Entry deterrence and dynamic competition*

The role of capacity reconsidered

Jean-Pierre Benoît
New York University, New York, NY 10003, USA

Vijay Krishna
Harvard Business School, Boston, MA 02163, USA

Final version received January 1991

We study a model with sequential capacity choice and entry by firms into an industry. Post-entry competition is long term and firms compete by choosing prices. The complex role played by the capacity choice of the first mover is highlighzd. In contrast to the conclusions derived from static or reduced form specifications, entry may be deterred only by choosing a low capacity level and charging a very high price. The arguments are used to provide an explanation of events in the U.S. phosphorus industry.

1. Introduction

It is not difficult to argue that the nature of imperfect competition is better understood by examining firm behavior in a dynamic rather than a static setting. Dynamic specifications are not only preferable because of greater realism but also seem necessary for explaining such phenomena as implicit collusion and the persistence of cartels. While this is generally recognized, for reasons of both greater tractability and predictability much of industrial organization theory has worked with either static specifications or used dynamic models in which most of the dynamic play is subsumed in a reduced form 'black box'.

This paper seeks to argue that such simplifications may be substantially misleading when compared to a more complete dynamic structure. Specifically, we reexamine the nature of strategic barriers to entry in this light. The manner in which these barriers are erected and their viability has occupied

*We wish to thank Martin Osborne, John Pratt and a referee for helpful comments. Research support was provided by the National Science Foundation (Grant No. SES-8611211) and the Division of Research at the Harvard Business School.
the attention of economists, business strategists and policy makers for a long time. A major theme in this regard has concerned the role of commitment on the part of incumbent firms. It is well understood that by pre-emptively committing to an aggressive course of action, and making this commitment public, firms can steer equilibrium outcomes in their favor.

Spence (1977) suggested that irrevocable capacity choice may act as a means of commitment – by pre-emptively investing in a large capacity and threatening to utilize it fully, an incumbent firm might be able to successfully deter entry by other firms. Thus, in equilibrium, the incumbent firm would carry unutilized capacity – its role that of a sheathed but visible sword. Although the equilibrium examined by Spence was not ‘perfect’, Dixit (1980) showed that even when the requirement of subgame perfection is imposed, capacity choice can be strategically used to deter entry. Under Dixit’s assumptions idle capacity is never observed in equilibrium. The threat of entry, however, may lead the incumbent to choose a capacity greater than it would choose had there been no potential entrant. In an extension of Dixit’s quantity setting model, Ware (1984) allowed the entrant to commit to a capacity also, pointing out that entry deterrence resulted from a first mover advantage alone and not from any other asymmetries. In another variant of Dixit’s model Bulow, Geanakoplos and Klemperer (1985) showed that idle capacity could indeed be found in a perfect equilibrium. Allen (1986) obtained the same conclusion in a price setting model.

While these models are basically static in nature, Spence (1979) and Fudenberg and Tirole (1983) incorporate dynamic elements into their analyses. They adopt the so called ‘state variable approach’ wherein the detailed aspects of post-entry competition are ‘black-boxed’ by making assumptions on the instantaneous profits resulting from capacity choices. Intuition about these assumptions – the inner workings of the black box – is derived from static models.

The reasoning behind these models is clear – increased capacity enables the incumbent firm to threaten the would-be entrant with lower profits in the post-entry equilibrium, thus discouraging entry. However, when one considers a more fully specified dynamic model the story changes considerably. Such a model typically allows for a richer set of equilibrium outcomes and for some of these equilibria, the assumptions made in the state variable approach are violated in essential ways. The thrust of our argument is that commitments that make predatory behavior in the post-entry game credible also increase the prospects for collusion. This is because in a dynamic setting, a greater degree of collusion may be supported by the increased severity of available threats. The entrant may then view the incumbent’s choices as a commitment to collude. When these considerations are made explicit, novel conclusions may emerge.

We begin by outlining our basic framework in section 2. We study price
setting models where firms are constrained to supply no more than installed capacities. In section 3 we examine a model where capacity choices are made sequentially by the incumbent and entrant respectively before a static one-shot price game ensues. We are able to re-derive Allen's (1986) results very quickly and show that for a range of parameter values, Spence's (1977) story can be verified. The conclusions of section 3 serve as a benchmark against which those of section 4 may be compared. In section 4, we extend the static model by allowing the post-entry game to be of infinite duration. Because of the possibility of multiple equilibria in the post-entry game and the fact that collusive duopoly profits are a non-monotonic function of capacities, equilibria in the entry deterrence game may take paradoxical forms. We demonstrate that entry may be 'best' deterred by installing little capacity—surprisingly, sometimes even smaller than monopoly capacity. Section 5 illustrates how the analysis may be extended to take account of non-stationary equilibria. In section 6 we use the arguments of the previous sections to analyze a 'case study' of the U.S. phosphorus industry in the period 1970–89. Section 7 contains some methodological and concluding remarks.

2. Preliminaries

We adopt the basic framework of Kreps and Scheinkman (1983) and Osborne and Pitchik (1986). Consider a market for a homogeneous product in which demand is given by the market demand function $D: R_+ \rightarrow R_+$. The demand function is assumed to be invertible and its inverse is written as $D^{-1} = p: R_+ \rightarrow R_+$. 

There are two identical firms. A firm with capacity $x$ can produce up to $x$ units of the product at zero marginal cost. The cost of installing capacity $x > 0$ is given by $F + b(x)$, where $F$ represents the fixed costs of operation and $b$ is an increasing function. For given capacity levels $x_1$ and $x_2$, firms play the capacity constrained price game $g(x_1, x_2)$ as follows. Firm $i$ chooses a price $p_i$ from the interval $[0, p(O)]$.

If $p_1 < p_2$, firm 1 sells up to its capacity or demand. Firm 2 then faces the resulting residual demand function. Formally, firm 1's sales are given by

$$z_1(p_1, p_2; x_1, x_2) = \min(x_1, D(p_1)) \tag{1}$$

and firm 2's sales are

$$z_2(p_1, p_2; x_1, x_2) = \min(x_2, \max(0, D(p_2) - z_1)). \tag{2}$$

If $p_1 = p_2$, firms share the market equally unless one of the firms has
insufficient capacity to meet one-half of the market demand. In that case the other firm sells to meet the rest of what is demanded up to its capacity.

Notice the 'rationing rule' implicit in (2). Suppose that there are a continuum of consumers each of whom wishes to purchase at most one unit of the good. Let $D(p)$ be the number of consumers whose reservation price is at least $p$. With this interpretation, (2) implies that consumers are rationed in descending order of reservation prices, that is, those with high reservation prices purchase from the firm with the lower price first. While other rationing rules are possible, (2) is salient in that it is the worst from the point of view of the firm that charges the higher price. This will have important implications later.

Throughout we make the following assumptions about demand.

**Assumption 1.** The inverse demand function $P: R_+ \rightarrow R_+$ is strictly positive on some bounded interval $(0, X)$ and is continuous and strictly decreasing on $[0, X]$. For $x \leq X$, $P(x) = 0$.

**Assumption 2.** The revenue function $R$ defined by $R(p) = pD(p)$ is single peaked, that is, there is a price $p^* > 0$ such that for all $p$ in $(0, p^*)$, $R$ is strictly increasing and for all $p$ in $(p^*, P(0))$, $R$ is strictly decreasing. ($p^*$ is the revenue maximizing price.)

Osborne and Pitchik (1986) require only that Assumption 1 be satisfied. Although our assumptions are somewhat stronger, Assumption 2 is much weaker than assuming the concavity of $R$ (or $P$) as in Kreps and Scheinkman (1983). Assumption 2 rules out the occurrence of certain degeneracies in the structure of equilibria.

In the game $g(x_1, x_2)$, let $v_i(x_1, x_2)$ denote firm $i$'s minmax revenue. $v_i(x_1, x_2)$ is the least amount firm $j \neq i$ can hold firm $i$'s revenues down to. It is easy to see that this occurs when firm $j$ charges a price of zero. Thus,

$$v_i(x_1, x_2) = \min_{p_1} \max_{p_2} p_1 z_i(p_1, p_2; x_1, x_2)$$

$$= \max_{p_1} p_1 z_i(p_1, 0; x_1, x_2)$$

and symmetrically for firm 2. Also, note that $x'_1 \geq x_1$ implies that $v_i(x'_1, x_2) \geq v_i(x_1, x_2)$ and that $x'_2 \geq x_2$ implies that $v_i(x_1, x'_2) \leq v_i(x_1, x_2)$.

3. **Excess capacity as an entry deterrent**

In this section we analyze a model with sequential capacity decisions
in which the post-entry game is a static one-shot affair. Models of this variety were first studied by Spence (1977) and then by Dixit (1980) and Bulow, Geanakoplos and Klemperer (1985) and others. These papers all concern quantity setting models. Following Kreps and Spence (1983) and others, we analyze a model in which firms make both quantity (capacity) decisions and price decisions. In this section, our model is the same as that of Allen (1986).

Consider a three stage game, $g^*$, where firms 1 and 2 choose capacity levels $x_1$ and $x_2$ in stages I and II, respectively. A capacity choice of zero is equivalent to a decision not to enter the market. In stage III, the firms compete by choosing prices taking the installed capacity levels as fixed – they play $g(x_1, x_2)$. Under the assumption of linear demand, Allen (1986) has shown that the equilibrium of this game may involve firm 1 picking a capacity larger than monopoly output, and firm 2 choosing not to enter. Using the following result from Benoit and Krishna (1987) we can quickly generalize this. Let $e_J(x_1, x_2)$ be the unique equilibrium revenues in $g(x_1, x_2)$.

**Proposition 1.** [Benoit and Krishna (1987)]: If $x_1 \geq x_1^*$ then for all $x_2$, $e_2(x_1, x_2) \leq e_2(x_1^*, x_2)$.

Thus, in $g(x_1, x_2)$, firm 2's equilibrium revenues decline as firm 1's capacity increases.

Since in any perfect equilibrium of $g^*$ firms must always play the unique equilibrium following any choice of capacities, firm 1 deters entry 'more effectively' the larger its capacity. That is, if $x_1' > x_1$, then entry will be deterred for a larger range of fixed costs, $F$, if firm 1 picks a capacity of $x_1'$ rather than $x_1$. If the variable component of capacity costs, $b(x)$, is zero for all $x$, then it is a dominant strategy for firm 1 to choose $x_1 = D(0)$.

Thus, once again we have confirmation for the by now familiar message: excess capacity deters entry. Idle capacity is not an idle threat.

4. Some dynamic considerations

In the model outlined in section 3, it was assumed that the post-entry game, $g(x_1, x_2)$, is a one-shot affair. This fact, together with an appropriate specification of market demand and technology, ensures that the post-entry game has a unique equilibrium outcome and hence that the overall game, $g^*$, can be unambiguously solved by using backwards induction.

In this section we modify the model of section 3. Suppose that firms pick capacities sequentially in stages I and II as before. However, stage III consists of an infinite horizon subgame, $G(x_1, x_3)$, which is a repeated version
of \( g(x_1, x_2) \). Firms evaluate the resulting infinite stream of revenues by using the discount factor \( \delta \). Call this game \( G^* \).

As we argued earlier, such a specification of the post-entry game is attractive as it seems to capture certain essential features of oligopolistic interaction. Competition is a long-term affair and the possibilities for intertemporal tradeoffs and implicit collusion are naturally admitted in a dynamic model.

It is well known that an infinitely repeated game like \( G(x_1, x_2) \) has numerous equilibria, especially if discount rates are low. Thus, in contrast to \( g^* \), it is not possible to solve for a unique equilibrium of \( G^* \) using backwards induction.

Consider the point of view of a firm contemplating entry in a market with an incumbent firm who has an installed capacity of \( x_1 \). For any choice of \( x_2 \), the resulting subgame \( G(x_1, x_2) \) may have many equilibria, some more profitable for the entrant than others. For some choices of \( x_1 \), it may happen that for all \( x_2 \) the equilibrium revenues in the worst equilibrium of \( G(x_1, x_2) \) for firm 2 are not enough to cover the fixed costs of entry. At the same time, revenues from 'good' equilibria may make entry worthwhile at some level \( x_2 \). Thus the question of entry deterrence is clouded by the issue of which equilibrium will be selected in the post-entry game. Informally, the answer may be said to hinge on what 'expectations' the potential entrant has about the nature of the post-entry situation. If these are 'optimistic' it would enter, otherwise not. Notice that the perfection of the equilibria is not an issue. Both outcomes, entry deterrence and entry, can result from imperfect equilibria and moreover for the same choice of the incumbent's capacity, \( x_1 \). The question of whether a choice of \( x_1 \) deters entry or not cannot be unambiguously answered.

There is a further twist. If the stage III game were static we showed that excess capacity held by an incumbent firm would act as a deterrent. Now in the dynamic game, consider the following scenario: an incumbent firm picks a capacity \( x_1 \) which greatly exceeds monopoly capacity. On the one hand, this greatly improves the incumbent's ability to punish the entrant by ensuring that for any choice of \( x_2 \), the revenues in the equilibrium of \( G(x_1, x_2) \) which is worst from its point of view are extremely low. However, the very fact that the worst equilibrium payoff is lowered, may increase the possibilities for collusion among firms. One can think of a collusive equilibrium as one where firms forego short-term profits because of the threat of retaliatory punishments. The higher the price being charged in equilibrium, the greater the short-term gain from undercutting one's rival. In order to sustain such high prices, a severe threat is necessary against any deviators. In a perfect equilibrium, worst equilibrium payoffs are the most severe credible threats, so the lower these are, the greater the possibility of collusion.
Thus, the same capacity which deters entry by a pessimistic firm may encourage entry by an optimistic firm. Given that the entering firm has already sunk its entry costs, it seems no worse to assume that the firms in the industry will now collude, rather than fight each other. Hence, we are led to the conclusion that a large capacity may actually foster entry! This immediately suggests that to deter entry, an incumbent firm might choose a capacity exceeding the monopoly level but smaller than it would if the post-entry game were one-shot. We will see by means of an example that a more extreme conclusion may emerge: entry deterrence may in fact involve cutting back on monopoly capacity rather than expanding.

Some organizing principles may be useful at this stage. We will say that a capacity choice of $x_1$ on the part of the incumbent firm weakly deters entry if for all choices of $x_2$, the revenues from the worst equilibrium for 2 in $G(x_1, x_2)$ cannot cover the fixed costs of entering the industry. We say that $x_1$ strongly deters entry if for all choices of $x_2$, no equilibrium of $G(x_1, x_2)$ can generate enough revenues to cover fixed costs. Clearly, to see if a particular choice of $x_1$ strongly deters entry it is sufficient to check that for all $x_2$, the best equilibrium of $G(x_1, x_2)$ from 2's point of view is not profitable enough. We turn to an examination of such equilibria. While determining the 'best' and 'worst' equilibria of $G(x_1, x_2)$ is a tricky task, our main result yields a substantial simplification of the problem. For any $x_1, x_2$ pair let $M_i(x_1, x_2)$ denote the maximum per period revenues that firm $i$ can obtain in a stationary perfect equilibrium of $G(x_1, x_2)$. For simplicity, we consider only stationary equilibria in this section. In addition, we assume that firms do not randomize except to play the one-shot equilibrium. Clearly, for all $x_1, x_2$, $M_i(x_1, x_2) \geq e_i(x_1, x_2)$, the revenues obtainable in the equilibrium of the one-shot game $g(x_1, x_2)$. This is because repeated play of this equilibrium would result in an equilibrium of $G(x_1, x_2)$. The result stated below says that if the entrant hopes to exceed the one-shot revenues $e_2(x_1, x_2)$ in $G(x_1, x_2)$, it is best for it to pick $x_2 = x_1$. Thus, in effect, if the two firms are in fact going to implicitly collude once entry has taken place, the entrant does best by picking the same capacity level as the incumbent. As will become apparent later, the resulting symmetry is extremely helpful for the purposes of analysis.

**Proposition 2.** Let $M_2(x_1, x_2)$ be firm 2's maximum revenues in a stationary subgame perfect equilibrium of $G(x_1, x_2)$. Then,

$$\max_{x_2} M_2(x_1, x_2) = \max \{ M_2(x_1, x_1), \max_{x_2} e_2(x_1, x_2) \}.$$

The proof of Proposition 2 has been relegated to an appendix but the basic underlying idea is that for any $x_2$ such that neither firm is capacity
constrained, the set of possible equilibrium prices of $G(x_1, x_1)$ contains the set of possible equilibrium prices of $G(x_1, x_2)$. This is because in the unconstrained case, any profitable way to deviate in $G(x_1, x_1)$ can be mimicked in $G(x_1, x_2)$. The formal proof provides details of this argument and also deals with cases where one or the other firm is capacity constrained.

As a result of Proposition 2, we are able to study the notion of strong deterrence in more detail: it is only necessary to examine the most profitable equilibria of symmetric subgames $G(x_1, x_1)$ and equilibria of the one-shot game $g(x_1, x_2)$ to check whether a particular capacity $x_1$ strongly deters entry. We illustrate the use of Proposition 2 by means of an example.

In fig. 1, we have shown, for each value of $x$, the function $M(x, x) = M_i(x, x)$ which gives the highest revenues obtainable in a stationary perfect equilibrium of $G(x, x)$. This is computed for the linear demand function $D(p) = 120 - p$ and the discount factor $\delta = 0.61$. For each value of $x$ the maximum sustainable revenues are computed by finding the highest price from which deviations can be prevented by the threat of the worst perfect equilibrium. In symmetric subgames $G(x, x)$ the worst perfect equilibrium path consists of repeated play of the equilibrium of the one-shot game $g(x, x)$. This is because $e_i(x, x) = v_i(x, x)$, and the payoffs in no equilibrium can be

---

Fig. 1. $M(x, x)$.
lower than the latter. [See Brock and Scheinkman (1985) for a similar analysis].

One might expect that as capacities increase, higher prices and revenues would be sustainable. However, as $x$ increases, two forces come into play. On the one hand, threats become more severe. On the other, the gains from deviating also increase as more capacity is available. These two forces balance to produce the highest sustainable price and the corresponding profits. Notice in fig. 1, the non-monotonic nature of $M$ and the fact that it may be discontinuous.

In fig. 2 we have displayed a graph of the function $\max_y M_2(x_1, y)$. As a result of Proposition 2, this is the maximum of (a) $M(x_1, x_1)$; and (b) $\max_y e_2(x_1, y)$, the highest one-shot equilibrium payoffs.$^2$ For small values of $x_1$, the latter exceeds $M(x_1, x_1)$ as collusive profits are low and firm 2 does better by picking a large capacity and receiving the corresponding one-shot equilibrium revenues. For larger capacities the opposite is true. (Compare figs. 1 and 2). To see if strong deterrence is possible, it is sufficient to compare the range of this function against the fixed costs of entry.

---

$^2$Explicit expressions for one-shot equilibrium payoffs for the case of linear demand may be found in earlier versions of both Kreps and Scheinkman (1983) and Osborne and Pitchik (1986). Published versions of both papers make weaker assumptions on demand.
Suppose that there are no variable costs of capacity (for all x, b(x)=0). First, if the fixed costs of entry per period, F, are lower than A (see the right axis of fig. 2), strong deterrence is not possible. For every x₁, there is at least one stationary equilibrium of G(x₁, x₁) in which revenues exceed F. If fixed costs exceed C, entry is strongly deterred by a choice of x₁ = c, which is the monopoly output level. In intermediate ranges strong deterrence is possible but requires firm 1 to choose its capacity strategically. If F is between B and C, strong deterrence requires firm 1 to choose a capacity in excess of c. In this case, a story not unlike that of section 3 emerges as excess capacity plays a significant role. However, firm 1 cannot pick too large a capacity since any capacity exceeding d would lead to entry. In contrast, recall that in section 3 it was a dominant strategy to pick x₁ = D(0) when the variable costs of capacity were zero. The final case, when F is between A and B, is also the most interesting. Let us examine this in more detail.

For concreteness, let us suppose that the fixed costs of entry are 1770 per period. (This is between A and B). First, it can be verified that if firm 1 chooses to install capacity x₁ equal to the monopoly output level xₘ=60, firm 2 is weakly deterred from entering. This is because for all x₂, firm 2 can receive no more than 1488 in the equilibrium of the one-shot game g(xₘ, x₂). Since, for all x₂, e₂(xₘ, x₂) < F, the threat of repeated play of the equilibrium of g(xₘ, x₂), no matter what capacity x₂ the entrant chooses, is enough to keep him out of the market. Notice, however, that choosing x₁ = xₘ is not enough to strongly deter entry since if firm 2 were to choose x₂ = xₘ also, each firm would make 1779 in the most profitable equilibrium of G(xₘ, xₘ). Thus, it would be in 2's interest to enter the market if the firms were to 'collude' in the post-entry game. In fact, inspection of fig. 2 shows that the only way for firm 1 to strongly deter entry is for it to choose a capacity less than b, that is, a capacity smaller than monopoly capacity. For F = 1770, firm 1 does best by choosing a capacity x₁ = 43, the largest capacity less than b which yields no more than F. But now notice that in equilibrium, since 1 is the only operating firm and has an installed capacity x₁ < xₘ, it does best by fully utilizing this capacity and charging a price of P(x₁) > P(xₘ). Furthermore, firm 1's revenues in this equilibrium are greater than if it allowed entry. Thus, in our example, deterring an 'optimistic' entrant involves picking a capacity smaller then the monopoly output level and the equilibrium involves charging a price that is higher than the monopoly price!

Some points are worth noting. First, there are many perfect equilibria of the three stage game G* in which entry is deterred. Most of these require pessimistic beliefs about the equilibria of the post entry game. The one we have calculated is salient in that it exhibits the phenomenon of strong deterrence: the entrant stays out no matter what its beliefs about the post-entry game. Second, note that in our example the only costs are the fixed costs of entry. If capacity were costly, we might expect firm 1 to pick a
capacity lower than it might otherwise (it may be too expensive to deter entry). In our example, since the marginal cost of additional capacity is zero, a small capacity is chosen for purely strategic reasons.

The discount rate in our example was deliberately chosen to exhibit the widest span of possible entry deterring behavior. In general, when discounting is not too severe, collusion is easier and the only way to deter entry strongly may be to choose low capacities (region BC may be empty). In those circumstances the message is even stronger than in the example presented.

We have assumed that the choice of capacity is irrevocable. Suppose instead that firms can invest in additional capacity up to a maximum amount I > 0 per period [as in Benoît and Krishna (1987)]. Call the subgame thus resulting after an initial capacity choice of \( x_1 \) and \( x_2 \), \( \Gamma(x_1, x_2) \). The essential features of the example discussed above survive as long as \( I \) is small. This follows from the fact that minmax payoff levels in the game \( \Gamma(x_1, x_2) \) continuously approach from below those in \( G(x_1, x_2) \). Recall that in our example entry is strongly deterred by a choice of \( x_1 < x^m \) because punishments in the game \( G(x_1, x_1) \) (which yield minmax levels exactly) are unable to prevent deviations from price paths profitable enough to cover the fixed costs of entry. When \( I \) is small, worst equilibrium payoffs in \( \Gamma(x_1, x_1) \) cannot be much lower than in \( G(x_1, x_1) \). Hence it is still the case that no profitable path is an equilibrium of \( \Gamma(x_1, x_1) \).

On the other hand for all \( x_1 > x^m \), any equilibrium price path of \( G(x_1, x_1) \) can still be sustained in \( \Gamma(x_1, x_1) \) with no additional investment. Consider the threat to meet any price deviation or unplanned investment by playing the corresponding one-shot equilibrium forever. To see that this threat is sufficient, first notice that for \( x_1 > x^m \),

\[
e_1(x_1 + I, x_1) = v_1(x_1 + I, x_1) = v_1(x_1, x_1) = e_1(x_1, x_1),
\] 

and similarly for firm 2. The first and third equalities are general properties of the function \( e_i \) and the second is true for values of \( x_1 \) greater than Cournot equilibrium levels. Eq. (4) implies that a threat by firm 2 to play the one-shot equilibrium no matter what the capacity levels yields the same payoffs in \( \Gamma(x_1, x_1) \) as it did in \( G(x_1, x_1) \) and hence is still a sufficient deterrent.

We have shown that the example is robust to a change in specification of the model that allows firms to increase capacity by small amounts in subsequent periods.

In light of the static model and the resulting conventional wisdom in the industrial organization literature, the example appears almost paradoxical: entry is deterred by choosing a small capacity level and as a result, a very high price emerges in equilibrium. In the static model excess capacity is
viewed by the entrant as a commitment on the part of the incumbent to predatory behavior if he enters. In a dynamic model a large capacity may be viewed by the entrant as a commitment to implicitly collude or, more to the point, a small capacity as a commitment not to collude.

5. Non-stationary paths

The analysis of the previous section made the simplifying assumption that equilibrium price paths were stationary. This was done in order to facilitate explicit calculation of the maximal collusive profit levels in the post-entry game (Proposition 2). In this section, we briefly argue that the basic message of the paper is quite general and survives a relaxation of this assumption.

For instance, it can be shown that a capacity choice of $x_1^* < D(0)$ by the incumbent may be preferable to a choice of $x_1 = D(0)$ even if non-stationary paths are considered.

First note that if both firms have a capacity of $D(0)$, the worst equilibrium payoff for both firms is zero. On the other hand, if firm 1 chooses $x_1^* = D(0) - \varepsilon$, then for all $x_2$, the worst equilibrium path (stationary or not) yields positive payoffs to both firms. The reason is that firm 2's minmax payoff in $G(x_1^*, x_2)$, is positive and hence any equilibrium involves firm 2 charging a positive price in some period. Since firm 1 can undercut firm 2 in such a period, its equilibrium payoff must also be positive. As a result it may be that 2's payoff in the best equilibrium in $G(x_1^*, x_2)$ is less than the best equilibrium payoff in $G(D(0), D(0))$.

To see that this is possible consider the case when $\delta = 0.5$. In $G(D(0), D(0))$ firm 2 can get a discounted sum of $\pi^m = \frac{p^m D(p^m)}{1 - \delta}$ along a path where both firms charge the monopoly price $p^m$ in every period. Any deviations are punished by going to the equilibrium that yields zero. However, we claim that for all $x_2$, a total of $\pi^m$ cannot be obtained in any equilibrium (stationary or not) of $G(x_1^*, x_2)$.

We argue by contradiction. Suppose there is an equilibrium of $G(x_1^*, x_2)$ in which firm 2's total payoff is $\pi^m$ or greater. Let $\pi_1^T$ denote firm 1's equilibrium payoff in period $t$ and let $\pi_1^T = \sum_{t=1}^{\infty} \delta^{t-1} \pi_1^T$ be the discounted sum of 1's payoff from period $T$ onwards. Suppose,

$$\pi_2^T + 0.5 \pi^*_2 \geq \pi^m. \quad (5)$$

If $\pi_2^T \geq \pi^m/2$, we claim that 1 can profitably deviate in period 1 and not be punished enough. Let 1's worst equilibrium payoff in $G(x_1^*, x_2)$ be $w_1 > 0$. By undercutting firm 2 in period 1, firm 1 can gain $\pi_1^T$ since for small $\varepsilon$ it has sufficient capacity to capture the whole market. In order for 1 to be deterred, it is required that
\[ \pi_2^1 \leq 0.5(\Pi_2^1 - w_1). \]  

But since \( \Pi_2^1 \leq 2\pi^m - \Pi_2^2 \) (recall that if \( \delta = 0.5 \) the discounted sum of industry profits is \( 2\pi^m \) at most), (6) implies that

\[ (\pi_2^1 + 0.5\Pi_2^2) - \pi^m + 0.5w_1 \leq 0. \]  

But (7) is contradicted by (5) since \( w_1 > 0 \).

On the other hand, if \( \Pi_2^2 < \pi^m/2 \), then (5) implies that \( \Pi_2^2 \geq \pi^m \) and that \( \pi_2^1 \geq \pi^m/2 \). The same reasoning as above now implies that firm 1 can deviate in period 2 and not be punished.

Hence no equilibrium of \( G(x_1, x_2) \) can yield \( \pi^m \) to firm 2 if \( \delta = 0.5 \). Thus for some parameter values (for instance if the per period entry costs \( F = \pi^m/2 \)) a capacity choice of \( x_1' < D(0) \) will strongly deter entry whereas \( x_1 = D(0) \) will not.

A consideration of non-stationary paths does not permit explicit calculations exhibiting the extreme anomalies of the previous section. However, the point that any action which harms a firm by lowering payoffs in some equilibrium, may by the same token, help the firm by raising payoffs in other equilibria is quite general, and follows simply from the nature of dynamic games. In the entry deterrence game, the same threats that deter entry in a static setting may invite entry in a dynamic setting – the capacity sword has two edges!

6. Some empirical evidence

We have drawn attention to the theoretical possibility of some counter-intuitive aspects of capacity choice when competition is long-term. In particular, we have pointed out that in a dynamic setting, the Spence–Dixit theory may be seriously questioned. Here we present a brief ‘case study’ of the phosphorus industry in the United States (in the period 1970–1989) which provides a neat example of the phenomena we have discussed earlier. Both the structure of the industry and its history fit our model rather well. In this industry, excess capacity does not seem to have deterred entry – indeed, the evidence suggests that it may even have acted as an invitation.

Phosphate rock is mined and melted down by an electro-thermal process to yield pure phosphorus. This is then used to manufacture various phosphorus compounds. Fixed costs of plant and equipment are high, exceeding \$100 million. The major component of variable costs is the cost of energy. The biggest single use of phosphorus chemicals is in synthetic detergent powders, usually in the form of sodium tripolyphosphate (STPP). Six firms account for almost all of the output of phosphorus and its primary derivatives. These are FMC (with approximately 30% market share), Mon-
santo (25%), Stauffer (15%), Occidental (10%), Olin (10%) and Albright & Wilson (5%). The remainder is accounted for by imports.

Since the early 1970s environmental regulations in the U.S. have led to a substantial decrease in the use of phosphates in detergents. Somewhat surprisingly, the drop in demand has not led to commensurate reductions in capacity. Exit costs in the industry, mainly environmental clean-up costs, are rather high.\(^2\) As depicted in fig. 3, this has led to substantial excess capacity in the phosphorus industry since the early 1970s. Actually fig. 3 overstates capacity utilization since about 20% of the production is devoted to low margin agricultural products, mainly fertilizers. Phosphorus compounds used in agriculture need not be free of impurities and can be manufactured by using less expensive manufacturing techniques. In particular, it is not necessary to obtain pure phosphorus via the thermal process to manufacture these compounds. The manufacture of other high value chemicals (including STPP), on the other hand, does require that no impurities be present. Since the 1970s U.S. firms have used pure phosphorus for the manufacture of agricultural products, selling these at close to marginal cost. For instance, in 1985, actual capacity utilization in the industry was about 65% though only about 50% was accounted for by high value products. A large portion of the capacity devoted to other products could easily have been diverted to more

\(^2\)Clean up costs at a plant closed by Occidental were estimated to be over $50 million.
J.-P. Benoif and K. Krishna, Entry deterrence and dynamic competition

![Phosphorus and STPP Prices](image)

**Fig. 4.** Phosphorus and STPP prices.

Lucrative products like STPP at low cost. Thus the capacity used for low margin products is effectively idle in economic terms.

Despite the fact that all firms have carried substantial excess capacity in this period, there has been very little in the way of price erosion. (See fig. 4 which depicts inflation adjusted prices for both phosphorus and STPP). Indeed it seems that the excess capacity has acted as an effective deterrent against price cutting and the industry has pretty much followed the largest firm, FMC, in setting prices. FMC has also carried greater amounts of idle capacity than other firms in the industry.

In 1988 Texasgulf and Albright and Wilson (A&W), a subsidiary of Tenneco, decided to invest in a completely new facility in a joint venture to make high grade phosphorus chemicals. The plant opened in early 1990, used a new, low-cost technique and was estimated to have capacity equivalent to 50,000 tons of phosphorus – increasing industry capacity by over 10%. Before this, A&W had been a minor player in the North American market with facilities in Canada whose output was mainly intended for export to the U.K. Furthermore, A&W’s Canadian operations suffered from a cost disadvantage relative to the other firms in the industry. Texasgulf is a major manufacturer of fertilizers.

*Data are obtained from the Current Industrial Reports published by the U.S. Bureau of the Census. Prices are deflated using the producer price index for industrial chemicals.  
\footnote{Chemical and Engineering News, March 12, 1990.}
It seems that the excess capacity carried in the industry did not act as a deterrent to new entry in the form of the Texasgulf and A&W joint venture. Indeed, given the history of the industry, it was very likely that the new entry (and the additional capacity) would be accommodated with little decrease in prices. In fact, despite projections that the new plant would have lower costs than the industry as a whole, only such an assumption on the part of Texasgulf and A&W would have made the investment seem worthwhile.

In a similar vein, Rosenbaum (1986) reports that the number of producers in the U.S. primary aluminum ingot industry grew from three to twelve in the period 1955–1981 while the industry carried substantial excess capacity. The data await a more careful analysis.

7. Concluding remarks

The fact that dynamic models yield a rich set of equilibrium outcomes is certainly well known. That this fact may have serious implications for some conventional industrial organization questions, however, does not seem to have been widely appreciated. In this paper, we have attempted to explore some of these implications for the theory of strategic entry barriers and shown how these may provide the best explanation of events in some industries.

In what circumstances might one expect these effects to be important? When can a potential entrant reasonably read the presence of excess capacity as a ‘signal’ that it will be accommodated? The analysis suggests that two factors will play an important role in determining the likelihood of such behavior. First, if there is a history of cooperation in the industry it makes it more likely that the new entrant can also become part of the cooperative arrangement. Some facilitating features are a stable (or predictable) demand function and easily monitored prices. Of course, the presence of an industry leader assisting price coordination and enforcing the arrangement helps. Second, high exit barriers also make accommodation more likely. Both of these are key characteristics of the phosphorus industry.

While we have argued in the context of the theory of strategic entry deterrence, we believe that the general lessons have wider applicability.

First, dynamic or repeated game models can lead to conclusions very different from those derived from static or reduced form analysis. The differences, sometimes dramatic, arise mainly because of the more complex

Although the joint venture was to have been between Texasgulf and A&W only, by the time the plant came on line they had been joined by Olin in a three-way partnership. The exact terms of Olin’s participation were still to be determined but the fact that a current player in the industry joined with the new entrant(s) seems to be an extreme case of accommodation.

Avinash Dixit refers to this as the ‘topsy turvy’ principle.
role played in dynamic situations by the means of commitment, in this case, irrevocable capacity choice.

Second, and in our view more important, while dynamic models typically suffer from a lack of predictability as they have numerous equilibria, in specific contexts it is possible to study properties shared by all equilibria. For instance, strong entry deterrence may be possible for some parameter values, as in the numerical example of section 4. In that case one may be able to draw strong inferences about equilibrium behavior without isolating a single equilibrium. We feel that this route is more satisfactory than narrowing the equilibrium set by means of ad hoc restrictions in order to derive strong predictions. The key feature that allows this is the relatively inflexible choice of capacity. In other contexts this might concern the choice of technology or product quality.

Appendix

In this appendix we provide a formal proof of Proposition 2 characterizing the largest equilibrium payoffs possible for the entrant.

**Proposition 2.** Let $M_2(x_1, x_2)$ be firm 2's maximum revenues in a stationary subgame perfect equilibrium of $G(x_1, x_2)$. Then,

$$\max_{x_2} M_2(x_1, x_2) = \max \left\{ M_2(x_1, x_1), \max_{x_2} e_2(x_1, x_2) \right\}$$

**Proof.** Let $x_2 \in \arg\max_y M_2(x_1, y)$.

1. Suppose, first that in the price path of $G(x_1, x_2)$ which yields $M_2(x_1, x_2)$, $p_1 = p_2 = p$.

1a. If firm $i$ is capacity constrained along the path, then $j \neq i$ must be getting its minmax revenues. Thus, $j$ must be playing a best-response and hence must also be capacity constrained. Otherwise, $j$ could do better by undercutting $p$. Thus, both firms must be capacity constrained and therefore playing an equilibrium of $g(x_1, x_2)$. This, of course, yields $e_2(x_1, x_2)$ to firm 2.

1b. If neither firm is capacity constrained, then both must be earning the same revenues. Furthermore, these revenues must be the same as they would have earned had they charged $p$ in $G(x_1, x_1)$. If $M_2(x_1, x_2) > M_2(x_1, x_1)$ then charging a price of $p$ could not have been an equilibrium path of $G(x_1, x_1)$. Therefore, both firms could profitably deviate in $G(x_1, x_1)$ from a price of $p$. Let $i$ be the firm with the larger capacity among $x_1$ and $x_2$. In $G(x_1, x_2)$, $i$'s deviations are at least as profitable as they were in $G(x_1, x_1)$. Furthermore, $i$
cannot be punished in \(G(x_1, x_2)\) any more severely than it can be in \(G(x_1, x_1)\). This is because \(v_i(x_1, x_2) \geq v_i(x_1, x_1)\) and since \(v_i(x_1, x_1) = e_i(x_1, x_1)\), \(i\) could be punished to minmax levels in \(G(x_1, x_1)\). Thus, if a profitable deviation were possible in \(G(x_1, x_1)\), \(i\) would find it profitable to deviate in the same manner in \(G(x_1, x_2)\). But \(p\) was an equilibrium path of \(G(x_1, x_2)\) and hence, \(M_2(x_1, x_2) \leq M_2(x_1, x_1)\).

Next, suppose that in the price path of \(G(x_1, x_2)\) that yields \(M_2(x_1, x_2)\), \(p_1 \neq p_2\). Clearly both \(p_1\) and \(p_2\) are less than or equal to \(p^m\).

(IIa) If \(p_2 > p_1\), firm 1 is selling up to its capacity. Thus, for this to be an equilibrium, firm 2's revenues must exactly equal \(v_2(x_1, x_2) = e_2(x_1, x_2)\).

(IIb) If \(p_2 < p_1\), then firm 2 is selling up to its capacity and 1's revenues are exactly \(v_1(x_1, x_2)\). Thus, if 1 is not to undercut 2, we must have that,

\[
p_2 \min(D(p_2), x_1) \leq v_1(x_1, x_2). \quad (8)
\]

Once again we argue that 2's revenues cannot exceed \(e_2(x_1, x_2)\). To see this, recall from Osborne and Pitchik (1986) that except in non-generic and degenerate cases ruled out by Assumption 2,

\[
e_i(x_1, x_2) = p \min(D(p), x_i) \quad (9)
\]

where \(p\) is the lowest price in the supports of the equilibrium mixed strategies of the two firms (it can be shown that the lowest prices in the supports coincide).

Using (8) and the fact that \(v_1(x_1, x_2) \leq e_1(x_1, x_2)\), we can write,

\[
p_2 \min(D(p_2), x_1) \leq p \min(D(p), x_1) \quad (10)
\]

and since by Assumption 2, the function \(L_i(p) = p \min(D(p), x_i)\) is monotonic if \(p < p^m\), we conclude that \(p_2 \leq p\). But this implies that,

\[
p_2 \min(D(p_2), x_2) \leq p \min(D(p), x_2) = e_2(x_1, x_2). \quad (11)
\]

Since the left hand side of the above inequality is exactly equal to 2's revenues when \(p_2 < p_1\), we are done. \(\Box\)

References

Kreps, D. and A.M. Spence, 1983, Modelling the role of history in industrial organization and competition, Mimeo. (Stanford University, Stanford, CA).
Rosenbaum, D.I., 1986, An empirical test on the effect of excess capacity in price-setting, capacity constrained supergames, Working paper no. 86-6 (Department of Economics, College of Business Administration, University of Nebraska-Lincoln, Lincoln, NE).