ORIGINAL PAPER

A dynamic approach to cartel formation

J. Kuipers · N. Olaizola

Abstract   We study the endogenous formation of cartels from a dynamic point of view. First, we review some stability analyses from existing literature, which usually predict a rather small cartel as the outcome. We then study a dynamic model, in which firms can freely leave the cartel if they wish, and firms can enter the cartel if both the cartel and the firm so desire. We show that the dynamic process converges to a strongly stable cartel if one exists. Otherwise, the process does not converge, but during the process the cartel-size will never go below any size predicted by a stability analysis.

Keywords   Group decisions · Firm organization · Market structure · Static versus dynamic

JEL Classification   L22 · C72

1 Introduction

It is generally assumed that the formation of a cartel induces positive externalities on firms outside the cartel and that there is an incentive to free-ride, so that collusive agreements tend to be unstable.[1] Most of the literature on this topic searches for cartels

---

[1] The fact that the existence of a cartel may be hindered by free-riding was observed by Stigler as early as 1950.

J. Kuipers (✉)
Faculty of Health and Sciences, MICC/Mathematics, Maastricht University,
P.O. Box 616, 6200 MD Maastricht, The Netherlands
e-mail: kuipers@micc.unimaas.nl

N. Olaizola
Dept. de Fundamentos del Análisis Económico I, Facultad de Ciencias Económicas y Empresariales,
University of the Basque Country, Avenida Lehendakari Agirre 83, 48015 Bilbao, Spain
e-mail: norma.olaizola@ehu.es

that are nevertheless stable in some sense, so the models proposed do not answer the question of what happens in the many cases where no stable agreement exists. The main contribution of this paper is to develop a model of endogenous formation of cartels that allows us to deal with such cases.

Selten (1973) was the first to apply a formal non-cooperative model of coalition formation to the study of cartels. His results suggest that cartels of at most five firms will be observed. Other early works on cartel formation are comparative static analyses, i.e., they compare the Nash equilibria of industry structures before and after a merger (see Salant et al. 1983 for a model à la Cournot and Deneckere and Davidson 1985 for a model à la Bertrand). These studies show that a merger may lead to losses for its members. This is a somewhat counter intuitive result, as the members of a merger have the option of leaving their production strategy unchanged after they merge, thereby obtaining the same profit. One is inclined to believe that profits at the new equilibrium should be higher, but that intuition is false in general. The concept of a stable cartel introduced by D' Aspremont et al. (1983) was the first to take into account the effects on profits at the new equilibrium after addition or loss of a cartel member. A cartel is called stable if no firm in it has an incentive to defect (internal stability) and no firm outside it has an incentive to join (external stability). Under the assumptions of identical firms and equal division of profits in the cartel, a stable cartel always exists, although it may be degenerate (i.e. consist of a single firm). When they applied this stability concept to specific economic models, D' Aspremont et al. (1983) found that stable cartels are usually relatively small compared to the size of the industry. Moreover, the fraction of firms in the cartel tends to zero as the size of the industry increases.

In the definition of stability due to D' Aspremont et al. (1983) only deviations by a single firm are considered. In order to allow the possibility of coalitional deviations, Thoron (1998) develops a non-cooperative model of cartel formation in which all firms have the option of cooperating (in a cartel) or not. It turns out that the set of Nash equilibrium outcomes of the Thoron's cartel formation model coincides with the set of stable cartels. Thoron (1998) also applies the concepts of a strong Nash equilibrium and a coalition-proof Nash equilibrium, and defines strong stability and coalition-proof stability accordingly. She shows that strongly stable cartels do not necessarily exist, and that there exists a unique size of coalition-proof stable cartels, which is the size of a largest stable cartel.

In our view, the inexistence of strongly stable cartels is consistent with what happens in real life, i.e. cartels break up and other configurations appear, but the stability concepts mentioned above do not explain how the movements take place. In order to explain those endogenous changes we introduce a dynamic model of cartel formation. In this dynamic context, convergence to a certain cartel is not guaranteed, as the cartel may keep changing indefinitely. Shenoy (1979), who was the first to use this approach for the analysis of coalition formation, introduced the concept of an *elementary dynamic solution*. This solution concept captures the dynamic aspect of negotiations: whenever the negotiation process leads to an alternative that belongs to an elementary dynamic solution, it may shift to any other alternative that belongs to the same set in a finite number of steps, and from then on any alternative that does not belong to this elementary dynamic solution is impossible to reach. The dynamic

approach thus allows us to understand those situations in which there is no strong Nash equilibrium. Moreover, since the elementary dynamic solution and the strong Nash equilibrium coincide when the latter exists, the solution concept we use is consistent with the more traditional approach. Specifically, we show that convergence to a strongly stable cartel (Thoron 1998) is guaranteed if such a cartel exists. Otherwise the dynamic process does not converge, but we show that there is a unique set of cartels to which that dynamic process becomes restricted. From the characterization of this set, it follows immediately that all cartels in it have at least the size of a largest stable cartel. The results are derived under the assumptions that a cartel generates positive externalities for independent firms and that there is an incentive for free riding, which are assumptions satisfied by most oligopoly models.

The paper is organized as follows. In Sect. 2, we review the results on stability of cartels by d'Aspremont and Thoron that are relevant to this paper. In Sect. 3, we introduce our dynamic model of cartel formation. We propose, discuss and analyze a mechanism that governs the possible moves from one cartel to another. We illustrate the results for a Cournot oligopoly model with a homogeneous product. Finally, Sect. 4 contains some conclusions and pointers for further research.

2 Preliminaries

We consider an oligopoly that consists of a set $N = \{1, \ldots, n\}$ of firms (with $n \in \mathbb{N}$). We assume that any subset of N can form a cartel, at most one cartel can exist at a given time, all firms in the oligopoly are ex-ante identical, and firms inside a cartel share their profits equally.[2] As a consequence, the profit of a firm depends only on the size of the cartel and on whether the firm is independent or a member of the cartel. Let $\pi^C(k)$ denote the profit of a firm that belongs to a cartel of size k, and let $\pi^I(k)$ denote the profit of an independent firm in the presence of a cartel of size k.

D' Aspremont et al. (1983) calls a cartel of size k *internally stable*

$$\text{if } k = 1 \text{ or if } k > 1 \text{ and } \pi^C(k) \geq \pi^I(k - 1),$$

and *externally stable*

$$\text{if } k = n \text{ or if } k < n \text{ and } \pi^I(k) \geq \pi^C(k + 1).$$

A cartel is *stable* if it is both internally stable and externally stable.

D'Aspremont et al. show that a stable cartel always exists, although it may be degenerate (that is, it may have only one member). In the following theorem, we provide a proof of d'Aspremont's existence result, and an explicit expression for the size of a largest stable cartel.

---

[2] This assumption is fulfilled if firms inside a cartel bargain "a la Nash." Since firms are symmetric, the distribution of gains will be symmetric.

Theorem 1 *A stable cartel always exists and the size k^* of a largest stable cartel is*

$$k^* := \max\{k \mid \pi^C(k) \geq \pi^I(k-1); k > 1\},$$

if we adopt the convention that the maximum over the empty set is 1.

Proof To see that a cartel of size k^* is stable, observe that the internal stability condition is trivially satisfied if $k^* = 1$, and if $k^* > 1$, it follows from $\pi^C(k^*) \geq \pi^I(k^* - 1)$, by definition of k^*. External stability is trivially satisfied if $k^* = n$, and if $k^* < n$, it follows from $\pi^C(k^* + 1) < \pi^I(k^*)$, by definition of k^*. To see that k^* is the size of a largest stable cartel, note that $\pi^C(k) < \pi^I(k-1)$ for all $k > k^*$ by definition of k^*. Hence the internal stability condition $\pi^C(k) \geq \pi^I(k-1)$ is violated for all $k > k^*$. $\square$

In the same framework, Thoron (1998) defines a non-cooperative game G in which the firms are the players and each player has the choice between two actions, C and I. If a firm chooses C, it will be member of the cartel. If it plays I, then it will be independent.

Thoron shows a one-to-one correspondence between stable cartels defined by D' Aspremont et al. (1983) and the Nash equilibria of G.

Theorem 2 (Thoron 1998) *A cartel of size k is stable if and only if it is the outcome of a Nash equilibrium in the non-cooperative game G.*

The definition of stability considers a possible deviation by individual firms. A firm can either leave the cartel or join it. Thoron calls a cartel *strongly stable* if it is the outcome of a strong Nash equilibrium in G, i.e. if no coalition of firms has an incentive to deviate from their strategies.

Thoron makes two assumptions on the profit functions π^I and π^C.

A1 $\pi^I(k) > \pi^I(k-1)$ for all $k \in \{2, \ldots, n-1\}$.
A2 $\pi^I(1) = \pi^C(1)$ and $\pi^I(k) > \pi^C(k)$ for all $k \in \{2, \ldots, n-1\}$.

Assumption **A1** says that the profit of an independent firm increases as the cartel size increases. Hence, collusion generates a positive externality on independent firms. Assumption **A2** says that, given a cartel, an independent firm receives a larger profit than a firm in the cartel. Therefore, firms that remain independent are free-riders.

Theorem 3 (Thoron 1998) *Under Assumption A2, a cartel of size k is strongly stable if and only if*

(i) $\pi^C(k) \geq \pi^I(k-l)$ *for all* $l \in \{1, \ldots, k-1\}$*, and*
(ii) $\pi^I(k) > \pi^C(k+l)$ *for all* $l \in \{1, \ldots, n-k\}$.

That is, a cartel is the result of a strong Nash equilibrium if it is strongly internally stable, and strongly externally stable. A strongly stable cartel satisfies a property of strong internal stability: no coalition of firms in the cartel has the incentive to jointly defect from it, and a property of strong external stability: no coalition of firms outside the cartel has the incentive to join it. However, it is not difficult to find examples for which no strongly stable cartel exists (Sect. 3).

The fact that strong Nash equilibria may fail to exist motivated (Bernheim et al. 1987; Bernheim and Whinston 1987) to introduce the concept of a coalition-proof Nash equilibrium. This concept also allows for coalitional deviations, but not all profitable coalitional deviations are considered as valid deviations here. Thoron calls a cartel *coalition-proof stable* if it is the outcome of a coalition-proof Nash equilibrium in G.

Theorem 4 *Under Assumptions A1 and A2, there is a unique size for coalition-proof stable cartels, which is k^*, the size of a largest stable cartel.*

3 Dynamic cartel formation

In this section, we introduce a dynamic model of cartel formation: we consider the possibility of a transition from one cartel to another and that an infinite sequence of such transitions may take place. Let us call a directed graph $\mathcal{G}$ a *transition graph* if its vertex set is the set of cartels. Its arc set is the set of possible transitions.

The term *abstract game* is sometimes used to denote a pair (X, dom), where X is a set of *outcomes* and dom is a binary relation defined on X. Thus, any transition graph may also be interpreted as an abstract game. As a solution concept for abstract games, Shenoy introduces the concept of an *elementary dynamic solution*.[3] Here, we give the definition in terms of a transition graph.

Definition 1 A nonempty set $\mathcal{A}$ of cartels is an *elementary dynamic solution* (*EDS*) of a transition graph $\mathcal{G}$ if

(i) for every $K \in \mathcal{A}$ and every $L \in \mathcal{A}$, there is a directed path in $\mathcal{G}$ from K to L, and

(ii) there is no $K \in \mathcal{A}$ and $L \notin \mathcal{A}$ such that there is a directed path in $\mathcal{G}$ from K to L.

The first condition says that for any two cartels there is a sequence of possible transitions from one to the other. This implies that whenever the transition process leads to a cartel in an EDS the process may subsequently lead to any other cartel in the same EDS. The second condition says that no cartel outside an EDS can be reached from a cartel inside the EDS. Thus, once the process leads to a cartel in an EDS, the process is restricted to that set, and there is no smaller set of cartels to which the process will become restricted. There is at least one EDS in our finite context (see Shenoy, Theorem 2). In general there may be more than one EDS, in which case these sets are obviously disjoint. In what follows we explore several restrictive assumptions on a transition graph and then investigate the EDSs under Assumptions **A1** and **A2**.

We assume that firms inside a cartel are free to leave, but cannot be forced to leave. We also assume that firms outside the cartel need the consent of firms inside the cartel to enter, but cannot be forced to enter.

Let $K \subseteq N$ be a cartel. We say that *a group of firms $M \subseteq N$ can move from K to another cartel L*, and we write $K \xrightarrow{M} L$, if one of the following two conditions holds:

---

[3] The union of all elementary dynamic solutions is called *dynamic solution*. This solution was previously defined by Kalai et al. (1976) under the name of *admissible set*. Schwartz (1974) also introduces an equivalent solution, called *top cycle solution*.

(i) $L \subseteq K$ and $K \setminus L \subseteq M$,

(ii) $K \cup L \subseteq M$.

Condition (i) means that a group of firms leaves the cartel K. The firms of L that remain in K do not have to agree with this transition: if all members of $K \setminus L$ want to leave, they can do so. Now, any group of firms M such that $M \supseteq K \setminus L$ would be able to force the transition from K to L. Condition (ii) means that the new cartel L contains at least one firm outside K. In this case, all members of K have to agree with the transition. All new members of the cartel also have to agree, so in fact the consent of everyone in $K \cup L$ is needed. Therefore, any group of firms $M \supseteq K \cup L$ would be able to force the transition from K to L.

We say that the move $K \xrightarrow{M} L$ is *profitable* if no firm in M loses and at least one firm in M strictly gains. We define $\mathcal{G}^P$ as the transition graph whose arc set consists of the cartel pairs (K, L) for which a profitable move of the type $K \xrightarrow{M} L$ exists.

Not all transitions in $\mathcal{G}^P$ seem equally reasonable. Suppose that a group M of firms is considering the move $K \xrightarrow{M} L$. Then each member of M is likely to take other possible moves into consideration before making its decision, and it may have better options. Say that the move $K \xrightarrow{M'} L'$ is a *countermove* for $K \xrightarrow{M} L$ if the following two conditions hold:

(i)   The move $K \xrightarrow{M'} L'$ is profitable;

(ii) For all $i \in M'$, i's payoff at L' is no less than the maximum of its two payoffs at K and L, and for at least one $\hat{\imath} \in M' \cap M$, $\hat{\imath}$'s payoff at L' is strictly higher than the maximum of its two payoffs at K and L.

We call a profitable move *robust* if it has no countermove, and we denote by $\mathcal{G}^R$ the transition graph that represents the set of robust moves. If a group of firms plan a move that is robust, then it seems likely that the plan will indeed be carried out, since it is profitable for everyone who has a say in it, and no one has a better alternative.

If firms indeed look around for other possibilities when a plan comes up, then it seems unlikely that a plan will be carried out if it has a robust countermove, since this is a better and credible alternative for at least one of the firms. The situation is less clear when a profitable move has only non-robust countermoves, since such countermoves may not be credible alternatives. Let $\mathcal{G}^{\mathrm{NRC}}$ denote the transition graph of profitable moves without a robust countermove (NRC move). Note that $\mathcal{G}^R \subseteq \mathcal{G}^{\mathrm{NRC}}$, where the notation $\mathcal{G} \subseteq \mathcal{H}$ means that the arc set of $\mathcal{G}$ is a subset of the arc set of $\mathcal{H}$.

Now, let $\mathcal{G}^*$ denote the actual transition graph. We have argued that every transition in $\mathcal{G}^R$ corresponds to a plan that is likely to be carried out. Therefore, we will assume that $\mathcal{G}^R \subseteq \mathcal{G}^*$. Also, every transition not in $\mathcal{G}^{\mathrm{NRC}}$ corresponds to a plan that is unlikely to be carried out. Therefore, we assume $\mathcal{G}^* \subseteq \mathcal{G}^{\mathrm{NRC}}$. In the following, we demonstrate that, under Assumptions **A1** and **A2**, the EDSs of $\mathcal{G}^R$ and $\mathcal{G}^{\mathrm{NRC}}$ coincide, so that we are able to characterize the EDSs of $\mathcal{G}^*$.

Define $\hat{k}$ as the cartel size for which the payoff to firms inside the cartel is maximal. If there are different cartel sizes for which this payoff is maximal, then choose the smallest. Formally,

$$\hat{k} := \min\{k \mid \pi^C(k) \geq \pi^C(l) \text{ for all } l \in N\}.$$

Further, define $\bar{k}$ as the minimum number of firms needed for a cartel such that an independent firm receives a profit that is higher than the profit of firms inside a cartel of size $\hat{k}$. Formally,

$$\bar{k} := \min\{k \mid \pi^I(k) > \pi^C(\hat{k})\}.$$

If $\hat{k} \geq 2$, then $\bar{k} \leq \hat{k}$, since then $\pi^I(\hat{k}) > \pi^C(\hat{k})$ by Assumption **A2**. As there is no incentive to form a cartel if $\hat{k} = 1$, we assume that $\hat{k} \geq 2$. Recall from Sect. 2 that the unique size of a largest stable cartel is

$$k^* := \max\{k \mid \pi^C(k) \geq \pi^I(k-1)\}.$$

We have $\pi^I(k-1) \geq \pi^I(\bar{k}) > \pi^C(\hat{k}) \geq \pi^C(k)$ for all $k > \bar{k}$, where the first inequality follows from **A1**, the second from the definition of $\bar{k}$, and the third from the definition of $\hat{k}$. Hence, it follows from the definition of k^* that $k^* \leq \bar{k}$. In summary, for the three special cartel sizes k^*, $\bar{k}$ and $\hat{k}$, we have

$$k^* \leq \bar{k} \leq \hat{k}.$$

Lemma 1 *Let K be a cartel of size $k > \bar{k}$. Then for all $i \in K$, the move $K \overset{\{i\}}{\to} K\backslash\{i\}$ is robust. Also, for any other profitable move at K, there is a robust countermove of the type $K \overset{\{i\}}{\to} K\backslash\{i\}$.*

Proof [4] Consider a profitable move of the type $K \overset{M}{\to} L$. Since $k > \bar{k}$, we have $\pi^I(k) > \pi^I(\bar{k}) > \pi^C(\hat{k}) \geq \pi^C(l)$, where the first inequality follows from **A1**, the second the definition of $\bar{k}$, and the third from the definition of $\hat{k}$. This shows that no independent firm has entered the cartel with this move, as the move was assumed to be profitable, so we have proved that $L \subset K$.

Choose $i \in K\backslash L$. We will now show that the move $K \overset{\{i\}}{\to} K\backslash\{i\}$ is robust and that, unless $L = K\backslash\{i\}$, it is a countermove for $K \overset{M}{\to} L$. Observe that the move $K \overset{\{i\}}{\to} K\backslash\{i\}$ is indeed profitable, since $k > \bar{k} \geq k^*$, hence $\pi^I(k-1) > \pi^C(k)$ by definition of k^*. To see that it is a countermove for $K \overset{M}{\to} L$ if $L \neq K\backslash\{i\}$, note that $\pi^I(l) < \pi^I(k-1)$ by **A1**, since $l < k-1$ in this case. Hence, is payoff will be strictly higher if it moves to $K\backslash\{i\}$. Finally, to see that $K \overset{\{i\}}{\to} K\backslash\{i\}$ is robust, it suffices to observe that this move is actually a countermove for all other profitable moves that involve firm i. Therefore, no other move can be a countermove for it. □

Lemma 2 *Let K be a cartel of size $k < \bar{k}$, and let $\hat{K}$ be a cartel of size $\hat{k}$. Then the move $K \overset{M}{\to} \hat{K}$ is robust for all M with $M \supseteq K \cup \hat{K}$.*

Proof We have $\pi^I(k) \leq \pi^C(\hat{k})$, by definition of $\bar{k}$ and the fact that $k < \bar{k}$. This shows that independent firms that join the cartel $\hat{K}$, i.e. the members of $\hat{K}\backslash K$ do not lose. We also have $\pi^C(k) < \pi^C(\hat{k})$, by definition of $\hat{k}$ and the fact that $k < \bar{k} \leq \hat{k}$. This

---

[4] We use upper-case letters to indicate cartels. For the size of a cartel, we use the corresponding lower-case letter.

shows that members of $K \cap \hat{K}$ strictly gain. Also members of $K \backslash \hat{K}$ strictly gain, since $\pi^C(k) < \pi^C(\hat{k}) < \pi^I(\hat{k})$, where the second inequality follows from **A2**. Finally, possible members of $M \backslash (K \cup \hat{K})$ strictly gain, since $\pi^I(k) \leq \pi^C(\hat{k}) < \pi^I(\hat{k})$. We have proved that the move $K \overset{M}{\to} \hat{K}$ is profitable, since no one in M loses and the members of K (a non-empty set) strictly gain.

To see that the move is robust, suppose that the move $K \overset{M'}{\to} L$ is a countermove for it. If $l \geq \hat{k}$ and $L \neq \hat{K}$, then at least one cartel member of L is independent in $\hat{K}$. Its payoff at L, however, is strictly less than it is at $\hat{K}$, since $\pi^C(l) \leq \pi^C(\hat{k}) < \pi^I(\hat{k})$, where the first inequality is by definition of $\hat{k}$, and the second by **A2** . Therefore, $l < \hat{k}$. However, for all $l < \hat{k}$, we have $\pi^C(l) < \pi^C(\hat{k}) = \min(\pi^C(\hat{k}), \pi^I(\hat{k}))$, where the inequality is by definition of $\hat{k}$ and $l < \hat{k}$, and the equality is by **A2**. This means that the payoff of all members of L is strictly less than it is at $\hat{K}$. So in fact no countermove exists. □

Lemma 3 *Let K be a cartel of size $k < \bar{k}$, and let L be a cartel of size l with $l \leq k$ or $l > \hat{k}$. Then, if the move $K \overset{M}{\to} L$ is profitable, it has a robust countermove.*

Proof Suppose the move $K \overset{M}{\to} L$ is profitable. This implies $l \neq k$. If $l < k$, then all firms of N receive a profit that is at most $\max(\pi^C(l), \pi^I(l)) = \pi^I(l)$, where the equality is from **A2**. At a cartel of size $\hat{k}$, any firm of N receives at least $\min(\pi^C(\hat{k}), \pi^I(\hat{k})) = \pi^C(\hat{k})$, where the equality is again from **A2**. Since $l < k < \bar{k}$, we have $\pi^I(l) < \pi^I(k) \leq \pi^C(\hat{k})$, where the first inequality is from **A1** and the second one is by definition of $\bar{k}$. It follows that $K \overset{N}{\to} \hat{K}$ is a countermove for $K \overset{M}{\to} L$, where $\hat{K}$ is any cartel of size $\hat{k}$. From Lemma 2, this countermove is robust.

If $l > \hat{k}$, then consider the move $K \overset{L}{\to} \hat{K}$, where $\hat{K}$ is chosen to be a subset of L of cardinality $\hat{k}$. This move is robust from Lemma 2. We have $\pi^C(l) \leq \pi^C(\hat{k})$ by definition of $\hat{k}$, so that the members of $\hat{K}$ do not lose compared to the move $K \overset{M}{\to} L$. The members of $L \backslash \hat{K}$ (a non-empty set) strictly gain, since $l > \hat{k} \geq \bar{k}$, hence $\pi^I(\hat{k}) > \pi^C(\hat{k}) \geq \pi^C(l)$, where the first inequality is by **A2** and the second is by definition of $\hat{k}$. This shows that $K \overset{L}{\to} \hat{K}$ is a robust countermove for $K \overset{M}{\to} L$. □

Lemmas 1 and 2 show that robust moves exist at every cartel, except possibly at cartels of size $\bar{k}$. To handle this case, we need to distinguish between the cases $k^* < \bar{k}$ and $k^* = \bar{k}$.

Lemma 4 *If $k^* = \bar{k}$, then no profitable move exists at a cartel of size $\bar{k}$. If $k^* < \bar{k}$ and if $\bar{K}$ is a cartel of size $\bar{k}$, then for all $i \in \bar{K}$, the move $\bar{K} \overset{\{i\}}{\to} \bar{K} \backslash \{i\}$ is robust, and for any other profitable move at $\bar{K}$, there is a robust countermove of the type $\bar{K} \overset{\{i\}}{\to} \bar{K} \backslash \{i\}$.*

Proof For the case $k^* < \bar{k}$, it suffices to read the proof of Lemma 1 once more, and to observe that the result is valid under the more general condition $k \geq \bar{k}$ and $k > k^*$.

Now assume $k^* = \bar{k}$, let $\bar{K}$ be a cartel of size $\bar{k}$, and consider a move $\bar{K} \overset{M}{\to} L$. If $l < \bar{k} = k^*$, then $\max(\pi^C(l), \pi^I(l)) = \pi^I(l) \leq \pi^I(k^* - 1) \leq \pi^C(k^*) = \pi^C(\bar{k}) = \min(\pi^C(\bar{k}), \pi^I(\bar{k}))$, where the equalities and inequalities are obtained by applying

respectively **A2**, **A1** and $l \leq k^* - 1$, the definition of k^*, the assumption $k^* = \overline{k}$, and **A2**. Since everyone's payoff at $\overline{K}$ is at least the payoff at L, the move is not profitable.

If $l \geq \overline{k}$ and $L \neq \overline{K}$, then at least one independent firm at $\overline{K}$ is a member of L. This firm is also a member of M and it loses, since $\pi^I(\overline{k}) > \pi^C(\hat{k}) \geq \pi^C(l)$, where the first inequality is by definition of $\overline{k}$, and the second one is by definition of $\hat{k}$. Hence, such a move is not profitable either. □

We now characterize the EDSs of the transition graph $\mathcal{G}^*$.

Proposition 1 *If $k^* < \overline{k}$, then the unique EDS of $\mathcal{G}^*$ is $\mathcal{A}^* = \{K \mid \overline{k} - 1 \leq |K| \leq \hat{k}\}$. If $k^* = \overline{k}$, then each EDS of $\mathcal{G}^*$ consists of a single cartel, and the size of such a cartel is $k^* = \overline{k}$.*

Proof Assume $k^* < \overline{k}$. It then follows from Lemmas 1, 2 and 4 that for every cartel, a sequence of robust moves exists, leading to cartels of size $\overline{k} - 1, \ldots, \hat{k}$, and in fact *any* cartel of such size can be reached. Since these are all transitions in $\mathcal{G}^R \subseteq \mathcal{G}^*$, there is a unique EDS, which contains $\mathcal{A}^*$. To prove that the unique EDS is equal to this set, we have to demonstrate that every move in $\mathcal{G}^*$ from a cartel inside $\mathcal{A}^*$ is again to a cartel inside $\mathcal{A}^*$. It follows from Lemmas 1 and 4 that, for cartels of sizes $\overline{k}, \ldots, \hat{k}$, every move in $\mathcal{G}^{NRC}$ is to a cartel with one firm fewer. Since $\mathcal{G}^* \subseteq \mathcal{G}^{NRC}$, this shows that one cannot leave $\mathcal{A}^*$ from cartels of sizes $\overline{k}, \ldots, \hat{k}$. Lemma 3 states that a transition in $\mathcal{G}^{NRC}$ at cartel of size $\overline{k} - 1$ is necessarily to a cartel of size $\overline{k}, \ldots, \hat{k}$. Hence, one cannot leave $\mathcal{A}^*$ from a cartel of size $\overline{k} - 1$ either.

Now assume $k^* = \overline{k}$. It then follows from Lemmas 1 and 2 that for every cartel, a sequence of robust moves exists, leading to a cartel of size $\overline{k}$. According to Lemma 4, no profitable move exists at cartels of size $\overline{k}$. It thus follows that there is a one-to-one correspondence between the cartels of size $\overline{k}$ and the EDSs of $\mathcal{G}^*$. □

Proposition 1 can be compared with the equilibrium stability results from existing literature by associating an appropriate transition graph with a given equilibrium concept. For example, call a move at cartel K *strongly Nash* if it is a profitable coalitional deviation from the strategy profile to form K, and denote the transition graph of strongly Nash moves by $\mathcal{G}^{SN}$. According to this definition, the strongly Nash equilibria (the strongly stable cartels) correspond to singleton EDSs in $\mathcal{G}^{SN}$. In general, there may be other (non-singleton) EDSs, and that is certainly the case when there is no strongly stable cartel. Under Assumptions **A1** and **A2**, the transition graph $\mathcal{G}^{SN}$ coincides with the graph $\mathcal{G}^P$ of profitable moves. The explanation is that under the assumptions, if it is profitable for a firm to join the cartel then this is also profitable for the cartel members. Hence, there is no difference between a profitable move under the open membership rule (i.e. a strongly Nash move) and a profitable move under the exclusive membership rule, as defined in this paper.

Call a move at K *coalition-proof* if it corresponds to a coalition-proof deviation from the strategy to form cartel K, and denote the transition graph of coalition-proof moves by $\mathcal{G}^{CP}$. By definition, the coalition-proof equilibria correspond to EDSs in $\mathcal{G}^{CP}$. The graph $\mathcal{G}^{CP}$ may have more elementary solutions, but under Assumptions **A1** and **A2**, the EDSs contain cartels of size at most k^* only. To see this, we need to know that a coalition-proof deviation is recursively defined as a deviation that is profitable

and for which no proper *subcoalition* exists that has a coalition-proof deviation from the deviation. Now, it is straightforward to prove that the move to a cartel with one firm fewer (as a result of one firm that leaves the cartel) is a profitable move whenever the size is larger than k^*. Since there is no proper subcoalition that can further deviate, it follows trivially that the transition is in $\mathcal{G}^{CP}$. Hence, for a cartel of size $k > k^*$, a sequence of coalition-proof moves exists to a cartel of size k^*. Since the last cartel in the sequence corresponds to an EDS, then all other cartels in the sequence do not belong to any EDS. So indeed, every cartel in an EDS has at most size k^*.

The question arises of why the dynamic solutions of $\mathcal{G}^{CP}$ and $\mathcal{G}^*$ do not intersect, except possibly at k^*. Part of the answer is given above, and lies in the fact that only subcoalitions can prevent a profitable deviation in the definition of a coalition-proof deviation. Hence, a profitable move by one firm is always in $\mathcal{G}^{CP}$, but not necessarily in $\mathcal{G}^*$, so that a cartel will generally break down further when transitions in $\mathcal{G}^{CP}$ are made, as compared to when transitions in $\mathcal{G}^*$ are made. This explains why the EDSs of $\mathcal{G}^{CP}$ contain small cartels, but it does not explain why all of them are small. We are able to deduce this from the fact that cartels of size k^* correspond, by definition, to EDSs in $\mathcal{G}^{CP}$ coalition-proof. To understand why, consider a deviation from a cartel of size $k \leq k^*$ to a cartel of size $k' > k^*$, and observe that the subsequent deviation to a cartel of size $k' - 1$ is a coalition-proof deviation from the deviation. Therefore, a move from $k \leq k^*$ to $k' > k^*$ is not coalition-proof. We see that the argument why no cartels larger than k^* form in $\mathcal{G}^{CP}$ is essentially a farsighted argument: no large cartel is formed, since it will break down later anyway.

The following example illustrates the possible gap between k^* and $\bar{k} - 1$.

Example 1 **A Cournot oligopoly model** Consider a single industry consisting of n ($n \geq 3$) firms that produce a homogeneous good. The demand function is $p = 1 - Q$, with p being the price of the good and Q being the total quantity produced in the industry. Marginal cost of production is constant and equal to c ($0 < c < 1$), for all firms. Given a cartel $K \subseteq \{1, \ldots, n\}$, the cartel and the independent firms compete in the market by setting quantities simultaneously, i.e. they play a non-cooperative game, where the action-space of the firms is determined by the quantities that they produce.

It can be shown that a unique Nash equilibrium exists, that the total quantity produced by the cartel in equilibrium equals the quantity produced by each individual firm, and that this quantity is $q = \frac{1-c}{n-k+2}$. Hence, the profit $\pi^I(k)$ of an independent firm and the profit $\pi^C(k)$ *per member* of a firm in the cartel are

$$\pi^I(k) = \frac{(1-c)^2}{(n-k+2)^2} \quad \text{and} \quad \pi^C(k) = \frac{1}{k}\pi^I(k).$$

These profit functions indeed show an incentive for free riding, i.e. the model satisfies **A2**, which is explained by the fact that the cartel produces as if it were a single firm, but shares the profit among all cartel members. Also independent firms receive higher profits when larger cartels form. Hence, the model also satisfies **A1**.

A simple calculation shows that the profit of a cartel member is maximal if the cartel consists of all firms. That is, for this example, $\hat{k} = n$. Further, it can be checked that $\bar{k} = \lfloor 3 + n - 2\sqrt{n} \rfloor$. Hence, the dynamic analysis suggests that, at a given time, at most

$\lceil 2\sqrt{n} - 2 \rceil$ firms are independent. On the other hand, it can be checked that $k^* = 1$ for all $n \neq 2$ and $k^* = 2$ if $n = 2$. Hence, stability analysis predicts the degenerate cartel of size 1 in all cases except a duopoly, for which it predicts a strongly stable cartel.

4 Concluding remarks

In our model a firm's profit is characterized by positive externalities of cartel formation (Assumption **A1**) and there is a free-rider incentive for independent firms (Assumption **A2**). These assumptions are fulfilled by many oligopolistic models, e.g. Cournot, Bertrand and price-leadership models. We restrict our study to the situation in which all firms are ex-ante identical and at most one cartel can form. The same situation has also been studied by others, but our approach differs in that we use a dynamic model to study and predict likely changes in the cartel, as opposed to investigating stability criteria. The dynamic approach suggests that, at any time, a larger cartel will be observed than predicted by stability analysis. As in D' Aspremont et al. (1983) and Thoron (1998), we assume that firms are myopic when they take a decision, i.e. they take into account the immediate consequences of their actions, but not the possible consequences of further actions provoked by the initial action. For a model à la D' Aspremont et al. (1983) with forward looking players see Diamantoudi (2005).

Assumptions **A1** and **A2** are descriptive of the economic setting in which the cartel operates. If these assumptions are not fulfilled, for example if cost-reducing alliances induce negative externalities on other firms, very different results should be expected. The assumptions that all firms are ex-ante identical and that at most one cartel can exist at a given time are restrictions of a different nature. They are introduced to simplify the analysis, and it is sensible to ask whether similar results would be obtained if these restrictions were relaxed. In Olaizola (2005), a generalization of the model to a heterogeneous group of players is analyzed. The results show that there is still a unique EDS, as long as the players are not "too heterogeneous." Otherwise, more than one EDS may exist, even when there are no strongly stable cartels. For a model à la D' Aspremont et al. (1983) with heterogeneous firms see Donsimoni (1985) and Donsimoni et al. (1986).

In the literature on cartel formation it is usually assumed that no more than one cartel can exist at a given time. An overview of games with externalities across coalitions in the context of general coalition formation can be found in Yi (1997) (see also Bloch 1996; Ray and Vohra 1997). In an approach similar to Thoron's, Espinosa and Inarra (2000) present a non-cooperative view of the case of multiple cartels. The analysis of a dynamic model that allows several cartels to exist simultaneously is complicated by the fact that in such a model robust moves do not necessarily exist when profitable moves exist. Further research is necessary in order to generalize our results, but some preliminary investigations that deal with this issue can be found in Olaizola (2005).

References

Bernheim BD, Peleg B, Whinston MD (1987) Coalition-proof nash equilibria, i concepts. J Econ Theor 42:1–12

Bernheim BD, Whinston MD (1987) Coalition-proof nash equilibria, II applications. J Econ Theor 42: 13–29

Bloch F (1996) Sequential formation of coalitions in games with fixed payoff division. Games Econ Behav 14:90–123

D' Aspremont CA, Jacquemin A, Gabszewicz JJ, Weymark J (1983) On the stability of collusive price leadership. Can J Econ 16:17–25

Deneckere R, Davidson C (1985) Incentives to form coalitions with bertrand competition. RAND J Econ 16:473–486

Diamantoudi E (2005) Stable cartels revisited. Econ Theor 26(4):907–921

Donsimoni MP (1985) Stable heterogeneous cartels. Inter J Ind Organ 3:451–467

Donsimoni MP, Economides N, Polemarchakis H (1986) Stable cartels. Int Econ Rev 27:317–327

Espinosa , Inarra E (2000) Von Neumann and morgenstern stable sets in a cournot merger system. Int Game Theor Rev 2(1):29–45

Kalai E, Pazner EA, Schmeidler D (1976) Admissible outcomes of social bargaining processes as collective choice correspondence. J Econ Theor 63:299–325

Olaizola N (2005) Strategic choices: formation of coalitions and internal organization of firms. Absorbing and generalized stable sets. Ph.D. thesis, University of the Basque Country

Ray D, Vohra R (1997) Equilibrium binding agreements. J Econ Theor 73:30–78

Salant SW, Switzer S, Reynolds RJ (1983) Losses from horizontal merger: the effects of an exogenous change in industry structure on cournot-nash equilibrium. Q J Econ 98(2):185–199

Selten R (1973) A simple model of imperfect competition, where 4 are few and 6 are many. Int J Game Theor 2:141–201

Schwartz T (1974) Notes on the abstract theory of collective choice. School of Urban and Public Affairs

Shenoy L (1979) On coalition formation: a game theoretical approach. Int J Game Theor 8:133–164

Stigler JG (1950) Monopoly and oligopoly by merger. Am Econ Rev 40:23–34

Thoron S (1998) Formation of a coalition-proof stable cartel. Can J Econ 31:63–76

von Neumann J, Morgenstern O (1944) Theory of games and economic behavior. Princeton University Press, Princeton

Yi SS (1997) Stable coalition structures with externalities. Games Econ Behav 20:201–237