

Dynamic competition with switching costs

Joseph Farrell*

and

Carl Shapiro**

We analyze an overlapping-generations model of duopolistic competition in the presence of consumer switching costs. Competition for established buyers is continually intermingled with competition for new, uncommitted buyers. In equilibrium the firm with attached customers typically specializes in serving them and concedes new buyers to its rival. This pattern of repeated entry persists even when there are economies of scale or network externalities. Switching costs thus can cause inefficiency in a surprising way: far from forming an entry barrier, they encourage entry to serve new customers, even when such entry is inefficient.

1. Introduction

■ In many markets buyers must make relationship-specific investments to deal with suppliers. For instance, learning to use a vendor's product takes time, and the skill may not be fully transferable to a competitor's product if there is inadequate standardization. Similarly, buyers may acquire complementary goods (such as computer software) that work only with one vendor's products. There also may be transaction costs associated with finding and establishing a working relationship with a new supplier.¹ Relationship-specific assets and institutional responses to them have been examined in detail by Williamson (1985).

These relationship-specific assets create *switching costs* for a buyer who changes from one supplier to another. Evidently, such brand loyalty gives the seller some monopoly power: in the absence of effective long-term contracts² a buyer is open to exploitation by an opportunistic seller who could raise his price above competitors' by an amount almost equal to the buyer's switching cost. And, if competitors are making similar calculations,

* University of California, Berkeley.

** Princeton University.

We are grateful to Avinash Dixit, Drew Fudenberg, Gene Grossman, and Michael Riordan for comments on an earlier draft, and to seminar participants at the University of California, Berkeley, the University of California, Los Angeles, California Institute of Technology, Massachusetts Institute of Technology, and Northwestern, Princeton, and Stanford Universities for their suggestions. Farrell acknowledges financial support from the National Science Foundation (IRI 87-12238). Shapiro acknowledges financial support from the Alfred P. Sloan Foundation as a Sloan Research Fellow and from the National Science Foundation (SES 86-06336). An earlier version of this research appeared as a working paper, Farrell and Shapiro (1987a).

¹ For more examples, see Klemperer (1987a, b).

² Clearly, *complete* long-term contracts would solve any problems associated with switching costs. In Farrell and Shapiro (1987b), we explore the role of *incomplete* long-term contracts in the presence of such costs.

prices might rise farther above cost than just the switching cost itself, as in Diamond's (1970) search model.

Switching costs may thus weaken competition among existing firms. One might fear that they would also subvert the competitive process of *entry*. For instance, Porter (1980, p. 10) states (approvingly), "A barrier to entry is created by switching costs." And in *U.S. v. IBM*, the government claimed that "the lack of transferability of software from machines of one manufacturer to those of another has constituted a substantial barrier to entry" (Fisher, McGowan, and Greenwood, 1983, p. 196). One of our goals in this article is to explore whether this conclusion is correct.

Clearly, if there are economies of scale, the incumbent (given that his existing customers are an assured market) has a cost advantage over the entrant in selling to new buyers. This observation correctly suggests that an incumbent *could* exclude rivals while still making positive economic profits. But it is not obvious that he *will* do so, rather than exploit existing buyers and leave to his rivals the market for new customers. In fact, we shall show below that switching costs tend to *promote* entry.

The force that we identify is the incumbent's unwillingness to cut his price to attract new buyers when he cannot discriminate between them and existing ones. Farrell (1986) and Klemperer (1987a) have shown in two-period models how this fat-cat effect (Fudenberg and Tirole, 1984) makes incumbent sellers with large market shares lose share by setting higher prices than their leaner rivals. In this article we extend this insight to a fully dynamic model and focus on the extreme case where one seller has all existing buyers and the other is an "entrant." We then assess how far this force overcomes the effect of switching costs as an entry barrier.

Existing work on switching costs largely discusses competition in static and two-period models.³ These two-period models (Klemperer, 1987a, 1987b) capture the basic structure of competition for an isolated cohort of consumers: sellers price aggressively (perhaps even below cost) in the first period to attract buyers, whom they can later exploit.⁴ Klemperer (1987c) has also explored the relationship between consumer switching costs and entry deterrence in a two-period model. While in some cases an incumbent firm may be able to create an entry barrier by capturing a large fraction of the market before the threat of entry, in other situations the incumbent will deter entry by keeping his customer base *small* to remain an aggressive competitor.

While two-period models are instructive, they omit important features of competition in markets with switching costs. Sellers constantly compete both for attached customers and for new, uncommitted buyers; and the presence of each group affects the competition for the other, so long as the sellers cannot discriminate between new and old customers.⁵ The intensity of competition today depends on the future value of market share, and that value itself depends on the intensity of competition at yet later dates. Two-period models conceal the subtleties of this interdependence.

In Section 2 we present a simple overlapping-generations duopoly model with infinitely many periods. We show that the unique Markov perfect equilibrium displays the proentry tendency we have described: the seller without a customer base (the "entrant") is willing to price more aggressively than the "incumbent," and so attracts the new, unattached buyers. The roles of incumbent and entrant are then reversed in the subsequent period.

³ See Scotchmer (1986) and Green and Scotchmer (1986) for static models (equivalent to allowing all agents' discount factors to approach zero, and looking for a stationary equilibrium). Green and Scotchmer do consider entry, but do not focus on entry barriers. Von Weizsäcker (1984) used a many-period dynamic model, but did not consider closed-loop equilibrium.

⁴ Intense initial competition followed by feeble *ex post* competition occurs more generally when a first-period customer base confers a second-period advantage, for instance when learning-by-doing or network externalities are important. See Katz and Shapiro (1986), for example.

⁵ In our model, sellers would like to offer discounts to new buyers. This is impossible if existing customers can pretend that they are new, possibly by establishing a fictitious new entity, or if the good can be resold.

We extend this basic model to examine entry barriers with switching costs. The proentry tendency persists even when there are moderate economies of scale (or the demand-side equivalent, network externalities). Far from forming an entry barrier, switching costs induce excessive entry: even when it would be more efficient for production to be concentrated in one firm, and the incumbent consequently has a cost advantage in serving unattached buyers, entry occurs. When economies of scale are great, however, there is no entry in equilibrium. The switching costs protect the incumbent from the entrant's competition for attached buyers, while the economies of scale make it unattractive for the entrant to enter and serve only the unattached buyers as well as unattractive for the incumbent to set a price that encourages him to do so. We show how large scale economies must be in comparison with switching costs for an incumbent to choose to deter entry.

2. A duopoly model of competition with switching costs

■ **Description of the model.** Our model has two sellers, each of whom can produce a good at a constant average cost which, for simplicity, we take to be zero.⁶ In each of infinitely many (discrete) periods a cohort of identical young buyers or “youngsters” enters the market. We normalize the size of each cohort to unity.⁷ Each consumer lives for two periods, so at every date a cohort of “oldsters” is in the market. Each youngster buys from the seller offering the lower price.⁸ Therefore, given our assumption of identical buyers, one of the firms sells to all of the youngsters: it is this simplification that makes our model tractable. At the beginning of any period we call the firm that sold to the young buyers in the previous period the *incumbent*, and the other firm the *entrant*.

If an oldster switches and meets his second-period requirements from the entrant, he incurs a switching cost s , whose value is the same for all buyers and is common knowledge. Each buyer requires one unit of the good in each of the two periods he is in the market,⁹ and tries to minimize his outlays, including switching costs, with second-period outlays discounted at rate $\delta < 1$. Each seller maximizes the present value of his profits and discounts at rate δ . We look for perfect Markov equilibria, i.e., those in which the strategies depend only on who is the incumbent, and not otherwise on history.¹⁰

In each period each firm sets its price. We assume that firms cannot discriminate between youngsters and oldsters (see footnote 5). The timing of price-setting within each period is endogenous: each firm may set its period- t price at any time within period t , but once a price is set, it cannot be changed in that period. Thus, either firm may take the lead and set its price first, or wait until the other does so. After both period- t prices are set, consumers make their selections for that period.

We show below in Proposition 1 that in equilibrium the incumbent takes the lead. Both the incumbent (weakly) and the entrant (strictly) prefer this; the entrant does not prefer to preempt the incumbent's price-setting, and the incumbent does not want to outwait

⁶ If average cost is constant and positive, prices will simply be that much higher. Section 3 deals with the case where average costs decline in output.

⁷ In a growing or shrinking market, the numbers of youngsters and oldsters will not be equal. We discuss this case below.

⁸ We shall show below that, in equilibrium, this simple and plausible rule is fully optimal. It may not, however, be optimal in response to deviations by firms. We also discuss consumers' response to equal prices—our tie-breaking rules—below.

⁹ For simplicity, we suppose that buyers' reservation prices are so high as never to bind. One can use our methods to calculate equilibrium prices when the buyer's reservation price does bind in equilibrium. In the basic model the qualitative results are little affected if the reservation price binds.

¹⁰ We thus rule out punishment threats, which (for δ near unity) could sustain any outcome in which both sellers make positive profits. The equilibrium we study, the unique perfect-Markov equilibrium, is also the unique limit of subgame-perfect equilibria in the finite-horizon game as the horizon becomes long.

the entrant.¹¹ Proposition 1 also shows that in equilibrium the entrant sells to the youngsters and the incumbent sells to the oldsters.

□ **Solution of the basic model: incumbent leads and alternation.** We analyze the model by using dynamic programming. Since we are considering only Markov equilibria, we can let W_0 and W_1 denote the net present value (at the beginning of a period) of being the entrant and the incumbent, respectively. Our first proposition characterizes the equilibrium.

Proposition 1. In the basic model, there is a unique Markov-perfect equilibrium. The incumbent moves first, and sells to his existing customers; the entrant serves the new buyers.

Proof. Our proof proceeds in three parts. First, we characterize the pricing, sales pattern, and profits that would prevail if the incumbent were to set his price before the entrant does so. Then we analyze what would happen if instead the entrant were to move first. Finally, comparing profits under these two patterns of timing, we show that both firms prefer the incumbent to move first.

Suppose that the incumbent sets his price, p , before the entrant during the current period. To determine what price the incumbent will set, we must first find the entrant's price response, q , to each possible p . In response to p the entrant will do one of three things: (i) abandon the market this period by setting $q > p$; (ii) abandon the oldsters but attract the youngsters by setting $q = p$; or (iii) sell to all current customers by setting $q = p - s$.¹² These responses yield E payoffs of δW_0 , $p + \delta W_1$, and $2(p - s) + \delta W_1$, respectively. As a result, the incumbent sells to both cohorts if $p < \min [-\delta(W_1 - W_0), s - \delta(W_1 - W_0)/2]$, and sells to the oldsters but not to the youngsters if $-\delta(W_1 - W_0) \leq p \leq 2s$. If $p > 2s$, the incumbent sells to neither cohort.

The incumbent can therefore enjoy a (present value) profit of $2s + \delta W_0$ by setting a price $2s$. In contrast, he can get no more than δW_0 by setting a price above $2s$. And if he were to exclude the entrant altogether, then since $p < s - \delta(W_1 - W_0)/2$, his payoff, $2p + \delta W_1$, is no larger than $2s + \delta W_0$, which he could earn by setting a price of $2s$ and conceding the youngsters to the entrant. (In fact, his profits from setting $p = 2s$ are strictly higher than what he could earn by excluding the entrant.) We conclude that the incumbent's unique best price is $p = 2s$, and his value is $2s + \delta W_0$. The entrant's best response is to match the incumbent's price of $2s$. Thus, he sells to the youngsters but not to the oldsters. His payoff is $2s + \delta W_1$.

Now suppose instead that the entrant moves first during the current period and sets price q . As above, the incumbent's best response must be either $p = q$, $p = q + s$, or $p > q + s$, capturing respectively both cohorts, just the oldsters, or neither cohort. The entrant cannot profitably exclude the incumbent entirely: the incumbent would bow out for the period only if $\max [q + s + \delta W_0, 2q + \delta W_1] < \delta W_0$, and then the entrant's payoff $2q + \delta W_1$ would be less than δW_0 , his payoff from conceding all sales this period. The entrant's payoff, if he sets so high a price that the incumbent responds by matching and selling to both cohorts of buyers, is also δW_0 . These payoffs certainly will not tempt the entrant to move first.

To complete the proof, we now show that the entrant cannot profitably move first and set an intermediate price. To sell to the youngsters but not to the oldsters, he must set a price q such that the incumbent's best response is to charge $q + s$ and divide the market.

¹¹ Having the current incumbent move first also represents the choice of timing that the sellers would unanimously pick at the beginning of the game if they could "choose the rules" once and for all. A similar preplay "vote" has been studied recently by Klemperer and Meyer (1986).

¹² In all cases we avoid (by choice of tie-breaking rules) the trivial open-set problems that arise here when one firm wants to set its price at "the highest price below p ." These same open-set problems arise, for example, in Bertrand competition with unequal costs.

This requires that $q + s + \delta W_0 \geq \max [2q + \delta W_1, \delta W_0]$, or $-s \leq q \leq s - \delta(W_1 - W_0)$. If this strategy is feasible, i.e., if $\delta(W_1 - W_0) \leq 2s$, then it gives the entrant a payoff of $s - \delta(W_1 - W_0) + \delta W_1 = s + \delta W_0 > \delta W_0$. But both this and the corresponding payoff of $2s - \delta(W_1 - W_0) + \delta W_0$ for the incumbent are less than their payoffs if the incumbent moves first, and so the unique endogenous-timing equilibrium is for the incumbent to move first. *Q.E.D.*

In equilibrium, then, all consumers pay the same price,¹³ and the firms alternate in selling to unattached buyers. The prices and payoffs are given by

$$\left. \begin{aligned} p &= 2s \\ W_0 &= \frac{2s}{1 - \delta} \\ W_1 &= \frac{2s}{1 - \delta} \end{aligned} \right\} \quad (1)$$

Intuitively, the incumbent does not wish to price so low as to exclude the entrant from the market: to do so would mean making no profits, since the entrant competes on equal terms for the youngsters. The incumbent prefers to set a higher price and to exploit the oldsters, even though he knows that the entrant will match the price and win the youngsters.¹⁴ In Section 3 we shall examine what happens if the incumbent has a cost advantage in selling to the youngsters.

Prices and profits unambiguously increase in the switching cost s . The firms thus have a joint interest in raising consumers' switching costs, for example, through product-design decisions. This contrasts with Klemperer's (1987b) finding (see also Katz and Shapiro (1986)) that increases in s may so exacerbate first-period competition that firms would prefer to keep s low. This is so since W_1 does not exceed W_0 , and therefore there is no incentive to fight hard over market share.

In fact, the asymmetry between the sellers enables them to divide the market, with one firm specializing in its existing customers and the other specializing in the new arrivals. The duopolists are less direct rivals than they would be in the absence of switching costs. We might say that the switching costs and the asymmetric positions of the two firms facilitate implicit collusion: the incumbent is a natural server of the oldsters, and this fact reduces competition for the youngsters as well.

□ Growing demand. We can easily extend our basic model to the case in which every generation of buyers is larger by a constant factor γ than the previous generation. Of course, $\gamma > 1$ corresponds to a growing market.

Arguments similar to those leading to Proposition 1 show that the incumbent still moves first, provided $\gamma \geq 1$. The incumbent sets p so that the entrant is indifferent between matching p and thus earning $\gamma p + \delta W_1$, and undercutting p and thus earning

¹³ We claimed above that it was rational (in equilibrium) for young buyers to patronize the cheaper seller. We can now see why. If a youngster were to buy from the more expensive seller, he would still pay $p = 2s$ when he grew old. Therefore, in contrast to Farrell (1986) or Klemperer (1987a), there is no incentive for a young buyer to flee the crowd.

¹⁴ If the entrant were to set his price first (Farrell and Shapiro, 1987a), the incumbent's reluctance to compete would take a more transparent form. Because the incumbent's locked-in buyers are also paying his price, his marginal gain from a price cut that brings in the youngsters is always strictly less than the entrant's. Consequently, the entrant can set a profitable price that the incumbent will not choose to match. We call this a *common-margin* phenomenon: the youngsters are a common margin over whom the sellers compete, but one of them has other considerations (the inframarginal oldsters) in setting his price. For similar two-period, common-margin results, see Farrell (1986) and Klemperer (1987a). This is also related to the "judo economics" of Gelman and Salop (1983).

$$(1 + \gamma)(p - s) + \delta W_1;$$

hence $p = (1 + \gamma)s$. In particular, the equilibrium price, p , is increasing in γ , the growth rate.

This finding is counterintuitive: one would expect that a larger value of γ , corresponding to a market in which relatively more buyers are unattached, would lead to more intense competition and a lower price. In our model, however, price is determined by the entrant's willingness to undercut and to attract the oldsters: the incumbent "limit-prices" against that threat. When γ is large, the entrant is less eager to undercut, because the oldsters are relatively less valuable compared with the youngsters; so prices are higher. In a growing market the incumbent knows that his rival will attach relatively great importance to the new, unattached buyers. This knowledge allows the incumbent to set a higher price for oldsters without tempting the entrant to respond aggressively, and this in turn permits the entrant to charge the youngsters a higher price. In other words, larger values of γ make the implicit collusion (dividing the market between new and old customers) more effective.

□ **Summary.** We can summarize the lesson of this section as follows. With identical firms a seller with attached buyers will exploit his consumers and leave the "new" market to firms without a customer base. More generally than in our model, we conjecture that (absent such offsetting factors as efficiency differences and economies of scale), firms with larger shares of attached customers will tend to attract *smaller* shares of new, unattached customers: a "dynamic fat-cat effect." In the duopoly case this tendency emerges in the form of an alternating equilibrium. In a competitive model incumbent firms specialize in their existing buyers and do not sell to unattached consumers.¹⁵ Our specialization result thus appears to be fairly robust.

In a more general model with several firms and product heterogeneity, we conjecture that our specialization/alternation result would find an analogy in the form of a *stabilizing* tendency: if one firm gained a lead in market share, it would set a higher price and let its share erode over time. In the Appendix we prove this result for a wide class of duopoly models with consumer inertia.

In our equilibrium there is no switching. Since our model assumed inelastic demand, this means that the equilibrium is efficient. In a more realistic model the fact that prices are above marginal cost would cause allocative inefficiency. We ignore this effect to focus on possible inefficiencies of entry.

3. Economies of scale or network externalities

■ As we have shown, when costs are constant and equal across firms, switching costs do not block entry. Indeed, they actually *encourage* entry: the entrant has a strict strategic advantage in selling to the youngsters. In this section we formalize this by showing that the alternating tendency persists even when the incumbent, given that he sells to the oldsters, has an advantage in selling to the youngsters too. We consider *economies of scale* in the form of a fixed cost. With a fixed cost there is a cost saving if the same seller supplies both cohorts in a given period.¹⁶

We assume that either firm must incur a fixed cost f if it produces a positive quantity. We retain our assumption of constant marginal costs, which are again set at zero for sim-

¹⁵ See Farrell and Shapiro (1987a) for a model with many firms and with buyers who are in the market for many periods. So long as there are no economies of scale, the firm or firms with attached buyers specialize in serving them, with entrants' attracting new buyers.

¹⁶ An alternative interpretation of the analysis below is that there are economies of scale on the demand side: each buyer is better off if the other current buyers also patronize the same seller: then youngsters, resigned to the fact that oldsters will not likely switch, will be willing to pay more to the incumbent than to the entrant. With our simple demand structure, such network externalities are equivalent to conventional economies of scale.

plicity. Since each cohort size is normalized to unity, f represents the fixed costs per member of a given cohort. We seek conditions under which the incumbent uses his cost advantage in selling to the youngsters to exclude the entrant completely, and conditions under which the alternating tendency persists.

As above, the incumbent will move first. The reason, as we show below, is essentially the same as before: the entrant is less inclined to undercut a price than the incumbent is to match one, so the incumbent can afford to set a higher price if he goes first than can the entrant if he does. This benefits both sellers; it is an implicit form of collusion.

With economies of scale, $f > 0$, it is always possible for the incumbent profitably to exclude the entrant. But it may not be optimal for the incumbent to pursue such a strategy. We proceed by determining for what values of f , s , and δ the incumbent finds it *optimal* to do so. A partial answer is given by the following lemma.

Lemma 1. Suppose that the firms expect continuation values of W_1 and W_0 beginning next period. Then the incumbent firm will set his price so as to exclude the entrant from the market if and only if $f > s + \delta(W_1 - W_0)/2$.

Proof. As above, the entrant's optimal response to the incumbent's price p is either $q = p - s$, $q = p$, or $q > p$. To exclude the entrant, the incumbent must set a price p such that the entrant's optimal response is $q > p$, which requires that

$$\max [2(p - s) - f + \delta W_1, p - f + \delta W_1] \leq \delta W_0.$$

The incumbent's profits, $2p - f + \delta W_1$, are therefore given by

$$\min [f - \delta W_1 + 2\delta W_0, 2s + \delta W_0]. \tag{2}$$

This exceeds δW_0 if and only if $f > \delta(W_1 - W_0)$, a condition which (as we shall see) always holds in equilibrium. In other words, the incumbent can always exclude the entrant and make pure profits (relative to conceding the market this period).

Does the incumbent have a better option than excluding the entrant? The former could set a price that the entrant will match but not undercut and thus share the market. Such a price must satisfy

$$p - f + \delta W_1 \geq \max [2(p - s) - f + \delta W_1, \delta W_0].$$

A necessary and sufficient condition for such a price p to exist is that

$$f - \delta(W_1 - W_0) \leq 2s. \tag{3}$$

Clearly, if this strategy is feasible (i.e., if (3) holds), then we can evaluate it by assuming that the incumbent sets the highest price that works, which is $2s$. His profits, therefore, are $2s - f + \delta W_0$.

Will the incumbent choose to exclude the entrant? To answer this we must compare the profits from exclusion and from market sharing. Exclusion is the more profitable if and only if $\min [f - \delta W_1 + 2\delta W_0, 2s + \delta W_0] > 2s - f + \delta W_0$, or $2f - 2s > \delta(W_1 - W_0)$, which is equivalent to

$$f > s + \delta(W_1 - W_0)/2. \tag{4}$$

Finally, note that if (4) fails, then (3) certainly holds. In other words, if the incumbent does not find it optimal to exclude, then he indeed can find a price at which the market will be shared. This means that (4) is in fact necessary and sufficient for the incumbent to choose to exclude the entrant. *Q.E.D.*

The incumbent is more likely to find exclusion optimal when f is large compared with s . This finding accords with intuition: when economies of scale are large, the incumbent firm's guaranteed customer base is more likely to be sufficient to keep out the entrant.

We can now characterize the incumbent's choices, given the prospective values of W_1 and W_0 . If (4) fails, then it does not pay the incumbent to exclude the entrant. Con-

sequently, the market is divided, and the price is $2s$. The entrant's payoff is $2s - f + \delta W_1$, and the incumbent's is $2s - f + \delta W_0$. This is our familiar *alternating equilibrium* from the previous section, with the modification that each firm incurs the fixed cost, f .

If (4) holds, then we have a *dominant equilibrium* in which the incumbent does exclude the entrant. The price charged differs according to what threat constrains the incumbent's pricing. The entrant might enter "rapidly," undercutting the incumbent's price p by s to exploit the economies of scale; this strategy affords the entrant a payoff of

$$2(p - s) - f + \delta W_1.$$

Alternatively, the entrant might enter "gradually" by matching p and selling initially only to the youngsters; this approach gives the entrant a payoff of $p - f + \delta W_1$. The incumbent sets the highest price p subject to neither of these entry strategies' being attractive to the entrant; the analysis differs according to which binds.

To find the relevant condition here, imagine p gradually increasing from a low level. If (3) holds, as in Figure 1, then gradual entry is the entry strategy that first becomes attractive. Conversely, if (3) fails, as shown in Figure 2, then rapid entry becomes attractive for some value of p lower than that which would make gradual entry attractive. Therefore, there are two types of dominant-equilibrium pricing: pricing against gradual entry if (3) holds, and pricing against rapid entry if (3) fails.

If (3) holds, then the incumbent must set a price p designed to deter gradual entry. This means that the constraint on p is $p - f + \delta W_1 \leq \delta W_0$, so $p = f - \delta(W_1 - W_0)$. In Figure 1 the incumbent sets the price corresponding to the intersection of the "no entry" and "gradual entry" lines at point A . This strategy gives the incumbent a payoff of $2p - f + \delta W_1 = f + 2\delta W_0 - \delta W_1$.

If (3) fails, then the incumbent must set p to deter rapid entry. The binding constraint is then $2(p - s) - f + \delta W_1 \leq \delta W_0$, so that $p = s + f/2 - \delta(W_1 - W_0)/2$. In Figure 2 the incumbent's price corresponds to the point B . The incumbent's payoff in this case is $2p - f + \delta W_1 = 2s + \delta W_0$.

FIGURE 1

ENTRANT'S PROFITS WHEN GRADUAL ENTRY CONSTRAINS THE INCUMBENT (EQUATION (3) HOLDS)

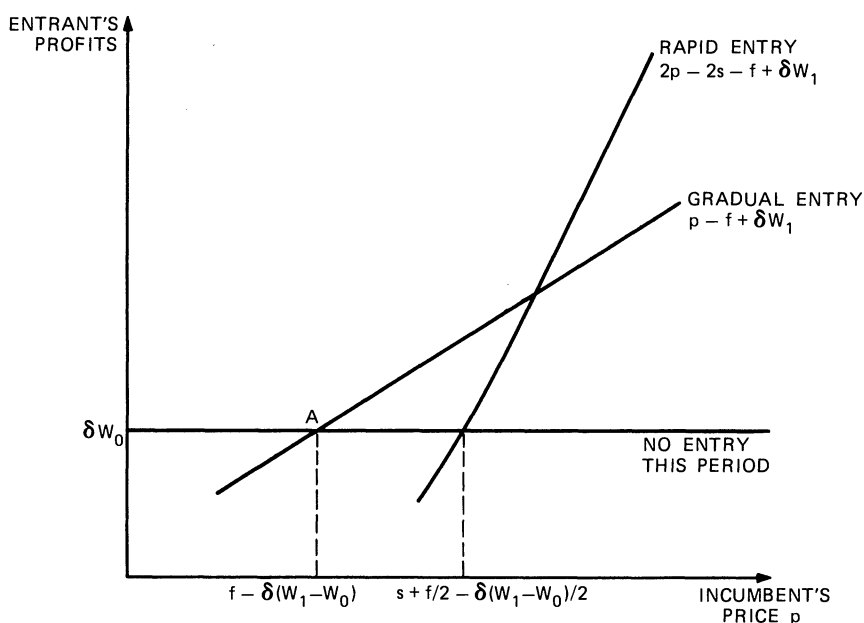
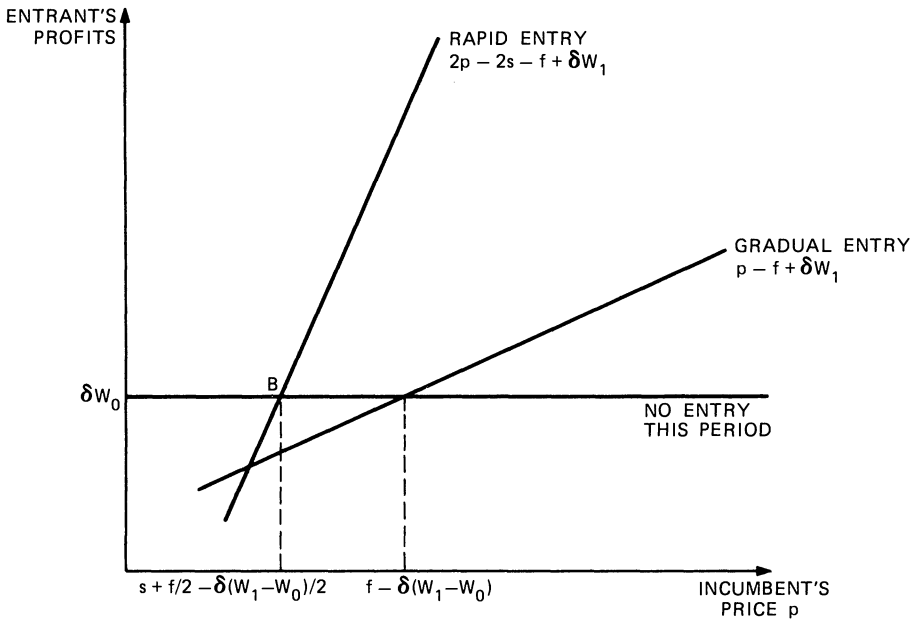


FIGURE 2
 ENTRANT'S PROFITS WHEN RAPID ENTRY CONSTRAINS THE INCUMBENT (EQUATION (3) FAILS)



□ **Characterization of stationary Markov equilibria.** We now complete our analysis by relating these payoffs to W_1 and W_0 : that is, by introducing the requirement that these hitherto arbitrary continuation values reflect the same behavior as we have predicted in the period under consideration.

Alternating equilibrium. An alternating equilibrium is one in which (4) fails. As in our basic model above, the firms' payoffs and the equilibrium price are given by

$$\begin{cases} W_0 = p - f + \delta W_1 \\ W_1 = p - f + \delta W_0 \\ W_0 = 2(p - s) - f + \delta W_1. \end{cases}$$

Solving these equations gives $W_1 = W_0 = 2s/(1 - \delta)$ and $p = 2s$. Hence, the condition that (4) fails becomes simply

$$f \leq s. \tag{5}$$

Thus, a necessary and sufficient for the existence of an alternating equilibrium is that the *per capita* fixed costs not exceed the switching cost. Of course, (5) was satisfied in our basic model, where $f = 0$. As in our basic model, observe that in the region where an alternating equilibrium prevails the firms each benefit from higher switching costs, as these reduce the competition between the two firms.

Dominant equilibrium. A dominant equilibrium is one in which condition (4) holds. Since the incumbent always excludes the entrant, $W_0 = 0$. Condition (4) then reduces to $2(f - s) \geq \delta W_1$. We must now calculate W_1 in terms of primitives to see for which parameter values this condition actually holds. This calculation differs, naturally, according to whether we are looking for a gradual-entry-constrained or a rapid-entry-constrained dominant equilibrium.

Gradual-entry dominant equilibrium. In a gradual-entry dominant equilibrium we found above that the incumbent's payoff is $f + 2\delta W_0 - \delta W_1 = f - \delta W_1$. Setting this expression

for the incumbent's payoff equal to W_1 , we find that $W_1 = f/(1 + \delta)$. The price charged by the incumbent is $p = f - \delta W_1$ or $p = f/(1 + \delta)$. In this type of equilibrium the incumbent's payoff is independent of the switching cost, and increasing in, but bounded above by, the size of the fixed costs.

We need to find for which parameter values conditions (3) and (4) hold, as they must in a gradual-entry dominant equilibrium. These conditions require that $f - 2s \leq \delta f/(1 + \delta)$ and $2f - 2s > \delta f/(1 + \delta)$. Thus, a gradual-entry-constrained dominant equilibrium exists if and only if

$$\frac{2(1 + \delta)}{2 + \delta} < f/s \leq 2(1 + \delta). \quad (6)$$

Rapid-entry dominant equilibrium. In a rapid-entry dominant equilibrium we found above that the incumbent's payoff is $2s + \delta W_0$. Setting $W_0 = 0$ gives us $W_1 = 2s$. The incumbent's price that supports this equilibrium is $p = f/2 + s(1 - \delta)$. With these values for W_0 and W_1 , our requirements for a rapid-entry equilibrium, namely that (4) but not (3) hold, give us

$$f/s > 2(1 + \delta). \quad (7)$$

In a rapid-entry dominant equilibrium the incumbent's payoff, $2s$, is increasing in the switching cost, independent of the size of the fixed cost, but again bounded above by the size of the fixed cost.

Mixed-strategy equilibrium. For $f/s \in (1, 2(1 + \delta)/(2 + \delta))$ there is no pure-strategy Markov equilibrium. This is so since condition (4) carries within it the seeds of its own destruction: if $W_1 - W_0$ is large, then (4) fails, and so the incumbent chooses to split the market with the entrant. But then $W_1 = W_0$, and (since $f > s$) this makes (4) hold. Intuitively, if incumbency is worth little or nothing in the next period, then the entrant is easy to bar, and so the incumbent can keep him out—incumbency is valuable. But if incumbency is valuable, then the entrant is relatively keen to enter, and so the incumbent must offer a price that the entrant will match, which equalizes their payoffs for this period.

A mixed-strategy equilibrium exists in which (4) holds with equality; the incumbent randomizes between excluding the entrant and sharing the market. The probabilities can be chosen to make W_0 have the correct value so that (4) does indeed hold with equality (and thus the incumbent is willing to randomize).

There do exist pure-strategy Markov equilibria, but none is stationary. Their nature can be seen by looking at our game with a long but finite horizon. With a finite horizon the equilibrium would be as follows (working back from the end): in the last period, since $W_1 - W_0 = 0$, (4) holds, so that the incumbent excludes the entrant and makes positive profits. Working backwards, therefore, so long as the incumbent excludes his rival, $W_1 - W_0$ is growing. At some point far enough from the end, that net value of incumbency is large enough that it no longer pays the incumbent to price as low as he must to exclude the entrant: (4) fails. A phase of alternating behavior ensues. During that phase, the two firms are making equal current profits, and (still working backwards in time) the end-game effect that makes incumbency valuable recedes into the mists of discounting. When it has receded far enough, $W_1 - W_0$ becomes small enough to make (4) hold again, and a phase of dominance ensues. And so on.

□ **Checking that the entrant will not move first.** Now we must check that neither firm would want the entrant to move first.

Could the entrant earn better profits by being the price leader? As in the basic model, he cannot profitably drive the incumbent from the market this period. And he would not strictly want to move first to set a price that the incumbent would match. So we need only consider one possible defection: the entrant moves first and sets a price q such that the incumbent chooses to charge $q + s$.

Such a strategy would require a price no greater than $s - \delta(W_1 - W_0)$, and would yield the entrant a payoff no greater than $s - \delta(W_1 - W_0) - f + \delta W_1 = s - f + \delta W_0$. Since a market-splitting incumbent prices at $2s$, this defection cannot be profitable for the entrant unless the incumbent would otherwise exclude him from the market, in other words, unless (4) holds. But if the defection is profitable, its payoff must exceed δW_0 , so that $s > f$, in which case (4) cannot hold. We conclude that the entrant would never (strictly) want to preempt the incumbent's price-setting.

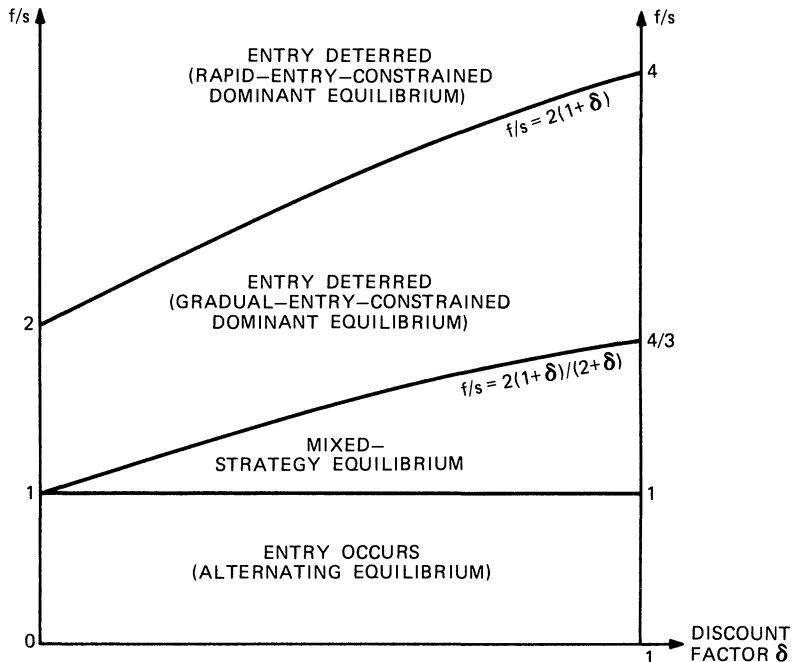
□ **Summary and interpretation.** We can summarize our findings in this section with the following proposition.

Proposition 2. With economies of scale the incumbent acts as a price leader. If *per capita* fixed costs are no larger than consumer switching costs ($f/s \leq 1$), the incumbent sets a price that splits the market, as in the absence of scale economies. For larger fixed costs his cost advantage in serving the youngsters leads to an equilibrium in which he excludes the entrant; this is the equilibrium if $f/s \geq 2(1 + \delta)/(2 + \delta)$. For f/s between 1 and $2(1 + \delta)/(2 + \delta)$, there is a mixed-strategy Markov equilibrium in which the incumbent randomizes between excluding entry and not.

The equilibrium structure of the model with economies of scale is shown in Figure 3. The key parameters are f/s and δ .

It is easy to reinterpret our results to apply them to markets with network externalities rather than economies of scale. Suppose that each consumer would pay a premium n for a product consumed by both cohorts, in comparison with one consumed only by his own cohort. With $n > 0$, efficiency calls for a single firm to serve all consumers, just as when $f > 0$. For moderate n , however, this will not happen: the incumbent will prefer to exploit switching costs, and unattached buyers will patronize the entrant.¹⁷ In other words, an

FIGURE 3
EQUILIBRIUM STRUCTURE WITH SCALE ECONOMIES



¹⁷ Of course, our discussion here applies only when there is market power in the established technology, i.e., when the established technology is sponsored.

established technology with an installed base of users may succumb to a new technology that is no better,¹⁸ merely because the established one is monopolized, and buyers find switching costly. This result contradicts the common intuition that an established technology is likely to survive, even in the face of superior alternatives. Although the owner of the incumbent technology *could* profitably exclude entry, this strategy is not necessarily or typically the optimal one.

We have identified two rather different patterns of pricing and sales in the presence of switching costs and scale economies. If switching costs are more significant than scale economies ($s > f$), the incumbent exploits his locked-in oldsters and concedes the youngsters to the entrant. In such an alternating equilibrium the fixed costs f have no effect on prices, although they do affect profits. The entry-inducing effect of switching costs causes a productive inefficiency, since it is costly to divide output between incumbent and entrant.¹⁹ In contrast, when scale economies are substantial ($f > s$), the incumbent firm keeps the entrant out, despite the entrant's willingness to price below (his) cost. The nature of the limit pricing varies according to whether the entrant is more willing to enter at a high price but an inefficiently small scale or to enter at a low price, tear the oldsters away from the incumbent, and exploit the scale economies.

4. Conclusion

■ We have studied a model of dynamic competition in which, because of specific investments, buyers become “locked-in” to sellers. We found that, as in two-period models, a firm with more existing users will be less aggressive in seeking out new customers. In our model this tendency produced an alternating equilibrium: the incumbent specialized in serving its attached buyers and left the unattached buyers to its rivals.

More generally, we find that in markets with switching costs, various factors—such as economies of scale, efficiency advantages, or network externalities—that make it *possible* for an incumbent firm to exclude entry profitably will typically not make it *optimal* to do so. Intuition based on what established firms *can* do, rather than what they will optimally do, can be misleading. So, for example, a somewhat inferior product may successfully enter a market in which consumers face switching costs and there are economies of scale. Despite what might appear to be three barriers to entry—switching costs, a product disadvantage, and economies of scale—excessive entry occurs in equilibrium. Although switching costs surely make it harder for entering firms to attract established buyers, they actually encourage entry into the market to serve unattached ones.

Entry occurred in our model even when efficiency in the presence of scale economies called for the incumbent to serve all buyers. The same proentry effect can be found when an incumbent firm is simply more efficient than the entrant (quite apart from any scale economies). In that case (Farrell and Shapiro, 1987a), the inefficient firm enters as the incumbent specializes in serving his existing customers. Thus, switching costs, by inducing excessive entry, can cause inefficiency even though switching never occurs in equilibrium and even though we assume inelastic demand. Of course, large economies of scale eliminate this inefficiency, since the incumbent sells to all customers. We have identified the circumstances in which this dominance by the incumbent occurs. In our model dominance is efficient, although in a richer model with elastic demand we would, of course, find an allocative inefficiency because the incumbent can price above marginal and average cost without inducing entry.

¹⁸ Indeed, even a somewhat inferior technology may successfully enter into a market with network externalities and switching costs.

¹⁹ Our finding is thus similar to that in Klemperer (1987d), where in a static model a less efficient firm may enter with knowledge that the incumbent will be tempted to set a high price to exploit his existing customers.

We conclude that switching costs alone *do not* form an entry barrier. In *combination* with economies of scale, however, they may enable an incumbent firm to exclude competitors and still make positive economic profits.²⁰ This is one definition of an entry barrier.

Another definition of *entry barrier* reserves the term for situations in which socially beneficial entry does not occur. Since in our model entry is inefficient, we have not identified entry barriers in this sense. We have, however, explored this possibility in separate calculations by including, as well as switching costs and scale economies, *cost differences* between the two firms. In particular, we have asked whether a higher-cost firm can and will retain its incumbency (on the basis of switching costs and scale economies), even when this outcome is inefficient.²¹ We have shown (by construction) that this type of barrier to entry can indeed occur. In our “network externalities” interpretation, this possibility has been called *excess inertia* (Farrell and Saloner, 1986; Katz and Shapiro, 1986).

Although the alternation tendency in our model is extreme, its continuous version—bigger firms’ losing market share to smaller rivals when there are consumer switching costs and turnover of buyers—seems likely to be present in a variety of dynamic models involving consumer switching costs, even when economies of scale are also present. We show in the Appendix that this conjecture holds for a wide class of models, so long as there are diminishing returns to market share in equilibrium.

Our findings suggest that in some markets with switching costs, such as those for computer systems, there will be a natural product (or firm) life cycle. Products enter by attracting unattached buyers, through offering an attractive price/quality mix. Later, the product may continue to attract new clients, but it will increasingly rely on established buyers as it becomes less attractive relative to younger products, either on account of a higher price or a lower quality (which, in the presence of technological change, may be directly related to vintage). In its final phase a product will almost exclusively serve locked-in buyers, since it will have lost its appeal to unattached buyers. As the locked-in buyers leave the market, or as they make new product-specific investments, the product loses its customer base and is withdrawn.

In other markets in which gradual adjustment of price is more important than generational technological leapfrogging, the tendency for firms with many existing consumers to set high prices and exploit these customers will lead to a stable steady state. If a product were to gain a market-share lead, its price would rise relative to its rivals’, and its share of new customers would decline and return its market share to some equilibrium value. This should lead to stability in market shares. The temptation to exploit existing customers will cause more efficient firms to maintain smaller market shares than would be efficient.²² In our formal model this tendency appeared in the form of entry by less efficient firms.

What do these strategic considerations imply for observed market behavior? Our findings suggest that there will be a tension between two effects: (1) firms (or products) with attached customers will tend to offer less attractive terms and thus will lose market share, but (2) these very firms may have achieved their position on the basis of superior efficiency or an ability to exploit scale economies, in which case they may not be replaced. Switching costs *per se* create a negative correlation between existing market share and share of new placements, but efficiency differences create a positive correlation between these two.

Appendix

■ In this Appendix we prove in a general duopoly model that a firm with a larger market share sets a higher price and sees its share erode over time. To show this we have been obliged to make two nonprimitive (but plausible)

²⁰ But the incumbent’s profits in any such dominant equilibrium are bounded above by the size of the fixed costs.

²¹ Exclusion of a lower-cost firm is not necessarily inefficient, since there are real, albeit transitory, social costs attached to entry: either oldsters must switch, or production must for a period be inefficiently divided between firms.

²² This is also true of conventional Cournot oligopoly.

assumptions: that market share is positively valuable, and that there are diminishing returns to market share, i.e., in equilibrium a firm's value is a concave, increasing function of its market share. Consider a symmetric, continuous-time duopoly model in which the state variable is firm 1's market share, σ , and each firm sets a price at each date. Let the total number of consumers be normalized to unity. Call $V(\sigma)$ firm 1's value; thus, firm 2's value is $V(1 - \sigma)$.

Assume that old consumers continue to buy from the firm they have patronized in the past. Then the flow of profits earned by a firm with σ consumers that sets price p are $\sigma\pi(p)$; here $\pi(p) \equiv pX(p) - c(x(p))$ is the per consumer profit function, where $x(p)$ is the demand function and $c(\cdot)$ is the cost function. Consumers leave the market at a rate of ϕ per unit of time, and are replaced by an equal number of unattached consumers.

We suppose that new consumers shop purely on the basis of price.²³ Denote by $f(p - q)$ the fraction of new buyers who select firm 1 when its price is p and its rival's is q ; of course, $f' < 0$. With these assumptions the equation of motion for a firm's market share is

$$\dot{\sigma} = -\phi\sigma + \phi f(p - q)$$

if it sets price p and its rival sets price q .

Write $p(\sigma)$ and $q(\sigma) = p(1 - \sigma)$ for the equilibrium prices. The Hamiltonian is

$$H = \sigma\pi(p) + \lambda(-\phi\sigma + \phi f(p - q)),$$

where λ gives the value of market share, $V(\sigma)$. The price $p(\sigma)$ satisfies $\partial H/\partial p = 0$, or

$$\sigma\pi'(p) + V'(\sigma)\phi f'(p - q) = 0. \quad (A1)$$

We begin by showing that for nearly equal market shares, the market leader sets the higher price, i.e., that $p'(1/2) > 0$. Differentiating (A1) with respect to σ and using subscripts to denote partial derivatives, we have

$$H_{pp}p'(\sigma) + H_{pq}q'(\sigma) + H_{p\sigma} = 0. \quad (A2)$$

At $\sigma = 1/2$, $q'(\sigma) = -p'(\sigma)$, so that (A2) becomes

$$(H_{pp} - H_{pq})p'(\sigma) + H_{p\sigma} = 0.$$

Now from (A1), assuming that $V'(\cdot) > 0$, we obtain $\pi'(p) > 0$. Using that and assuming that $V''(\cdot) \leq 0$, we find that $H_{p\sigma} > 0$. Hence, by (A2) $p'(1/2)$ has the same sign as $H_{pq} - H_{pp}$. By the second-order condition, $H_{pp} < 0$. Moreover, $f(\cdot)$ is symmetric (i.e., $f(-x) = 1 - f(x)$), so that $f''(0) = 0$, and hence $H_{pq} \equiv -V'\phi f'' = 0$. Therefore, $p'(1/2) > 0$.

This already establishes the important fact that the game is locally stable around $\sigma = 1/2$: small advantages in market share tend to be dissipated in high pricing. With only a little more work, however, we can prove the stronger result (implying global stability) that the "premium" $p(\sigma) - q(\sigma) \equiv p(\sigma) - p(1 - \sigma)$ is everywhere increasing in σ .

Mathematically, this claims that $p'(\sigma) + p'(1 - \sigma) > 0$ for all σ . We have shown that this is true for $\sigma = 1/2$, so that we need only show that that expression never changes sign. Were it to do so, we would have

$$p'(\sigma) + p'(1 - \sigma) = 0.$$

Combining this with (A2) (and using the fact that $q'(\sigma) = -p'(1 - \sigma)$) shows that $p'(\sigma) > 0$. But we made no assumptions distinguishing σ from its complement $1 - \sigma$, so we must also have $p'(1 - \sigma) > 0$. This contradicts our assumption that $p'(\sigma) + p'(1 - \sigma) = 0$, and proves our assertion.

We have shown that the tendency for large firms to charge higher prices is no artifact of our discrete model, although we have had to make two nonprimitive (though plausible) assumptions about the equilibrium value function $V(\cdot)$. A complete analysis would require solving for the $V(\cdot)$ function in terms of exogenous variables, and checking that indeed this value function is concave.

References

- DIAMOND, P. "A Model of Price Adjustment." *Journal of Economic Theory*, Vol. 3 (1970), pp. 156-168.
- FARRELL, J. "A Note on Inertia in Market Share." *Economics Letters*, Vol. 21 (1986), pp. 73-75.
- AND SALONER, G. "Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation." *American Economic Review*, Vol. 76 (1986), pp. 940-955.
- AND SHAPIRO, C. "Dynamic Competition with Lock-In." Woodrow Wilson School Discussion Paper No. 121, Princeton University, February 1987a.
- AND —. "Optimal Contracts with Lock-In." Woodrow Wilson School Discussion Paper No. 130, Princeton University, August 1987b.

²³ Consumers may have no other information on which to base their decisions. In particular, they may not know the firms' market shares. But note that in any market in which our result holds, consumers would, at equal prices, rather buy from the firm with the smaller market share—since this firm can be expected to price more aggressively in the future.

- FISHER, F., MCGOWAN, J., AND GREENWOOD, J. *Folded, Spindled, and Mutilated: Economic Analysis and U.S. v. I.B.M.* Cambridge: MIT Press, 1983.
- FUDENBERG, D. AND TIROLE, J. "The Fat-Cat Effect, the Puppy-Dog Ploy, and the Lean and Hungry Look." *American Economic Review*, Vol. 74 (1984), pp. 361-366.
- GELMAN, J. AND SALOP, S. "Judo Economics: Capacity Limitation and Coupon Competition." *Bell Journal of Economics*, Vol. 14 (1983), pp. 315-325.
- GREEN, J. AND SCOTCHMER, S. "Switching Costs as an Explanation of Stable Price Dispersion." Harvard Institute of Economic Research Discussion Paper 1260, 1986.
- KATZ, M.L. AND SHAPIRO, C. "Technology Adoption in the Presence of Network Externalities." *Journal of Political Economy*, Vol. 94 (1986), pp. 822-841.
- KLEMPERER, P. "Markets with Consumer Switching Costs." *Quarterly Journal of Economics*, Vol. 102 (1987a), pp. 375-394.
- . "The Competitiveness of Markets with Consumer Switching Costs." *RAND Journal of Economics*, Vol. 18 (1987b), pp. 138-150.
- . "Entry Deterrence in Markets with Consumer Switching Costs." *Economic Journal, Supplement*, Vol. 97 (1987c), pp. 99-117.
- . "Welfare Effects of Entry into Markets with Consumer Switching Costs." St. Catherine's College, Oxford University, 1987d.
- AND MEYER, M. "Price Competition versus Quantity Competition." *RAND Journal of Economics*, Vol. 17 (1986), pp. 618-638.
- PORTER, M. *Competitive Strategy*. New York: Free Press, 1980.
- SCOTCHMER, S. "Market Share Inertia with More Than Two Firms." *Economics Letters*, Vol. 21 (1986), pp. 77-79.
- VON WEIZSÄCKER, C.C. "The Costs of Substitution." *Econometrica*, Vol. 52 (1984), pp. 1085-1116.
- WILLIAMSON, O. *The Economic Institutions of Capitalism*. New York: Free Press, 1985.