Cartel Stability Under Capacity Constraints:
The Traditional View Restored

by
Volker Nocke*
STICERD, London School of Economics

Abstract

The existence of a negative relationship between cartel stability and the level of excess capacity in an industry has for a long time been the dominant view in the traditional IO literature. Recent supergame-theoretic contributions (e.g. Brock and Scheinkman 1985) appear to show that this view is ill-founded. Focussing on the issue of enforcement of cartel rules ("incentive constraints"), however, this literature completely ignores firms' "participation constraints". Reverting the focus of attention, the present paper restores the traditional view: large cartels will not be sustainable in periods of high excess capacity (low demand). In contrast to the supergame-theoretic literature, it predicts a negative relationship between excess capacity and the collusive price.

Keywords: collusion, excess capacity, business cycle, cartel stability
JEL Classifications: L13, D43

1 Introduction

For a long time, economists have believed that cartels will tend to break down under the pressure of low demand and high excess capacity. In the second edition of his famous work, "Industrial Market Structure and Economic Performance", Scherer (1980) states:

> There is evidence that industries characterized by high overhead costs are particularly susceptible to pricing discipline breakdowns when a cyclical or secular decline in demand forces member firms to operate well below designed plant capacity.[1]

Examples of industries mentioned by Scherer include chemicals, steel, aluminium, cement and mining. According to this view, we should thus expect to find a negative correlation between the level of excess capacity and the degree of collusion.[2] This, in turn, should reinforce the tendency of prices to be high during booms and low during recessions.

However, recent supergame-theoretic contributions appear to show that the traditional view is theoretically ill-founded. For instance, Brock and Scheinkman (1985) analyse an infinitely repeated price-setting game, in which each firm faces the same capacity constraint. They investigate the situation where all firms in the industry try to sustain "full collusion", i.e. the monopoly price, by the threat of an infinite Nash reversion in case of deviation. In particular, Brock and Scheinkman analyse how the threshold discount factor, above which the monopoly price can be sustained in equilibrium, varies with the level of capacity in the industry. They show that an increase in the industry level of capacity has two opposing effects. On the one hand, it will make cheating more profitable in terms of current profits; on the other, it implies that a more severe punishment can be inflicted on a deviating firm. As a result, the threshold discount factor is generally a nonmonotonic function of industry capacity. Therefore, the model does not give any sharp predictions as to the relationship between the level of capacity (or demand) and the equilibrium price under collusion.

For all its merits, the following two questions arise regarding such a supergame-theoretic approach. First, in supergames, there exists a continuum of more or less collusive equilibria, among which the literature usually selects a particular one. However, there may be other, equally plausible, equilibria with quite different properties.[3] Second, the supergame-theoretic approach to collusion focusses on the problem of enforcement of

---

[1] See Scherer (1980), page 206.

[2] See Phlips (1995), page 152, for another statement of this traditional view.

[3] In particular, the literature mostly confines attention to the "best" symmetric subgame perfect equilibrium, which requires firms to inflict the worst possible punishment on any deviator. The well-documented cases of price wars do not appear to support the existence of such severe punishments. See, for instance, Genesove and Mullin (1998) on the Sugar Institute, and Levenstein (1997) on the Bromine Cartel.

collusive behaviour, that is, on firms' "incentive constraints". What this approach leaves out, are firms' "participation constraints": it can not explain why many real world cartels do not comprise all firms in the industry, the OPEC being a famous example.[4]

There is another strand in the literature on cartel stability, which takes a quite different route. The seminal papers in this literature are Selten (1973) and d'Aspremont et al. (1983). These papers investigate cartel stability in static models. By their very nature, these models leave unexplained why cartel members do not cheat on a cartel agreement; they may, therefore, be viewed as models of explicit rather than implicit collusion. In contrast to the supergame-theoretic literature, these papers focus on firms' "participation constraints". At the heart of this literature lies the trade off between participation and nonparticipation in a cartel: on the one hand, a firm has an incentive to join the cartel so as to achieve a more collusive outcome; on the other, it has an incentive to stay out of the cartel so as to take a free ride on the cartel's effort to restrict output.[5]

The main aim of the present paper is to develop a theoretical foundation of the traditional view on the relationship between cartel stability and capacity levels. The paper follows the tradition of Selten (1973) and d'Aspremont et al. (1983) in focussing on the issue of participation in a cartel. As in all static models of cartel stability, it is simply assumed that cartel rules on pricing or output setting can be enforced.[6] Alternatively, when joining the cartel, firms are assumed to delegate output decisions to some "cartel manager". This leaves the study of the interaction between incentive and participation constraints for future research.

We consider a two-stage game in which firms, each being subject to the same capacity constraint, first decide whether or not to join the cartel, and then compete in quantities.[7] Since cartel members internalise the externalities they impose on each other, they will tend to produce less than fringe firms, which take a free ride on the cartel's effort to restrict output. Now, the size of the cartel and, hence, the degree of collusion, is deter-

---

[4]The continuum of equilibria in supergames allows us to select equilibria such that only a subset of firms collude. But mere selection of equilibria does not explain why the number of colluding firms is larger in some situations than in other. In fact, the literature usually selects the best symmetric subgame perfect equilibrium. By selecting a *symmetric* equilibrium, it is simply assumed that, in equilibrium, either all firms participate in the cartel or none. Clearly, the *best* symmetric equilibrium is such that all participate.

[5]Here, "participation constraints" refer to firms' incentive to join the cartel rather than the "fringe", whereas "incentive constraints" refer to the incentives of cartel members only to cheat on the cartel agreement.

[6]This applies also to models of adverse selection in cartel formation, such as Roberts (1985) and Kihlstrom and Vives (1992). In these models, colluding firms may not truthfully announce their costs; cheating at the output stage is not considered.

[7]The main reason why we consider quantity rather than price competition is the nonexistence of pure strategy equilibria in games of price competition under capacity constraints. The notion of a market price is not well-defined in these games. Hence, it is not quite clear what is meant by the relationship between demand or capacity and the collusive equilibrium price.

mined by the relative incentives to join the cartel rather than the fringe. Intuitively, if the level of demand is rather high, or the industry level of capacity rather low, then free-riding on the cartel is not very profitable since fringe firms will face a binding capacity constraint in equilibrium. Therefore, the lower (higher) is the level of demand (capacity), the more attractive it is to join the fringe, and the more unstable the cartel becomes. The model thus predicts a positive (negative) relationship between the equilibrium cartel size and the level of demand (capacity). This reinforces the positive (negative) relationship between the collusive equilibrium price and demand (capacity) for a given degree of collusion. Introducing heterogeneity in capacities, we show that firms with large capacities have stronger incentives to join the cartel then small firms. The model thus predicts that firm sizes and capacity utilisation rates are negatively correlated across firms.

In their paper, Brock and Scheinkman have studied the comparative statics of cartel stability with respect to the level of capacity; all exogenous variables are constant over time. Following the paper by Rotemberg and Saloner (1986), there also exists a large literature on collusion over the business cycle where the level of demand is assumed to follow some dynamic process. Most of these papers do not consider capacity constraints. This literature investigates, in particular, whether collusive prices tend to vary procyclically or rather countercyclically. However, the results appear to depend quite delicately on the time-series properties of the assumed process. To address the issue of cartel stability over the business cycle, we analyse a very simple dynamic extension of the basic two-stage game. We show that collusive prices will tend to vary procyclically, independently of the assumed stochastic process. This, again, restores the traditional view.

2 A Simple Model of Cartel Formation and Cournot Competition

Consider the following two-stage game of cartel formation and quantity competition. There are n (identical) firms, each seeking to maximise its own profit; let N denote the set of firms. Each firm in the industry produces the same homogenous good. Firm i's cost of producing quantity Q_i is given by

$$C(Q_i) = \begin{cases} cQ_i & \text{if } Q_i \in [0, K] \\ \infty & \text{if } Q_i > K; \end{cases}$$

that is, all firms face a common unit cost of c up to the capacity level K.

Market demand can be represented by the (twice continuously differentiable) inverse demand function $P(Q/S)$ with $P'(\cdot) < 0$, where $Q \equiv \sum_{i=1}^{n} Q_i$ is total industry output, and S is a measure of market size (or the level of demand). An increase in the size of the market is understood to mean a replication of the population of consumers, leaving the distribution of incomes and tastes unchanged. We put the following, rather mild,

restriction on the shape of the demand curve:

$$P'(q) + qP''(q) < 0 \text{ for all } q \in (0, nK/S],\tag{1}$$

which implies that, under quantity competition, each firm's best-reply function is downward-sloping; that is, quantities are strategic substitutes. Inequality (1) holds, for instance, if demand is concave, i.e. $P''(\cdot) \leq 0$. To exclude the trivial case in which production is not viable, we posit $P(0) > c$. For notational convenience, we define output and capacity levels per unit of market size: $q_i \equiv Q_i/S$, $i = 1, ..., n$, and $k \equiv K/S$.

The timing of the game is as follows. At the first stage ("participation stage"), firms simultaneously decide whether to join the cartel (or coalition) M or the "fringe" $N\backslash M$. Formally, each firm i, $i \in N$, selects a zero-one variable z_i:

$$z_i = \begin{cases} 1 & \text{iff firm } i \text{ joins the cartel } M \\ 0 & \text{iff firm } i \text{ joins the fringe } N\backslash M. \end{cases}$$

In modelling cartel formation as a noncooperative simultaneous-move game, we follow Selten (1973).

At the second stage ("output stage"), firms simultaneously set quantities; that is, each firm i chooses q_i, $q_i \in [0, \infty)$, as a function of the vector of participation decisions $\mathbf{z}$, $\mathbf{z} = (z_1, ..., z_n)$. The difference between fringe firms and cartel members is the following. Each fringe firm i, $i \in N\backslash M$, sets q_i so as to maximise its own profit. In contrast, each cartel member i, $i \in M$, is constrained (by cartel rules) to set q_i so as to maximise the cartel's joint profit. This is the key assumption of the paper. One way to make cartel firms' behaviour consistent with individual profit maximisation is the following. By joining the cartel, each firm has to delegate its output decision to some "cartel manager" whose objective it is to maximise the cartel's joint profit. Alternatively, one might assume that cartel members can write a binding contract which prevents them from deviating.[8] All this is to say that we simply assume that cartel rules (which stipulate that firms set output so as to maximise the cartel's joint profit) can be enforced; this allows us to focus

---

[8] Another way of interpreting the assumption of joint profit maximisation of cartel members is the following. The output stage of the two-stage game is simply the "reduced form" of a quantity-setting supergame in which cartel members sustain joint profit maximisation using trigger strategies. According to these trigger strategies, cartel members will revert forever to the static (noncollusive) Nash equilibrium in case of deviation by a cartel member. The difference to the usual supergame models is that firms that choose not to participate in the cartel do not get punished. If firms do not discount future payoffs, then joint profit maximisation by cartel members can be sustained in an SPE of this supergame as long as joint profit maximisation gives each firm a higher per-period payoff than it could get in the static (noncollusive) Nash equilibrium. The latter condition does not necessarily hold. Hence, under this interpretation of the output stage, we would have to modify firms' payoffs slightly. (This would not change the results: one can show that, in equilibrium, cartel size is always such that each cartel member makes at least the same profit as under noncollusion.) Given the appropriate modification of payoffs, the participation stage simply "selects" (induces) an SPE of the ensuing quantity-setting supergame, which has quite different properties than the SPE usually considered in the literature.

on firms' participation decisions, and follows Selten (1973), d'Aspremont et al. (1983), and others.[9]

How a given joint cartel output q_M, $q_M \equiv \sum_{i \in M} q_i$, is divided among cartel members does not affect joint cartel profit; this is due to the assumption of constant returns to scale. Since firms are symmetric, we assume throughout an equal output sharing rule: $q_i = q_j$, $\forall i, j \in M$. Side payments between firms are not allowed. (Alternatively, we could directly assume that each cartel member receives the same profit in equilibrium, independently of how cartel output is divided among firms.)

3 Equilibrium Analysis

In this central section of the paper, we seek the (pure strategy) subgame perfect equilibrium (SPE) of the simple two-stage game. We proceed by backward induction. First, we take the size of the coalition as given, and show existence and uniqueness of a Nash equilibrium in quantities. We then investigate how changes in capacity and the size of the cartel affect the stage-2 Nash equilibrium. In the third subsection, we first determine necessary and sufficient conditions for a cartel size to be "stable", i.e. to be supportable in an SPE. We then show the existence of an SPE, and investigate the set of cartel sizes than can be supported in equilibrium. Finally, we turn to the main question of this paper, namely the effects of capacity and market size on equilibrium cartel size and, in particular, equilibrium price.

3.1 Equilibrium Analysis for a Given Cartel Size

In this subsection, we take firms' participation decisions, described by the vector $\mathbf{z} = (z_1, ..., z_n)$, as given. Suppose there are $m \equiv \#M \in \{0, 1, ..., n\}$ cartel members and $n - m \equiv \#N\backslash M$ fringe members. We now seek the associated Nash equilibrium in quantities.

Let $g(x, q) \equiv P(q) - c + xP'(q)$. Since $P'(q) < 0$, there exists a unique $x(q)$, $x(q) \in (-\infty, \infty)$, such that $g(x(q), q) = 0$. From inequality (1), it follows that $x'(q) = -\left[P'(q) + x(q)P''(q)\right]/P'(q) < 0$ for all $x(q) \in [0, q]$. Define $\overline{x}(q) \equiv \max\{0, x(q)\}$. Then, a vector of (normalised) quantities, $((q_i)_{i \in N})$, forms a stage-2 Nash equilibrium if and only if

$$q_i = \min\{\overline{x}(q), k\}, \ i \in N\backslash M, \tag{2}$$

and

$$q_M \equiv \sum_{i \in M} q_i = \min\{\overline{x}(q), mk\}, \tag{3}$$

---

[9]Notice that any renegotiation-proof contract between cartel members must involve joint profit maximisation if renegotiation is allowed between cartel members only.

where $q \equiv \sum_{i \in N} q_i$ is industry output (per unit of market size). To see this, notice that each fringe firm's profit function is strictly concave in its own output; similarly, the cartel's joint profit function is strictly concave in joint cartel output. Equations (2) and (3), combined with the cartel's output sharing rule, uniquely determine each firm's output for a given value of q. It follows immediately that all $n - m$ fringe firms produce the same quantity in equilibrium.

Define the function $h(q) \equiv (n - m) \min\{\overline{x}(q), k\} + \min\{\overline{x}(q), mk\} - q$. An output vector $((q_i)_{i \in N})$ is then an equilibrium if and only if $h(q) = 0$. Observe that $h(0) > 0$ since $P(0) > c$, and $\lim_{q \to \infty} h(q) = -\infty$. Moreover, $h'(q)|_{h(q)=0} < 0$, where this derivative exists (otherwise, both the right- and left-hand derivatives are negative), since $\min\{\overline{x}(q), mk\}|_{h(q)=0} < q$. Hence, $h(q)$ has a unique nonnegative root. This shows that there exists a unique Nash equilibrium in quantities. In equilibrium, each firm produces a positive output, i.e. $\overline{x}(q) = x(q) > 0$.

Denote by $q_f(m; k)$ and $q_c(m; k)$ a fringe firm's and a cartel member's equilibrium output, respectively; the cartel's joint equilibrium output is $q_M(m; k) = mq_c(m; k)$. We have $q_f(1; k) = q_f(0; k)$ since $q_M(1; k) = q_f(1; k)$. In the following characterisation of equilibrium quantities, we therefore restrict attention to $m \in \{1, ..., n\}$. Depending on the values of k and m, three different cases can arise.

Case (i): If

$$P((n - m + 1)k) - c + kP'((n - m + 1)k) \leq 0, \tag{4}$$

then capacity constraints are nonbinding for all firms in equilibrium. In this case, equilibrium quantities are given by

$$q_M(m; k) = q_f(m; k) = x((n - m + 1)q_f(m; k)) \in (0, k]. \tag{5}$$

Equilibrium profits are

$$\pi_c(m; k) = [P((n - m + 1)q_f(m; k)) - c] \frac{q_f(m; k)}{m} \tag{6}$$

and

$$\pi_f(m; k) = [P((n - m + 1)q_f(m; k)) - c] q_f(m; k). \tag{7}$$

Case (ii): If

$$P((n - m + 1)k) - c + kP'((n - m + 1)k) > 0 \tag{8}$$

and

$$P(nk) - c + mkP'(nk) \leq 0, \tag{9}$$

then fringe firms only face a binding capacity constraint. Hence, $q_f(m; k) = k$, and

$$q_M(m; k) = x(q_M(m; k) + (n - m)k) \in (k, mk]. \tag{10}$$

6

Equilibrium profits are then given by

$$\pi_c(m; k) = [P(mq_c(m; k) + (n - m)k) - c] \, q_c(m; k) \tag{11}$$

and

$$\pi_f(m; k) = [P(mq_c(m; k) + (n - m)k) - c] \, k. \tag{12}$$

Notice that this case can only arise if cartel size $m \geq 2$.

Case (iii): Capacity constraints are binding for all firms in equilibrium if and only if

$$P(nk) - c + mkP'(nk) > 0. \tag{13}$$

Then, each firm sets its output equal to its capacity level. Equilibrium profit is

$$\pi_c(m; k) = \pi_f(m; k) = [P(nk) - c] \, k. \tag{14}$$

Let us summarise the results in the following proposition.

Proposition 1 *There exists a unique Nash equilibrium in quantities. In equilibrium,* $\pi_c(m; k) < \pi_f(m; k)$ *if* $P(nk) - c + mkP'(nk) < 0$, *and* $\pi_c(m; k) = \pi_f(m; k)$ *otherwise.*

In equilibrium, fringe firms are at least as well off as cartel members. This follows from the fact that fringe firms maximise their own profit whereas cartel members maximise joint cartel profit, whereby they internalise the externality they impose on fellow cartel firms.

In this paper, the (m, k)-space plays a crucial role. The following lemma gives a useful result on the division of this space into different regions.

Lemma 1 *There exist real-valued functions* $m_c(k)$ *and* $m_f(k)$, *with* $dm_c(k)/dk < 0$ *and* $dm_f(k)/dk > 0$, *such that*

- *cartel members face a binding capacity constraint if and only if cartel size* m, $m \in [1, n]$, *is such that* $m < m_c(k)$;

- *fringe firms face a binding capacity constraint if and only if* $m > m_f(k)$ *for* $m \in [1, n)$.

Moreover, if k *such that* $m_c(k) > 0$, *then* $m_f(k) < 1$.

The proof can be found in the appendix. See figure 1 for an illustration. Lemma 1 implies, in particular, that if a cartel member faces a binding capacity constraint when the cartel size is m, then the capacity constraint is binding for all smaller cartel sizes m', $m' \in \{1, ..., m - 1\}$.
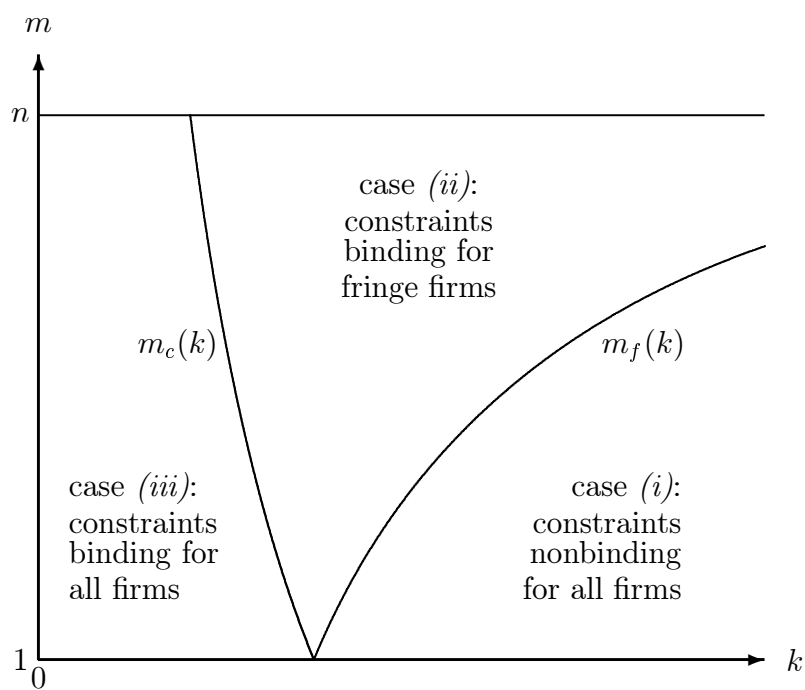
Figure 1: The Division of the (m, k)-space: The Functions $m_c(k)$ and $m_f(k)$.

8

3.2 Comparative Statics

The main aim of this paper is to investigate how the equilibrium price of the two-stage game varies with capacity K and demand S. There are two potential channels of influence. For any given coalition structure, capacity per unit of market size, k, has a direct impact on the stage-2 equilibrium price. But k can also influence the equilibrium cartel size, which in turn affects the stage-2 equilibrium price. Therefore, having established existence and uniqueness of a Nash equilibrium in quantities, we now want to investigate the comparative statics of this stage-2 equilibrium price with respect to k and m. In subsection 3.4, we will analyse how the equilibrium value of m varies with k.

Observe that equilibrium output, $q_c(m; k)$ and $q_f(m; k)$, and equilibrium profit, $\pi_c(m; k)$ and $\pi_f(m; k)$, are continuous functions of k, and of cartel size m if one treats m as a continuous variable. The following result shows how equilibrium output, and hence equilibrium price, varies with cartel size.

Lemma 2 *Total industry output, $q_M(m; k) + (n - m)q_f(m; k)$, is weakly decreasing in cartel size m. Joint cartel output, $q_M(m; k)$, and each fringe firm's output, $q_f(m; k)$, are weakly increasing in cartel size m.*

The proof can be found in the appendix; the intuition for the result is as follows. Decompose the effect of an increase in cartel size m into a (converging) sequence of output changes – neglecting, for ease of exposition, the existence of capacity constraints. As the first step, let the ("old" and "new") cartel members "myopically" adjust their output decisions, holding fixed the output level of the remaining fringe firms. Since cartel members optimally internalise the externalities they impose on each other, this will lead to a lower industry output, and hence a higher price. As the second step, let the remaining fringe firms myopically change their output, keeping constant the cartel's output. As a response to the higher price, fringe firms will increase their output. This, in turn, will induce the cartel firms to reduce their output (step 3) since quantities are strategic substitutes, and so on. In equilibrium, the remaining fringe firms will have increased their output. However, this joint output increase will be less than the decrease in output by the (old and new) cartel members since each firm's best-reply function has a slope larger than -1. In the absence of capacity constraints, joint cartel output is equal to each fringe firm's output. Hence, joint cartel output will increase as well.

Since industry output is a (weakly) decreasing function of the size of the coalition, the equilibrium price is (weakly) increasing in m. The following corollary is then immediate.

Corollary 1 *A fringe firm's equilibrium profit, $\pi_f(m; k)$, and joint cartel profit, $m\pi_c(m; k)$, are weakly increasing functions of cartel size m.*

The relationship between an individual cartel member's output (and profit) and cartel size is, however, generally nonmonotonic. This is a well-known result from the literature on horizontal mergers under Cournot competition without capacity constraints; see

Salant, Switzer, and Reynolds (1983). The reason is that an increase in cartel size will induce the remaining fringe firms to increase their output, provided that they are not capacity constrained. Since quantities are strategic substitutes, each cartel member's output and profit may decrease. If, however, fringe firms face a binding constraint in equilibrium, then an increase in cartel size will not have this adverse effect on a cartel member's profit.

Lemma 3 *If $m_c(k) > 1$, then $\pi_c(m; k)$ is constant on $[0, m_c(k)]$, and strictly increasing in m on $(m_c(k), n)$.*

Proof. First, notice that $\pi_c(m; k) = k[P(nk) - c]$ for all $m \in [0, \max\{m_c(k), 1\}]$. Second, if $m \in (m_c(k), n)$, $m_c(k) > 1$, then $\pi_c(m; k)$ is given by (11). Using the envelope theorem, we get

$$\frac{\partial}{\partial m}\pi_c(m; k) = -(k - q_c(m; k))q_c(m; k)P'(mq_c(m; k) + (n - m)k),$$

which is strictly positive since $k > q_c(m; k)$ if $m > m_c(k) > 1$. ∎

We now turn to comparative statics – for a given cartel size – with respect to k.

Lemma 4 *For a given cartel size m, equilibrium price $P(q_M(m; k) + (n - m)q_f(m; k))$ is a (weakly) decreasing function of the capacity level per unit of market size, k.*

Proof. As in the proofs of proposition 1 and lemma 2, we have to distinguish between three cases; again, we confine attention to $m \in [1, n]$.

Case (i): Since capacity constraints are nonbinding for all firms, equilibrium output is unaffected by a marginal change in k.

Case (ii): Note first that $\partial q_M(m; k)/\partial k > -(n - m)$. Taking the derivative of the equilibrium price with respect to k yields

$$\frac{dP(q_M(m; k) + (n - m)k)}{dk} = P'(q_M(m; k) + (n - m)k)\left[\partial q_M(m; k)/\partial k + n - m\right] < 0.$$

Case (iii): When all firms face a binding constraint, equilibrium output is equal to $P(nk)$. We thus have $dP(nk)/dk = nP'(nk) < 0$. ∎

We have thus established that an increase in the level of capacity or a decrease in demand (market size), will lower price for a given size of the coalition. From lemma 2 we already know that an increase in the size of the cartel will increase price. In order to determine the overall effect on equilibrium price, we, therefore, need to analyse the effect of changes in K and S on the equilibrium cartel size. But before doing so, we have to determine the conditions that need to hold for a cartel size to be supportable in an SPE of the two-stage game; this is the purpose of the next subsection.

3.3 Stable Cartel Size

So far, we have analysed firms' stage-2 output decisions for a given size of the coalition M. We now turn to the analysis of the participation decision stage (stage 1) at which firms simultaneously decide whether to join the cartel M or rather the fringe $N \backslash M$. The equilibrium cartel size is determined by the tension between a firm's incentive to free ride on the cartel's effort to restrain output (by staying out of the cartel) and to achieve a more collusive outcome (by joining the cartel).

The following lemma gives the necessary and sufficient conditions for a cartel size to be sustainable in an SPE of the two-stage game. It turns out that these conditions are identical to the notions of internal and external stability, which are used in the (not explicitly game-theoretic) literature on static cartel stability. The seminal contribution in this literature is d'Aspremont et al. (1983); other papers include Donsimoni (1985), Donsimoni, Economides, and Polemarchakis (1986), and Shaffer (1995).

Lemma 5 *A cartel of size m can be supported in an SPE of the two-stage game if and only if the following two conditions hold:*

(internal stability) $m = 0$ or

$$\pi_c(m; k) \geq \pi_f(m - 1; k) \text{ for } m \in \{1, ..., n\} \tag{15}$$

and

(external stability) $m = n$ or

$$\pi_c(m + 1; k) \leq \pi_f(m; k) \text{ for } m \in \{0, ..., n - 1\}. \tag{16}$$

Proof. The condition of internal stability implies that no firm has an incentive to leave the cartel. Similarly, the condition of external stability ensures that no firm can profitably deviate at stage 1 by joining the cartel. Hence, these conditions are sufficient. Conversely, if one of these conditions is not satisfied, then there exists a profitable deviation for some firm. Hence, the conditions are necessary for an SPE. ∎

In the following, we will say that a cartel size is *stable* whenever it can be supported in an SPE. Having defined the necessary and sufficient conditions for an equilibrium cartel size, we want to be sure that the set of equilibria is nonempty for all possible parameter values. For this purpose, we introduce some notation. Let $[x]$ denote the integer part of x, and define the step function

$$\overline{m}_c(k) \equiv \begin{cases} n & \text{if } m_c(k) \in (n, \infty) \\ [m_c(k)] & \text{if } m_c(k) \in [1, n] \\ 1 & \text{otherwise.} \end{cases}$$

Notice that $\overline{m}_c(k)$ is weakly decreasing in k.

Proposition 2 *The set of cartel sizes that can be supported in an SPE, $M^*(k)$, has at least two elements. In particular, $\{0, ..., \overline{m}_c(k) - 1\} \subset M^*(k)$; hence, noncollusion $(m = 0)$ can always be supported in equilibrium. Moreover, there exists at least one other stable cartel size $m^*(k)$, $m^*(k) \in \{\overline{m}_c(k), ..., n\}$.*

Proof. To see that $\{0, ..., \overline{m}_c(k) - 1\} \subset M^*(k)$, notice first that $\pi_f(0; k) = \pi_c(1; k)$. Hence, $m = 0$ is a stable cartel size. Furthermore, $\pi_c(m; k) = \pi_f(m - 1; k)$ for all $m \in \{1, ..., \overline{m}_c(k)\}$. This proves the assertion. To show that there exists a stable cartel size $m^*(k)$, $m^*(k) \in \{\overline{m}_c(k), ..., n\}$, we can simply apply the algorithm by d'Aspremont et al. (1983). Let us start with a cartel of size $m = \overline{m}_c(k)$. It is internally stable since $\pi_c(\overline{m}_c(k); k) = \pi_f(\overline{m}_c(k) - 1; k)$; if it is externally stable, then the algorithm stops since we have found a stable cartel. Otherwise, we continue with a cartel of size $m = \overline{m}_c(k) + 1$, which is internally stable since the cartel of size $m = \overline{m}_c(k)$ was found to be externally unstable. If $m = \overline{m}_c(k) + 1$ is externally stable, then the algorithm stops; otherwise, it continues with $m = \overline{m}_c(k) + 2$, and so on. By induction, the algorithm reaches size $m = n$ if and only if $m = n - 1$ was found to be externally unstable. But, then, $m = n$ must be internally stable, and it is externally stable by definition; hence, it is stable. This concludes the proof. ∎

In order to conduct comparative statics, ideally, we want to have uniqueness of a stable cartel size. Unfortunately, there is no hope of finding a unique equilibrium cartel size, as proposition 2 indicates. Notice, however, that the equilibrium price is the same for all $m \in \{0, ..., \overline{m}_c(k)\}$. Arguably, *this* kind of multiplicity should not be a cause of concern. Below, we will further characterise the set of stable cartel sizes; however, a full characterisation for a general demand function is beyond the scope of this paper.

Due to the multiplicity of equilibria, we will focus on the maximum and minimum sustainable cartel sizes. From proposition 2, we already know that the empty cartel is the smallest stable cartel. The following result gives an instructive and useful result regarding the largest stable cartel.

Lemma 6 *All cartel sizes above the maximum stable cartel size, $\overline{m}^*(k)$, are internally unstable; that is,*

$$\pi_c(m; k) < \pi_f(m - 1; k) \text{ for all } m \in \{\overline{m}^*(k) + 1, ..., n\}.$$

Proof. Suppose $\overline{m}^*(k) = n - 1 - l$, $l \in \{0, 1, ..., n - 2\}$. Since n is externally stable by definition, it must be internally unstable. Assume now that $l \geq 1$. Since n is internally unstable, $n - 1$ is externally stable; hence, $n - 1$ must be internally unstable. The proof proceeds in the same fashion. ∎

Proposition 2 leaves the possibility open that all stable cartel sizes are "degenerate" cartels. Indeed, for all $m \in \{1, ..., \overline{m}_c(k)\}$, the equilibrium price is the noncollusive one (when $m = 0$). Using lemma 6, we are now in the position to further characterise the set of stable cartel sizes. The following result shows that "nondegenerate" cartels exist for all k such that $m_c(k) \in [1, n)$.

Proposition 3 *If $m_c(k) \in [1, n)$, then the maximum sustainable cartel size, $\overline{m}^*(k)$, is strictly larger than $m_c(k)$, i.e. $\overline{m}^*(k) \geq \overline{m}_c(k) + 1$; moreover, if $m_c(k) \in (1, n)$, then $\overline{m}_c(k)$ is externally unstable.*

Proof. Notice first that $\pi_c(\overline{m}_c(k); k) = \pi_f(\overline{m}_c(k); k)$. Furthermore, if $m_c(k) \in (1, n)$, then $\pi_c(\overline{m}_c(k) + 1; k) > \pi_c(\overline{m}_c(k); k)$ by lemma 3. It follows immediately that $\pi_c(\overline{m}_c(k) + 1; k) > \pi_f(\overline{m}_c(k); k)$. Hence, $\overline{m}_c(k)$ is externally unstable, and $\overline{m}_c(k) + 1$ internally stable, if $m_c(k) \in (1, n)$. By continuity, $\overline{m}_c(k) + 1$ must also be internally stable if $m_c(k) = 1$. Suppose now the assertion is false; that is, suppose there exists a capacity level k such that $m_c(k) \in [1, n)$ and $\overline{m}^*(k) < \overline{m}_c(k) + 1$. Then, $\overline{m}_c(k) + 1$ is internally unstable by lemma 6. But this can not be true as we have just shown. ∎

Proposition 3 implies that if $m_c(k) \in (n-1, n)$, then there exists an equilibrium such that all firms participate in the cartel, i.e. $\overline{m}^*(k) = n$, although no firm faces a binding capacity constraint. Nevertheless, even in this case, the existence and level of capacity constraints have a major impact on the equilibrium outcome in that a further rise in capacity or a fall in demand can cause the collapse of collusion. This will become clear from the analysis below. Before we turn to the analysis of changes in capacity or demand on the equilibrium cartel size, let us make two further remarks on stability. First, suppose the smallest stable cartel in $\{\overline{m}_c(k), ..., \overline{m}^*(k)\}$ is strictly larger than $\overline{m}_c(k)$. Then, $\overline{m}_c(k)$ and all cartel sizes below the smallest stable one are externally unstable. The argument is similar to that in the proof of proposition 3. Second, recall that, from proposition 2, if k is sufficiently small such that $m_c(k) \geq n$, then all cartel sizes can be supported in equilibrium, i.e. $\overline{m}^*(k) = n$.

3.4 The Effects of Capacity and Demand on Cartel Size and Price

We now want to investigate how the equilibrium cartel size varies with k, that is, with capacity level K and market size S. Consider a stable cartel size, say $m^*(k)$. Since $\pi_c(m; k)$ and $\pi_f(m; k)$ are continuous in k, a marginal change in k will have no effect on $m^*(k)$ unless the condition for internal stability, (15), or the condition for external stability, (16), hold with equality, that is, unless $\pi_c(m^*(k); k) = \pi_f(m^*(k) - 1; k)$ or $\pi_c(m^*(k) + 1; k) = \pi_f(m^*(k); k)$ hold.

Therefore, we will, in turn, analyse the effect of a small change in k on equilibrium profits for cartel and fringe members when these conditions hold with equality. The proof of the following lemma is quite involved and can be found in the appendix.

Lemma 7 *Suppose cartel size m, $m \in \{\overline{m}_c(k) + 1, ..., n\}$, and capacity k are such that the condition for internal stability is binding, i.e. $\pi_c(m; k) = \pi_f(m - 1; k)$. Then,*

$$\frac{\partial}{\partial k} \{\pi_c(m; k) - \pi_f(m - 1; k)\} \leq 0, \tag{17}$$

where this derivative exists; otherwise, the inequality holds for both the derivatives from the right and the left. That is, a marginal increase in k either implies that the condition for internal stability continues to be binding or that it will be violated.

A direct consequence of lemma 7 is the following result.

Corollary 2 *Suppose cartel size m, $m \in \{\overline{m}_c(k), ..., n-1\}$, is such that the condition for external stability is binding, i.e. $\pi_c(m+1; k) = \pi_f(m; k)$. Then,*

$$\frac{\partial}{\partial k} \left\{ \pi_c(m+1; k) - \pi_f(m; k) \right\} \leq 0,$$

where this derivative exists; otherwise, the inequality holds for both the derivatives from the right and the left. That is, a marginal increase in k either implies that the condition for external stability continues to be binding or that it will become slack.

Proof. This follows directly from lemma 7, substituting m for $m-1$, and $m+1$ for m. ∎

The idea behing lemma 7 is the following. If a firm decides to join the fringe rather than the cartel, its equilibrium output will tend to be higher since it will take a free ride on the cartel's attempt to restrict output. For very high levels of capacity, a firm will be unconstrained, independently of whether or not it joins the cartel. If it is initially indifferent, then a small increase in capacity has no effect on the relative incentives to join the cartel. Now, for moderate levels of capacity k, a firm will face a binding capacity constraint if it decides to join the fringe, but not if it becomes a member of the cartel. In this case, then, an increase in capacity will make it relatively more attractive to join the fringe since the rise in k relaxes the fringe's constraint. Notice, however, that inequality (17) does not hold if $m = m_c(\overline{k}) \in \{2, ..., n\}$ at some capacity level $\overline{k}$. For all $k \leq \overline{k}$, we have $\pi_c(m; k) = \pi_f(m-1; k)$, whereas for capacity levels k, $k > \overline{k}$, such that $m > m_c(k) > m-1$, we have $\pi_c(m; k) > \pi_f(m-1; k)$; see proposition 3. That is, the right-hand derivative of $\pi_c(m; \overline{k}) - \pi_f(m-1; \overline{k})$ with respect to k is strictly positive. The reason is that at capacity $\overline{k}$ and cartel size $m-1$, fringe firms *and* cartel members are constrained so that a rise in capacity will induce *all* firms to increase output. In contrast, at capacity $\overline{k}$ and cartel size m, cartel members are just unconstrained so that fringe firms only will raise output as k increases.

Lemma 7 and corollary 2 show that an increase in capacity (or a decrease in demand) can cause a hitherto stable cartel size m to become internally unstable but not externally unstable, given that $m > \overline{m}_c(k)$. This suggests that stable cartels tend to become smaller the higher is the level of capacity per unit of market size. Due to the multiplicity of equilibria, we focus on the maximum stable cartel size, $\overline{m}^*(k)$. (From proposition 2, we already know that the smallest stable cartel is the empty one.)

We are now in the position to state and prove one of the main results of this paper.

Proposition 4 *The maximum stable cartel size $\overline{m}^*(k)$, $\overline{m}^*(k) \in \{1,...,n\}$, is weakly decreasing in capacity K, and weakly increasing in market size S.*

Proof. Suppose that the assertion of the proposition is false. Then, there exist some capacity levels k_1 and k_2, $k_2 > k_1$, such that $\overline{m}^*(k_1) \equiv m_1 < m_2 \equiv \overline{m}^*(k_2)$. By proposition 2, $m_1 \geq \overline{m}_c(k_1)$; moreover, $\overline{m}_c(k_1) \geq \overline{m}_c(k)$, for all $k \in [k_1, k_2]$, since $\overline{m}_c(k)$ is weakly decreasing in k. Hence, $m_2 > \overline{m}_c(k)$ for all $k \in [k_1, k_2]$. From lemma 6, we know that m_2 is internally instable at k_1, but it is, by assumption, internally stable at k_2. Hence, there exists a capacity level k^*, $k^* \in [k_1, k_2]$, such that $\pi_c(m_2; k^*) = \pi_f(m_2 - 1; k^*)$, and

$$\frac{\partial}{\partial k} \left\{ \pi_c(m_2; k^*) - \pi_f(m_2 - 1; k^*) \right\} > 0,$$

where this derivative exists (otherwise, the inequality must hold for the corresponding derivative from the left). But this contradicts lemma 7. Hence, the assertion of the proposition can not be false. ∎

Proposition 4 restores the traditional view in that it shows that large cartels tend to break down in periods of high (excess) capacity and low demand. Although proposition 4 states an important result, eventually, we are more interested in price rather than in the size of the cartel. This is not only for the obvious reason that price directly affects consumer surplus, whereas cartel size has an impact on consumer surplus only via price (and output). More importantly, in empirical applications, we can hope to observe or measure price, but usually not the size of a coalition.

The following proposition states the central prediction of this paper, which only involves variables that are, in principle, measurable.

Proposition 5 *The maximum and minimum sustainable equilibrium prices, $\overline{p}^*(k)$ and $\underline{p}^*(k)$, are weakly decreasing in capacity K, and weakly increasing in market size S.*

Proof. Observe that, given a size m of the cartel, there exists a unique equilibrium price (proposition 1). Moreover, this equilibrium price is weakly increasing in m, given k (lemma 2). Hence, for any k, the maximum stable equilibrium price $\overline{p}^*(k)$ is the unique equilibrium price given the maximum stable cartel size. Now, the maximum sustainable cartel size is (weakly) decreasing in k (proposition 4). Furthermore, the equilibrium price is weakly increasing in m, given k (lemma 2), and weakly decreasing in k, given m (lemma 4). Hence, $\overline{p}^*(k)$ is weakly decreasing in k. Similarly, for any k, the minimum sustainable cartel size $\underline{p}^*(k)$ is the unique equilibrium price given the minimum stable cartel size. From proposition 2, we know that the empty cartel ($m = 0$) can be supported in an SPE for all values of k. Hence, $\underline{p}^*(k)$ is the unique equilibrium price given $m = 0$. The assertion of the proposition then follows directly from lemma 4. ∎

15

3.5 Robustness of Results

Quantity Competition. An important question is whether our results hinge on the assumption of quantity competition at the output stage. It would, in particular, be interesting to study the case of price competition. However, as is well known, pure strategy equilibria fail to exist in games of price competition under capacity constraints so that the notion of a market price is not well defined. Now, following the seminal contribution by d'Aspremont et al. (1983), much of the literature on cartel stability in static models assumes that the cartel acts as price leader, whereas all firms in the fringe are price takers (when deciding upon output); see, for instance, Donsimoni (1985) and Donsimoni, Economides, and Polemarchakis (1986).[10] It is quite straightforward to show that the main conclusions of our Cournot model carry over to the price leadership model. To see this, notice that, for any price above marginal cost, each fringe firm sets its output equal to its capacity. Given the aggregate supply of the fringe, the cartel faces a residual demand curve, and sets price so as to maximise its joint profit. But in this case, price and quantity setting are equivalent. Hence, we are back in the Cournot world, namely in the subcase where fringe firms face a binding capacity constraint. In contrast to the previous Cournot model, however, an empty cartel ($m = 0$) may *not* be supported in equilibrium. This is due to the fact that fringe firms are assumed to take price as given; that is, they neglect the effect of their output decisions on price. Most other previous results go through. In particular, the maximum stable cartel size, $\overline{m}^*(k)$, is weakly decreasing, and the maximum sustainable equilibrium price, $\overline{p}^*(k)$, is even *strictly* decreasing in k.[11]

Cartel Formation. Following Selten (1973), we have modelled cartel formation as a noncooperative simultaneous-move game. Let us denote this game by Γ. To investigate the robustness of our predictions, we now want to study what happens if we assume that firms decide sequentially, rather than simultaneously, whether or not to join the cartel. That is, suppose that, at the participation stage, first firm 1 selects z_1, then firm 2 selects z_2, and so on; the labelling of firms is arbitrary. Denote the new game by Γ'. It is straightforward to show that there exists a pure strategy subgame perfect equilibrium (SPE) in the sequential cartel formation game Γ'. Moreover, any cartel size m that can be supported in an SPE of Γ' can also be supported in an SPE of Γ. To see this, notice that the conditions for internal and external stability of cartel size m, equations (15) and (16), are necessary conditions for m to be sustainable in an SPE of Γ', since, otherwise, firm n could profitably deviate. Indeed, the sequential-move game selects the "most efficient" equilibrium of the simultaneous-move game in that it solves the "coordination problem" in the formation of the cartel. Hence, the maximum sustainable equilibrium

---

[10]These models invite an obvious criticism: Why are firms in the fringe so "naive" not to take into account the effect of their output decision on price but, at the same time, so "sophisticated" to compute the exact price effect of their participation decision? Putting this theoretical issue aside, these models might nevertheless give a good empirical description of the nature of competition in some markets.

[11]See the proof of lemma 4.

size in Γ, $\overline{m}^*(k)$, can always be supported in Γ'. Moreover, we prove the following result in the appendix.

Proposition 6 *If the maximum sustainable equilibrium size in the simultaneous cartel formation game* Γ, $\overline{m}^*(k)$, *is such that* $\pi_c(\overline{m}^*(k); k) > \pi_f(\overline{m}^*(k) - 1; k)$, *then* $\overline{m}^*(k)$ *is the unique equilibrium cartel size in the sequential cartel formation game* Γ'. *In the unique equilibrium of* Γ', $z_i = 0$ *if* $i \in \{1, ..., n - \overline{m}^*(k)\}$, *and* $z_i = 1$ *if* $i \in \{n - \overline{m}^*(k) + 1, ..., n\}$.

It is possible to show that the condition of the proposition is satisfied for "almost all" k such that $m_c(k) \in (1, n)$. This result is reassuring in that it shows the robustness of our predictions, which are even sharper under the assumption of sequential cartel formation.

We have modelled cartel formation as an open membership game; that is, cartel members can not prevent other firms from joining the cartel. However, under quantity competition, profit per cartel member is not monotonic in the number of cartel members. Does this imply that, in equilibrium, a cartel member may be better off by restricting membership? The following proposition gives a reassuring answer.

Proposition 7 *In equilibrium, a cartel member is never better off by restricting cartel membership. Formally, if* m^* *is a stable cartel size, then*

$$\pi_c(m^*; k) \geq \pi_c(m^* - l; k) \text{ for any } l \in \{1, ..., m^* - 1\}.$$

Proof. Fix $l \in \{1, ..., m^* - 1\}$. For a given cartel size, a fringe firm is always at least as well off as a cartel member, i.e. $\pi_c(m^* - l; k) \leq \pi_f(m^* - l; k)$. From corollary 1, the profit of a fringe member is (weakly) increasing in the size of the cartel; accordingly, $\pi_f(m^* - l; k) \leq \pi_f(m^* - 1; k)$. Finally, we have $\pi_f(m^* - 1; k) \leq \pi_c(m^*; k)$ by stability of m^*. This proves the assertion. ∎

4 Heterogeneity in Capacity

Following Brock and Scheinkman (1985), we have assumed so far that all firms face the same capacity constraint. The aim of this section is to sharpen the predictions of the basic two-stage game. In particular, we address the following question. Allowing for heterogeneity in capacity levels among firms, do "larger" firms have stronger incentives to join the cartel than their "smaller" rivals? If the answer were yes, then this would be reinsuring for two reasons. First, it would reinforce the tendency for fringe firms to face a binding capacity constraint in equilibrium, and for cartel firms not to be constrained. As we have shown, it is in this case that a change in demand will have a major impact on the equilibrium cartel size. Second, due to their free-riding behaviour, fringe firms tend to be much larger than cartel members in terms of production. However, economists usually think of "fringe firms" as being rather small. The pure existence of capacity

constraints alleviates this "problem" somewhat. But if firms with large capacities had stronger incentives to join the cartel than firms with small capacities, then cartel firms would tend to be large, and fringe firms would tend to be small. Such an equilibrium outcome would thus be consistent with the common notion of small fringe firms. Below, we show that the answer to the question is indeed yes. Moreover, we show that the model predicts that firm size and the degree of capacity utilisation are negatively correlated across firms.

We generalise the model of section 2 in that we allow for arbitrary heterogeneity in capacity levels among firms. Each firm i, $i \in N$, is equipped with a capacity per unit of market size of k_i, $k_i \in (0, \infty)$. Inequality (1) is now supposed to hold for all $q \in [0, \sum_{i \in N} k_i]$. Since firms are no longer symmetric, we have to chose an output sharing rule for the cartel that reflects the different firm sizes. The most natural assumption appears to be that each cartel member's share α_i of joint cartel output, and hence of joint profit, is equal to its share of joint cartel capacity, $\alpha_i \equiv k_i / k_M$, where $k_M \equiv \sum_{i \in M} k_i$.[12] All other features of the basic model remain unchanged.

Turning to the equilibrium analysis of the generalised two-stage game, we proceed again by backward induction. We first seek the stage-2 Nash equilibrium, given any vector $\mathbf{z}$ of participation decisions. Using the same notation as in subsection 3.1, a vector of quantities, $((q_i)_{i \in N})$, forms a Nash equilibrium if and only if

$$q_i = \min\{\overline{x}(q), k_i\}, \ i \in N \backslash M,$$

and

$$q_M \equiv \sum_{i \in M} q_i = \min\left\{\overline{x}(q), k_M\right\},$$

where $q \equiv \sum_{i \in N} q_i$ is industry output, and $\overline{x}(q)$ the backward reaction mapping. Let $h(q) \equiv \sum_{i \in N \backslash M} \min\{\overline{x}(q), k_i\} + \min\left\{\overline{x}(q), k_M\right\} - q$. As before, there exists a unique non-negative q such that $h(q) = 0$. Hence, there is a unique Nash equilibrium in quantities. In equilibrium,

$$q_i = \begin{cases} \overline{x}(q) & \text{if } k_i \geq \overline{x}(q) \\ k_i & \text{otherwise,} \end{cases} \ i \in N \backslash M,$$

and

$$q_M = \begin{cases} \overline{x}(q) & \text{if } k_M \geq \overline{x}(q) \\ k_M & \text{otherwise.} \end{cases}$$

Denote by $q_i((z_i, \mathbf{z}_{-i}); \mathbf{k})$ firm i's equilibrium output, where $\mathbf{z}_{-i} = (z_1, ..., z_{i-1}, z_{i+1}, ..., z_n)$ are the participation decisions of i's rivals, and $\mathbf{k} = (k_1, ..., k_n)$ gives the distribution of capacities. Similarly, $\pi_i((z_i, \mathbf{z}_{-i}); \mathbf{k})$ denotes firm i's equilibrium profit; the cartel's joint equilibrium profit is $\pi_M(\mathbf{z}; \mathbf{k})$.

Before analysing firms' participation decisions, we consider some useful properties of the stage-2 equilibrium. The proof of the following results can be found in the appendix.

---

[12]There is evidence that cartels use such a rule in setting output quotas; see Davidson and Deneckere (1990) and Scherer (1980).

Lemma 8 *A fringe firm's equilibrium profit $\pi_i(\mathbf{z}; \mathbf{k})$, $i \in N\backslash M$, is weakly decreasing in k_j, $j \neq i$, and weakly increasing in k_i. Similarly, joint cartel profit is weakly decreasing in rival capacity, and weakly increasing in its own joint capacity.*

Inducing backwards, let us now turn to firms' participation decisions. From the definition of subgame perfection, it follows that the vector $\mathbf{z}$ can be supported in an SPE if and only if

$$\pi_i((1, \mathbf{z}_{-i}); \mathbf{k}) \geq \pi_i((0, \mathbf{z}_{-i}); \mathbf{k}) \text{ if } z_i = 1 \text{ (``internal stability''),}$$

and

$$\pi_i((1, \mathbf{z}_{-i}); \mathbf{k}) \leq \pi_i((0, \mathbf{z}_{-i}); \mathbf{k}) \text{ if } z_i = 0 \text{ (``external stability'').}$$

This implies that noncollusion, i.e. $z_i = 0$ for all $i \in N$, can again be sustained in equilibrium since $\pi_i((1, 0, ..., 0); \mathbf{k}) = \pi_i((0, 0, ..., 0); \mathbf{k})$. The following lemma considers firms' incentives to join the cartel. For the proof, the reader is referred to the appendix.

Lemma 9 *Fix an arbitrary vector $\mathbf{z}$ of participation decisions, and consider the two largest fringe firms i and j, $z_i = z_j = 0$, with $k_i \geq k_j$. Then,*

$$\pi_i((1, \mathbf{z}_{-i}); \mathbf{k}) < \pi_i((0, \mathbf{z}_{-i}); \mathbf{k}) \Longrightarrow \pi_j((1, \mathbf{z}_{-j}); \mathbf{k}) < \pi_j((0, \mathbf{z}_{-j}); \mathbf{k}).$$

That is, the larger firm has the stronger incentive to join the cartel.

Note that this result is not trivial. To see this, suppose firm i's capacity is twice that of firm j's. Then, firm i's share of cartel output in case of firm i's deviation will be *less* than twice the share of firm j in case firm j deviates. The following lemma considers the incentives for firms to leave the cartel.

Lemma 10 *Fix an arbitrary vector $\mathbf{z}$ of participation decisions such that $k_h \geq k_l$ for all $h \in M$ and $l \in N\backslash M$. Consider any two cartel members i and j, $z_i = z_j = 1$, with $k_i \geq k_j$. Then,*

$$\pi_j((1, \mathbf{z}_{-j}); \mathbf{k}) \geq \pi_j((0, \mathbf{z}_{-j}); \mathbf{k}) \Longrightarrow \pi_i((1, \mathbf{z}_{-i}); \mathbf{k}) \geq \pi_i((0, \mathbf{z}_{-i}); \mathbf{k}).$$

That is, the smaller firm has the stronger incentive to leave the cartel.

Proof. The proof is similar to that of the previous lemma. ∎

We are now in the position to state and prove the main result of this section.

Proposition 8 *There always exists a nonempty stable cartel M^* such that $k_i \geq k_j$ for all $i \in M^*$ and $j \in N\backslash M^*$.*

Proof. We proceed by applying the algorithm in the proof of proposition 2 in decreasing order of capacity. Let us start with the participation decision of the largest firm. Since $\pi_i((1, 0, ..., 0); \mathbf{k}) = \pi_i((0, 0, ..., 0); \mathbf{k})$, a cartel consisting of the largest firm only is internally stable. If the next largest firm has no incentive to join the cartel, then, from lemma 9 no smaller firm has an incentive to do so. Hence, the cartel consisting of the largest firm is externally stable. In this case, the algorithm stops since we have found a nonempty stable cartel. Suppose now that the second largest firm has an incentive to join the cartel. But this means that it has no incentive to leave the cartel, nor has, from lemma 10, the largest firm. The cartel consisting of the two largest firms is thus internally stable. If the next largest firm does not want to join the cartel, then it is externally stable as well. In this case, the algorithm stops; otherwise, it continues in the same fashion. The algorithm is finite since either it stops before all firms have decided to join the cartel or it does not, in which case the cartel consisting of all firms in the industry is found to be stable. ■

Proposition 8 thus shows that there always exists an equilibrium in which the cartel consists of the largest firms in the industry, and the fringe of the smallest.[13] The mechanism behind lemmas 9 and 10, and proposition 8, may be more general. A large firm has more incentives to restrict output than a small firm. The reason is that the output restriction leads to a higher price which is more beneficial to larger firms. It is this mechanism that drives, for instance, the prediction of Ghemawat and Nalebuff (1990), namely that, in declining industries, large firms will start decreasing their output first.

Proposition 8 shows that, in equilibrium, cartel members tend to be larger than fringe firms, given the assumed output sharing rule. We already know that a fringe firm's unconstrained equilibrium output is larger than that of a cartel member. Hence, it appears to be quite "likely" that, in equilibrium, fringe firms face a binding capacity constraint, while cartel members are unconstrained. Moreover, the model makes the following empirical prediction.

Proposition 9 *Consider an equilibrium such that $k_i \geq k_j$ for all $i \in M$ and $j \in N \backslash M$. There exists a nonpositive cross-sectional correlation between firm size (as measured by capacity) and the degree of capacity utilisation.*

Proof. Given the assumed output sharing rule of the cartel, all cartel members produce at the same degree of capacity utilisation. All fringe firms have at least the same degree of capacity utilisation, namely for two reasons. First, their unconstrained equilibrium output is larger than that of a cartel member. Second, fringe firms tend to be smaller. Hence, any fringe firm that is smaller than a cartel member produces at a weakly larger degree of capacity utilisation. Since the unconstrained equilibrium output is the same for all fringe firms, the assertion holds also when comparing fringe firms of different sizes. ■

---

[13]A similar result has been shown by Donsimoni (1985) in the context of a (non-game-theoretic) model of price leadership with a competitive fringe. In her model, demand is assumed to be linear, and firms differ in the slope of their linearly increasing marginal cost curve; there are no capacity constraints.

Hence, firms with larger capacities will tend to have lower capacity utilisation rates.

5 Cartel Stability over the Business Cycle

Oligopolistic price and quantity setting over the business cycle has been a major topic of research in theoretical and empirical IO. At the heart of this literature lies the question whether prices in oligopolistic markets tend to vary procyclically or rather countercyclically. Standard oligopoly and monopoly models usually predict either a positive or no correlation between prices and demand. The seminal contribution by Rotemberg and Saloner (1986) was to show that this relationship can be reversed in collusive equilibria because firms tend to have a higher incentive to deviate in states of high demand. Their paper triggered a large and still growing literature which has been mainly concerned with the robustness of Rotemberg and Saloner's predictions; see, for instance, Halti-wanger and Harrington (1991), Kandori (1991), and Bagwell and Staiger (1997).[14] The literature commonly uses an infinitely repeated game where demand follows some (deterministic or stochastic) process; using trigger strategies, all firms in the industry together try to sustain the joint profit maximising price subject to the constraint that no firm has an incentive to deviate. That is, these papers focus on firms' "incentive constraints" for maximum sustainable collusion, and neglect the "participation constraints". Most papers do not consider capacity constraints.[15] Results appear to depend quite delicately on the time-series properties of the assumed process, and the level of the discount factor.

In our static two-stage game, we have investigated the comparative statics of cartel stability and the collusive equilibrium price with respect to capacity K and market size (or demand) S. One might think that this model does not lend itself easily to the analysis of collusion over the business cycle. There is, however, an extremely simple dynamic extension of the model in which the issue can be investigated in a meaningful way; it goes as follows.

Time is discrete, and indexed by t $(t = 1, 2, ...)$; the time horizon is infinite. There are n firms whose objective it is to maximise the sum of discounted profits. In each period, the dynamic game consists of the two stages described in section 2: the participation stage and the output stage. If a firm decides to join the cartel in a given period, then it delegates its output decision for that period to some cartel manager who seeks to maximise the cartel's joint profit in that period. As before, the cartel is assumed to share profit equally. In contrast to the static model, demand (market size) S now follows an arbitrary (discrete-time) stochastic process; capacity level K is assumed to be constant over time. At the start of each period, before firms decide upon joining the cartel, the realisation of S becomes common knowledge. Notice that the only tangible state

---

[14]The influential paper by Green and Porter (1984) is often cited in this context. This is somewhat misleading since their model deals with demand fluctuations that are even ex post unobservable.

[15]An exception is the paper by Staiger and Wolak (1992), where firms have to build capacity from scratch in each period.

variable just prior to the participation decisions is thus given by the current realisation of S; at the start of the output stage, the tangible state can be described by the tuple $(S, \mathbf{z})$, where $\mathbf{z} = (z_1, ..., z_n)$ is again the vector of participation decisions. It is then immediate to see that, in a given period, a vector of participation and output decisions that can be sustained in an SPE of the static two-stage game can also be sustained in a Markov perfect equilibrium of the dynamic game (for the same realisation of S). (Recall that Markov perfection is a refinement of subgame perfection.) Hence, the comparative statics results with respect to S carry over from the static to the dynamic model. From propositions 4 and 5, it follows in particular that the minimum and maximum sustainable cartel sizes and prices are (weakly) procyclical, in the sense of a nonnegative correlation between S and these variables.[16]

It is interesting to compare our results with those of Rotemberg and Saloner (1986). They consider an infinitely repeated price setting game where demand follows an i.i.d. process, and firms face constant marginal cost; there are no capacity constraints. Rotemberg and Saloner investigate the dynamic properties of a particular equilibrium, namely the *symmetric* SPE which attains the highest profit (by the threat of an infinite Nash reversion). To the extent that the static monopoly price tends to be procyclical, there are two opposing forces at work.[17] On the one hand, collusive prices will tend to be procyclical since monopoly price is positively correlated with demand. On the other, the gain from cheating is positively related to the state of demand. (Due to the i.i.d. assumption, the expected future loss from punishment is the same in each state.) In the case of intermediate discount factors, collusive prices will be procyclical in periods of low demand, and countercyclical in high demand states. The reason is that the monopoly price can be sustained only when demand is low. In high demand states, collusive profits are optimally constant so as to keep firms just indifferent between deviating and sticking to the collusive price; to achieve this, prices must be set the lower, the higher is demand.[18]

In the dynamic extension of our model, an increase in demand has two effects as well; however, in contrast to Rotemberg and Saloner's paper, they reinforce each other. For a given degree of collusion (i.e. cartel size), equilibrium price is procyclical; see lemma 4. Now, the maximum degree of collusion is itself procyclical (proposition 4), which reinforces the positve correlation between price and demand (lemma 2).

The (theoretical and empirical) study of different kinds and causes of price wars is a very active area of research, even though there does not appear to be a generally accepted

---

[16]It should be noted, however, that this does not necessarily imply that equilibrium price varies procyclically. For instance, if there is some exogenous "regime" shift from a collusive equilibrium ($m \geq 2$) to a noncollusive equilibrium ($m = 0$), then the equilibrium price might fall despite an increase in demand.

[17]Note, however, that in Rotemberg and Saloner's leading example the monopoly price is independent of the state of demand.

[18]For large discount factors, the monopoly price can be sustained in each state so that prices will tend to vary procyclically; for small discount factors, collusion can not be sustained in any state.

definition of a "price war" in the literature.[19] Let us (somewhat loosely) call a "price war" a situation in which some or all firms in the industry cease "colluding". In this sense, price wars never take place in the model by Rotemberg and Saloner. In our model, however, two quite different kinds of price wars can occur along the equilibrium path. First, a sufficient fall in demand will lead to (more) excess capacity in the industry, which in turn causes a cartel to become unstable. In equilibrium, some or all firms will leave the cartel, and price will collapse. Second, due to the possible multiplicity of equilibria, an exogenous "regime" shift from a more to a less collusive equilibrium can occur. Such a price war might be caused by some change in beliefs of industry participants.

6 Conclusion

The main aim of this paper has been to develop a theoretical foundation of the traditional view in industrial organisation, according to which there exists a negative relationship between the level of excess capacity and cartel stability. In contrast to the supergame-theoretic approach to collusion, we have focussed on firms' incentives to participate in a cartel, rather than on the issue of enforcement of cartel rules. The basic two-stage game predicts a positive (negative) correlation between the equilibrium cartel size and the level of demand (resp. capacity). This reinforces the tendency of prices to be low in periods of high excess capacity. In a simple dynamic extension of the model, we have shown that collusive equilibrium prices will vary procyclically, independently of the assumed stochastic process for demand. Allowing for arbitrary heterogeneity in capacity among firms, the analysis has revealed that smaller firms have stronger incentives to take a free ride on the cartel's effort to restrict output. Hence, the model predicts that firm size and the degree of capacity utilisation tend to be negatively correlated across firms.

Although the model has been cast as a model of cartel formation, it can also be seen as a contribution to the still underdeveloped literature on endogenous horizontal mergers.[20] If mergers are driven by market power rather than efficiency considerations, then the model predicts merger waves to occur during booms.

Several issues have been left open for future research. First, following Selten (1973), we have modelled cartel formation as a noncooperative simultaneous-move game; we have shown that our results are even sharper in the case of sequential cartel formation. It may be fruitful to consider different ways of modelling cartel formation, or to allow for more general coalition structures. Second, as in Brock and Scheinkman (1985), each firm's capacity level was assumed to be exogenously given. Endogenising capacity decisions may provide new insights, as Davidson and Deneckere (1990) and Staiger and Wolak (1992) have shown in the context of supergames. Third, it would be interesting to study the interaction of firms' incentive and participation constraints, thereby combining the

---

[19]An outstanding example is Levenstein's work on the Bromine Cartel; see Levenstein (1996,1997).

[20]The seminal paper in this literature is Kamien and Zang (1990).

supergame-theoretic literature with the present approach.

7 Appendix

Proof of lemma 1. Let $r(m,k) \equiv P(nk) - c + mkP'(nk)$ and $s(m,k) \equiv P((n-m+1)k) - c + kP'((n-m+1)k)$. At cartel size m, $m \in [1,n]$, a cartel member (fringe firm) faces a binding capacity constraint if and only if $r(m,k) > 0$ ($s(m,k) > 0$). Note that $dr(m,k)/dm < 0$, $dr(m,k)/dk < 0$, $ds(m,k)/dm > 0$, and $ds(m,k)/dk < 0$, for all real numbers $m \in [1,n]$ and $k \in (0,\infty)$. Hence, if k is such that $r(n,k) \leq 0 \leq r(1,k)$, then $m_c(k)$, $m_c(k) \in [1,n]$, is uniquely defined by $r(m_c(k),k) = 0$. We thus have $r(m,k) > 0$ if and only if $m > m_c(k)$. From the implicit function theorem, it follows that $dm_c(k)/dk < 0$ in this domain. Outside this domain, the function $m_c(k)$ can be extended in an arbitrary way as long as it satisfies $dm_c(k)/dk < 0$ for all $k \in (0,\infty)$. Similarly, if k is such that $s(1,k) \leq 0 \leq s(n,k)$, then $m_f(k)$, $m_f(k) \in [1,n]$, is uniquely defined by $s(m_f(k),k) = 0$; we have $dm_f(k)/dk > 0$. Outside this domain, $m_f(k)$ can be extended in an obvious way. Finally, observe that $r(1,k) = s(1,k)$, which implies the last assertion of the lemma. ∎

Proof of lemma 2. As in subsection 3.1, we have to distinguish between three cases, depending on whether the fringe firms and cartel members face a binding capacity constraint. Since $q_f(0;k) = q_f(1;k) = q_M(1;k)$, we restrict attention to real numbers $m \in [1,n)$.

Case (i): Suppose capacity constraints are nonbinding for all firms, i.e. condition (4) holds. Then, equilibrium output is such that $q_c(m;k) = q_f(m;k)/m$, where $q_f(m;k)$ is defined by (5). Implicit differentiation yields

$$\frac{\partial q_f(m;k)}{\partial m} = \frac{q_f}{n-m+2} \cdot \frac{P' + q_f P''}{P' + \frac{n-m+1}{n-m+2}q_f P''} \in \left(0, \frac{q_f}{n-m+1}\right).$$

Hence, a fringe firm's output and joint cartel output are increasing in m. Differentiating total industry output with respect to m gives

$$\frac{\partial}{\partial m}(n-m+1)q_f(m;k) = -q_f + (n-m+1)\frac{\partial q_f}{\partial m} < 0;$$

that is, total industry output is negatively related to cartel size.

Case (ii): Suppose now that capacity constraints are binding for fringe firms but nonbinding for cartel members. Then, $q_f(m;k) = k$, which is independent of m, and $q_c(m;k)$ is defined by (10). Implicit differentiation of joint cartel output gives

$$\frac{\partial}{\partial m}q_M(m;k) = k\frac{P' + q_M P''}{2P' + q_M P''} \in (0,k);$$

24

joint cartel output is thus an increasing function of m. To see that total industry output is negatively related to m, note that

$$\frac{\partial}{\partial m}\{q_M(m;k) + (n-m)k\} = \frac{\partial}{\partial m}q_M - k < 0.$$

Case (iii): If capacity constraints are binding for all firms, then a marginal change in m has no effect on output. ∎

Proof of lemma 7. In the following, we distinguish between different cases, depending on whether fringe firms face binding capacity constraints when cartel size is m or $m-1$, and whether cartel members are constrained when cartel size is $m-1$. Notice that the assumption $m \geq \overline{m}_c(k)$ implies that $m > m_c(k)$, i.e. cartel members are unconstrained when cartel size is m. For expositional clarity, we suppose below that $m, m-1 \neq m_c(k), m_f(k)$ so that the derivative of profit with respect to k exists. Let us first discuss the cases where the derivative does not exist. Suppose $m-1 = m_c(k)$. Since we have assumed that $m \in [\overline{m}_c(k)+1, n]$, this is possible only if $m_c(k) \in [1, n-1]$. Now, from the proof of lemma 3, we know that $\pi_c(m;k) > \pi_f(m-1;k)$ if $m-1 = m_c(k) \in (1, n-1]$; but this contradicts the assumption $\pi_c(m;k) > \pi_f(m-1;k)$. By continuity, $\pi_c(m;k) \geq \pi_f(m-1;k)$ if $m-1 = m_c(k) = 1 = m_f(k)$. Hence, if indeed $\pi_c(2;k) = \pi_f(1;k)$ at this value of k, then the left-hand derivative of $\pi_c(2;k) - \pi_f(1;k)$ with respect to k must be strictly negative. The sign of the derivative from the right at this point will be dealt with below in case *(b)*, where $m-1 > m_c(k)$ and $m > m_f(k) > m-1$. Suppose now that $m = m_f(k)$ $(m-1 = m_f(k))$. Since $m_f(k)$ is strictly increasing in k, the derivate from the left will be dealt with in the case where $m > m_f(k)$ (resp. $m-1 > m_f(k)$), and the derivative from the right in the case where $m < m_f(k)$ (resp. $m-1 < m_f(k)$).

Case (a): Suppose $m-1 > m_c(k)$ and $m < m_f(k)$. This implies that capacity constraints are nonbinding for both fringe and cartel members when there are $m-1$ or m cartel members. Then, obviously, a marginal increase in the capacity level k has no effect on profits.

Case (b): Suppose that $m-1 > m_c(k)$ and $m > m_f(k) > m-1$. In this case, capacity constraints do not matter when cartel size is $m-1$ so that $\partial \pi_f(m-1;k)/\partial k = 0$. When there are m cartel members, equilibrium profit is given by (11). Using the envelope theorem, we get $\partial \pi_c(m;k)/\partial k = (n-m)q_c(m;k)P'(mq_c(m;k) + (n-m)k) \leq 0$, and hence $\partial \{\pi_c(m;k) - \pi_f(m-1;k)\}/\partial k \leq 0$.

Case (c): Suppose now that $m-1 > m_c(k)$ and $m-1 > m_f(k)$. Then, equilibrium profits are, from (11) and (12),

$$\pi_c(m;k) = [P(mq_c(m;k) + (n-m)k) - c]q_c(m;k)$$

and

$$\pi_f(m-1;k) = [P((m-1)q_c(m-1;k) + (n-m+1)k) - c]k,$$

25

where $q_c(m; k)$ is implicitly defined by (10), and $q_c(m-1; k)$ is defined analogously. Let us now make two observations, namely

$$
\begin{aligned}
\frac{\partial q_c(m-1; k)}{\partial k} &= -\frac{n-m+1}{m-1} \cdot \frac{P' + (m-1)q_c(m-1; k)P''}{2P' + (m-1)q_c(m-1; k)P''} \\
&< -\frac{(n-m+1)^2 k}{(m-1)[(m-1)q_c(m-1; k) + 2(n-m+1)k]}
\end{aligned} \tag{18}
$$

and

$$
mq_c(m; k) \le (m-1)q_c(m-1; k) + k. \tag{19}
$$

To see inequality (18), observe first that inequality (1) implies

$$
(m-1)q_c(m-1; k)P'' < -\frac{(m-1)q_c(m-1; k)}{(m-1)q_c(m-1; k) + (n-m+1)k}P',
$$

where both P' and P'' have to be evalued at $(m-1)q_c(m-1; k) + (n-m+1)k$. We then get

$$
\frac{P' + (m-1)q_c(m-1; k)P''}{2P' + (m-1)q_c(m-1; k)P''} > \frac{(n-m+1)k}{(m-1)q_c(m-1; k) + 2(n-m+1)k},
$$

using the fact that the ratio on the l.h.s. is decreasing in $(m-1)q_c(m-1; k)P''$. Inequality (18) follows immediately.

To see inequality (19), note that, by assumption, we have $q_c(m; k) \le k$ and $\pi_c(m; k) = \pi_f(m-1; k)$ so that $P(mq_c(m; k) + (n-m)k) \ge P((m-1)q_c(m-1; k) + (n-m+1)k)$, and hence the result follows.

Let us now take the derivative of $\Delta\pi(k) \equiv \{\pi_c(m; k) - \pi_f(m-1; k)\}$ with respect to k:

$$
\begin{aligned}
\Delta\pi'(k) &= (n-m)q_c(m; k)P'(mq_c(m; k) + (n-m)k) \\
&\quad - [P((m-1)q_c(m-1; k) + (n-m+1)k) - c] \\
&\quad - P'((m-1)q_c(m-1; k) + (n-m+1)k)k \\
&\quad \times \left[(m-1)\frac{\partial q_c(m-1; k)}{\partial k} + n - m + 1\right],
\end{aligned}
$$

where we have made use of the envelope theorem. Now, using (18) and the fact that, by assumption, m and k are such that $\pi_c(m; k) = \pi_f(m-1; k)$, we get

$$
\begin{aligned}
\Delta\pi'(k) &\le [P((m-1)q_c(m-1; k) + (n-m+1)k) - c] \\
&\quad \times \left[-\frac{(n-m)k}{mq_c(m; k)} - 1\right. \\
&\quad \left. +\frac{(n-m+1)^2\left[k + \frac{(m-1)q_c(m-1; k)}{(n-m+1)}\right]k}{(m-1)^2 q_c^2(m-1; k) + 2(m-1)(n-m+1)kq_c(m-1; k)}\right].
\end{aligned}
$$

26

Using (19), and $(m-1)q_c(m-1;k) \geq k$, one can show that the expression on the r.h.s. is negative. Thus, we obtain $\partial \{\pi_c(m;k) - \pi_f(m-1;k)\}/\partial k < 0$ (where this derivative exists).

Case (d): Suppose $m-1 < m_c(k) < m$ and $m-1 > m_f(k)$. Equilibrium profits are then given by $\pi_c(m;k) = [P(mq_c(m;k) + (n-m)k) - c]\, q_c(m;k)$ and $\pi_f(m-1;k) = [P(nk) - c]\, k$, where $q_c(m;k)$ is implicitly defined by (10). Since $m > m_c(k)$ by assumption, we must have $q_c(m;k) < k$. Then, joint cartel profit must be strictly higher at a joint cartel output of $mq_c(m;k)$ than at mk. That is, $[P(mq_c(m;k) + (n-m)k) - c]\, mq_c(m;k) > [P(nk) - c]\, mk$, but this contradicts the assumption $\pi_c(m;k) = \pi_f(m-1;k)$. Hence, this case can not occur. ∎

Proof of proposition 6. Suppose we are at the start of "substage" l, $l \in \{1,...,n\}$, i.e. just prior to firm l's participation decision. Define $z(l) \equiv \sum_{i=1}^{l} z_i$ for $l \in \{1,...,n\}$, and $z(0) \equiv 0$. Assume $z(l-1) \geq \overline{m}^*(k) - (n-l+1)$, i.e. if firms l to n all choose to join the cartel, then the resulting cartel size is at least $\overline{m}^*(k)$. Then, there exists a unique SPE of the ensuing subgame such that $z(l-1) + \sum_{i=l}^{n} z_i^* = \min\{\overline{m}^*(k), z(l-1)\}$, where z_i^* is firm i's participation decision along the equilibrium path. In equilibrium,

$$z_l^* = \begin{cases} 1 & \text{if } z(l-1) = \overline{m}^*(k) - (n-l+1) \\ 0 & \text{if } z(l-1) > \overline{m}^*(k) - (n-l+1). \end{cases}$$

The proposition follows by setting $l = 1$. We prove the assertion by induction.

Assume $l = n$. If $z(n-1) \geq \overline{m}^*(k)$, then $z_n^* = 0$ since all cartel sizes above $\overline{m}^*(k)$ are internally unstable (lemma 6), i.e. $\pi_f(m;k) > \pi_c(m+1;k)$ for all $m \in \{\overline{m}^*(k)+1,...,n\}$. If, however, $z(n-1) = \overline{m}^*(k) - 1$, then $z_i^* = 1$ since $\pi_c(\overline{m}^*(k);k) > \pi_f(\overline{m}^*(k)-1;k)$ by assumption. Hence, the assertion holds for $l = n$. Assume now that the assertion is satisfied for all $l \in \{\widehat{l}+1,...,n\}$, where $\widehat{l} \in \{1,...,n-1\}$. We now want to prove that it is then still satisfied for $l = \widehat{l}$. Now, if $z(\widehat{l}-1) \geq \overline{m}^*(k)$, then clearly $z_{\widehat{l}}^* = 0$ since $\pi_f(m;k) > \pi_c(m+1;k)$ for all $m \in \{\overline{m}^*(k)+1,...,n\}$. If $z(\widehat{l}-1) = \overline{m}^*(k) - (n-\widehat{l}+1)$, then firm $\widehat{l}$'s payoff from joining the cartel is $\pi_c(\overline{m}^*(k);k)$, whereas from joining the fringe it is *at most* $\pi_f(\overline{m}^*(k)-1;k)$ recalling that $\pi_f(m;k)$ is strictly increasing in m. By assumption, $\pi_c(\overline{m}^*(k);k) > \pi_f(\overline{m}^*(k)-1;k)$. Hence, $z_{\widehat{l}}^* = 1$ in this case. If, however, $z(\widehat{l}-1) \in \{\overline{m}^*(k) - (n-\widehat{l}), \overline{m}^*(k) - 1\}$, then firm $\widehat{l}$'s profit from joining the cartel is $\pi_c(\overline{m}^*(k);k)$; the payoff from joining the fringe is $\pi_f(\overline{m}^*(k);k)$, which is strictly more. In equilibrium, $z_{\widehat{l}}^* = 0$ in this case. This proves the assertion. ∎

Proof of lemma 8. Suppose first that $k_i \geq \overline{x}(\mathbf{z};\mathbf{k})$. Then, an increase in k_i has no effect on $q(\mathbf{z};\mathbf{k})$, and hence none on firm i's profit. A rise in k_j, $j \neq i$, weakly increases $q(\mathbf{z};\mathbf{k})$, as can be seen by examining the implicit function $h(q(\mathbf{z};\mathbf{k})) = 0$. Assuming the derivatives exist,

$$\frac{d\pi_i(\mathbf{z};\mathbf{k})}{dk_j} = \frac{d}{dq} \{\overline{x}(q(\mathbf{z};\mathbf{k}))\, [P(q(\mathbf{z};\mathbf{k})) - c]\} \frac{dq(\mathbf{z};\mathbf{k})}{dk_j}$$

$$= \{\overline{x}'(q(\mathbf{z};\mathbf{k}))\left[P(q(\mathbf{z};\mathbf{k})) - c\right] + \overline{x}(q(\mathbf{z};\mathbf{k}))P'(q(\mathbf{z};\mathbf{k}))\}\frac{dq(\mathbf{z};\mathbf{k})}{dk_j} \le 0,$$

since $\overline{x}'(q(\mathbf{z};\mathbf{k})) < 0$. Suppose now that $k_i < \overline{x}(\mathbf{z};\mathbf{k})$. Again assuming the derivatives exist, we have

$$
\begin{aligned}
\frac{d\pi_i(\mathbf{z};\mathbf{k})}{dk_i} &= \frac{d}{dk_i}k_i\left[P(q(\mathbf{z};\mathbf{k})) - c\right] \\
&= \left[P(q(\mathbf{z};\mathbf{k})) - c\right] + k_i P'(q(\mathbf{z};\mathbf{k}))\frac{dq(\mathbf{z};\mathbf{k})}{dk_i} \\
&\ge \left[P(q(\mathbf{z};\mathbf{k})) - c\right] + k_i P'(q(\mathbf{z};\mathbf{k})) > 0,
\end{aligned}
$$

where the first inequality follows from $dq(\mathbf{z};\mathbf{k})/dk_i \in (0,1]$, and the second from the fact that $k_i < \overline{x}(\mathbf{z};\mathbf{k})$. Similarly,

$$\frac{d\pi_i(\mathbf{z};\mathbf{k})}{dk_j} = k_i P'(q(\mathbf{z};\mathbf{k}))\frac{dq(\mathbf{z};\mathbf{k})}{dk_j} \le 0.$$

The same analysis applies to the cartel's joint profit by simply changing the indices. ∎

Proof of lemma 9. Suppose first that neither firm i nor firm j face a binding capacity constraint if $z_i = z_j = 0$; hence, $q_i((0,\mathbf{z}_{-i});\mathbf{k}) = q_j((0,\mathbf{z}_{-j});\mathbf{k}) = \overline{x}(q(\mathbf{z};\mathbf{k}))$, where $q(\mathbf{z};\mathbf{k}) \equiv \sum_{i \in N} q_i((z_i,\mathbf{z}_{-i});\mathbf{k})$ is aggregate industry output. This implies that both firms make the same profit in the fringe, i.e. $\pi_i(\mathbf{z};\mathbf{k}) = \pi_j(\mathbf{z};\mathbf{k})$. But then, firm i has a (weakly) higher incentive to deviate since $\pi_i((1,\mathbf{z}_{-i});\mathbf{k}) \ge \pi_j((1,\mathbf{z}_{-j});\mathbf{k})$. To see this notice that $\pi_i((1,\mathbf{z}_{-i});\mathbf{k}) = k_i/\left(k_{M(\mathbf{z})} + k_i\right)\pi_M((1,\mathbf{z}_{-i});\mathbf{k})$, where $k_{M(\mathbf{z})}$ is the cartel's joint capacity given $\mathbf{z}$, i.e. excluding firms i and j. Moreover, $\pi_M((1,\mathbf{z}_{-i});\mathbf{k}) \ge \pi_M((1,\mathbf{z}_{-j});\mathbf{k})$ since the cartel's joint profit is weakly increasing in its own capacity, and weakly decreasing in the capacity of a fringe firm; see lemma 8.

Suppose now that both firms i and j face a binding capacity constraint if $z_i = z_j = 0$; hence, $q_i((0,\mathbf{z}_{-i});\mathbf{k}) = k_i$, and $q_j((0,\mathbf{z}_{-j});\mathbf{k}) = k_j$. Since i and j are the largest fringe firms, all other fringe firms are then constrained as well. We need to show that if

$$\frac{k_i}{k_{M(\mathbf{z})} + k_i}\pi_M((1,\mathbf{z}_{-i});\mathbf{k}) - k_i\left[P(q(\mathbf{z};\mathbf{k})) - c\right] < 0,$$

then

$$\frac{k_j}{k_{M(\mathbf{z})} + k_j}\pi_M((1,\mathbf{z}_{-j});\mathbf{k}) - k_j\left[P(q(\mathbf{z};\mathbf{k})) - c\right] < 0. \tag{20}$$

Comparing the two inequalities, one sees that it is sufficient to show the following inequality:

$$\frac{\pi_M((1,\mathbf{z}_{-i});\mathbf{k})}{k_{M(\mathbf{z})} + k_i} \ge \frac{\pi_M((1,\mathbf{z}_{-j});\mathbf{k})}{k_{M(\mathbf{z})} + k_j}. \tag{21}$$

28

Notice that firms i and j continue to face a binding constraint in the fringe if the other firm deviates and joins the cartel. To see this notice that, from the implicit function $h(q(\mathbf{z}; \mathbf{k})) = 0$, such a deviation will weakly decrease industry output $q(\mathbf{z}; \mathbf{k})$, and hence weakly increase $\overline{x}(q(\mathbf{z}; \mathbf{k}))$. In order to show inequality (21), we adopt a differential approach. If $q_M((1, \mathbf{z}_{-i}); \mathbf{k}) = \overline{x}(q((1, \mathbf{z}_{-i}); \mathbf{k}))$, then one can prove that

$$\frac{d}{dk_i} \left(\frac{\pi_M((1, \mathbf{z}_{-i}); \mathbf{k})}{k_{M(1, \mathbf{z}_{-i})}} \right)\bigg|_{k_i + k_j = const.}$$

$$\geq \frac{k_{M(1, \mathbf{z}_{-i})} \left(-\overline{x}(q((1, \mathbf{z}_{-i}); \mathbf{k})) P'(q((1, \mathbf{z}_{-i}); \mathbf{k})) \right) - \pi_M((1, \mathbf{z}_{-i}); \mathbf{k})}{\left(k_{M(1, \mathbf{z}_{-i})} \right)^2}$$

$$= \frac{\left(k_{M(1, \mathbf{z}_{-i})} - \overline{x}(q((1, \mathbf{z}_{-i}); \mathbf{k})) \right) \left[P(q((1, \mathbf{z}_{-i}); \mathbf{k})) - c \right]}{\left(k_{M(1, \mathbf{z}_{-i})} \right)^2} \geq 0$$

which implies inequality (21). If, instead, $q_M((1, \mathbf{z}_{-i}); \mathbf{k}) = k_{M(1, \mathbf{z}_{-i})} < \overline{x}(q((1, \mathbf{z}_{-i}); \mathbf{k}))$, then this derivative is zero, since aggregate output is independent of how capacity is distributed between a fringe firm and the cartel if both face a binding capacity constraint.

Finally, suppose that firm j, but not firm i, faces a binding capacity constraint if $z_i = z_j = 0$; hence, $q_i((0, \mathbf{z}_{-i}); \mathbf{k}) = \overline{x}(q(\mathbf{z}; \mathbf{k}))$, and $q_j((0, \mathbf{z}_{-j}); \mathbf{k}) = k_j$. First, observe that firm j's incentive to deviate, as represented by the l.h.s. of (20), is weakly decreasing in k_i for $k_i \geq \overline{x}(q(\mathbf{z}; \mathbf{k}))$; the first term in (20) is weakly decreasing in k_i, and the second term is independent of k_i for $k_i \geq \overline{x}(q(\mathbf{z}; \mathbf{k}))$. Second, notice that firm i's incentive to deviate,

$$\frac{k_i}{k_{M(\mathbf{z})} + k_i} \pi_M((1, \mathbf{z}_{-i}); \mathbf{k}) - \overline{x}(q(\mathbf{z}; \mathbf{k})) \left[P(q(\mathbf{z}; \mathbf{k})) - c \right],$$

is strictly increasing in k_i for $k_i \geq \overline{x}(q(\mathbf{z}; \mathbf{k}))$. To see this, observe that the ratio of capacities is strictly increasing in k_i, whereas all other terms are independent of k_i for $k_i \geq \overline{x}(q(\mathbf{z}; \mathbf{k}))$. Hence, it is sufficient to prove that firm i has a stronger incentive to join the cartel than firm j if $k_i = \overline{x}(q(\mathbf{z}; \mathbf{k}))$. But this has already been shown. ∎

References

Bagwell, K. and R. W. Staiger (1997). Collusion over the Business Cycle. *Rand Journal of Economics 28*(1), 82–106.

Brock, W. A. and J. A. Scheinkman (1985). Price-setting Supergames with Capacity Constraints. *Review of Economic Studies 52*(3), 371–382.

d'Aspremont, C., A. Jacquemin, J. J. Gabszewicz, and J. A. Weymark (1983). On the Stability of Collusive Price Leadership. *Canadian Journal of Economics 16*(1), 17–25.

Davidson, C. and R. Deneckere (1990). Excess Capacity and Collusion. *International Economic Review 31*(3), 521–541.

Donsimoni, M.-P. (1985). Stable Heterogenous Cartels. *International Journal of Industrial Organization 3*(4), 451–467.

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

Genesove, D. and W. Mullin (1998). Narrative Evidence on the Dynamics of Collusion: The Sugar Institute Case. Mimeo. MIT and Michigan State University.

Ghemawat, P. and B. Nalebuff (1990). The Devolution of Declining Industries. *Quarterly Journal of Economics 105*(1), 165–186.

Green, E. J. and R. H. Porter (1984). Noncooperative Collusion under Imperfect Price Information. *Econometrica 52*(1), 87–100.

Haltiwanger, J. and J. E. Harrington (1991). The Impact of Cyclical Demand Movements on Collusive Behavior. *Rand Journal of Economics 22*(1), 89–106.

Kamien, M. I. and I. Zang (1990). The Limits of Monopolization through Acquisition. *Quarterly Journal of Economics 105*(2), 465–499.

Kandori, M. (1991). Correlated Demand Shocks and Price Wars During Booms. *Review of Economic Studies 58*(1), 171–180.

Kihlstrom, R. and X. Vives (1992). Collusion by Asymmetrically Informed Firms. *Journal of Economics & Management Strategy 1*(2), 371–396.

Levenstein, M. C. (1996). Do Price Wars Facilitate Collusion? A Study of the Bromine Cartel before World War I. *Explorations in Economic History 33*(1), 107–137.

Levenstein, M. C. (1997). Price Wars and the Stability of Collusion: A Study of the Pre-World War I Bromine Industry. *Journal of Industrial Economics 45*(2), 117–137.

Phlips, L. (1995). *Competition Policy: A Game-Theoretic Perspective.* Cambridge: Cambridge University Press.

Roberts, K. (1985). Cartel Behaviour and Adverse Selection. *Journal of Industrial Economics 33*(4), 401–413.

Rotemberg, J. J. and G. Saloner (1986). A Supergame-Theoretic Model of Price Wars during Booms. *American Economic Review 76*(3), 390–407.

Salant, S. W., S. Switzer, and R. J. Reynolds (1983). Losses from Horizontal Merger: The Effects of an Exogenous Change in Industry Structure on Cournot-Nash Equilibrium. *Quarterly Journal of Economics 98*(2), 185–199.

Scherer, F. M. (1980). *Industrial Market Structure and Economic Performance* (Second ed.). Boston: Houghton Mifflin.

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*(3), 744–754.

Staiger, R. W. and F. A. Wolak (1992). Collusive Pricing with Capacity Constraints in the Presence of Demand Uncertainty. *Rand Journal of Economics 23*(2), 203–220.