cartel and a fringe depends on the discount factor. If the discount factor is large enough, the cartel and the fringe simultaneously choose quantities. The intuition is that firms may prefer to stay outside the leader cartel and free-ride on the high price induced by collusion. Conversely, under simultaneous play the incentives to free-ride are reduced since by not participating in the cartel, firms risk the complete collapse of collusion. The main economic implication is that the antitrust authorities may be extremely wary to take the quantity leadership as a symptom of collusive behavior. Conversely, when tacit collusion is easily sustained, cartel firms may have strong incentives not to behave as a Stackelberg leader.

The rest of the paper is structured as follows. In section 2, we present the model and we study cartel stability. In a subsection, we study the production timing between the cartel and the fringe. We conclude in section 3. All proofs are grouped together in the appendix.

---

<sup>2</sup>The assumption of a cartel involving a subset of firms is based on the fact that some of the best known examples of cartels involve only a part of the industry. Some significant cases are the citric acid, the carbonless paper or the North Atlantic shipping industries (see Levenstein and Suslow (2006)).

## 2 The model and results

We consider an industry with  $N > 2$  firms, indexed by  $i = 1, \dots, N$ . Each firm produces a quantity of a homogeneous product with a linear cost function  $c(q_i) = cq_i$ , where  $q_i$  is the output produced by firm  $i$ . The industry inverse demand is given by the piecewise linear function  $p(Q) = \max(0, a - bQ)$  where  $Q = \sum_{i=1}^N q_i$  is the industry output,  $p$  is the output price and  $a > 0$  with  $a > c$ . We assume that  $K \in [2, N]$  firms, indexed by  $k = 1, \dots, K$  — henceforth, cartel firms — behave cooperatively so as to maximize their joint profits. The remaining  $(N - K)$  firms constitute the fringe and choose their output in a non-cooperative way. We assume that only one cartel is formed and that firms compete repeatedly over an infinite horizon with complete information (i.e. each of the firms either fringe or cartel observes the whole history of actions) and discount the future using a discount factor  $\delta \in (0, 1)$ . Time is discrete and dates are denoted by  $t = 1, 2, \dots$ . In this framework, a pure strategy for firm  $k$  is an infinite sequence of functions  $\{S_k^t\}_{t=1}^\infty$  with  $S_k^t : \sum^{t-1} \rightarrow \mathcal{Q}$  where  $\sum^{t-1}$  is the set of all possible histories of actions (output choices) of all cartel firms up to  $t - 1$ , with typical element  $\sigma_j^\tau$ ,  $j = 1, \dots, K$ ,  $\tau = 1, \dots, t - 1$ , and  $\mathcal{Q}$  is the set of output choices available to each cartel firm. Following Friedman (1971), we restrict our attention to the case where each cartel firm is only allowed to follow grim trigger strategies. In words, these strategies are such that cartel firms adhere to the collusive agreement until there is a defection, in which case they revert forever to the static  $N$ -firm Nash equilibrium. Since firms are symmetric, each cartel firm produces the same amount of output that we denote by  $q$ . The output corresponding to noncooperative behavior is denoted by  $q_n$ . Since we restrict attention to trigger strategies,  $\{S_k^t\}_{t=1}^\infty$  can be specified as follows. At  $t = 1$ ,  $S_k^1 = q$ , while at  $t = 2, 3, \dots$

$$S_k^t(\sigma_j^\tau) = \begin{cases} q & \text{if } \sigma_j^\tau = q \text{ for all } j = 1, \dots, K \text{ and } \tau = 1, \dots, t - 1 \\ q_n & \text{otherwise} \end{cases} \quad (1)$$

Regarding fringe firms, their optimal response consists of maximizing their current period's payoff. We denote the output produced by each fringe firm by  $q_f$ . We denote by  $\Pi^c(N, K)$  and  $\Pi^f(N, K)$  the profit function of a cartel firm and that of a fringe firm respectively. As shown by Friedman (1971), cartel firms colluding in each period can be

sustained as a SPNE of the repeated game with the strategy profile (1) if and only if for given values of  $N, K$  and  $\delta$ , the following condition is satisfied

$$\frac{\Pi^c(N, K)}{1 - \delta} \geq \Pi^d(N, K) + \frac{\delta \Pi(N)}{1 - \delta} \quad (2)$$

where  $\Pi^d(N, K)$  denotes the profits attained by an optimal deviation from the collusive output, and  $\Pi(N)$  denotes the Nash equilibrium profits. If  $\delta$  exceeds a certain critical level, (2) is not a binding constraint. For given values of  $N$  and  $K$ , we denote by  $\delta_K$  the minimum  $\delta$  required for the condition (2) to be satisfied.

**Definition 1**  $\delta_K$  is said to be the minimum discount factor required for the cartel of  $K$  firms to be sustainable as a SPNE. Then, a cartel of  $K$  firms is said to be sustainable if  $\delta \geq \delta_K$  and  $\delta_K \in (0, 1)$ .

## 2.1 Cartel stability

In this subsection we endogeneize cartel formation. We assume that firms can coordinate in the different outcomes by showing their willingness to participate in a collusive agreement. This pre-communication play is modelled as a stage prior to market competition. Those decisions will not affect the payoff of firms, but they will only be used as a coordination device: if  $K$  firms decide to participate in a cartel agreement, a cartel of size  $K$  can only be observed in the repeated game if the discount factor allows a cartel of  $K$  firms to be active ( $\delta \geq \delta_K$ ). Hence, the discounted payoffs attained by cartel and fringe firms respectively at  $t = 0$  are:

$$\begin{aligned} \Pi_0^c(N, K) &= \begin{cases} \frac{1}{1-\delta} \Pi^c(N, K) & \text{if } \delta \geq \delta_K \\ \frac{1}{1-\delta} \Pi(N) & \text{otherwise} \end{cases} \\ \Pi_0^f(N, K) &= \begin{cases} \frac{1}{1-\delta} \Pi^f(N, K) & \text{if } \delta \geq \delta_K \\ \frac{1}{1-\delta} \Pi(N) & \text{otherwise.} \end{cases} \end{aligned} \quad (3)$$

In other words, we add to the game an initial stage in which firms simultaneously choose whether or not to join the (unique) cartel and subsequently all firms produce for an

infinite number of periods.<sup>3</sup> This repeated game setting exhibits multiple SPNE collusive agreements. Therefore, to select among those equilibria, we adopt the particular criterion of restricting attention to the strict Nash equilibria. At this initial stage, a strict Nash equilibrium is defined as follows: (i) a number  $0 < K \leq N$  of firms join the cartel if and only if

$$\Pi_0^c(N, K) > \Pi_0^f(N, K - 1). \quad (4)$$

and

$$\Pi_0^f(N, K) > \Pi_0^c(N, K + 1). \quad (5)$$

We note that these conditions are equivalent to the standard stability concept where no individual move is desirable and where firms hypothesize that no other firm will change its strategy concerning its membership in the cartel.<sup>4</sup> Apart from the degenerate case of  $K = 1$ , a cartel can be defined as internally stable if it is not profitable for a cartel member to defect to the fringe (condition (4)). Likewise, apart from the degenerate case of  $K = N$ , a cartel is said to be externally stable if it is not profitable for a fringe firm to join the cartel (condition (5)).

**Definition 2** *A cartel is said to be stable if it is both internally and externally stable.*

We analyze two different variations of the model described above: after the initial participation stage in which firms choose whether or not to collude, cartel and fringe firms produce for an infinite number of periods (i) simultaneously choosing quantities (ii) with cartel firms being a Stackelberg leader with respect to the fringe in each period. We are now in the position to obtain the strict Nash equilibrium (namely, the number of firms in the stable cartel) in each of the two different production timings. We focus on the equilibria where a number  $0 < K \leq N$  of firms join the cartel.<sup>5</sup>

---

<sup>3</sup>It should be noted that like previous studies, we do not solve which  $K$  of  $N$  ex ante identical firms will precommit to collude.

<sup>4</sup>Thoron (1998) proves the correspondence between this stability concept and the Nash equilibria

<sup>5</sup>We focus on the *strict* Nash equilibria since the analysis of the Nash equilibria would lead us to uninteresting results. For instance, when for each  $K$ ,  $\delta < \delta_K$ , then firms obtain the same profits by joining the cartel than by staying at the fringe. Therefore, joining the cartel is always a Nash equilibrium. We note also that the equilibria in which all firms decide not to join the cartel are not considered.

**Proposition 1** *When cartel and fringe firms simultaneously choose quantities the smallest cartel among those that are sustainable is stable.*

The intuition behind Proposition 1 is as follows. As shown in the appendix,  $\Pi^c(N, K) < \Pi^f(N, K - 1)$  and thus, internal stability is not satisfied in the one-shot game. We note that this is related to a well-known result in the literature that, in a Cournot setting, although mergers (or any other collusive agreement) increase price, they are (generally) not profitable because non-participating firms react by increasing their production.<sup>6</sup> Then, it can be checked from (3) and (4) that if  $\delta_K \leq \delta < 1$  firms inside the cartel of  $K$  firms find it desirable to exit. On the other hand, if  $\delta_K < 1$ , it is strictly decreasing in  $K$ .<sup>7</sup> Then, if  $\delta_K \leq \delta < \delta_{K-1}$  a cartel of  $K$  or more firms is sustainable but a cartel of more than  $K$  firms is not stable because internal stability does not hold. However, in this case a cartel of  $K$  firms is stable because by leaving the cartel a firm would obtain lower profits (namely, the Nash equilibrium profits) since the cartel of  $K - 1$  firms is not sustainable. In other words, only the smallest sustainable cartel is stable because although firms have incentives to exit the cartel, they collude when the sustainability of collusion depends on their participation.

**Proposition 2** *When cartel firms behave as a Stackelberg leader with respect to the fringe and  $N \geq 4$ , a cartel of  $K$  firms is stable whenever  $\delta \geq \delta_K$  and  $K \in [f(N), f(N) + 1]$  where  $f(N) = \frac{1}{4}(1 + 3N - \sqrt{(N - 2)N - 7})$ . When  $N < 4$ , the unique stable cartel is joint monopoly whenever  $\delta \geq \delta_N$ . Otherwise no cartel is stable.*

Proposition 2 indicates that a stable leader cartel contains just over half the firms in the industry. This naturally extends previous results on cartel stability (see for instance Shaffer (1995)) to the case of implicit collusion. In other words, when a cartel of  $K$  firms is sustainable as a SPNE of the repeated game is stable whenever it is also stable in the one-shot game. From Propositions 1 and 2 it is immediate to obtain the following.

---

<sup>6</sup>This paradoxical result that has been referred to as the “merger paradox” (see Salant et al. (1983)) is valid only in Cournot environments, and generally fails to hold in differentiated Bertrand models. This is an issue not raised here and left for future research.

<sup>7</sup>Remark also the similarity with the result in Salant et al. (1983) that if a merger of  $K$  firms is profitable, a merger with more firms is also profitable.

**Corollary 1** *When  $N \geq 4$ , there exists  $\bar{\delta} \in (0, 1)$  such that if  $\delta \geq \bar{\delta}$  the stable cartel contains more firms when all firms simultaneously choose quantities, than when cartel firms behave as a Stackelberg leader with respect to the fringe.*

## 2.2 Decision-making of production timing by the cartel

In this subsection we examine and discuss the equilibrium sequence of moves between the cartel and the fringe. One can check that for a given  $N$  and  $K$ , (i) a cartel firm earns higher profits when the cartel leads than under simultaneous play and (ii)  $\forall \delta_K^C, \delta_K^S \in (0, 1)$ ,  $\delta_K^C > \delta_K^S$ . From this analysis one may conclude that, in our model, an endogenous timing would suggest a Stackelberg-type sequential moves model. However, this is not the end of the story. Since for each  $\delta$ , the number of firms constituting the cartel with simultaneous play and leadership is the one derived in Propositions 1 and 2 respectively, we must also consider that firms foresee that the size of the stable cartel depends on the production timing. For simplicity and as in all cartel and fringe literature cited in the introduction, we assume that the cartel is able to impose its most preferred timing.<sup>8</sup> Thus, with complete information in  $t = 0$  all firms are perfectly informed about the fact that the cartel will impose its preferable production timing. Then, as a function of  $\delta$  firms may predict the size of the cartel and consequently the sequence of play at equilibrium. Thus, the game runs as follows. In  $t = 0$  firms decide simultaneously whether or not to join the cartel. Afterwards, and prior to market competition, the cartel chooses either Stackelberg behavior or simultaneous play with the fringe depending on the number of firms that would constitute the cartel in each case. In  $t = 1$  market competition begins. By backwards induction we have:

**Proposition 3** *When  $N \geq 4$ , there exists  $\tilde{\delta} \in (0, 1)$  such that if  $\delta \geq \tilde{\delta}$  the sequence of play*

---

<sup>8</sup>This assumption is based on the fact that partial cartels often involve a big market share. As an example, the citric acid industry where three North-American and five European firm were fined for fixing prices and allocating sales in the worldwide market. Their joint market share was around 60 percent. The fringe included a variety of minor companies based in Eastern Europe, Russia and China that could unlikely coordinate their behavior or counteract the production timing imposed by the cartel (see Levenstein and Suslow (2006)).

*chosen by the cartel implies that all firms simultaneously choose quantities. Otherwise, a leader's role to the cartel and a follower's role to the fringe is assigned.*

This finding casts doubt on the cartel literature's assumption of a Stackelberg cartel for a quantity-setting model. Intuitively, when  $\delta$  is large enough, collusion with leadership and with simultaneous play can be both sustained as a SPNE of the repeated game. However, the size of the stable cartel is smaller when the cartel is allowed to lead because firms inside the cartel find it desirable to exit. In other words, it is precisely the success of the leader cartel what may reduce the incentives to participate in it. However, when firms simultaneously choose quantities, the incentives to free-ride from a cartel by defecting from it are reduced since by leaving the cartel, firms risk the complete collapse of collusion. Consequently, the cartel decides to behave as a Stackelberg leader with respect to the fringe only when collusion is not easily sustained.

### **3 Concluding comments**

Several economists have investigated the endogenous timing in oligopoly games (see for instance Robson (1990), Albæk (1992) or Mailath (1993)). These papers suggested that the Stackelberg model is more plausible than the Cournot model if firms can choose when to take their actions. We addressed the question of whether the Stackelberg sequence is also appropriate in a quantity-setting cartelistic model. To that extent, we have developed a theoretical framework to study how in a repeated game setting, the sequence of play between the cartel and the fringe affects cartel stability. We show that the size of the stable cartel crucially depends on the sequence of each firm's choice: when the cartel behaves as a Stackelberg leader with respect to the fringe a stable leader cartel contains just over half the firms in the industry. On the other hand, when all firms simultaneously choose quantities the smallest cartel among those that are sustainable is stable. As a consequence, we obtain that the results of an endogenous sequence of play between a stable cartel and a fringe are mixed since a leader's role to the cartel and a follower's role to the fringe should only be assigned when the discount factor is not large. Then,

our model enables us to raise serious doubts on the cartel literature's assumption of a Stackelberg cartel for a quantity-setting model.

Several interesting issues have not been addressed in this paper. First of all, although the empirical evidence provides support for the exercise of market power by cartels, the assumption that the production timing is also imposed by the cartel could be relaxed. Second, the concept proposed by d'Aspremont et al. (1983) for cartel stability conjectures that the rest of the cartel will remain together and that no independent firm will react. This assumption is relaxed in Thoron (1998) and Diamantoudi (2005) and taking into account coalitional deviations and foresight respectively, a non-empty set of stable cartels is obtained under certain conditions. On the other hand, in Morasch (2000) strategic alliances are considered. They differ from traditional cartels in that member firms are aware of the fact that fringe firms react to collusion by expanding their output, and may use the alliance contract as a strategic commitment device. He obtains that in a linear Cournot oligopoly with at most five firms only one alliance forms whereas with more than five firms an alliance structure with at least two alliances results. To summarize, the framework we have worked with is, admittedly, a particular one. To analyze real-world cartels, additional research is required and the issues mentioned above should also be considered. We believe that those are subjects for future research.

## **Acknowledgements**

I am grateful to my supervisor Ramon Faulí-Oller for his advice and encouragement and to Pedro Barros, Luís Corchón, Inés Macho, Joel Sandonís and George Symeonidis for numerous valuable suggestions. I am also grateful to two anonymous referees for helpful comments which led to substantial improvements of the paper. Financial support by the “Ministerio Español de Educación y Ciencia” through its project “Gobierno Corporativo” (Ref: SEJ2005-08462) is gratefully acknowledged. The usual disclaimer applies.

## Appendix

Proof of Proposition 1. First, we prove that in the one-shot game when all firms simultaneously choose quantities, no cartel is stable, namely  $\Pi^c(N, K) < \Pi^f(N, K - 1)$ . It can be verified that when firms simultaneously choose quantities

$$\Pi^c(N, K) = \frac{(a-c)^2}{bK(N-K+2)^2} \text{ and } \Pi^f(N, K) = \frac{(a-c)^2}{b(N-K+2)^2}.$$

Then, internal stability holds if  $\Pi^c(N, K) - \Pi^f(N, K - 1) = \frac{1}{K(N-K+2)^2} - \frac{1}{(N-K+3)^2} \geq 0$  which can be easily proved that is true if  $N \in [-2 - \frac{1}{1+\sqrt{K}} + K, -2 + \frac{1}{1+\sqrt{K}} + K]$  but this cannot hold when  $K \leq N$ . Therefore, firms have always incentives to leave the cartel in the one-shot game. We denote by  $\delta_K^C$  the critical level of the discount factor above which a cartel of size  $K$  is sustainable with simultaneous play. It can be easily verified that

$$\delta_K^C = \frac{(-1+K)(1+N)^2}{4K^2(3+2N)-4K^3-(1+N)^2-3K(1+N)^2}$$

Then, it is also easy to check that  $\delta_K^C \leq 1$  is only true when  $K \geq \frac{3}{2} + N - \frac{1}{2}\sqrt{5 + 4N}$ . Note that this condition is exactly the same in Salant et al. (1983) which says that the minimum profitable agreement under Cournot oligopoly involves at least 80 percent of the firms in the industry (note for instance that in the last inequality  $N = 5 \implies K \geq 4$ ). Also,  $\frac{\partial \delta_K^C}{\partial K} = -\frac{4(1+N)^2(1-2K(3+(-3+K)K)+2N+2(-2+K)KN+N^2)}{(1-2K+N)^2(1+N+K(5-2K+3N))^2}$ . We can check that  $\frac{\partial \delta_K^C}{\partial K} < 0 \forall K \geq \frac{3}{2} + N - \frac{1}{2}\sqrt{5 + 4N}$ . Thus,  $\delta_K^C$  is strictly decreasing in  $K$  when  $\delta_K^C \in (0, 1)$ . Then, if  $\delta_K \leq \delta < \delta_{K-1}$  only a cartel of  $K$  or more firms can be sustained as a SPNE of the repeated game (that is, the cutoff of the discount factor decreases with the size of the cartel) and can thus be stable since with less than  $K$  firms, by joining the cartel firms would obtain the Cournot profits and condition (4) would not be satisfied. On the other hand, the cartel containing more than  $K$  firms is not stable either because smaller cartels are sustainable and we have just shown that  $\Pi^c(N, K) - \Pi^f(N, K - 1) < 0$  (namely, when cartels can be enforced like in the one shot game, and in this case this is true for the smaller cartel that can be sustained, firms have an incentive to leave the cartel). Therefore, when  $K \geq \frac{3}{2} + N - \frac{1}{2}\sqrt{5 + 4N}$  (namely,  $\delta_K \in (0, 1)$ ) if  $\delta \geq \delta_K$  it is immediate to verify that condition (4) is satisfied:  $\Pi_0^c(N, K) = \frac{1}{1-\delta}\Pi^c(N, K) > \Pi_0^f(N, K - 1) = \frac{1}{1-\delta}\Pi(N)$  (remember that obviously,  $\Pi^c(N, K) > \Pi(N) \forall K \geq 2$ ). On the other hand, it can be verified that external stability is satisfied in the one-shot game ( $\Pi^f(N, K) \geq \Pi^c(N, K+1)$ )

if  $\frac{1}{(2+N-K)^2} - \frac{1}{(1+K)(1+N-K)^2} \geq 0$  which is always true if  $K \leq N$ . Then, condition (5) is also satisfied for the cartel of  $K$  firms if  $\delta_K \leq \delta < \delta_{K-1}$  since  $\Pi_0^f(N, K) = \frac{1}{1-\delta} \Pi^f(N, K) > \Pi_0^c(N, K+1) = \frac{1}{1-\delta} \Pi^c(N, K+1)$ . In other words, in the latter case the cartel of  $K+1$  firms is sustainable and regarding the external stability, the analysis of the one-shot game applies and cartels are also stable in the repeated game. ■

**Proof of Proposition 2.** We proceed like in the proof of Proposition 1 and we first analyze the stability in the one-shot game: when the cartel behaves as a Stackelberg leader  $\Pi^c(N, K) = \frac{(a-c)^2}{4bK(N-K+1)}$  and  $\Pi^f(N, K) = \frac{(a-c)^2}{4b(N-K+1)^2}$ . It can be easily verified that internal and external stability become

$\frac{(2K^2+(2+N)^2-K(5+3N))}{K(1+N-K)(2-K+N)^2} \geq 0$  and  $\frac{(1+K(2K-1)+N(N-3K+1))}{(1+K)(K-N)(1-K+N)^2} \geq 0$  respectively. If  $N < 4$  external stability is never met unless  $K = N$ . In this case, internal stability holds if  $N \leq 4$ . If  $N \geq 4$  internal stability holds if  $K \leq \frac{3}{4}N - \frac{1}{4}\sqrt{(N-2)N-7} + \frac{5}{4}$  and external stability holds if  $K > \frac{1}{4}(1+3N - \sqrt{(N-2)N-7})$ . Hence, it follows that since  $(\frac{3}{4}N - \frac{1}{4}\sqrt{(N-2)N-7} + \frac{5}{4}) - \frac{1}{4}(1+3N - \sqrt{(N-2)N-7}) = 1$ , stable cartel exists only when both conditions hold and this is true for  $K \in [f(N), f(N)+1]$  where  $f(N) \equiv \frac{1}{4}(1+3N - \sqrt{(N-2)N-7})$ . Note that this parallels the results in Shaffer (1995) (Proposition 4, page 746).

Regarding the repeated game, we denote by  $\delta_K^S$  the critical level of the discount factor above which a cartel of size  $K$  is sustainable when the cartel behaves as a Stackelberg leader,

$$\delta_K^S = \frac{(1+K)^2(1-2K+N)(-1+2K(1-K+N)-N)}{2K^4(9+N)-2K^3(1+N)(14+N)-2K(1+N)(1+2N)+K^2(1+N)(17+11N)-(1+N)^2}$$

where  $\delta_K^S$  is calculated assuming that if a cartel firm deviates fringe firms optimally respond to the deviation. It is easy to check that  $\delta_K^S \in (0, 1)$  and  $\delta_K^S$  strictly increasing with  $K$  are true iff  $K \geq \frac{N+1}{2}$ . Also, one can check that  $f(N) > \frac{N+1}{2}$ . Then, we can apply exactly the same argument of the proof of Proposition 1 for the case where  $N \geq 4$ : When a cartel of size  $K > f(N) + 1$  is sustainable, only a cartel of  $K$  or less firms can also be sustained as a SPNE of the repeated game. However, as we showed above, in this case firms have an incentive to leave the cartel since a smaller cartel can be sustained and thus condition (4) is not satisfied:  $\Pi_0^c(N, K) = \frac{1}{1-\delta} \Pi^c(N, K) < \Pi_0^f(N, K-1) = \frac{1}{1-\delta} \Pi^f(N, K-$

1). In words, if a firm leaves the cartel, collusion can still be sustained and therefore, the cartel is not stable. On the other hand, when the a cartel of size  $K < f(N)$  can be sustained, only a cartel of  $K$  or less firms can be sustained as a SPNE of the repeated game (recall that the case where  $K < \frac{N+1}{2}$  is ruled out since  $\delta_K^S \notin (0, 1)$ ). However, as we showed in Proposition 1, firms have incentives to join the cartel and thus condition (5) is not satisfied:  $\Pi_0^f(N, K) = \frac{1}{1-\delta}\Pi^f(N, K) < \Pi_0^c(N, K+1) = \frac{1}{1-\delta}\Pi^c(N, K+1)$ . Therefore, when the cartel is sustainable (namely,  $\delta \geq \delta_K$ ) for all  $K > f(N) + 1$  internal stability does not hold and when  $K < f(N)$  external stability does not hold either. Thus, iff when  $K \in [f(N), f(N) + 1]$  and the cartel is sustainable, it can also be stable. Note also that if the cartel is not sustainable, (for instance  $\delta = \delta_K - \epsilon < \delta_K$ ) internal stability does not hold since  $\Pi_0^c(N, K) = \frac{1}{1-\delta}\Pi^c(N, K) < \Pi_0^f(N, K-1) = \frac{1}{1-\delta}\Pi^f(N, K-1)$ . Finally, when  $N < 4$  (in our model, since we assume  $N > 2$ , this case is only  $N = 3$ ) it is immediate to see that if  $K < N$ , the cartel cannot be stable for any  $\delta$  since the external stability condition mentioned above does not hold:  $\frac{1}{(-4+K)^2} + \frac{1}{K(K-2)-3} < 0$  if  $K = 2$ . When  $\delta \geq \delta_N$  internal and external stability are satisfied in the repeated game for  $K = N$  since  $\Pi_0^c(N, K) = \frac{1}{1-\delta}\Pi^c(N, K)$  and  $\Pi_0^f(N, K-1) = \frac{1}{1-\delta}\Pi^f(N, K-1)$  and as we showed above, in this case the cartel is stable in the one-shot game and consequently also in the repeated game. ■

Proof of Corollary 1. The result is directly implied by the proof of Propositions 1 and 2 since  $\frac{3}{2} + N - \frac{1}{2}\sqrt{5+4N} > f(N) = \frac{1}{4}(1 + 3N - \sqrt{(N-2)N-7})$ : when  $K \geq \frac{3}{2} + N - \frac{1}{2}\sqrt{5+4N}$ ,  $\delta_K^C$  is strictly decreasing in  $K$  and smaller than 1, thus there exists  $\bar{\delta} \in (0, 1) \geq \delta_{\bar{K}}$  such that when  $\delta \geq \bar{\delta}$ ,  $\bar{K} > f(N)+1$  is the stable cartel under simultaneous play. ■

Proof of Proposition 3. We denote by  $\Pi_c^c(N, K)$  the profit function of a cartel firm when cartel and fringe firms simultaneously choose quantities. Equivalently, let  $\Pi_s^c(N, K)$  be the profit function of a cartel firm when cartel firms behave as a Stackelberg leader. Then,  $\Pi_c^c(N, K) = \frac{(a-c)^2}{bK(N-K+2)^2}$  and  $\Pi_s^c(N, K) = \frac{(a-c)^2}{4bK(N-K+1)^2}$ . We assume that at  $t = 0$ , firms foresee and are perfectly informed about which is going to be the size of the stable cartel for any  $\delta$  with leadership and with simultaneous play. We assume that the number of firms constituting the cartel will be the stable one in each competition structure. It is easy

to check that  $\frac{\partial \Pi_s^c(N,K)}{\partial K}$  and  $\frac{\partial \Pi_c^c(N,K)}{\partial K}$  are positive and thus, both are monotonous increasing functions of  $K$ . On the other hand, the following is true:  $\Pi_s^c(N, \frac{3}{2} + N - \frac{1}{2}\sqrt{5 + 4N}) > \Pi_c^c(N, \frac{3}{2} + N - \frac{1}{2}\sqrt{5 + 4N})$  and  $\Pi_s^c(N, \frac{3}{2} + N - \frac{1}{2}\sqrt{5 + 4N}) < \Pi_c^c(N, N)$ . Therefore, there exists  $\tilde{K} \in (\frac{3}{2} + N - \frac{1}{2}\sqrt{5 + 4N}, N]$  such that  $\Pi_s^c(N, \frac{3}{2} + N - \frac{1}{2}\sqrt{5 + 4N}) = \Pi_c^c(N, \tilde{K})$ . Since  $\forall K \geq \frac{3}{2} + N - \frac{1}{2}\sqrt{5 + 4N}$ ,  $\delta_K^C$  is strictly decreasing in  $K$  and  $\delta_K^C > \delta_K^S$ , there exists also  $\tilde{\delta} \equiv \delta_{\tilde{K}} < 1$  such that when  $\delta \geq \tilde{\delta}$ , the size of the stable cartel under simultaneous play is  $K \geq \tilde{K}$  whereas under cartel leadership is  $\hat{K} \in [f(N), f(N) + 1] < \tilde{K}$ . Thus, when  $\delta \geq \tilde{\delta}$  cartel firms would prefer the simultaneous play and would not behave a Stackelberg leader since  $\Pi_s^c(N, \hat{K}) < \Pi_c^c(N, K)$ , that is, cartel firms would obtain higher profits in the stable cartel under simultaneous play than under cartel leadership. The reverse is true when  $\delta < \tilde{\delta}$ . ■

## References

- Albæk, S. (1990). “Stackelberg Leadership as a Natural Solution under Cost Uncertainty”. *Journal of Industrial Economics* 38, 335–347.
- d’Aspremont, C., Jacquemin, A., Gabszewick, J.J. and Weymark, J.A. (1983). “On the Stability of Collusive Price Leadership”. *Canadian Journal of Economics* 16 (1), 17-25.
- Davidson, C., Deneckere, R. (1984). “Horizontal Mergers and Collusive Behavior”. *International Journal of Industrial Organization* 2, 117-132.
- Diamantoudi, E. (2005). “Stable cartels revisited”. *Economic Theory* 26, 907–921
- Donsimoni, M.-P. (1985). “Stable Heterogenous Cartels”. *International Journal of Industrial Organization* 3(4), 451-467.
- Donsimoni, M.-P., Economides, N. and Polemarchakis, H. (1986). “Stable Cartels”. *International Economic Review* 27, 317-327.
- Friedman, J.W. (1971). “A Non-cooperative Equilibrium for Supergames”. *Review of Economic Studies* 28, 1-12.
- Levenstein, M. and Suslow, V. Y. (2006). “What Determines Cartel Success?”. *Journal of Economic Literature*, 44(1), 43-95.
- Lofaro, A. (1999). “When Imperfect Collusion Is Profitable”. *Journal of Economics*

70(3), 235-259.

-Mailath, G. J. (1993). "Endogenous Sequencing of Firm Decisions". *Journal of Economic Theory* 59, 169-182.

-Martin, S. (1993). "Advanced Industrial Economics". *Blackwell Publishers*.

-Morasch, K. (2000). "Strategic alliances as Stackelberg cartels – concept and equilibrium alliance structure". *International Journal of Industrial Organization* 18, 257–282.

-Robson, A. (1990). "Stackelberg and Marshall". *American Economic Review* 80, 69-82.

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

-Salant, Stephen. Switzer, Sheldon and Reynolds, Robert (1983). "The Effects of an Exogenous Change in Industry Structure on Cournot-Nash Equilibrium". *Quarterly Journal of Economics* 98(2), 185-99.

-Selten, R. (1973). "A simple model of Imperfect Competition, where 4 are Few and 6 are Many". *International Journal of Game Theory* 2(3), 141-201.

-Shaffer, S. (1995). "Stable Cartels with a Cournot Fringe". *Southern Economic Journal* 61, 744-754.

-Thoron, S. (1998). "Formation of a Coalition Proof Stable Cartel". *Canadian Journal of Economics* 31, 63-76.

-Thoron, S. (2004). "Which acceptable agreements are equilibria?". *Mathematical Social Sciences* 47, 111–134.