On the Stability of Collusion*

Claude d'Aspremont and Jean Jaskold Gabszewicz†

1 Introduction

The literature on oligopolistic competition abounds with various implicit statements about the 'stability' of collusive arrangements. A well known example is provided by comments on price arrangements between the sellers in a given industry. It is asserted that any oligopolistic configuration must be unstable with respect to monopolistic collusion: 'the combined profits of the entire set of firms in an industry are maximised when they act together as a monopolist, and the result holds for any number of firms' (Stigler, 1950, p. 24). At the same time however, it is recognized that 'when the group of firms agrees to fix and abide by a price approaching monopoly levels, strong incentives are created for individual members to chisel – that is, to increase their profits by undercutting the fixed price slightly, gaining additional orders at a price that still exceeds marginal cost' (Scherer, 1980, p. 171). On the other hand, in the recent literature on the core of an exchange market, it has been shown that, sometimes, monopolistic collusion can be disadvantageous to the traders involved when compared to the competitive outcome (Aumann, 1973). As for the collusive price-leadership model, it is stressed that the outsiders of a merger agreement may be better off than the insiders:

*Reprinted from *New Developments in the Analysis of Market Structure*, Joseph E. Stiglitz and G. Frank Mathewson (eds.), Cambridge (Mass.), MIT Press, 243-264, 1986. We are grateful to A. Shaked for his comments and suggestions.

†Center for Operations Research and Econometrics, Louvain-la-Neuve

> the major difficulty in forming a merger is that it is more profitable to be outside a merger than to be a participant. The outsider sells at the same price but at the much larger output at which marginal cost equals price. Hence the promoter of a merger is likely to receive much encouragement from each firm – almost every encouragement, in fact, except participation (Stigler, 1950).

All these statements reveal implicitly the existence of two opposite tendencies in the coordination of group decision processes. The first tendency comes from the recognition by individual participants of the collective advantage of coordination. The other comes from the evidence that this coordination may not be immune against advantageous recontracting by some of these participants. However, it seems that no attempt has been made to provide a unified theoretical framework for analysing the effectiveness of coordination. As a first step in this direction, this chapter aims to define some alternative stability concepts for collusive arrangements and to illustrate, in the light of these concepts, how collusive coordination can be effective in two different economic contexts.

To illustrate the first stability concept we deal with the problem of imputing the output resulting from a productive activity, between the owners of the factors involved in this activity. We assume that some of the factor owners collude so as to orient the choice of the imputation to their own advantage. We show that, under a constant returns-to-scale technology, they cannot succeed in this endeavour if the imputation is chosen in the core of the market.

The second illustration is devoted to an alternative concept of stability which we use in the context of the price-leadership model. A (dominant) group of firms collude to quote a price which a competitive fringe accepts. We examine the extent to which, with such an arrangement, it is to the advantage of an individual member of the group to move outside it, or of a single member of the fringe to join the group. We show, using an example, that there is always a division of the firms between a cartel and a competitive fringe where no such individual advantage exists.

2 The stability of collusive arrangements: an abstract framework

The departure point of our analysis is a given list of individual decision-makers involved in a group decision process – where they have partially parallel, and partially conflicting, individual interests. The collusive mechanism operates when these individuals, instead of acting independently, tend to guide the results of the process, by forming *groups*. These groups delegate to a single decision unit the task of representing their economic interests. The formal effect of this mechanism consists, first, in reducing the initial set of individual decision-makers to a new set of decision units, which now consists of the groups formed through collusion. Second, the bias in the outcome of the decision process is introduced by the substitution for its expected outcome (when the original decision units are acting independently) of a new outcome, or a new class of outcomes: namely, those which are expected to prevail if groups are effectively acting in place of the individuals themselves. Given a set $N = \{\cdots, i, \cdots, n\}$ of individual decision-makers, we call a *collusive scheme*, denoted C, a partition of N into subsets N_k, with all individuals in N_k acting in unison; N_k is called a *group*. The simplest example of a collusive scheme is the finest partition of N, namely

$$C_0 \stackrel{\text{def}}{=} \{(1), \cdots, \{i\}, \cdots, \{n\}\};$$

C_0 is the scheme resulting from a collusive process where no collusion has been successful, or even initiated; accordingly, we call C_0 the *disagreement collusive scheme.* At the opposite extreme, if a single decision unit is substituted for the n initial ones, total agreement is reached through the collusive process and we obtain $\overline{C}$ defined by

$$\overline{C} = \{N\}.$$

We call $\overline{C}$ the *total agreement collusive scheme.* An economic example of a total agreement collusive scheme is of course provided by a set of ostensibly independent firms which collude to determine jointly the output or the price which is to prevail in a given industry. All intermediate forms of collusion, like collusive oligopolies or 'leader-ship' situations, consisting of one 'big' collusive cartel and many small independent firms, can easily be captured by a particular collusive scheme. In the latter case, for instance, if the cartel groups firms 1 to k, and the 'competitive fringe' consists of firms $\{k+1, \cdots, n\}$, the corresponding collusive scheme is $C = \{\{1, \cdots, k\}, \{k+1\}, \cdots, \{n\}\}$.

Given a particular group decision process, either institutional factors, or the very nature of the process, impose self-evident restrictions on the class of collusive schemes which are feasible.[1] The most extreme example is the case of an 'anti-trust law' which would forbid any explicit or implicit oligopolistic coordination between the firms in a given industry: no other collusive scheme than C_0 is then feasible. Other examples of such restrictions are provided by negotiations involving trade unions, professional associations, bidders at auctions, syndicates of property owners, etc. Such 'groups' share the property that the individual members of the group are all of the same 'type': a trade union includes only workers, and no other type of economic agent. Consequently, the very nature of the decision process excludes any collusive scheme which would embody a group consisting of individuals of different types. Let us formally capture this idea. Let N be subdivided into m disjoint subsets, or *types*, i.e. $N = \{N_1, \cdots, N_i, \cdots, N_m\}$, with all individuals ij in N_i identical. Let A_i be a subset of $N_i, i = 1, \cdots, m\}$. We call A_i *the syndicate of type* i and $A = \{A_1, \cdots, A_i, \cdots, A_m\}$ *a syndicate structure*.[2] Thus a syndicate is a group involving only identical decision units. A *syndicate collusive scheme* is defined by

$$\begin{aligned} C &= \{A_1, \cdots, A_i, \cdots, A_m, \\ &\quad \{1j\}_{1j \in N_1 \setminus A_1}, \cdots, \{ij\}_{ij \in N_i \setminus A_i}, \cdots, \{mj\}_{mj \in N_m \setminus A_m}\}, \end{aligned}$$

where the notation $\{ij\}_{ij \in N_i \setminus A_i}$ represents the set of singletons of $N_i \setminus A_i$.

Intuitively, a syndicate collusive scheme is simply a partition of the m types of decision-makers into m syndicates of different types and the 'isolated' individuals of the various types who are *not* members of their corresponding syndicates. To provide an economic illustration, which will be developed below, consider the group decision process, involving a given set N_1 of

[1] This kind of admissibility restriction on the set of feasible collusive schemes is also at the root of the Ψ-stability concept (see Luce and Raiffa, 1957, Chapter 10).

[2] This terminology is borrowed from J. Jaskold Gabszewicz and J.H. Drèze (1971).

workers and a given set (N_2) of capital owners. The process consists of choosing collectively an imputation of the output resulting from their joint activity (imputation of social output). If, in this process, a trade union is formed (call it A_1), and if some capital owners collude (say, in A_2) we obtain a syndicate collusive scheme embodying: (i) the trade union A_1; (ii) the 'unorganised' workers; (iii) the syndicate of capitalists A_2; (iv) the 'unorganised' capital owners. A simpler illustration is also provided by the leadership model, described above, when all the n firms are identical. Then, the number of types reduces to one, and the collusive scheme $\{A, \{i\}_{\{i \in N \setminus A\}}\}$, where A denotes the cartel (or the syndicate) and $\{i\}_{i \in N \setminus A}$, the competitive fringe, is a syndicate collusive scheme.

As was stated earlier, collusion is intended to guide the collective decision process, by enforcing outcomes which would not otherwise prevail. The outcome, or class of outcomes, to be taken into consideration in a given collusive scheme, depends on the nature of the collective decision process. Either the process is designed to bring a bout a non-cooperative outcome, like a Nash equilibrium among the 'players', or it envisages a leadership solution analogous to a Stackelberg point. Or a co-operative outcome must be expected, like an imputation in the core. As soon as a particular solution concept is selected to describe the outcome, or class of outcomes, of a given decision process, it is easy to specify how collusion guides the mechanism of collective choice. To illustrate, let us consider again the example of an industry consisting initially of a large number of identical firms. If no collusion occurs, i.e. if the disagreement scheme C_0 is realised, then it is natural to take as outcome relative to this scheme the vector of competitive payoffs. By contrast, if all the firms collude in a single cartel, i.e. if $C = \overline{C}$, we may take as the class of outcomes relative to $\overline{C}$ the set of all possible imputations of the monopoly profit between the firms, or its uniform imputation. Now, suppose this set of firms splits' into two cartels, C_1 and C_2, then $C = \{C_1, C_2\}$. We can then pick as outcomes relative to C the set of payoffs corresponding to the Nash equilibrium pairs of strategies of the game with two players C_1 and C_2 (recall that all firms in C_1 and C_2 act in unison). In the price leadership model, for the collusive scheme $\{\{1, \cdots, k\}, \{k+1\}, \cdots, \{n\}\}$, we may take as the outcome the uniform imputation of the price leadership profit inside $\{1, \cdots, k\}$ and, for each firm in the competitive fringe, the profit level resulting from profit maximisation at the price chosen by the cartel. If we want to consider cooperative outcomes, it may be natural to use the concept of core.[3] Given a set N of decision units, the core is usually defined as the set of imputations feasible for the 'grand' coalition N, which cannot be blocked by any coalition S, where a coalition S is simply defined as a subset of N. However if a collusive scheme $C = \{C_1, \cdots, C_h, \cdots, C_m\}$ that is different from C_0 becomes effective, any coalition which would include a *proper* subset of any element C_h of the corresponding partition can no longer be considered: such a coalition would 'split' an indivisible decision unit, and is thereby forbidden. Consequently, to each collusive scheme, there corresponds the core relative to C, namely the set of imputations which is not blocked by any *permissible* coalition, i.e. a coalition S in the set $\{S \mid \forall\, h \text{ either } S \cap C_h = \emptyset, \text{ or } S \cap C_h = C_h\}$. With this terminology, the core relative to C_0 is the core as usually defined, where *any* coalition is permissible.

All the above examples show that the cohesiveness of collusive agreements must be evaluated by the outcome, or class of outcomes, which must be expected from a particular collusive scheme. Moreover, to achieve the cohesiveness of the groups observed in this collusive scheme, collusion must bring about outcomes which are, loosely speaking, 'advantageous' to their members: otherwise they might, for instance, be tempted to 'cheat' by secretly recontracting with individuals outside their own group, and thereby breaking the effectiveness of the collusive scheme. Intuitively, one can think of two sorts of requirement of the resulting outcome, which may protect the stability of a given collusive scheme. The first sort refers to comparisons between individual outcomes for a *given* collusive scheme $C = \{N_1, \cdots, N_i, \cdots, N_m\}$, in particular, individual outcomes of N_h-members compared with individual outcomes of N_k-members, $k \neq h$. Such comparisons are particularly relevant in the context of a syndicate structure $\{A_1, \cdots, A_i, \cdots, A_m\}, A_i \subset N_i$, for outcomes across the *same* type. Then comparisons are made, for a given collusive scheme, between the payoffs received by the members of the syndicate, A_i, and the payoffs of those who remained 'unorganised', and are of the same type (in $N_i \setminus A_i$). The second sort refers to a

[3]It is clear that the core is only one possibility. Other solution concepts like the bargaining set, or the Shapley value, might be more appropriate in other contexts.

comparison of the individual outcomes accruing to the decision makers if a collusive scheme C is formed, with the outcomes they would face if some alternative collusive scheme were formed.

These two types of comparisons lead to the following stability concepts for collusive schemes (C) relative to a given class of outcomes. Denote by

$$\Pi(C) = \{\pi(C) \mid \pi(C) = (\pi_1(C), \cdots, \pi_i(C), \cdots, \pi_n(C))\}$$

the class of outcomes associated with the collusive scheme C, where $\pi_i(C)$ denotes the payoff to individual i if outcome $\pi(C)$ is selected.[4] The first definition refers to the stability of syndicate collusive schemes. Let N be subdivided into m types, i.e. $N = \{N_1, \cdots, N_i, \cdots, N_m\}$; let $A = \{A_1, \cdots, A_i, \cdots, A_m\}$ be a given syndicate structure, and C be a particular syndicate collusive scheme.

2.1 Internal stability of syndicates

The syndicate collusive scheme C is internally stable if

$$\forall\, \pi(C) \in \Pi(C),\ \forall\, ij \in A_i \text{ and } ik \in N_i \setminus A_i, \pi_{ij}(C) \geq \pi_{ik}(C),$$

with strict inequality for some ij and ik.

If a syndicate collusive scheme is internally stable, there may be no tendency for the syndicate members to break the agreement which binds them to the syndicate i: they enjoy a more favourable treatment than do their similar companions, who remained outside the syndicate.

By contrast we may consider stability concepts for a given collusive scheme which refer explicitly to the outcomes received under one, or several, alternative collusive schemes. From this viewpoint, our first definition is a strong stability concept, but relative to a given class of collusive schemes.

[4] $\Pi(\cdot)$ is defined here as a *correspondence*; in some applications, Π is a function, i.e. the set $\Pi(C)$ reduces to a single element $\pi(C)$. This remark is important, since most of the stability concepts proposed below may be given alternative (stronger or weaker) definitions based on the multiplicity of outcomes. For simplicity of exposition, we do not consider all such variations.

2.2 Unanimous stability

Given a class $\{C\}$ of collusive schemes, a particular scheme $\widetilde{C}$ is *unanimously stable relative to* $\{C\}$ if $\widetilde{C} \in \{C\}$ and if

$$\begin{aligned}
&\forall\, \pi(\widetilde{C}) \in \Pi(\widetilde{C}),\ \forall\, C \in \{C\},\ \forall\, \pi(C) \in \Pi(C),\\
&\forall\, i \in N,\ \pi_i(\widetilde{C}) \geq \pi_i(C),
\end{aligned}$$

with strict inequality for at least some i.

The concept of the 'unanimous stability' of a collusive scheme $\widetilde{C}$ in a given class of schemes thus asserts that the payoffs to the agents under this collusive scheme are at least as high as for any other in the class, and unambiguously higher for at least one agent. Generally, there exists no collusive scheme which is unanimously stable in the class of all collusive schemes. Collusive schemes may, however, be stable in restricted, but 'natural', reference classes of collusive schemes. Given a collusive scheme C, consider for instance the class $\{C, C_0\}$. If C is stable in this binary class, the payoffs to the agents, if C forms, dominate their corresponding payoffs under the disagreement scheme C_0. This property motivates the following definition.

A collusive scheme C is *unanimously stable relative to the disagreement scheme* C_0 if it is unanimously stable in the class $\{C, C_0\}$. By analogy we define: a collusive scheme C is unanimously stable relative to the agreement scheme $\overline{C}$ if it is unanimously stable in the class $\{C, \overline{C}\}$.

Given a collusive scheme $\widetilde{C} = \{\widetilde{C}_1, \cdots, \widetilde{C}_n, \cdots, \widetilde{C}_m\}$, another 'natural' reference class obtains by considering *individual moves* across the elements of the partition $\widetilde{C}$. To be more precise, assume that some individual i leaves an element $\widetilde{C}_h$ to join $\widetilde{C}_k$. Then a new collusive scheme obtains, namely $C = \{\widetilde{C}_1, \cdots, \widetilde{C}_h \setminus \{i\}, \cdots, \widetilde{C}_k \cup \{i\}, \cdots, \widetilde{C}_m\}$. It is, loosely speaking, a 'neighbour' of $\widetilde{C}$.

More formally, consider a class $\{C\}$ of collusive schemes and a given collusive scheme $\widetilde{C} = \{\widetilde{C}_1, \cdots, \widetilde{C}_h, \cdots, \widetilde{C}_k\}$ in $\{C\}$; we may then define $V_i(\widetilde{C})$ as the set of all collusive schemes $C =$

$\{C_1, \cdots, C_h, \cdots, C_{k'}\}$ in $\{C\}$ such that, for every h, we have either

$$\begin{array}{lll} & C_h = \widetilde{C}_m, & \text{for some } \widetilde{C}_m \text{ in } \widetilde{C}, \\ \text{or} & C_h = \widetilde{C}_m \setminus \{i\}, & \text{with } \widetilde{C}_m \text{ the element of } \widetilde{C} \text{ that } i \text{ leaves}, \\ \text{or} & C_h = \widetilde{C}_m \cup \{i\}, & \text{with } \widetilde{C}_m \text{ the element of } \widetilde{C} \text{ that } i \text{ joins}, \\ \text{or} & C_h = \{i\}. & \end{array}$$

If we consider, for some individual i, a collusive scheme C in $V_i(\widetilde{C})$, we see that there is only an individual, unilateral change between $\widetilde{C}$ and C. It seems natural to introduce the following concept of stability for a given collusive scheme relative to its 'neighbours'.

2.3 Individual stability

Consider a class $\{C\}$ of collusive schemes. A collusive scheme $\widetilde{C}$ is individually stable relative to $\{C\}$ if, for some $\pi \in \Pi(\widetilde{C})$, there is no individual i and no collusive scheme $C \in V_i(\widetilde{C})$ such that

$$\pi_i(C) > \pi(\widetilde{C}), \text{ for some } \pi(C) \in \Pi(C).$$

Under a collusive scheme $(\widetilde{C})$ which is individually stable, no individual move is desired. It is clear that, because Π is a correspondence, several alternative definitions can be given. However, where Π is a function, then this concept, which involves only individual unilateral moves, may be seen as a Nash equilibrium in a non-cooperative game. In that game, the choice by every player of some strategy would determine a particular collusive scheme C, where the resulting payoffs would be given by $\Pi(C)$. Hence, a 'strong equilibrium' notion, for which no subgroup could find an advantageous unilateral move, could also be considered. With this framework, that would amount to comparing collusive schemes which are not necessarily 'neighbours'. However, we shall not introduce these alternative concepts in this chapter.[5]

To illustrate individual stability, consider n identical firms, with zero production cost, facing a market-demand function $P(q) = 1 - q$. If the n firms collude, then $\pi_i(\overline{C}) = 1/4n$. Assume k firms, say firms $\{1, \cdots, k\}$ leave the cartel, resulting in the formation of two cartels, then the

[5]For such considerations, in the context of an extension of the Shapley value, see Hart and Kurz (1981).

new collusive scheme is $C = \{\{1, \cdots, k\}, \{k+1, \cdots, n\}\}$. If both cartels act non-cooperatively, we obtain the Cournot-Nash equilibrium payoff for each firm:

$$\begin{aligned} \pi_i(C) &= \frac{1}{9k} && 1 \leq i \leq k, \\ \pi_i(C) &= \frac{1}{9(n-k)} && k+1 \leq i \leq n. \end{aligned}$$

Putting $k = 1$, shows immediately that $\overline{C}$ is not individually stable for $n \geq 3$. For such cases it is easily seen that the individually-stable collusive schemes are such that

$$k \geq \frac{n-1}{2} \text{ and } (n-k) \geq \frac{n-1}{2}.$$

Another illustration of individual stability can be obtained in the context of the theory of local public goods as formulated by Tiebout (1956) and others.[6] Consider a set of individuals who have to be 'allocated' among a number of distinct 'communities'. The resulting partition would create an individually-stable collusive scheme, whenever no individual would gain from moving from his assigned community to another, taking into account the adjustment in the local provision of public goods.

Equipped with the above framework, we may now turn to detailed illustration of our stability concepts.

3 Two illustrations

3.1 Collusion and the imputation of social output

The first illustration is devoted to an application of our concepts of stability in the context of a cooperative decision process: the solution concept selected to describe the outcome from a given collusive scheme C is the core, $\Pi(C)$. This example is borrowed from Hansen and Gabszewicz (1972). We consider an economy in which a single output is produced under a constant-returns-to-scale, differentiable production function $F(z_1, z_2)$, where z_1 is labour and z_2 capital. Labour (or capital) is distributed among r labour owners (r capital owners), each labourer (capitalist)

[6]See, for example, Westhoff (1977).

owning exactly one single unit of labour (or capital). We denote by N_r the set of factor owners; thus $N_r = N_{1r} \cup N_{2r}$, with $N_{1r} = \{11, \cdots, 1j, \cdots, 1r\}$ and $N_{2r} = \{21, \cdots, 2j, \cdots, 2r\}$. We normalise the production function (F) in such a way that $F(1,1) = 1$. Total output is thus equal to $F(r,r) = rF(1,1) = r$. An *imputation* is a $2r$-tuple of numbers $\pi = (\pi_{11}, \cdots, \pi_{1r}; \pi_{21}, \cdots, \pi_{2r})$, with $\sum_{i=1}^{2} \sum_{j=1}^{r} \pi_{ij} = r$. A *coalition* is a subset of N_r. The aggregate factor-endowment of a coalition S is equal to $|S \cap N_{1r}|$ units of labour and $|S \cap N_{2r}|$ units of capital.[7] Accordingly, a coalition S can produce, by its own means,

$$F(|S \cap N_{1r}|, |S \cap N_{2r}|) \stackrel{\text{def}}{=} F(S).$$

Consider then a proposed imputation π of the total output r among the factor owners. If, for some coalition S, $F(S) > \sum_{ij \in S} \pi_{ij}$, then the coalition S *blocks* the proposed imputation. The *core* is the set of all unblocked imputations.

We denote a particular collusive scheme by C^r if there are r factor owners of each type. If no collusion takes place, i.e. if $C^r = C_0^r$, then *any* coalition S which is a subset of N_r may form. However, if a syndicate A_{1r} (or A_{2r}) forms among the labour owners (capital owners), then the syndicate collusive scheme $C^r = \{A_{1r}, A_{2r}; \{1j\}_{1j \in N_r \setminus A_{1r}}, \{2j\}_{2j \in N_r \setminus A_{2r}}\}$ obtains, and any coalition (S) which would include a proper subset of A_{1r} or A_{2r} is forbidden. The class of *permissible* coalitions is thus reduced to the set:

$$\{S \mid \text{either } S \cap A_{ir} = \emptyset, \text{or } S \cap A_{ir} = A_{ir}, i = 1, 2\}.$$

The core, $\Pi(C^r)$, is the set of all imputations which are not blocked by any permissible coalition. Denote by k_1 (or k_2) the fraction of factor owners of type 1 (type 2) which are members of the syndicate $A_{1r}(A_{2r})$, i.e. $k_1 = |A_1|/r$ (or $k_2 = |A_{2r}|/r$). We study now the internal stability of the syndicate collusive scheme (C^r) under the assumptions that: (i) both k_1 and k_2 are unambiguously smaller than one (no type is 'fully organised'); (ii) the amount of output obtained by each syndicate A_{ir} is uniformly distributed among its members.

Proposition 1 *No syndicate collusive scheme C^r is internally stable.*

[7]For any set T, $|T|$ denotes the cardinal of T.

Proof: Let $\pi \in \Pi(C^r)$. First, let us show that, $\forall\, ij, ik \in N_{ir} \setminus A_{ir}, i = 1, 2$,

$$\pi_{ij} = \pi_{ik};$$

namely the core does not discriminate among the 'unorganised' factor owners of type i (notice that $N_{ir} \setminus A_{ir}$ is non-empty, by the assumption $k_i < 1$). Define

$$\pi_{i\,\min} = \min_{ij \in N_{ir} \setminus A_{ir}} \{\pi_{ij}\}; \pi_{i\,\max} = \max_{ij \in N_{ir} \setminus A_{ir}} \{\pi_{ij}\}.$$

First, we must have $\pi_{1\,\min} + \pi_{2\,\min} \geq 1$: otherwise the coalition S consisting of a single, unorganised factor owner of each type receiving $\pi_{i\,\min}$ would block the arrangement, since they can produce together one unit of output. (This coalition is permissible, since it does not intersect either A_{1r}, or A_{2r}). On the other hand, $\pi_{1\,\max} + \pi_{2\,\max} \leq 1$; otherwise the coalition consisting of all factor owners, except the labourer receiving $\pi_{1\,\max}$ and the capitalist receiving $\pi_{2\,\max}$, would block the imputation. This coalition can indeed produce $F(r-1, r-1) = r - 1$, and would receive *less* than $r - 1$, if $\pi_{1\,\max} + \pi_{2\,\max} > 1$. Furthermore this coalition is permissible, since it includes both A_{1r} and A_{2r}. Consequently,

$$(\pi_{1\,\max} - \pi_{1\,\min}) + (\pi_{2\,\max} - \pi_{2\,\min}) \leq 0,$$

which implies

$$\pi_{1\,\max} = \pi_{1\,\min} \text{ and } \pi_{2\,\max} = \pi_{2\,\min},$$

and the desired conclusion follows. Thus, any imputation π in $\Pi(C^r)$ is represented by 4 numbers, namely π_{11} (or π_{21}): the amount received by each syndicate member of A_1 (A_2) and π_{12} (π_{22}): the amount received by each 'unorganised' factor owner of type 1 (type 2). Now we prove that no syndicate collusive scheme C^r is internally stable. Suppose on the contrary that there exist C^r and $\pi \in \Pi(C^r)$, such that $\pi_{11} \geq \pi_{12}$, $\pi_{21} \geq \pi_{22}$, with strict inequality for at least one i, $i = \{1, 2\}$. Since π is an imputation, the equality

$$k_1 r \pi_{11} + (1 - k_1) r \pi_{12} + k_2 r \pi_{21} + (1 - k_2) r \pi_{22} = r$$

must hold. This equality may be rewritten as

$$k_1(\pi_{11} - \pi_{12}) + k_2(\pi_{21} - \pi_{22}) = 1 - (\pi_{12} + \pi_{22}) > 0,$$

where the last inequality follows from the fact that $\pi_{11} \geq \pi_{12}$, $\pi_{21} \geq \pi_{22}$, with strict inequality for at least one i. But then consider a coalition S consisting of a single unorganised worker and a single unorganised capitalist. This is a permissible coalition which can produce $F(S) = 1$, and this coalition receives $\pi_{12} + \pi_{22}$, and is strictly less than one, which is a contradiction. ■

According to the above proposition, for all values of r, no syndicate collusive process is internally stable. *A controrio*, any imputation in $\Pi(C^r)$ which would discriminate between syndicate members and non-members, must necessarily give privilege to the non-members of at least one type. Does this mean that the syndicates' cohesiveness must necessarily slacken off ? This is not certain: if the syndicate members are better treated at $\Pi(C^r)$ than they would be at $\Pi(C_0^r)$, i.e., in the core if no collusion exists, it might be advantageous for them to keep collusion running. This would be so despite the fact that some syndicate members are worse off than their similar companions who have remained outside the syndicate. But we are then led to study the *external* stability of C^r in the class (C^r, C_0^r). To proceed in that way, let us show that, *as r becomes large*, only the imputation which assigns their marginal product, i.e. the competitive payoff, to the unorganised factor owners can remain in the core $\Pi(C^r)$. To that end, denote by π^r an imputation in $\Pi(C^r)$ and by π_i^r the amount assigned by this imputation to the unorganised factor owners of type i, $i = 1, 2$.

Proposition 2 *Let $\pi^r \in \Pi(C^r)$ for all r. Then*

$$\lim_{r \to \infty} \{\pi_i^r\} = \left.\frac{\partial F}{\partial z_1}\right|_{(1,1)}.$$

Proof: First, for all r,

$$\pi_1^r + \pi_2^r \geq 1. \tag{1}$$

Otherwise, a coalition consisting of a single labourer and a single capitalist would block, contrary to the assumption that $\pi^r \in \Pi(C^r)$. Furthermore, consider for all r the coalition S consisting of all factor owners, except for a single, unorganised factor owner of type 1. This coalition can produce exactly $F(r-1, r)$, and is a permissible coalition. Furthermore, it receives an amount of output equal to $r - \pi_1^r$. Accordingly, if $r - \pi_1^r < F(r-1, r)$, the coalition S would block, contrary to the assumption that $\pi^r \in \Pi(C^r)$. Thus, for all r,

$$F(r-1,r) \leq r - \pi_1^r,$$

or by homogeneity of degree one of F,

$$\frac{F(1,1) - F\left(1-\frac{1}{r},1\right)}{1/r} \geq \pi_1^r.$$

Moving to the limit we would obtain

$$\left.\frac{\partial F}{\partial z_1}\right|_{(1,1)} \geq \lim_{r\to\infty}\{\pi_1^r\}. \tag{2}$$

By a perfectly symmetric argument, using the coalition consisting of all factor owners except a single unorganised factor owner of type 2, we would get

$$\left.\frac{\partial F}{\partial z_2}\right|_{(1,1)} \geq \lim_{r\to\infty}\{\pi_2^r\}. \tag{3}$$

Accordingly, by EUler's theorem,

$$1 \geq \left.\frac{\partial F}{\partial z_1}\right|_{(1,1)} + \frac{\partial F}{\partial z_2}_{(1,1)} \geq \lim_{r\to\infty}\{\pi_1^r + \pi_2^r\} \geq 1 \tag{4}$$

where the last inequality follows from (1). Consequently, combining (2), (3) and (4), we obtain

$$\left.\frac{\partial F}{\partial z_i}\right|_{(1,1)} = \lim_{r\to\infty}\{\pi_i^r\},\ i=1,2.$$

■

Thus, combining Propositions 1 and 2, we can conclude that, if r is large enough, *no syndicate collusive scheme C^r is unanimously stable relative to the disagreement scheme C_0^r.* Indeed, by Proposition 2, applied to the case $k_1 = k_2 = 0$, $\Pi(C_0^r)$ consists asymptotically of the sole competitive imputation. By the same proposition the unorganised factor owners approximately get their competitive payoff at an imputation in $\Pi(C^r)$ for an r which is large enough. Since, by Proposition 1, there is at least one syndicate whose members are worse off under this imputation than the corresponding unorganised factor owners (and since not all factor owners can receive more than their marginal product), the members of one syndicate must receive *less* than their marginal product, which is the amount they would receive in the collusive scheme C_0^r; the conclusion therefore follows.

3.2 An example of collusive price leadership

In this section, we return to the price-leadership model which was called to mind above. The price-leadership arrangement in an industry has been formulated (see Markham, 1951) as a particularly useful practice of tacit coordination among business firms. As emphasised by Scherer (1980), it appears to be compatible with most anti-trust legislation. The particular form of this kind of arrangement that we want to illustrate is the one where a dominant group of firms act as a 'leader' in the choice of the industry price, and the other firms are supposed to react competitively to this given price. The usual argument in favour of this type of arrangement is that the set of firms outside the dominant cartel forms a 'fringe of competitors' and that each one of them is too small individually to expect to have any influence on the price. However, the stability analysis must be carried out when the possibility that a firm may quit (or join) the dominant group is introduced. We examine this question through a simple example, which we developed in common with Jacquemin and Weymark (1981).[8]

Suppose we are given an industry for an homogeneous product, in which each of a set N of n firms faces the same total cost $C(q) = q^2/2$ for a quantity produced q. The firms are partitioned according to the following collusive scheme:

$$C_A = \{A, \{i\}_{i \in N \setminus A}\}, \quad |A| = a \geq 2,$$

where the set A represents the dominant cartel. We see that, by varying a, we may vary the size of the dominant cartel – and hence generate a whole class of collusive schemes $\{C_A; a \geq 2\}$. The stability analysis which will follow the presentation of this example, will be relative to this class. For simplicity, again, we assume that total demand at price p is $D(p) = n(1 - p)$. Furthermore, we suppose that for a price chosen by the cartel A the firms in $N \setminus A$ maximise their individual profit, taking this price as given (they equate marginal cost to price). On the other hand, the dominant cartel is assumed to choose the price (p) which maximises the joint

[8] For the application of this argument in a more general framework and for further analysis of the present example, the reader is referred to d'Aspremont *et al.* (1981). An extension of this example, where conjectural variations are introduced, is analysed in Donsimoni *et al.* (1981).

profit of its members given the quantity produced $S(p)$, at that price, by every firm in the competitive fringe. Hence, for every price p, they face only a 'residual' demand, which depends on the collusive scheme C_A, and is given by

$$D(p) - (n-a)S(p) = n - (2n-a)p,$$

since $S(p) = p$. Because all firms in the cartel are identical, maximising joint profit amounts to maximising profit for each firm in the cartel, which is

$$\pi_i(C_A, p) = \frac{p}{a}[n - (2n-a)p] - \frac{1}{2}\left[\frac{n-(2n-a)p}{a}\right]^2, \ i \in A.$$

The optimal price can easily be computed as being

$$p^* = \frac{2}{4 - \left(\frac{a}{n}\right)^2}.$$

Finally, the maximum profit for each firm outside the cartel (for a given price p) is simply $p^2/2$. So we get the following outcome function:

$$\begin{aligned}\pi_i(C_A) &= \frac{1}{2\left(4-\left(\frac{a}{n}\right)^2\right)}, \ i \in A \\ \pi_i(C_A) &= \frac{2}{\left[4-\left(\frac{a}{n}\right)^2\right]^2}, \ i \in N \setminus A.\end{aligned}$$

Under the disagreement scheme (C_0) the outcome $\pi(C_0)$ is assumed to be the competitive outcome, which is obtained by letting $a = 0$ in the above expressions (or $\pi_i(C_0) = \frac{1}{8}, i = 1, \cdots, n$).

Three points are to be stressed with respect to this outcome function:

(i) The profit of a firm outside the cartel is larger than the profit of a firm inside the cartel;

(ii) The profit of any firm (outside or inside the cartel) is increasing in a, the size of the cartel;

(iii) For $a \geq 2$, the profit of any firm is larger than the competitive profit.

These facts allow us to prove the following proposition, concerning cartel stability.

Proposition 3 *In the price-leadership example,* *(a) No (syndicate) collusive scheme C_A is internally stable.*

(b) There exists a collusive scheme $\widetilde{C}$, in the class $\{C_A\}$, which is individually stable relative to this class.

Proof: The first statement is a direct consequence of (i). For statement *(b)*, let us denote by (A_a) any cartel of size a, $a = 2, 3, \cdots, n$, and suppose that none of the corresponding collusive schemes is individually stable. Since, by (iii) no firm would gain by leaving a cartel A_2, we must have the result that one firm, and hence all firms, outside A_2, would gain by joining the cartel. But, then, by the same token, no firm would gain by leaving a cartel A_3. We must therefore have the result that some firm, and hence all firms, outside A_3 would gain by joining the cartel. Continuing in this fashion, we reach the conclusion that no firm would gain by leaving $A_n = N$, and hence that C_N is individually stable. The result follows by contradiction. ■

The proof given for statement *(b)* above is based on a general argument.[9] In fact, with the present, particular example, a sharper result may be demonstrated.

Proposition 4 *In the price-leadership example, and for all $n \geq 3$, the collusive schemes C_A with $a = 3$ are the only individually-stable collusive schemes relative to the class $\{C_A\}$.*

Proof: To be individually stable, a collusive scheme C_A should satisfy:

$$\frac{n^2}{2[4n^2 - (a+1)^2]} \leq \frac{2n^4}{(4n^2 - a^2)^2} \tag{5}$$

or, equivalently, $4n^2(a^2 - 2a - 1) \geq a^4$;

$$\frac{2n^4}{[4n^2 - (a-1)^2]^2} \leq \frac{n^2}{2[4n^2 - a^2]} \tag{6}$$

or, equivalently, $4n^2(a^2 - 4a + 2) \leq (a-1)^4$, where (5) should hold only if $a < n$ and (6) only if $a > 1$.

It is easy to verify that (5) holds for $a = 3$ and $n \geq 4$ and that (6) holds for $a = 3$ and $n \geq 3$. Hence if $a = 3$, C_A, is individually stable for all $n \geq 3$.

Moreover (5) is not satisfied for $n \geq 3$ and $a < 3$, since then $(a^2 - 2a - 1) < 0$. Finally, let us analyse (6) when $n \geq a \geq 4$. Since $(a-1)^4/(a^2 - 4a + 2)$ is positive and increasing in a (with $a \geq 4$), we have to show only that, for all $n \geq 4$, condition (6) is violated with $a = n$, i.e.,

$$4n^2 > \frac{(n-1)^4}{(n^2 - 4n + 2)}$$

or, equivalently,

$$n(3n^2 - 12n + 2) > \frac{1}{n} - 4.$$

Because $(3n^2 - 12n + 2) > 0$, for $n \geq 4$, this last inequality holds for all $n \geq 4$. ■

It is interesting to consider this result in the case where we allow n to be very large: that is the case for which the assumption concerning the reaction of the competitive fringe seems more reasonable. Indeed, in that case, for any dominant cartel A representing a sufficient proportion of the total industry (i.e., $a/n > 3/n$), a firm inside the cartel would gain by leaving.

This conclusion is to be compared with the statement of Proposition 3, asserting the non-existence of an internally-stable (syndicate) collusive scheme (C_A).[9] This statement was based on the fact that the profit of a firm outside the cartel was larger than the profit of a firm inside the cartel, for any collusive scheme (C_A). Similarly here, when n is large enough, the negative impact of one firm leaving the dominant cartel on the profit of every firm (and especially every outsider) is negligible in comparison to the advantage of leaving. This is so whenever the size of the given cartel is sufficient to maintain the proportion of dominant firms almost unchanged. Under such conditions, the violation of individual stability almost coincides with the violation of internal stability. On the other hand, consider the individually- stable collusive schemes, those C_A for which $a = 3$. Then, examining the individual profits where 3 is substituted for a, we see that those profits all converge to the competitive profit 1/8 when we let n grow. In other words, the advantage of price-leadership coordination vanishes when the number of firms becomes large.

[9] It is also to be compared with the assertions in Postlewaite and Roberts (1977).

4 Conclusion

In this chapter, we have presented, in a general form, some alternative stability concepts for collusive arrangements, based on two kinds of payoff comparison: 'internal' comparisons for a given collusive scheme, and 'external' comparisons to different collusive schemes. Of course, one can imagine another conceptual basis and, indeed, other criteria to explain the stability of particular schemes: the existence of threats, the value of commitments, etc. We think, however, that the present approach provides a unifying framework for existing contributions on the subject. Hopefully, it could stimulate new applications. One can think, for instance, of alternative situations or models where these concepts would fit, and for which stability properties would be established. In particular, interesting phenomena, like entry, can be captured, using our framework. In many market situations, indeed, it turns out that, even if 'at the start' a large number of individual decision makers are involved in the group decision process, only a small subset of them could be considered as 'active' decision units. That happens because the others have no interest in entering the market. More specifically, the partitioning between 'active' or 'non-active' individual decision units leads to stability considerations of the kind discussed above.[10]

An essential element in these considerations is the introduction of a two-stage sequence of moves by the individual decision units. At the first stage, their decisions result in the determination of a particular collusive scheme; at the second stage, other kinds of decisions are made taking into account the chosen collusive scheme. This sequential element, and the deterrent implications that it has for individual behaviour, is common to all the numerous analyses of different types of pre-emptive strategies that can be used in industrial competition. They are treated elsewhere in this volume.

[10] For instance, assume that n identical firms are candidates to enter the market, but that only a subset of k of them, $k < n$, can make a positive profit under some outcome function. The collusive scheme $\{\{1, \cdots, k\}, \{k+1\}, \cdots, \{n\}\}$ is individually stable since no excluded firm has an interest in entering, and no active firm has an interest in leaving. For an analysis of entry in this spirit see Selten and Güth (1982).

References

[1] Aumann, R.J. (1973). Disadvantageous monopolies. *Journal of Economic Theory*, 6, 1–11.

[2] d'Aspremont, C., Jacquemin, A., Gabszewicz, J. and J. Weymark (1981). On the stability of collusive price leadership. *Canadian Journal of Economics*, 16(1), 17–25.

[3] Donsimoni, M., Ekonomides, N.S. and H.M. Polemarchakis (1981). Stable cartels. Working Paper, 8112 ENSAE, France.

[4] Hansen, T. and J. Jaskold Gabszewicz (1972). Collusion of factor owners and distribution of social output. *Journal of Economic Theory*, 4, 1–18.

[5] Jaskold Gabszewicz, J. and J.H. Drèze (1971). Syndicates of traders in an exchange economy. In H.W. Kuhn and G. Szegö (eds.), *Differential Games and Related Topics*, Amsterdam, North-Holland, 399–414.

[6] Luce, R.D. and H. Raiffa (1957). *Games and Decisions*, New York, John Wiley.

[7] Markham, J.W. (1951). The nature and significance of price-leadership. *American Economic Review*, 41, 891-905.

[8] Postlewaite, A. and J. Roberts (1977). A note of the stability of large cartels. *Econometrica*, 45, 1877-1878.

[9] Scherer, F.M. (1980). *Industrial Market Structure and Economic Performance*, 2nd edn, New York, Rand McNally.

[10] Selten, R. and W. Güth (1982). Equilibrium point selection in a class of market entry games. In M. Deistler, E. Fürst and G. Schwödiauer (eds), *Games, Economic Dynamics, and Time Series Analysis*, Wurzburg-Vienna, Physica-Verlag.

[11] Stigler, G.J. (1950). Monopoly and oligopoly by merger. *American Economic Review, Papers and Proceedings*, 40, 23–34.

[12] Stigler, G.J. (1964). A theory of oligopoly. *The Journal of Political Economy*, 72, 44–61.

[13] Tiebout, C.M. (1956). A pure theory of local expenditures. *The Journal of Political Economy*, 64, 416–424.

[14] Westhoff, F. (1977). Existence of equilibria in economies with a local public good. *Journal of Economic Theory*, 14, 84–112.

Discussion of the Paper by d'Aspremont and Gabszewicz

In discussing this paper, *Shaked* made a number of comments. First, he pointed out that stability in this paper required that no individual player in the game be better off in any new equilibrium that might arise should that player move between syndicates, or from a syndicate to the fringe group. The payoff to player movement then depended critically on the solution concept used to describe the nature of the game after the movement had occurred. This solution concept should depend on the structure of the syndicate. For example, with many small syndicates, a non-cooperative concept would be appropriate, while if the structure consisted of a few large syndicates, a cooperative solution would be more appropriate. Shaked added that he understood the difficulties associated with attempting to resolve this problem. Next, Shaked commented that individual stability ignored the possibility that a group of players might jointly make themselves better off by leaving a syndicate. He wondered whether a stability concept could be usefully defined that was more closely related to the notion of a strong Nash equilibrium. Further, Shaked thought that some attention should be paid to the question of whether a syndicate would accept a new member, should someone decide to move. Finally, Shaked suggested that the concept of internal stability could best be applied when the agents in an economy were small. Otherwise, stability was best directed towards determining whether a coalition would disband completely if there was some better alternative available for the individual players.

Reinhard Selten proposed that the determination of cooperation within a game was an important issue that had not been analysed properly by game theorists. He praised this paper as a first step in this process. He suggested, however, that it might be inappropriate to apply cooperative solution concepts like the core to the second stage of a game when the analysis of stability in the first stage assumed non cooperation. Selten added that the title 'internal stability' should be changed to something like 'syndicate advantage'. The reason was that 'internal stability' was not a stability concept in the usual sense, so that this title was misleading.

Dasgupta responded to Shaked's suggestion by noting that strong Nash equilibria need not always exist. A stability concept based on that notion might be of limited usefulness, since a stable syndicate structure could never be found.

Jacquemin suggested that oligopolies worked very hard at devising practices that would make coordination of strategies easier. He questioned whether the framework of the paper would allow one to analyse these issues. *Gabszewicz* replied that the paper did not analyse these issues directly, but that these problems could be incorporated into the analysis by enlarging the strategy spaces open to firms, or specifying the rules of the game more precisely.

Stiglitz proposed that players in an economic game were not intrinsically interested in co-operation, rather they cooperated in non-cooperative behaviour. This idea, he suggested, had been captured in the literature in repeated games, where cooperative solutions were achieved as the non-cooperative equilibria of games when players used appropriate threat strategies. Stiglitz wondered why d'Aspremont and Gabszewicz had taken the approach to the problem embodied in their paper.

Gabszewicz responded that he and d'Aspremont had wished to exploit the kind of structure suggested by the entry-deterrence literature, where the pre-entry equilibrium was conditional on the knowledge of players and that the post-entry equilibrium occurred as a result of their own actions. In the d'Aspremont-Gabszewicz paper, players foresaw the equilibrium that would occur after they deviated from the cooperative solution. This concept was in the entry-deterrence spirit since the deviator was similar to the first mover in any entry-deterrence game.

Stiglitz suggested that this specification did not prevent members of the cartel from making threats against firms who deviated. In repeated games, he suggested, these threats could even be made credible. For example, supergame strategies might involve commitments to punish firms who did not punish firms who violated agreements. In infinitely repeated games, such strategies could become Nash equilibria.

Gabszewicz responded that he had recently written a paper with Claude d'Aspremont looking at a related story. In this paper, they began with a continuum of firms all having unit capacity. If the entire market formed a cartel, then these firms would set a monopoly price. Of course, it was always possible for some small subset of all firms to deviate from this monopoly solution. The question then became whether the reduction in the cartels' profits would be large enough to cause them to respond. Since firms had limited capacity, however, their supply would not depress the cartel price very much. Under some conditions, the cartel would not find it profitable to respond to this deviation.

Curtis Eaton suggested that any coalition structure that was individually stable in the sense of the d'Aspremont-Gabszewicz paper, would be stable with respect to threats.

Spence suggested that the best way to view the problem of cartel stability was to consider competition on two levels – competition in the formation of groups and competition in the determination of prices once these groups had been formed. The ways that firms interacted on the first of these competitive levels was unspecified, and as a consequence, allowed threat strategies. Individual stability, then, constituted a necessary but not sufficient condition for equilibrium in the first level of competition.

Selten referred to the example given in the paper of the dominant-firm competitive-fringe and asked why the only stable syndicate structure involved only three firms, and in particular, why this number should be independent of the number of firms in the market. d'Aspremont responded that the example was from a more general paper with Jacquemin and Weymark (1982) where the existence of a stable cartel was established. There was no reason to expect the number of firms in this cartel to be 3, or to be independent of the number of firms in the market. Shaked explained that this particular result flowed from a specification of market demand which was linear in prices and multiplicative in the number of consumers.

References

[1] d'Aspremont, C., Jacquemin, A., Gabszewicz, J. and J. Weymark (1982). On the stability of dominant cartels. *Canadian Journal of Economics*, 14(1), 17–25.