Government-Enforced Cartels, Output and Welfare *

Julio J. Rotemberg

June 14, 1998

Abstract

Cartels tend to raise prices to a greater extent when demand is low than when it is high. With random demand, the expectation of higher prices in low demand states leads cartels to have higher capacity than the corresponding perfectly competitive firms if, in line with actual historical experience, cartels are unable to restrict entry. As a result, sales can actually increase when demand is high. This may raise the expected utility of consumers if there are costs of financial distress because, in this case, perfect competition induces an inefficiently low level entry. In effect, cartels stabilize the cash flow of firms and this can have a socially favorable effect on entry if financial markets are incomplete. In addition, the extra stability of cash flow brought about by cartels can reduce the incidence of inefficient bankruptcies when demand is low. The work of Liefmann (1932) suggests that these considerations were important in the German experience with cartels.

*Harvard Business School. I wish to thank Ken Judd, David Moss, Thomas McCraw, Robert Pindyck, David Sharfstein, Deborah Spar and seminar participants at MIT and Stanford for helpful comments and suggestions. I am also grateful to the Harvard Business School Division of Research for research support.
Collusion among competing firms is not always illegal. On the contrary, contracts among firms whose aim is to curtail industry output have been enforced for a long time by many governments on the European continent, including those of Germany, Sweden and Switzerland. Even in the United States, cartel contracts of this sort are countenanced in limited circumstances. As economists, our basic theoretical presumption is that the enforcement of these contracts is a bad idea because it facilitates increases in prices and reductions in output below the efficient competitive level. This theoretical indictment of cartels seems bolstered by the empirical work on Swedish cartels of Fölster and Peltzman (1997) which suggests that, indeed, these cartels have raised average prices while also lowering average productivity.

It turns out, however, that whether cartels are socially desirable or not is a somewhat subtle empirical question. The reason is that cartels also yield a potential social benefit: they can reduce the risk faced by firms. In the presence of incomplete financial markets, it is theoretically possible for this benefit to be so large that consumers and firms are both better off when cartel contracts are enforceable.

Further empirical work is obviously needed before one can know the extent to which these theoretical benefits have been provided by actual cartels. What is notable, however, is that some contemporary academic observers of German cartels regarded these benefits as important. Liefmann (1932, p. 56), in particular, says “... the diminution of capital risks which the cartels render possible through an improved adjustment of production to the fluctuations of demand is of great advantage.” This reduction in risk results from changing output more when demand fluctuates so that price fluctuates less. As he goes on to say on p. 146, “With the help of cartels, cyclical movements of prices have to some extent been rendered less violent...” He particularly emphasizes the cartels’ ability to raise prices when demand is low. As he says (p. 104) “... we may draw the general conclusion that the main effect on price, which differentiate cartel operation from open competition, appear in times

\[1\] In addition, United States agricultural policy takes the form of inducing certain producers to act as if they had such a contract (see Rhodes (1978 p. 324-332 and Cave and Salant 1995).
of decreasing trade. For it is at such times that cartels prevent a fall in price to correspond to a fall in demand” (italics in the original).

My model of cartel behavior captures these effects by supposing that firms must acquire capacity before demand is known while ex post marginal cost is low and constant until firms reach their capacity constraint. The result is that competition involves large changes in price when demand changes. Cartels, like monopolies, raise price relative to competition when demand is low. When demand is sufficiently high, they behave like both perfectly competitive firms and monopolies by selling their entire capacity.

The similarity with monopoly ends there because the equilibrium capacity of legal cartels differs from the capacities that either perfectly competitive firms or monopolies would acquire in the same industry. Unlike monopolies, both the cartels in my model and those in Germany were unable to restrict the growth of capacity. As a result, Liefmann (1932) says that in Germany (p. 107) “... experience has hitherto proved that all cartels where they have been some time in existence have brought about a notable increase in the quantity produced and offered for sale; if this has not been due to the stimulus they have afforded to the founding of new firms, it has been caused by the members enlarging their own works.” Given that this expansion in output is not predicted by standard models of cartels while it is part of what makes cartels valuable in my model, it is interesting to note that Liefmann (1932, p. 65) sees it in an unfavorable light. He calls the “expansion of production and ... the entrance of new firms” an “unfavorable effect of the cartels.”

This expansion in output takes place in my model, at least for certain parameter values, when demand is high. The reason is that the constriction of output when demand is low raises expected profits. These profits attract entry so that total capacity is higher. Because ex post marginal cost is relatively low, the cartel may then sell all this extra capacity even though a monopolist would actually curtail output relative to the competitive level.2

Matsui (1989) also argues that the presence of a cartel can sometimes increase output. His logic is quite different from mine, however. His analysis is deterministic and shows not that cartels can increase output relative to perfect competition but that they can do so relative to the Cournot outcome. This occurs in his setting if the method used by the cartel to allocate output among its members creates incentives for members to acquire more capacity than they do in the Cournot equilibrium. If, for some reason, the cartel

---

2Matsui (1989) also argues that the presence of a cartel can sometimes increase output. His logic is quite different from mine, however. His analysis is deterministic and shows not that cartels can increase output relative to perfect competition but that they can do so relative to the Cournot outcome. This occurs in his setting if the method used by the cartel to allocate output among its members creates incentives for members to acquire more capacity than they do in the Cournot equilibrium. If, for some reason, the cartel
Even if cartels reduce the risk faced by firms by restricting output when demand is low and raising it when it is high, this need not imply that they are socially beneficial. If there were complete financial markets the competitive allocation would be Pareto optimal so that cartels could not possibly improve matters for all parties even if they did reduce risk for firms. On the other hand there are good reasons to suppose that, at least in Germany prior to World War I, there were limitations on the financial contracts into which owners of firms could enter.

This leads me to analyze settings where financial markets are incomplete so that, in effect, firm owners lose disproportionately when their cash flow is negative. This can lead industry capacity to be inefficiently low. Cartels can then increase the expected utility of consumers while keeping the utility of firm owners constant, or even increasing it, by encouraging a socially desirable industry expansion.

One related property of cartelization that has been repeatedly emphasized by its defenders is that it curtails inefficient bankruptcies in industry downturns. As Liefmann (1932, p. 143) writes “The cartels are often blamed for preventing the weak producers from being weeded out during depression.” He then immediately goes on to say, “But this [the weeding out] is not desirable, at least when, in the period of rising trade which follows, the necessary production is only possible through the erection of new works at a heavy cost involving selling prices at which even the weaker works could have survived.” One of the objectives of this paper is to demonstrate this benefit formally. While I simplify the analysis by making all firms identical, I show that cartels can reduce the closing and reopening of firms and that, given the financial market frictions I assume, this can be socially desirable.

The basic reason for this finding is once again that cartels can improve cash flow in bad states of demand. I show that this extra cash flow can prevent the inefficient closing of plants that comes about when financial markets work poorly and there is perfect competition. In chooses such an allocation rule, cartelization makes consumers better off while making firms worse off. In my analysis, firms do not lose profits as a result of cartelization and, sometimes, their expected profits even rise together with consumer welfare.

Lecraw (1977), for example, justifies the uranium cartel along these lines.
effect, this extra cash flow makes it possible for firms to remain in business when, without the cartel, they would be forced to close.

Since the reduction of price fluctuations is responsible for the welfare benefits of cartels in my model, the literature on price stabilization is an important antecedent. For example, Newbery and Stiglitz (1981) show that with incomplete risk markets, it may be socially beneficial to stabilize prices. Closer still, Judd (1980) shows that a minimum price floor which effectively stabilizes the price received by producers can be socially beneficial. This can occur in his setting because, in the absence of insurance, it induces more entry into the activity so that sales are higher in what would otherwise be high-price states.

While price stabilization and cartelization are similar from the point of view of their potential welfare benefits, they impose rather different requirements on governments who want to carry out these policies. For a government to stabilize prices, it must stipulate both a set of prices and a great many details about the commodity whose price it is setting. It thus imposes a heavier informational burden on government than the policy of enforcing cartel agreements. In the absence of breach, the government can then delegate price (and output) setting to whoever the industry designates. In the presence of breach, the government must acquire more information though even there, the relevant information may be relatively easy to obtain.

The paper proceeds as follows. Section 1 lays out a static deterministic model and shows that, in this case, cartels can only reduce output relative to competition. This result is reversed in Section 2 which shows that, when demand is random, cartels can increase output when demand is high. In this section, I assume that firms are risk neutral while financial markets are complete so that, in spite of their effects on output, cartels are worse than competition from a social point of view. Section 3 then shows that cartels can improve on the competitive allocation from a welfare point of view if firms face a large cost of financial distress. This cost implies that capacity is inefficiently low when there is perfect competition so that the increase in output brought about by cartels can increase consumer welfare. Section 4 compares cartels to monopolies facing the same conditions and shows that the
former raise output by more.

Section 5 considers the role of cartels in preventing the inefficient turnover of firms. For this purpose, I set up a model of inefficient firm shut-downs which may be of some independent interest. I suppose that there are large costs of financial distress, that firms need a minimum level of capacity to be operational and that firm owners have a limited amount of personal wealth that they can invest in their own firms. If this endowment is limited, firm entry is limited by the fear of financial distress. More interestingly, if this endowment is limited but not too small, it leads firms to exit the industry when demand is low and be immediately replaced by new firms. This occurs even though continuing with the firm’s operations costs less than starting a new firm (or equivalently the benefits of scrapping plants are lower than the costs of starting operation under new ownership). The exit is caused by the willingness of new owners with fresh endowments of resources to compete aggressively with existing owners whose endowment has already been exhausted as a result of a bad realization of demand.

Cartelized firms can be subject to similar problems. However, if cartelization leads firms to have a higher cash flow when demand is low, cartelized firms can survive even in situations where competitive firms would exit. In particular, I show that there is a range of owner endowments such that cartelized firms remain in business after a low realization of demand while competitive firms do not. Given that this implies that cartels produce at lower cost than competitive firms, it is not surprising that I am able to show that, in some cases, this leads cartels to be superior to competition from a welfare point of view. Section 6 concludes.

1 The Static Model

I assume that there are constant returns to scale but that there is a two-step production process.\footnote{The assumption of constant returns to scale is not critical for what follows except that it allows the equilibrium without cartels to take the usual price-taking form. With increasing returns to scale, the equilibrium without cartels must take some other form. However, much of the analysis goes through in the presence of fixed costs if one assumes that firms compete in Cournot fashion when cartels are outlawed.} In particular, a firm must first spend $vk$ to get a level of capacity equal to $k$. Then,
in the second step, the firm can set its output \( q \) anywhere between zero and \( k \) as long as its spends an additional \( cq \). Thus, a firm that plans to produce its entire capacity has total costs equal to \((c + v)q\) though, more generally, the cost of producing \( q \) is

\[
C(q) = cq + vk; \quad q \leq k. \tag{1}
\]

I consider a game in several stages. In the first stage, anyone can decide to acquire capacity and thereby become a firm in this industry. In other words, the game starts out with an infinite number of potential firms, each of which can spend \( vk \) and thereby acquire a productive capacity of \( k \). In the second stage, after these firms have sunk the costs of capacity, these firms can join a cartel.

Because my motivation is different, my model of cartel formation differs from more standard ones such as Selten (1973) (see also the exposition in Phlips (1995) and the references cited therein). Since I aim to discuss the welfare consequences of successful cartels, I introduce a model where cartels form in equilibrium regardless of the number of firms involved. Thus the model is clearly abstracting from anti-cartel forces that are relevant in practice. It is worth noting, however, that some German cartels involved a very large number of firms.\(^5\) One additional potential benefit of this model is that it may clarify why cooperation breaks down in more standard models as soon as the industry contains more than a handful of firms.\(^6\)

The standard approach is to suppose that firms who join a cartel cooperate with each other while simply accepting passively that those who do not join the cartel act as a competitive fringe. I suppose instead that the failure of some firms to join the cartel scuttles all cooperation. Empirically, there is something to be said for this approach because failure

\(^5\) Liefmann (1932, p. 26) says: “... the National Federation of Quadrangular Wire-netting Works, which is a loose price-convention, includes over 200 members ... Hexagonal Wire-netting, on the other hand, is only made by a dozen biggish firms, and these have a firm syndicate, which - with a few interruptions - is one of the oldest in the whole industry.” This paper does not distinguish between looser and tighter cartels though the nature of this distinction clearly deserves further research.

\(^6\) As can be seen in Shaffer (1995) and the references cited therein, results in the more standard treatment depend on whether firms use prices or quantities as a strategic variable. Nonetheless, monopoly outcomes of the sort I consider do not seem achievable in these models unless the number of firms is quite small.
of a sufficient number of firms to join in any particular German cartel did indeed lead the whole cartel to break down.\footnote{On the other hand, unlike what is required in my model, 100 percent participation seems not to have been necessary for cartels to be effective. As Liefmann (1932 p. 9-10) says “Experience has shown that normally about three quarters of the firms concerned must participate, otherwise monopolistic action becomes impossible.” From a theoretical point of view one could reasonably expect the failure of a small firm to join not to stop other, larger firms from colluding. Insofar this is true, small firms would of course prefer to be part of the “competitive fringe” rather than being part of the cartel. At the same time, as will be clear below, firms clearly benefit by their commitment to allow the cartel to collapse if even one firm fails to participate in it. Since this paper does not really explain the source of this commitment, and since in practice this commitment appears to have been present only partially, my model has to be seen as a simplification.}

I suppose that the act of joining a cartel consists of making a conditional offer. In particular, a firm that joins offers to be subject to large penalties if it fails to set its price and output in the way determined by a central entity (the cartel) \textit{as long as} all other firms make the same offer. There are two key features about this offer that deserve comment. First, because the offer is conditional on all firms making the same offer, it is irrelevant if any firm fails to make this offer. Second, the offer has a profound effect if all firms make it and the penalties for breach are enforced. In particular, if all firms make such an offer and the penalties for breach embedded in this offer are sufficiently large, the firms all find it in their interest to abide by the cartel’s decisions.

As in Cave and Salant (1995), I simplify the analysis by supposing that the cartel allocates output quotas in proportion to each firm’s capacity. Thus, the cartel chooses a single ratio $\bar{q}$ of individual firm output to individual firm capacity. In some parts of the analysis below, the constancy of this ratio of output to capacity emerges naturally because all firms are assumed to have the same size and it then seems natural to focus on symmetric equilibria. Even where I do not suppose that all firms are identical, the assumption of a constant ratio of output to capacity seems appealing because it eliminates any incentive on the part of firms to merge or split up in order to increase their output allocation. One additional advantage of this assumption is that, coupled with my assumptions on costs, it ensures that cartel members are unanimous with respect to the output that they wish the cartel to sell.

If all firms join the cartel, the third and last stage consists of choosing $\bar{q}$ to maximize individual profits. To do this, the cartel takes into account the demand function
\[ Q = D(P) \]  

where \( P \) is the industry price, \( Q \) is aggregate output and \( D \) is a decreasing function.

If any firm fails to join the cartel, I suppose that each firm maximizes profits by treating the price as outside its control. The price is then set by a Walrasian auctioneer in such a way that the resulting supply is equal to demand.\(^8\) For a given aggregate capacity \( K \), the resulting equilibrium quantity is

\[ Q = \min (K, D(c)). \]  

If the quantity demanded at a price of \( c \) is less than capacity, the market clearing price is simply the *ex post* marginal cost \( c \).

I now consider the choice of capacity \( K \) in the case where no cartel is ever expected to exist. In choosing their capacity, firms maximize profits taking as given their expectation of the price that will prevail when goods are finally sold. If they expect this price \( P \) to exceed \( c \), they expect to sell their entire capacity so that the expected profits of a firm that acquires \( k \) units of capacity equal

\[ (P - c - v)k \]

Free entry ensures that this expression is zero so that the equilibrium price \( P \) must be expected to equal \( c + v \). As a result, the equilibrium level of capacity \( K \) must be \( D(c + v) \) which implies that the first term in (3) is binding.

I now turn to the analysis of a cartel. To understand whether firms would join, I first study what a cartel would do if all firms that had entered in the first stage did indeed sign up. Supposing they do, the cartel sets \( \bar{q} \) to maximize individual firm profits. Given an aggregate

\(^8\)When demand is perfectly foreseen as in this section, the analysis is essentially unchanged if the absence of a cartel leads firms to set their individual outputs in Cournot fashion, *i.e.*, by treating other firm’s output choices as fixed. In either case the equilibrium price is \( c + v \) and all capacity is utilized. The reason is that, if the price were expected to be higher, firms would have an incentive to enter and expect to sell their entire capacity. By the same token, if this were the price, firms would not enter unless they could be sure to sell their entire capacity.
capacity $K$, aggregate output equals $\bar{q}K$ so that profits of a firm with capacity $k$ are equal to:
\[
(D^{-1}(\bar{q}K) - c)\bar{q}K \frac{k}{K}.
\]  (4)

The quantity $\bar{q}K$ that maximizes this expression depends just on $c$ and on the properties of the demand curve, so I denote it by $M(c)$. Since the cartel cannot set output above $K$, industry output $Q$ must equal
\[
Q = \min(K, M(c)).
\]  (5)

Two things are worth noting about this solution. First, all firms regardless of their capacity $k$ agree about what they would like aggregate output to be: they want it to be the level that maximizes industry profits.\(^9\) As a result output is the minimum of the level of capacity and the ex post monopoly level of output. Secondly, the monopoly level that is relevant is the one in which marginal cost is $c$ because, by the time the cartel sets $\bar{q}$ the capacity costs have already been sunk.

The question now arises whether the ex post monopoly level of output $M(c)$ is larger or smaller than the level of output without cartels, $D(c + v)$. If $v$ is sufficiently large that
\[
D(c + v) < M(c)
\]  (6)
the cartel does not change the equilibrium level of output. To see this, suppose (6) holds and the aggregate of firms' initial capacity is equal to $D(c + v)$. Then, the cartel would ask them all to produce at full capacity (since its preferred level of output given a marginal cost of $c$ is even higher) and the outcome would be the same as without the cartel. If firms enter with more total capacity, then the cartel would set output at a higher level than $D(c + v)$. The resulting price would be below $c + v$ so that firms would lose by installing capacity. Thus, no more than $D(c + v)$ units of capacity are installed. It is also easy to see that, the lure of profits in the case when the price exceeds $c + v$ ensures that no less than $D(c + v)$ units of capacity are installed. Thus, when inequality (6) is satisfied, firms are indifferent as to whether they join the cartel or not but output is the same in either case.

\(^9\)This contradicts the claim of section III A. of Cave and Salant (1995). Nonetheless, their analysis makes it clear that unanimity does not prevail in the case where marginal costs differ across firms.
Now suppose that (6) is violated so that the ex post monopoly output is less than perfectly competitive output with free entry. Since \( D(c) \) exceeds \( M(c) \) this occurs whenever \( v \) is sufficiently low. The result is that, if firms’ aggregate capacity were equal to \( D(c + v) \), the cartel would reduce output below capacity and thereby increase both aggregate and individual profits. Thus, all firms would profit from joining the cartel and they would all do so (remember that the failure of any firm to join leads the cartel not to be effective so that each firm is pivotal). Since there are positive profits in equilibrium if firms enter with \( D(c + v) \) units of capacity, more capacity is built initially. Thus the violation of (6) implies that the cartel reduces output below the competitive level while, at the same time, increasing capacity. Cartels are then worse for consumers while simultaneously increasing the resources absorbed by their industry.

Before closing this section, it is worth pointing out that the cartel outcome with free entry is different from the monopoly outcome. A monopoly would produce at a level equal to \( M(c + v) \) which is strictly less than both \( D(c + v) \) and \( M(c) \). So, the monopoly always curtails output while the cartel does so only in certain circumstances. Moreover, even when the cartel reduces output, free entry at the capacity building stage ensures that this curtailment of output is less severe than in the case of monopoly. On the other hand, the cartel may end up with excess capacity and this does not happen when there is a monopoly, at least when demand is deterministic as in this section. I consider random demand next.

2 Random Demand When Firms Maximize Expected Profits

The purpose of this section is to show that, under certain circumstances, the presence of an enforceable cartel actually increases output when demand is high. The reason this can occur is that the cartel raises profits when demand is low and this encourages firms to enter and/or increase their capacity. The resulting increase in aggregate capacity can then lead to increased sales. To show this in a simple model, I suppose that there are two states of demand which I denote by \( h \) and \( \ell \) and I let state \( h \) occurs with probability \( \lambda \). In state \( i \),
the quantity demanded is given by

\[ Q^i = D(P^i, \xi^i), \quad i = \ell, h \]  

(7)

where I suppose without loss of generality that \( \xi^h > \xi^\ell \) and that \( \frac{\partial D}{\partial \xi} \) is positive. I also assume that \( D(P^i, \xi^i) + P^i \frac{\partial D(P^i, \xi^i)}{\partial P^i} \) is increasing in \( \xi^i \) to ensure that the ex post monopoly level of output \( M(c, \xi^i) \) is increasing in \( \xi \).

Free entry coupled with the assumption that firms are risk neutral ensures that the industry’s ex post profits equal the aggregate costs of capacity or that

\[ v = \lambda \left\{ \frac{Q^h}{K}(P^h - c) \right\} + (1 - \lambda) \left\{ \frac{Q^\ell}{K}(P^\ell - c) \right\}. \]

(8)

This zero profit condition is common to both competition and cartels. In the case of perfect competition, it reduces to the requirement that the expected price be equal \( c + v \), the “full” marginal cost of a unit of output. To see this, let \( D^{-1}(x, \xi^i) \) represent the price at which the quantity \( x \) is demanded in state \( i \) while \( K_o^S \) denotes the equilibrium competitive capacity (where the \( S \) subscript stands for the fact that demand is stochastic). Equation (8) then implies that

\[ \lambda D^{-1}(K_o^S, \xi^h) + (1 - \lambda)[\max(D^{-1}(K_o^S, \xi^\ell), c)] = c + v. \]

(9)

where the terms multiplying the probabilities \( \lambda \) and \( 1 - \lambda \) are, respectively, the prices in the high and the low state so that the left hand side is the expected price. I now demonstrate this, as well as the fact that we can replace \( Q^i/K \) by one in (8) so that the two equations are equivalent. If the price when demand is low exceeds \( c \), firms sell their entire capacity so that \( Q^\ell \) is equal to \( K_o^S \) and the price is \( D^{-1}(K_o^S, \xi^\ell) \). If \( D^{-1}(K_o^S, \xi^\ell) \) is less than \( c \), however, the firms charge \( c \) in the low state because they would not be willing to sell anything for less. In this case, we can still replace \( Q^\ell/K \) by one in (8) because \( Q^\ell/K \) is irrelevant since it multiplies the ex post profits which are zero. Thus, in either case, we can replace \( Q^\ell/K \) by one and the term multiplying \( 1 - \lambda \) in (9) is the price in the low state. In the high demand state, the price must exceed \( c \) so that \( Q^h/K \) is one because, otherwise, firms would not recoup their capacity costs. This implies that all of capacity is sold in state \( h \) so that the price in this state is indeed \( D^{-1}(K_o^S, \xi^h) \). Equation (9) is thus indeed equivalent to (8).
I now turn to the case where there is a cartel in stage two and I denote the resulting equilibrium level of capacity by \( K^S_0 \). For a given level of capacity \( K \), total output is now given by (5). Because \( M(c, \xi^h) > M(c, \xi^\ell) \), it follows that if the capacity constraint \( K \) is binding in state \( \ell \), because \( M(c, \xi^\ell) \) is larger than \( K \), it is also binding in state \( h \). This leaves three possible types of outcomes. The capacity constraint can be binding in both states, in neither state, or just in state \( h \). The first two are analogous to the two outcomes we saw in the deterministic case. The randomness in demand makes the third kind of outcome possible and this allows output to rise when demand is high.

The capacity constraint binds in both states so that the existence of the cartel has no effect on output when

\[
K^S_0 < M(c, \xi^\ell). \tag{10}
\]

The cartel would then choose ex post a level of output higher than the competitive capacity level \( K^S_0 \) even when demand is low. As in the static case in which (6) holds, equilibrium capacity and output are then the same with and without cartels.

The capacity constraint binds in neither state when \( v \) is sufficiently low that

\[
K^S_0 > M(c, \xi^h) \tag{11}
\]

The cartel then sells \( M(c, \xi^i) \) in state \( i \) so that output is below \( K^S_0 \) (and below the competitive level) in both states. Since this increases profits when capacity is \( K^S_0 \) capacity is increased further. Thus, as in the static case when (6) is violated, the cartel produces less (in both states of nature) than do competitive firms even though its capacity is higher.

The randomness in demand makes it possible for the analogue to (6) to hold when demand is high - so that (11) is violated - while the analogue to (6) is violated when demand is low - so that (10) is violated as well. This occurs when

\[
M(c, \xi^\ell) < K^S_0 < M(c, \xi^h). \tag{12}
\]

so that the capacity constraint only binds in state \( h \). I now show that, this implies that the cartel expands output in this state. Because (10) is violated, firms in a cartel with a
capacity of $K^S_a$ make more profits in state $\ell$ than do firms that act competitively. The result of this increased profitability is that capacity expands beyond $K^S_o$. Because (11) is violated the cartel would prefer, ex post, a level of output higher than $K^S_o$ in state $h$. This implies that this expansion in capacity raises output in state $h$. Therefore, (8) implies that $K^S_a$ must satisfy

$$vK^S_a = \lambda \left\{ [P^h - c]Q^h \right\} + (1 - \lambda) \left\{ [D^{-1}(M(c, \xi^h), \xi^h) - c]M(c, \xi^h) \right\}$$

(13)

where

$$Q^h = \min(M(c, \xi^h), K^S_a) \quad P^h = D^{-1}(Q^h, \xi^h).$$

The extent to which output increases in state $h$ relative to the competitive level of output depends on the parameters. It is equal to the increased level of capacity $(K^S_a - K^S_o)$ when $K^S_a \leq M(c, \xi^h)$ but it is less than this increase in capacity when $K^S_a$ is greater than $M(c, \xi^h)$. In any event, the output increase in state $h$ makes cartelization with free entry different from monopoly. A monopolist would maximize

$$\lambda K \left[ D^{-1}(K, \xi^h) - c \right] + (1 - \lambda) \min(K, M(c, \xi^h)) \left[ D^{-1} \left( \min(K, M(c, \xi^h)) \right) - c \right] - vK$$

(14)

with respect to $K$. This leads to a capacity which is smaller than $D(c + v, \xi^h)$ since the optimal price is greater than “full” marginal cost $(c + v)$ in the high demand state. Thus, as usual, a monopolist always reduces output below the competitive level.

The increased output of the cartel in state $h$ implies that, in state $h$, ex post profits per unit of capacity, $(P^h - c)Q^h/K$, are lower than under competition. Given (8), this implies that, relative to competition, the cartel raises ex post profits per unit of capacity in state $\ell$ so that it stabilizes ex post profits. This result will prove important below.

The result that output increases in the high demand state as a result of cartelization is reminiscent of the result in Deneckere, Marvel and Peck that resale price maintenance (RPM) can raise output when demand is high. What resale price maintenance and cartelization have in common is that they prevent prices from falling when demand is low. They thus increase profits in low demand states and encourage additional capacity. In Deneckere et. al., this additional capacity takes the form of additional purchasers by retailers from the upstream
monopolist who imposes RPM whereas here it is the cartelized firms that increase their productive capacity. The two results thus have some common elements. However, there is a crucial difference. The difference is that the conditions under which Deneckere et. al. find that RPM raises output are ones where the monopolist imposing RPM lowers the price at which he sells his good to the retailers. This reduction in the upstream price brought on by RPM is also the reason why, in their model, RPM can increase consumer surplus even though, as in this section, consumers and firms are risk neutral.

By contrast, the assumption that costs are independent of cartelization implies that cartelization with risk neutrality cannot make consumers better off. The reason is simple. Without cartelization, the outcome is Pareto optimal by the first welfare theorem. Cartelization keeps profits of producers equal to zero. This means that cartelization cannot increase the welfare of consumers (for that would contradict the Pareto optimality of the competitive allocation). This negative result about cartelization extends well beyond the confines of the simple example I have been developing. In particular, it holds true for arbitrary cost functions and arbitrary randomness. It also holds for arbitrary levels of risk aversion as long as there are complete financial markets.

This leads me to consider a quite different reason why cartels might be Pareto desirable in the next section. In particular, I focus on a setting with financial market imperfections.

3 Random Demand with Incomplete Financial Markets

In this section I consider a simple variant of the model in which the legality of cartels does indeed lead to a Pareto improvement in the allocation of resources. The basic advantage of cartelization is that, as in the last section, output increases when demand is high because firms install more capacity. The difference is that I start from a situation where competition leads to a level of capacity that is too low so that this increase in capacity can increase overall welfare. The particular model I consider has a level of capacity that is too low because financial markets are incomplete and firm owners cannot insure themselves against
random changes in demand.

I consider two settings with very similar implications. In the first, owners finance their operations from unlimited internal funds but they are risk averse while consumers marginal utility of income is the same in both states of nature (so that they are effectively risk neutral). In the second, owners are risk neutral but their internal resources are limited so they must borrow to finance their investment in capacity. Because there are huge penalties for default, firms borrow only the amount that they can repay with probability one and this keeps equilibrium capacity low.

I simplify the analysis of the risk averse self-financing firm by supposing that the firms’ risk aversion is extreme: the owner’s utility is the minimum of total ex post profits in both states of demand. This minimum of profits affects utility in the same way that reductions in the amount spent on capacity. Thus, a firm with capacity $k$ which faces a price $P^i$ in state $i$ has an expected utility of

$$U^F = -vk + k \min\left(\left(P^h - c\right)q^h, \left(P^\ell - c\right)q^\ell\right)$$

(15)

where $q^i$ represents the fraction of its capacity that it sells in state $i$. In addition to capturing risk aversion, this utility function ensures that firms do not invest in capacity unless they earn a nonnegative expected return.

For investment to be worthwhile in (15), both $P^h$ and $P^\ell$ must be bigger than $c$ so $q^i$ is equal to $k$. Thus, in choosing $k$, owners maximize

$$k\left[P^\ell - c - v\right].$$

(16)

Free entry then implies that $P^\ell$ is equal to $c + v$ so that equilibrium capacity is $D(c + v, \xi^\ell)$. I also use $K^F_0(0)$ to denote this capacity so that the financial market incompleteness is reflected in the superscript $F$.

An alternative specification of owner preferences with similar implications is to suppose that the owner’s resources are limited while the cost of bankruptcy is high. Suppose in particular that the owner’s initial resources equal $R$. If $R$ is strictly positive and firms can
enter with infinitesimal capacity this boundedness of the owner’s endowment does not matter given my assumption that average costs are independent of scale. I thus assume also that there is a minimum level of capacity which, by normalizing output, I set equal to the capacity needed to produce one unit of output.

Then, if \( R < v \), firms who wish to acquire capacity must raise funds from financial markets.\(^{10}\) I assume that lenders in these markets demand an expected return of zero (as I have been assuming implicitly by not discounting revenues relative to the capacity cost \( v \)) but that they have extremely poor information about the firm and its opportunities. Thus, after the firm raises funds, lenders (and contract enforcers) know nothing about the firm other than whether it has repaid the amount that is specified at the time the firm borrows. Repaying even slightly less than this amount leads to a large utility penalty of \( \bar{L} \). Thus, the only feasible financial contract is a debt contract with large penalties for default.\(^{11}\) I denote the amount that needs to be repaid by \( B' \) (which obviously depends on \( B \)).

The owner of the firm then has the following choice. He can get \( R \) units of utility by not investing in capacity at all. Alternatively, he can buy one unit of capacity using borrowed funds equal to \( B \) as well as \( Y \) of his own resources where

\[
B = v - Y \quad Y \leq R. \tag{17}
\]

Supposing that any funds not invested in the project earn a zero rate of return and denoting by \( B' \) the contractually stipulated repayment, the resulting utility of the owner is

\[
U^R = R - Y - B' + \sum_{i=\ell,h} \pi^i \left\{ (P^i - c)q^i - L\left( (P^i - c)q^i - B' + R - Y \right) \right\} \tag{18}
\]

where

\[
L(x) = \begin{cases} 
\bar{L} & \text{if } x < 0 \\
0 & \text{otherwise}
\end{cases}
\]

\(^{10}\) I could have assumed that, in addition, the firm needs to borrow for its “working capital”, which equals \( cqk \) where \( q \) is the ratio of output to capacity. This would have made no difference to the analysis, however, in part because I make the rate of return demanded by lenders equal to zero.

\(^{11}\) This debt contract is arising under assumptions which are much stronger, though somewhat related to those of Townsend (1979) and Gale and Hellwig (1985).
and where I use $\pi^i$ to denote the probability of state $i$ (so that $\pi^h = \lambda$). With $\bar{L}$ sufficiently large, the firm makes sure that it repays $B'$. This has two consequences. First, the firm ensures that its resources on hand are sufficient to make this payment so that

$$[P^i - c]q^i - B' + R - Y \geq 0 \quad i = \ell, h. \quad (19)$$

Second, the fact that $B'$ is always repaid while lenders demand a zero rate of return implies that $B'$ is equal to $B$. Because $q^h$ is one and $P^h > P^\ell$, it must be that case that, if the constraint (19) binds at all, it does so in the low demand state. Using (17), (19) then implies that

$$q^\ell(P^\ell - c) - v + R \geq 0. \quad (20)$$

If $R$ equals or exceeds the cost of capacity $v$, the financial market imperfection is irrelevant. Otherwise, (20) implies that the price $P^\ell$ is no smaller than $c + v - R$ so that, ignoring integer constraints as I do throughout, equilibrium capacity is $K^F_o(R)$ where

$$D(c + v - R, \xi^\ell) = K^F_o(R) \quad (21)$$

When $R$ is equal to zero $K^F_o(R)$ is the same as the capacity with infinitely risk averse owners. The capacity $K^F_o(0)$ requires that the price be equal to $c + v$ in the low demand state so that the average price exceeds $c + v$. Thus, the financial imperfection lowers equilibrium capacity below $K^S_o$ in this case. More generally, capacity falls below $K^S_o$ if $R$ is low enough and, because financial imperfections only matter in this case, I focus on it from now on.\(^{12}\)

Subsuming the situation where firms maximize $U^F$ into the one where $R$ equals zero, the earlier analysis should make it clear that the enforceability of cartel agreements has no effect if

$$M(c, \xi^\ell) > D(c + v - R, \xi^\ell) \equiv K^F_o(R) \quad (22)$$

\(^{12}\)One interesting aspect of this model is that, expected profits must be positive when $K^F_o(R)$ is smaller than $K^S_o$ since the average price is higher. Thus, when the financial constraint is binding, owners who succeed in getting loans get an increase in utility relative to owners who do not. This is, in effect, a form of credit rationing since identical individuals are not treated symmetrically and some end up better off than others. If the minimum capacity is small, $c + v$ and $P^h$ are small so that this surplus is small. If the minimum capacity is substantial, however, these profits can be large.
The cartel, like competitive firms, then sells its entire capacity of \( D(c + v - R, \xi^\ell) \) in state \( \ell \) and cannot afford to acquire more capacity.

If (22) is violated, the cartel produces only \( M(c, \xi^\ell) \) in state \( \ell \) so that its output falls relative to the competitive case (and the price is higher than \( c + v - R \)). If, in addition,

\[
M(c, \xi^h) < D(c + v - R, \xi^\ell),
\]

the cartel also reduces output in the high demand state. Thus, the cartel reduces welfare when (23) holds.

I now show that, when both (22) and (23) are violated, the cartel increases output in state \( h \). The violation of (22) ensures that, in state \( \ell \), the cartel sells \( M(c, \xi^\ell) \) and increases its profits by charging more than \( c + v - R \). If aggregate capacity were equal to \( K_F(R) \), the financing constraint (20) would therefore not be binding in state \( \ell \). Since even competitive firms make positive expected profits with this capacity, further entry takes place and output rises.

The constraint (20) becomes binding only when the ex post profits in state \( \ell \) are just sufficient to pay for \( v - R \). In other words, the financial market imperfection prevents capacity from exceeding \( K_F^*(R) \) which satisfies

\[
\left[D^{-1}(M(c, \xi^\ell), \xi^\ell) - c\right] \frac{M(c, \xi^\ell)}{K^F(R)} = v - R.
\]

Equilibrium capacity with a cartel is then equal to the minimum of \( K_F^*(R) \) and \( K_S \).

I now show that, when (22) and (23) are violated, welfare can be higher with a cartel than with perfect competition. Consider first the firms. If owner utility is given by \( U^F \), free entry ensures that \( v \) is equal to second term in (15). Thus means that ex ante owners are indifferent whether there is a cartel or not. Even so, owners prefer to get larger profits ex post so that a higher level of \( K[(P^h - c)q^h - v] \) makes them better off in the aggregate. The same is true when owner utility is given by \( U^B \). The financial constraint then ensures that (20) is zero so that, ignoring constants, \( U^B \) is equal to

\[
U^B = \lambda[(P^h - c)q^h - v].
\]
Rather than analyzing how the surplus of each firm changes when one moves from competition to a cartelized outcome, I study the aggregate surplus received by all firms

\[(P^h - c)Q^h - vK. \tag{26}\]

With perfect competition, \(Q^h\) and \(K\) equal \(D(c + v - R, \xi^\ell)\) so that this surplus equals

\[D(c + v - R, \xi^\ell)[D^{-1}(D(c + v - R, \xi^\ell), \xi^h) - c - v]\]

whereas it equals

\[\min(M(c, \xi^h), K^F_a(R))\left[D^{-1}\left(\min(M(c, \xi^h), K^F_a(R))\right) - c\right] - vK^F_a\]

in the case where the industry is cartelized. Whether the former is greater than the latter depends on the parameters. If the monopoly output in the high state is smaller then the quantity demanded in the low state when the price is \((c + v - R)\) so that

\[M(c + v, \xi^h) < D(c + v - R, \xi^\ell) \tag{27}\]

overall producer surplus is unambiguously smaller when there is a cartel. The reason is that, assuming \(Q^h\) equals \(K\), (26) reaches a maximum when \(K\) is equal to \(M(c + v, \xi^h)\). If the competitive outcome involves a higher capital stock it involves a lower surplus and the surplus of the cartel equilibrium, whose \(K\) is even higher, must be smaller still. If, instead, (27) is violated, an increase in overall capacity starting from the competitive outcome increases aggregate profits in the high state.\(^{13}\) Even in this case, the cartel increases profits only if it does not raise capacity far beyond \(M(c + v, \xi^\ell)\).

I now turn to consumer welfare. Cartelization clearly leads to consumer losses in the low demand state since the violation of (22) implies that output is lower in this state. On the other hand, the increase in output in the high demand state is good for consumers. To compare these effects, I suppose that there is a representative consumer with a constant marginal utility of income and that his overall utility in state \(i\) can be written as

\[W(Q^i, \xi^i) + I^i \quad \text{where} \quad I^i = I^i_0 - P^iQ^i \tag{28}\]

\(^{13}\)Note that, for \(\xi^h\) sufficiently close to \(\xi^\ell\), (27) is sure to be violated; it is only if the states of demand are quite different from one another that this condition can hold.
and \( I_0^i \) is their level of assets (or income) before he has paid for the purchases of \( Q^i \). Ignoring variables that are independent of industry actions, expected utility is then

\[
\lambda \{ W(Q^h, \xi^h) - P^h Q^h \} + (1 - \lambda) \{ W(Q^\ell, \xi^\ell) - P^\ell Q^\ell \}.
\] (29)

Since the reduction in \( P^h \) raises the first term in brackets, expected utility must rise if \( \lambda \) is sufficiently close to one. Thus both producer and consumer welfare can increase as a result of industry cartelization. This occurs when the parameters are such that cartelization raises output when it matters a great deal (i.e., when consumer welfare and firm profit are raised significantly by increases in output) whereas it reduces output when it matters less.

Before closing this section, it is worth analyzing the effect of cartelization on average price and productivity, two measures of economic performance closely related to those studied empirically in the study of Swedish cartels by Fölster and Peltzman (1997). One reason for doing so is to show that cartels can be welfare improving even if they reduce productivity and raise average prices. In my models of financial market imperfection, cartels unambiguously reduce productivity relative to perfect competition as long as they have any allocative effect whatsoever. To see this, note that the level of output per unit of inputs is \( Q / (vK + cQ) \).

Under perfect competition, the capital market constraint ensures that firms sell their entire capacity in both states of demand so that this measure of productivity equals \( 1/(c + v) \). If cartels matter, they reduce output below the competitive level when demand is low. Since this also leads capacity to increase, productivity in the low demand state must fall below \( 1/(c + v) \). In the high demand state productivity stays the same if the increase in output is the same as the increase in capacity whereas it falls otherwise. In any event, average productivity must fall.\(^{14}\)

The effect of cartelization on average price is ambiguous.\(^{15}\) The reason is that cartels,

\(^{14}\) This is somewhat consistent with the finding of Fölster and Peltzman (1997). They find that cartels tend to reduce the growth in total factor productivity. In my setup, this growth decline would follow the creation of a cartel but it would be more appropriate to look at the effects on the level of productivity. While Fölster and Peltzman (1997) also study the level of labor productivity, this is affected by the intensity of materials use for which they do not have data.

\(^{15}\) This ambiguity is consistent with the weakness of the effect of Swedish cartels on Swedish prices (as compared to EEC prices for the same good) found by Fölster and Peltzman (1997). They find much stronger
if they matter, raise price when demand is low but lower it when demand is high. This ambiguity is reminiscent of the effect of cartelization on consumer welfare because this too falls when demand is low and rises when demand is high. However, the effect of cartelization on average welfare does not necessarily have the same sign as the effect on average price. To see this, it is worth analyzing a particular example, namely one where demand is linear so that price in state $i$ is given by

$$P^i = a^i - bQ^i \quad i = \ell, h \quad \text{and} \quad a^h > a^\ell. \quad (30)$$

Since the inverse demand function $D^{-1}(Q^i, \xi^i)$ must equal the partial derivative of $W$ with respect to $Q^i$, consumer welfare (28) must then be

$$a^iQ^i - \frac{b}{2}(Q^i)^2 + I^i - P^iQ^i \quad (31)$$

where I have ignored the constant of integration. Using (30) in (31) implies that, ignoring constants, consumer welfare is

$$\frac{b}{2}(Q^i)^2. \quad (32)$$

That consumer welfare depends on the square of output when demand is linear is not surprising since the area under the demand curve is a triangle whose height is $Q$. What is somewhat more interesting is that this dependence on the square of output is unaffected by vertical demand shifts so that (32) implies that expected utility is proportional to the average of the squared values of output.$^{16}$ By contrast, (30) implies that the average price depends on the average level of output. Since cartels tend to stabilize prices in response to demand fluctuations, they destabilize quantities. This effect tends to raise consumer welfare for any given change in the expected price. I now illustrate this possibility with a numerical example.

---

16 This says that the variability in consumption of a particular good is desirable for a given mean of this consumption. This conclusion hinges in part on letting consumers adjust their consumption of other goods as one varies tastes and the price of this particular good.

---
With the financial market imperfections I have considered, competitive capacity $K^F_o$ with this linear demand is $(a_\ell - c - \nu)/b$. The cartel matters only if (22) is violated, which requires that the ex post monopoly price $(a_\ell + c)/2$ be in excess of the competitive price $(c + \nu)$. I start by supposing that $a^h = 20$, $a^\ell = 9$, $c$ is zero, $\nu$ is 3 and $\lambda$ is .65. Competitive capacity and output are then equal to 6 while cartel capacity is 6.75. Because (22) is violated at these parameters, cartel output in state $\ell$ is lower than competitive output and equals only 4.5. The effect of this large decline in output is that the average price of the cartel is slightly higher; it is 10.19 instead of the 10.15 which prevails with perfect competition. Nonetheless, the average of the square of output is higher with the cartel; it equals 36.7 instead of 36 under competition. Thus consumer welfare is higher with the cartel. Moreover, all inclusive firm profits in state 1 defined as $(P - c)Q - \nu K$ are higher with the cartel; they equal 66.19 instead of the competitive value of 66.17. Thus, overall welfare is definitely larger as well.

This example illustrates that average price can rise with the cartel (and average quantity fall) even though consumer welfare increases. This by no means implies that welfare increases require that average prices rise, nor that cartels necessarily raise average prices (average prices fall when $\nu$ is set equal to 2.5 instead of 3 while the other parameters stay unchanged). All that I have shown is that the connection between the effect of cartelization on welfare and its effect on average prices and quantities is somewhat loose.

### 4 Comparison with Monopoly

This section shows that, from the point of view of consumers, cartelization with free entry is never worse and is often superior to full monopolization of the industry even with the financial constraints considered in the previous section. When firms can only use their owners’ funds and the owners are infinitely risk averse, I show that monopolists always produces less than cartels. Even when the problem is that the amount that either type of firm can borrow is dictated by cash flow in state $\ell$, monopolists often produce less, and never

---

17 Obviously, this increase in profits is not required for the other results to hold. Keeping all other parameters the same and lowering $a^h$ so that it equals 15 leads profits in state $h$ to fall even though average price and expected consumer utility are still higher with the cartel.
produce more. The reason for this is the following. This financial constraints limits the amount of capacity that can be invested in the industry. If a monopolist wants to invest more than this feasible capacity, it finds itself with the same capacity as the corresponding cartel. However, monopolists often desire to invest less than this maximum feasible capacity for the reasons highlighted in section 1. In these cases, cartels produce more output in at least one state of demand.

The effect of monopoly is different depending on whether owners maximizes $U^F$ or $U^B$. Consider first the case where they maximize $U^F$. No matter what capacity $K$ a monopolist installs, the full profits, $(PQ - cQ - vK)$, are smaller in state $\ell$ than in state $h$ given that $Q$ is chosen optimally ex post. This means that a monopolist who maximizes $U^F$ maximizes $Q[D^{-1}(Q, \xi^\ell) - c] - vK$ subject to $Q \leq K$

He thus sets $K$ and $Q$ equal to $M(c + v, \xi^\ell)$ which is less than the competitive quantity $D(c + v, \xi^\ell)$. Thus a monopolist unambiguously lowers output in this case.

Now suppose a monopolist maximizes $U^B$ where I assume again that $R$ is zero. If the financial constraint were absent, the monopolist would install the capacity $K^S_m$ which maximizes (14). If

$$K^S_m < D(c + v, \xi^\ell)$$

(33)

so that this capacity is less than the competitive output in state $\ell$, the monopoly sells strictly less than the cartel in the high demand state and no more in the low demand state. In the high demand state, the cartel sells the lesser of $D(c + v, \xi^\ell)$ and $M(c, \xi^h)$. The sales of the monopolist are less than the former because of (33) and are less than the latter because the maximand of (14) is less than $M(c + v, \xi_h) which is itself smaller than $M(c, \xi^h)$. In the low demand state, the monopolist sells either his capacity $K^S_m$ or $M(c, \xi^\ell)$, whichever is less while the cartel sells the lesser of $D(c + v, \xi^h)$ and $M(c, \xi^\ell)$. Therefore, when (33) is satisfied, the monopolist sells no more than the cartel and sometimes sells less.

I now suppose that (33) is violated and show that the monopoly still never sells more, and often sells less than the corresponding cartel. If (22) holds then the monopolist, like the
cartel, installs the same capacity as perfectly competitive firms. The reason is that this is the highest capacity consistent with the financial constraint. When, instead, (22) is violated, the cartel increases its capacity to \( K^F_a(R) \), which is given by (24). If \( K^S_m \) is larger than \( K^F_a(R) \) then the increase in capacity of the monopolist is the same as the increase in capacity of the cartel. This is the maximum capacity that can be financed given the revenues that accrue in the low demand state when output is set at \( M(c, \xi^l) \). But, as long as \( K^S_m \) is less than \( K^F_a(R) \) (which is itself strictly more than \( D(c+v, \xi^l) \) when (22) is violated), the monopolist increases his capacity by less than the cartel. The cartel then unambiguously sells more in the high demand state because it sells the lesser of its capacity (which exceeds the monopolist’s) and \( M(c, \xi^h) \) which, by the argument above, exceeds the monopolist’s unconstrained capacity \( K^S_m \).

5 Avoiding the Inefficient Closing of Plants

This section demonstrates that the presence of cartels can reduce firm turnover and that this can be socially beneficial. In particular, cartels can reduce the extent to which firm exit is matched by the creation of new, more costly, firms. To demonstrate this, I first need to construct a model in which this socially costly turnover actually takes place.

To analyze firm turnover, a dynamic model is necessary since exit at a point in time must sometimes be followed by new entry at a later date. I consider the simplest model where this is possible, namely one with only two periods. In each period demand is given by (7). The states of demand in the two periods are independent and the probability that \( i \) is equal to \( h \) in any given period remains \( \lambda \). In each period, owners can spend \( v \) and build the minimum capacity that is needed for effective production. I assume once again that there is an infinite potential supply of owners, each of which starts out with resources equal to \( R \). The utility of all owners is given by a function such as \( U^B \), suitably extended to take into account the existence of two periods.

Entrants in the first period have an advantage over entrants in the second period. In particular, they can spend \( v' < v \) in the first period and thereby keep their capacity oper-
ational in the second period. The model can also be interpreted as one in which firms pay $v + v'$ per unit of capacity initially and where this capacity is productive for two periods unless the owner scraps the capacity and thereby receives $v'$ at the end of the first period. Because I continue to abstract from discounting, the expenditures in the two periods are strictly comparable. Therefore, this difference in costs implies that no owner would fail to spend $v'$ in the first period and then pay $v$ to get his operations restarted in the second period. Therefore, I deem a firm to exit the industry if it enters in the first period and then fails to spend $v'$ (or scraps the capacity for $v'$) so that its plant is not operational in the second period. This exit often leads to new entry in the second period. One can think of the situation in which an existing owner fails either fails to maintain or scraps his plant while an entrant pays $v$ as a situation where a new owner buys the plant from the old owner and where this transfer of control is costly.

To understand the equilibrium, I proceed backwards and start with the entry decisions in the second period. In this period, owners who enter the industry with the minimum level of capacity get a utility of $U_{2B}$ which is now given by

$$U_{2B} = R - v + \sum_{i=\ell,h} \pi^i \left\{ \left[ P_i^2 - c \right] q_i^2 - L \left( \left[ P_i^2 - c \right] q_i^2 - v + R \right) \right\}$$

which is identical to (18) except that the subscript indicates the period in which $P$ and $q$ are in effect. Also, this equation already takes into account that both the amount borrowed $B$ and the amount repaid $B'$ equal $v - Y$. If no capacity is inherited from period one, the outcome in the second period is the same as in the one period stochastic model I considered in section 3. With perfect competition, for example, capacity in period 2, $K_{o2}$ is given by

$$K_{o2} = \min(K_o^F(R), K_o^S).$$

If inherited capacity is less than $K_{o2}$, entry in period 2 brings capacity up to this level. If inherited capacity is larger, no firm enters in period 2. The second period behavior of firms in a cartel is analogous.

I now analyze capacity investment in the first period. I let $Y_j$ denote the amount of its own funds that a firm spends on capacity in period $j$, where neither $Y_j$ nor the sum $(Y_1 + Y_2)$
can exceed the available resources $R$. I also let $B_{ij}$ denote the amount the firm borrows in state $i$ of period $j$. Given that the owner earns a zero rate of return on any resources he does not invest in the project and that this is also the return demanded by lenders, there is no loss of generality in supposing that no repayment of debt takes place in the first period and that this debt is simply rolled over. However, the resulting period 2 debt must be paid back in full to avoid the cost $L$.

I denote the initial capacity by $k_1$ and the capacity that is maintained by suitable expenditure of $v'$ in state $i$ of period 1 by $k_{i2}$. I capture the minimum capacity constraint by supposing that $k_{ij}$ can only equal either one or zero. Overall utility of the owner is then

$$U_1 = R - Y_1 + \sum_{i=\ell,h} \sum_{m=\ell,h} \pi^i \pi^m \left\{ -Y^i_2 - B^i_2 + k_{i2}^m [P^m_2 - c] q^m_2 
- L \left( k_{i2}^m [P^m_2 - c] q^m_2 - B^i_2 + R - Y_1 - Y^i_2 \right) \right\}$$

while

$$B_1 = v k_1 - Y_1 \quad \quad B^i_2 = v' k_{i2} - k_1 [P^i_1 - c] q^i_1 + B_1 - Y^i_2.$$  

Substituting the expressions in (37) in (36), we have

$$U_1 = R - \sum_{i=\ell,h} \sum_{m=\ell,h} \pi^i \pi^m \left\{ k_1 ([P^i_1 - c] q^i_1 - v) + k_{i2}^m ([P^m_2 - c] q^m_2 - v') 
- L \left( k_1 ([P^i_1 - c] q^i_1 - v) + k_{i2}^m ([P^m_2 - c] q^m_2 - v') + R \right) \right\}.$$  

If one could ignore the terms involving $L$, and every firm sells the same fraction of its capacity (as required by the cartel arrangement), owners are indifferent between staying out and building capacity for two periods as long as

$$2\lambda \left\{ \frac{Q^h}{K} (P^h - c) \right\} + 2(1-\lambda) \left\{ \frac{Q^l}{K} (P^l - c) \right\} - (v + v') = 0.$$  

In the case of perfect competition, we can set $Q^i/K$ equal to one in (39) because firms sell their entire capacity unless the price is $c$, at which point the level of $Q^i/K$ is irrelevant. Thus, this zero profit condition would lead perfectly competitive firms to set a capacity $K^U_0$ such that the expected price in each period is $c + \frac{v+v'}{2}$. Because the average capacity cost is
lower when this capacity is built for two periods, $K^U_o$ exceeds the unconstrained one period capacity level $K^S_o$ (which leads to an average price of $c + v$). On the other hand, $K^U_o$ is less than the capacity that would yield zero profits with perfect competition if the capacity cost were $v'$. Since, at the end of the first period firms perceive the cost of their capacity to be $v'$, this means that they would all choose to maintain (or fail to scrap) their capacity at this point. Thus, when the financial constraints are absent, $K^U_o$ is built in the first period and maintained into the second. That no capacity is scrapped is not surprising since capacity costs are lower in the second period and the benefits per unit of capacity are the same. Because $K^S_o$ is smaller than $K^U_o$, no new capacity is then built in the second period.

The term in (36) involving $L$ may prevent $K^U_o$ from being maintained, however, even if it has been built initially in period 1. This term is most likely to have a costly effect on the firms’ decisions if the state of demand is $\ell$ in both periods. The reason is that low demand at time 1 means that the firm needs a large additional loan to maintain its capacity while low demand at time 2 makes such a loan difficult to repay. If $i$ and $m$ are both equal to $\ell$, the last term in (38) is nonnegative only if

$$k_1([P_1^\ell - c]q_1^\ell - v) + k_2^L([P_2^\ell - c]q_2^\ell - v') + R \geq 0.$$  

(40)

For firms to satisfy (40) while, simultaneously having the aggregate capacity that would prevail if the $L$ term were absent, their resources must equal at least $R^U_o$ where

$$R^U_o = v + v' - 2\max(0, D^{-1}(K^U_o, \xi^\ell) - c).$$  

(41)

If $R$ is lower than $R^U_o$, the constraint (40) implies that equilibrium capacity is different from the unconstrained level $K^U_o$.

A key effect of cartelization is that it can reduce the level of resources owners need to have in order to ensure that the $L$ terms are irrelevant. Owners of cartelized firms require strictly less than $R^U_o$ to be unconstrained in this way if

$$M(c, \xi^\ell) < K^U_o.$$  

(42)
When this condition is violated a cartel with a capacity of $K_o^U$ sells the same as a competitive industry in state $\ell$. Thus, just as when (10) is satisfied in Section 2, the competitive outcome remains the equilibrium when cartel contracts are enforceable. When (42) holds, the cartels sells less in state $\ell$ than a competitive industry with capacity $K_o^U$ and its ex post profits are higher.\footnote{Notice that this occurs whether the competitive industry sells $K_o^U$ or not because, if it doesn’t, the equilibrium price must be $c$ and the cartel would charge more.} This leads firms to build extra capacity so that either $P^h$ or $Q^h/K$, or both are lower in the cartelized equilibrium with zero profits. The zero profit condition (39) then implies that the ex post profit per unit of capacity in state $\ell$, $(P^\ell - c)Q^\ell/K$ is higher with a cartel, where I denote it by $\pi_{\alpha \ell}$, than with perfect competition, where it equals $\max(0, D^{-1}(K_o^U, \xi^\ell) - c)$. The minimum $R$ for the cartel to be unaffected by $L$ is $R_a^U$ where

$$R_a^U = v + v' - 2\pi_{\alpha \ell}. \tag{43}$$

which is now less than $R_o^U$. Thus cartels need fewer resources to be unconstrained as long as (42) holds. Of course, the attractiveness of the cartel from a welfare point of view increases if, in addition,

$$M(c, \xi^h) > K_o^U \tag{44}$$

because this means the cartel also sells more than the corresponding competitive industry when demand is high.

Having established that the cartel needs fewer resources to be unconstrained, I now show that, for certain levels of firm resources, competitive firms shut down when demand is low. Afterwards, I provide numerical examples where the competitive outcome does indeed involve turnover of firms while the cartel does not and where, in addition, this leads the cartel to be superior from a welfare point of view.

**Lemma 1** Let $K_o^R$ be the capacity which, under perfect competition, makes the average price be $c + \frac{v + \lambda v'}{1 + \lambda}$. In other words, it is the value of $K_o^S$ which satisfies (9) when $v$ is replaced by $\frac{v + \lambda v'}{1 + \lambda}$. Then, as long as the owners’ individual endowment satisfies

$$v + c - \max(D^{-1}(K_o^R, \xi^\ell), c) \leq R < v + v' + 2c - \max(D^{-1}(K_o^R, \xi^\ell), c) - \max(D^{-1}(K_o^S, \xi^\ell), c) \tag{45}$$


competition leads first period capacity to equal $K^R_o$. If the first period state is $h$, this capacity is maintained. Otherwise, it is scrapped and $K^S_o$ is built for the second period.

**Proof:** Since the definition of $K^R_o$ implies that it exceeds $K^S_o$, the first inequality in (45) implies that $K^S_o$ is larger than $K^F_o(R)$ so that entrants in period 2 prevent capacity from being smaller than $K^S_o$. The second inequality in (45) implies that, if first period capacity is $K^R_o$, firms run the risk of losing $\bar{L}$ if they set $k_2^F$ equal to one because doing so leads (40) to be violated. Thus, if capacity is $K^R_o$, firms do not maintain their capacity when the state is $\ell$. The first inequality implies that if industry capacity is $K^R_o$ and the state in the first period is $\ell$ firms do not lose $\bar{L}$ as long as they set $k_2^F$ equal to zero.

Ignoring the $L$ term, firms that set $k_2^h$ equal $k_1$ while $k_2^\ell$ is zero expect to earn nonnegative profits by setting $k_1$ equal to one as long as industry capacity is no larger than $K^R_o$. This still leaves the question of whether a capacity of $K^R_o$ leads to a loss of $\bar{L}$ if the state is $\ell$ in the second period after having been $h$ in the first. Firms avoid this cost with this sequence of events as long as

$$R \geq v + c - \max(D^{-1}(K^R_o, \xi^\ell), c) + \{v' + c - D^{-1}(K^R_o, \xi^h)\}$$

The definition of $K^R_o$ implies that the term in curly brackets is negative so that the first inequality in (45) ensures that the inequality is satisfied. Thus, $K^R_o$ is the equilibrium level of capacity. Lower levels of capacity would lead to more entry because further entry is both feasible without loss of $\bar{L}$ and gives positive expected profits. Higher levels of capacity would lead to expected losses. QED

The basic intuition for the result that firms shut down and reopen when $R$ is in this range is that the willingness of two owners (the present and the future ones) to sacrifice their resources $R$ when the state of demand is low in both periods allows for lower prices than would prevail if only one owner was willing to sacrifice one $R$ in this way. The reason two owners are willing to carry out this sacrifice, of course, is that the financial market imperfection curtails entry to such an extent that prices are quite high in the high demand state. Another way of seeing this is from the perspective of an owner who builds capacity for
time one. The reason he sometimes fails to maintain it is because there are eager potential entrants who are ready to sacrifice $R$ if the state is low in period 1. It is this competition by cash-rich future firms that leads initial investors to abandon the industry even though they are more efficient than the entrants.

This interpretation makes it clear that firm exit is predicated on the wealth of future entrants. If this wealth is low, i.e. if $R$ is low enough, the willingness to sacrifice these resources does not suffice to allow the potential entrants to charge a lower price than the more efficient incumbent firms. Thus, if $R$ is low enough, there is no firm turnover. This can be seen in the following lemma.

**Lemma 2:** If

$$R < v + c - \max(D^{-1}(K_o^S, \xi^c), c) \quad \text{and} \quad R < v - v'$$

(46)

equilibrium capacity in period 1 is given by $\hat{K}_{o1}$ where

$$\hat{K}_{o1} = D\left(c + \frac{v + v' - R}{2}, \xi^c\right)$$

(47)

Moreover, all this capacity is maintained regardless of the state in period 1 while no new capacity is added by new firms in period 2.

**Proof:** The first inequality in (46) implies that $K_o^F(R)$ is lower than $K_o^S$ so that the former is the most that firms contribute to by building capacity in period 2. The definition of $\hat{K}_{o1}$ implies that, if $\hat{K}_{o1}$ is built in period one, all this capacity can be maintained while satisfying (40) as long as no further capacity is built in period 2. The second inequality in (46) ensures $\hat{K}_{o1}$ is larger than $K_o^F(R)$. As a result, firms in period 2 do not increase capacity beyond $\hat{K}_{o1}$ so that all of $\hat{K}_{o1}$ can be maintained if it is built in the first place. The combination of inequalities in (46) also guarantee that $\hat{K}_{o1}$ is smaller than $K_o^U$. Thus firms who build and maintain the minimum capacity when aggregate capacity is $\hat{K}_{o1}$ gain positive expected profits ignoring the terms in $L$ by doing so. Since they also fail to lose $\bar{L}$ when capacity is $\hat{K}_{o1}$, equilibrium capacity in period 1 can be no smaller. It also can be no bigger because this would lead to a loss of $\bar{L}$ whether this capacity is maintained or not. QED
To gain some further intuition for these results, Figure 1 draws the equilibrium levels of capacity for different levels of $R$ in the special case where $D(c, \xi\ell)$ is less than $K_o^S$. Since $K_o^R$ and $K_o^U$ are even larger than $K_o^S$, these capacity levels would also lead to a price of $c$ in the low state of demand.

When $R$ is greater than or equal to $v + v'$, equilibrium capacity is $K_o^U$ and this capacity is held for two periods. Because owners can afford to buy capacity for two periods with their endowment, the financial market imperfection is not binding and the price is $c$ in the low demand state. As soon as $R$ falls below $v + v'$, owners refrain from buying capacity for two periods unless the price in the low state is above $c$ (so that profits in the low state help pay for capacity). With $R$ between $v$ and $v + v'$, however, owners are willing to let the price in the low state be $c$ as long as they do not maintain their first period capacity. The result is that equilibrium capacity is unconstrained by the financial restrictions except for the fact that $k_2^\ell$ is zero. Capacity then equals $K_o^S$ in the second period if the state in the first was $\ell$ and equals $K_o^R$ otherwise.

For $R$ below $v$, owners cannot pay even for one period of capacity with their endowment so that the price must be above $c$ in the low demand state. There is thus another discontinuous fall in capacity at this level of $R$. For $R$ between $v - v'$ and $v$, firms still do not maintain their capacity if the state is $\ell$. The reason is that, when $R$ is in this range, the price that a sequence of firms needs to charge to avoid the cost $\bar{L}$ is lower than the price that a firm that stays for two periods needs to charge to avoid this cost. The former is $(c + v - R)$ while the latter is $c + \frac{v + v' - R}{2}$. Note that, when firms stay for two periods, they are willing to lose only $R/2$ in "variable profits" in each of the two periods.

Finally, for $R$ below $v - v'$ the equilibrium contains only firms that last for two periods because $c + \frac{v + v' - R}{2}$ is smaller than $(c + v - R)$. In this region, a one dollar reduction in $R$ has a smaller effect on capacity than when $R$ is above $v + v'$. The reason is that, in the former case, the loss of $\bar{L}$ is avoided if the equilibrium price in each period rises by half of the reduction in $R$ whereas it needs to increase by the full reduction in $R$ in the latter.

I now show numerically that, in this setting, the cartelization of an industry has the
potential for increasing consumer welfare. I consider an example in which both (42) and the conditions of Lemma 1 hold. In this example there exist levels of resources $R$ such that the cartel acts as if the terms in $L$ did not exist while the competitive allocation leads firms to shut down in state $\ell$. Since this represents a productive inefficiency, it is not surprising that the cartel then achieves a superior allocation of resources.

For this analysis, I suppose that demand is linear as in (30) so that expected utility is proportional to the sum of the expected square of output in the two periods. I let the probability of state $h$ be .5, while the marginal cost $c$ is zero. The parameters $a^h$, $a^\ell$, $b$, $v$ and $v'$ equal 8, 2, 1, 3 and 1 respectively. These parameters imply that $K^S_o$ equals 2.0 while $K^R_o$ equals 3.33. Whether capacity is equal to either of these, the price in state $\ell$ is equal to $c$ (which is zero), because 2 units are demanded at the price of zero. Thus both inequalities in (45) are satisfied as long as $R$ is between $v$ (which equals 3) and $(v+v')$ (which equals 4). So, with $R$ in this range, equilibrium competitive capacity in the first period is $K^R_o$. When the state is $\ell$, these firms do not maintain their capacity, so $K^S_o$ is built in the second period. Firm’s expected profits are zero in both periods while the sum of the expected squares of output is 13.3.

Since the competitive price is equal to $c$ when demand is low, (42) is satisfied and the cartel restricts output when demand is low. Its price in state $\ell$ is 1 (and its sales are equal to 1 as well). The resulting profits induce entry and $K^U_o$, the equilibrium capacity when the cartel is unrestricted equals 4.23. Thus, profits per unit of capacity in state $\ell$ equal 1/4.23 or .24. Using (43), this means that firms in the cartel can acquire this level of capacity without worrying about $\bar{L}$ as long as $R$ is no smaller than $(4-2\times .24)=3.52$. Thus, for $R$ between 3.52 and 4, the cartel equilibrium involves the unrestricted level of capacity 4.23 while competition involves a capacity of only 3.33 in the first period and of only 2 in the second period if demand is low in the first. Admittedly, cartel output when demand is high is only 4 (because the ex post monopoly price in state $h$ is 4) while it is only 1 when demand is low. Still, the sum of the expected square of output is 17 which is exceeds the competitive level. Thus, the cartel’s outcome is superior from a welfare point of view.
6 Conclusions

The result that cartels can be Pareto superior to perfect competition with imperfect financial markets raises several new questions. The first issue is whether the result is robust to alternate (and weaker) specifications of financial market imperfections. A full exploration of this issue requires further work. However, what appears to be necessary for the result is that managers (or owners) be especially averse to outcomes with low profits. Assuming that bankruptcy and reorganization are costly, this is an almost necessary consequence of the fixed-payment nature of debt obligations since the firm has more difficulty meeting these obligations when profits are low.

From a theoretical point of view, this leaves the question of why debt obligations take this form. Here I have taken the simplest approach to explaining debt coupled with costly bankruptcy, namely that it economizes on the knowledge that creditors need to acquire about the value of the firm. Alternative views are provided in Aghion and Bolton (1992) and Hart and Moore (1994) and it would be worth knowing whether the benefits of cartels I have explored extend to their settings. Aghion and Bolton's (1992) model of debt and bankruptcy emphasize the benefits of pre-specifying that creditors gain control over the firm's decisions when realized profits are low. In their model, this loss of control by the original owner in the states in which the firm declares bankruptcy is painful for the initial owner. Thus, cartels which smooth industry cash flow might also encourage efficient entry in a variant of their model.

The second, related, question is the extent to which financial market imperfections led actual cartels to be socially valuable. As is well known, equity markets were weak in Germany in the period of cartels and many of the most important German firms of the period were privately held. At the same time, the costs of bankruptcy were quite high. As Vorbrugg (1986, p. 74) says “Unlike other jurisdictions, German law does not provide for any possibility of discharge of debts” (Vorbrugg, 1986, p. 74). Finally, firms appear to have generally been unable to hedge the risk of demand fluctuations in futures markets (except
by paying substantial premia). As Liefmann (1932, p. 143) describes it, while forward sales contracts were desirable to the selling firm from a risk-reduction point of view, they required “price concessions to the purchasers, especially where the latter cannot count on a steady, continuous delivery to their customers.” Thus, my model mimics some of the institutional features of Germany. On the other, it is simply unclear whether German banks (which were important intermediaries) were prepared to impose large costs on borrowers who were unable to meet their obligations. While universal banks backed the signing of cartel contracts among their customers (Fear 1998), this may or may not be related to the issues that are important in this paper.

The third issue is the extent to which the analysis would be affected (and the desirability of cartels either increased or reduced) if one considered a more realistic model of the outcome that prevails when cartel contracts are unenforceable. In this paper I have kept the analysis simple by supposing that this alternative is perfect competition. However, John D. Rockefeller’s response to what he regarded as “ruinous competition” was to consolidate the oil industry (Chernow 1998, p. 130). More generally, United States firms ended up being larger than German ones in many industries (Fear 1998).

Understanding whether these mergers led to the sort of cash-flow stabilization I have demonstrated for cartels is a complicated question in its own right. Such mergers might well have facilitated implicit collusion. When marginal cost are constant, Rotemberg and Saloner (1986) show that implicitly collusive oligopolists who cannot sustain the monopoly outcome have prices that rise relative to marginal cost when demand rises and this certainly stabilizes cash flow. However, as Staiger and Wolak (1992) show, this property of implicitly collusive prices does not carry over neatly to a situation with the sort of capacity constraints that play a role in this paper. Thus the question of whether industry consolidation stabilizes cash flow as much (or more than) cartels seems to defy easy answers. Perhaps more importantly, the response of an industry to entry might be very different in the case of a cartelized industry than in the case of a consolidated industry in which cartels are illegal. If such differences did exist, they might weigh heavily in one’s assessment of the social desirability of cartels.
7 References


Selten, Reinhardt, “A Simple Model of Imperfect Competition where Four are Few and Six are Many,” *International Journal of Game Theory*, 2, 141-201.

