RAND Journal of Economics
Vol. 35, No. 1, Spring 2004
pp. 23–49

# Capacity dynamics and endogenous asymmetries in firm size

**David Besanko***

**and**

**Ulrich Doraszelski****

*Empirical evidence suggests that there are substantial and persistent differences in the sizes of firms in most industries. We propose a dynamic model of capacity accumulation that is consistent with the observed facts. The model highlights the mode of product market competition and the extent of investment reversibility as key determinants of the size distribution of firms in an industry. In particular, if firms compete in prices and the rate of depreciation is large, then the industry moves toward an outcome with one dominant firm and one small firm. Industry dynamics in this case resemble a preemption race. Contrary to the usual intuition, this preemption race becomes more brutal as investment becomes more reversible.*

## 1. Introduction

■ Empirical evidence suggests that there are substantial and persistent differences in the sizes of firms in most industries. For example, Gort (1963) showed that in 74% of U.S. manufacturing industries, the intertemporal correlation of market shares over a seven-year period was .8 or higher, and that in only 10% of the industries was the correlation less than .5. More recently, Mueller (1986) found that in 44% of 350 U.S. manufacturing industries, the identity of the industry leader remained unchanged over a 22-year period and that the correlation between market shares over this period was .66. This high degree of persistence seems especially remarkable in light of evidence that a significant portion of the variation in firm performance over time is accounted for by firm-specific shocks (McGahan and Porter, 1997).

In this article we propose a dynamic model of capacity accumulation that is able to explain such substantial and persistent differences in the sizes of firms. The pattern of capacity

* Northwestern University; d-besanko@kellogg.northwestern.edu.

** Stanford University; doraszelski@hoover.stanford.edu.

We would like to thank Pat Bajari, Lanier Benkard, Michaela Draganska, J.P. Dubé, Ken Judd, Dan Kovenock, Patricia Langohr, Ariel Pakes, Rob Porter, Mark Satterthwaite, Karl Schmedders, Andy Skrzypacz, and Dmitry Vedenov for helpful comments. This article has also benefited from comments of participants in seminars at the Freie Universität Berlin, Humboldt Universtität zu Berlin, Ludwig-Maximilians-Universität München, University of Pennsylvania, Clarence W. Tow Conference on Computational Economics (2002), the 2002 North American Summer Meetings of the Econometric Society at UCLA, the 2002 International Conference of the Society for Computational Economics in Aix-en-Provence, and the Stanford-Berkeley IO Fest (2002). Finally, we are grateful to the Editor Joe Harrington and two anonymous reviewers for their constructive suggestions.

accumulation is a critical determinant of the firm size distribution. The capacity decisions made by industry participants have potentially profound intertemporal consequences because they entail nontrivial lead times, stickiness, and lumpiness. As a result, capacity decisions can shape the firm size distribution for many years beyond the point at which they are actually made. For example, the current asymmetric structure of the North American titanium dioxide industry—an industry dominated by a single large firm, DuPont—can be traced back to the preemptive strategy of capacity accumulation that DuPont initiated in the early 1970s (Ghemawat, 1984 and 1987; Hall, 1990).

Asymmetric industry structures can, of course, arise as the outcome of a game in which firms differ in their economic fundamentals (e.g., cost structures) or their strategic positions at the outset of the game (e.g., first versus late mover). However, this raises the following question: How do such differences in initial conditions arise in the first place?

Several articles have shown that asymmetric industry structures can arise as the outcome of a capacity accumulation game played by *ex ante* identical firms (e.g., Saloner, 1987; Maggi, 1996; Reynolds and Wilson, 2000). Though valuable in highlighting that substantial differences in firm size can arise endogenously for strategic reasons, these models are less satisfactory in explaining the persistence of these differences over time, which is highlighted by the evidence. This is because these articles consider an unchanging competitive environment. Once an equilibrium has been reached, nothing further happens to upset the positions of firms. That is, asymmetric industry structures persist in these models by default. Ideally, however, one would like to understand whether there are circumstances under which asymmetric industry structures persist in a competitive environment that changes over time, for example, because of firm-specific shocks. The competitive environment can also change due to feedback effects. For example, if large firms invest more aggressively than small firms, then a small initial asymmetry may become larger over time, but it may vanish otherwise. In general, one would expect feedback effects to play a role whenever the time horizon under consideration is long enough to allow firms to interact repeatedly.

For these reasons, an important question is whether substantial and persistent size differences can arise endogenously in equilibrium in a market in which *ex ante* identical firms interact repeatedly and are subject to firm-specific shocks that continuously alter their positions. To address this question, we adapt the Markov-perfect-equilibrium (MPE) framework presented in Ericson and Pakes (1995) to track the evolution of an oligopolistic industry. In our model, firms accumulate capacity over time and compete repeatedly in the product market. We assume that capacity is lumpy and that investment outcomes and depreciation of capacity levels are subject to idiosyncratic shocks. In each period, firms first make their investment decisions and then compete in the product market. At the end of the period, the outcome of a firm's investment decision and the depreciation of its capital stock are realized. We solve numerically for the symmetric MPE in order to characterize industry dynamics and identify circumstances under which asymmetric industry structures arise and persist over time.

We demonstrate that industry dynamics depend critically on the mode of product market competition and on the degree to which investment is reversible, as measured by the rate of depreciation. Under quantity competition, each firm accumulates enough capacity to supply the Cournot quantities, leading to an industry structure of equal-sized firms independent of whether investment is irreversible (zero depreciation) or reversible (positive depreciation). With positive depreciation, firms tend to hold idle capacity out of a precautionary motive. By contrast, under price competition, there are forces that propel the industry toward asymmetric structures. In particular, if investment is reversible, then the industry evolves toward an outcome with one dominant firm and one small firm. Industry dynamics in this latter case resemble a rather fierce preemption race. During this race, firms invest heavily as long as they are of equal size even though this leads to significant industrywide overcapacity. Once one of the firms manages to pull slightly ahead in this race, however, the smaller firm "gives up," thereby propelling the larger firm eventually into a position of dominance.

This article sheds new light on the relationship between preemption and reversibility. We show that under price competition, the preemption race between contending firms becomes more brutal as investment becomes more reversible. This stands in marked contrast to the usual intuition that depreciation reduces the commitment power of capacity and that capacity accumulation can therefore lead to only a temporary advantage. The key insight underlying our result is that with reversible investment, the consequences of a firm's falling behind its rival are not fatal: the firm can allow its capacity to depreciate and assume the more profitable posture of a "puppy dog." The higher the rate of depreciation, the easier it is for a lagging firm to "disengage" from the preemption race and hence the more attractive it is for firms to enter in such a race in the first place.

Our article is related to a number of literatures that seek to explain intraindustry heterogeneity as an equilibrium phenomenon. One such literature studies the implications of adding a time dimension to models of capacity choice. Saloner (1987) considers a three-stage model in which firms are allowed to invest twice in capacity before competing in the product market. The opportunity to observe and react to the initial capacity choice of one's rival induces multiple subgame-perfect equilibria, including all outcomes on the outer envelope of the static best-reply functions between the firms' Stackelberg outcomes. Subsequent work by Pal (1991), Maggi (1996), and Kovenock and Roy (1998) strives to eliminate this continuum of equilibria to arrive at more definitive predictions about the size of firms.

This literature stands in contrast to another that studies once-and-for-all capacity choice under demand uncertainty. Gabszewicz and Poddar (1997) consider a two-stage model of capacity choice followed by capacity-constrained quantity competition, and they show that each firm chooses the certainty-equivalent Cournot capacity, thereby leading to a symmetric industry structure. By contrast, Reynolds and Wilson (2000) demonstrate that if capacity-constrained price competition is assumed instead of quantity competition, then there does not exist a symmetric equilibrium in pure strategies for capacity choices, provided that demand is sufficiently volatile.

Our article builds upon these literatures by embedding capacity accumulation in a competitive environment that changes over time because of both idiosyncratic shocks and feedback effects. In contrast to Saloner (1987), Pal (1991), and Maggi (1996), we employ standard models of product market competition (capacity-constrained quantity competition and capacity-constrained price competition) rather than representing a firm's profit from product market competition by a reduced-form specification. The sign of the cross-partial derivative of this profit function with respect to firms' capital stocks is assumed to be either negative (i.e., capacities are strategic substitutes) or positive (i.e., capacities are strategic complements) for all possible combinations of capital stocks, and it governs the nature of the equilibrium. In Maggi (1996), for example, asymmetric equilibria emerge when capital stocks are strategic substitutes, whereas the equilibrium is always symmetric if capital stocks are strategic complements. Restricting the sign of the cross-partial derivative of the reduced-form profit function to be always either negative or positive, however, is problematic because the profit functions arising from standard models of product market competition like capacity-constrained quantity competition or capacity-constrained price competition fail to satisfy this restriction.[1] We avoid this pitfall by explicitly specifying how firms compete in the product market (as do Kovenock and Roy (1998)) rather than using reduced-form profit functions.

In addition, we draw upon the recent literature on models of industry evolution. The articles by Jovanovic (1982), Hopenhayn (1992), and Mitchell (2000) point out the important role that idiosyncratic shocks play in explaining the great variation in the fate of similar firms over time. We incorporate this insight by making the law of motion of a firm's capital stock stochastic. The firm-specific shocks can generate small asymmetries among firms, and we provide conditions under which these asymmetries evolve into substantial and persistent size differences. Unlike these models of industry evolution, however, we do not assume a continuum of infinitesimally

[1] Details are available from the authors upon request.

small, perfectly competitive firms but instead allow oligopolistic firms to interact strategically (similar to Ericson and Pakes (1995)).[2]

Our article also relates to the extensive literature on dynamic models of capacity accumulation, which analyze firms' investment behavior over time (Spence, 1979; Fudenberg and Tirole, 1983; Hanig, 1985; Reynolds, 1987, 1991; and Dockner, 1992). A limitation of these models is that, contrary to what we observe, no asymmetries emerge in equilibrium provided that firms are *ex ante* identical. Moreover, in the Hanig (1985), Reynolds (1987, 1991), and Dockner (1992) models, the steady state is independent of the initial conditions. Hence, any asymmetry in firms' capital stocks vanishes over time. Thus, these articles cannot explain how differences in the sizes of firms emerge in their own right nor how they are sustained over extended periods. Finally, the articles by Hanig (1985), Reynolds (1987, 1991), and Dockner (1992) specify a linear-quadratic single-period profit function. While this makes the models tractable, it also assumes that firms in each period dump their entire capacities into the market and that price is established to clear the market. Hence, linear-quadratic specifications are not necessarily consistent with optimizing behavior by the firms. Our numerical methods enable us to use standard models of product market competition that are more plausible from a behavioral perspective.

The remainder of this article is organized as follows. In Section 2 we set up the model. To keep things simple, we spell out the model for two firms and homogeneous goods, but we relax both assumptions later on. Sections 3–6 present the results. In Section 7, we first turn to cost/benefit considerations as an alternative rationale for asymmetric industry structures. Then we extend our model from homogeneous goods to differentiated products and allow for more than two firms. Finally, we briefly discuss entry and exit as well as demand uncertainty and report on further robustness checks. Section 8 concludes.

## 2. Model

■ We model the evolution of an oligopolistic industry. The model is cast in discrete time and has an infinite horizon to avoid end effects. There are two firms with identical marginal costs of production but potentially different capacities.[3]

□ **Setup and timing.** We think of production as taking place in a plant with capacity $\overline{q} \geq 0$. The plant's capacity takes on one of $M$ values, and we set $\overline{q}_0 < \overline{q}_1 < \dots < \overline{q}_{M-1}$. In each period, a firm decides how much to invest in order to add capacity to its plant. At the same time, the firm is bound to lose capacity due to depreciation. The outcome of the investment and depreciation processes is assumed to be stochastic. Thus, even if a firm invests, it is not guaranteed that its capacity increases. Moreover, the firm's capacity might decrease due to depreciation in spite of its investment.

After making their investment decisions (but before the outcome of a firm's investment decision and the depreciation of its capital stock are realized), firms compete in the product market. The firms' profits from product market competition are determined by their capacities $(\overline{q}_i, \overline{q}_j)$. To simplify notation we take $(i, j)$ to mean that firm 1 has a capacity of $\overline{q}_i$ and firm 2 a capacity of $\overline{q}_j$, and we denote the single-period profit functions of firm 1 and firm 2 by $\pi_1(i, j)$ and $\pi_2(i, j)$, respectively. We first give details on the product market competition and then turn to the dynamic framework.

□ **Demand.** We employ two models of product market competition in our analysis: capacity-constrained quantity competition and capacity-constrained price competition (Bertrand-Edgeworth competition). To ensure that the two models are comparable, we base them on the same

[2] Moreover, as Maggi (1996) points out, in these models, small firms are small either because they are growing or because they are shrinking: a firm is never persistently small. In this sense, these models of industry evolution cannot explain the persistence of asymmetric industry structures.

[3] It is straightforward to accommodate different marginal costs.

demand specification. The market demand function is $Q(P)$ with corresponding inverse demand function $P(Q)$. Here $Q$ denotes market demand and $P$ market price. To keep things simple, we use a linear demand specification:

$$Q(P) = a - bP \Leftrightarrow P(Q) = \frac{a}{b} - \frac{Q}{b}. \tag{1}$$

This allows us to normalize the marginal costs of production to zero without loss of generality.

□ **Quantity competition.** Suppose that firms' capacities are given by $(\overline{q}_i, \overline{q}_j)$ and that they compete in the product market by setting quantities $(q_1, q_2)$. The profit-maximization problem for, say, firm 1 is given by

$$\max_{0 \le q_1 \le \overline{q}_i} P(q_1 + q_2)q_1.$$

Firm 1's profit-maximization problem and the corresponding problem for firm 2 give rise to (symmetric) reaction functions $q_1(q_2; i, j)$ and $q_2(q_1; i, j)$. It can be shown that there exists a unique fixed point $\left(q_1^*(i, j), q_2^*(i, j)\right)$ of the system of reaction functions (see, e.g., Vives, 1999). The single-period profit function of firm 1 in the Nash equilibrium of the capacity-constrained quantity-setting game is thus

$$\pi_1(i, j) \equiv P(q_1^*(i, j) + q_2^*(i, j))q_1^*(i, j).$$

The profit function for firm 2 is defined analogously.

□ **Price competition.** We next consider a product market game in which two capacity-constrained firms simultaneously set prices $(p_1, p_2)$. This model was first analyzed by Kreps and Scheinkman (1983). Our treatment follows the more general model of Deneckere and Kovenock (1996) and Allen et al. (2000). As is well known, with capacity-constrained price competition, a Nash equilibrium in pure strategies does not always exist, thus making it necessary to consider mixed strategies. Furthermore, the nature of the Nash equilibrium depends on the rationing rule (Davidson and Deneckere, 1986). Following Kreps and Scheinkman (1983), we consider the efficient rationing rule. To illustrate, suppose that $p_1 > p_2$. Then the low-price firm (firm 2) either serves the entire market or hits its capacity constraint, and the residual demand function facing the high-price firm (firm 1) is $\max\{0, Q(p_1) - \overline{q}_j\}$. Given this, firm 1's profit when it is the high-price firm is

$$H_1(p_1; i, j) \equiv p_1 \min\left\{\overline{q}_i, \max\{0, Q(p_1) - \overline{q}_j\}\right\}, \qquad p_1 > p_2,$$

and firm 1's profit when it is the low-price firm is

$$L_1(p_1; i, j) \equiv p_1 \min\left\{\overline{q}_i, Q(p_1)\right\}, \qquad p_1 < p_2.$$

In case of a tie, we take firm 1 to be the low-price firm and firm 2 to be the high-price firm.

We draw on Theorems 1–3 of Deneckere and Kovenock (1996) to characterize the equilibrium profits under capacity-constrained price competition. To do so, let $H_1^*(i, j) \equiv \max_{p_1} H_1(p_1; i, j)$ be firm 1's minmax profit. Further, let $\underline{p}_1(i, j) \equiv \min\{p_1 : L_1(p_1; i, j) = H_1^*(i, j)\}$ denote the lowest price such that firm 1's profit, when it is the low-price firm, equals its minmax profit. $H_2^*(i, j)$ and $\underline{p}_2(i, j)$ are defined analogously for firm 2. Now define the following regions of capacity space:

$$A \equiv \{(\overline{q}_i, \overline{q}_j) \in \mathbb{R}_+^2 : \overline{q}_i \le q_1(\overline{q}_j) \wedge \overline{q}_j \le q_2(\overline{q}_i)\},$$
$$C \equiv \{(\overline{q}_i, \overline{q}_j) \in \mathbb{R}_+^2 : \overline{q}_i \ge Q(0) \wedge \overline{q}_j \ge Q(0)\},$$
$$B_1 \equiv \{(\overline{q}_i, \overline{q}_j) \in \mathbb{R}_+^2 \backslash (A \cup C) : \overline{q}_i \le \overline{q}_j\},$$
$$B_2 \equiv \{(\overline{q}_i, \overline{q}_j) \in \mathbb{R}_+^2 \backslash (A \cup C) : \overline{q}_i \ge \overline{q}_j\},$$

where $q_1(\overline{q}_j)$ and $q_2(\overline{q}_i)$ denote the static reaction functions for firm 1 and firm 2, respectively, in a Cournot game with marginal costs of zero and unlimited capacities. Firm 1's single-period profit function in the Nash equilibrium of the capacity-constrained price-setting game then is

$$\pi_1(i, j) \equiv \begin{cases} \overline{q}_i P(\overline{q}_i + \overline{q}_j) & \text{if } (\overline{q}_i, \overline{q}_j) \in A, \\ 0 & \text{if } (\overline{q}_i, \overline{q}_j) \in C, \\ L_1(\underline{p}_2(i, j); i, j) & \text{if } (\overline{q}_i, \overline{q}_j) \in B_1, \\ H_1^*(i, j) & \text{if } (\overline{q}_i, \overline{q}_j) \in B_2. \end{cases}$$

The profit function for firm 2 is defined analogously.

The product market equilibrium is as follows: In region $A$, the lower envelope of the reaction functions, both firms sell at full capacity and charge the market-clearing price. In region $C$, each firm has sufficient capacity to serve the entire market. There is classic Bertrand competition: both firms set price equal to marginal costs and make zero profits. Outside of regions $A$ and $C$, the product market equilibrium is in mixed strategies. In region $B_1$, firm 1 is smaller than firm 2; in region $B_2$, firm 1 is larger than firm 2. The price distribution of the larger firm first-order stochastically dominates that of the smaller firm. In this sense, the larger firm holds a "price umbrella" over the smaller firm. In regions $B_1$ and $B_2$, the larger firm earns its Stackelberg follower profit (taking the capacity of the smaller firm as given), whereas the smaller firm earns less than its Stackelberg leader profit.

□ **State-to-state transitions.** We now turn to the dynamic framework. The industry is completely described by the tuple $(i, j) \in \{0, \ldots, M-1\}^2$, and we call $(i, j)$ the state of the industry. Given that the industry is in state $(i, j)$ today, it will be in state $(i', j')$ tomorrow. Our next task is to specify the probability distribution that governs the state-to-state transitions.

We treat capacity as lumpy. In fact, we think about adding a block of capacity as adding a new assembly line to an existing plant or building a new plant. Investment projects of this magnitude involve uncertainty, as witnessed by cost overruns and extensive delays. For example, Nucor had to resolve a large number of technical difficulties in building its Crawfordsville, Indiana steel mill. By the time it was operational, the cost of the project had grown well beyond the originally budgeted \$225 million to \$260 million (Preston, 1991). Pindyck (1993) more generally shows that the *actual* construction time for nuclear power plants varied from six to as long as sixteen years during the late 1970s and 1980s, whereas the *expected* construction time clustered around ten years during the early 1980s. We capture this uncertainty by assuming that an investment project is either successful or unsuccessful and that the probability of success is higher the more resources are devoted to it. Specifically, if firm 1 invests $x_1 \geq 0$, then the probability that its investment project succeeds is $\alpha x_1/(1 + \alpha x_1)$, where the parameter $\alpha > 0$ measures the effectiveness of investment (at $x_1 = 0$, to be precise). Note that this can alternatively be interpreted as a time-to-build specification (Kydland and Prescott, 1982): If firm 1 invests $x_1$ period after period, then the time until a block of capacity comes online is geometrically distributed with mean $(1 + \alpha x_1)/\alpha x_1$ and variance $(1 + \alpha x_1)/(\alpha x_1)^2$. Hence, the more resources a firm devotes to the investment project, the sooner (in the sense of first-order stochastic dominance) the capacity will come online. In particular, by investing more, the firm can reduce the expected time to build and enhance its ability to come to terms with the inherent uncertainty of the investment project.[4]

While a firm invests to add capacity to its plant, it is bound to lose capacity due to depreciation. That is, depreciation tends to offset investment. Similar to the investment process, we take the outcome of the depreciation process to be random and assume that a firm is hit by a depreciation shock with probability $\delta \geq 0$. We think of depreciation shocks as being caused by things like machine breakdowns, industrial fires, and other accidents. They are therefore specific to a firm.[5]

[4] Note that a firm does not have to commit to a stream of investments. We differ from a time-to-build specification in that we allow each firm to adjust to changes in the capacity levels of both firms. For example, a firm may change its investment if its rival succeeds first in bringing additional capacity online.

[5] To ensure that our results are robust, we have repeated the computations with an industrywide depreciation shock. The differences are minor.

Combing the investment and depreciation processes, firm 1 transits between capacity levels according to the transition function

$$\theta(i' \mid i, x_1) = \begin{cases} \dfrac{(1-\delta)\alpha x_1}{1+\alpha x_1} & \text{if } i' = i+1, \\ \dfrac{1-\delta+\delta\alpha x_1}{1+\alpha x_1} & \text{if } i' = i, \\ \dfrac{\delta}{1+\alpha x_1} & \text{if } i' = i-1 \end{cases}$$

if $i \in \{1, \ldots, M-2\}$. That is, $\theta(i' \mid i, x_1)$ is the probability that firm 1 will be state $i'$ tomorrow given that it is in state $i$ today. Clearly, firm 1 cannot move further down (up) from the lowest (highest) state. We therefore set

$$\theta(i' \mid i, x_1) = \begin{cases} \dfrac{\alpha x_1}{1+\alpha x_1} & \text{if } i' = i+1, \\ \dfrac{1}{1+\alpha x_1} & \text{if } i' = i \end{cases}$$

if $i = 0$, and

$$\theta(i' \mid i, x_1) = \begin{cases} \dfrac{1-\delta+\alpha x_1}{1+\alpha x_1} & \text{if } i' = i, \\ \dfrac{\delta}{1+\alpha x_1} & \text{if } i' = i-1 \end{cases}$$

if $i = M-1$. Because we interpret the lowest state as zero capacity, it is natural to assume the absence of depreciation in the transition function for $i = 0$.

Note that although we have opted for a simple specification of the depreciation process, alternative specifications could be readily incorporated into the state-to-state transitions. For example, it is straightforward to let the rate of depreciation depend on a firm's capacity level. Such a specification may be appropriate if we think about a plant's capacity as being determined by the number of machines in the plant. Given that $\mu$ is the probability that a machine breaks down, $1-(1-\mu)^i$ is the probability that at least one of $i$ machines breaks down. This could be taken into account by replacing the constant rate of depreciation $\delta$ with the increasing rate of depreciation $\delta(i) = 1-(1-\mu)^i$.[6]

□ **Bellman equation.** Let $V_1(i, j)$ denote the expected net present value to firm 1 of being in the industry given that firm 1 holds $\overline{q}_i$ units of capacity and firm 2 holds $\overline{q}_j$ units of capacity. In what follows, we first characterize the value function $V_1(i, j)$ under the presumption that the firm behaves optimally. In a second step, we derive the policy function $x_1(i, j)$. Throughout we take firm 2's investment strategy $x_2(i, j)$ as given.

The Bellman equation is

$$V_1(i, j) = \max_{x_1 \geq 0} \pi_1(i, j) - x_1 + \beta \sum_{i'=0}^{M-1} W_1(i')\theta(i' \mid i, x_1), \tag{2}$$

where $0 < \beta < 1$ is the discount factor and

$$W_1(i') = \sum_{j'=0}^{M-1} V_1(i', j')\theta(j' \mid j, x_2(i, j)). \tag{3}$$

[6] Because the alternative specification makes it more expensive for the larger firm to maintain/expand its capacity, it tends to reduce any differences in the size of firms. Yet, the conclusions that we draw below also emerge under the alternative specification provided that investment remains reasonably cost-effective for the larger firm.

The Bellman equation adds the firm's current cash flow $\pi_1(i, j) - x_1$ and its discounted expected future cash flow. Note that $\sum_{i'=0}^{M-1} W_1(i')\theta(i' \mid i, x_1)$ is the expectation over all possible future states $(i', j')$ calculated under the presumption that firm 1 invests $x_1$ and firm 2 invests $x_2(i, j)$ in the current state $(i, j)$.

□ **Investment strategy.** The first-order condition (FOC) for an interior solution is

$$-1 + \beta \sum_{i'=0}^{M-1} W_1(i') \frac{\partial \theta(i' \mid i, x_1)}{\partial x_1} = 0.$$

Consider $i \in \{1, \ldots, M - 2\}$. Solving the FOC for $x_1$ yields

$$\frac{-1 + \sqrt{\beta\alpha\left((1-\delta)(W_1(i+1) - W_1(i)) + \delta(W_1(i) - W_1(i-1))\right)}}{\alpha}.$$

The second-order condition (SOC) reduces to

$$-\left((1-\delta)(W_1(i+1) - W_1(i)) + \delta(W_1(i) - W_1(i-1))\right) < 0.$$

Hence, the SOC is satisfied whenever a solution to the FOC exists. Moreover, the objective function equals $\pi_1(i, j) + \beta((1 - \delta)W_1(i) + \delta W_1(i - 1))$ at $x_1 = 0$ and approaches $-\infty$ as $x_1$ approaches $\infty$. Hence, the objective function is decreasing when a solution to the FOC fails to exist. Thus,

$$x_1(i, j) = \max\left\{0, \frac{-1 + \sqrt{\beta\alpha\left((1-\delta)(W_1(i+1) - W_1(i)) + \delta(W_1(i) - W_1(i-1))\right)}}{\alpha}\right\} \quad (4)$$

if this is well defined, and $x_1(i, j) = 0$ otherwise. If $i = 0$ or $i = M - 1$, the investment strategy of firm 1 can be derived using similar arguments.

□ **Equilibrium.** The two models of product market competition that we employ in our analysis give rise to symmetric profit functions, i.e., $\pi_1(i, j) = \pi_2(j, i)$. We therefore define $\pi(i, j) \equiv \pi_1(i, j)$, note that $\pi_2(i, j) = \pi(j, i)$, and restrict attention to symmetric Markov-perfect equilibria (MPE). Hence, if $V(i, j) \equiv V_1(i, j)$ denotes firm 1's value function, then firm 2's value function is given by $V_2(i, j) = V(j, i)$. Similarly, if $x(i, j) \equiv x_1(i, j)$ denotes firm 1's policy function, then firm 2's policy function is given by $x_2(i, j) = x(j, i)$.[7] Existence of a symmetric MPE in pure investment strategies follows from the arguments in Doraszelski and Satterthwaite (2003) if we impose an upper bound on investment. While uniqueness cannot in general be guaranteed, our computations always led to the same value and policy functions irrespective of the starting point and the particulars of the algorithm.

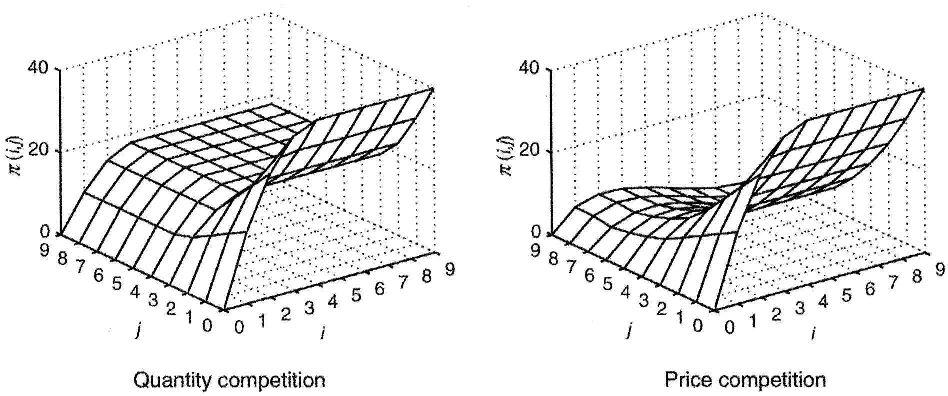
□ **Computation.** To compute the symmetric MPE, we use a variant of the algorithm described in Pakes and McGuire (1994). The algorithm works iteratively. It takes a value function $\tilde{V}(i, j)$ and a policy function $\tilde{x}(i, j)$ as its input and generates updated value and policy functions as its output. Each iteration proceeds as follows: First, we use equation (4) to compute firm 1's investment strategy $x(i, j)$ taking firm 2's investment strategy to be given by $\tilde{x}(j, i)$. In doing so, we use $\tilde{V}(i, j)$ and $\tilde{x}(j, i)$ to compute $W(i')$ (as defined in equation (3)). Second, we compute the payoff $V(i, j) = \pi(i, j) - x(i, j) + \beta \sum_{i'=0}^{M-1} W(i')\theta(i' \mid i, x(i, j))$ associated with firm 1 using $x(i, j)$ as its investment strategy and firm 2 using $x(j, i)$ (see equation (2)).[8] In this step, we use $\tilde{V}(i, j)$ and $x(j, i)$ to obtain $W(i')$. The iteration is completed by assigning $V(i, j)$ to $\tilde{V}(i, j)$

[7] Firms compete not only with their investments, but also with their quantities and prices (depending on the mode of product market competition). Similar to the investment strategy, the quantity sold or price charged depends on firms' capacities, which in turn gives rise to state-specific per-period profits $\pi(i, j)$.

[8] In general, we compute the payoff that results when firm 1 adheres to $x(i, j)$ and firm 2 to $x(j, i)$ for $K \geq 1$ periods into the future.

FIGURE 1

PROFITS $\pi(i, j)$ FROM PRODUCT MARKET COMPETITION




and $x(i, j)$ to $\tilde{x}(i, j)$.[9] The algorithm terminates once the relative change in the value and the policy functions from one iteration to the next are below a prespecified tolerance. All programs are written in Matlab 5.3.[10]

□ **Parameterization.** Increasing the rate of depreciation $\delta$ gives rise to two opposing forces. First, the probability of success, and thus the marginal benefit to investment, declines. Second, additional investment is needed to maintain installed capacity. To counteract the first effect somewhat, we choose $\alpha$ given $\delta$ such that the probability of success equals $\bar{\theta}$ at an investment of $\bar{x}$. That is,

$$\bar{\theta} = \frac{(1-\delta)\alpha\bar{x}}{1+\alpha\bar{x}} \Leftrightarrow \alpha = \frac{\bar{\theta}}{(1-\delta-\bar{\theta})\bar{x}}.$$

This in turn requires that $1 - \delta - \bar{\theta} > 0$. We pick $\delta \in \{0, .01, .1, .3\}$.

While we vary the rate of depreciation over a wide range, we hold the remaining parameters constant. Their values are $\bar{x} = 20$, $\bar{\theta} = .5$, and $\beta = 1/1.05$. Note that the choice of $\beta$ corresponds to a yearly interest rate of 5%. We set $M = 10$ with $\bar{q}_0 = 0$, $\bar{q}_1 = 5$ up to $\bar{q}_9 = 45$. That is, if the industry is in state $(i, j)$, then firm 1 has $i$ blocks or $5i$ units of capacity and firm 2 has $j$ blocks or $5j$ units of capacity.

The demand parameters are $a = 40$ and $b = 10$. Figure 1 illustrates the resulting single-period profit functions for quantity competition (left panel) and price competition (right panel). Note that since $\bar{q}_9 = 45 > 40 = a$, a single firm is potentially able to serve the entire market. For future reference, we note that given these demand parameters, the monopoly quantity is $q^M = 20$ (4 blocks), the Cournot quantity is $q^D = 40/3 \approx 13.33$ (between 2 and 3 blocks), and the Stackelberg quantity is $q^L = 20$ (4 blocks) and $q^F = 10$ (2 blocks) for the leader and the follower, respectively.

## 3. Capacity dynamics and industry structure under quantity competition

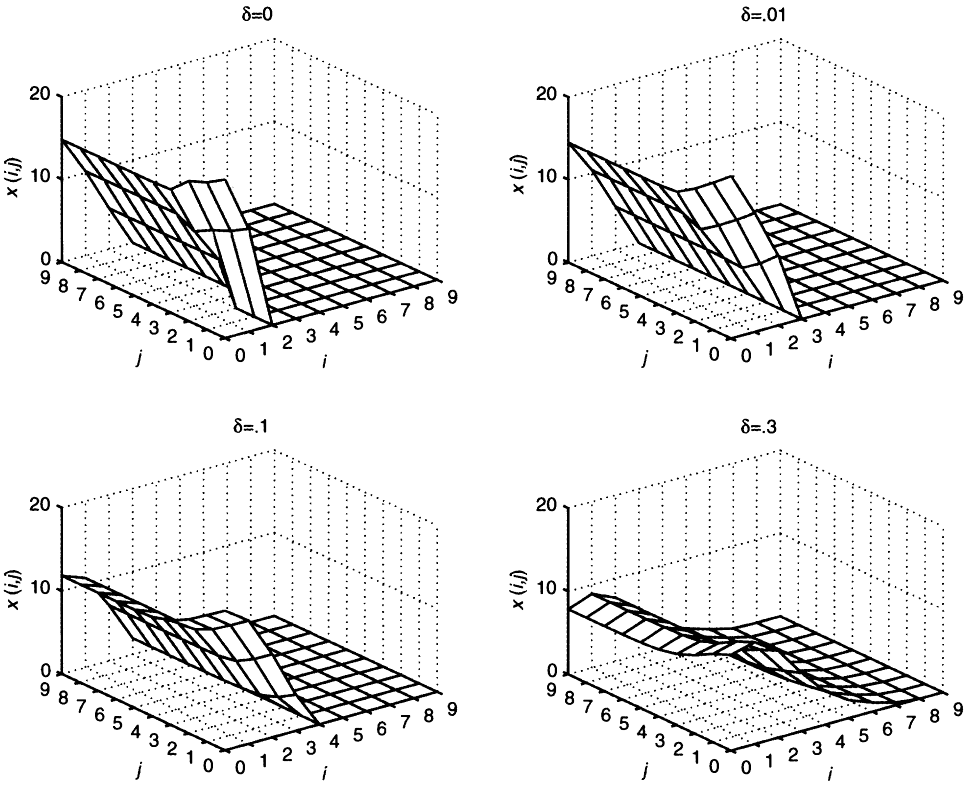
■ In this and the next section, we first present the results for the two modes of product market competition described above: capacity-constrained quantity competition and capacity-constrained price competition. In Sections 5 and 6, we then compare our results to the stylized facts about persistent size differences and unchanging industry leadership and discuss the relationship between investment reversibility and preemption races.

[9] More precisely, we assign a weighted average of $V(i, j)$ and $\tilde{V}(i, j)$ to $\tilde{V}(i, j)$ to help the algorithm converge.

[10] Details are available from the authors upon request.

FIGURE 2

POLICY FUNCTION $x(i, j)$: QUANTITY COMPETITION




Throughout, our approach is to use the equilibrium policy functions to construct the probability distribution over tomorrow's state $(i', j')$ given today's state $(i, j)$, i.e., the transition matrix that characterizes industry dynamics. This allows us to use tools from stochastic process theory to analyze the Markov process of industry dynamics (rather than relying on simulation).

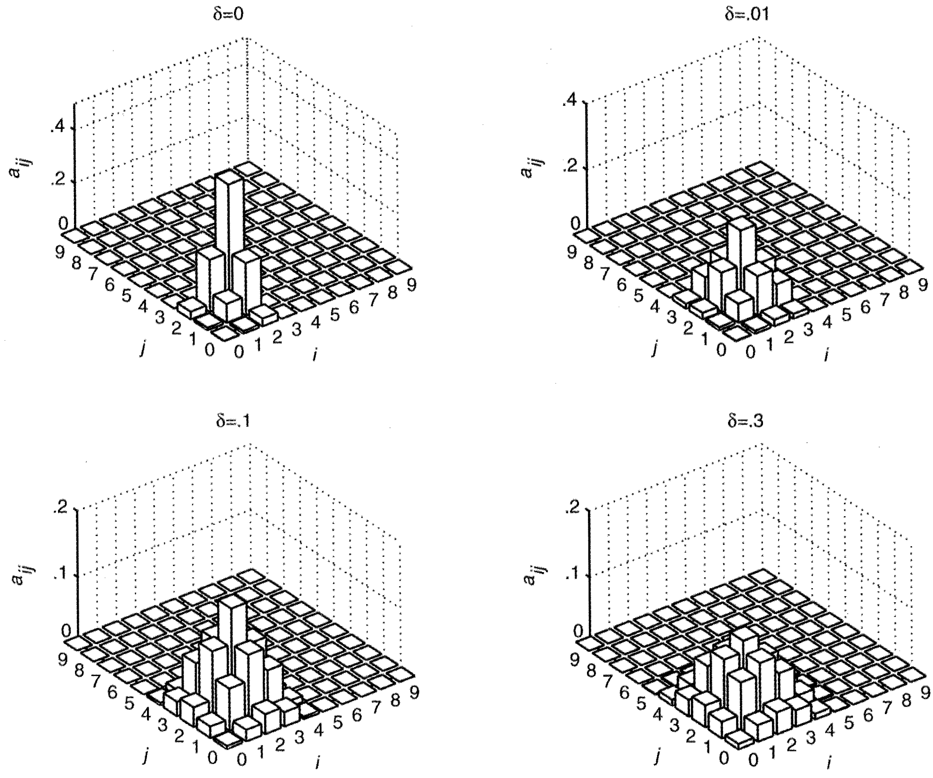
We discuss the transitory (short-run) dynamics of the Markov process of industry dynamics first and then turn to its steady-state (long-run) dynamics. We find that quantity competition leads to symmetric industry structures, whereas price competition leads to asymmetric industry structures. Moreover, under price competition, differences in firm size are more pronounced when investment is reversible than when it is irreversible. Next, we examine the transition matrix itself in order to describe the governing forces of industry evolution (resultant force) and show that the symmetric industry structures under quantity competition are supported by global catch-up forces, whereas local forces of increasing dominance give the industry a natural tendency toward firms of unequal size under price competition. Lastly, we investigate the long-run performance of the industry.

□ **Transitory dynamics.** When depreciation is zero ($\delta = 0$), investment is irreversible. In this case, as can be seen from the top left panel of Figure 2, the policy function takes on a simple form. More or less irrespective of its opponent's capacity, a firm invests until it has built up its capacity to a level that roughly allows it to supply its Cournot quantity.

Given the simple form of the policy function, one expects the industry to evolve toward the Cournot point. This is indeed the case, as seen in the top left panels of Figures 3 and 4. The top left panel of Figure 3 depicts the marginal distribution of states $(i, j)$ after $T = 5$ periods, starting from state (0, 0), and the top left panel of Figure 4 depicts the same after $T = 25$ periods. (In addition, we

FIGURE 3

TRANSIENT DISTRIBUTION AFTER $T = 5$ PERIODS WITH INITIAL STATE $i_0 = j_0 = 1$: QUANTITY COMPETITION



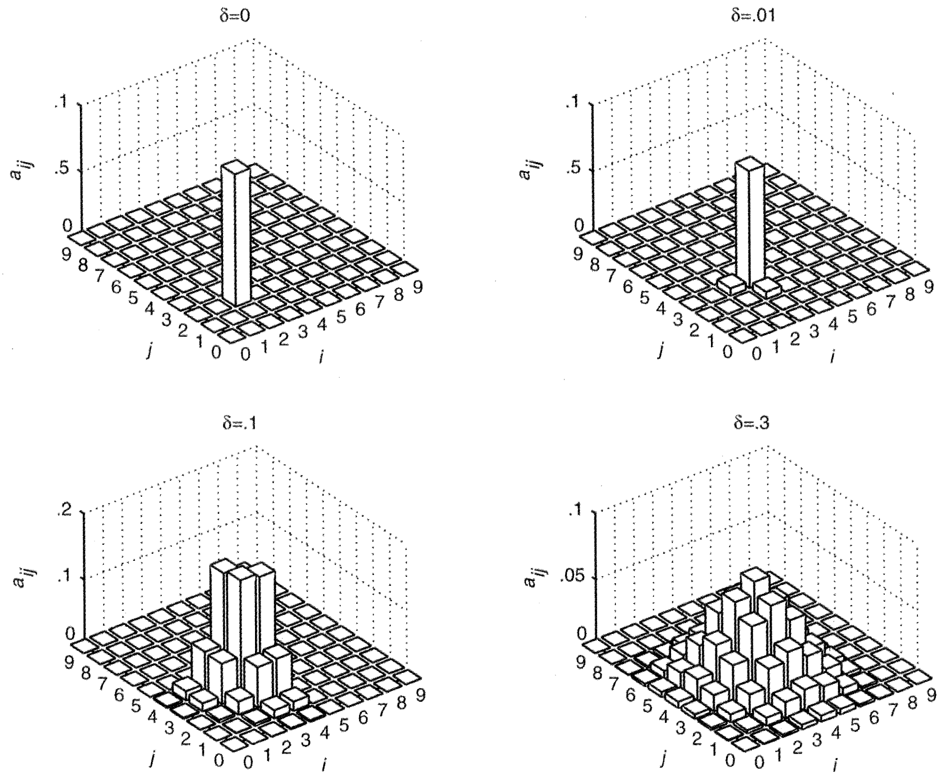

computed the marginal distribution after $T = 15$ and $T = 50$ periods but omitted the graphs.) As can be seen, the industry converges to state (2, 2) over time. Specifically, state (2, 2) is the mode of the marginal distribution and has a probability of .45, .99, 1.00, and 1.00 after $T = 5, 15, 25, 50$ periods, respectively. While asymmetric states are possible if one firm's investment fails and the other's succeeds, asymmetric states become less likely over time. For example, states (1, 2) and (2, 1) each have a probability of .20, .01, .00, and .00 after $T = 5, 15, 25, 50$ periods, respectively. The reason for this is that the lagging firm strives to catch up with the leading firm: $x(1, 2) = 10.22 > 0 = x(2, 1)$. This restores symmetry.

When depreciation is positive ($\delta > 0$), investment is reversible, and the degree of reversibility is increasing with the rate of depreciation. Depreciation provides the firm with an additional incentive to invest in order to maintain its installed capacity. Because of this "maintenance" or "precautionary" incentive, the critical level of capacity at which a firm finally stops investing rises as the rate of depreciation rises (top right and bottom panels of Figure 2). Consequently, the mode of the marginal distribution of states $(i, j)$ shifts out (Figures 3 and 4). After $T = 25$ periods, for example, it is state (3, 3) for $\delta \in \{.01, .1\}$ and state (4, 4) for $\delta = .3$.

Since the depreciation of firms' capital stocks is driven by random shocks, the evolution of the industry becomes more uncertain as investment becomes more reversible (i.e., as $\delta$ increases) in the sense that the support of the marginal distribution after any given number of periods $T$ becomes more spread out (Figures 3 and 4). However, because a firm's investment continues to be fairly insensitive to its rival's capacity (Figure 2), the transitory distribution remains concentrated on states in which firms are of equal size. In other words, the pattern of strategic interaction between firms tends to restore symmetry.

FIGURE 4

TRANSIENT DISTRIBUTION AFTER $T = 25$ PERIODS WITH INITIAL STATE $i_0 = j_0 = 1$: QUANTITY COMPETITION




□ **Steady-state dynamics.** When $\delta = 0$, neither firm invests in state (2, 2), so neither firm's capacity changes. Hence, with irreversible investment, state (2, 2) is a steady state or, more precisely, an absorbing state. But are there other steady states? To answer this question, we have partitioned the state space into closed communicating classes (recurrent sets) and the remaining states and computed the limiting distribution (ergodic distribution) for each of these closed communicating classes. In addition to state (2, 2), each state $(i, j)$ with $i \geq 3$ and $j \geq 3$ constitutes a closed communicating class because neither firm invests in these states, so that once the industry enters one of them, it remains there. Since each closed communicating class consists of a single state, the steady-state distribution is trivial (top left panel of Figure 5).

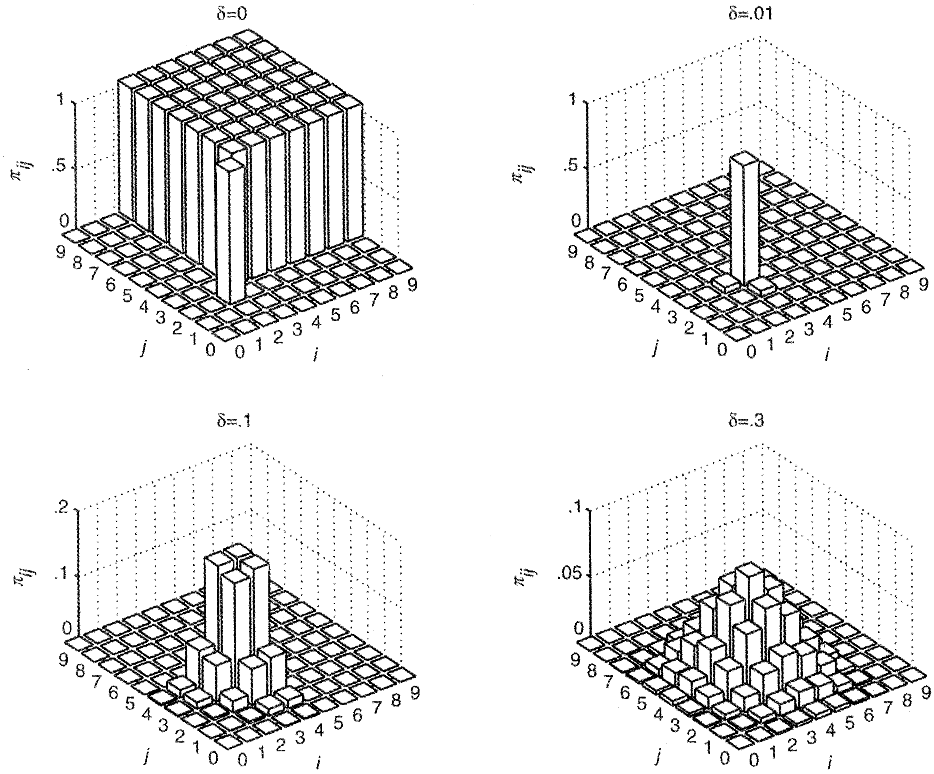
As we move from irreversible to reversible investment, there is a "discontinuity" in the closed communicating classes. In particular, the states in a rectangular region around the origin form a single recurrent set. This is the case because in the presence of depreciation, it is always possible to move down a block. At the same time, it is possible to move up a block as long as there is investment. Consequently, the states around the origin are communicating.[11] Moreover, the recurrent set becomes larger with the rate of depreciation because the critical level of capacity at which a firm eventually ceases to invest rises.

Since there is a single recurrent set, the industry eventually ends up within it, regardless of the initial state. Once it enters the closed communicating class, the fraction of time that the Markov process of industry dynamics spends in each state is given by the ergodic distribution.

[11] Formally, consider state $(i, j)$ and suppose that $\delta > 0$. Then states $(i-1, j)$, $(i, j-1)$, and $(i-1, j-1)$ are accessible from state $(i, j)$. Suppose next that $x(i, j) > 0$ and $x(j, i) > 0$. Then states $(i+1, j)$, $(i, j+1)$, and $(i+1, j+1)$ are accessible from state $(i, j)$. Finally suppose that $x(i, j) > 0$ but $x(j, i) = 0$. Then state $(i+1, j)$ is accessible from state $(i, j)$. Two states communicate if and only if they are mutually accessible.

FIGURE 5

LIMITING DISTRIBUTION: QUANTITY COMPETITION




As can be seen from Figure 5, the shape of the ergodic distribution changes "smoothly" with the rate of depreciation. More specifically, it remains unimodal while the mode shifts out (due to maintenance investment) and the modal probability decreases (due to additional uncertainty). The mode is state (3, 3) for $\delta = .01$ and state (4, 4) for $\delta \in \{.1, .3\}$, and the fraction of time spent at the mode is .91, .18, .07 for $\delta = .01, .1, .3$, respectively.

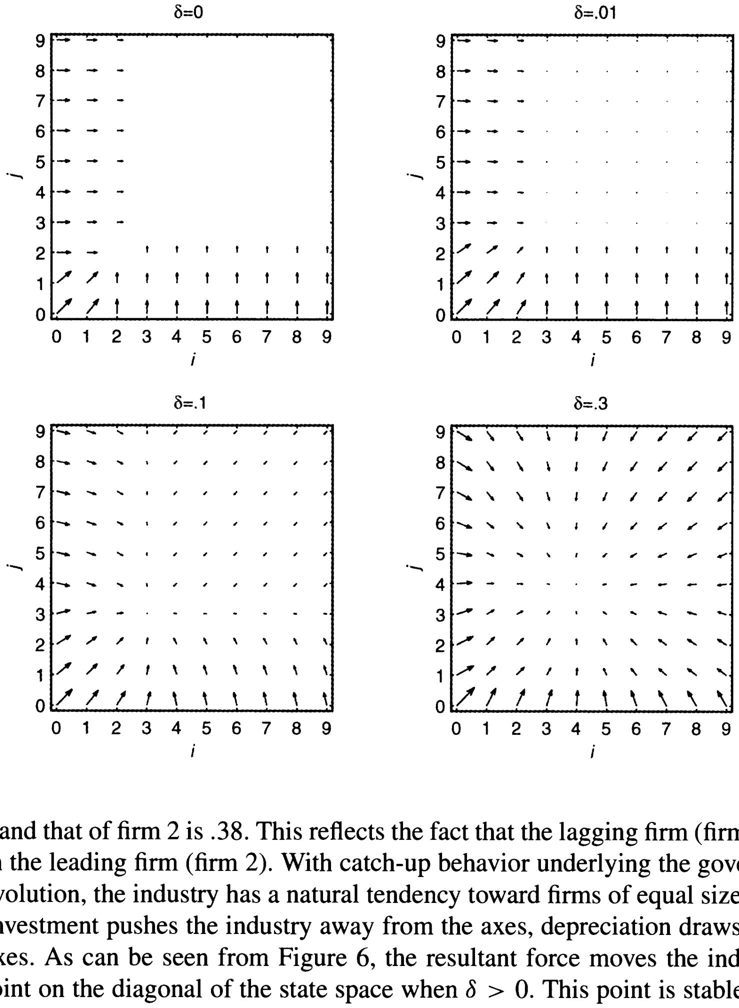
□ **Resultant force.** Given that the industry is in state $(i, j)$ today, it will be in state $(i', j')$ tomorrow, where $i' - i \in \{-1, 0, 1\}$ and $j' - j \in \{-1, 0, 1\}$. We determine the expected movement of the state by computing the probability-weighted average of $(i' - i, j' - j)$. Reinterpreting $(i' - i, j' - j)$ as a direction and the associated probability as the force operating in that direction, the expected movement becomes the resultant force. The arrows in Figure 6 thus point in the direction in which the industry is expected to evolve, and their length indicates the speed at which this is expected to happen.[12]

When $\delta = 0$, in states $(i, j)$ with $i \leq 1$ and $j \leq 1$, the resultant force indicates that firms accumulate capacity roughly in lockstep; in states $(i, j)$ with $i \geq 2$ or $j \geq 2$, the smaller firm grows and the larger firm remains the same. A closer inspection of the resultant force reveals that the expected movement of firm 1 is (weakly) larger than the expected movement of firm 2 in states $(i, j)$ with $i < j$. Indeed, even around the origin (in states $(i, j)$ with $i \leq 1$ and $j \leq 1$) there is a slight drift toward the diagonal. For example, in state (0, 1) the expected movement of

[12] More formally, given a state $(i, j)$, we first compute the resultant force as $E((i' - i, j' - j) \mid (i, j)) = (\sum_{i'=i-1}^{i+1} (i' - i)\theta(i' \mid i, x(i, j)), \sum_{j'=j-1}^{j+1} (j' - j)\theta(j' \mid j, x(j, i)))$. Then we plot an arrow with foot at $(i, j)$ and head at $(i, j) + E((i' - i, j' - j) \mid (i, j))$.

FIGURE 6

RESULTANT FORCE: QUANTITY COMPETITION




firm 1 is .47, and that of firm 2 is .38. This reflects the fact that the lagging firm (firm 1) strives to catch up with the leading firm (firm 2). With catch-up behavior underlying the governing forces of industry evolution, the industry has a natural tendency toward firms of equal size.

While investment pushes the industry away from the axes, depreciation draws the industry toward the axes. As can be seen from Figure 6, the resultant force moves the industry toward an interior point on the diagonal of the state space when $\delta > 0$. This point is stable because the resultant force offsets a displacement.

As in the case of irreversible investment, the expected movement of firm 1 is (weakly) larger than the expected movement of firm 2 in states $(i, j)$ with $i < j$. Due to these catch-up forces, the industry gravitates toward firms of equal size.

□ **Industry performance.** To evaluate the long-run performance of the industry, we use the ergodic distribution to compute the expected value of the combined profits from product market competition, $E(\pi(i, j) + \pi(j, i))$, and of the combined capacity, $E(\overline{q}_i + \overline{q}_j)$. Table 1

**TABLE 1** **Industry Performance: Quantity Competition**

| $\delta$ | $E(\pi(i, j) + \pi(j, i))$ | $E(\overline{q}_i + \overline{q}_j)$ |
|---|---|---|
| 0 | 40.00 | 20.00 |
| .01 | 35.75 | 29.51 |
| .1 | 36.28 | 32.32 |
| .3 | 36.35 | 38.88 |

Note: Expected combined profits and capacity.

presents the results for different rates of depreciation. For comparison purposes, the corresponding quantities for $\delta = 0$ and the steady state (2, 2) are added. Note that the combined profit function $\pi(i, j) + \pi(j, i)$ achieves its maximum of 40 in state (2, 2).[13] Turning from $\delta = 0$ to $\delta > 0$, Table 1 shows that under quantity competition with reversible investment, over time, the industry grows to a point where firms reap less than maximal gains from product market competition and that the industry suffers from overcapacity. Yet, although expected combined capacity increases quickly, expected combined profits increase slowly with $\delta$. This is because quantity competition significantly mutes the deleterious effects of overcapacity on profits.

□ **Summary.** Quantity competition results in an industry structure with equal-sized firms. Differences in firm size are temporary because catch-up behavior tends to restore symmetry. In fact, the equilibrium strategies give rise to global catch-up forces. That is, the expected movement of the laggard exceeds that of the leader for all possible combinations of capacity levels. The industry therefore has a natural tendency toward firms of equal size.

With quantity competition, the rate of depreciation affects the investment levels and the resulting capacity levels of firms. In particular, if the rate of depreciation increases, both firms tend to accumulate more capacity. This accumulation of capacity is attributable to precautionary motives as opposed to preemption or deterrence (strategic) motives. This is because a firm's equilibrium investment is relatively insensitive to its rival's capacity. Hence, a firm cannot deter its rival from investing by growing large.[14]

## 4. Capacity dynamics and industry structure under price competition

■ We now assume that firms compete as capacity-constrained price setters in the product market. As before, we consider the cases of irreversible ($\delta = 0$) and reversible ($\delta > 0$) investment.

□ **Transitory dynamics.** As the top left panel of Figure 7 shows, under price competition with irreversible investment the policy function takes on a somewhat more complicated form than under capacity-constrained quantity competition. More specifically, while a firm again invests until it has built up a capacity that roughly allows it to supply the Cournot quantity, the magnitude of the firm's investment depends critically on its opponent's capacity. To characterize the sensitivity of firm 1's investment to firm 2's capacity, fix $i$ and consider $\Delta^x(i) = \max_j x(i, j) - \min_j x(i, j)$. Let $\Delta^x \equiv (\Delta^x(0), \ldots, \Delta^x(9))$. Under price competition, $\mathbf{\Delta}^x = (16.62, 19.97, 20.77, 0, 0, 0, 0, 0, 0, 0)$, compared to $\mathbf{\Delta}^x = (4.36, 3.36, 4.94, 0, 0, 0, 0, 0, 0, 0)$ under quantity competition.

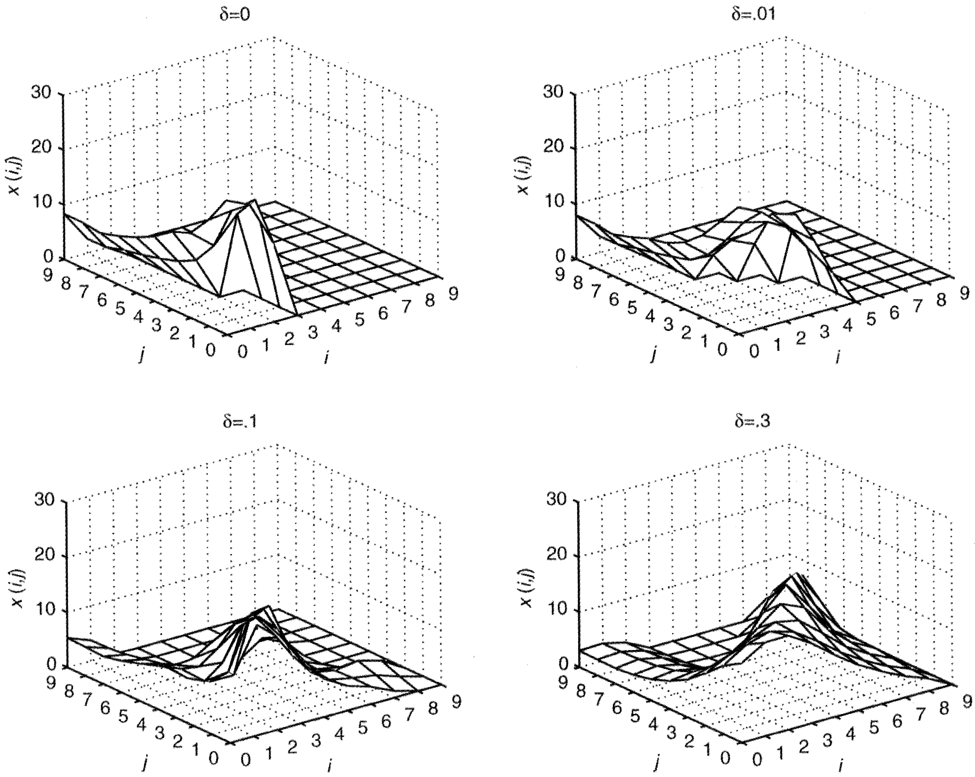
The industry evolves to a distribution of firm sizes that is more asymmetric than the distribution under quantity competition with irreversible investment. Recall that under quantity competition the symmetric state (2, 2) was the mode of the marginal distribution and had a probability of .45, .99, 1.00, and 1.00 after $T = 5, 15, 25, 50$ periods, respectively. By contrast, under price competition, the asymmetric states (2, 3) and (3, 2) are the modes of the marginal distribution, and each has a probability of .20, .44, .45, and .45 after $T = 5, 15, 25, 50$ periods, respectively (top left panels of Figures 8 and 9). In other words, the most likely industry structure is one in which one firm has 15 units of capacity and the other has 10 units. The symmetric state (3, 3) has a probability of just .07, .11, .11, and .11 after $T = 5, 15, 25, 50$ periods, respectively, because, in contrast to quantity competition, the smaller firm under price competition does not strive to catch up with the larger firm. For example, if firm 1 has 5 units of capacity and firm 2 has 10 units, then firm 1 invests $x(1, 2) = 12.01$ while firm 2 invests $x(2, 1) = 20.77$. Because the laggard invests less than the leader, the initial asymmetry between firms tends to be further

[13] This is an artifact of the capacity levels we use. A separate appendix available at www.rje.org/main/sup-mat.html gives details.

[14] We talk of deterrence (inducement) if a firm stops (starts) to invest as its rival grows. Usually this term is used to describe that a firm decreases (increases) its investment as its rival grows.

FIGURE 7

POLICY FUNCTION $x(i, j)$: PRICE COMPETITION




exaggerated. Consequently, symmetric industry structures are much less likely than asymmetric ones, as they essentially presuppose that firms accumulate capacity in lockstep.

As we turn from $\delta = 0$ to $\delta > 0$, a firm has a precautionary motive to invest, similar to the case of quantity competition. Different from the case of quantity competition, however, the firm also has a strategic motive, as it can now deter its rival from investing by growing large.[15] For $\delta = .3$, for example, $x(7, j) > 0$ if $j \leq 8$ and $x(7, j) = 0$ if $j \geq 9$; $x(6, j) > 0$ if $j \leq 7$ and $x(6, j) = 0$ if $j \geq 8$; and $x(i, j) > 0$ if $4 \leq i \leq 5$ and $j \leq 6$ and $x(i, j) = 0$ if $4 \leq i \leq 5$ and $j \geq 7$ (bottom right panel of Figure 7 and Table 3). On the other hand, for $\delta = .3$, a firm always invests until it has built up at least 4 blocks of capacity. The reason is that a small firm, irrespective of its rival's capacity, can increase its profit from product market competition by adding a modest amount of capacity (right panel of Figure 1). Taken together, these two features of the policy function imply that the larger firm has a strategic advantage over a not-too-small rival because the smaller firm "gives up" if it is sufficiently far behind.

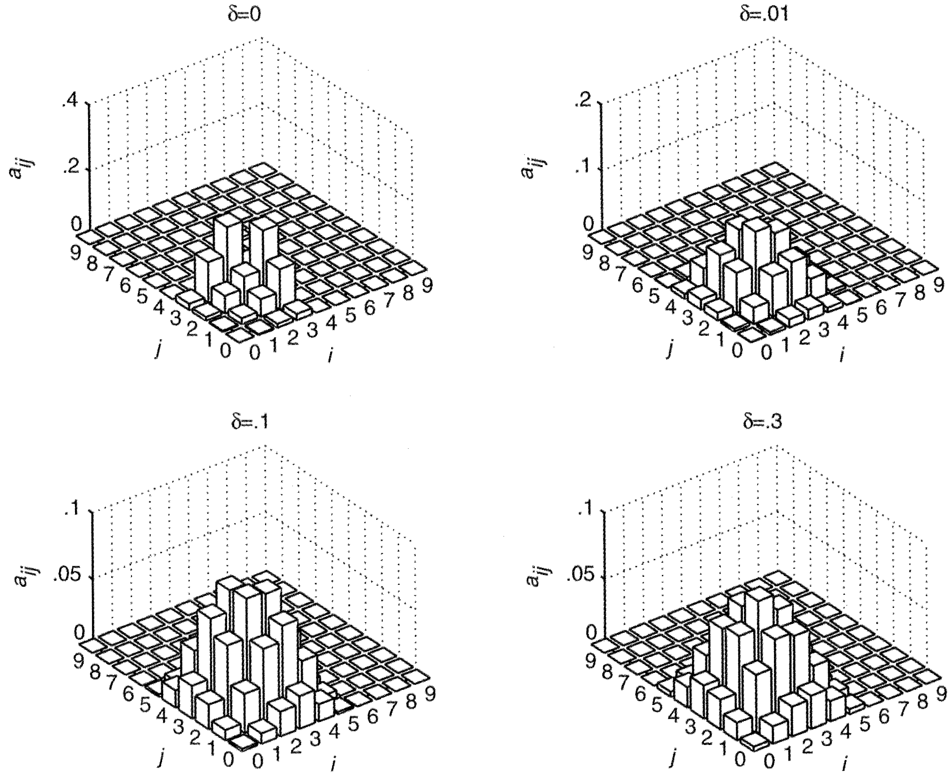
The possibility of gaining a strategic advantage leads to industry dynamics that resemble a preemption race. This can be seen most clearly by examining the policy functions for $\delta = .1$ and $\delta = .3$ in Tables 2 and 3, respectively (see also the bottom panels of Figure 7). In this race, both firms start off investing heavily. Moreover, as long as they are of equal size, both firms continue investing heavily. For $\delta = .3$, for example, both firms are investing 14.40 in state (0, 0) and 11.37 in state (6, 6). This is astonishingly large, especially given that there is significant industrywide overcapacity in state (6, 6) (i.e., industry capacity equals 60 units compared to a market demand of, at most, 40 units).

---

[15] This deterrence effect is already present for $\delta = 0$: $x(2, j) > 0$ if $j \leq 2$ and $x(2, j) = 0$ if $j \geq 3$.

FIGURE 8

TRANSIENT DISTRIBUTION AFTER $T$ = 5 PERIODS WITH INITIAL STATE $i_0 = j_0 = 1$: PRICE COMPETITION



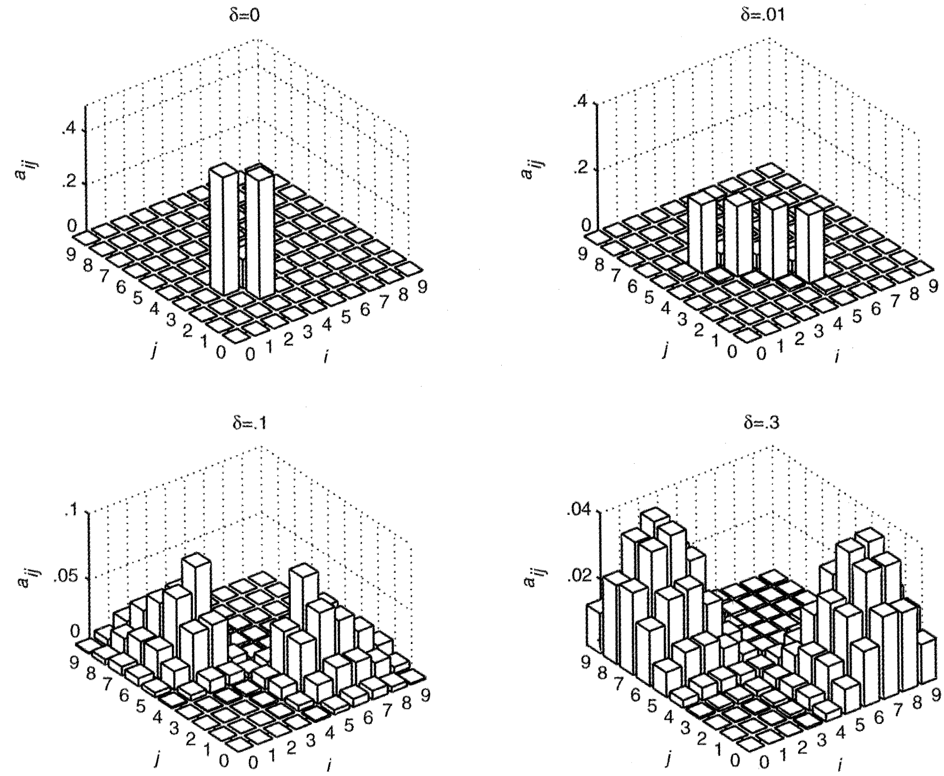

On the other hand, there is a marked drop in investment activity once one firm gains a size advantage over its rival. Continuing with the above example, if firm 1 pulls even slightly ahead in the race (the industry moves to state (7, 6)), then firm 2 scales back its investment to 2.73 while firm 1 ratchets up its investment to 13.46. This tends to further enhance the asymmetry between firms. But once the industry has reached state (7, 5) or state (8, 6), firm 2 gives up. Since investment is reversible, this ensures that firm 2 shrinks, thereby propelling firm 1 eventually into a position of dominance.

Due to additional uncertainty from idiosyncratic shocks, industry dynamics become more volatile as investment becomes more reversible. Sooner or later, therefore, one of the firms gains a size advantage over its rival who will then, in effect, surrender. Consequently, industry structures become more asymmetric as investment becomes more reversible. After $T = 25$ periods, for example, the modes of the marginal distribution are states (3, 4) and (4, 3) for $\delta = .01$, states (3, 6) and (6, 3) for $\delta = .1$, and states (2, 8) and (8, 2) for $\delta = .3$. That is, with $\delta = .3$, the most likely outcome after $T = 25$ periods is that one firm has 40 units of capacity while the other has but 10 units. By contrast, under quantity competition with $\delta = .3$, the modal outcome is that each firm has 20 units of capacity. Moreover, under price competition, differences in firm size are dramatically larger when investment is reversible than when it is irreversible. For example, firms most likely differ by 15 units of capacity after $T = 50$ periods if $\delta = .01$ as opposed to by 5 units if $\delta = 0$.

□ **Steady-state dynamics.** With irreversible investment, neither firm invests in states (2, 3), (3, 2), and (3, 3), so these states are absorbing. As in the case of quantity competition with irreversible investment, there is a number of additional steady states. In fact, each state $(i, j)$ with

FIGURE 9

TRANSIENT DISTRIBUTION AFTER $T = 25$ PERIODS WITH INITIAL STATE $i_0 = j_0 = 1$: PRICE COMPETITION




$i \geq 2$ and $j \geq 3$ or $i \geq 3$ and $j \geq 2$ by itself constitutes a closed communicating class. Since each closed communicating class consists of a single state, the steady-state distribution is trivial (top left panel of Figure 10).

As we move from $\delta = 0$ to $\delta > 0$, there is again a "discontinuity" in the closed communicating classes at $\delta = 0$ in that the states around the origin form a single recurrent set. This recurrent set becomes larger with the rate of depreciation. As the above discussion of firms' investment

**TABLE 2** **Investment Policy $x(i, j)$: Price Competition with $\delta = .1$**

| $x(i, j)$ | | $j = 0$ $\bar{q}_j = 0$ | $j = 1$ $\bar{q}_j = 5$ | $j = 2$ $\bar{q}_j = 10$ | $j = 3$ $\bar{q}_j = 15$ | $j = 4$ $\bar{q}_j = 20$ | $j = 5$ $\bar{q}_j = 25$ | $j = 6$ $\bar{q}_j = 30$ | $j = 7$ $\bar{q}_j = 35$ | $j = 8$ $\bar{q}_j = 40$ | $j = 9$ $\bar{q}_j = 45$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $i = 0$ | $\bar{q}_i = 0$ | 21.02 | 16.75 | 12.37 | 9.56 | 7.94 | 6.81 | 5.75 | 5.03 | 5.26 | 5.32 |
| $i = 1$ | $\bar{q}_i = 5$ | 22.10 | 18.91 | 12.88 | 8.14 | 6.23 | 5.55 | 4.88 | 4.21 | 3.83 | 3.75 |
| $i = 2$ | $\bar{q}_i = 10$ | 18.36 | 19.13 | 15.83 | 7.69 | 2.74 | 1.88 | 2.12 | 1.55 | .62 | .15 |
| $i = 3$ | $\bar{q}_i = 15$ | 12.09 | 14.06 | 17.70 | 14.66 | 3.27 | 0 | 0 | 0 | 0 | 0 |
| $i = 4$ | $\bar{q}_i = 20$ | 6.99 | 7.78 | 10.09 | 15.66 | 12.76 | 0 | 0 | 0 | 0 | 0 |
| $i = 5$ | $\bar{q}_i = 25$ | 4.47 | 4.11 | 3.92 | 5.34 | 7.60 | 7.92 | 0 | 0 | 0 | 0 |
| $i = 6$ | $\bar{q}_i = 30$ | 3.14 | 3.23 | 3.11 | 1.80 | 1.73 | 4.50 | 2.86 | 0 | 0 | 0 |
| $i = 7$ | $\bar{q}_i = 35$ | 1.22 | 2.57 | 3.64 | 1.03 | 0 | 0 | .86 | 0 | 0 | 0 |
| $i = 8$ | $\bar{q}_i = 40$ | 0 | 1.18 | 2.54 | .21 | 0 | 0 | 0 | 0 | 0 | 0 |
| $i = 9$ | $\bar{q}_i = 45$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**TABLE 3** **Investment Policy $x(i, j)$: Price Competition with $\delta = .3$**

| $x(i, j)$ | | $j = 0$ $\bar{q}_j = 0$ | $j = 1$ $\bar{q}_j = 5$ | $j = 2$ $\bar{q}_j = 10$ | $j = 3$ $\bar{q}_j = 15$ | $j = 4$ $\bar{q}_j = 20$ | $j = 5$ $\bar{q}_j = 25$ | $j = 6$ $\bar{q}_j = 30$ | $j = 7$ $\bar{q}_j = 35$ | $j = 8$ $\bar{q}_j = 40$ | $j = 9$ $\bar{q}_j = 45$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $i = 0$ | $\bar{q}_i = 0$ | 14.40 | 12.50 | 9.54 | 6.91 | 5.06 | 3.89 | 3.17 | 2.85 | 2.90 | 2.97 |
| $i = 1$ | $\bar{q}_i = 5$ | 16.99 | 15.14 | 11.66 | 8.11 | 5.64 | 4.32 | 3.66 | 3.32 | 3.17 | 3.13 |
| $i = 2$ | $\bar{q}_i = 10$ | 17.12 | 16.32 | 13.68 | 9.40 | 5.55 | 3.49 | 2.73 | 2.42 | 2.20 | 2.10 |
| $i = 3$ | $\bar{q}_i = 15$ | 14.52 | 15.10 | 15.31 | 12.69 | 7.31 | 2.85 | .91 | .36 | .14 | .02 |
| $i = 4$ | $\bar{q}_i = 20$ | 11.21 | 12.28 | 14.45 | 15.93 | 12.74 | 5.71 | .25 | 0 | 0 | 0 |
| $i = 5$ | $\bar{q}_i = 25$ | 8.29 | 9.08 | 10.97 | 14.14 | 16.78 | 12.47 | 4.09 | 0 | 0 | 0 |
| $i = 6$ | $\bar{q}_i = 30$ | 5.89 | 6.35 | 7.40 | 9.01 | 11.59 | 15.63 | 11.37 | 2.73 | 0 | 0 |
| $i = 7$ | $\bar{q}_i = 35$ | 3.85 | 4.22 | 4.82 | 5.47 | 6.02 | 9.15 | 13.46 | 9.39 | 2.58 | 0 |
| $i = 8$ | $\bar{q}_i = 40$ | 2.23 | 2.51 | 2.90 | 3.25 | 3.29 | 3.93 | 7.20 | 10.39 | 6.12 | 3.04 |
| $i = 9$ | $\bar{q}_i = 45$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.75 | .25 | 0 |

behavior and the induced transitory dynamics suggests, the ergodic distribution is indeed bimodal. Hence, in sharp contrast to the case of quantity competition, the industry consists, most of the time, of a large firm and a small firm. As the rate of depreciation increases, the modes are increasingly asymmetric and the modal probability decreases. The modes are states (2, 5) and (5, 2) for $\delta = .01$, states (2, 8) and (8, 2) for $\delta = .1$, and states (1, 8) and (8, 1) for $\delta = .3$, and the fraction of time spent at any one of the modes is .28, .07, .05 for $\delta = .01, .1, .3$, respectively (Figure 10).

FIGURE 10

LIMITING DISTRIBUTION: PRICE COMPETITION

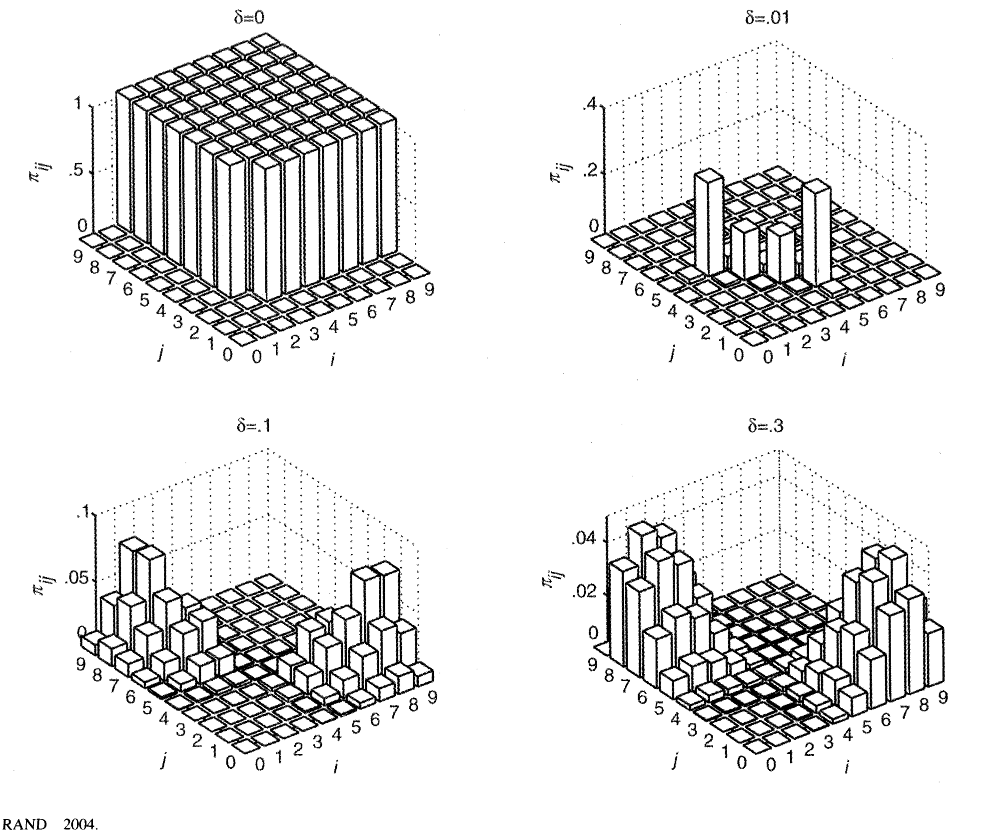
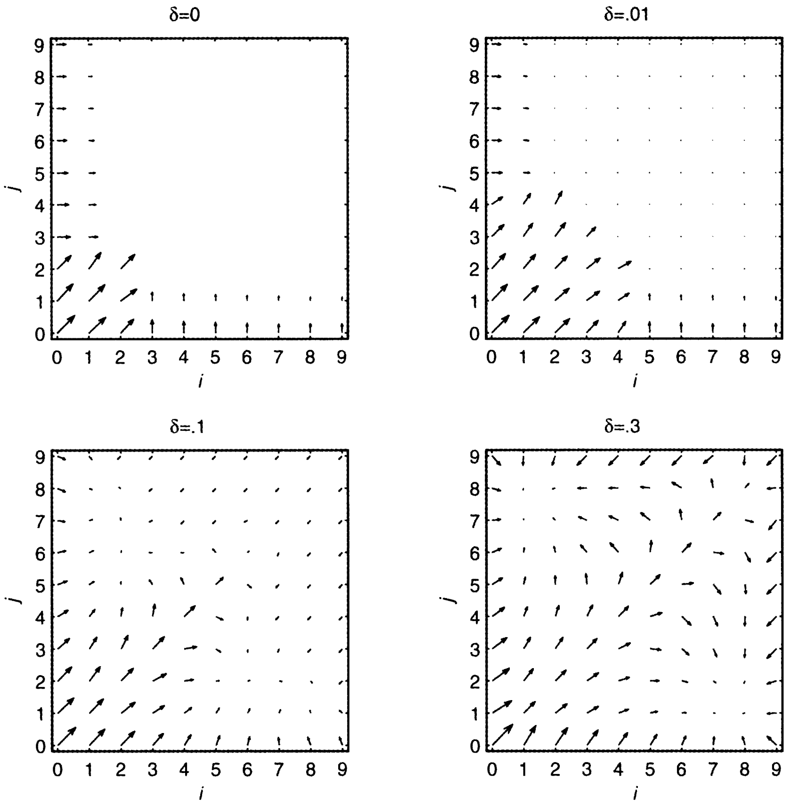


FIGURE 11

RESULTANT FORCE: PRICE COMPETITION




□ **Resultant force.** When $\delta = 0$, the resultant force under price competition exhibits some similarities to those in evidence under quantity competition. In particular, in states $(i, j)$ with $i \leq 2$ and $j \leq 2$, the resultant force operates roughly parallel to the diagonal; in states $(i, j)$ with $i \geq 3$ or $j \geq 3$, it operates along one of the axes (top left panel of Figure 11). However, there is also a notable difference in that the expected movement of firm 1 is (strictly) smaller than the expected movement of firm 2 in states (0, 1) and (1, 2). For example, in state (0, 1) the expected movement of firm 1 is .49, and that of firm 2 is .53. Thus, rather than catching up with the leading firm, there is a tendency for the lagging firm to fall further behind.

When $\delta > 0$, the resultant force moves the industry toward one of two interior points of the state space, as can be seen from Figure 11. One of these points is below the diagonal, and the other one is above the diagonal. Both points are locally stable in the sense that the resultant force offsets a small displacement.

Compared to the case of irreversible investment, the local forces of increasing dominance are decidedly more prominent in the case of reversible investment. For example, for $\delta = .1$, the expected movement of firm 1 is (strictly) smaller than the expected movement of firm 2 in states $(1, j)$ with $2 \leq j \leq 4$, states $(2, j)$ with $3 \leq j \leq 8$, states $(3, j)$ with $4 \leq j \leq 8$, states $(4, j)$ with $5 \leq j \leq 6$, state (5, 6), and state (6, 7). These local forces of increasing dominance ensure that small differences between firms tend to develop into large differences. In other words, the industry gravitates toward firms of unequal size.

□ **Industry performance.** Table 4 presents the expected value of the combined profits from product market competition, $E\left(\pi(i, j) + \pi(j, i)\right)$, and of the combined capacity, $E\left(\overline{q}_i + \overline{q}_j\right)$. For comparison purposes we add the corresponding quantities for $\delta = 0$ and the steady state (2, 3).

**TABLE 4 Industry Performance: Price Competition**

| $\delta$ | $E(\pi(i,j)+\pi(j,i))$ | $E(\overline{q}_i+\overline{q}_j)$ | $E(\pi^L(\cdot))$ | $E(\pi^S(\cdot))$ | $E(\pi^L(\cdot)/\overline{q}^L(\cdot))$ | $E(\pi^S(\cdot)/\overline{q}^S(\cdot))$ |
|---|---|---|---|---|---|---|
| 0 | 37.50 | 25.00 | 22.50 | 15.00 | 1.50 | 1.50 |
| .01 | 29.73 | 35.24 | 19.91 | 9.82 | .85 | .85 |
| .1 | 30.14 | 46.25 | 24.02 | 6.12 | .67 | .67 |
| .3 | 32.25 | 43.90 | 27.48 | 4.78 | .78 | .58 |

Note: Expected combined profits and capacity as well as expected profits and profit-to-capacity ratio.

Note that $\pi(2,3)+\pi(3,2) = 37.50$, which compares to a maximum of 40 of the combined profit function $\pi(i,j)+\pi(j,i)$. Hence, under price competition with irreversible investment, over time, the industry grows to a point where firms earn less than maximal profits in the product market. Moreover, the industry suffers from overcapacity. As Table 4 shows, the same is true under price competition with reversible investment. Compared to quantity competition (Table 1), expected combined profits are lower and expected combined capacity is higher.

Given that maximal joint profits are 40, firms are apparently quite far from the joint profit maximum (at least whenever $\delta > 0$). This stands in contrast to the tendency to move toward the joint profit maximum that arises in Budd, Harris, and Vickers's (1993) model of dynamic competition between two firms. There are at least two main differences between our approach and theirs: First, in their model, profits from product market competition depend on the *difference* between firms' states. In other words, their state space is one-dimensional, while ours is two-dimensional. Second, they use asymptotic expansions (perturbation methods) to approximate the value and policy functions around the special cases of infinite discounting and infinite uncertainty. As Budd, Harris, and Vickers note, away from these special cases, additional effects operate and may well dominate the joint-profit effect that comes out of the asymptotic expansions.

Table 4 also presents the expected profits of the large firm, $E(\pi^L(i,j))$, and the small firm, $E(\pi^S(i,j))$, along with their respective profit-to-capacity ratios $E(\pi^L(i,j)/\overline{q}^L(i,j))$ and $E(\pi^S(i,j)/\overline{q}^S(i,j))$.[16] The profit gap between the large firm and the small firm is large to begin with and becomes even larger as the rate of depreciation increases. The reason is that the large firm dominates the industry more and more, while its rival is marginalized. The rates of return are nevertheless roughly similar between the two firms: if $\delta = 0$, both firms in fact make the same profits per unit of capacity; if $\delta > 0$, the profit-to-capacity ratios are alike although the large firm charges a higher price (in a stochastic sense) because the small firm enjoys a higher capacity utilization rate.

□ **Summary.** In contrast to quantity competition, price competition leads to firms of unequal size. Moreover, under price competition, differences in firm size are more pronounced when investment is reversible than when it is irreversible. These differences are persistent because increasing-dominance behavior tends to aggravate asymmetries that arise accidentally. The equilibrium strategies under price competition give rise to local forces of increasing dominance as opposed to the global catch-up forces under quantity competition.

If the rate of depreciation increases, then one firm accumulates more capacity, the other less. Hence, firms become more asymmetric. In fact, one of the firms is able to achieve a lasting advantage over the other. The reason for this is that a larger firm has a strategic advantage over its not-too-small rival because the smaller firm gives up once it is sufficiently far behind. This locks the smaller firm into a marginal position. Which of the two firms eventually dominates the market is determined in a rather fierce preemption race. During this race, firms fight hard to gain an initial edge over their rival, knowing that once they attain size superiority, the rival will, in effect, surrender.

[16] More formally, we define $\pi^L(i,j) = \pi(\max(i,j), \min(i,j))$, $\pi^S(i,j) = \pi(\min(i,j), \max(i,j))$, $\overline{q}^L(i,j) = \overline{q}_{\max(i,j)}$, and $\overline{q}^S(i,j) = \overline{q}_{\min(i,j)}$. In addition, we take the profit-to-capacity ratio to be zero whenever capacity is zero.

## 5. Industry leadership

■ We nowask how our results on industry structure and dynamics under the different modes of product market competition compare to the stylized facts about persistent size differences and unchanging industry leadership (Gort, 1963; Mueller, 1986).

We start with capacity-constrained quantity competition. Table 5 shows the contemporaneous and intertemporal correlations of capacity levels. The contemporaneous correlation $\rho(\overline{q}_{i,t}, \overline{q}_{j,t})$ between firms' capacities at time $t$ measures the strength of the link between firms in equilibrium. Irrespective of the degree of reversibility, the contemporaneous correlation is small, reflecting the fact that a firm's investment is insensitive to its rival's capacity. The intertemporal correlation $\rho(\overline{q}_{i,t}, \overline{q}_{i,t-h})$ between a firm's capacity at time $t$ and its capacity at time $t - h$, where $h \geq 1$, is a measure of the degree of persistence in a firm's capacity. The intertemporal correlation is declining rapidly in the lag $h$. For example, for $\delta = .01$, the intertemporal correlation decreases from .80 over .32 to .00 as the lag $h$ increases from 1 over 5 to 25 periods. Hence, past firm size is a weak predictor of current firm size. Taken together, the contemporaneous and intertemporal correlations indicate that it is improbable that a firm gains a lasting advantage over its rival. Rather, industry leadership quite frequently changes hands under capacity-constrained quantity competition.

Next we turn to capacity-constrained price competition (Table 6). The contemporaneous correlation $\rho(\overline{q}_{i,t}, \overline{q}_{j,t})$ between firms' capacities is large (in absolute value), irrespective of the degree of reversibility, because a firm's investment is sensitive to its rival's capacity. Moreover, firms' fortunes are negatively correlated. The intertemporal correlation $\rho(\overline{q}_{i,t}, \overline{q}_{i,t-h})$ of a firm's capacity is declining slowly in the lag $h$, indicating that past firm size is a strong predictor for current firm size. For example, the intertemporal correlation to the lag of 25 periods is .94, .89, .84 for $\delta = .01, .1, .3$, respectively. This is, of course, more in line with the evidence on long-lasting industry leadership than the rapidly declining intertemporal correlation under quantity competition.

Note that once a firm reaches a position of dominance, that position is not guaranteed to persist indefinitely. Because both modes of the ergodic distribution are contained in a single recurrent set, it follows that role reversals must occur from time to time. Still, role reversals are unlikely. For example, the expected time it takes the industry to move from one mode to the other is 1,715, 3,023, 1,219 periods for $\delta = .01, .1, .3$, respectively. This is reflected in the large gap between the equilibrium payoff of the leading and the lagging firm. For example, for $\delta = .01$, $V(5, 2) = 464.55$ and $V(2, 5) = 184.22$; for $\delta = .1$, $V(8, 2) = 474.75$ and $V(2, 8) = 193.07$; and for $\delta = .3$, $V(8, 1) = 526.26$ and $V(1, 8) = 40.94$.[17] To see why role reversals are unlikely, suppose that $\delta = .1$ and consider state (2, 8) in which firm 1 has 10 units of capacity and firm 2 has 40 units. Firm 1 stops investing as soon as it adds one more block of capacity. In fact, as long as firm 2 has at least 25 units of capacity, firm 1 cannot accumulate more than 15 units of capacity of its own accord. It can do so only once firm 2 has fewer than 25 units of capacity (Table 2). Hence, what is needed for a role reversal is a long string of bad luck for firm 2 in order to "bring the leader back to the pack," followed by some good luck for firm 1.[18] This is unlikely to happen.

**TABLE 5 Contemporaneous and Intertemporal Correlations: Quantity Competition**

| $\delta$ | $\rho(\overline{q}_{i,t}, \overline{q}_{j,t})$ | $\rho(\overline{q}_{i,t}, \overline{q}_{i,t-1})$ | $\rho(\overline{q}_{i,t}, \overline{q}_{i,t-5})$ | $\rho(\overline{q}_{i,t}, \overline{q}_{i,t-25})$ |
|---|---|---|---|---|
| .01 | .00 | .80 | .32 | .00 |
| .1 | −.04 | .86 | .47 | .02 |
| .3 | −.06 | .90 | .59 | .08 |

[17] Recall that the modes of the ergodic distribution are states (2, 5) and (5, 2) for $\delta = .01$, states (2, 8) and (8, 2) for $\delta = .1$, and states (1, 8) and (8, 1) for $\delta = .3$.

[18] The empirical evidence (see Caves (1998) and the references given therein) appears to support such a reversion to the mean. However, the model also suggests that an industry is generally dominated by one firm (although the identity of the leader may change from time to time). It is not clear whether this latter prediction is supported by the data.

TABLE 6 Contemporaneous and Intertemporal Correlations: Price Competition

| $\delta$ | $\rho(\bar{q}_{i,t}, \bar{q}_{j,t})$ | $\rho(\bar{q}_{i,t}, \bar{q}_{i,t-1})$ | $\rho(\bar{q}_{i,t}, \bar{q}_{i,t-5})$ | $\rho(\bar{q}_{i,t}, \bar{q}_{i,t-25})$ |
|---|---|---|---|---|
| .01 | −.95 | .99 | .98 | .94 |
| .1 | −.89 | .99 | .96 | .89 |
| .3 | −.86 | .98 | .93 | .84 |

Our dynamic model of capacity accumulation yields testable predictions. In particular, the implied correlations in Tables 5 and 6 could be contrasted with their empirical counterparts. Of course, this requires us to determine the mode of product market competition. This can be done in several ways, including direct observation of industry practice (i.e., "production to stock" versus "production to order" (Friedman, 1982)) and customer behavior (i.e., can a firm write binding contracts with customers or not (Grossman, 1981)). In addition, the empirical IO literature has attempted to infer market conduct in oligopolistic industries, e.g., by estimating a conjectural variations parameter or by using a menu approach in combination with nonnested tests (see Bresnahan (1989) for a survey of the literature). Once industries are classified into quantity-setting and price-setting ones, we could also apply our model toward explaining interindustry differences in firm size distributions.

## 6. Preemption races

■ Capacity-constrained price competition leads to asymmetric industry structures. As we have shown above, with reversible investment the roles of dominant and marginal firm are determined by the outcome of the preemption race. In the course of this race, firms invest heavily even though this puts them in a situation of significant industrywide overcapacity, with all its deleterious effects on profits. Moreover, the preemption race becomes more brutal as the rate of depreciation increases. In what follows, we further discuss these surprising results.

It has been suggested that depreciation reduces the commitment value of capacity and therefore that capacity accumulation can only lead to a short-run advantage. The intuition is that when capital can depreciate, it loses (some of) its commitment power, making preemption less feasible and thus less attractive (Tirole, 1988). Our analysis of price competition with reversible investment seems to run counter to this intuition. We find that when investment is reversible, one firm is able to achieve a substantial and persistent advantage over its rival. Moreover, the preemption race itself becomes more brutal as investment becomes more reversible: firms now invest in states in which they would not have invested before. In particular, investment now continues in states where the contending firms are similar in size although the industry is in a situation of overcapacity. For example, $x(7, 7) = 0$ if $\delta = .1$ but $x(7, 7) > 0$ if $\delta = .3$ (Tables 2 and 3). In fact, put somewhat loosely, investment activity increasingly concentrates in states where the two firms are of equal size as the rate of depreciation increases (Figure 7).

To see why this is happening, consider the single-period profit function shown in Figure 1. Under capacity-constrained price competition, a firm's profit from product market competition *peaks* in its own capacity provided that its rival has 20 or more units of capacity (see the right panel of Figure 1). Hence, it is often better for the smaller firm to be considerably smaller than the larger firm rather than slightly smaller. For example, if the smaller firm (firm 1) has 25 units of capacity compared to the 30 units of the larger firm (firm 2), it earns a per-period profit of 4.69. On the other hand, if the smaller firm were to scale back to 15 units of capacity, it would earn a profit of 7.81. In other words, under price competition, there is a benefit to assuming the posture of a "puppy dog" while allowing one's rival to be a "top dog." The reason for this is that, for intermediate capacity levels, the top dog charges higher prices (in a stochastic sense) and in this way extends a price umbrella over the puppy dog.[19] By contrast, under capacity-constrained quantity competition, a firm's profit from product market competition *plateaus* in its own capacity (left panel of Figure 1). Hence, matching one's rival never hurts because whenever firms' capacities are at or in excess of the Cournot quantity, both firms sell at the Cournot quantity (and thus hold idle capacity).

Consequently, under price competition with reversible investment, it is in the self-interest of a not-too-small firm to withdraw from the race once its rival has gained a size advantage. This, in effect, "tips" the market in favor of the leader. Once this happens, the laggard, by dramatically scaling back its investment, puts itself on a path whereby it can shrink. Along this path, price competition in the product market softens, and the laggard's (as well as the leader's) profitability increases.

This, in turn, explains why firms invest heavily during a preemption race even though this leads to significant overcapacity. By building up its capacity, a firm hopes to gain an initial edge over its rival and to tip the market in its favor, thereby inducing its rival to surrender. This logic becomes more powerful as investment becomes more reversible because the laggard, by stopping to invest, can become a "puppy dog" more rapidly. In contrast, if investment is irreversible, then this logic does not apply. This explains why even small rates of depreciation lead to large differences in firm size.

The more general point highlighted by our analysis is that firms' incentives to participate in a preemption race grow stronger as investment becomes more reversible. While this defies the notion that reversibility makes preemption less attractive, the intuition is straightforward: When investment is reversible, each firm knows that its rival has the opportunity to back down. Knowing this, each firm rightfully anticipates that it could eventually be propelled into a position of dominance. By contrast, when investment is irreversible, the excess capacity that is accumulated during a preemption race cannot be withdrawn from the industry. In this case, preemption races are inevitably bloody and therefore not worth entering into in the first place.

## 7. Extensions and generalizations

■ In this section, we briefly report on a number of extensions and generalizations. A separate appendix available at www.rje.org/main/sup-mat.html gives details.

□ **Cost/benefit considerations and asymmetric industry structures.** Asymmetries arise and persist provided that one firm has a strategic advantage over the other. The tangible form of this advantage is that one firm can get the other to stop investing. In Section 4 we have shown that this is the case under price competition because a firm's profits from product market competition peak in its own capacity. In our online Appendix we show that cost/benefit considerations can also give rise to a strategic advantage and hence asymmetries. This occurs irrespective of the mode of product market competition if the benefit of adding a block of capacity is rather low or the cost is rather high. The dynamics of the industry, however, hinge on the source of the strategic advantage. In particular, the possibility of gaining a strategic advantage based on cost/benefit considerations does not lead to a preemption race.

□ **Product differentiation.** Up to this point, we have focused on product market competition with homogeneous products. Turning to the opposite extreme of independent goods, it should be clear that asymmetric industry structures can no longer arise: After all, with independent goods, both firms are monopolists, and each firm therefore accumulates enough capacity to supply the monopoly quantity. The question then is: What happens in intermediate cases?

Unfortunately, there is no "off-the-shelf" model of capacity-constrained price competition with differentiated products. In our online Appendix we therefore derive the single-period profit

[19] This is reminiscent of Gelman and Salop's (1983) model. There the entrant makes itself less threatening by credibly committing itself to a lower capacity, thereby inducing the incumbent to maintain an umbrella and to accommodate entry. We differ in a number of ways: First, in our model there is little a firm can do to commit itself to limiting its capacity. Second, their model is about forcing entry into a market, ours about battling for market leadership (and knowing when to withdraw from the battle).

functions from first principles. In doing so, we replace the "hard" capacity constraints of Kreps and Scheinkman (1983), Deneckere and Kovenock (1996), and Allen et al. (2002) with "soft" capacity constraints. Hence, a firm can produce any quantity, albeit at an exploding cost. Soft capacity constraints allow us to assume that a firm cannot turn away customers and is obliged to satisfy all of its demand. This "common carrier requirement" gives rise to a Nash equilibrium in pure strategies in the product market game.

Our computations indicate that the industry evolves toward a symmetric structure if the degree of product differentiation is high. In contrast, industry dynamics resemble a preemption race if the degree of product differentiation is low, and the industry evolves toward an asymmetric structure. These size differences are declining slowly as we move away from homogeneous goods and toward independent goods. In sum, asymmetric industry structures arise and persist as long as products are not too differentiated.

□ **Number of firms.** In contrast to the Kreps and Scheinkman (1983) model of capacity-constrained price competition with homogeneous products, our model of capacity-constrained price competition with differentiated products is easily extended from $N = 2$ to $N > 2$ firms.[20] With $N > 2$ firms, price competition still leads to firms of unequal size. In general, the most likely industry structure consists of one medium-sized and $N - 1$ small firms. Moreover, industry dynamics continue to resemble a preemption race. Once a firm falls behind, it stops investing, whereas its rivals continue to invest. In short, firms drop out of the race one by one, thus propelling the last remaining firm into a position of dominance.

□ **Other extensions and robustness checks.** Entry and exit as well as demand uncertainty can be incorporated in straightforward ways into our model, but they do not significantly affect our results. Similarly, the main insights described above continue to hold as the capacity grid is changed.

## 8. Conclusion

■ In this article we identify circumstances under which asymmetric industry structures arise and persist. We apply the Markov-perfect equilibrium framework presented in Ericson and Pakes (1995) to study the evolution of an oligopolistic industry, and we highlight the mode of product market competition and the extent of investment reversibility as key determinants of the size distribution of firms in the industry.

More specifically, under quantity competition, each firm accumulates enough capacity to supply the Cournot quantities, leading to an industry structure of equal-sized firms independent of whether investment is irreversible (zero depreciation) or reversible (positive depreciation). With positive depreciation, firms tend to hold idle capacity out of a precautionary motive.

By contrast, under price competition, the industry tends to evolve toward asymmetric structures. In particular, if investment is reversible, then the industry moves toward an outcome with one dominant firm and one small firm. Industry dynamics in this latter case take the form of a preemption race in which firms invest heavily as long as they are of equal size. Once one of the firms succeeds in moving slightly ahead of the other, however, the smaller firm gives up, thereby propelling the larger firm eventually into a position of dominance. Along this path, some of the excess capacity that has been accumulated during the preemption race is withdrawn from the industry, which in turn improves the profitability of both firms. Since this capacity withdrawal can proceed at a faster pace as investment becomes more reversible, gaining an initial edge becomes more valuable and consequently the preemption race itself more brutal.

[20] For a discussion of the difficulties involved in extending the Kreps and Scheinkman model of capacity-constrained price competition with homogeneous products to $N > 2$ firms, see Deneckere, Doraszelski, and Kovenock (2003).

## References

ALLEN, B., DENECKERE, R., FAITH, T., AND KOVENOCK, D. "Capacity Precommitment as a Barrier to Entry: A Bertrand-Edgeworth Approach." *Economic Theory*, Vol. 15 (2000), pp. 501–530.

BRESNAHAN, T.F. "Empirical Studies in Industries with Market Power." In R. Schmalensee and R.D. Willig, eds., *Handbook of Industrial Organization*, Vol. 2. Amsterdam: North-Holland, 1989.

BUDD, C., HARRIS, C., AND VICKERS, J. "A Model of the Evolution of Duopoly: Does the Asymmetry between Firms Tend to Increase or Decrease?" *Review of Economic Studies*, Vol. 60 (1993), pp. 543–573.

CAVES, R.E. "Industrial Organization and New Findings on the Turnover and Mobility of Firms." *Journal of Economic Literature*, Vol. 36 (1998), pp. 1947–1982.

DAVIDSON, C. AND DENECKERE, R. "Long-Run Competition in Capacity, Short-Run Competition in Price, and the Cournot Model." *RAND Journal of Economics*, Vol. 17 (1986), pp. 404–415.

DENECKERE, R.J. AND KOVENOCK, D. "Bertrand-Edgeworth Duopoly with Unit Cost Asymmetry." *Economic Theory*, Vol. 8 (1996), pp. 1–25.

———, DORASZELSKI, U., AND KOVENOCK, D. "Bertrand-Edgeworth Triopoly." Working Paper, University of Wisconsin, Madison, 2003.

DOCKNER, E.J. "A Dynamic Theory of Conjectural Variations." *Journal of Industrial Economics*, Vol. 40 (1992), pp. 377–395.

DORASZELSKI, U. AND SATTERTHWAITE, M. "Foundations of Markov-Perfect Industry Dynamics: Existence, Purification, and Multiplicity." Working Paper, Hoover Institution, 2003.

ERICSON, R. AND PAKES, A. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies*, Vol. 62 (1995), pp. 53–82.

FRIEDMAN, J. *Oligopoly Theory*. New York: Cambridge University Press, 1982.

FUDENBERG, D. AND TIROLE, J. "Capital as a Commitment: Strategic Investment to Deter Mobility." *Journal of Economic Theory*, Vol. 31 (1983), pp. 227–250.

GABSZEWICZ, J.J. AND PODDAR, S. "Demand Fluctuations and Capacity Utilization under Duopoly." *Economic Theory*, Vol. 10 (1997), pp. 131–146.

GELMAN, J.R. AND SALOP, S.C. "Judo Economics: Capacity Limitation and Coupon Competition." *Bell Journal of Economics*, Vol. 14 (1983), pp. 315–325.

GHEMAWAT, P. "Capacity Expansion in the Titanium Dioxide Industry." *Journal of Industrial Economics*, Vol. 33 (1984), pp. 145–163.

———. *Games Businesses Play: Cases and Models*. Cambridge, Mass.: MIT Press, 1997.

GORT, M. "Analysis of Stability and Changes in Market Shares." *Journal of Political Economy*, Vol. 71 (1963), pp. 51–63.

GROSSMAN, S.J. "Nash Equilibrium and the Industrial Organization of Markets with Large Fixed Costs." *Econometrica*, Vol. 49 (1981), pp. 1149–1172.

HALL, E.A. "An Analysis of Preemptive Behavior in the Titanium Dioxide Industry." *International Journal of Industrial Organization*, Vol. 8 (1990), pp. 469–484.

HANIG, M. *Differential Gaming Models of Oligopoly*. Ph.D. thesis, MIT, 1985.

HOPENHAYN, H.A. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica*, Vol. 60 (1992), pp. 1127–1150.

JOVANOVIC, B. "Selection and the Evolution of Industry." *Econometrica*, Vol. 50 (1982), pp. 649–670.

KOVENOCK D. AND ROY, S. "Dynamic Capacity Choice in a Bertrand-Edgeworth Framework." *Journal of Mathematical Economics*, Vol. 29 (1998), pp. 135–160.

KREPS, D.M. AND SCHEINKMAN, J.A. "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes." *Bell Journal of Economics*, Vol. 14 (1983), pp. 326–337.

KYDLAND, F.E. AND PRESCOTT, E.C. "Time to Build and Aggregate Fluctuations." *Econometrica*, Vol. 50 (1982), pp. 1345–1370.

MAGGI, G. "Endogenous Leadership in a New Market." *RAND Journal of Economics*, Vol. 27 (1996), pp. 641–659.

MCGAHAN, A. AND PORTER, M. "How Much Does Industry Matter, Really?" *Strategic Management Journal*, Vol. 18 (1997), pp. 15–30.

MITCHELL, M.F. "The Scope and Organization of Production: Firm Dynamics over the Learning Curve." *RAND Journal of Economics*, Vol. 31 (2000), pp. 180–205.

MUELLER, D.C. *Profits in the Long Run*. New York: Cambridge University Press, 1986.

PAKES, A. AND MCGUIRE, P. "Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated-Product Model." *RAND Journal of Economics*, Vol. 25 (1994), pp. 555–589.

PAL, D. "Cournot Duopoly with Two Production Periods and Cost Differentials." *Journal of Economic Theory*, Vol. 55 (1991), pp. 441–448.

PINDYCK, R.S. "Investments of Uncertain Cost." *Journal of Financial Economics*, Vol. 34 (1993), pp. 53–76.

PRESTON, R. *American Steel: Hot Metal Men and the Resurrection of the Rust Belt*. New York: Prentice Hall, 1991.

REYNOLDS, S. "Capacity Investment, Preemption and Commitment in an Infinite Horizon Model." *International Economic Review*, Vol. 28 (1987), pp. 69–88.

———. "Dynamic Oligopoly with Capacity Adjustment Costs." *Journal of Economic Dynamics and Control*, Vol. 15 (1991), pp. 491–514.

——— AND WILSON, B. "Bertrand-Edgeworth Competition, Demand Uncertainty, and Asymmetric Outcomes." *Journal of Economic Theory*, Vol. 92 (2000), pp. 122–141.

SALONER, G. "Cournot Duopoly with Two Production Periods." *Journal of Economic Theory*, Vol. 42 (1987), pp. 183–187.

———. "Investment Strategy and Growth in a New Market." *Bell Journal of Economics*, Vol. 10 (1979), pp. 1–19.

TIROLE, J. *The Theory of Industrial Organization*. Cambridge, Mass.: MIT Press, 1988.

VIVES, X. *Oligopoly Pricing: Old Ideas and New Tools*. Cambridge, Mass.: MIT Press, 1999.