

HOW MARKET FRAGMENTATION CAN FACILITATE COLLUSION

Kai-Uwe Kühn

European Commission
and University of Michigan

Abstract

Economists have recommended the fragmentation of capacities before regulated markets are liberalized because static oligopoly models imply that outcomes approximate perfect competition with a fragmented enough market structure. This intuition fails under collusion. When individual firms are capacity constrained relative to total demand, the fragmentation of capacity facilitates collusion and increases the highest sustainable collusive price. Collusive outcomes remain feasible even for arbitrarily fragmented capacity. These results can explain the finding in Sweeting (2007, *Economic Journal*, 117, 654–685), that dramatic fragmentation of generation capacity in the English electricity industry did not reduce price cost margins. (JEL: J1, J11)

1. Introduction

One of the most robust results in economic theory is that market power tends to be reduced when the supply side of a market becomes fragmented relative to the demand side. In static homogeneous goods models a replication of the industry (replicating both the demand and the supply side) will lead to price moving closer to marginal cost. At the limit of infinite replication, price converges to marginal cost.¹ As long as splitting up given capacity among more firms does not have efficiency effects, we should therefore expect lower prices when an industry becomes more fragmented. Following Scherer and Ross (1990) market fragmentation is also taken as the most important factor undermining the ability of firms to collude. A recent antitrust textbook

The editor in charge of this paper was Xavier Vives.

Acknowledgments: I would like to thank Chris House, Ryan Kellogg, Pierre Regibeau, Steve Salant, Andrew Sweeting, Xavier Vives, and two anonymous referees for helpful comments and valuable advice on this paper. I have also benefited from comments of participants at the 2007 Workshop in Industrial Organization and Management Strategy at SHUFE, Shanghai, and IIOC 2009 at Northeastern University, Boston. The paper reflects only the views of the author and not those of the European Commission.

E-mail: kai-uwe.kuhn@ec.europa.eu

1. Amir and Lambson (2000) seemingly contradict this statement. They derive a result that under Cournot competition with decreasing marginal costs an increase in the number of firms leads to higher prices. The reason is that each remaining firm will exploit economies of scale less. However, the result is driven by the fact that the demand side is not replicated.

(Motta 2004, p. 142) states unambiguously: “Other things being equal, collusion is the more likely the smaller the number of firms in the industry.”²

Such strong predictions of theory invariably find their expression in policy advice. Especially in industries that have recently undergone regulatory reform and liberalization, it has been a mantra of economists to call for the breakup of firms before liberalization. One example is the pioneering liberalization of electricity generation in England and Wales, which was criticized for failing to fragment capacity holdings before liberalization (see Henney 1987; Sykes and Robinson 1987).³ Calibrations of supply function equilibrium in the electricity spot market for England and Wales by Green and Newbery (1992) concluded that the creation of a fifth competitor would have significantly lowered price cost margins and noticeably enhanced welfare (Green 1996). These criticisms together with evidence for persistent market power finally made the electricity regulator for England and Wales force divestitures on the two largest firms in the industry, and to the progressive fragmentation of electricity generation in England and Wales between 1995 and 2000. The industry saw the divestiture of significant assets to three new players in the market and a dramatic fall in the Herfindahl–Hirshman Index from 2,264 to 1,000. Surprisingly, these developments did not have the expected impact. Sweeting (2007) shows that market power measured in terms of price–cost margins did not fall, but slightly increased, during this period. He conjectures that one possible reason for this effect is the impact of tacit collusion.⁴

In this paper we resolve this apparent puzzle by re-examining homogeneous goods models with capacity constraints. We show that convergence of all collusive equilibria to the competitive outcome always fails in these models: in the limit of an arbitrarily fragmented industry structure there is always a collusive equilibrium with strictly less production than the competitive output. Furthermore, in Bertrand–Edgeworth models, collusion is facilitated when a given amount of capacity is divided among a larger number of firms as long as a single firm cannot serve the whole market at

2. It is often not made clear what the *ceteris-paribus* condition in such a statement means. In some papers total capacity is increased as the number of firms is increased while keeping demand constant (see Brock and Scheinkman 1985). This is sometimes seen as satisfying the *ceteris-paribus* condition because the size of the firms does not change. However, this comparative static does not look at the pure effect of the number of players because capacity relative to demand is increased. To avoid this effect and isolate the effect of having a large number of firms, both the demand and the supply side of the market are usually replicated at the same rate when economists study the convergence of behavior to perfect competition in work on core convergence, static oligopoly models (see Novshek and Sonnenschein 1978), and dynamic oligopoly models (see Green 1980). In our paper we change the number of firms, keeping the total capacity constant. This generates the same incentives for oligopolists as the standard approach when technology has constant returns to scale. Our approach to convergence turns out to be the relevant one for the comparative statics in n for the analysis of merger policy, divestitures, and the effect of endogenous entry.

3. Similar concerns were expressed by US economists in the early 1980s, who advocated a cautious approach to electricity deregulation pointing to persistent high effective concentration in generation (see Schmalensee and Golub 1984). The Spanish electricity liberalization in 1998 was severely criticized because the government allowed considerable consolidation by a state-owned generation company before privatization and liberalization occurred (see Kühn and Regibeau 1998; Arocena, Kühn, and Regibeau 1999).

4. Dissatisfaction with the performance of the industry has led to further reform. In 2001, the England and Wales spot market was replaced by the New Electricity Trading Arrangement (NETA), which was hoped to produce more competitive outcomes.

the collusive price. In fact, under the most realistic assumptions on the relationship between capacity and demand, the highest achievable price will monotonically rise as the industry becomes more fragmented.⁵

We derive these results for a set of simple models with n symmetric firms, each owning $1/n$ th of the total capacity K in the market. Firms compete in prices but may be constrained by their capacity so they cannot serve all of the demand at the competitive price. The intuition for the result that increased market fragmentation at given total capacity enhances collusion is relatively simple: Suppose that no firm in the industry can serve all the demand at the monopoly price. Then any firm deviating from a collusive price will be capacity constrained. The gain from deviation in a symmetric model is then proportional to the capacity share of the firm: $1/n$. If the punishment technology is such that profits in the punishment continuation equilibrium decrease at a rate faster than $1/n$, the benefits of deviating from collusion decrease relative to the losses from increased punishment for deviations.

In almost all models in which collusion is generated from a switch from a collusive to a punishment continuation equilibrium, punishment profits will decline in n at a rate faster than $1/n$. In Section 3 we show that this also holds for the infinitely repeated Bertrand–Edgeworth game with efficient rationing (see Brock and Scheinkman 1985; Davidson and Deneckere 1990).⁶ The reason is that competition is intensified in the punishment equilibrium when n is higher: when a smaller firm deviates from punishments a larger proportion of the total capital K will be fully utilized. Hence, punishment profits fall more than proportionately as n increases and market fragmentation strictly facilitates collusion. The highest sustainable collusive price will therefore increase as markets become more fragmented as long as there is no firm that can serve the whole market at the monopoly price. In contrast, capacity fragmentation will *reduce* the ability to collude when a single firm *can* serve the full monopoly output because the above argument breaks down. However, in any market collusion will eventually become easier as capacity is more fragmented because small enough firms are always constrained relative to market demand at the monopoly price. Indeed, we show that there are many cases in which collusion in a market with very fragmented capacity is easier to sustain than in a homogeneous goods duopoly model *without* capacity constraints.

Our results generate qualitatively the opposite results to those derived by Brock and Scheinkman (1985). First, in their paper the ability to collude disappears as the number of firms gets large. Second, increases in the number of firms facilitate collusion only when the number of firms is small. Eventually, for a large enough number of firms, an increase in the number of firms decreases the ability to collude. The difference arises because they consider a fundamentally different comparative statics result: they

5. Green (1980) was the first to show that for dynamic oligopoly models with perfect observability convergence to competition can fail. We show here that it robustly does so for models with capacity constraints.

6. This property extends to the general class of rationing rules in Lambson (1987) so that all of the results of Section 3 can be generalized to this class of rationing rules. Proofs are available from the author upon request.

increase the number of firms holding capacity *per firm* fixed. This means that capacity increases relative to the demand in the market as n increases. As the number of firms increases, the market share of a firm in any collusive equilibrium goes down while its available capacity stays the same. The incentive to deviate therefore increases in n , in contrast to our model. More generally, the difference between their model and ours arises because our result obtains when capacity relative to the size of the market weakly declines in n while their result obtains when capacity strictly increases relative to the size of the market when n is increased.

One of the contributions of this paper is to point out that for policy issues and the interpretation of cross-sectional work on collusion the case in which capacity (weakly) decreases relative to the size of the market as n increases is the relevant case. This is obvious for assessing divestitures (or mergers) because in such situations a *given amount of total capacity* is broken up (or consolidated) while leaving market size unchanged. But ours is also the appropriate theory to understand the implications of endogenous entry on the interpretation of the results from cross-sectional studies about the relationship between market fragmentation and collusion. With endogenous entry, a firm with given capacity can enter the market when market size increases the right amount, leaving capacity relative to the size of the market unchanged. We show in the paper that in a free-entry equilibrium that is followed by an infinitely repeated game, total capacity decreases relative to the size of the market so that larger markets with more firms will exhibit more collusion. This rationalizes the puzzling results in the literature on cross-sectional studies of collusion that the duration of cartels increases with their size (Posner 1970; Dick 1996; Levenstein and Suslow 2006).

In Section 2 we use the simplest conceivable framework for a collusion model to give an intuition for the result that capacity fragmentation will lead to more collusive outcomes. Section 3 specifically looks at a standard infinitely repeated Bertrand–Edgeworth model with efficient rationing to derive complete comparative statics results for prices with respect to capacity fragmentation. In Section 4, we discuss the robustness of these results to model specification and show that for high-capacity fragmentation there are always collusive outcomes, independently of the competitive variable assumed. However, we also show that convergence does not have to be monotonic. Section 5 concludes by drawing out the implications of the analysis for issues in antitrust and regulatory policy design as well as the implications for cross-sectional empirical studies of collusion.

2. How Capacity Fragmentation Facilitates Collusion

Consider a single homogeneous good with a market demand function $D(p)$, which is strictly decreasing, log-concave, and has $\lim_{p \rightarrow \infty} D(p) = 0$. For the production of the homogeneous good there is a total amount of K units of capital available. There are n firms in the industry, each owning K/n units of capital. Each firm has the same cost function with constant marginal cost c up to capacity K/n . The unconstrained monopoly price in the industry, p^m , solves $\max_p \pi(p) = \max_p (p - c)D(p)$. To make

the problem nontrivial we assume that there is sufficient capacity in the market to serve demand at the monopoly price: $D(p^m) < K$. We denote the competitive price by \bar{p} . It solves $D(\bar{p}) = K$ when $D(c) > K$ and is equal to c when $K \geq D(c)$.

We consider simultaneous-move Bertrand–Edgeworth games in which every firm i sets price p_i . The firm is committed to serve all customers up to its capacity constraint at price p_i . Buyers purchase from the firm with the lowest price first. If the demand for firm i exceeds capacity, buyers at that firm are rationed according to the efficient rationing rule.⁷ If several firms set the lowest price, total demand is allocated equally among the firms charging that price.

To develop the basic intuition for the results we first consider a very stylized model of collusion in which prices are set only once. A deviation from collusive price p^c leads to a reduction in the continuation value of the firm of an amount $F(n)$, which we call the “punishment”. In collusion models based on infinitely repeated noncooperative games, the punishment $F(n)$ is endogenously determined by the loss incurred when switching from a collusive continuation equilibrium to the lowest value continuation equilibrium. In this case we have $F(n) = [\delta/(1 - \delta)] [(1/n)\pi(p^c) - \underline{v}(n)]$, where $\underline{v}(n)$ is the average per period payoff in the lowest value continuation equilibrium and δ is the discount factor for future profits.⁸ To focus on the basic intuition we abstract at this point from the complication of the repeated game.

If capacity constraints are never binding (in the sense that a single firm can serve the whole market at the collusive price for all n), we can derive a standard *convergence result* to a competitive outcome. In this case a firm makes profits $(1/n)\pi(p^c)$ when setting p^c . It earns up to $\pi(p^c) - F(n)$ if it undercuts p^c slightly, captures the entire market, and incurs the punishment $F(n)$. Hence, collusion is sustainable if and only if

$$\frac{n-1}{n}\pi(p^c) \leq F(n). \quad (1)$$

The gains from deviating and winning over the market share of all other competitors, $[(n-1)/n]\pi(p^c)$, has to be smaller than the punishment $F(n)$ in order to sustain collusion. Note that $(n-1)/n$ is increasing in n . The more firms are in the market, the smaller the initial market share of a single firm and the greater the market share that is gained from a deviation. The smallest punishment that can prevent a deviation from a collusive price p^c , $\underline{F}(n)$, is therefore *increasing* in n . Conversely, for any given fixed punishment $F < \pi(p^c)$, collusion at p^c breaks down for a sufficiently fragmented market structure. Since $\pi(p^c)$ is increasing on $(0, p^m)$ (by log-concavity of $D(p)$), the most profitable sustainable collusive price under a fixed fine F strictly falls in n and converges to marginal cost as the number of firms becomes arbitrarily large.

This result of convergence to the competitive outcome is reversed when individual firms are capacity constrained relative to market demand at the collusive price. Instead

7. All results can be generalized to a large class of rationing rules that includes all the ones used in the literature.

8. We follow the convention in the repeated games literature to express all terms in “average values”. Dividing average values by $(1 - \delta)$ yields the present value of a continuation equilibrium.

of $\pi(p^c) - F(n)$, a deviating firm can make a profit of at most $(p^c - c)(K/n) - F(n)$. Hence, the incentive constraint becomes

$$\frac{1}{n}(p^c - c)K - \frac{1}{n}\pi(p^c) \leq F(n). \tag{2}$$

The incentive to deviate from collusion is now proportional to $1/n$. This means that the incentive to deviate from collusion falls as the market becomes more fragmented. The lowest fine at which collusion at p^m can be sustained, $\underline{F}(n)$, now *decreases* in n . In fact, $n\underline{F}(n)$ is constant. Conversely, for any fixed fine F that is smaller than half the monopoly profits, that is, when $[(p^m - c)K - \pi(p^m)]/2 > F > 0$, collusion is impossible for $n = 2$. But it becomes possible if the market becomes fragmented enough. Note that there is always some p^c strictly exceeding \bar{p} , for which the inequality in (2) is reversed for given F so that some collusion is always feasible. The highest collusive price that satisfies (2) is therefore increasing in n .⁹

Why is there such a stark difference between binding and nonbinding capacity constraints? In both models a firm that undercuts attracts the customers from all other firms. However, with capacity constraints the deviating firm can only serve a fixed number of consumers K/n . When n increases, total capacity K is distributed among more and more firms. Each firm becomes smaller relative to the market so that the firm can only serve a smaller fraction of the customers. A deviation thus becomes less attractive in face of a fixed punishment or even a punishment that declines less than proportionately in $1/n$. Intuitively, capacity constraints mean that a single firm becomes small not only when playing the collusive strategy but also when deviating from it.

Under binding capacity constraints, the behavior of $nF(n)$ determines whether increased fragmentation of capacity facilitates collusion ($nF(n)$ increases in n) or makes collusion more difficult ($nF(n)$ decreases in n). Under the specific assumptions of an infinitely repeated game this property is determined by the behavior of aggregate average profits in the punishment equilibrium, $n\underline{v}(n)$, since punishment is given by $nF(n) = [\delta/(1 - \delta)][\pi(p^c) - n\underline{v}(n)]$. Intuitively we would expect $n\underline{v}(n)$ to be weakly decreasing. To see this consider punishments generated by infinite reversion to a one-shot Nash equilibrium. In the one-shot Nash equilibrium per firm sales are (on average) $1/n$ of the total. But the average price cost margin converges to the competitive margin, so that $n\underline{v}(n)$ converges to the competitive profit at a rate of at least $1/n$. Essentially, this result follows from the standard results on convergence to the competitive outcome for static oligopoly games. Collusive equilibria sustained by optimal punishments are at least as severe as punishment generated from infinite reversion to a static Nash equilibrium. Aggregate optimal punishment profits therefore tend to inherit the property of convergence at a rate of at least $1/n$. We prove this result formally in the next section for the infinitely repeated Bertrand–Edgeworth model with efficient

9. To see this, consider the left-hand side of incentive condition (2). Differentiating in p^c yields $(1/n)[K - D(p^c) - (p^c - c)D'(p^c)]$, which is strictly positive on (c, p^m) , unless the firms are capacity constrained at p^c . Hence, for a given punishment F , the incentive constraint can be strictly relaxed by lowering the collusive price.

rationing. The basic effect from the fragmentation of capacity is a very robust feature of any collusion model with price setting firms.¹⁰

3. Fragmentation in the Standard Collusion Model

We now model collusion in the standard framework of an infinitely repeated game. In every period of the game firms play the Bertrand–Edgeworth game described in the previous section. We first show that the aggregate average value in the worst equilibrium is strictly decreasing in n . This result on the behavior of punishment profits then immediately leads to our main results:

1. There is always a collusive equilibrium with production strictly below the competitive level—even in the limit of a completely fragmented industry.
2. In Bertrand–Edgeworth industries in which all firms are capacity constrained relative to the monopoly output, the best sustainable collusive price (weakly) increases as total capacity is distributed among more firms.

3.1. The Structure of Optimal Punishments

In a Bertrand–Edgeworth game with efficient rationing it is easy to determine the lowest feasible equilibrium profit because it achieves the “security level payoffs”. This payoff can be derived by assuming that all rivals sell all of their capacity to the buyers with the highest valuation so that residual demand becomes

$$d(p, [(n - 1)/n]K) = \{0, D(p) - [(n - 1)/n]K\}.$$

This is the worst possible scenario for any firm, so that profits cannot be lower than

$$\pi^{\min} = \max_{p \in [\bar{p}, p^m]} (p - c) \left[\min \left\{ \max \left\{ 0, D(p) - \frac{n - 1}{n} K \right\}, \frac{K}{n} \right\} \right]. \quad (3)$$

As is well known, this is precisely the expected profit at the (mixed strategy) Nash equilibrium of the one-shot game. The lower bound on profits π^{\min} can therefore be attained as an average value of the infinitely repeated game by reverting to the one-shot Nash equilibrium of the static game forever. Hence, $n\underline{v}(n) = n\pi^{\min}$.

It is then straightforward to show that $n\underline{v}(n)$ is decreasing in n . Let $p^*(n)$ be the (unique) price that solves (3).¹¹ Then there is either a corner maximum at $p^*(n) = \bar{p}$

10. Indeed, this feature of punishment payoffs appears fairly robust beyond the framework of repeated games. For example, one form of a collusive contract captured by the stylized model in this section would be that each firm posts a bond that is forfeited after a deviation from the collusive price. If the firm can borrow up to expected future earnings, the maximal bond that could be posted would be constrained by future earnings under collusion: $(1/n)\pi(p^c)$, so that $nF(n) = \pi(p^c)$. If on the other hand, the firm could borrow against an asset with fixed value only, the fine would be independent of n . There are other realistic scenarios in which we would expect $nF(n)$ to be approximately constant. For example, if all that can be seized from a firm after a deviation are its assets, the liquidation value of a single firm should be expected to be approximately $1/n$ of the total liquidation value of all assets in the industry.

11. This is the highest price in the support of one-shot Nash equilibrium mixed strategy.

or $p^*(n) > \bar{p}$ achieves an interior maximum of $(p - c)(D(p) - [(n - 1)/n]K)$. Note that $p^*(n)$ is strictly decreasing in n in the latter case.¹² But we also have that $D(\lim_{n \rightarrow \infty} p^*(n)) - K \leq 0$, because there would be strictly positive excess demand in the limit otherwise. The latter would violate profit maximization for a firm that has arbitrarily small capacity. Hence, there exists \bar{n} , such that $n\pi^{\min}$ is equal to the competitive industry profit (i.e. $p^*(n) = \bar{p}$) if and only if $n \geq \bar{n}$.

If $D(c) < K$ the competitive industry profit is attained when $n - 1$ firms can serve the whole market, that is, \bar{n} is defined by $D(c) = [(\bar{n} - 1)/\bar{n}]K$. If $D(c) \geq K$, then \bar{n} is defined as the number of firms for which the Cournot best response to $[(\bar{n} - 1)/\bar{n}]K$ is exactly K/\bar{n} . In both cases we clearly have $\bar{n} < \infty$.¹³ If $n \geq \bar{n}$, then $n\nu(n)$ is constant at the competitive level. If $n < \bar{n}$, then $p^*(n) > \bar{p}$ and, by the envelope theorem, the static aggregate Nash equilibrium profits are strictly decreasing in n . These observations directly yield Lemma 1.

LEMMA 1. *Total industry profit in the lowest profit equilibrium, $n\nu(n)$, declines in n . It strictly declines if $n < \bar{n}$. For $n \geq \bar{n}$, it is constant and equals the perfectly competitive industry profits.*

Proof. See the Appendix.

Lemma 1 is a standard convergence result for static games. As the capacity is fragmented, behavior converges to the competitive outcome, which requires that industry profits fall in n as long as firms can exercise some market power. For collusive models this means that the punishments for deviating from collusion can be made more severe the more fragmented capacity is in the industry up to fragmentation level \bar{n} . This insight is central to the main message of the paper because in capacity-constrained Bertrand–Edgeworth oligopoly the punishment effect is the only effect at work when firms are sufficiently capacity constrained.

3.2. Capacity Fragmentation does not lead to Competitive Outcomes

We now derive the result that the highest sustainable collusive price does not converge to the competitive price as capacity becomes arbitrarily fragmented. This result is in stark contrast to the convergence result for the static game that we just derived. It is easiest to approach this analysis indirectly by characterizing two critical discount factors: First, the lowest discount factor at which a collusive price p^c can be sustained in equilibrium, $\bar{\delta}(p^c, n)$. Second, the lowest discount factor $\underline{\delta}(n)$, such that for every $\delta > \underline{\delta}(n)$ an average profit strictly exceeding the one-shot Nash equilibrium profit can be sustained. Based on the analysis of these critical discount factors, the nonconvergence result follows directly.

12. This follows from the fact the the cross-derivative of $(p - c)(D(p) - (n - 1)/nK)$ with respect to n and p is given by $-(1/n^2)K$.

13. This feature does not carry over to general rationing rules. With proportional rationing one can show that $\bar{n} = \infty$.

We first characterize $\bar{\delta}(p^c, n)$. We can write the incentive constraint for any collusive price p^c as

$$(p^c - c)D(p^c) - (1 - \delta)(p^c - c) \min\{nD(p^c), K\} \geq \delta n\underline{v}(n), \quad (4)$$

where, by definition, the inequality is binding for $\bar{\delta}(p^c, n)$. The comparative statics of $\bar{\delta}(p^c, n)$ depend on whether a firm is capacity constrained when it deviates from p^c or not. Let $\hat{n}(p^c) = K/D(p^c)$ be the critical number of firms such that a single firm has just enough capacity to serve all demand at the price p^c . For $n > \hat{n}(p^c)$ a firm is capacity constrained if it deviates to any price $p < p^c$. Solving (4) for $\bar{\delta}(p^c, n)$ yields

$$\bar{\delta}(p^c, n) = \begin{cases} \frac{n-1}{n} \left[1 - \frac{\underline{v}(n)}{\pi(p^c)} \right]^{-1} & \text{if } n \leq \hat{n}(p^c), \\ \frac{\hat{n}(p^c) - 1}{\hat{n}(p^c)} \left[1 - \frac{n\underline{v}(n)}{(p^c - c)K} \right]^{-1} & \text{if } n > \hat{n}(p^c), \end{cases} \quad (5)$$

where $(\hat{n}(p^c) - 1)/\hat{n}(p^c) = 1 - D(p^c)/K$.

When a deviating firm is capacity constrained, that is, $n > \hat{n}(p^c)$, the number of firms affects the incentive condition only through $n\underline{v}(n)$, which is decreasing in n by Lemma 1. Hence, the incentive constraint is relaxed and $\bar{\delta}(p^c, n)$ falls for all $n > \hat{n}$. Whenever a single firm cannot serve all the demand at price p^c , collusion at p^c becomes easier when a given level of capacity is distributed among more firms.¹⁴

When $n < \hat{n}(p^c)$, a firm deviating from the collusive price is not capacity constrained. It can serve all the demand at a price arbitrarily close to p^c . If punishment profits are zero, the critical discount factor is the same as in a homogeneous goods model without capacity constraints: $(n - 1)/n$. Otherwise, $\bar{\delta}(p^c, n) > (n - 1)/n$ because punishments are less severe. Since $(n - 1)/n$ is increasing in n and $\underline{v}(n)$ is decreasing, there are countervailing effects on $\bar{\delta}(p^c, n)$ in this case. However, we show in the proof of Proposition 1 that the net effect is for collusion at price p^c to get more difficult when $n < \hat{n}(p^c)$. Hence, $\bar{\delta}(p^c, n)$ is increasing on $[2, \hat{n}(p^c))$ and decreasing on $[\hat{n}(p^c), \infty)$ as shown in Figure 1.¹⁵

In a homogeneous goods model without capacity constraints the critical discount factor $\bar{\delta}(n) \equiv \delta(p^m, n)$ fully characterizes the scope for collusion in the model. In such a setting, a collusive price $p^c < p^m$ can never be sustained if p^m cannot be sustained (i.e. $\bar{\delta}(p^c, n) = \bar{\delta}(n)$ for all $p^c > c$).¹⁶ Our model behaves in the same way as a model without capacity constraints when a single firm can serve the whole market at marginal cost. This is the case if $n \leq K/D(c) \equiv n_c < \hat{n}$. The one-shot equilibrium therefore

14. The critical discount factor $\bar{\delta}(p^c, n)$ is strictly decreasing on (\hat{n}, \bar{n}) when $\bar{n} > \hat{n}$.

15. Note that this is the opposite pattern to Brock and Scheinkman (1985). In their case the critical discount factor initially falls and then rises in the number of firms.

16. Note that this is different from our simple static collusive model of Section 2. The reason is that punishments depend on the best collusive price p^c in the standard supgame model of collusion.

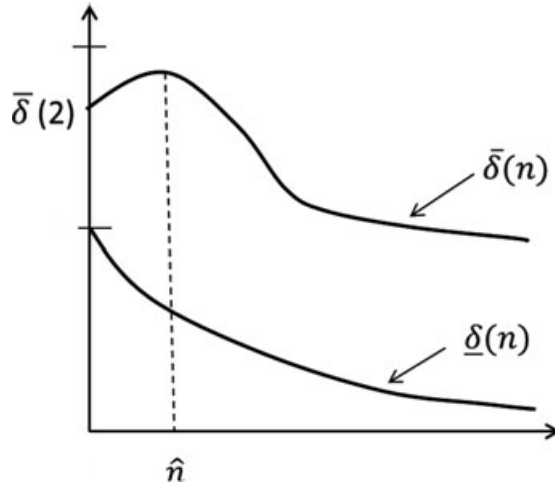


FIGURE 1.

involves marginal cost pricing and $n\underline{v}(n) = 0$ and the incentive constraint for collusion at p^c is given by

$$\pi(p^c)[1 - (1 - \delta)n] \geq 0. \tag{6}$$

Collusion at $p^c < p^m$ can therefore only be sustained if it can also be sustained at p^m . Hence, for $n \leq n^c$, $\underline{\delta}(n) = \bar{\delta}(n) = (n - 1/n)$ as in the case without capacity constraints.

However, this feature does not carry over to when capacity constraints do not allow a single firm to serve the market at marginal cost, that is $n > n^c$. Then there is some price $p^c > \bar{p}$ that can be sustained at discount factors for which full collusion at p^m is impossible. To characterize the set of discount factors for which *some* collusion is possible, let $\underline{\delta}(n)$ denote the lowest discount factor such that, for every $\delta > \underline{\delta}(n)$, the highest achievable average equilibrium profit strictly exceeds $\pi^N(n)$. If $n < \hat{n} \equiv \hat{n}(p^m)$ each firm can still serve the whole demand at the monopoly price p^m . For prices p^c just below p^m the firm would still not be capacity constrained. Therefore it is harder to sustain collusion at that price. However, when firms do not have enough capacity so that price is equal to marginal cost in the one-shot equilibrium (i.e. $n > n^c$), the price can always be lowered far enough so that an undercutting firm becomes capacity constrained at some price p strictly above $p^*(n)$. In the proof of Proposition 1 we show that there exist prices strictly exceeding $p^*(n)$ that can be sustained in a collusive equilibrium if and only if $\delta > \max\{1/n, 1/\bar{n}\}$. It follows that

$$\underline{\delta}(n) = \begin{cases} \frac{n - 1}{n} & \text{if } n \leq n_c, \\ \max\left\{\frac{1}{n}, \frac{1}{\bar{n}}\right\} & \text{if } n > n_c. \end{cases} \tag{7}$$

Note that equations (5) and (7) quite dramatically contradict the common wisdom that market fragmentation facilitates collusion. If a duopolist cannot serve all of the demand at the monopoly price, namely $n_c < \hat{n} \leq 2$, then the scope for collusion monotonically increases as capacity K is progressively distributed among more and more firms. This is true both in the sense that the critical discount factor for sustaining the monopoly price falls and in the sense that the lowest discount factor at which collusion at some level above the Nash equilibrium profit can be sustained falls. Since $\hat{n} \leq 2$ implies that $\bar{n} > \hat{n}$, the decrease in the two critical discount factors must be strict up to \bar{n} . It follows immediately that there is no convergence of equilibria to the competitive outcome in contrast to static oligopoly games.

PROPOSITION 1. *Suppose that $\delta > 1/\bar{n}$, then there exists a set of collusive equilibria with prices strictly exceeding the competitive price for all $n \geq \bar{n}$.*

Proof. See the Appendix.

Consider the market at the limit (i.e. $n > \bar{n}$) and with $\delta > 1/\bar{n}$. If there is excess capacity at the competitive price, namely $K > D(c)$, the highest sustainable price for $n > \bar{n}$, $p^c(\bar{n})$, can be inferred directly from equation (5) because the discount factor δ has to be equal to the critical discount factor for that price, that is, $\delta = 1 - D(p^c(\bar{n}))/K$. Hence, $p^c(\bar{n})$ is given by $D^{-1}((1 - \delta)K)$. Note that collusion gets harder when K is increased, because punishments are unaffected by K but any deviation captures a larger quantity as K is increased.

For $K < D(c)$ it is harder to determine the optimal collusive price. An increase in K does not only affect the incentives to deviate but also makes punishments more severe because competitive profits fall. It can be shown that the latter effect dominates so that collusion is easiest at $K = D(c)$. In fact, *some* collusion is sustainable for *any* discount factor in this case as long as n is large enough. Proposition 1 therefore implies that convergence to competitive pricing is obtained in the infinitely repeated game *only* when there are no capacity constraints at all. This is the strongest possible violation of the convergence result for the static Bertrand–Edgeworth game.

But the failure of the standard convergence result is even more dramatic than Proposition 1 suggests. Suppose, for example, that there is enough capacity in the market to serve all demand at marginal cost, namely $K > D(c)$. Assume that, at the same time, half of total capacity is not enough to serve all demand at the monopoly price, so that $D(p^m) > K/2$. Then collusion at the monopoly price is easier in a market with $n \geq \bar{n}$ firms than in a duopoly without a capacity constraint since $\bar{\delta}(\bar{n}) = 1 - D(p^m)/K \leq 1/2$. This means that full collusion is easier to sustain in a highly fragmented market than in a duopoly without capacity constraints. The reason for this is very simple. Very fragmented markets allow for very severe punishments, while capacity constraints limit the incentives for deviation from a collusive price.

3.3. The Highest Sustainable Price Increases with Capacity Fragmentation

Our characterization of the two critical discount factors also makes it easy to derive the change in the highest sustainable collusive price as the market becomes more fragmented for a given discount factor δ . Consider first the case $\bar{n} > \hat{n}$.¹⁷ Then the capacity constraint becomes binding for a firm deviating from the monopoly price for smaller n than the number of firms necessary to generate competitive behavior. For this case the result can be directly read off from Figure 1. Suppose first that $\delta \geq \bar{\delta}(\hat{n})$. Then the monopoly price is sustainable for any n . For $\delta \in [\bar{\delta}(2), \bar{\delta}(\hat{n})]$, the monopoly price can be sustained up to some n_L such that $\delta = \bar{\delta}(n_L)$. Since at this point $\bar{\delta}(n_L) > \underline{\delta}(n_L)$, this means that there will also be some smaller profit $\pi(p^c) > \pi^N$ sustainable at δ . However, since a deviating firm is not capacity constrained at p^m , there will exist $\varepsilon > 0$, such that $\pi(p^m - \varepsilon)$ is not sustainable at δ for a slightly more fragmented market. Hence, at n_L firms the highest sustainable collusive price drops to some price strictly below p^m when n is increased. After that the price must strictly increase until it reaches $p^c(\bar{n}) \leq p^m$. To see that, note that at $p^c(n) < p^m$ the capacity constraint must be strictly binding when a firm deviates from the collusive price. Hence, the incentive constraint only depends on $n \underline{v}(n)$, which is strictly decreasing in n , since $n < \hat{n} < \bar{n}$. These comparative statics are illustrated in Figure 2.

If $\delta \in (1/2, \bar{\delta}(2))$, some collusion can be sustained at some price $p^c(n) < p^m$ since $\delta > 1/n$ for all $n \geq 2$. Consequently, $p^c(n)$ increases monotonically on $[2, \infty)$, with the increase strict until the price reaches $p^c(\bar{n}) \leq p^m$. However, when $\delta \in (1/\bar{n}, 1/2)$, there will generally be a range of n such that no collusion is sustainable because $\delta < 1/n$. Then there only exists the mixed strategy equilibrium of the one-shot game. The highest price in the support of the mixed-strategy equilibrium will be $p^*(n)$. For some n_L we will have $\delta = 1/n_L$. At this point $\lim_{n \downarrow 1/\delta} p^c(n) = p^*(n_L)$ and then the highest sustainable collusive price strictly increases until $p^c(\bar{n}) \leq p^m$ is reached. With the exception of the range $(\bar{\delta}(2), \bar{\delta}(\hat{n}))$, price therefore always weakly increases. Under the condition $K/2 < D(p^m)$, $p^c(n)$ is monotonically increasing for all δ . Under the realistic condition that in duopoly neither firm can serve all of the demand at the monopoly price, our model therefore generates a highest sustainable collusive price that always increases as capacity becomes more fragmented.

The main difference between the case $\bar{n} > \hat{n}$ and the case $\bar{n} < \hat{n}$ is that in the latter case there exists a region of discount factors $\delta \in (\bar{\delta}(\bar{n}), \bar{\delta}(\hat{n}))$ for which the price

17. This should be seen as the much more likely case. For $D(c) > K$, we always have $\bar{n} > \hat{n}$. Consider the case $K \geq D(c)$. Since

$$\hat{n} = \frac{K}{D(p^m)} \quad \text{and} \quad \bar{n} = \frac{1}{1 - D(c)/K}$$

(the second is obtained by rearranging $D(c) = ((\bar{n} - 1)/\bar{n})K$), we have that $\hat{n} < \bar{n}$ implies

$$\frac{K}{D(p^m)} > \frac{1}{1 - D(c)/K}.$$

Hence, $D(p^m) + D(c) > K$. The case $\hat{n} > \bar{n}$ therefore requires fairly large excess capacities in the market.

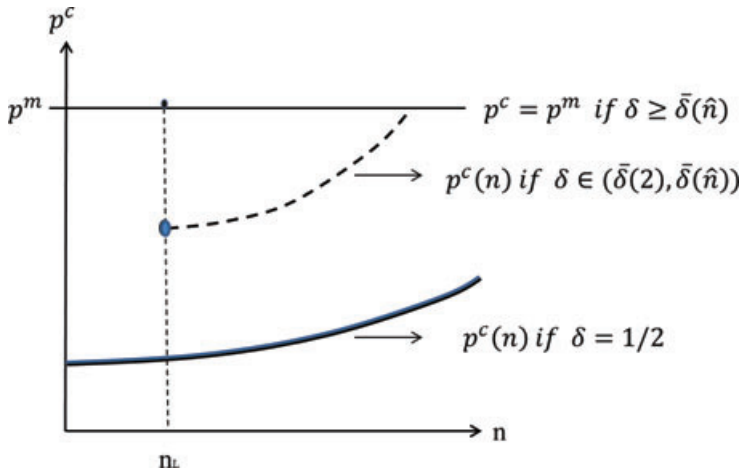


FIGURE 2.

$p^c(n)$ never strictly increases in n . The reason is that a deviating firm does not become capacity constrained before punishment profits reach the competitive level when n is increased. Except for this range of discount factors, the pattern of the comparative static in n is exactly the same in the two cases: For a low enough number of firms either the monopoly price can be sustained or no collusion is sustainable. If the monopoly price can be sustained for low n , $p^c(n)$ will discretely drop at n_L . If the monopoly price cannot be sustained initially, then some collusion can be sustained for n larger than some critical value $n_L \geq 2$ as long as $\delta > 1/\bar{n}$. For $n > n_L$ the price $p^c(n)$ monotonically increases in n until it becomes constant at $n = n_H$. To simplify notation in the following result, let $n^+ = \min\{\hat{n}, \bar{n}\}$. We can then summarize these observations formally in Proposition 2.

PROPOSITION 2. For any $\delta \in (0, 1)$, there exists $n_L(\delta) \in [2, \max\{2, n^+\}]$ and $n_H(\delta) \in [\max\{2, n^+\}, \bar{n}]$, such that

- (a) $p^c(n) = p^m$ or no collusion is sustainable on $[2, n_L(\delta))$,
- (b) $p^c(n)$ is strictly increasing on $(n_L(\delta), n_H(\delta))$,
- (c) $p^c(n)$ is constant for $n \geq n_H(\delta)$.

There is a downward jump in $p^c(n)$ at $n_L(\delta)$ if and only if $p^c(n_L(\delta)) = p^m$. Furthermore, if $n_H(\delta) < \bar{n}$, then $p^c(n_H(\delta)) = p^m$, otherwise $p^c(n_H(\delta)) < p^m$.

Proof. See the Appendix.

4. Robustness of the Results

The results derived in Section 3 are highly robust to variations on the basic infinitely repeated Bertrand–Edgeworth game. For example, Lemma 1 and Propositions 1 and 2

can be extended to the class of rationing schemes discussed by Lambson (1987). Essentially, all that has to be shown is that for general rationing schemes $n\underline{v}(n)$ is nonincreasing in order to prove that these results carry over. Since the security level profits are attained in the lowest-value equilibrium, this result is intuitively straightforward. Introduction of demand uncertainty (and state contingent pricing) preserves the result of nonconvergence to the competitive outcome because $n\underline{v}(n)$ continues to converge to the competitive outcome as n gets large.¹⁸ Indeed, since essentially any one-shot oligopoly model with capacity constraints has the property that behavior converges to competitive behavior as capacity gets arbitrarily fragmented, this necessary property for the nonconvergence result will hold almost universally. What is not robust more generally is the property of the Bertrand–Edgeworth model that deviation incentives are proportional to $1/n$. This property arises because a firm undercutting a collusive price will face the whole demand at that price in a Bertrand–Edgeworth model. We now show that the nonconvergence result nevertheless remains robust for other oligopoly models, because deviation incentives remain approximately proportional to $1/n$ when the number of firms is large. However, this analysis also implies that it is harder to generate price increases when capacity is fragmented among more firms.

What is different with oligopoly models that are not homogeneous goods price-setting models is that a deviation does not take all of the demand away from rivals as in Bertrand–Edgeworth competition. To see this point, consider the incentive condition for a general infinitely repeated homogeneous goods oligopoly model. To sustain a collusive price p^c with total production $D(p^c)$ at a symmetric collusive equilibrium, we need

$$(p^c - c) D(p^c) \geq (1 - \delta) \max_p \left[(p - c) \min \left\{ \max \left\{ 0, \left(D(p) - \frac{n - 1}{n} \sigma^c(p) \right) \right\}, K \right\} \right] + \delta n \underline{v}(n), \tag{8}$$

where the symmetric collusive strategy of all other firms is given by $(1/n)\sigma^c(p)$, which represents the firm-specific collusive supply function. It has the property $\sigma^c(p^c) = D(p^c)$. Cournot competition is a special case where $\sigma^c(p) = \sigma^c(p^c)$ for all p . Bertrand–Edgeworth competition is a special case where $\sigma^c(p) = 0$ for all $p < p^c$ and $\sigma^c(p) = K$ for $p > p^c$. The general model of supply function competition has $\sigma^c(p)$ as an increasing function between these extremes.¹⁹ The difference between the Bertrand–Edgeworth model and other models like the Cournot model is that a firm that undercuts p^c at price p does not gain $D(p)$ but only $D(p) - [(n - 1)/n]\sigma^c(p)$. A small price cut will only generate a small increase in sales. For any n , the price, $p^D(n, p^c)$, that maximizes $(p - c)(D(p) - [(n - 1)/n]\sigma^c(p))$ is therefore strictly smaller than p^c .

18. Formal proofs for these claims can be obtained from the author upon request.

19. When the firm takes the actions of others as given it does not matter for the analysis whether the firm chooses price or quantity along the residual demand function. We find it convenient for the comparison of models to write the problem as one of maximization over prices.

Note that this effect is smaller, the less elastic is $\sigma^c(p)$. The Cournot model is therefore a boundary case where price has to be reduced most in order to increase market share so that it is most difficult to generate our results. To simplify the discussion we will therefore restrict the discussion to the Cournot model, although it should be clear that it applies more generally.

The fact that strict price cuts are needed to win over market share has two consequences. First, it is harder to make capacity constraints binding when a firm deviates. It remains true that the deviating firm is capacity constrained when capacity becomes sufficiently fragmented because then only a small price cut is needed to fill capacity. However, the degree of fragmentation must be higher for capacity constraints to matter. Second, when capacity constraints are binding (i.e. $n > \hat{n}(p^c)$), the gains from deviation are no longer proportional to $1/n$ but fall more slowly. This implies that there will be a trade-off between the change in the incentives to deviate and the increase in the available punishments. To see this point note that when a deviating firm is capacity constrained, the optimal deviation cuts the price only to the point where residual demand is just equal to its capacity. In the Cournot case, this implies

$$D(p^D(p^c, n)) = \frac{K}{n} + \frac{(n-1)}{n} D(p^c). \quad (9)$$

It follows that $p^D(p^c, n)$ is increasing in n for $n > \hat{n}(p^c)$ and $\lim_{n \rightarrow \infty} p^D(p^c, n) = p^c$. The deviation benefit when the capacity constraint is binding thus becomes

$$(1 - \delta)(p^D(n, p^c) - c) \frac{K}{n} - (p^c - c) \frac{D(p^c)}{n}. \quad (10)$$

Hence, after pre-multiplying equation (10) by n as in the Bertrand–Edgeworth model, this expression is strictly increasing in n . We thus get a trade-off between the deviation and punishment effects of fragmenting capacity:

$$(1 - \delta) \frac{\partial p^D(p^c, n)}{\partial n} K + \delta \frac{\partial}{\partial n} (n \underline{v}(n)). \quad (11)$$

Note that for $K < D(c)$ the firms will be capacity constrained even in the one-shot Cournot equilibrium for some $n < \infty$. Hence, punishment profits become equal to the competitive profits for some $n_H < \infty$ and the second term becomes zero. It follows that the first term must eventually dominate and the best achievable collusive price must eventually decrease. Conversely, if $K \geq D(c)$, capacity does not constrain firms at the one shot Cournot equilibrium. If $n \underline{v}(n)$ is generated from reversion to one-period Cournot behavior forever, then both terms are of order $1/n^2$. It can then be shown that the second term dominates the first for high enough n . It is therefore clear that in an environment with Cournot competition it is not a robust result that capacity fragmentation leads to higher collusive prices. However, the price-increasing effect of capacity fragmentation can still occur for a wide range of reasonable capacity configurations. Furthermore, the scope for increasing collusive prices due to fragmentation will be greater in models like the supply function model, where the effect of n on deviation profits is smaller.

Despite this ambiguous result on the effect of fragmentation on prices, the more fundamental property of nonconvergence of collusive prices to the competitive benchmark is preserved in the Cournot model. To see this, note that the critical discount factor at which price p^c can just be sustained for $n > \hat{n}(p^c)$ is given by

$$\bar{\delta}(p^c, n) = \left[1 - \frac{(p^c - c) D(p^c)}{(p^D(p^c, n) - c) K} \right] \left[1 - \frac{n\underline{v}(n)}{(p^D(p^c, n) - c) K} \right]^{-1}. \quad (12)$$

Taking the limit in n immediately gives the nonconvergence result.

PROPOSITION 3. *For any*

$$\delta > \lim_{n \rightarrow \infty} \bar{\delta}(p^c, n) = \left[1 - \frac{D(p^c)}{K} \right] \left[1 - \frac{(\bar{p} - c)}{(p^c - c)} \right]^{-1},$$

there exists $n_H(\delta)$ such that p^c can be sustained as a collusive price for all $n > n_H(\delta)$.

Proof. See the Appendix.

Note that the critical discount factor at the limit of perfect capacity fragmentation, $\lim_{n \rightarrow \infty} \bar{\delta}(p^c, n)$, is the same as that of the Bertrand–Edgeworth model. In the limit the convergence result of static models fails and it fails quantitatively in exactly the same way as in the Bertrand–Edgeworth model. Nonconvergence to the competitive outcome is therefore a robust result for collusion models with capacity constraints and perfect monitoring.

The main caveat to this conclusion of general robustness of nonconvergence to the competitive outcome relates to our assumption of perfect monitoring. Imperfect monitoring clearly provides a potential channel for an important countervailing effect. When firms become very small even a maximal deviation to producing at full capacity will have a very small effect on any observed demand realizations for other firms. Hence, it may become difficult to detect unilateral deviations from collusion if capacity is very fragmented. However, when demand uncertainty is small this effect will only become important for very fragmented markets. Then the analysis in this paper will still be relevant for highly transparent markets like the former centralized electricity spot market in the UK and intermediate degrees of market concentration as we observe in such markets. Similarly, coordinated effects analysis in merger control will still be driven by the effect we have analyzed as long as the market is not extremely fragmented and market transparency is high. In contrast, when market transparency is low, collusion will not be a concern for either a regulatory or merger policy intervention. So the impact of fragmentation on collusion would be an irrelevant issue anyway.

5. Implications for Regulatory Policy and Empirical Work on Collusion

In this section we discuss the insights one can gain from the analysis in the paper for a number of regulatory policy issues as well as the empirical assessment of collusion. While each one of these issues could be addressed more rigorously, this section focuses

on the economic intuition derived from our previous analysis rather than deriving formal theoretical results.

Our initial analysis was motivated by the result of Sweeting (2007) that prices did not decline in the England and Wales electricity spot market after the existing capacity was very significantly fragmented among a larger set of firms. Proposition 2 shows that the empirical result of Sweeting (2007) that capacity fragmentation led to an increase in price cost margins in the electricity spot market in England and Wales is not as surprising as it might seem. The result is precisely what collusion theory predicts for Bertrand–Edgeworth markets when individual firms cannot serve all of the demand at the collusive price. The paper thus gives a theoretical explanation for the empirical results obtained by Sweeting (2007), although the analysis also reveals that a result of increasing prices is not necessarily robust under supply function competition (which appears a good approximation for the market institution at the time). However, the paper does provide some insights for policy towards industries like the electricity industry in which restructuring of capacities is an option in the process of regulatory reform. First, the paper makes clear that capacity fragmentation is not necessarily an effective remedy when collusion is a serious concern because significant collusion can occur at any level of capacity fragmentation. Nevertheless, the paper does not imply that such fragmentation should not be undertaken.²⁰ However, when there are serious concerns about the potential for collusion in an industry, changes in the market mechanism may be of greater importance to improve the performance of the industry than the fragmentation of assets. It should also be noted that this paper does not imply that entry into an industry would not be desirable if one wishes to limit collusion. What drives the results in this paper is that *given capacity* is redistributed among firms. Policies that lead to entry in the industry, increasing overall capacity, will typically mitigate the collusion problem.

Secondly, the paper contributes to our understanding of the recent literature on the coordinated effects of mergers. Compte, Jenny, and Rey (2002) have previously derived the result that a merger that concentrates capacity can make collusion more difficult in Bertrand–Edgeworth markets. They attributed this result to changes in the asymmetries between the firms. However, our results for the symmetric model suggests that it is the fragmentation of market share itself that facilitates collusion and makes mergers by the largest firm undermine collusion. This is in contrast to the coordinated effects of mergers model by Kühn (2004) where product lines of differentiated products are transferred between firms in a merger. In Kühn's model it is asymmetry in product lines alone that drives the results because fragmentation of product lines that preserves symmetry between firms makes collusion harder (see Kühn and Rimler 2006). The result that fragmentation of assets facilitates collusion is thus specific to models with capacity constraints. For practical purposes our result implies that for mergers in which the merging firms remain capacity constrained relative to the monopoly output there

20. In fact, in the Spanish electricity industry, in which there is considerable vertical integration between activities upstream and downstream from the spot market, collusion may be a lesser concern. Then market fragmentation can greatly enhance the performance of the industry.

is an even stronger case that no anticompetitive coordinated effects can arise from a merger of a large firm compared to a setting in which only an asymmetry argument applies (as in Kühn's paper).

Third, the analysis in our paper also suggests a reassessment of the empirical literature that has attempted to identify industry features that facilitate collusion from cross-sectional studies of cartels.²¹ It has been considered a puzzle in this literature that a number of cross-sectional studies (Posner 1970; Dick 1996) have found cartel duration to be increasing in the number of cartel members. Posner (1970) attributed this result in his paper to a selection bias of using data from DoJ enforcement actions. He argued that large cartels are easier to detect and that this led to a bias towards cartels with larger numbers of firms when studying enforcement data. However, Dick (1996) focuses on a data set of legal export cartels that does not suffer from this kind of selection bias. Nevertheless, he finds a positive relationship between the number of firms in the cartel and cartel duration. Most of the remaining literature has found little evidence for a systematic relationship between cartel duration and number of firms in the market (See Levenstein and Suslow 2006; Hay and Kelley 1974; Asch and Seneca 1975; Suslow 2005), while Fraas and Greer (1977) find that a larger number of firms reduces the likelihood of collusion.²²

Assuming that cartel duration is a good proxy for the ease of collusion in a market, our model explains why such puzzling results might have been obtained in the empirical literature. To see this it is useful to note that what is crucial for our result is that total capacity does not increase relative to the size of the market as the number of firms increases. This relationship between market size, capacities, and the number of firms will be preserved in a model with endogenous entry when every new firm brings the same capacity K to the market. To see how this property can lead to the observed positive correlation between the number of firms and collusion it is useful to slightly formalize the argument in a highly stylized model. Let demand be given by $SD(p)$, where S is a shift parameter that captures "market size". We assume for the sake of argument that cross-sectionally markets only differ in market size, but that the size per firm K , the function $D(p)$, the marginal cost c , and entry cost E are the same across markets. We consider a two-stage game. In the first stage, firms decide to enter the market or not. In a second stage, they play the best symmetric equilibrium of the infinite horizon collusion game of Section 3. To simplify the argument as much as possible, consider an initial situation in which any deviating firm from a collusive price would be capacity constrained, namely $SD(p^c) > K$, and let $\tau \equiv S/n$. When the incentive condition for the highest sustainable collusive price is binding at the best

21. Levenstein and Suslow (2006) provide a recent survey.

22. Levenstein and Suslow (2006) suggest that large cartels with more than ten members tend to survive only when there is active involvement of a trade association. While this is plausible given that the coordination problems for a large cartel are difficult to resolve, the presence of a trade association would not explain why so many very large cartels are successful. The results from a collusion model with homogeneous goods but without capacity constraints would make such outcomes appear highly unlikely.

achievable equilibrium it follows that

$$\begin{aligned} & \tau(p^c - c)D(p^c) \\ & = (1 - \delta)(p^c - c)K + \delta(p^* - c) \min \{0, \max \{n\tau D(p^*) - (n - 1)K, K\}\} \end{aligned} \quad (13)$$

and by free entry, it follows that

$$\tau(p^c - c)D(p^c) = E. \quad (14)$$

Suppose that the highest price sustainable in a market with size S is p^c . Now consider a market with $S' > S$. Suppose for the moment that we want to sustain the same collusive price p^c in the market S' . Then entry would guarantee that $\tau' = \tau$ and the number of firms in market S' is $n'(p^c) > n$. Note that the impact on the incentive constraint then only comes from the punishment profits. If $\tau D(p^*) = K$ in market S , then this will be the same in market S' with the number of firms raised to n' . Hence, firms are and remain capacity constrained in the punishment equilibrium and there is no effect from the increase in the number of firms on the incentive constraint. In this case the degree of collusion is the same in two markets with different numbers of firms. Now suppose that $\tau D(p^*) < K$, but that there are positive punishment profits. Then punishment profits are by $[n' - n][K - \tau D(p^*)]$ lower in market S' . This means that the incentive constraint becomes strictly slack and the best collusive price can be raised above p^c . This in turn will lead to more entry into the market anticipating that collusive prices are higher given the greater demand. Hence, in this demand region we would see a positive relationship between the collusive price (and therefore the ease of collusion) and the number of firms in the market.

The purpose of the model is to illustrate how endogenous entry combined with our result that fragmentation of assets relative to demand can lead to a positive correlation between the number of firms and the success of collusion observed in the data.²³ All that is significant for the argument is the basic intuition that the degree of entry is endogenously determined by the level of demand for given entry costs. At higher demand levels there is more entry and therefore capacity is more fragmented relative to the demand. Hence, as the market gets bigger, market fragmentation is endogenously increased, leading to more collusion in larger markets with more firms. As we have seen in the paper, this effect is driven by capacity constraints being binding on deviating firms. This property will be true for some markets but not for all. It is therefore not surprising in light of this theory that empirical work has not established an unambiguous relationship between the number of firms and various proxies for the ease of collusion in a market. Our analysis suggests that the number of firms is a bad measure for a cross-sectional study. What really matters is to what extent the firms are constrained by their individual capacities to serve the market and to what extent overall capacity is sufficient to drive down the price in a punishment equilibrium.

23. This game is not supposed to be a realistic model of collusion with entry as in Harrington (1989).

Note also that the intuition gained from Brock and Scheinkman (1985) would be misleading for a setting with endogenous entry. In their paper more “entry” implies more capacity relative to demand and (except for very small numbers of firms) a reduction in collusion when the number of firms in the market is greater. However, as we have just shown, this logic does not apply when entry is endogenous. Instead the comparative static result we focused on in this paper is the one that can rationalize the empirical pattern. We have thus shown that our result about the fragmentation of capacity is the relevant one over a wide range of applications.

Appendix: Proofs

A.1. Proof of Lemma 1

For $D(c) < K$, $\bar{p} = c$, and it is immediately verified that $p^*(n) = \bar{p} = c$ if and only if $n \geq \bar{n} = K/[K - D(c)]$. If $D(c) > K$, then $D(\bar{p}) = K$. Hence, any price below \bar{p} would lead to the same sales as \bar{p} and profit maximization requires $p^* \geq \bar{p} > c$. For this case define \bar{n} as the n for which $\bar{p} = \arg \max(p - c)(D(p) - [(n - 1)/n]K)$. Note that the cross-derivative of $(p - c)(D(p) - [(n - 1)/n]K)$ between p and n is strictly negative. Thus, $\arg \max(p - c)(D(p) - [(n - 1)/n]K)$ is strictly decreasing in n , is $p^m > \bar{p}$ for $n = 1$ and strictly smaller than \bar{p} in the limit for $n \rightarrow \infty$. It follows that \bar{n} is unique and, for any $n > \bar{n}$, $\arg \max(p - c)(D(p) - [(n - 1)/n]K) < \bar{p}$, so that the capacity constraint is binding. Hence industry profits are constant at $(\bar{p} - c)K$ for all $n \geq \bar{n}$. Now suppose $n < \bar{n}$ so that $p^*(n) > \bar{p}$. Then industry profits are given by $n(p^* - c)(D(p^*) - [(n - 1)/n]K)$. By the envelope theorem, the first derivative of industry profits with respect to n is given by $(p^* - c)(D(p^*) - K)$, which is strictly negative for $p^* > \bar{p} \geq c$. It follows that industry profits strictly decline on $[2, \bar{n})$ and are constant at the competitive profits for higher n .

A.2. Proof of Proposition 1

The proof of this proposition follows directly from the proof of the following lemma.

LEMMA A.1. *For any n there exist critical discount factors $\bar{\delta}(n) < 1$ and $\underline{\delta}(n) \leq \bar{\delta}(n)$, such that the monopoly price can be sustained for high enough discount factors (i.e. $\delta \in [\bar{\delta}(n), 1]$), a collusive price level below the monopoly price (i.e. $p^c \in (p^*(n), p^m)$) can be sustained for intermediate discount factors (i.e. $\delta \in (\underline{\delta}(n), \bar{\delta}(n))$), and no collusion is sustainable for low discount factors (i.e. $\delta < \underline{\delta}(n)$). Furthermore:*

- (a) $\bar{\delta}(n)$ is strictly increasing on $[2, \hat{n})$ and constant on $(\max\{\hat{n}, \bar{n}\}, \infty)$. If $\bar{n} > \hat{n}$, $\bar{\delta}(n)$ is strictly decreasing on (\hat{n}, \bar{n}) ;
- (b) $\underline{\delta}(n)$ is strictly increasing on $[2, n^c)$, strictly decreasing on $[n^c, \bar{n}]$, and constant at $n > \bar{n}$.

A.3. Proof of Lemma A.1.

Part (a). Clearly, for $p^c = p^m$, equation (4) is strictly satisfied at $\delta = 1$ and is not satisfied at $\delta = 0$. Since the incentive constraint is strictly relaxed by increasing δ there exists a unique $\bar{\delta}(n)$ at which equation (4) is just binding. It is given by

$$\bar{\delta}(n) = \begin{cases} \frac{n-1}{n} \left[1 - \frac{\underline{v}(n)}{\pi(p^m)} \right]^{-1} & \text{if } n \leq \hat{n}, \\ \frac{\hat{n}-1}{\hat{n}} \left[1 - \frac{n\underline{v}(n)}{(p^m - c)K} \right]^{-1} & \text{if } n > \hat{n}. \end{cases} \quad (\text{A.1})$$

If $n > \hat{n}$, $\bar{\delta}(n)$ is decreasing from Lemma 1 because $n\underline{v}(n)$ is decreasing. If $n > \max\{\hat{n}, \bar{n}\}$ it is constant. If $\bar{n} > \hat{n}$ it is strictly decreasing on (\hat{n}, \bar{n}) . Suppose that $n < \hat{n}$. If $n > \bar{n}$, then $p^*(n) = \bar{p}$, and $\bar{\delta}(n)$ is strictly increasing. If $n < \bar{n}$, then $p^*(n) > \bar{p}$ and

$$\begin{aligned} \frac{\partial}{\partial n} \Big|_{\delta=\bar{\delta}(n)} & \left[(p^m - c)D(p^m) - (1 - \delta)(p^m - c)nD(p^m) - \delta n\underline{v}(n) \right] \\ &= -(1 - \bar{\delta}(n))(p^m - c)D(p^m) - \bar{\delta}(n) [\underline{v}(n) + n\underline{v}'(n)] \\ &= -\frac{\pi(p^m)}{n} - \bar{\delta}(n)n\underline{v}'(n) \\ &= -\frac{\pi(p^m)}{n} \left[1 - \frac{n-1}{n} \frac{(p^* - c)K}{\pi(p^m) - \underline{v}(n)} \right] \\ &= -\frac{\pi(p^m)}{n [\pi(p^m) - n\underline{v}(n)]} \left[\pi(p^m) - \underline{v}(n) - \frac{n-1}{n} (p^* - c)K \right] \\ &= -\frac{\pi(p^m)}{n [\pi(p^m) - n\underline{v}(n)]} \\ & \quad \times \left[\pi(p^m) - (p^* - c) \left[D(p^*) - \frac{n-1}{n} K \right] - \frac{n-1}{n} (p^* - c)K \right] \\ &= -\frac{\pi(p^m)}{n [\pi(p^m) - n\underline{v}(n)]} [\pi(p^m) - \pi(p^*)] < 0, \end{aligned} \quad (\text{A.2})$$

where the second equality comes from substituting in from the incentive constraint equation (4) and the third equality comes from substituting in from equation (A.1) and factoring out $\pi(p^m)/n$. Hence, $\bar{\delta}(n)$ is strictly increasing on $(2, \hat{n})$. This proves part (a) of Lemma A.1.

Part (b). Note first that for discount factors below $\bar{\delta}(n)$ a firm deviating from a collusive price p^c must be strictly capacity constrained. The collusive price therefore must be in the interval $(\bar{p}, D^{-1}(K/n))$. From equation (4) it follows that the collusive price that maximally relaxes the incentive constraint for a colluding firm solves

$$\max_{p \in [\bar{p}, D^{-1}(K/n)]} (p - c) \max\{[D(p) - (1 - \delta)K], K/n\}. \quad (\text{A.3})$$

For any $\delta \in (0, 1)$,

$$p^*(\delta) = \arg \max_{p \in [\bar{p}, D^{-1}(K/n)]} (p - c) \max\{[D(p) - (1 - \delta)K], K/n\}$$

is increasing in δ with $\underline{\delta} < 1$ the largest δ for which $p^*(\delta) = \bar{p}$. Then $\bar{n} \equiv 1/\underline{\delta}$ is the smallest n such that, for all $n > \bar{n}$,

$$\bar{p} = \arg \max_{p \in [\bar{p}, D^{-1}(\frac{n-1}{n}K)]} (p - c)(D(p) - [(n - 1)/n]K).$$

By the envelope theorem, it follows that

$$\max_{p \in [\bar{p}, D^{-1}(K/n)]} (p - c) \max\{[D(p) - (1 - \delta)K], K/n\}$$

strictly exceeds

$$n\underline{v}(n) = \max_{p \geq \bar{p}} (p - c) \max\{[D(p) - [(n - 1)/n]K], 0\}$$

if and only if $\delta > \max\{1/n, 1/\bar{n}\}$. Since $p^*(\delta)$ is the price that maximally relaxes the incentive constraint, it follows that, for $\delta \leq \max\{1/n, 1/\bar{n}\}$, no collusive profits above $\pi^N(n)$ can be obtained. Furthermore, for any $\delta > \underline{\delta}$, $p^*(\delta)$ is strictly increasing. Hence, for any $\delta > \max\{1/n, 1/\bar{n}\}$,

$$p^c > p^*(\delta) = \arg \max (p - c)[D(p) - (1 - \delta)K] > p^*(n), \tag{A.4}$$

where the second inequality follows from the monotonicity of $p^*(\delta)$ in δ . The first inequality follows because the incentive constraint is strictly slack at $p^*(\delta)$ and the profit function is quasi-concave in p .

A.4. Proof of Proposition 2

The proof of Proposition 2 is in two parts. First we show the result for $\hat{n} < \bar{n}$. Then we show the result for $\hat{n} > \bar{n}$.

Part 1 ($\hat{n} < \bar{n}$). First note that $\bar{\delta}(n)$ increases in n up to \hat{n} and then strictly decreases in n up to \bar{n} after which it stays constant. For any $\delta \in (\bar{\delta}(2), \bar{\delta}(\hat{n}))$ there then must be exactly one $n < \hat{n}$ such that $\delta = \bar{\delta}(n)$. Call this $n_L(\delta)$. Similarly, there can at most be one $n > \hat{n}$ for which $\delta = \delta(n)$. If it exists call this $n_H(\delta) < \bar{n}$, if it does not exist set $n_H(\delta) = \bar{n}$. If $\delta \geq \bar{\delta}(\hat{n}) = \max_n \bar{\delta}(n)$, then the monopoly price can be sustained for all n . For $\delta \in (\bar{\delta}(2), \bar{\delta}(\hat{n}))$, let $n < n_L(\delta)$. Then $\delta > \bar{\delta}(n)$ and the monopoly price can be sustained. For $n \in (n_L(\delta), n_H(\delta))$, the monopoly price cannot be sustained and the capacity constraint must be binding for a deviating firm at the highest sustainable collusive price. Furthermore, $n < \bar{n}$, so that the incentive constraint is strictly slackened by increasing the number of firms. Hence, a slightly higher price p^c can be sustained. The collusive price is therefore strictly increasing on $(n_L(\delta), n_H(\delta))$. Since $p^c < p^m$ for all $n \in (n_L(\delta), n_H(\delta))$ and price is strictly increasing on that range it follows that

$\lim_{n \downarrow n_L(\delta)} p^c(n) < p^m$. By definition $p^c = p^m$ for all $n \geq n_H(\delta)$ if $n_H(\delta) < \bar{n}$. Otherwise, $p^c(n)$ becomes constant at a price strictly below p^m for $n > \bar{n}$. Now consider the range $\delta \in (1/\bar{n}, \bar{\delta}(2))$. For this region there is no $n \leq \hat{n}$ such that p^m can be sustained. Define $n_L(\delta)$ in this case as the smallest n such that $\delta \geq 1/n$ and define $n_H(\delta)$ as before. For $n < n_L(\delta)$ no collusion can be sustained by Lemma A1. For $n \in (n_L(\delta), n_H(\delta))$, $\delta < \bar{\delta}(n)$ and therefore a deviating firm must be capacity constrained at the best collusive price. Since $n < n_H(\delta) \leq \bar{n}$, punishment profits are strictly decreasing so the highest collusive price is strictly increasing on this range. Beyond $n_H(\delta)$ the price is constant and the same argument about the level of the price applies as previously discussed. Note that for $\delta \in (1/2, \bar{\delta}(2))$ the critical $n_L(\delta) = 2$ and collusion is possible for all $n \geq 2$.

Part 2 ($\hat{n} < \bar{n}$). For $\delta \geq \bar{\delta}(\hat{n})$ the same argument as in Part 1 applies. For $\delta \in (\bar{\delta}(2), \bar{\delta}(\hat{n}))$ we now have to distinguish between two cases because $\bar{n} < \hat{n}$. First consider $\delta \in (\bar{\delta}(\bar{n}), \bar{\delta}(\hat{n}))$. Since $\bar{\delta}(n)$ is strictly decreasing for $n < \hat{n}$, there exists a unique $n_L(\delta) \in (\bar{n}, \hat{n})$, such that $\delta > \bar{\delta}(n)$ for all $n < n_L(\delta)$ and $\delta < \bar{\delta}(n)$ for all $n > n_L(\delta)$. Hence, p^m can be sustained for all $n < n_L(\delta)$. Beyond that, the most collusive price has a deviating firm capacity constrained and, since $n > \bar{n}$, punishment profits are at the competitive profits. Hence, price must drop to some $p^c(n) < p^m$ at $n_L(\delta)$ and then stay constant. For $\delta \in (\bar{\delta}(2), \bar{\delta}(\bar{n}))$, monotonicity of $\bar{\delta}(n)$ on $(2, \hat{n})$ again implies that there is a unique $n_L(\delta)$ such that $\delta > \bar{\delta}(n)$ for $n < n_L(\delta)$ and the reverse for $n > n_L(\delta)$. Hence, the best collusive price is p^m up to $n_L(\delta)$. For larger n the price must be below p^m to make the capacity constraint binding for deviators. Since $n < \bar{n}$ for n just above $n_H(\delta)$ the incentive constraint is relaxed by increasing n , since punishment profits fall in n . Hence, $\lim_{n \downarrow n_L(\delta)} p^c(n) < p^m$. Since a firm deviating from the best collusive price must be capacity constrained and the competitive profits are reached at \bar{n} , the highest collusive price must be increasing on $(n_L(\delta), \bar{n})$ and is constant at $p^c(\bar{n}) \leq p^m$ on (\bar{n}, ∞) . For the range $\delta \in (1/\bar{n}, \bar{\delta}(2))$, define $n_L(\delta) = 1/\delta$. By the same argument as before no collusion can be sustained for $n < n_L(\delta)$. By analogous argument to the proof of Proposition 2 the highest collusive price is strictly increasing on $(n_L(\delta), \bar{n})$, and constant afterwards at a price strictly less than the monopoly price.

A.5. Proof of Proposition 3

We first derive the limiting discount factor at which price p^c can just be sustained in the limit of infinite fragmentation of capacity. Then for any discount factor strictly exceeding the limiting discount factor there must exist, by continuity of the discount factor in n , some $n(\delta) < \infty$ such that price p^c can still be sustained. Note that

$$\begin{aligned} \lim_{n \rightarrow \infty} \bar{\delta}(p^c, n) &= \lim_{n \rightarrow \infty} \left[1 - \frac{(p^c - c) D(p^c)}{(p^D(p^c, n) - c) K} \right] \left[1 - \frac{n \underline{v}(n)}{(p^D(p^c, n) - c) K} \right]^{-1} \\ &= \left[1 - \frac{D(p^c)}{K} \right] \left[1 - \frac{(\bar{p} - c)}{(p^c - c)} \right]^{-1}, \end{aligned} \quad (\text{A.5})$$

where the second equality follows because

$$\lim_{n \rightarrow \infty} p^D(p^c, n) = p^c \quad \text{and} \quad \lim_{n \rightarrow \infty} n\underline{v}(n) = (\bar{p} - c)K.$$

The proposition follows.

References

- Amir, Rabah and Val E. Lambson (2000). "On the Effects of Entry in Cournot Markets." *Review of Economic Studies*, 67, 235–254.
- Arocena, Pablo, Kai-Uwe Kühn, and Pierre Regibeau (1999). "Regulatory Reform in the Spanish Electricity Industry: a Missed Opportunity for Competition." *Energy Policy* 22, 387–399.
- Asch, Peter and Joseph Seneca (1975). "Characteristics of Collusive Firms." *Journal of Industrial Economics*, 23, 223–237.
- Brock, William A. and José Scheinkman (1985). "Price Setting Supergames with Capacity Constraints." *Review of Economic Studies*, 52, 371–382.
- Compte, Olivier, Frederic Jenny and Patrick Rey (2002). "Capacity Constraints, Mergers, and Collusion." *European Economic Review*, 46, 1–29.
- Davidson, Carl and Raymond J. Deneckere (1990). "Excess Capacity and Collusion." *International Economic Review*, 31, 521–541.
- Dick, Andrew (1996). "When are Cartels Stable Outcomes?" *Journal of Law and Economics*, 39, 241–283.
- Fraas, Arthur G. and Douglas F. Greer (1977). "Market Structure and Price Collusion: an Empirical Analysis." *Journal of Industrial Economics*, 26, 21–44.
- Green, Edward J. (1980). "Noncooperative Price Taking in Large Dynamic Markets." *Journal of Economic Theory*, 22, 155–182.
- Green, Richard J. (1996). "Increasing Competition in the British Electricity Spot Market." *Journal of Industrial Economics*, 44, 205–216.
- Green, Richard J. and David M. Newbery (1992). "Competition in the British Electricity Spot Market." *Journal of Political Economy*, 100, 929–953.
- Harrington, Joseph E. (1989). "Collusion and Predation under (almost) Free Entry." *International Journal of Industrial Organization*, 7, 381–401.
- Hay, George A. and Daniel Kelley (1974). "An Empirical Survey of Price-Fixing Conspiracies." *Journal of Law and Economics*, 17, 13–38.
- Henney, Alex (1987). "Privatize Power: Restructuring the Electricity Supply Industry." Policy Study 83, Center Policy Studies, London.
- Kühn, Kai-Uwe (1997). "Nonlinear Pricing in Vertically Related Duopolies." *Rand Journal of Economics*, 28, 37–62.
- Kühn, Kai-Uwe (2004). "The Coordinated Effects of Mergers in Differentiated Goods Markets." CEPR Discussion Paper 4769, London.
- Kühn, Kai-Uwe and Pierre Regibeau (1998). "¿Ha Llegado la Competencia? Un Análisis Económico de la Reforma de la Regulación del Sector Eléctrico en España." *Informes del IAE*, Instituto de Análisis Económico, Barcelona.
- Kühn, Kai-Uwe and Michael S. Rimpler (2006). "The Comparative Statics of Collusion Models." CEPR Discussion Paper 5742, London.
- Lambson, Val E. (1987). "Optimal Penal Codes in Price Setting Supergames with Capacity Constraints." *Review of Economic Studies*, 54, 385–397.
- Margaret, Levenstein and Valerie Suslow (2006). "What Determines Cartel Success." *Journal of Economic Literature*, 44(1), 43–95.
- Motta, Massimo (2004). *Competition Policy: Theory and Practice*. Cambridge University Press.
- Novshek, William and Hugo Sonnenschein (1978). "Cournot and Walras Equilibrium." *Journal of Economic Theory*, 19, 223–266.

- Posner, Richard (1970). "A Statistical Study of Antitrust Enforcement." *Journal of Law and Economics*, 13, 365–419.
- Rotemberg, Julio J. and Garth Saloner (1986). "A Supergame-Theoretic Model of Price Wars During Booms." *American Economic Review*, 76, 390–407.
- Scherer, Frederic M. and David R. Ross (1990). *Industrial Market Structure and Economic Performance*, 3rd ed. Houghton Mifflin, Boston.
- Schmalensee, Richard and Bennett W. Golub (1984). "Estimating Effective Concentration in Deregulated Wholesale Electricity Markets." *RAND Journal of Economics*, 15, 12–26.
- Suslow, Valerie Y. (2005), "Cartel Contract Duration: Empirical Evidence from Inter-war International Cartels." *Industrial and Corporate Change*, 14, 705–744.
- Sweeting, Andrew (2007). "Market Power in the England and Wales Wholesale Electricity Market 1995–2000." *Economic Journal*, 117, 654–685.
- Sykes, Allen and Colin Robinson (1987). "Current Choices: Good Ways and Bad to Privatise Electricity." Policy Study 87, Center for Policy Studies, London.