Volume Title: Trade and Protectionism, NBER-EASE Volume 2

Volume Author/Editor: Takatoshi Ito and Anne O. Krueger, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-38668-6

Volume URL: http://www.nber.org/books/ito_93-2

Conference Date: June 19-21, 1991

Publication Date: January 1993

Chapter Title: Forward Pricing versus Fair Value: An Analytic Assessment of "Dumping" in DRAMS

Chapter Author: Kenneth Flamm

Chapter URL: http://www.nber.org/chapters/c8071

Chapter pages in book: (p. 47 - 94)

3 Forward Pricing versus Fair Value: An Analytic Assessment of "Dumping" in DRAMs

Kenneth Flamm

Since the mid-1970s, the concept of sales at a cost less than a constructed "fair value" has become an alternative standard for findings under the U.S. trade laws that imports are being "dumped" in the U.S. market.[1] It has been estimated that, since 1980, about 60 percent of all dumping cases have been based on charges of selling at a price below some constructed average cost (Horlick 1989, 136). Perhaps the most widely publicized application of this standard can be found in the case of imports of the largest category (by value) of semiconductor device sold, dynamic random access memory (DRAM) chips. A U.S. firm's petition for relief from dumping, brought against Japanese 64,000-bit (64K) DRAM imports in 1985, specifically acknowledges that prices for these chips in the U.S. market may actually have been marginally higher than prevailing prices in the Japanese market, the exact opposite of the traditional concept of dumping as sales abroad at less than home market prices.[2] Instead, the U.S. complainant charged that Japanese chips were being sold at prices not covering the full costs of production, the new definition of dumping in the U.S. trade laws.

The investigation of DRAM dumping was expanded by the U.S. govern-

Kenneth Flamm is senior fellow in the Foreign Policy Studies Program at the Brookings Institution in Washington, D.C.

The views expressed in this paper are the author's alone and do not represent those of other staff, officers, or trustees of the Brookings Institution. The author thanks Yuko Iida Frost for her very helpful research assistance and Ann Ziegler for wrestling the mathematical appendices into WordPerfect's equation format. Without implicating them in his errors, the author thanks Dan Hutcheson, Takatoshi Ito, Motoshige Itoh, Anne Krueger, Douglas Nelson, Gary Saxonhouse, and Yun-wing Sung for helpful comments and suggestions.

1. A brief history of the origins of this new standard may be found in Nivola (1990, 229–30) and Horlick (1989, 133–34).

2. See the Micron Technology (1985, 11–14) petition to the U.S. Commerce Department and the U.S. International Trade Commission.

ment to include 256K and 1 megabit (1M) DRAMS and, folded into investigations of U.S. industry charges of dumping of erasable programmable read only memory (EPROM) chips, ultimately culminated in the controversial U.S.-Japan Semiconductor Trade Arrangement (STA) of 1986. One of the outcomes of the STA was a system of floor prices for Japanese DRAM and EPROM imports administered by the U.S. Commerce Department, based on the calculation of something called "foreign market value" (FMV), derived from the "fair value"–constructed cost comparisons enshrined in the dumping provisions of U.S. trade law.

Although the FMV calculations have been dropped from the 1991 successor to the STA, Japanese producers are required to continue to collect the same data, in order to facilitate a "fast response" dumping investigation. Thus, one may surmise that the implicit threat of a dumping investigation continues to give the FMV calculation a significant—if shadowy—role in determining lower bounds on pricing of Japanese chip exports to U.S. (and possibly third-country) markets.

While the idea of requiring producers always to maintain a price at or above some concept of full long-run average cost is hard to defend, either as a positive description of what a profit-maximizing producer in a "competitive" market would choose to do or as a normative guide for efficient resource allocation, it is possible to construct an economically coherent argument that pricing below *marginal* cost can serve as a warning signal of "strategic" behavior by producers that in some circumstances can justify policy intervention by the government. However, in an industry subject to so-called learning economies (where unit production cost falls with cumulative production experience), it is possible that producers may rationally choose to "forward price," that is, sell at a price below current marginal cost, for completely "competitive," nonstrategic reasons.

Is below-marginal-cost pricing for nonstrategic reasons empirically relevant in the semiconductor industry? Is it reasonable to defend even some revised version of a constructed cost test for dumping, based on a constructed *marginal* cost, as a reasonable trip wire for government scrutiny of possible strategic behavior by foreign producers? Perhaps the most interesting question is, What can we deduce about the relation between price and production costs using a minimally realistic model of the product life cycle when large up-front investments in capacity constrain output, large and relatively fixed investments in research and development (R&D) create economies of scale, and learning economies are likely to be significant? How is an "FMV-like" system likely to constrain producer behavior in these circumstances? Because these characteristics are typical not just of semiconductor manufacture but of a broad range of high-technology products, the answers to these questions, and the methodology used in the inquiry, are of some importance.

This paper is intended to provide a simple analytic framework that can be used to compare the time path of output and prices in a nonstrategic, competitive (open-loop Cournot-Nash equilibrium) semiconductor industry with various variants of constructed "fair value" that would be associated with the same path for output. The model is applied with empirically based parameters associated with 1M DRAM chip production, in order to explore how pricing of semiconductors is likely to be constrained, over the product life cycle, by constructed values—FMV-type pricing rules. The basic model should also prove useful in analyzing many other interesting questions about the potential effect of public policies affecting high-technology industries with scale and learning economies.

3.1 Economic Rationality of Below-Marginal-Cost Pricing

To a first approximation, stripped of a variety of practically important cost allocation and accounting issues, the U.S. Commerce Department's procedures for constructing FMVs resemble an economist's concept of average cost of production, plus a fixed 8 percent markup that ostensibly reflects "normal" profit. (This completely arbitrary 8 percent markup is ignored in further discussion.) A result taught in most any introductory economics course is the fact that, under some circumstances (e.g., a downturn in demand), it can be economically rational for a producer in a competitive industry to sell at a price less than full average cost, just as long as short-run marginal cost is covered by price. As long as a firm at least covers the variable costs of running a production line and the marginal cost of producing an incremental unit on that line, it makes economic sense to continue operating a factory, even if revenues received are insufficient to recover the full historical cost of an initial investment in developing the product and building the factory.

Thus, most economists would find it entirely normal that over at least some periods, observed prices would fall short of the full (long-run) average cost of production.[3] Any policy measure that prohibits marginal cost pricing by foreign exporters, while leaving domestic producers unaffected, will—if it actually affects market outcomes—increase domestic production at the expense of domestic consumers and (possibly) foreign producers. It will also arguably deny foreign producers national treatment, forbidding them the right to economic behavior permitted domestic firms.

If price falls short of the full average cost of production continuously, of course, one may safely predict that some firms will exit the industry and that the industry will shrink to the point that full average costs are at least recovered over the life of sunk investments by the remaining firms. Thus, if sustained constructed "fair value" dumping (pricing below long-run average cost) is observed in a competitive industry, one may generally infer that excess ca-

3. For further elaboration on this point, see, e.g., Deardorff (1989, 30–33).

pacity exists and that exit will follow. But observed pricing behavior may still reflect "normal," competitive behavior on the part of the firms pricing below FMVs.

Can one imagine any economic justification for remedial policies triggered by selling below a constructed FMV? The point at which many economists would agree that something other than "competitive," nonstrategic behavior might be suspected is when a firm's price falls short of its short-run marginal cost or, even more obviously, average variable cost (which bounds short-run marginal cost from below over the relevant range).[4]

In considering why a firm might rationally choose to produce and sell a product at a price not covering the current marginal cost of production, it is helpful to distinguish between "strategic" and nonstrategic behavior. I shall label a firm's behavior "strategic" when it explicitly takes account of effects of its decisions on the behavior of other economic agents. This contrasts with what I will call "nonstrategic" behavior, decisions taken considering the actions or choices of other agents as fixed, unaffected by one's own.

One possible explanation for producers pricing below marginal cost, consistent with nonstrategic behavior, is that current production may lower a firm's future production costs. In this case, measured current marginal cost overstates "true" marginal cost, which should take into account the cost-reducing effects of current production on future output.[5]

But another possible explanation for behavior of this sort is a strategic motive on the part of the "dumper": either predation (actions intended to encourage other firms to exit from the industry), limit pricing (intended to discourage entry by others), or a defensive response against predatory behavior by others.[6] In this case, the rents received from the exercise of monopoly power later must be forthcoming to justify absorption of a temporary loss on output shipped now.

Many forms of strategic behavior by firms, like predation, are regulated within a national market by antitrust laws. Thus, a constructed cost test, used

4. For a detailed survey of the literature on tests for predatory behavior, see Ordover and Saloner (1989, 579–90).

5. Note that such learning economies can also be used as a strategic instrument, with a firm's production decisions taking into account the effect of its learning on the actions of its rivals. For such a model, see Fudenberg and Tirole (1983). Deardorff (1989, 37–38) points out that low-priced sales designed to build brand loyalty or otherwise alter consumer preferences might also rationally lead a producer to sacrifice current profitability for future rents and price below marginal cost. In effect, greater current output shifts future demand schedules, and current marginal revenue understates "true" marginal revenue. Such "demand-side learning effects," however, may be considered a form of "strategic" behavior since they are designed to alter the behavioral response to price of other economic agents (i.e., consumers).

6. The modern rehabilitation of the theory of predation focuses on its effect on rival firms' expectations about future profitability: as an exit-inducing investment in "disinformation" about the predator's cost structure, e.g., or as the consequence of asymmetric financial constraints among competing firms created by imperfections in capital markets. The basic references are Milgrom and Roberts (1982) and Kreps and Wilson (1982); useful interpretations are found in Milgrom (1987) and Tirole (1988, 367–80).

in the framework of the dumping laws, might be interpreted as a second-best attempt to remedy behavior by foreign firms that, if carried out on a purely domestic basis, would be considered the domain of antitrust policy. Lacking the ability to impose domestic policy standards on a foreign firm's behavior outside the national market, a national government can instead impose controls on the manifestations of that behavior—that is, pricing of sales to importers—in the domestic market.

Since, absent learning effects, pricing below short-run marginal cost is sufficient (but not necessary)[7] to conclude that a firm is acting strategically in its pricing policies, it may seem reasonable at that point to review its activities and to take corrective action if the intent is deemed to be predation and the potential effect significant. It is at least possible that increased monopoly rents paid out later by national consumers to foreign producers, and deadweight losses, could more than offset the windfall to national consumers created by a temporary episode of low import prices, justifying some policy intervention (as noted by Deardorff [1989, 35–36]).

From this point of view, the economic problem with cost-based definitions of dumping is not necessarily their existence but their use of the wrong cost concept (long-run average cost instead of short-run marginal cost) as the prima facie trigger for consideration of possible intervention. This perspective also leads one to focus on the close relation between "fair trade" laws and competition and antitrust policy. It might be argued that some binding international standards for competitive business behavior (and their enforcement) might be offered as a constructive alternative to national fair value dumping tests based on constructed costs, as remedies for predation.

This paper will not attempt to evaluate whether predation is a plausible description of what was going on in the DRAM marketplace in the 1980s. I merely note that predatory behavior was one of the allegations made by the U.S. industry in pressing its case for protection. However, the modern theory of predation has been interpreted to suggest that high-technology industries are particularly important places to look for such behavior.[8]

3.2 Costs and Pricing in the Semiconductor Industry

High-technology industries, facing large sunk costs in research and development relative to sales, along with highly capital-intensive industries, are by nature particularly prone to trade friction involving charges of dumping based

7. Criticism of a short-run marginal cost test for predation generally argues that the rule is not stringent enough; prices above short-run marginal cost may still be associated with socially costly predatory activity (see Tirole 1988, 372–3; Ordover and Saloner 1989, 579–80).

8. Paul Milgrom argues that "policymakers should be especially sensitive to predatory pricing in growing, technologically advanced industries, where the temptation to discourage entry is large, and the costs of curtailed entry even larger" (1987, 938). For further consideration of the plausibility of strategic behavior in semiconductor competition, see Flamm (in press, b).

on constructed cost tests. When fixed investments in R&D or factories are very large in relation to a firm's sales, a significant gap between average variable cost and long-run average cost will exist, and short-run marginal cost may fall significantly below long-run average cost for a substantial range of economically rational output levels. In such a case, perfectly competitive behavior may often trigger pricing below long-run average cost—and dumping charges—in a downturn.

High-tech industries are also particularly prone to dumping cases because of the peculiar way in which R&D investments are treated by trade law (and many companies') accounting principles. An investment in a capital facility, for example, is not charged immediately against company revenues when construction is begun, or completed, but spread over the period in which it is to be used through the use of depreciation charges. One may argue that accounting depreciation is at least an attempt to approximate the profile of true economic depreciation charges. An R&D investment, by way of contrast, is generally charged against revenues at the moment it is incurred, not spread over its economically useful life.

It is sometimes argued that, when processed through constructed cost calculations, this "front loading" of R&D leads to artificially high prices for high-tech imports (like DRAMs) when initially shipped, in effect retarding technological progress. Defenders of this practice argue that, since R&D charges are often allocated on the basis of sales rather than identified with some particular product, the practical effect is to spread R&D charges over generations of products, through time (although it clearly remains true that a company just entering an industry after making a fixed R&D investment will necessarily have to charge an initially high price).

The semiconductor industry is both technology intensive and capital intensive: it spends almost 15 percent of sales on R&D; it also typically spends an even larger fraction of sales (15–20 percent annually) on capital investments. Demand for semiconductors is also notoriously cyclical, and it is not, therefore, surprising to find that constructed cost tests for dumping were invoked in the 1985 industry downturn.

In addition, semiconductor production is believed to be characterized by so-called learning economies. Unit production costs are believed to fall sharply with accumulated production experience. This further complicates our discussion of the borderline between nonstrategic pricing behavior and strategic activities. The key result due to Spence (1981) is that, with learning economies but no strategic interactions with its rivals, a rational firm will generally equate marginal revenue to a value below its current short-run marginal cost of production, as it takes into account the cost-reducing effect of current production on future production costs.

While the Spence model is rather unsuitable for analyzing production decisions in the semiconductor industry, the point it makes greatly complicates the issue of whether constructed cost-dumping tests—amended perhaps to use short-run marginal cost rather than long-run average cost as the trip wire for possible intervention—can be justified as a reasonable safeguard against predation by foreign producers. For in the Spence model, even with nonstrategic behavior—that is, with a firm taking production decisions by competitors as given, independent of its actions—economically rational firms will engage in "forward pricing," that is, choose output levels where marginal revenue lies below their current short-run marginal cost.

3.3 Modeling the Semiconductor Product Life Cycle

In my somewhat stylized depiction of the industry, a DRAM producer will be assumed to produce a homogeneous commodity, perfectly substitutable for that of other producers.[9] Difficult issues concerning the timing of the switchover from one generation of DRAM to another, and intergenerational externalities, are ignored by assuming that a DRAM producer faces a fixed period over which the DRAM is sold and that costs for developing and producing his product are relevant to that generation of DRAM alone. The product life cycle begins at time 0 and ends at time 1 (hence, the unit of time is the "product life cycle"). Every producer faces revenue function R, giving total revenues at any moment t as a function of his own production, $y(t)$, and the aggregate output of all other producers, $x(t)$. All revenues and costs are measured in constant dollar terms. Following Spence, for simplicity, I ignore discounting on the grounds that product life cycles are short (typically, a new generation of DRAM is introduced every three years) and the additional complexity introduced by discounting over time substantial.

In semiconductor production, plant capacity may be measured in terms of "wafer starts," the number of slices of silicon, on which integrated circuits are etched, that can be processed per unit time. At any moment t, $w[E(t)]$ functioning chips are yielded per wafer processed, where w is an increasing function of $E(t)$, "experience" through time t. How one defines relevant "experience" is a subject explored below. I will parametrize the effect of output, y, on relevant experience, E, as

$$\dot{E} = \frac{dE}{dt} = \frac{y}{K^{\gamma}},$$

where K is capacity, and γ is a parameter taking on a value between zero and one. For notational simplicity, time will sometimes be suppressed as an argument of time-varying variables.

Some of the variable cost of producing a chip is incurred with every wafer processed, and some of the cost is incurred only with good, yielded chips (assembly and final test, e.g.). If a wafer-processing facility is utilized at rate

9. This is not an unreasonable approximation. For more detailed discussion of this issue in the context of semiconductor price indexes, see Flamm (in press, a).

$u(t)$ (u between 0 and 1), total variable costs at any moment are $dy + cuK$; d is manufacturing cost per good, yielded chip, c processing cost per wafer start. Note that $y(t, K) = w[E(t)]u(t)K$.

Up-front, sunk costs independent of output levels (like R&D) are equal to F, and fixed capital investment costs required for a facility processing K wafer starts are equal to r per wafer start. For the moment, take K (wafer-processing capacity) as a given. The producer's problem is to maximize

$$
\begin{aligned}
&\max_{u(t),K} \int_0^1 \{R[x(t), y(t)] - dy(t) - cu(t)K - rK\}dt - F, \\
(1) \qquad &\text{with} \quad y(t) = w[E(t)]u(t)K, \\
&\text{s.t.} \quad \dot{E} = \frac{y}{K^{\gamma}} = w[E(t)]u(t)K^{1-\gamma}.
\end{aligned}
$$

Firms will be assumed to simultaneously choose initial capacity investments K and a time path for utilization rates, which give rise to a path for output over time, which they then proceed to follow. My assumption that capacity investments in DRAMs are committed at the beginning of the product cycle is not terribly unrealistic: it typically takes a year or more to get a new fabrication facility up and running, and a new generation of DRAM is introduced roughly every three years.[10]

For the moment, take γ to equal zero (i.e., absolute cumulative production is the relevant measure of experience). The model that I present in appendix A will, like the Spence model, assume a Nash equilibrium in output paths; that is, given rivals' actual choices of output paths, (1) is maximized by every firm. Firms' behavior in this static game is *nonstrategic* since they take their rivals' output choices as given.[11]

3.3.1 Spence's Model

If wafer-processing capacity K is not fixed over the life cycle but is continuously variable, as is implicit in Spence's formulation, then we have a special case of the above model in which r is zero (capital costs are included in wafer-

10. The world record for bringing a new fabrication facility on line seems to be held by NMB Semiconductor, which claims that it took only nine months to go from initial ground breaking for a new factory to initial production of 256K DRAMs in 1985 (see Waller 1988).

11. An alternative would be to set up a two-stage competition among rival firms, with capacity investment as the initial phase, followed by a second stage in which firms choose output paths subject to capacity constraints. The solution of the static game presented here corresponds to the open-loop (nonstrategic) equilibrium of this two-stage game, in which a firm's first-period choice of capacity takes its rivals' choices in both periods as given. An alternative equilibrium concept would assume second-period subgame perfectness, i.e., that firms take into account the effect of their first-period capacity choices on their rivals' second-period output paths. This creates *strategic* interactions among firms (see Dixit 1986, 114; Shapiro 1989, 383–86). Flamm (in press, b) extends the model presented in this paper to include strategic capacity investments by firms. Note that, despite its nonstrategic nature, the present model is developed using the conjectural variations framework—which (somewhat controversially) permits strategic behavior—to allow greater generality in its application and to preserve comparability with the Baldwin-Krugman model.

processing cost c and some arbitrary initial scale for capacity K is set), capital is a completely variable input, and a producer is free to choose any nonnegative u—that is, u is unbounded above, not bounded by one—and produce any yielded chip output desired. Under these circumstances, as is easily shown in appendix A, formal maximization of objective function (1) yields the first-order condition

$$(2) \qquad R_y = d + \frac{c}{w} - \frac{\delta}{K^{\gamma}};$$

that is, u is chosen so that marginal revenue is set equal to current marginal cost ($d + c/w$) less a term proportional to nonnegative adjoint variable δ, which captures the future cost-reducing effects of current production. Adjoint variable δ, in turn, is determined by the transversality condition,

$$(3) \qquad \delta(1) = 0,$$

and equation of motion,

$$(4) \qquad \dot{\delta} = -\frac{c}{w}uKw_E.$$

By differentiating both sides of equation (2) with respect to time, we immediately see that marginal revenue, R_y, must be constant over time and therefore, by (3), equal to current marginal cost at the end of the product cycle, $d + c/w[E(1)]$.

In short, with continuously variable capacity, a profit-maximizing producer will choose his output so that marginal revenue equals his terminal (not current!) marginal cost. This is so-called forward pricing. With a constant elasticity and autonomous demand, a constant price proportional to terminal marginal cost will result.

Now this does not necessarily mean that price falls below current marginal cost since price will in general exceed marginal revenue. Whether constructed "fair values" based on current marginal cost will serve as binding constraints on pricing will depend on many factors, including market structure and the elasticity of industry demand.

While this model provides an appealing explanation of the phenomenon of forward pricing, a notable empirical feature of business practice within the semiconductor industry, the actual trajectory of pricing suggested by this model (with a constant elasticity demand, price is fixed at some constant level over the entire product cycle) is quite inconsistent with observed behavior.[12] Chip prices typically drop very quickly over the first part of the product cycle, drop less quickly as the product approaches maturity, and fall very slowly, if

12. Dick (1991), e.g., invokes the Spence model to motivate his assumptions about the time path of semiconductor prices over the product life cycle but ignores the constant pricing prediction of the Spence model.

at all, at the end. As shall be seen in a moment, a more realistic treatment of capacity constraints yields a more plausible trajectory for prices.

3.3.2 The Baldwin-Krugman Approach

The pioneering attempt to incorporate learning economies into a stylized, empirical model of the semiconductor industry is that of Baldwin and Krugman (B-K) (1988).[13] The B-K focuses on regional segmentation of the U.S. and Japanese semiconductor markets, in order to simulate the effect of market closure policies, and takes an approach to producer behavior that differs significantly from that of Spence. B-K constrain firms to operate at full capacity over the entire product cycle; the choice variable for the firm is initial capacity, which, once set, determines output levels over the entire product life cycle. The first-order condition for an optimum is that the life-cycle revenue created from the addition of a marginal unit of wafer-processing capacity just equals the cost of building and operating that marginal unit of wafer-processing capacity (since all capacity is always fully utilized, the distinction I am drawing between investment costs and wafer-processing costs is immaterial).

Firms in the Spence model are never capacity constrained; firms in the B-K model always operate at their capacity constraint. The Spence model has firms forward pricing—maintaining marginal revenue constant over the life cycle, equal to their terminal marginal cost. The B-K model has marginal revenue—and price—falling smoothly over the life cycle. Thus, while the striking forward pricing behavior of the Spence model has disappeared, a more empirically plausible path for prices has replaced it.

As Krishna (1988) notes, however, the algebraic tractability created by the simplicity of the B-K specification of firm behavior has been purchased by excluding the possibility of some interesting forms of strategic competition. (Because B-K empirically calibrate conjectural variations, strategic interactions among firms exist.) Investments in capacity may be undertaken with strategic objectives, to convince rivals to exit or dissuade them from entering an industry, creating additional monopoly power that can then be exploited. Constraining firms to operate at full capacity over the entire product life cycle may restrict them to suboptimal output paths, where monopoly power is not fully exploited. It also hinders analysis of interesting policy questions regarding the potential welfare effect of strategic government policies that may foster the creation and exercise of monopoly power.

A variant of the B-K model can be fit into the framework outlined above for the Spence model, after suitable amendments. Utilization rate u is constrained

13. A somewhat different exposition of this model is given in Helpman and Krugman (H-K) (1989, chap. 8). This later interpretation differs in some significant respects from B-K. For example, the learning curve in B-K has yields improving with cumulative wafers processed (i.e., faulty chips have the same yield-enhancing effects as good ones), while H-K presents a more conventional view of the learning curve, with yield rates rising with cumulative output of *yielded* (i.e., good) chips. While the B-K assumption on yields is not the accepted approach to modeling yield improvement within the industry, it simplifies the mathematical structure of the model.

to equal one at all times, and objective function (1) is maximized with respect to K alone The right-hand side of equation (4) is replaced by the more complex variant shown in appendix A (corresponding to $u = 1$), and a new question determining optimal capacity choice is added:

$$(5) \qquad \int_0^1 \left[\left(R_y - d - \frac{c}{w} + \frac{\delta}{K^\gamma} \right) uw - \gamma \frac{\delta}{K^\gamma} uw - r \right] dt = 0,$$

where the B-K specification fixes u equal to one and γ equal to zero.

3.3.3 A More Realistic Model of the Semiconductor Product Cycle

It is possible to create a more realistic model of firm behavior, in which firms can continuously adjust output, as in the Spence model, yet also face capacity constraints on output, as in the B-K model.

I briefly summarize the more detailed exposition laid out in appendix A to this paper. The firm's problem is to maximize (1) by choosing both an initial level of capacity K and time-varying utilization rates $u(t)$ for that capacity that determine output at any moment in time. The optimal level of capacity chosen satisfies equation (5) above; the left-hand side of this equation can be interpreted as the net marginal return on additional investment in capacity. It also must be true that the optimal path must be capacity constrained over some interval (i.e., $u[t] = 1$).

In general, the optimal path for $u(t)$ will be made up of three types of segments: interior segments, where $0 > u > 1$; lower-boundary segments, where $u = 0$; and upper-boundary segments, where $u = 1$. Within an interior segment, equations (2) and (4) will hold, as in the Spence model, as will a form of forward pricing: marginal revenue will be held constant, set equal to current marginal cost less δ/K^γ—the marginal cost-reducing value (over the remainder of the product life cycle) of an additional unit of output—at the endpoint of this interval.

With additional assumptions, one can further sharpen the characterization of the optimal behavior of a profit-maximizing firm. I shall assume a symmetric industry equilibrium with N identical firms, an autonomous demand (i.e., not an explicit function of time), and concavity of total industry revenues in industry output (as would be the case, e.g., with a constant elasticity demand function and price elasticity exceeding unity). Although a nonstrategic, Nash equilibrium in output paths is assumed for the remainder of this paper, for expositional purposes I will parametrize a firm's perceptions of other firms' reactions to changes in its output in terms of a constant, nonnegative conjectural variation. (The two interesting cases that motivate this parametrization are Cournot-Nash equilibrium [conjectural variation equal to zero] and a collusive, constant market share cartel [conjectural variation equal to $N - 1$.)[14]

14. The major behavioral assumption excluded by a nonnegative conjectural variation is Bertrand competition in prices. Because DRAMs are essentially a homogeneous commodity sold in

Under these assumptions, optimal u must decline over an interior segment, and u must be continuous in time. Therefore, the optimal path of the utilization rate must look like an upper-boundary segment, possibly followed by an interior segment, possibly then followed by a lower-boundary segment. Along lower-boundary segments, where $u = 0$, δ will be constant and therefore equal to its terminal value. Thus, the Spence forward pricing result of marginal revenue being set equal to terminal marginal cost will hold whenever we are producing but are not capacity constrained (i.e., $0 < u < 1$, along an interior segment).

If we further assume that firm marginal revenue exceeds the initial value of current marginal cost (so some production will always be profitable) as industry output approaches zero (as must be the case with a constant elasticity demand), we can exclude the possibility of lower-boundary segments occurring along the optimal path. Note that nothing about the specific shape of the learning curve (function w) beyond the fact that it is increasing in experience ($w_E > 0$) has been assumed in arriving at this characterization of optimal policy.

In short, with this simple description of the semiconductor product life cycle, we derive a more realistic specification of firm behavior that captures both the importance of capacity investments and the ability of firms fully to exploit what monopoly power they enjoy by varying utilization rates over time. It is simple enough to be empirically tractable. Firms will make some capacity investment, run at that capacity full blast for some period of time, then possibly switch to a constant output path (with constant marginal revenue but decreasing utilization of capacity as yields rise) over the remainder of the product life cycle.

3.4 Preliminary Observations

Even with its relatively general structure, the analysis presented above provides a couple of insights into the question of "dumping" over the product life cycle. First, below-current-marginal-cost pricing will *never* be observed near the end of the product cycle among competitive, nonstrategic, profit-maximizing firms. This follows immediately from the fact that, if any output is being produced, marginal revenue will never be less than the right-hand side of equation (2) (see app. A), which at time 1—the end of the product cycle—equals current marginal cost ($d + c/w$). Since price exceeds marginal revenue, price must also exceed current marginal cost in some neighborhood of time 1, the end of the product cycle.

well-developed secondary spot markets, specifying that producers sell at a single market price and choose quantities sold is the natural assumption. Moreover, Kreps and Scheinkman (1983) have shown that, in a two-stage game, where first-stage capacity investments are followed by a second-stage Bertrand game in prices and a particular ("efficient") rationing rule, the outcome is a Cournot equilibrium in output.

3.4.1 Closing the Model

Can we say anything about the relation between price and long-run average cost? The model sketched out thus far takes the number of firms in the industry—which will affect profitability and pricing—as given. One "natural" way to close the model is to specify that firms enter the industry until rents earned by producers, that is, the integrand in equation (1), just equal zero. The zero-profit condition then determines N, the number of firms entering the industry (I will ignore the difficulties created by insisting that N be an integer).

Zero profits mean that total life-cycle revenues just equal total life-cycle costs. Therefore (after dividing both concepts by total output over the product life cycle) average "life-cycle" price must equal average "life-cycle" cost per unit.

But is there any clear relation between current price and current "fully allocated" average cost at any given moment? Current short-run marginal cost (SRMC) is a relatively clear concept: the additional current cost saved by producing one less unit at any given moment. This is the incremental cost saved when output is reduced by one unit.[15] In my model, current short-run marginal cost—$d + c/w$—is constant at any moment and equal to current average variable cost.

To define a current average cost, however, it is first necessary to define an intertemporal cost allocation rule to spread fixed entry costs F and capital costs rK over the product life cycle. Dividing the capital and entry costs allocated to some instant in time by output produced at that moment yields a current average fixed cost per unit produced. If this current average fixed cost is added to current average variable cost (identical to short-run marginal cost in my model), we have a long-run average cost (LRAC) concept that satisfies the basic requirements of a long-run average cost: when multiplied by output at that moment and summed over all moments, total costs of production over the entire product life cycle are given.

Now, because my assumption about entry means that total life-cycle costs are exactly equal to total life-cycle revenues, price less the fully allocated long-run average cost defined above (i.e., profit per unit), multiplied by output, and summed over every moment of the product cycle must be exactly equal to zero. Thus, if price exceeds the fully allocated average cost concept at any instant, it must fall below fully allocated average cost at some other instant over the product cycle, and vice versa. Therefore, if my assumption that entry drives long-run profits to zero is a realistic one, below-LRAC "dumping" *must* be occurring sometime during the product cycle, unless the cost allocation rule for fixed costs defines an average fixed cost that, when added to current average variable cost, is exactly equal to actual price *at every moment.*

15. When a firm operates at less than full capacity, this is identical to the increased cost incurred in producing one more unit. When operating at full capacity, the incremental cost of an additional unit is effectively infinite; marginal cost is "L-shaped" with a kink at full-capacity output.

It is easy to see that, for a cost allocation rule to satisfy this requirement, with learning economies present, it must generally be a function of *all* the parameters of the control problem and will, in general, take on negative values as well as positive values. Since the cost allocation rules actually used to spread fixed costs over the product cycle—by firms or by the U.S. Commerce Department—are generally functions only of the size of the fixed costs and time and produce only nonnegative values, it is essentially guaranteed that there will be an episode of below-LRAC dumping if learning economies are present and the industry is in a symmetric, zero-profit equilibrium.[16]

3.5 Some Further Assumptions

My next step is to take this simple control model and solve it to explicitly derive an individual firm's behavior over time. Let the time at which a firm switches from full blast production to constant output production be t_s (with full blast production over the entire product life cycle an important possibility). To sharpen my characterization of a profit-maximizing firm's optimal policy, I must address some additional issues.

16. Define a cost allocation rule $g(Z, t)$, where Z is a vector of arguments, t is time, such that

$$\int_0^1 g(Z, t)dt = F + rK.$$

Define fully allocated average cost (FAAC) by

$$\text{FAAC} = \frac{c}{w} + d + \frac{g}{y},$$

i.e., current average variable cost plus average fixed cost. We know that the optimal path must contain a capacity-constrained segment and that along this portion of the optimal path

$$\dot{P} = P'N\dot{y} = P'NKw_E\dot{E}$$

in symmetric industry equilibrium. If price P is always to equal FAAC along this segment, however, differentiating the expression for FAAC with respect to t, and setting this equal to the last expression, we must at every moment of this interval have

$$\begin{aligned}\frac{dg}{dt} &= w_E\dot{E}(P'NK + \frac{c}{w^2} + \frac{g}{Kw^2})y \\ &= Kw_E\dot{E}[(1 + \beta)P - d].\end{aligned}$$

Since equilibrium N, and therefore P, will generally be functions of all the parameters of the optimal control problem, a function g that satisfies this last equation must generally include all parameters of the control problem as arguments, unless $w_E = 0$, in which case g is constant. (In this latter case, I note at the end of app. A that all capacity is utilized and output is constant over the entire product life cycle.) Thus, if there are learning economies (w_E not equal to zero), a cost allocation rule g varying only with F, r, K, and t cannot satisfy the requirement that P = FAAC, for arbitrary values of the parameters of the control problem, over this capacity-constrained interval.

Also, we have already noted that it is possible for P less than current marginal cost to be optimal in the presence of learning economies. (Indeed, the simulations reported below contain examples of such behavior.) Reexamining the definition of FAAC, it is clear that g must be negative for P = FAAC to hold true over such an interval.

3.5.1 Learning Economies

I shall approximate the learning curve by specifying that

$$w(E) = \phi E^{\varepsilon}, \quad \text{with } E(0) = E_0, 0 \leq \varepsilon \leq 1.$$

This gives yielded chips per wafer as a function of experience, E. This functional form is best regarded as an approximation: mass production typically starts at initially low yields; yields then rise quickly and flatten out at the end of the product cycle in a pattern closer to a logistic curve. Analytic tractability is the grounds for selecting this approximation. Note that a "dummy" value E_0 is used as an argument in the function to specify some initial nonzero yield—without this constant, yields would stay "stuck" at zero forever.[17]

This approximation to the "true" learning curve is shown in figure 3.1. If (as is believed in the industry) the "true" learning curve behaves more like a logistic function in its early stages, my approximation somewhat distorts yields, output, and pricing in the very earliest portion of the product cycle.

Defining *experience* raises additional issues. It is customary to use cumulative output as a proxy for experience in empirical studies, and most published empirical studies of learning economies have taken this approach. But using absolute, company-wide production experience as the determinant of any single facility's productivity implies that running, say, ten facilities in parallel produces the same yields at the end of a period as running a single facility to produce the same output over a much longer period. In the semiconductor industry, it is widely believed that improved manufacturing yields come from two main sources—iterative refinements of the operation of the production line (with each new refinement building on previous experience) and "die shrinks" (reductions in the feature size for chip designs made possible by improved use of existing process equipment)—that are iterative and sequential in nature. That is, lessons learned from running a line over some period of time are then applied to refine the operation of that line over a subsequent period.

However, by this logic, if numerous identical production lines are run in an identical fashion over the same period of time, then the same "lessons" are being learned, in parallel, on each line, and yields at the end of the period should be no higher than if only a single line were being run. Of course, if a new line (one with less experience and lower yields) were put into operation after an older line had been running for some time and it were possible completely to transfer the fruits of greater experience across facilities, then the maximum experience on any one line would be the "experience" variable determining production yields. Because all investment occurs at a single initial moment in my simple model, all lines will have identical amounts of produc-

17. B-K use the same functional form but do not face the "stuck" yield problem because the argument in their learning curve is gross wafers processed, not net good chips yielded. The latter specification is generally industry practice in estimating learning curves.

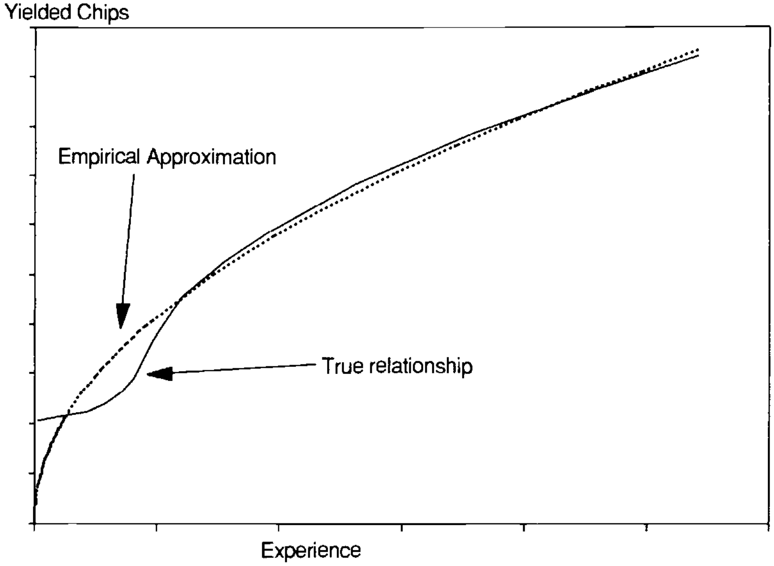

Fig. 3.1 Empirical approximation to hypothetical true yield curve

tion experience at any subsequent moment in time, and cumulative output per facility is the desired measure of experience.

It is possible that the lessons learned on different lines are not the same ones if completely different "experiments" in production refinement are being conducted at every production facility. If, once again, experience can be completely transferred across facilities and there is no duplication in "lessons learned" in different facilities, then it might be argued that company-wide, absolute cumulative output, rather than cumulative output per unit capacity, is the relevant experience variable.[18]

One way to parametrize these differences in the conceptualization of how learning economies work is to define *experience* as cumulative output divided by K^{γ}, where γ takes on value 0 if absolute, company-wide cumulative output is the correct experience variable, 1 if experience per facility (or unit capacity) is what is relevant. This means that

$$(7) \qquad \dot{E} = \frac{y}{K^{\gamma}} \quad \text{with some initial } E(0) = E_0$$

defines $E(t)$.[19] My approach will be an agnostic one: I will solve the model using both 0 and 1 as possible values for γ and then ask which seems to predict

18. Or perhaps even industry-wide cumulative output, if complete cross-company diffusion of the lessons of production experience occurs.

19. Note that an alternative specification might make cumulative output, or cumulative output per unit capacity, the state variable, subject to some initial value, and make this alternative state

more empirically plausible behavior. While the "true" value almost certainly lies somewhere between these two extremes, it is my prior belief that it should be substantially closer to one. With plausible empirical assumptions, it turns out that parameter γ plays a critical role in defining the nature of an industry equilibrium.

Note that the existing empirical literature on learning curves gives us little help in deciding the correct specification. If data on cumulative output from a given facility, or aggregate data from a group of facilities with fixed capacities, are used to estimate the relation $y = wK$ using (6), we get an equation like

$$\ln[y(t)] = a + \varepsilon \ln[Q(t)],$$

that is, giving the natural log of total output as a linear function of the natural log of cumulative output Q, even if cumulative output per unit capacity is the relevant experience variable. The effects of capacity size, K, have been absorbed into constant a. Data from different facilities of varying size within a single company, or from different companies, along with an additional variable controlling for capacity size, are required to identify and estimate γ.

An exact solution for $E(t)$ will be useful in what follows. Substituting (6) into the above differential equation giving $\dot{E}$, and solving for E as a function of time (assuming capacity-constrained output),

$$E(t, K) = [E_0^{1-\varepsilon} + K^{1-\gamma}\phi t(1 - \varepsilon)]^{1/1-\varepsilon} \tag{8}$$

describes the time path of $E(t, K)$ through time t_s, the endpoint of the period of full blast production.

3.5.2 Final Test and Assembly Yields

A tested, just-fabricated "good" die is not yet a finished integrated circuit (IC). The dice produced on the wafer fabrication line must then be assembled into a sealed package, then subjected to a rigorous final testing process. While yields of good, tested chips assembled from "good" dice may also show some evidence of a "learning curve," the effect of learning in this stage of the IC production process is thought to be quite small relative to learning economies in the wafer fabrication phase of IC manufacturing.

I will model assembly and final test yields by assuming a fixed yield of final good chips from "good" dice produced on the wafer fab line; that is, $v = \xi y$; where v is "net" good, assembled and tested ICs produced from quantity y of "gross" good dice yielded by wafer fabrication. If we denote DRAM consumers' inverse demand function for finished chips by $\tilde{P}(\xi x, \xi y)$, then the produc-

variable times K to some power the argument of w, the function giving yield per wafer. Such a specification, however, makes initial yield (with no experience) a function of the scale of capacity investment, which is undesirable. (In that case, increasing or decreasing capacity simply to raise initial yield on every line will play an entirely artificial role in determining optimal capacity.)

er's maximization problem, taking into account assembly and final test yield losses, is to

$$\max_{u(t),K} \int_0^1 \tilde{P}[\xi x(t), \xi y(t)]\xi y(t) - \frac{d}{\xi}\xi y(t) - \frac{c}{w\xi}\xi y(t) - rK,$$
$$\text{with} \quad \xi y(t) = \xi uwK,$$
$$\dot{E} = uwK^{1-\gamma}.$$

Now, if we define an inverse demand function for "gross" chips (including product ruined in assembly and final test) by

$$P[x(t), y(t)] = \tilde{P}[\xi x(t), \xi y(t)]\xi$$

and substitute, we get exactly the maximization problem given earlier in equation (1), where

$$R[x(t), y(t)] = P[x(t), y(t)]y(t).$$

Thus, after converting net, finished IC demand to a gross (defect-inclusive) demand for fabricated chips, we can pose the optimization problem in terms of choosing a time path for wafer fab output y (as opposed to net output ξy) and otherwise ignore the additional yield losses in the assembly and final test stages of production.[20] In interpreting the results, we must only remember to divide all "gross" per unit cost and revenue measures (like price, marginal revenue, marginal cost, etc.) emerging from the optimization analysis by ξ, in order to get the "net" cost and revenue measures per good unit observed in the chip marketplace.

3.5.3 DRAM Demand

We must specify a demand function for DRAMs, and an industry structure, in order to calculate marginal revenue R_y. I shall assume a constant elasticity demand function of the form

$$z = \alpha P^{\beta}, \tag{9}$$

with z aggregate demand for DRAMs, P DRAM price, and an industry made up of N identical firms. With this specification, we have

$$R_y = \left(\frac{Ny}{\alpha}\right)^{1/\beta}\left(\frac{\sigma}{\beta} + 1\right), \tag{10}$$

where parameter σ equals the conjectural variation plus one, divided by N, the number of firms. With Cournot competition, σ is $1/N$; with a constant market share cartel, it is 1.

20. The critical assumption is that all "good" chips coming off the wafer fab line incur all the costs of assembly and final test before being culled.

3.6 Model Solution

Next, I briefly summarize the method used to solve numerically for an optimal policy. Full details are given in appendix B. It is useful to categorize optimal policies in terms of two possibilities. One possibility is that full blast production is followed by an "interior segment" where a firm is producing at less than full capacity. In this case, an optimal policy boils down to picking both an optimal capacity K and some optimal time t_s to switch from full blast production to constant output production. The other possibility is that the firm runs at full capacity throughout the product cycle. In this latter regime, necessary conditions for the firm determine only an optimal capacity.

3.6.1 Optimal Output Decisions—with Interior Segments

Appendix B shows that, when the firm produces at less than full capacity, an optimal, profit-maximizing policy must set the difference between marginal cost and marginal revenue equal to δ/K^{γ}, the value of an additional unit of current production in reducing future production costs over the remainder of the product cycle. Therefore, at the optimal switchpoint t_s to an interior segment, we can solve a differential equation determining δ and derive an equation giving t_s as a function of K, N, and other parameters of the control problem.

A second equation giving optimal capacity may be derived from equation (5). After solving for δ over both interior and boundary segments and substituting into (5), we have an expression implicitly giving K as a function of optimal t_s and N. Together with the previous equation, for given N, and various other parameters, we have two equations in two unknowns. An optimal t_s and K pair must solve these two equations.

3.6.2 Optimal Capacity—with No Interior Segment

In many important cases, the optimal path may not contain an interior segment. In this case, $u(t)$ will always equal one. The transversality condition will still hold, and, using this boundary value, we can solve the equation of motion for δ given in appendix B.

Since, however, this expression gives us optimal K conditional on full-capacity utilization over the entire product cycle, we must be careful to ensure that such a path is in fact a Cournot equilibrium. In searching for Cournot equilibria, then, attempts were made to solve both the two-equation system characterizing an optimal policy with interior segments, for a t_s and K pair, and the single equation giving optimal K assuming full-capacity utilization throughout the product cycle. Solutions found were then checked as possible Cournot equilibria, by perturbing both firm capacity K and switching time t_s (if relevant) by .01 in all feasible directions, while maintaining the hypothesized equilibrium output path for all other firms, and calculating the effect on firm profitability (which should necessarily be negative in a Cournot equilibrium).

3.7 Plausible Parameter Values

The final step in this simulation of firm behavior is to decide on empirically plausible parameter values to be used in this model.

3.7.1 Learning Economies

While I am unaware of any published studies of experience curves in the semiconductor industry that control for the effects of varying facility capacities (i.e., estimate γ), there are numerous published estimates of learning curve elasticity ε based on the relation between log output (or log cost) and log cumulative output. In DRAMs, there are several published reports of an empirical 72 percent "learning curve," meaning that current unit cost drops by 28 percent with every doubling of output, corresponding to $\varepsilon = .47$.[21]

To specifically estimate the parameters of the learning curve for 1M DRAMs, estimates of "typical" wafer yields based on historical data and projections for the last four years of a five-year product life cycle were used to derive nonlinear least squares estimates of parameters corresponding to E_0, ϕ, and ε in equation (6).[22] Parameter γ was assumed to equal one; because the unconstrained estimate of E_0 was a small number very close to zero, I imposed a value of .01 for E_0. This was the largest power of ten, which substituted into (8) to constrain parameter estimation, left other parameter estimates unchanged from values produced by the unconstrained estimation procedure. Learning elasticity ε was estimated to be .49, while ϕK^γ had an estimated value of 31.[23]

To further check whether this critical parameter seems to reflect the reality of 1M DRAM production accurately, actual company-specific quarterly production estimates for the six largest 1M DRAM manufacturers were used to estimate learning elasticity ε. "Experience" at time t is given by

21. With a constant wafer-processing cost as the only cost element (the model that underlies these studies), we have

$$\text{unit cost} = \frac{c}{w} = \left(\frac{c}{\phi}\right)E^{-\varepsilon}.$$

A learning elasticity ε equal to .47 is solved from the 72 percent learning curve, since $2^{-\varepsilon} = .72$ (see Noyce 1977; U.S. Congress 1983, 76). On the basis of studies of production costs for IBM bipolar integrated circuits in the 1960s and 1970s, engineers at IBM derived a virtually identical 71 percent learning curve (see Harding 1981, 652). Webbink's 1977 survey of the integrated circuit industry notes that interviewed companies believed ε to lie generally in the .32–.52 range, depending on type of devices (Webbink 1977, 52). Note that Baldwin and Krugman appear to have erred in interpreting the report in U.S. Congress (1983) of a 72 percent learning curve—their basis for assuming that $\varepsilon = .28$ when it actually corresponds to $\varepsilon = .47$!

22. Since the unit of time is the (assumed five-year) product cycle, yields after two years correspond to time .4, after three years .6, etc. The data are given in VLSI Research (1990, addendum A). The data in this addendum correspond to a "typical" wafer fab running twenty-five hundred wafer starts per week, run at full capacity over the product life of the 1M DRAM (conversation with Dan Hutcheson, 19 August 1991).

23. If instead γ was set equal to zero, the estimate of ϕ would have risen from thirty-one to thirty-six, but the estimated ε would not have changed.

$$E(t) = E_0 + \frac{Q(t)}{K^{\gamma}},$$

where $Q(t)$ is cumulative production through time t, and K is capacity. If E_0 is small relative to $Q(t)/K^{\gamma}$, then

$$\ln[y(t)] = \ln[\phi K^{1-\gamma\varepsilon}] + \varepsilon \ln[Q(t)] + (\varepsilon E_0 K^{\gamma})\frac{1}{Q(t)}$$

must hold true.[24] If we choose a period of time in which capacity is approximately constant and fully utilized, then the expressions in K in the above equation may be regarded as part of firm-specific coefficients on two variables—a constant and the inverse of cumulative output—and ε as the coefficient of the log of cumulative output in a regression equation. The equation also provides a simple test for the hypothesis that γ equals zero since, in that case, the coefficient of inverse cumulative output should be constant across firms.

The above equation was estimated using data on cumulative output and current production for the six largest 1M DRAM producers (Toshiba, Hitachi, Fujitsu, NEC, Mitsubishi, and Samsung) over the quarters from 1988:3 to 1989:2, a period of booming demand when trade press accounts suggest that DRAM output was capacity constrained. The point estimate of ε was .65, corresponding to a 36 percent learning curve, confirming other evidence suggesting substantial learning economies.[25] Constraining the coefficient of inverse cumulative output to be the same for all companies reduced the estimate of ε to .51, but a formal statistical test of the corresponding hypothesis that $\gamma = 0$ was inconclusive.[26] In summary, all available data seem to point to a

24. Making use of the fact that $\ln(1 + x) = x$ approximately, for x small.

25. The data used are Dataquest estimates of quarterly output. Reported shipments by Motorola have been added to Toshiba's output and reported shipments by Intel to Samsung's output (since it is believed that most Motorola chips were fabricated by Toshiba and Intel chips "private-labeled" Samsung output during this period). The regression estimated was $\ln y = a_i + \varepsilon \ln Q + b_i 1/Q$, with coefficients a and b varying by producer. The estimate of ε was .52, with a standard error of .07.

26. The results were as follows:

Variable	Unconstrained		With Constraints	
	Estimated Coefficient	Standard Error	Estimated Coefficient	Standard Error
Ln cum. output	.67	.17	.52	.07
Inverse cum. output:				
Toshiba	−.77	1.62	−.13	.05
Hitachi	.27	.42		
NEC	.21	.46		
Fujitsu	.13	.30		

(*continued*)

large yield elasticity with respect to production experience, close to .5, and at least some evidence suggests that absolute cumulative output is not an appropriate choice of "experience" variable.[27] I shall use .49 as my estimate of ε, 31 as my estimate of ϕ.

3.7.2 The Demand for 1M DRAMs

There is little reliable information on the price elasticity of demand for DRAMs. Wilson, Ashton, and Egan (1980, 126–27) estimate that this price elasticity ranges between -1.8 and -2.3 on the basis of a graph of log bit price versus log bits sold. Finan and Amundsen (1986a, C-18; 1986b, 321) report a -1.8 price elasticity on the basis of a simple regression of log bit price on log bits sold worldwide. Neither of these estimates makes any attempt to control for the effect of variation in the overall level of economic activity on chip demand. Flamm (1985, 130–31) estimates an overall price elasticity of demand for semiconductors used in the computer industry of -1.6, assuming a quality adjustment equivalent to the improvement in bit density observed in DRAMs and chip use in computers fixed in proportion to computer output.

To get as reliable an estimate as possible for 1M DRAM demand, I estimated a loglinear demand function giving quantity shipped of 1M DRAMs as a loglinear function of real 1M DRAM price, real prices for 64K and 256K DRAMs (as possible substitutes), real GNP, and a linear trend included to capture intergenerational "transition" effects.[28] The implicit GNP price defla-

Mitsubishi	.75	.44		
Samsung	−.06	.09		
Constant terms:				
Toshiba	2.69	2.04	4.33	5.78
Hitachi	2.21	1.83	3.94	5.78
NEC	2.42	1.85	4.12	5.99
Fujitsu	2.40	1.79	4.06	6.05
Mitsubishi	2.02	1.84	3.96	5.78
Samsung	2.94	1.62	4.42	6.66
	R^2 = .99, SE = .083		R^2 = .98, SE = .094	

The test statistics for the hypothesis of a common coefficient on inverse cumulative output—F-statistic = 1.89 with 5 and 11 df; Wald chi-square statistic = 9.45 with 5 df—lead us to reject the hypothesis at the 10 percent, but not reject at the 5 percent, significance levels.

27. Estimation of a learning elasticity requires data on either current and cumulative output or current average variable cost and cumulative output. The dubious practice of using price as a proxy for current unit cost—as in Dick (1991)—will almost certainly lead to incorrect results since the simple models of pricing behavior reviewed above suggest that market prices will diverge from either current average or marginal cost.

28. The data on quantity cover quarterly worldwide shipments from 1985:2 to 1989:4 by "merchant" producers and are unpublished Dataquest estimates. Data on DRAM prices are also unpublished quarterly Dataquest estimates of average sales price over this same period. Real (deflated) GNP and the implicit GNP price deflator are taken from Council of Economic Advisers, *Economic Report of the President* (various years).

tor, rebased so that the fourth quarter of 1989 was equal to 1, was used to deflate all monetary values to "real" 1989:4 levels. Deflated GNP and substitute DRAM prices were converted to indices taking on value 1 in 1989:4; as a result, the constant in a regression equation may be interpreted as the "level" of DRAM demand corresponding to 1989:4 values for these variables. The estimated regression equation (with estimated standard errors underneath the various coefficients) was

$$\begin{aligned}\ln(Q) = \underset{(1.02)}{22.97} &+ \underset{(2.43)}{1.23}\ln(P_{64K}) + \underset{(1.68)}{.63}\ln(P_{256K}) - \underset{(.49)}{1.47}\ln(P_{1M}) \\ &+ \underset{(.27)}{.29}T - \underset{(36.26)}{1.75}\ln(\mathrm{GNP})\end{aligned}$$

and the estimated price elasticity about -1.5. Dropping the linear time trend variable as a proxy for transitional "generational shift" effects had little effect on the estimated own price elasticity, raising it to -1.55. Interestingly, dropping both GNP and the time trend substantially raised the estimated price elasticity, to -2.1.

On the basis of these results, -1.5 was used as an estimate of 1M DRAM own price elasticity β, and the value 190,000 was used as an estimate of product life-cycle demand "level" α.[29] To transform this demand function to a demand for "gross" fabricated dice (prior to test and assembly losses), it was assumed that net output of tested and finished chips equals .9 times good dice produced in wafer fab.[30] With the functional form assumed, a simple transformation of α is merely substituted for its original value in order to derive the appropriate inverse demand function.[31]

3.7.3 Cost Parameters

Based on estimated 1989 values found in VLSI Research (1990), I estimated r (capital cost per unit product cycle wafer capacity) to be \$240. Variable cost per wafer processed (including materials, labor, and wafer probe test) was estimated to be \$390.[32] Test and assembly costs were assumed to equal \$0.23 for the IC package and about \$0.52 for assembly and final test, for a total of \$0.75 per device produced.[33]

29. Exp(22.97) multiplied by 20 (= 190,000 million) gives demand that would be observed at a 1M DRAM price of \$1.00 over a twenty-quarter (five-year) product cycle, given real output and substitute price levels prevailing in 1989:4.

30. For estimated test and assembly yields in this general neighborhood, see VLSI Research (1990, addendum A) and ICE (1988, 7-16–7-17).

31. That is, $P(\xi z)\xi = (\xi Ny/\alpha)^{1/\beta}\xi = (Ny/\alpha')^{1/\beta}$, where $\alpha' = \alpha\xi^{-(1+\beta)}$.

32. VLSI Research (1990, addendum A) puts material and labor cost at \$380 per wafer processed; I add on a \$10.00 wafer probe test cost based on ICE (1988, 7-9).

33. The package cost comes from conversations with Dan Hutcheson of VLSI Research; the assembly cost is estimated to range from \$0.07 to \$0.20 offshore, or from \$0.10 to \$0.50 per device onshore (in the United States, Europe, and Japan), in ICE (1988, 7-16–7-18). I have used a "typical" value of \$0.32. Final test cost is estimated to be \$0.20 per unit in ICE (1988, 7-18), for a grand total of \$0.75 for package, assembly, and final test.

Overhead is normally a significant part of semiconductor cost. On the basis of aggregate historical data for the period 1981–87, I have assumed \$0.36 in general, administrative, and selling costs for every dollar of direct manufacturing cost.[34] Thus, the estimates for c, d, and r given above were marked up an additional 36 percent. Table 3.1 shows the assumed empirical parameter values used.

3.8 Baseline Simulations

Table 3.2 gives the optimal values of t_s and K derived from numerical solution of the optimal control problem described above. The roots of a system of two nonlinear equations in two unknowns (eqq. [B1] and [B2] in app. B), or one equation in one unknown (in the case where "full blast" production over the entire product life cycle is the optimal policy, eq. [B2′] in app. B), were sought. Table 3.2 also shows a "gross rent," that is, profits net of all costs other than fixed entry cost F, received by each producer. The columns of table 3.2 correspond to different assumed numbers of firms in the industry, the rows to differing assumptions about parameter gamma (γ), which defines the experience variable relevant to learning economies.

Since identical firms are assumed to make up the industry in equilibrium, one may "close" the model by assuming free entry, that firms enter the industry up to the point where gross rent per firm just covers the fixed cost of entry (F). Because we are restricted to an integer number of firms, I define the *equilibrium* as the number of firms where one more entrant reduces rent per firm below entry cost F. As a consequence of the integer number of firms, the symmetric equilibrium so defined will generally be characterized by some small, positive rent (net of entry cost F).

I shall assume that the fixed entry cost (primarily total R&D costs for the 1M DRAM) that must be invested prior to mass production of the 1M DRAM runs between roughly \$250 and \$500 million. Thus, for $\gamma = 1$, if entry costs F amounted to \$250 million, we would expect to find fourteen identical firms in the industry, each with facilities capable of producing 4.66 million wafer starts over a five-year product life cycle. With entry costs F of \$500 million, we would expect nine producers, each with the capacity to produce 6.94 million product cycle wafer starts. In either case, the optimal policy would involve full blast production over the entire life cycle. Thus, one immediate observation that emerges from table 3.2 is that, with $\gamma = 1$ (which I argued earlier is a heuristically appealing specification), small differences in fixed entry costs can make a large difference in the industrial structure of the industry (number of firms observed). The same cannot be said for γ much less than 1.

34. The data on which this calculation is based are found in ICE (1988, 7-20). I have excluded R&D and interest expense as elements of "overhead."

Table 3.1 **Empirical Parameter Values**

Parameter		Assumed Value
α_0	"Level" of life-cycle demand for assembled and tested units at $1.00 per chip	190,000 million units
β	Price elasticity of demand	-1.5
ξ	Share of good, yielded chips as fraction of good dice after assembly and final test	.9
α	Level of demand for "gross" fabricated dice (including units rejected at final test)	$\alpha_0\, \xi^{-(1+\beta)}$
ϕ	Learning curve wafer fab yield "level" parameter	31
E_0	Initial "experience" at time 0	.01
ε	Experience elasticity of wafer fab yield	.49
γ	(Gamma) parameter determining experience variable	0–1
m	Overhead expense per dollar direct manufacturing cost	.36
d	Package, assembly, and final test cost per fabricated unit	$\$0.75 \times (1 + m)$
c	Fabrication cost per processed wafer	$\$390 \times (1 + m)$
r	Capital cost per unit life-cycle wafer-processing capacity	$\$240 \times (1 + m)$

Table 3.3 summarizes some characteristics of industry equilibria derived from table 3.2 under differing assumptions about fixed entry costs F. I have taken F as either \$500 or \$250 million; these values are best interpreted as bracketing a range of feasible values. Alongside the equilibrium number of firms, the Hirschman-Herfindahl index of concentration is also shown.[35]

In order to get at the issue of whether "dumping" is observed, I have calculated observed prices and various cost concepts at one hundred equally spaced points over the product life cycle. One useful cost concept is current short-run marginal cost (SRMC), which in my model happens to be constant at any moment in time, coincides with average variable cost, and is equal to $d + c/w$. This is the incremental cost saved when output is reduced by one unit. Another important cost concept is fully allocated, long-run average cost (LRAC). To define this concept, I have assumed straight-line depreciation in spreading capital and fixed entry costs over the product life cycle: an equal amount of these fixed costs is allocated to every moment in time. Capital and fixed entry costs per unit are then calculated by dividing fixed costs corresponding to time t by the number of units $y(t)$ produced at that moment. Adding average variable cost to average fixed cost, I then have LRAC $= d + c/w + F/y + r/uw$. Multiplying LRAC by output at any instant, and summing these costs at every instant over the product cycle, gives the total cost of producing some time-varying path of output over the entire product cycle.

Table 3.3 shows that, assuming $\gamma = 1$, price falls short of short-run mar-

35. This index is defined as $HHI = \sum_{i=1}^{n} s_i^2$, where s_i is the market share of company i. The index ranges in value from 1, with monopoly, to 0, with a competitive industry composed of an infinite number of equally sized firms. In the special case of N identical firms, this index is just equal to $1/N$.

Table 3.2 **Baseline Simulation Results**

	Number of Firms															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
gamma = 1																
K (mil. wafer starts)	13.49	19.08	16.03	13.34	11.31	9.79	8.62	7.69	6.94	6.32	5.8	5.36	4.98	4.66	4.37	4.11
ts	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Gross rent (mil. $)	29,370	10,380	4,985	2,902	1,894	1,332	988	761	604	491	407	343	293	253	221	195
gamma = .9																
K (mil. wafer starts)	15.46	21.92	18.4	15.73	12.95	11.19	9.84	8.77	7.91	7.2	6.6	6.09				
ts	1	1	1	1	1	1	1	1	1	1	1	1				
Gross rent (mil. $)	32,360	10,370	4,360	2,180	1,191	676	382	204	90	15	−36	−70				
gamma = .8																
K (mil. wafer starts)	16.96	23.91	20.12	16.76	14.23	12.32	10.86	9.68	8.74	7.96	7.3	6.74				
ts	1	1	1	1	1	1	1	1	1	1	1	1				
Gross rent (mil. $)	35,570	10,410	3,752	1,455	477	5	−240	−371	−442	−479	−496	−500				
gamma = .7																
K (mil. wafer starts)	17.683	24.59	20.84	17.48	14.94	13	11.492	10.291								
ts	1	1	1	1	1	1	1	1								
Gross rent (mil. $)	38,840	10,630	3,296	846	−151	−601	−811	−907								

gamma = .68						
K (mil. wafer starts)	17.73	24.56	20.86	17.53	15	13.07
ts	.98	.95	.96	.98	1	1
Gross rent (mil. $)	39,480	10,700	3,232	747	−258	−707
gamma = .6						
K (mil. wafer starts)	17.77	24.4	20.81	17.57	15.11	13.22
ts	.86	.82	.84	.86	.88	.9
Gross rent (mil. $)	41,950	11,050	3,061	428	−619	−1,074
gamma = .5						
K (mil. wafer starts)	17.57	23.82	20.45	17.39	15.04	13.23
ts	.72	.68	.7	.72	.75	.77
Gross rent (mil. $)	44,820	11,630	3,029	197	−924	−1,406
gamma = .3						
K (mil. wafer starts)	16.69	22.06	19.18	16.53	14.46	12.86
ts	.51	.47	.49	.51	.53	.55
Gross rent (mil. $)	49,660	13,020	3,414	187	−1,120	−1,697
gamma = 0						
K (mil. wafer starts)	14.88	19.02	16.82	14.76	13.12	11.83
ts	.31	.28	.29	.31	.33	.34
Gross rent (mil. $)	54,880	15,140	4,492	782	−796	−1,543

Table 3.3 **Characteristics of Symmetric Industry Equilibria**

	F = \$500 Mil.				F = \$250 Mil.			
			Segments of Product Cycle				Segments of Product Cycle	
	No. Firms	Hirschman-Herfindahl	P < SRMC	P < LRAC	No. Firms	Hirschman-Herfindahl	P < SRMC	P < LRAC
gamma = 1	9	.1111	0–.03	0–.32	14	.0714	0–.03	0–.35
gamma = .9	6	.1667	0–.03	0–.29	7	.1429	0–.03	0–.29
gamma = .8	4	.2500	0–.02	0–.19	5	.2000	0–.02	0–.24
gamma = .7	4	.2500	0–.01	0–.17 .93–1	4	.2500	0–.01	0–.15 .97–1
gamma = .68	4	.2500	0–.01	0–.16 .87–1	4	.2500	0–.01	0–.15 .91–1
gamma = .6	3	.3333	0–.01	0–.07	4	.2500	0–.01	0–.12 .71–1
gamma = .5	3	.3333	0–0	0–.05	3	.3333	0–0	0–.05
gamma = .3	3	.3333	0–0	0–.02 .48–.50	3	.3333	0–0	0–.02 .49–.49
gamma = 0	4	.2500	0–0	0–.01 .24–.63	4	.2500	0–0	0–.01 .25–.58

ginal cost over the first 3 percent of the product life cycle and falls short of average cost over roughly the first third of the product cycle. With $\gamma = 0$, by way of contrast, price is less than marginal cost only at the very beginning of the product cycle; price is less than average cost over two distinct periods—at the very beginning and over roughly the second quarter of the product life cycle. Indeed, given my assumptions about other parameter values, for all values of γ price falls short of marginal cost only at the very beginning of the product cycle. Further perusal of this table makes clear, however, that the timing of periods of sales at less than average cost is quite sensitive to the specification of the experience variable—depending on γ, such episodes can occur at the beginning of the product cycle, the middle, or the end or in some combination of these sequences.

Table 3.3 also shows that the value of γ makes a big difference in the structure of a symmetric industry equilibrium. With cumulative output per facility ($\gamma = 1$) the relevant experience variable, a relatively large number of firms (nine to fourteen) populate the industry. With γ much below .9, no more than three or four firms make up the industry.

Figure 3.2 shows the path of price, marginal revenue, marginal cost, and average cost over time in the case where entry costs are \$250 million and $\gamma = 1$. Figure 3.3 shows the time path for these variables over the product cycle when $\gamma = 0$ instead.

Ironically, the specification of firm behavior in the B-K model—full blast production over the entire product cycle—turns out to be optimal if parameter γ is close to 1 (see table 3.2). The irony arises because the B-K model also specifies absolute cumulative output ($\gamma = 0$) as the experience variable, and, given realistic choices for other parameters, optimal behavior would then require cutting back production to levels below capacity after about the first third of the product cycle.

3.8.1 Reality Checks

How plausible are these simulations, and do they suggest anything about the realism of various assumptions about parameters? One straightforward way to evaluate the model is to compare the predicted industry structure with observed industry structure. Figure 3.4 shows Hirschman-Herfindahl concentration indexes constructed from Dataquest estimates of annual producer shipments of various generations of DRAMs.[36] For virtually all generations of

36. These indexes are calculated from unpublished Dataquest estimates of DRAMs shipped from 1974 through the end of 1989. Note that there were two distinct varieties of 16K DRAM, one with a single-voltage power source, the other requiring dual voltages; each is treated as a separate product in this figure. In calculating concentration indexes for 1M DRAMs, I have allocated Motorola-labeled product to Toshiba (since virtually all Motorola's product over this period is believed to have been assembled from Toshiba-fabricated dice or produced by a Toshiba-Motorola joint venture); 1M DRAMs bearing the Intel label have been assigned to Samsung since it is believed that virtually all Intel's sales over this period were "private labeled" Samsung product. Neither of these adjustments has a particularly significant effect on the pattern of concentration.

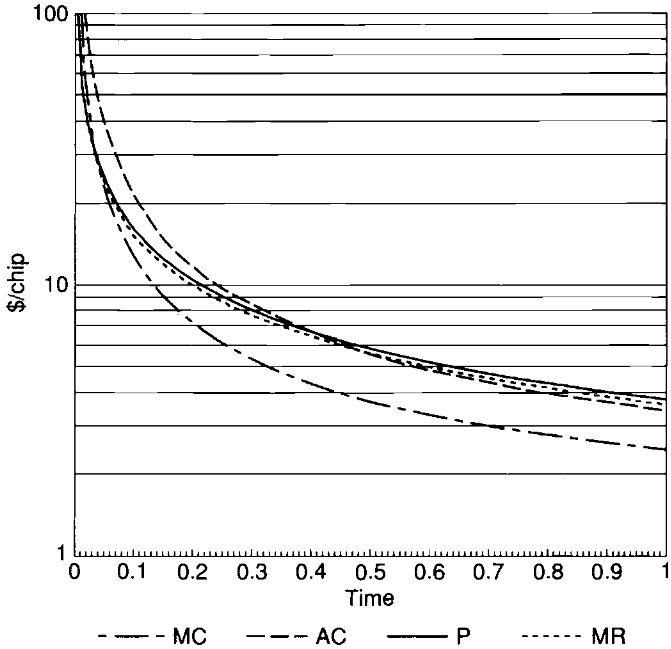

Fig. 3.2 Time profile of costs and prices, simulated equilibrium with gamma = 1 (MC = short-run marginal cost; AC = long-run average cost; P = price; MR = marginal revenue).

DRAM, the concentration index declines sharply from an initially very high level, as one producer after another comes on line with volume production. The index then levels off near .1, rising sharply at the end of the product cycle as producers drop the product line one after another. Although the early phases of the 256K and 1M DRAM may have been somewhat more concentrated than in earlier generations' life cycle, they too seem destined to follow this pattern eventually.

Comparing the Hirschman-Herfindahl indexes associated with my simulations to the pattern depicted in figure 3.4, only the results associated with the specification of cumulative output per facility ($\gamma = 1$) as the experience variable fit reasonably closely. Note that my assumption of symmetric firms means that the associated Hirschman-Herfindahl index of concentration must be constant over time. While conceding that my model is at best an approximation to reality, I conclude that only a γ close to 1 yields predicted behavior that is reasonably close to industrial reality.

Another cut at this question may be had by comparing predicted with actual paths for DRAM prices over time. To do so, I have assumed that a five-year product cycle for the 1M DRAM effectively began in 1988 (although small quantities were produced as far back as late 1985, quantity production did not

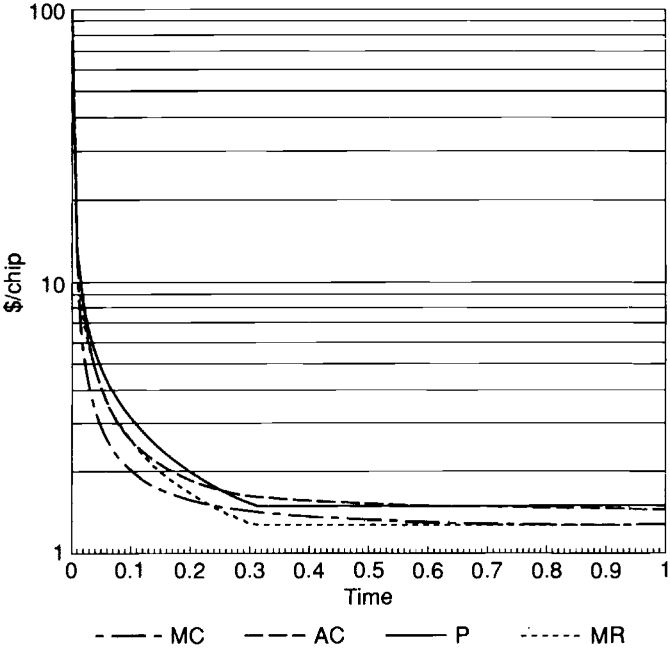

Fig. 3.3 Time profile of costs and prices, simulated equilibrium with gamma = 0 (MC = short-run marginal cost; AC = long-run average cost; P = price; MR = marginal revenue).

really ramp up until 1988). Figure 3.5 charts the actual behavior of one set of estimates of large volume contract prices for 1M DRAMs in the U.S. and Japanese markets through September 1991, along with simulated 1M DRAM price levels associated with an assumed γ equal to 1 and 0, respectively.[37] Note that the period from 1988 through the first quarter of 1989 was a period of extreme shortage in real-world DRAM markets, while the period after late 1989 was one marked by lackluster demand. Given that the early portion of my empirical approximation to the learning curve is probably poorer than in later periods (see the discussion above) and that my assumption of symmetric firms is probably least appropriate in the early stages of the product cycle, I am not surprised to find that the very earliest part of the predicted time path for prices seems least accurate. All things considered, the simulation with $\gamma = 1$ seems to do a reasonable job of tracking real 1M DRAM prices! The simulation with $\gamma = 0$ clearly does not.

Thus, two pieces of evidence—observed and predicted concentration indexes and the time path of DRAM prices—seem to suggest that a value of γ close to 1 provides significantly more realistic predictions than a value close to 0.

A final point to consider is that, historically, the industry folklore holds that

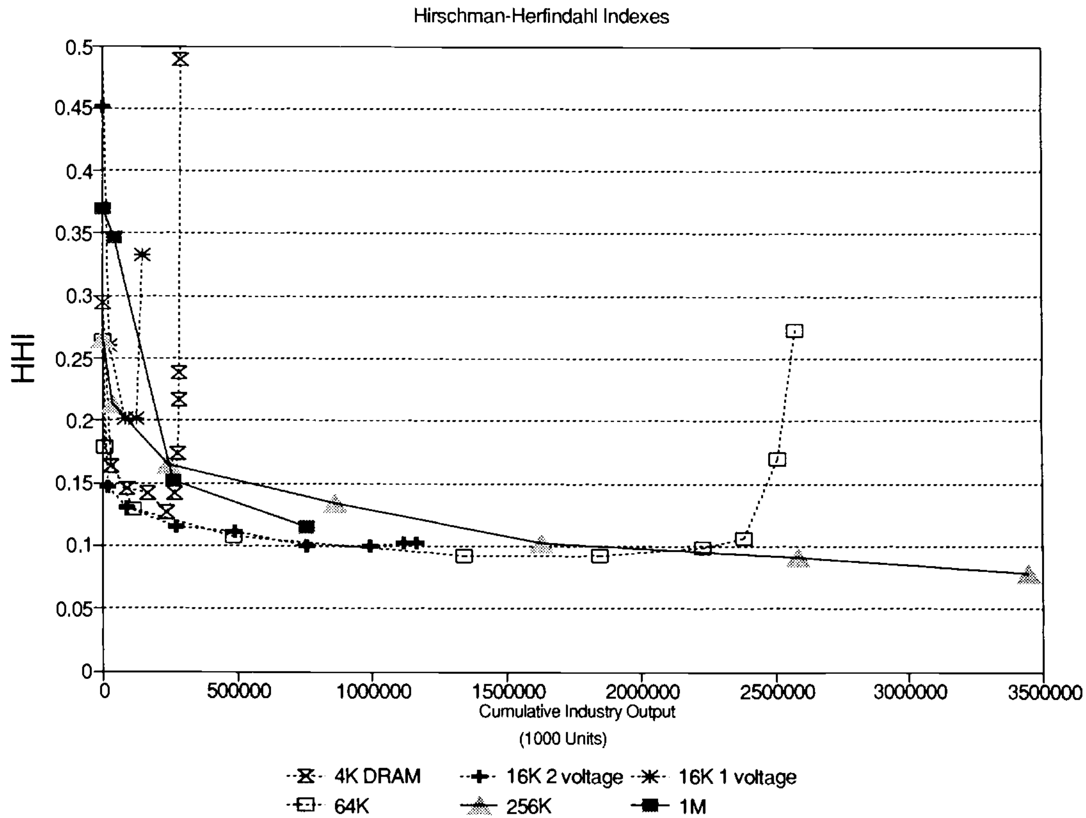

Fig. 3.4 Historical pattern of concentration in DRAM supply

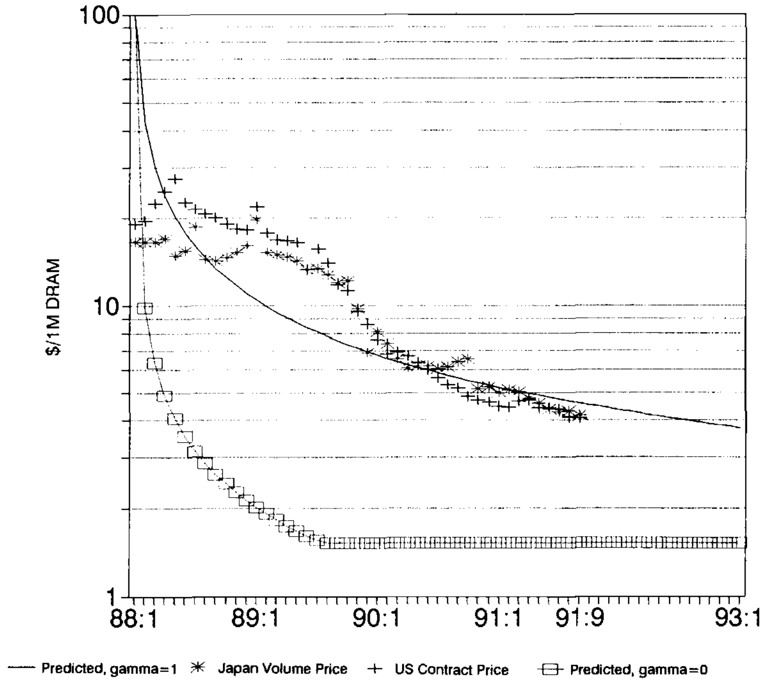

Fig. 3.5 Historical pricing compared with simulated equilibria

DRAM producers have traditionally run their plants at full blast while keeping them in operation, a behavior that is consistent with the simulations presented here. However, beginning in mid-1989, Japanese DRAM producers have announced production cutbacks for DRAMs. This raises three issues that I mention but do not explore in this paper. First, DRAM capacity may be shifted, at some cost, to production of other types of integrated circuits, a possibility not explicitly incorporated into my model. Second, DRAM demand is notoriously cyclical, and the consequences of shifts in demand for optimal producer behavior is, again, not explicitly explored here. Third, production of DRAMs after the conclusion of the 1986 Semiconductor Trade Arrangement was clearly affected by political constraints, may have led to a degree of collusive behavior among producers, and otherwise involved political economic factors not incorporated into my model.

3.8.2 The Dumping Issue

Given empirical values deemed to be plausible in the case of 1M DRAMs, the exercises portrayed in tables 3.2 and 3.3 suggest that a short-run marginal cost test for dumping, as a screening test for potentially predatory behavior, is likely to give only "false positives" (pricing below current marginal cost absent strategic behavior) in the very earliest stages of the product cycle. One might interpret this to mean that a marginal-cost-based dumping test might be defensible if some sort of "exception" to a marginal-cost-based pricing standard is granted when a new product is first introduced. But it is not clear how robust this conclusion is to changes in empirical parameters used in my simulations; further sensitivity analysis might shed greater light on this question.

The same cannot be said for an average cost test for predation. Depending on parameter values, episodes of below-average-cost pricing can pop up in virtually any part of the product life cycle, even when producer behavior is entirely nonstrategic.

Indeed, while the simulations depicted in tables 3.2 and 3.3 all show an episode of below-average-cost pricing at the beginning of the product cycle, possibly followed by a later episode, it would be incorrect to assert that below-average-cost pricing will always necessarily be observed at the beginning of the product life cycle.[38] Figure 3.6 shows that, by artfully changing a single parameter (in this case, by greatly raising initial yields, making $E_0 = 500$), assuming $\gamma = 1$ and $F = \$250$ million, one arrives at a symmetric industry equilibrium where price *never* falls below marginal cost and price falls below average cost only during the last half of the product cycle.[39]

3.9 Conclusions

In recent years, pricing below a constructed long-run average cost has become the principal grounds for applying the "dumping" laws to U.S. imports of foreign products. While this practice has little obvious economic defense, it is possible to argue that a test based on marginal cost might serve as a useful screen for potentially predatory behavior by foreign exporters. However, in the presence of learning economies, such as are thought to be present in many high-tech industries, including semiconductors, below-marginal-cost pricing can be rational even in the absence of strategic behavior, such as predation.

In this paper, I have developed a more realistic model of pricing over the life cycle of a product in which both fixed costs and learning economies are significant. Using empirically plausible parameters for production of 1M DRAMs, and assuming nonstrategic producer behavior, I have found that below-marginal-cost pricing is likely to be observed only in the very earliest stages of the product cycle.

The analysis has also shed considerable light on other facets of pricing and production over the product life cycle. A specification of learning economies

37. These data are monthly averages of Dataquest estimates of average contract prices in these markets. The data are reported in *Computer Reseller News* (various issues). For more on the strengths and weaknesses of these data and a thorough discussion of the segmented spot and contract markets in which DRAMs are sold, see Flamm (in press, a).

38. Dick (1991, 144–46) proposes this behavior.

39. The symmetric equilibrium depicted in this figure corresponds to eighteen producers, each with a capacity of 3.88 million life-cycle wafer starts, producing full blast over the entire product cycle.

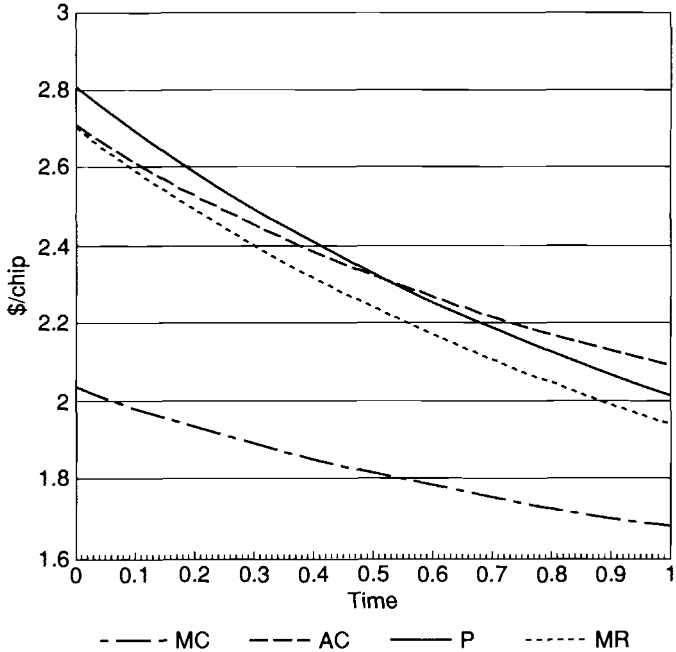

Fig. 3.6 Time profile of costs and prices, simulation with very high initial yield (MC = short-run marginal cost; AC = long-run average cost; P = price; MR = marginal revenue).

based on cumulative output per facility as the "experience" variable was found to yield results that were considerably more realistic than other possible—and popular—specifications. Contrary to popular belief, below-average-cost pricing does not necessarily have to occur near the beginning of the product cycle.

The model presented here appears to produce fairly realistic predictions of industry structure and pricing behavior when used with empirically plausible parameters. Interesting extensions of this work include considering the possibility of strategic, noncooperative behavior on the part of producers as well as cooperative or collusive behavior. Policy issues that might be explored include the effect of various government policies on industry structure and some measure of aggregate national welfare.[40]

Appendix A

Let $y(t)$ be company's output at time t, $x(t)$ output of other companies, F fixed cost of entry, $R[y(t), x(t)]$ company's revenues at time t, d assembly and test

40. These issues are explored in Flamm (in press, b).

cost per chip, c wafer-processing cost, $u(t)$ utilization rate at time t, K fixed capacity, r capital cost per unit capacity, w yielded good chips per wafer, Q cumulative company output, γ experience specification parameter $(0 \leq \gamma \leq 1)$, and E "yield relevant" experience $(E_0 + Q/K^\gamma)$.

1. The general problem described in the text is

$$\max_{u(t),K} \int_0^1 \{R[y(t), x(t)] - F - dy(t) - cu(t)K - rK\}dt,$$
$$\text{where} \quad y(t) = u(t)w(E)K,$$
$$\text{s.t.} \quad \dot{E} = \frac{y(t)}{K^\gamma} = u(t)w(E)K^{1-\gamma}, \quad w_E > 0,$$
$$u \in [0, 1], \quad w > 0.$$

Form the Hamiltonian (suppressing the arguments of functions for notational simplicity)

$$H = R - F - dy - cuK - rK + \delta\frac{y}{K^\gamma},$$

with

$$\dot{\delta} = -\frac{\partial H}{\partial E} = -\frac{\partial H}{\partial y}\frac{\partial y}{\partial E} = -(R_y - d + \frac{\delta}{K^\gamma})uw_E K,$$

where R_y denotes $\partial R/\partial y$, marginal revenue. By the maximum principle, choose $u(t)$ to maximize H, given $u \in [0, 1]$, with $\delta(1) = 0$ (transversality condition). There are three possible cases to consider.

$$0 < u < 1, \quad \frac{\partial H}{\partial u} = \left(R_y - d + \frac{\delta}{K^\gamma}\right)\frac{\partial y}{\partial u} - cK = 0,$$

(A1a)
$$\left(R_y - d + \frac{\delta}{K^\gamma} - \frac{c}{w}\right)Kw = 0,$$
$$R_y - d - \frac{c}{w} + \frac{\delta}{K^\gamma} = 0, \quad \dot{\delta} = -\frac{c}{w}uw_E K < 0.$$

At an interior maximum, we also have a second-order necessary condition:

$$\frac{\partial^2 H}{\partial u^2} = R_{yy}(Kw)^2 \leq 0.$$

(A1b)
$$u = 1, \quad R_y - d - \frac{c}{w} + \frac{\delta}{K^\gamma} \geq 0,$$
$$\dot{\delta} = -\left(R_y - d + \frac{\delta}{K^\gamma}\right)w_E K < 0,$$

(A1c)
$$u = 0, \quad R_y - d - \frac{c}{w} + \frac{\delta}{K^\gamma} \leq 0, \quad \dot{\delta} = 0.$$

Together with the transversality condition, this implies that $\delta \geq 0$ everywhere.

Along an interior segment, that is, possibility (A1a),

$$R_y = d + \frac{c}{w} - \frac{\delta}{K^\gamma},$$

(A1d) $$(\dot{R}_y) = \left(-\frac{c}{w^2}w_E\frac{y}{K^\gamma}\right) - \frac{\dot{\delta}}{K^\gamma} = -c\frac{w_E}{w^2}uwK^{1-\gamma} - \left(-cu\frac{w_E K^{1-\gamma}}{w}\right) = 0,$$

which means that R_y is constant along an interior segment.

The above analysis holds for any given K. Full optimization requires that optimal parameter K must be chosen to satisfy (see Leitmann 1966, 98–100):

$$\int_0^1 \frac{\partial H[\delta(t), K, u^*(t), E^*(t)]}{\partial K} dt = 0,$$

where $u^*(t)$ is the optimal utilization rate and E^* the trajectory of experience variable E corresponding to $u^*(t)$. Then

$$\int_0^1 \left[\left(R_y - d + \frac{\delta}{K^\gamma}\right)\frac{\partial y}{\partial K} + R_x\frac{\partial x}{\partial K} - \gamma\frac{\delta}{K^{1+\gamma}}y - cu - r\right]dt = 0.$$

I will assume nonstrategic behavior—the firm perceives $\partial x/\partial K = 0$—so

$$\int_0^1 \left[\left(R_y - d - \frac{c}{w} + \frac{\delta}{K^\gamma}\right)uw - \gamma\frac{\delta}{K^\gamma}uw - r\right]dt = 0.$$

The expression in brackets above $= 0$ if u interior, ≥ 0 if $u = 1$, and ≤ 0 if $uz = 0$.

Note that this shows that the expression in brackets must be positive, and $u = 1$, over some interval if optimal $K > 0$ (and any output is produced). The identical trajectory of output (and variable profit) could otherwise be produced at lower cost by choosing some smaller $\bar{K}$, then choosing utilization rate $\bar{u}(t) = (K/\bar{K})u(t) < 1$.

2. The above is perfectly general. To explicitly solve for an optimal path, I add additional structure and make the following further assumptions: (i) Industry inverse demand (and firm revenue function R) and learning function w are twice continuously differentiable functions in all their arguments. (ii) There are N firms. (iii) Industry revenues are an autonomous, strictly concave function of industry output.

Let $\bar{R}(z)$ be industry revenues, $z(t)$ industry output ($z[t] = x[t] + y[t]$). Then $\bar{R}(z) = P(z)z$, where $P(z)$ is industry inverse demand.

I will assume $\bar{R}_{zz} = P''z + 2P' < 0$ for $z \geq 0$ (i.e., $\bar{R}$ strictly concave). So

(A2a) $$P'' < -\frac{2P'}{z} \quad \text{for } z \geq 0.$$

I have also assumed that $R[y(t), x(t)]$ is autonomous, not a function of time other than through $x(t)$ and $y(t)$. Since output is a perfectly homogeneous commodity, with a single market price assumed,

$$R[y(t), x(t)] = P(x + y)y,$$
$$R_y = P'\frac{dz}{dy}y \text{ p}+ P = P'(1 + \lambda)y + P,$$

with λ ($= dx/dy$) the conjectural variation perceived by the firm. I shall regard λ as a constant varying between 0 and $N - 1$. These limits parametrize λ as lying between two useful limiting cases of industrial organization: (i) $\lambda = 0$ with Cournot competition; (ii) $\lambda = N - 1$ with a constant market share cartel made up of N identical firms.

I will be assuming $\lambda = 0$ for the moment, but I will develop my analysis of optimal utilization rates including the case of a constant market share cartel.

Note that

$$R_{yy} = P''\left(\frac{dz}{dy}\right)^2 y + 2P'\frac{dz}{dy} = (1 + \lambda)[P''(1 + \lambda)y + 2P']$$
$$< (1 + \lambda)2P'\left[1 - \frac{y}{z}(1 + \lambda)\right],$$

using (A2a). Now, consider two cases: (i) Under Cournot competition, $\lambda = 0$, since

$$\frac{y}{z} \leq 1, \text{ then } 1 - \frac{y}{z}(1 + \lambda) > 0, \quad \text{and} \quad R_{yy} < 0.$$

(ii) Under N identical firms in a collusive, fixed-market-share cartel, $\lambda = N - 1$,

$$\frac{y}{z} = \frac{1}{N}, \quad 1 - \frac{y}{z}(1 + \lambda) = 0, \quad \text{and} \quad R_{yy} < 0.$$

In these two cases, firm revenue R is strictly concave in y, with $R_{yy} < 0$ everywhere. Now, since functions R and w are assumed twice continuously differentiable in their arguments, so too will be the Hamiltonian H, with $H_{uu} = R_{yy}(Kw)^2 < 0$. H is strictly concave in u. For given K, then, the necessary conditions for optimal u are also sufficient to guarantee that we are maximizing H. More important, the strict concavity of H in u and the constant bounds constraining feasible u mean that we may invoke an appropriate theorem to conclude that optimal u is continuous in other arguments of H.[41] Since

41. For example, a result proved by Debreu, found in Lancaster (1987, 349–50), or a theorem due to Fiacco, cited in McCormick (1983, 245–46). More directly, we may note that the strict concavity of H in u means that, at every moment, the $u(t)$ that maximizes H is unique. Since the set of feasible values from which $u(t)$ is chosen is compact, we may invoke a theorem (6.1) from Fleming and Rishel (1975, 75) to conclude that $u(t)$ is a continuous function of time.

the other arguments of H are continuous functions of time, u must also be a continuous function of time as well.

Note also that $u(t)$ is a continuously differentiable function of time within an interior segment. This is a consequence of the linearity of $\dot{E}$ in u and the strict concavity of the Hamiltonian.[42]

Next I restrict discussion to symmetric industry equilibria, that is, with the industry made up of identical firms. In this case, I will show that, along an interior segment of such an equilibrium, $\dot{u} < 0$.

We already know (see [A1d]) that, along an interior segment,

$$\begin{aligned}(\dot{R}_y) &= 0 \\ &= \frac{d}{dt}[P'(1 + \lambda)y + P] \\ &= \dot{y}[P''(1 + \lambda)Ny + P'(1 + \lambda) + P'N].\end{aligned}$$

Is it possible that the expression in brackets equals zero? If

$$P''(1 + \lambda)Ny + P'(1 + \lambda) + P'N = 0,$$

$$P'' = -\frac{P'\left(1 + \dfrac{N}{1 + \lambda}\right)}{z} > -\frac{P'}{z}(2)$$

since $1 + N/(1 + \lambda) \geq 2$. But this contradicts (A2a) and our assumption that $\bar{R}_{zz} < 0$. So we must have $\dot{y} = 0$.

Now

$$y = uKw, \quad \dot{y} = \dot{u}Kw + uKw_E\frac{y}{K\gamma} = 0,$$

which can only be true if $\dot{u} < 0$.

Since u is continuous over time, when $u = 0$, it cannot jump to 1. Indeed, u cannot even become positive since $\dot{u} < 0$ as soon as $u > 0$. So, when $u = 0$, the optimal policy must remain $u = 0$.

If we add the further assumption that $\bar{R}_z(0)$ exceeds current marginal cost at time 0 (as must be true with constant elasticity < -1), then, because $y/z = 1/N$,

42. Consider optimal control $u^*(t)$ over some interior segment $t_1 < t < t_2$. Because of the strict concavity of the Hamiltonian, a local interior maximum of H must also be a global maximum of H with no constraints on control $u(t)$. That is, optimal control $u^*(t)$ over this interval must also be the optimal control for the problem

$$\max_{u(t)}\int_{t_1}^{t_2}[R(x, y) - F - dy - cuK - rK]dt,$$

subject to the initial and terminal conditions that $E(t)$ take on the values at times t_1 and t_2 associated with optimal control $u^*(t)$ in the original problem, but with no bounds on control $u(t)$. As Fleming and Rishel (1975, corollary 6.1, p. 77) note, the linearity of the equations of motion in the control variable and the concavity of the Hamiltonian are sufficient to guarantee that the optimal control is a continuously differentiable function of time for this subproblem (with no constraints on $u[t]$).

$$R_y = \bar{R}_z - P'z\left[1 - \frac{y}{z}(1 + \lambda)\right] \geq \bar{R}_z,$$

and (A1c) cannot hold, it can never be optimal for $u = 0$.

Note the role of learning economies in this model. If $w_E = 0$ everywhere, (no learning) δ must be constant and equal to zero and optimal u constant over the product cycle. Optimal capacity choice implies that $R_y - d - c/w$ must be greater than zero, and therefore $u = 1$. The first-order conditions then require that K be chosen so variable profit generated by a marginal unit of capital $[R_y - d - c/w]w$ just covers the cost of capital (r), and all available capacity is fully utilized over the entire product life cycle.

Appendix B

Specification of the Demand Curve

Let inverse demand be given by $P = (z/\alpha)^{1/\beta}$ with $z = x + y$ total industry output and β the price elasticity of demand. Consider industry sales, given by

$$\begin{aligned}
\bar{R} &= \left(\frac{z}{\alpha}\right)^{1/\beta} z, \quad -1 < \frac{1}{\beta} < 0, \\
\bar{R}_z &= \left(\frac{1}{\beta} + 1\right)\left(\frac{z}{\alpha}\right)^{1/\beta} > 0, \\
\bar{R}_{zz} &= \left(\frac{1}{\beta} + 1\right)\frac{1}{\beta}\alpha^{-1/\beta}z^{1/\beta - 1} < 0,
\end{aligned}$$

and note that $\lim_{z \to 0} \bar{R}_z = \infty$.

Now, marginal revenue for any individual firm is given by

$$\begin{aligned}
R_y = P'\left(1 + \frac{dx}{dy}\right)y + P &= P\left[\frac{P'zy}{Pz}\left(1 + \frac{dx}{dy}\right) + 1\right] \\
&= P\left[\frac{1}{\beta}\frac{y}{z}(1 + \lambda) + 1\right],
\end{aligned}$$

with $\lambda = \dfrac{dx}{dy}$ defining the conjectural variation:

> if $\lambda = N - 1$, we have a constant market share cartel;
> if $\lambda = 0$, we have Cournot competition.

In addition, I assume that the industry is made up of N identical firms. In symmetric industry equilibrium, each firm has market share $y/z = 1/N$. Let $\sigma = (1 + \lambda)/N$. So

$$R_y = P\left(\frac{\sigma}{\beta} + 1\right),$$
$$\text{where} \quad \sigma = \begin{cases} 1, \text{ cartel;} \\ 1/N, \text{ Cournot.} \end{cases}$$

Thus, in an industry made up of N identical firms,

$$R_y = \left(\frac{N\ y}{\alpha}\right)^{1/\beta}\left(\frac{\sigma}{\beta} + 1\right).$$

Specification of Learning Economies and Output

Let $w = \phi E^{\varepsilon}$, with

$$\dot{E} = \frac{y}{K^{\gamma}}, \quad E(0) = E_0, \quad 0 \leq \varepsilon \leq 1, \quad 0 \leq \gamma \leq 1.$$

Initial yield (ϕE_0^{ε}) is independent of capacity choice. We are mainly interested in two specific cases: $\gamma = 0$ (learning depends on absolute production experience) and $\gamma = 1$ (learning depends on experience per unit capacity, or facility). Then $y = \phi E^{\varepsilon} u K$.

a) Over the interval from 0 to t_s, where $u = 1$,

$$\dot{E} = \frac{y}{K^{\gamma}} = \phi E^{\varepsilon} K^{1-\gamma}.$$

Solving this differential equation, we have

$$E(t, K) = \left[E_0^{1-\varepsilon} + K^{1-\gamma}\,\phi t(1 - \varepsilon)\right]^{1/1-\varepsilon},$$
$$y(t, K) = \phi E(t, K)^{\varepsilon} K.$$

We also know

$$\dot{\delta} = -\frac{\partial H}{\partial E} = -\left(R_y - d + \frac{\delta}{K^{\gamma}}\right) w_E K = -\left(R_y - d + \frac{\delta}{K^{\gamma}}\right) \varepsilon \phi E^{\varepsilon - 1} K,$$

which can be rewritten as the linear monic differential equation

$$\dot{\delta} + f_1(t, K)\delta = f_2(t, K),$$

with

$$f_1(t, K) = \varepsilon \phi E(t, K)^{\varepsilon - 1} K^{1-\gamma}, \quad f_2(t, K) = -[R_y(t, K) - d] f_1(t, K) K^{\gamma}.$$

Solving for δ, for some t in the interval $(0, t_s)$, given boundary value $\delta(t_s) = \delta_{ts}$, we have

$$\delta(t) = \frac{\delta_{ts} + \int_{ts}^{t} f_2(\tau) e^{\int_{t_s}^{\tau} f_1(\mu) d\mu} d\tau}{e^{\int_{t_s}^{t} f_1(\tau) d\tau}}.$$

Define function $\delta_B(t, t_s, K, \delta_{ts})$ by the right-hand side of this equation. With some difficulty, and a great deal of tedious algebra, we can then integrate this expression and have

$$\delta_B(t, t_s, K, \delta_{ts}) = \frac{\delta_{ts} + \frac{\beta + \sigma}{\beta + 1}\left(\frac{\phi K N}{\alpha}\right)^{1/\beta} E(t_s, K)^{\varepsilon/\beta}\left\{1 - \left[\frac{E(t, K)}{E(t_s, K)}\right]^{\varepsilon/\beta + \varepsilon}\right\}K^{\gamma}}{\left[\frac{E(t, K)}{E(t_s, K)}\right]^{\varepsilon}} - \frac{d\left(1 - \left[\frac{E(t, K)}{E(t_s, K)}\right]^{\varepsilon}\right)K^{\gamma}}{\left[\frac{E(t, K)}{E(t_s, K)}\right]^{\varepsilon}}$$

b) *Over the interval from t_s to 1*, optimal u is set such that $y(t, k) = y(t_s, k)$, that is, constant output. $E(t_s, k)$ and $y(t_s, k)$ are given by expressions in the last section. Over the interval from t_s to 1, then,

$$\dot{E} = \frac{y(t_s, K)}{K^{\gamma}}, \quad \text{so } E(t, k) = (t - t_s)\frac{y(t_s, K)}{K^{\gamma}} + E(t_s, K);$$

$$E(t, K) = (t - t_s)\phi E(t_s, K)^{\varepsilon} K^{1-\gamma} + E(t_s, K) = E(t_s, K)[1 + Q(t, t_s, K)],$$
$$\text{where } \quad Q(t, t_s, K) = \phi E(t_s, K)^{\varepsilon - 1} K^{1-\gamma}(t - t_s),$$

incremental experience from time t_s to time t, relative to experience at t_s;

$$\dot{\delta} = -\frac{cuw_E K}{w} = -\frac{cyw_E}{w^2} = -c\frac{\phi^2 E(t_s, K)^{\varepsilon} K \varepsilon E(t, K)^{\varepsilon - 1}}{\phi^2 E(t, K)^{2\varepsilon}} = f_3(t; t_s, K),$$
$$\text{where } f_3(t; t_s, K) = \frac{-c\varepsilon E(t_s, K)^{\varepsilon} K}{E(t, K)^{\varepsilon + 1}}$$

Conditional on assumed switchpoint t_s, the solution to this differential equation can be written as

$$\delta(t) - \delta(1) = \int_1^t f_3(\tau; t_s, K) d\tau,$$

or, making use of the fact that $\delta(1) = 0$ (transversality condition), define function δ_E as

$$\delta_E(t;\, t_s,\, K) = \int_1^t f_3(\tau;\, t_s,\, K)d\tau$$

$$= \frac{cK^\gamma}{\phi[E(t_s,\, K) + \phi K^{1-\gamma}(t - t_s)E(t_s,\, K)^\varepsilon]^\varepsilon}$$

$$- \frac{cK^\gamma}{\phi[E(t_s,\, K) + \phi K^{1-\gamma}(1 - t_s)E(t_s,\, K)^\varepsilon]^\varepsilon}$$

$$= \frac{cK^\gamma}{\phi E(t_s,\, K)^\varepsilon}\left\{\frac{1}{[1 + Q(t,\, t_s,\, K)]^\varepsilon} - \frac{1}{[1 + Q(1,\, t_s,\, K)]^\varepsilon}\right\}.$$

In particular, we can solve for $\delta_E(t_s;\, t_s,\, k)$, that is, the value of δ at t_s, which, given some assumed t_s, solves the equation of motion over an interior segment and the transversality condition at time 1. In this special case, we have

$$\delta_E(t_s;\, t_s,\, K) = \frac{cK^\gamma}{\phi E(t_s,\, K)^\varepsilon}\left\{1 - \frac{1}{[1 + Q(1,\, t_s,\, K]^\varepsilon}\right\}.$$

Solution of the Model

Our specification has ruled out the possibility that optimal $u = 0$. The optimal path for u will consist of a capacity-constrained ($u = 1$) segment through time t_s, possibly followed by an interior segment. For the moment, assume that $t_s < 1$.

Case 1: $t_s < 1$

At t_s, we also know that $u(t)$ will be entering an interior segment, so

$$\frac{\delta}{K^\gamma} = -\left[R_y(t_s,\, K) - d - \frac{c}{\phi E(t_s,\, K)^\varepsilon}\right]$$

must be true at t_s. Thus, for given K, optimal t_s must satisfy

$$\frac{\delta_E(t_s;\, t_s,\, K)}{K^\gamma} + \left[R_y(t_s,\, K) - d - \frac{c}{\phi E(t_s,\, K)^\varepsilon}\right] = 0,$$

or, substituting,

$$\text{(B1)}\quad \frac{-c}{\phi[E(t_s,\, K) + \phi K^{1-\gamma}(1 - t_s)E(t_s,\, K)^\varepsilon]^\varepsilon} + \left[\frac{N\phi E(t_s,\, K)^\varepsilon K}{\alpha}\right]^{1/\beta}\left(\frac{\sigma}{\beta} + 1\right) - d = 0.$$

This is just the condition that marginal revenue equal current marginal cost at terminal time 1. If K is chosen to nonstrategically maximize profits, optimal K must satisfy

$$\int_0^1 \left\{ \left[R_y(t, K) - d - \frac{c}{\phi E(t, K)^\varepsilon} + \frac{\delta}{K^\gamma} \right] u \phi E(t, K)^\varepsilon - \gamma \frac{\delta}{K^\gamma} u \phi E(t, K)^\varepsilon \right\} dt - r = 0.$$

Since the expression in brackets is 0 after t_s and u is 1 before t_s, this can be written as

$$\int_{ts}^1 -\gamma \frac{\delta}{K^\gamma} u \phi E(t, K)^\varepsilon dt + \int_0^{ts} \left\{ [R_y(t, K) - d] \phi E(t, K)^\varepsilon + \frac{\delta}{K^\gamma} \phi E(t, K)^\varepsilon (1 - \gamma) \right\} dt - ct_s - r = 0.$$

We may substitute for δ with the function δ_B and δ_E, described earlier. Noting that

$$u(t) = \frac{y(t_s, K)}{K \phi E(t, K)^\varepsilon} = \frac{\phi E(t_s, K)^\varepsilon K}{K \phi E(t, K)^\varepsilon} = \frac{E(t_s, K)^\varepsilon}{E(t, K)^\varepsilon} = \frac{1}{[1 + Q(t, t_s, K)]^\varepsilon}$$

over the interval from t_s to 1, we can solve analytically for the first integral above, so, substituting,

$$\text{(B2)} \qquad -\frac{c\gamma(1 - t_s)}{(1 - \varepsilon)} \left\{ \frac{\frac{1}{Q(1, t_s, K)} + \varepsilon}{[1 + Q(1, t_s, K)]^\varepsilon} - \frac{1}{Q(1, t_s, K)} \right\} + \int_0^{t_s} \left([R_y(t, K) - d] \cdot \phi E(t, K)^\varepsilon + \delta_B \left\{ t, t_s, K, -\left[R_y(t_s, K) + d + \frac{c}{\phi E(t_s, K)^\varepsilon} \right] K^\gamma \right\} \cdot (1 - \gamma) \frac{\phi E(t, K)^\varepsilon}{K^\gamma} \right) dt - ct_s - r = 0.$$

An optimal t_s and K choice, then, must solve equations (B1) and (B2).

Case 2: $t_s = 1$

The other possibility is that $t_s = 1$ and producers fully utilize available capacity throughout the product life cycle. Since there is no interior segment, equation (B1) does not have to hold at t_s. Instead, the transversality condition means that $\delta(t_s) = 0$, so $\delta_B(t, 1, K, 0)$ gives the value of $\delta(t)$ at any time t. Incorporating this into the first-order condition for optimal K, I then have

$$\text{(B2}'\text{)} \quad \int_0^1 \Big\{ [R_y(t, K) - d] \cdot \phi E(t, K)^{\varepsilon} + \delta_B(t, 1, K, 0) \cdot (1 - \gamma)\, \phi \frac{E(t, K)^{\varepsilon}}{K} \Big\} dt - c - r = 0,$$

which can be solved for optimal K.

In searching for Cournot equilibria, then, attempts were made to solve (B1) and (B2) for optimal (t_s, K) and (B2′) for optimal K (assuming $t_s = 1$).

Industry Profits

Total rents per firm, gross of fixed entry cost F, earned in an industry made up of N identical firms can be calculated as (for given optimal t_s and K)

$$\int_0^1 \left(\left\{ \left[\frac{Ny(t, K)}{\alpha} \right]^{1/\beta} - d \right\} y(t, K) - cu(t)K - rK \right) dt$$

or

$$\int_0^{t_s} \left(\left\{ \left[\frac{Ny(t, K)}{\alpha} \right]^{1/\beta} - d \right\} y(t, K) \right) dt + \left\{ \left[\frac{Ny(t_s, K)}{\alpha} \right]^{1/\beta} - d \right\} y(t_s, K)(1 - t_s)$$

$$- cKt_s - \frac{cK(1 - t_s)}{(1 - \varepsilon)} \left\{ \frac{\frac{1}{Q(1, t_s, K)} + 1}{[1 + Q(1, t_s, K)]^{\varepsilon}} - \frac{1}{Q(1, t_s, K)} \right\} 3 - rK.$$

References

Baldwin, Richard E., and Paul R. Krugman. 1988. Market access and international competition: A simulation study of 16K random access memories. In *Empirical methods for international trade*, ed. Robert C. Feenstra. Cambridge, Mass.: MIT Press.

Deardorff, Alan V. 1989. Economic perspectives on antidumping law. In *Antidumping law and practice: A comparative study*, ed. J. H. Jackson and E. A. Vermulst. Ann Arbor: University of Michigan Press.

Dick, Andrew R. 1991. Learning by doing and dumping in the semiconductor industry. *Journal of Law and Economics* 34 (April): 133–59.

Dixit, Avinash K. 1986. Comparative statics for oligopoly. *International Economic Review* 27 (February): 107–22.

Finan, William F., and Chris B. Amundsen. 1986a. An analysis of the effects of targeting on the competitiveness of the U.S. semiconductor industry. Study prepared for the Office of the U.S. Special Trade Representative, the U.S. Department of Commerce, and the U.S. Department of Labor by Quick Finan & Associates, Washington, D.C.

———. 1986b. Modelling U.S.-Japan competition in semiconductors. *Journal of Policy Modeling* 8(3): 305–26.

Flamm, Kenneth. 1985. Internationalization in semiconductors. In *The global factory: Foreign assembly in international trade*, ed. J. Grunwald and K. Flamm. Washington, D.C.: Brookings.

———. In press, a. Measurement of DRAM prices: Technology and market structure. In *Price measurements and their uses*, ed. M. F. Foss, M. Manser, and A. Young. Chicago: University of Chicago Press.

———. In press, b. *Mismanaged trade? Strategic policy and the semiconductor industry.* Washington, D.C.: Brookings.

Fleming, Wendell H., and Raymond W. Rishel. 1975. *Deterministic and stochastic optimal control.* New York: Springer.

Fudenberg, Drew, and Jean Tirole. 1983. Learning by doing and market performance. *Bell Journal of Economics* 14 (Autumn): 522–30.

Harding, William E. 1981. Semiconductor manufacturing in IBM, 1957 to the present: A perspective. *IBM Journal of Research and Development* 25 (September): 647–58.

Helpman, Elhanan, and Paul R. Krugman. 1989. *Trade policy and market structure.* Cambridge, Mass.: MIT Press.

Horlick, Gary N. 1989. The United States antidumping system. In *Antidumping law and practice: A comparative study*, ed. J. H. Jackson and E. A. Vermulst. Ann Arbor: University of Michigan Press.

Integrated Circuit Engineering (ICE). 1988. *Mid-term 1988.* Scottsdale, Ariz.: ICE.

Kreps, D. M., and J. A. Scheinkman. 1983. Quantity precommitment and Bertrand competition yield Cournot outcomes. *Bell Journal of Economics* (Autumn): 326–37.

Kreps, D. M., and R. Wilson. 1982. Reputation and imperfect information. *Journal of Economic Theory* 27:253–79.

Krishna, Kala 1988. Market access and international competition: A simulation study of the 16K random access memories: Comment. In *Empirical methods for international trade*, ed. Robert C. Feenstra. Cambridge, Mass.: MIT Press.

Lancaster, Kelvin. 1987. *Mathematical economics.* New York: Dover.

Leitmann, George. 1966. *An introduction to optimal control.* New York: McGraw-Hill.

McCormick, Garth P. 1983. *Nonlinear programming: Theory, algorithms, and applications.* New York: Wiley.

Micron Technology. 1985. Petition for the imposition of antidumping duty. Submission to the U.S. Department of Commerce and the U.S. International Trade Commission, Washington, D.C., 21 June.

Milgrom, P. 1987. "Predatory pricing." In *The new Palgrave: A dictionary of economics*, ed. J. Eatwell, M. Milgate, and P. Newman. London: Macmillan.

Milgrom, P., and J. Roberts. 1982. Limit pricing and entry under incomplete information. *Econometrica* 50:443–59.

Nivola, Pietro S. 1990. Trade policy: Refereeing the playing field. In *A question of balance: The president, Congress, and foreign policy*, ed. Thomas E. Mann. Washington, D.C.: Brookings.

Noyce, Robert N. 1977. Microelectronics. *Scientific American* 237 (September): 62–69.

Ordover, Janusz A., and Garth Saloner. 1989. Predation, monopolization, and antitrust. In *Handbook of industrial organization*, vol. 1, ed. R. Schmalensee and R. D. Willig. Amsterdam: North-Holland.

Shapiro, Carl. 1989. Theories of oligopoly behavior. In *handbook of industrial organization*, vol. 1, ed. R. Schmalensee and R. D. Willig. Amsterdam: North-Holland.

Spence, A. Michael. 1981. The learning curve and competition. *Bell Journal of Economics* 12:49–70.

Tirole, Jean. 1988. *The theory of industrial organization.* Cambridge, Mass.: MIT Press.
U.S. Congress. Office of Technology Assessment. 1983. *International competitiveness in electronics.* Washington, D.C.: U.S. Government Printing Office.
VLSI Research. 1990. *Megabit manufacturing economics.* San Jose, Calif.: VLSI Research.
Waller, Larry. 1988. DRAM users and makers: Shotgun marriages kick in. *Electronics,* November, 29–30.
Webbink, Douglas W. 1977. *Staff report on the semiconductor industry.* Washington, D.C.: Federal Trade Commission, Bureau of Economics.
Wilson, Robert W., Peter K. Ashton, and Thomas P. Egan. 1980. *Innovation, competition, and government policy in the semiconductor industry.* Lexington, Mass.: Lexington.

Comment Yun-Wing Sung

During my flight to the conference, I was reading *Newsweek* magazine, in which I came across Michael Boskin's comment on potato chips and computer chips: "Potato chips, computer chips, they're all chips. What's the difference?" I agree wholeheartedly with Boskin because I know nothing about computer chips, and I beg your forgiveness if I mix them up with potato chips.

I enjoy Kenneth Flamm's paper very much. Flamm's model has the advantages of both the Spence model, which explains forward pricing, and the Baldwin and Krugman model, which gives a realistic time path of prices. In Flamm's model, firms first make some optimal capacity investment. Initially, they run that capacity full blast and then switch to a policy of maintaining constant output but decreasing utilization of capacity. In other words, firms exploit their monopoly power at the latter stages of the product cycle. The driving force behind the time paths of falling prices and decreasing capacity utilization is learning economies.

Although Flamm's model undoubtedly represents a big step forward in modeling, it does not take into account the effects of entry, and this is a significant weakness. According to the product-cycle theory, firms producing high-tech products typically enjoy temporary monopoly powers at the first stage of the product cycle, and they earn substantial economic profits. Such profits attract imitators, and the economic profits are competed away. In contrast to Flamm's model, firms exploit their monopoly powers at the initial rather than the latter stages of the product cycle. The fall in prices over time is a result of entry as well as learning economies.

In the case of the time path of prices, the effects of entry and learning economies work in the same direction: both factors lead to a fall in prices. How-

Yun-Wing Sung is senior lecturer in economics at the Chinese University of Hong Kong.

ever, in the case of the time paths of capacity utilization and exploitation of market power, the two factors work in opposite directions. Entry leads to a weakening of market power and increasing capacity utilization over time, but learning economies have the opposite effects.

It might be argued that, unlike potato chips, entry in the computer chips industry is much more difficult and learning economies in the computer chips industry much more significant, with the result that the effects of learning economies on capacity utilization and exploitation of market power predominate over those of entry. However, this is an empirical matter that cannot be presumed. Moreover, both casual empiricism and the product-cycle theory suggest that entry is an extremely important factor that cannot be ignored.

Besides entry, temporal optimization on the part of buyers would also counteract the decrease in capacity utilization over time. Buyers expect the prices of chips to fall over the product cycle, and they postpone their purchase as a result. This shifts purchases from the initial stages to the latter stages of the product cycle and tends to increase capacity utilization over time. It should be noted that Flamm's model assumes a constant elasticity of demand over time and thus cannot take into account the postponement of purchases on the part of buyers.