Coordination and Lock-In: Competition with Switching Costs and Network Effects

Part II of IV Switching Costs Section

December 2004

the latest version of this paper, and related material, will be at
www.paulklemperer.org

Joseph Farrell
University of California, Berkeley
549 Evans Hall # 3880
Berkeley, CA 94720-3880
USA

email: farrell@econ.berkeley.edu

and

Paul Klemperer
Nuffield College, Oxford University
Oxford OX1 1NF
England

Int Tel: +44 1865 278588

email: paul.klemperer@economics.ox.ac.uk

First draft: 1997
This draft: 2004

PRELIMINARY DRAFT: PLEASE SEND COMMENTS

1 Introduction

2 Switching Costs and Competition

2.1 Introduction

A consumer faces a *switching cost* between sellers when an investment specific to his current seller must be duplicated for a new seller.[1] That investment might be in equipment, in setting up a relationship, in learning how to use a product, or in buying a high-priced first unit that then allows one to buy subsequent units more cheaply (when firms' prices are non-linear). Switching costs may even be psychological.[2] Klemperer (1995) gives many examples of each of these kinds of switching costs, and Section 2.2 discusses empirical evidence for switching costs.

Switching costs may be *learning* costs, in which case consumer who switches from firm A to firm B has no switching cost of later buying from either firm. Alternatively, switching costs may be *transactional*, in which case a consumer who switches from A to B would incur an additional switching cost if he reswitched back to A (an example is the cost of returning rented equipment and renting from a new supplier). Of course, many switching costs have both learning and transactional aspects.

We will generally assume that switching costs are real social costs, but there can also be *contractual* or pecuniary switching costs (that are not social costs). Examples include airlines' "frequent-flyer" programs, and "loyalty contracts" that rebate a fraction of past payments to consumers who continue to patronise the firm. These pecuniary switching costs are a form of quantity discount or bundling. Lars Stole (forthcoming) discusses such price discrimination strategies elsewhere in this Volume, so we will focus mainly on "real" switching costs.

We assume consumers have perfect information about the existence and

[1] There can also be switching costs among different products of a single firm, as there were among IBM computers until the internally compatible System/360 family. But we (following the economics literature) focus on switching costs between firms.

[2] Social psychologists have shown that consumers change their own preferences in favour of products that they have previously chosen or been given, in order to reduce "cognitive dissonance" (Brehm, 1956).

qualities of all firms' products, even before purchasing any. So "new" consumers who have not yet developed an attachment to any particular product are especially important in markets with switching costs. In contrast, "search costs" directly affect even consumers' initial purchases. But search costs and switching costs have much in common, and models of the effects of switching costs can also apply to search costs. For example, either kind of friction makes a firm's market share important for its future profitability (see Section 2.6) and much empirical work does not distinguish between search and switching costs.[3] For a survey of search costs, see, for example, Stiglitz (1989) in Volume 1 of this Series.

"Experience-good" markets in which each consumer needs to purchase a product to determine its quality (see Nelson, 1970) and so prefers to repurchase a brand he tried and liked rather than try a new brand of unknown quality, also have much in common with switching-cost markets. But with experience goods, unlike with switching costs, complications can arise from the possibility of prices signaling qualities, and from the existence of consumers who disliked the product they last purchased.[4] [5]

Switching costs not only apply to repeat-purchases of identical goods. An important class of examples involves "follow on" goods, such as spare

[3]For example, empirical findings about the credit card (Ausubel (1991) etc. – see note CC) and telecommunications (see, e.g., Knittel (1997)) markets, and about the effects of firms' discount rates on prices (Froot and Klemperer (1989), Chevalier and Sharfstein (1996), Fitoussi and Phelps (1988) etc.) could be the result of either switching or search costs. On the other hand, Moshkin and Schachar (2000) develop a discrete-choice empirical model to estimate how many consumers behave as if they have switching costs and search costs respectively. Their test is based on the fact that whereas the switching probability of a consumer facing search costs depends on the match between his tastes and the attributes of the alternative he last chose, the switching probability of a consumer facing switching costs depends on the match between his tastes and the attributes of all available alternatives. Using panel data on television viewing choices, they suggest 72% of viewers act as if they have switching costs between TV channels, while 28% act as if they have search costs.

[4][Note V] Schmalensee (1982) and Villas Boas (2000) analyse models of experience goods that show similarities to switching costs models. Hakanes and Peitz (2003) and Doganoglu (2004) model experience goods when there are also learning or transactional switching costs; Doganoglu shows that adding small switching costs to Villas Boas' (2000) model can sometimes reduce price levels.

[5]For related models in which consumers differ in their "quality" from firms' point of view, and firms are uncertain about consumers they have not supplied and can exploit those they know to be of "high quality", see, for example, Nilssen (2000) and Cohen (2001) on insurance markets and Sharpe (1990) and Zephirin (1994) on bank loan markets.

parts and repair services, bought in "aftermarkets": buyers face additional "switching" costs if the follow-on goods are not compatible with the original purchase, as may be the case if they are not bought from the same firm.[6] [7] [8] [9]

Consumers may also incur switching costs, or "shopping costs", at a single date by buying related products from multiple suppliers rather than from a single supplier. (These "shopping costs" can be real social costs or contractual costs created by quantity discounts and "bundling".) However, most of the literature focuses on dynamic problems and emphasises the resulting commitment problems.

In the simplest cases, when firms can commit to future prices and qualities, a market with switching costs is closely analogous to a market with economies of scope in production; with switching costs each individual consumer can be viewed as a "market" with economies of scope between "purchases now" and "purchases later". Just as a market with production economies of scope is entirely captured by the firm with the lowest total costs in the simplest price-competition model, so in a simple model with complete contracts each individual buyer's lifetime's requirements in a market with

[6] Aftermarkets have been much studied since a US Supreme Court decision (ITS v. Kodak, [cite]) held that it was conceptually possible for ITS, an independent repair firm, to prove that Kodak had illegally monopolized the aftermarket for servicing Kodak photocopiers: see e.g. Shapiro (1995), Shapiro and Teece (1998), MacKie-Mason and Metzler (1999), and Borenstein, MacKie-Mason, and Netz (1995, 2000).

[7] [Note LM] Similar issues arise when retailers each advertise the prices of only some of their products (often the "loss leaders"), but expect consumers who enter their stores to buy other products also. See, for example, Lal and Matutes (1994) and Lee and Png (2004). (The unadvertised products may actually be the "quality" of the advertised product – see Ellison (2003) and Vickers (2003).) Gabaix and Laibson (2004) analyse this case when only some consumers are rational.

[8] Typically, a consumer who has not previously bought from any firm incurs a start-up cost similar to (or greater than) the new investment (switching cost) that a brand switcher must make. We will use the term "switching cost" to include these start-up costs. So a consumer may have a "switching cost" of making a first purchase. In many models consumers have high enough willingnesses to pay that this cost has little consequence since it does not affect consumers' preferences between firms.

[9] Sometimes costs of forming a new relationship fall upon the supplier, not (or as well as) on the customer, and firms' costs of serving new customers have parallels to consumers' switching costs (see Klemperer (1995)). Firms' switching costs have been less studied, but in some contexts, such as the simple model of the next subsection, the total prices (including any switching costs) paid by consumers are unaffected by whether firms or consumers actually pay the switching costs. Thus the equilibrium incidence need not coincide with the apparent incidence of switching costs.

switching costs are filled by the lowest-cost supplier of those requirements. That is, firms compete on "lifecycle" prices and the market lifecycle price is determined by lifecycle costs, with any subdivision of the lifecycle price being arbitrary and meaningless. In this case, the outcome is efficient and switching costs confer no market power on firms.

This simple analogy—including the efficiency of the outcome—can survive even if firms cannot credibly commit to future prices or qualities. But even small steps outside the simplest story suggest ways in which the analogy and the efficiency break down (Section 2.3). The analogy is still weaker if firms cannot discriminate between different customers (Section 2.4), or consumers use multiple suppliers (Section 2.5). After treating these cases (and having discussed empirical evidence in Section 2.2), we analyze the "market share" competition that switching costs generate (Section 2.6). All this discussion takes both the switching costs and the number of firms as exogenous, so we then consider entry (Section 2.7) and endogenous switching costs (Section 2.8), before addressing implications for competition policy (Section 2.9).

2.2 Empirical Evidence

The empirical literature on switching costs is much smaller and more recent than the theoretical literature. Some studies test specific aspects of the theory (see later sections), but only a few studies directly attempt to measure switching costs.[10]

Where micro data on individual consumers' purchases are available, a discrete choice approach can be used to explore the determinants of a consumer's probability of purchasing from a particular firm. Greenstein (1993) analyses federal procurement of commercial mainframe computer systems during the 1970s, and finds that an agency is likely to acquire a system from an incumbent vendor, even when controlling for factors other than the buyer's purchase history that may have influenced the vendor-buyer match;

[10]Experimental studies are even fewer and more recent, but include Cason and Friedman (2002), and Cason, Friedman, and Milam (2003). See note EX. The theoretical literature, by contrast, arguably began with Selten's (1965) model of "demand inertia" (which assumed a firm's current sales depended in part on history, even though it did not explicitly model consumers' behavior in the presence of switching costs), and then took off in the 1980s.

he suggests switching costs were an important source of incumbent advantage in this market.[11] Shum (1999) analyzes panel data on breakfast cereal purchases, and finds that households switching brands incur average implicit switching costs of $3.43—which exceeds every brand's price! (However he also finds advertising can be effective in attracting customers currently loyal to rival brands).

Because switching costs are usually both consumer-specific and not directly observable, and micro data on individual consumers' purchase histories are seldom available, less direct methods of assessing the level of switching costs are often needed. Kim et al. (2003) estimate a first-order condition and demand and supply equations in a Bertrand oligopoly to extract information on the magnitude and significance of switching costs from highly aggregated panel data which do not contain customer-specific information. Their point estimate of switching costs in the market for Norwegian bank loans is 4.12% of the customer's loan, which seems substantial in this market, and their results also suggest that switching costs are even larger for smaller, retail customers.[12] Shy (2002) argues that data on prices and market shares reveal that the cost of switching between banks varies from 0 to 11% of the average balance in the Finnish market for bank accounts. He also uses similar kinds of evidence to argue that switching costs in the Israeli cellular phone market approximately equal the price of an average phone.

One defect of all these studies is that none of them models the dynamic effects of switching costs that (as we discuss below) are the main focus of the theoretical literature; in effect, these empirical studies assume consumers myopically maximise current utility without considering the future effects of their choices.[13]

Other empirical studies, many of which we will discuss below in the context of specific theories, provide evidence for the importance of switching

[11] Breuhan (1997) studies the switching costs associated with the windows and DOS operating systems for personal computers.

[12] Sharpe (1997) studies the bank retail deposit market and argues that the data support the model of Klemperer (1987b). See also Waterson (2003).

[13] But Viard (2003) studies the impact of number portability on prices in the US market for toll-free numbers using a dynamic model in which consumers consider the future effects of their choices.

costs for credit cards (Ausubel (1991), Calem and Mester (1995), Stango (2002)); cigarettes (Elzinga and Mills (1998, 1999)); computer software (Larkin (2004)); supermarkets (Chevalier and Scharfstein (1996)); air travel, and alliances of airlines in different frequent-flyer programmes (Fernandes (2001), Carlsson and Löfgren (2004)); individual airlines for different flight-segments of a single trip (Carlton, Landes, and Posner (1980)); phone services (Knittel (1997), Gruber and Verboven (2001), Viard (2003)); television viewing choices (Moshkin and Schachar (2000)); online brokerage services (Chen and Hitt (2002)); electricity suppliers (Waterson (2003)); bookstores (Lee and Png (2004)); and automobile insurance (Schlesinger and von der Schulenberg (1993), Israel (2001), Waterson (2003)).

There is also an extensive empirical marketing literature on brand loyalty (or "state dependence") which often reflects, or has equivalent effects to, switching costs. Seetharam et al (1999) summarise this literature; a widely cited paper is Guadagni and Little's (1983) analysis of the coffee market.[14] Finally, Klemperer (1995) gives many other examples of markets with switching costs, and U.K. Office of Fair Trading (2003) presents useful case studies.

2.3 Firms who Cannot Commit to Future Prices

2.3.1 Bargains Followed by Ripoffs

The core model of the switching costs literature posits that firms cannot commit to future prices.

The simplest model has two periods and two symmetric firms, with costs c_t in periods $t = 1, 2$.[15] A single consumer has a switching cost s and reservation price $r_t > c_t + s$ for one unit of the period$-t$ good, firms set prices, and there is no discounting. Then in period 2 the firm that sold in period 1 will exercise its ex post market power by pricing (just below) $c_2 + s$ (the rival firm will offer price c_2 but make no sale). Foreseeing this,

[14] Jacoby and Chestnut (1978) survey earlier attempts in the marketing literature to measure brand loyalty. Theoretical marketing papers include Wernerfelt (1991) (see note W($\approx$28)), Villas Boas (2000) (see note V($\approx$10)), and Kim et al (2001) who study incentives to offer reward programs that create pecuniary switching costs.

[15] $c_2 \neq c_1$ is especially natural if the second-period good is spare parts/repair services/consumables for a first-period capital good.

It makes no difference if there are $n > 2$ firms.

firms are willing to price below cost in period 1 to acquire the customer who will become a valuable follow-on purchaser in period 2; undifferentiated competition to win the customer drives period-1 prices down to $c_1 - s$.

Note that in this simple model the consumer's expectations do not matter. Competition among non-myopic firms makes buyer myopia irrelevant.[16]

Although first-period prices are below cost, there is nothing predatory about them, and this pattern of low "introductory offers", or "penetration pricing" (see section 2.6), followed by higher prices to exploit locked-in customers is familiar in many markets. For example, banks offer gifts to induce customers to open new accounts, and Klemperer (1995) gives more examples.[17] This "bargains-then-ripoffs" pattern is a main theme of many two-period models in the switching-costs literature, including Klemperer (1987a, b, 1995 [Sec 3.2]), Basu and Bell (1991), Padilla (1992), Basu (1993), Ahtiala (1998), Lal and Matutes (1994), Pereira (2000), Gehrig and Stenbacka (2002), Ellison (2003), and Lee and Png (2004). Of these models, Klemperer (1995; Section 3.2) is particularly easy to work with and to extend for other purposes.[18]

Although the switching cost strikingly affects price in each period, it does not affect the life-cycle price $c_1 + c_2$ that the consumer pays in the simple model of this subsection. As in the case of full commitment noted in section 2.1, we can here view the life-cycle (the bundle consisting of the period-1 good

[16][Note 2.2.1]. Because firms are symmetric and so charge the same price in period 2, the consumer is indifferent in period 1. If firms A, B had different costs c_{A2} and c_{B2} in period 2, then if A made the period-1 sale, its period-2 price would be $p_{A2} = c_{B2} + s$ (that is, constrained by B), while if B made the period-2 sale, its period-2 price would be $p_{B2} = c_{A2} + s$. In this case, the prices that firms charge in period 1 (and hence also firms' incentives to invest in cost reduction, etc.) depend on whether the consumer has rational expectations about the period-2 prices it will face or whether the consumer acts myopically. We discuss the role of expectations in Section 2.4.5. Other simple models such as that in Klemperer (1995, Section 3.2) sidestep the issue of consumers' expectations by assuming period-2 prices are constrained by consumers' reservation price r_2, hence independent of consumers' period-1 choice. The distinction between these modelling approaches is crucial in some analyses of network effects (see Section 3.7.3).

It is important for the modelling that the customer buys from just one firm in period 1. If a unit mass of consumers splits evenly betwen the firms in period 1, there may be no pure-strategy equilibrium in period 2. See note MSE.

[17]Skott and Jepsen (2000) argue that a tough drug policy may encourage the aggressive marketing of illegal drugs to new users, by increasing the costs of switching between dealers.

[18]For example, the many-period extension of this model is Beggs and Klemperer (1992).

and the period-2 good) as the real locus of competition, and competition in *that* product has worked exactly as one would hope. In particular, the absence of price commitment did not lead to any inefficiency in this very simple model.

2.3.2 Inefficiency of the Price-Path

Although the outcome above is socially efficient, the inability to contract in period 1 on period-2 prices in general leads to inefficiencies, even if firms still earn zero profits over the two periods. Even slight generalizations of the simple model above show this:

In particular, if the consumer has downward-sloping demand in each period and firms are restricted to linear pricing (i.e. no two-part pricing), or if firms face downward-sloping demands because there are many heterogeneous consumers with different reservation prices among whom they cannot distinguish, then there will be excessive sales in period 1 and too few sales in period 2 (Klemperer (1987a)).[19]

Our simple model also assumed that ex-post profits can feed through into better early deals for the consumers. In practice this may not be possible. For example, setting very low introductory prices may attract worthless customers who will not buy after the introductory period.[20] If for this or other reasons firms dissipate their future profits in unproductive activities (e.g., excessive advertising and marketing) rather than by offering first-period customers truly better deals, or if, for example, risk-aversion and liquidity concerns limit the extent to which firms charge low introductory-period prices to the consumers whom they will exploit later, then consumers are made worse off by switching costs, even if competition ensures that firms are no better off.

In our simple model firms make zero profits with or without switching

[19]Thus discussions of aftermarket power point out the possibility of sub-optimal trade-offs between aftermarket maintenance services, self-supplied repair, and replacement of machines. See Borenstein, MacKie-Mason, and Netz (2000), for instance.

[20]This is a particular problem if the introductory price would have to be negative to fully dissipate the ex-post rents. There may also be limits on firms' ability to price discriminate in favour of new customers without, for example, antagonizing their "regular" customers. See Section 2.4 for the case in which price-discrimination is infeasible.

costs. But switching costs and the higher ex-post prices and lower ex-ante prices that they create can either raise or lower oligopolists' profits. The reason is that, in cutting its first-period price, each firm sets its *marginal* first-period profit sacrifice equal to its marginal second-period gain, so the *total* first-period profit given up can be greater or less than the total second-period gain (see, especially, Klemperer (1987a,b)). However, the arguments we will review in Section 2.4 (which also apply to two-period models) suggest firms typically gain from switching costs.[21]

Finally note that while we (and the literature) primarily discuss firms exploiting locked-in consumers with high prices, consumers can equally be exploited with low qualities. And if it is hard to contract on future quality, contracting on price does not easily resolve the inefficiencies discussed above.[22]

2.4 Firms who Cannot Discriminate Between Cohorts of Consumers

In our first bargains-then-ripoffs model, we assumed that there was just one customer. It is easy to see that the basic lessons extend to the case where there are many customers but firms can charge different prices to "old" and "new" consumers, perhaps because "old" consumers are buying "follow on" goods such as spare parts. But when old consumers buy the same good as new consumers, it can be difficult for firms to distinguish between them. We now consider this case when a new generation of consumers arrives in the market in each of many periods.

2.4.1 Free-Entry Model

Even if firms cannot distinguish between cohorts of consumers, we may get the same pricing pattern if firms specialize sufficiently. In particular, in a simple model with free entry of identical firms and constant returns to scale, in each period some firm(s) will specialise in selling to new consumers

[21] See, especially, Klemperer (1987b). Ellison (2003) argues that firms gain from switching costs for a natural type of demand structure.

[22] Farrell and Shapiro (1989) show that price commitments may actually be worse than pointless. See note FS [Section 2.8].

while any firm with any old locked-in customers will sell only to those old customers.

If consumers have constant probability ϕ of surviving into each subsequent period, new-entrant firms with constant marginal costs c and discount factor δ offer price $c - \phi\delta s$ and sell to any new consumers, while established firms charge s more, i.e., charge $c + (1 - \phi\delta)s$ in every period.[23] That is, established firms charge the highest price such that no "old" consumers want to switch, and new entrants' expected discounted profits are zero. Thus the price paths consumers face are exactly as if firms could perfectly discriminate between them. In either case one can think of every (new and old) consumer as getting a "discount" of $\phi\delta s$ in each period reflecting the present value of the extent to which he can be exploited in the future, given his option of paying s to switch to an entrant; simultaneously, every "old" consumer is indeed exploited by s in every period. The outcome is socially efficient.

2.4.2 Do Oligopolists Hold Simultaneous Sales?, or Staggered Sales?, or No Sales?

Just as in the free-entry model, if there is a small number of firms who face no threat of entry and who cannot distinguish between cohorts of consumers, it is possible that in every period one firm might hold a "sale", setting a low price to attract new consumers, while the other(s) set a higher price to exploit their old consumers. Farrell and Shapiro (1988) explore such an equilibrium in a model that has just one new and one old consumer in each period. Since this assumption implies that in any period one firm has no customer base while the other already has half the market "locked-in", it is not surprising that this model predicts asynchronous sales. However, Padilla's (1995) many-customer model yields somewhat similar results: firms mix across prices but a firm with more locked-in customers has more incentive

[23] This assumes all consumers have reservation prices exceeding $c + (1 - \phi\delta)s$ for a single unit in each period, and that all consumers survive into the next period with the same probability, ϕ, so a consumer's value is independent of his age. If consumers live for exactly two periods the price paths in general depend on whether firms can directly distinguish between old and new consumers (as in the previous subsection) or cannot do this (as in this section).

to charge a high price to exploit them, and so sets high prices with greater probabilities than its rival.[24] These papers illustrate how switching costs can segment an otherwise undifferentiated products market as firms focus on their established customers and do not compete aggressively for their rivals' buyers, letting oligopolists extract positive profits.

More generally it is unclear whether oligopolists will hold sales simultaneously or will stagger them. On the one hand, it might make most sense to forgo short run profits to go after new customers when your rivals are not doing so. On the other hand, if switching costs are learning costs, then staggered sales cause switching and create a pool of highly mobile consumers who have no further switching costs, intensifying future competition (see Section 2.5). Klemperer (1983, 1989) and the extension of the latter model in Elzinga and Mills (1999) all have simultaneous sales.[25] [26]

Another possibility is that rather than holding occasional sales, each oligopolist in every period sets a single intermediate price that trades off its incentive to attract new consumers and its incentive to exploit its old customers. In a steady state, each firm's price will be the same in every period. Such an equilibrium could break down in several ways: if the flow

[24]Farrell and Shapiro assume firms set price sequentially in each period, but Padilla assumes firms set prices simultaneously.

[25]Elzinga and Mills' model fits with observed behaviour in the cigarette market. See also Elzinga and Mills (1998).

[26][Note MSE] In a single-period model in which all consumers have the same switching cost, s, the incentive to either undercut a rival's price by s or to overcut the rival's price by just less than s generally eliminates the possibility of pure-strategy equilibria if s is not too large: numerous papers (Baye et al (1992), Padilla (1992), Deneckere et al (1992), Fisher and Wilson (1995), Green and Scotchmer (1986), Rosenthal (1980), Shilony (1977), Varian (1980)), analyse single-period models of switching costs (or models that can be interpreted in this way) that yield mixed strategy equilibria, and Padilla (1995) finds mixed-strategy equilibria in a multi-period model. However, adding more real-world features to some of these models yields either asymmetric pure-strategy equilibria or symmetric pure-strategy Bayesian-Nash equilibria (if information is incomplete) rather than mixed-strategy equilibria.

Asymmetric pure-strategy equilibrium can be interpreted as asynchronous sales. Like Farrell and Shapiro (1988), Deneckere et al find that if firms can choose when to set their prices, the firm with fewer locked-in customers sets price second and holds a "sale".

Symmetric Bayesian equilibria correspond to "tradeoff pricing" of the kind discussed in the next paragraph of the text. Bulow and Klemperer (1998, Appendix B) give an example of this by incorporating incomplete information about firms' costs into a one-period model with switching costs that would otherwise yield mixed-strategy equilibria.

Gabrielson and Vagstad (2003) analyse when a pure-strategy equilibrium that looks like monopoly pricing exists in a single-period duopoly with heterogeneous switching costs.

of new consumers is too large, a firm would deviate by cutting price significantly to specialise in new consumers. If some consumers' switching costs and reservation prices are too large, a firm would deviate by raising price significantly to exploit old customers while giving up on new ones. And if firms' products are undifferentiated except by switching costs, a firm might deviate to undercut the other slightly and win all the new consumers.[27] But when none of these breakdowns occurs, there may be a stationary "no-sales" equilibrium: much of the literature examines such equilibria.[28]

Beggs and Klemperer (1992) explore a no-sales equilibrium in which in period t, firm i sets price

$$p_t^i = c^i + \alpha + \beta\sigma_{t-1}^i + \gamma(c^j - c^i) \tag{1}$$

where c^i is i's cost, σ_{t-1}^i is i's previous-period market share (i.e., the fraction of consumers i sold to in the previous period) and α, β, and γ are positive constants. These constants depend on four parameters: the discount factor, the market growth rate, the rate at which individual consumers leave the market, and the extent to which the firms' products are functionally differentiated; when firms are symmetric, the steady-state equilibrium price increases in the last of these four variables and decreases in the other three.[29]

2.4.3 Oligopoly Dynamics

We have seen that sometimes a lean and hungry firm with few locked-in customers holds a sale while its rivals with larger customer bases do not. Similarly, in no-sale models in which all firms sell to both old and new consumers, a firm with more old locked-in customers has a greater incentive

[27] However, if consumers have rational expectations about future prices, a small price cut may win only a fraction of new consumers: see section 2.4.5 below.

[28] Even if there are occasional "sales", firms will balance exploiting the old with attracting the new in "ordinary" periods, and this literature is relevant to these ordinary periods.

In the case of monopoly, both stationary "no-sales" models (see Holmes, 1990) and models in which periodic sales arise in equilibrium (see Gallini and Karp, 1989) can be constructed.

[29] [Note W] Klemperer (1995) discusses this model further: variants are in Chow, 1995, and To, 1995. Other important "no-sales" models are von Weizsäcker, 1984 and Wernerfelt (1991); Phelps and Winter's (1970) and Sutton's (1980) models of search costs, and Radner's (2003) model of "viscous demand", are related.

to exploit them, so will usually price higher and win fewer new unattached consumers. In both cases, the result is stable industry dynamics as more aggressive smaller firms catch up with larger ones.

In the equilibrium of Beggs' and Klemperer's (1992) no-sale duopoly model, described in (1) above, for example, $\beta > 0$, so larger firms charge higher prices, yielding stable dynamics. Indeed, it can be shown that $\sigma_t^i = \sigma^i + (\mu)^t(\sigma_0^i - \sigma^i)$ in which σ^i is firm i's steady-state market share and $0 < \mu << \frac{1}{2}$, so the duopoly converges rapidly and monotonically back to a stable steady state after any shock. Chen and Rosenthal (1996) likewise demonstrate a tendency for market shares to return to a given value, while in Taylor (2003) any initial asymmetries in market shares between otherwise symmetric firms may persist to some extent but are dampened over time.

However, the opposite is possible. If larger firms have lower marginal costs, and especially if economies of scale make it possible to drive smaller firms completely out of the market, then a larger firm may charge a lower price than its smaller rivals. In this case, any small advantage one firm obtains can be magnified and the positive-feedback dynamics can result in complete dominance by that firm. This is just as is typical with network effects (see Section 3.5.3)—indeed, switching costs create positive network effects in this case, because it is more attractive to buy from a firm that other consumers buy from (Beggs, 1989).

So switching-costs markets *can* "tip" like network-effects markets. But the simple models suggest a presumption that markets with switching costs are stable, with larger firms acting as less-aggressive "fat cats".[30]

2.4.4 The Level of Profits

A central question in policy and in the literature is whether switching costs raise or lower oligopoly profits. The simple two-period model of Section 2.3.1 suggests they do neither, but many non-theorist commentators, notably

[30]In the terminology introduced by Fudenberg and Tirole (1984). In the terminology introduced by Bulow, Geanakoplos, and Klemperer (1985), there is strategic complementarity between a firm's current price and its competitors' future prices. See also Farrell (1986a).

Porter (1980, 1985), believe switching costs raise profits.[31] As we discuss next, most models that are richer than the simple model tend to confirm the popular idea that switching costs raise profits.

If duopolists who cannot discriminate between old and new buyers hold asynchronous sales, they can earn positive profits in price competition even if their products are undifferentiated except by switching costs. The switching costs segment the market, and when one firm (generally the firm with the larger customer base) charges a high price to exploit its locked-in customers, the other firm then has market power even over new consumers because it can operate under the price umbrella of its fat-cat rival (see Farrell and Shapiro (1988) and Padilla (1995)). So in these models, a duopolist earns positive profits even in a period in which it starts with no locked-in customers. (However, if there were two identical new firms entering in every period, they would not generally earn any profits).

Furthermore, if switching costs are heterogenous, a similar effect means even duopolists who can (and do) discriminate between old and new customers can earn positive profits in price competition with products that are undifferentiated except by switching costs—see our discussion of Chen (1997) and Taylor (2003) in Section 2.5.1, below.

In addition, the symmetric stationary price of a "no-sales" equilibrium of the kind described in the previous section is also usually higher than if there were no switching costs. There are two reasons:

First, the "fat cat" effect applies here too, though in the indirect way discussed in the previous subsection; firms price less aggressively because they recognise that if they win fewer customers today, their rivals will be bigger and (in simple models with switching costs) less aggressive tomorrow.

Second, when consumers face switching costs, they care about expected future prices as well as current prices. Depending on how expectations of future prices react to current prices, this may make new customers (not yet locked into any firm), react either more or less elastically to price differences. However, as we now discuss, the presumption is that it makes their response

[31] We know of no convincing empirical evidence, but Cason and Friedman (2002) provide supportive laboratory evidence.

less elastic than absent switching costs, thus raising firms' prices and profits.

2.4.5 The Effect of Consumers' Expectations on Prices

How consumers' expectations about future prices depend on current prices critically affects competition and the price level—just as in other parts of the lock-in literature.[32] Consumers' expectations about their own future tastes also matter in a market with real (functional) product differentation; we assume consumers expect some positive correlation between their current and future tastes.

In a market without switching costs, a consumer compares differences between products' prices with differences between how well they match his current tastes. But with switching costs, he recognises that whichever product he buys today he will, very likely buy again tomorrow. So switching costs make him more willing to change brands in response to a price cut if, roughly speaking, he expects that price cut to be more permanent than his tastes; they will lower his willingness to change in reponse to a price cut if he expects the price cut to be less permanent than his tastes.

(i) *Consumers who Assume any Price Cut below their Expected Price will be Maintained in the Future*

If consumers expect a firm that cuts price today to maintain that price cut forever then, relative to the case of no switching costs, they are more influenced by such a price cut than by their current (impermanent) product preferences.[33] (In the limit with infinite switching costs, a consumer's prod-

[32] Consumers' expectations about how future prices depend on costs are, of course, also important in determining whether firms have the correct incentives to invest in future cost reduction. This issue does not seem to have been directly addressed by the switching-costs literature, but we discuss in Section 3.7 how a network-effects model can be reinterpreted to address it. See also [note 2.3.1].

[33] A related model with these expectations is Borenstein, Mackie-Mason and Netz (2000). In their model, consumers buy a differentiated durable good ("equipment") from one of two firms and must then buy an aftermarket product ("service") in the next period from the same firm. High service prices generate profits from locked-in customers, but deter new customers from buying equipment because they expect high service prices in the following period. So the stationary equilibrium service price lies between marginal cost and the monopoly price, even if firms' products are undifferentiated except by switching costs.

uct choice is forever, so unless his preferences are also permanent, products are in effect less differentiated.) So switching costs then *lower* equilibrium prices; see von Weizsäcker's (1984) model in which each firm chooses a single price (and quality) to which it is committed forever, but in which consumers are uncertain about their future tastes.[34]

We will see below (see Section 3.7) that a similar effect arises when there are strong proprietary network effects and differentiated products. Then, consumers' desire to be compatible with others overwhelms their differences in tastes and drives firms whose networks are incompatible towards undifferentiated Bertrand competition. Here, with switching costs, each consumer's desire to be compatible with his future self (who in expectation has tastes closer to the average) likewise reduces effective differentiation and drives the firms towards undifferentiated Bertrand competition.

(ii) *Consumers whose Expectations about Future Prices are Unaffected by Current Prices*

If consumers expect that a firm that unexpectedly cuts price this period will return to setting the expected price next period, then price changes are less permanent than, and so influence consumers less than, taste differences. So switching costs raise price levels. Each consumer is making a product choice that his future selves must live with, and his future selves' preferences (while possibly different from his own) are likely to be closer to his currently-preferred product than to other products. Consumers are therefore less attracted by a current price cut than absent switching costs.

(iii) *Consumers with Rational Expectations*

If consumers have fully rational expectations they will recognise that a lower price today generally presages a higher price tomorrow. As we discussed above, a firm that wins more new consumers today will be a "fatter cat" with relatively greater incentive to price high tomorrow; and we ex-

[34]The effect we discussed in the previous subsection, 2.4.4—that firms moderate price competition in order to fatten and so soften their opponents—is also eliminated by von Weizsäcker's commitment assumption.

pect that this will typically be the main effect, although other effects are possible.[35] So consumers with rational expectations will be even less sensitive than in (ii) to price cutting, and switching costs thus raise prices.[36]

In summary, while there is no unambiguous conclusion, under either economists' standard rational-expectations assumption ((iii)), or a more myopic assumption ((ii)), switching costs raise prices overall. Only if consumers believe unanticipated price changes are more permanent than their product preferences do switching costs lower prices. For these reasons, Beggs and Klemperer (1992) argue that switching costs tend to raise prices when new and old customers are charged a common price. There is therefore also a more general presumption that switching costs usually raise oligopolists' total profits.

2.4.6 Collusive Behavior

Like most of the literature, the discussion above assumes noncooperative behavior by firms, without strategic threats of punishment if others compete too hard.[37]

One should also ask whether switching costs hinder or facilitate collusion,

[35]See e.g., Beggs and Klemperer (1992), Klemperer (1987a,b,c), Padilla (1992, 1995). As discussed above, the fat cat effect can be reversed if e.g. economics of scale or network effects are strong enough. (Doganoglu and Grzybowski (2004) show how appending network benefits to Klemperer's (1987b) model lowers prices.) Another caveat is that with incomplete information about firms' costs a lower price might signal lower costs, so consumers might rationally expect a lower price today to presage a lower price tomorrow. But if there is incomplete information about costs, firms might price high in order to signal high costs and thus soften future competition. (A search-costs model that is suggestive about how firm-specific cost shocks might affect pricing in a switching-costs model is Fishman and Rob (1995).) Furthermore, if firms differ in the extent that they can or wish to exploit locked-in customers, consumers will expect that a lower price today means a higher price tomorrow, which will also be a force for higher prices.

[36]Holmes (1990) analyses price-setting by a monopolist facing overlapping generations of consumers who must sink set-up costs before using the monopolist's good. He finds that if consumers have rational expectations, then prices are higher than those that would prevail if the firm could commit to future prices. The reason is similar: rational consumers are insensitive to price cuts because they understand that a low price today will encourage other consumers to sink more costs which in turn results in higher future prices.

[37]For example, Beggs and Klemperer assume each firm's price depends only on its current market share and not otherwise on history, and rule out the kind of strategies described by, for example, Abreu (1988) or Green and Porter (1984) that support collusive outcomes in contexts without switching costs.

in which high prices are supported by firms punishing any other firm thought to have deviated. While many people's intuition is that switching costs support collusion, this remains unclear as a theoretical matter:

Switching costs make deviating from a collusive agreement less profitable in the short run, because it is harder to quickly "steal" another firm's customers. But, for the same reason, switching costs make it more costly to punish a deviating firm. So it is not obvious whether collusion is easier or harder on balance, and in Padilla's (1995) model, which incorporates both these effects, switching costs actually make collusion more difficult.

Switching costs may also make it easier for firms to monitor collusion, because the large price changes necessary to win away a rival's locked-in customers may be easy to observe. And switching costs may additionally facilitate tacit collusion by providing "focal points" for market division, breaking a market into well-defined submarkets of customers who have bought from different firms. However, while these arguments are discussed in Stigler (1964) and Klemperer (1987a), they have not yet been well explored in the literature, and do not seem easy to formalise satisfactorily. Furthermore, if collusion is only easier after most customers are already locked-in, this is likely to induce fiercer competition prior to lock-in, as in the simple bargain-then-ripoff model.

2.5 Consumers Who Use Multiple Suppliers

In the models above, as in most leading models of switching costs, switching costs affect prices but there is no switching in equilibrium. In reality a consumer may actually switch, and use different suppliers in different periods, either because firms' products are differentiated and his tastes change, or because firms' relative prices to him change over time, as they will, in particular, when each firm charges new customers less than existing customers.

Furthermore, although we assumed above that each consumer buys one unit from one firm in each period, a consumer who values variety may buy multiple products even in a single period. Consumers may therefore use multiple suppliers in a period or, as we will discuss, each firm may produce a range of products.

2.5.1 Paying Consumers to Switch

Most of the switching costs literature assumes a firm offers the same price to all consumers in any given period. However, as the bargains-then-rip-offs theme stresses, firms would often like to price discriminate between their old locked-in customers, unattached (new) customers, and customers locked-in to a rival. And firms often do pay consumers to switch to them from rivals. For example, long-distance phone carriers make one-time payments to customers switching from a rival; credit card issuers offer lower interest rates for balance transfers from another provider; and economics departments pay higher salaries to faculty members moving from other departments. How does the possibility of such discrimination affect pricing?

Chen (1997) analyses a two-period, two-firm, model in which each firm can charge one price to its old customers and another to other consumers in the same period. In effect, second-priced consumers are in two separate markets according to which firm they bought from in the first period. Each of these "markets" is like the second period of our core (Section 2.3.1) two-period model. In that model all consumers had the same switching costs, s, so the period-2 incumbent charged a price just low enough to forestall actual switching.[38] But in Chen's model, old consumers have heterogenous switching costs (and firms cannot discriminate between them, perhaps because they cannot observe individual consumers' switching costs), so firms charge higher prices than their rivals to their old consumers but consumers with low switching costs switch firms. In Chen's model firms' profits are lower than if they could not discriminate between old and new customers, and consumers might also be worse off, because of the costs of actually switching. Firms'

[38] Likewise, the simple model of section 2.4.1 shows that if firms can price discriminate, the price will be $c + (1 - \phi\delta)s$ to all old consumers, and will be s lower to new consumers, but no consumers will ever actually switch. Similarly, Nilssen (1992) observes that if each firm can charge a different price to each consumer, there will be no actual switching. Nilssen showed that transactional switching costs give consumers less incentives to switch than do learning switching costs. Thus transactional costs lead to lower prices for new consumers, higher prices for loyal consumers, and so also a bigger within-period quantity distortion if there is downward-sloping demand in each period. (Gabrielsen and Vagstad (2002) note that two-part tariffs can in theory avoid this distortion.)

total discounted profits are nevertheless higher than absent switching costs because (as in Section 2.4.4) the switching costs segment the market, so firms have some market power even over customers who are new to them in the second period.[39],[40]

In Chen's two-firm model, consumers who leave their current supplier have only one firm to switch to, so this other firm can make positive profits even on new customers, and the duopolists earn positive profits in equilibrium. But with three or more firms, there are always at least two firms vying for any consumer willing to leave his current supplier and, if products are undifferentiated, these firms will bid away their expected lifetime profits from serving those consumers in their competition to attract them. So, as Taylor (2003) shows, with three or more firms, firms earn positive rents only on their current customers, and these rents are competed away ex ante, as in our core model.

These models of "paying customers to switch" suggest repeat buyers pay higher rather than lower prices. While this is often observed, we also often observe the opposite pattern in which customers are rewarded for loyalty. Taylor's model provides one possible explanation. He shows that if switching costs are transactional, consumers may move between suppliers to signal that they have low switching costs and so improve their terms of trade. Because this switching is socially costly, equilibrium contracts may discourage it through "loyal customer" pricing policies that give better terms to loyal customers than to those who patronised other firms in the past. But Taylor nevertheless finds that firms charge the lowest prices to new customers.

Shaffer and Zhang (2000) study a single-period model that is similar to the second period of Chen's model but in which the distributions of switching

[39]Because in this model a firm's old and new customers are effectively in unconnected markets, both of the firm's prices are independent of its previous-period market share, by contrast with the no-price-discrimination models discussed above. This feature allows Taylor (2003) to extend Chen's model to many periods and many firms, but Arbatskaya (2000) shows that the "independence" result does not persist if there is functional product differentiation as well as switching costs.

[40]Gehrig and Stenbacka (2004a) develop a model in which the last two periods are similar to Chen's model, and in which profits are increasing in the size of switching costs; in Gehrig and Stenbacka's *three*-period model firms therefore (non-cooperatively) make product choices that maximize the switching costs between them. See also Gehrig and Stenbacka (2004b)

costs from the two firms are different. If firm A's customers have lower and more uniform switching costs than firm B's, then A's loyal-customer demand is more elastic than its new-customer demand, so it may charge a lower price to its loyal customers than to customers switching from B. But this rationale is asymmetric, and this model never results in both firms charging lower prices to loyal customers than to switching customers.[41]

There are also models of contractual switching costs that result in lower effective prices to repeat customers than to new customers, and contracts that favour repeat customers arise endogeously in some of these models (see Section 2.8.3). But the literature has found it hard to explain how real switching costs might generate discrimination in favor of old customers.

2.5.2 Is There Too Much Switching?

Consumers decide whether or when to switch, and pay the switching costs. So there will generally be the wrong amount of switching if (i) firms' relative prices to a consumer fail to reflect their relative marginal costs[42], or (ii) consumers switch (or not) in order to affect firms' future prices, or (iii) consumers' switching costs are not real social costs. Most simple models recognize no efficiency role for switching, so any switching in such models is inefficient.

(i) *price differences don't reflect cost differences*

The bargains-then-ripoffs theme predicts that, when they can do so, firms charge lower prices to their new consumers. As a result, a given consumer will face different prices from different firms that do not reflect any cost differences

[41] Lee (1997) also studies a one-period switching-cost model similar to the second period of Chen's model. Fudenberg and Tirole (2000) explore a two-period model with some similar features to Chen's, in which firms price discriminate between consumers based on their past demands, but with real functional product differentiation between firms and without real (socially costly) switching costs; they too find that loyal customers are charged higher prices than switchers. However, they also show that firms may wish to offer long-term contracts that offer consumers a high period-one price in return for a guaranteed low period-two price (see Section 2.8.3). (Villas-Boas (1999) analyses a many-period model similar to Fudenberg and Tirole's but does not consider long-term contracts.) Acquisti and Varian (2002) present a related two-period monopoly model which can be interpreted as being of consumers with switching costs.

[42] Consumers must also have rational expectations about future price differences, etc.

between firms. This is true even when all firms symmetrically charge high prices to old customers and lower prices to new customers. Although some simple models such as our core (Section 2.3.1) model predict no switching, in general inefficient switching results.[43]

When firms do not price discriminate between new and old consumers, the same result applies for a slightly different reason. As we saw in section 2.4, a firm with a larger customer base will then charge a larger markup over its marginal cost. So if consumers have differing switching costs, such a firm's price exploits its old high switching-cost customers and induces its low switching-cost consumers to switch to a smaller firm or entrant. Thus Gabszewicz, Pepall and Thisse (1992), Farrell and Shapiro (1988), and Wang and Wen (1998) also predict excessive switching to smaller firms and entrants.

(ii) *consumers switch in order to affect prices*

If a consumer is a large fraction of the market, or if firms can discriminate between consumers (so each consumer is, in effect, a separate market), a consumer may switch to affect future prices.

If switching costs are learning costs, switching strengthens a consumer's outside option, so he may switch in order to strengthen his bargaining position—by switching he effectively creates a second source of supply and thereby increases the competition to supply him in the future (Lewis and Yildirim, 2003). And even if switching costs are transactional (and firms are imperfectly informed about their magnitude), we saw in section 2.5.1 that consumers may switch to signal that their switching costs are low and so improve their terms of trade.

Strategic consumers may also commit to ignore switching costs (or acting as if their switching costs were lower than they truly are) in their future

[43][note k88] Even if all consumers have the same switching cost, if an entrant's production cost plus that switching cost exceeds the incumbent's production cost, then in a *quantity*-competition model the entrant will sell to some of them, thus inducing inefficient switching (Klemperer (1988)). This result is just the standard oligopoly result that a higher-cost firm wins a socially excessive market share (though at a smaller markup).

A caveat is that these excessive-switching results take the number of firms as given. If the switching costs mean there is too little entry from the social viewpoint (see Section 2.7.2) then there may for this reason be too little switching.

purchase decisions, in order to force the incumbent supplier to price more competitively (Cabral and Greenstein (1990));[44] this strategy will generally increase the amount of switching.

In all these cases, socially costly switching in order to affect prices is inefficient to the extent that it merely shifts rents from firms to the customer who switches. On the other hand, if firms cannot discriminate between consumers, such switching usually lowers prices and so improves the efficiency of other consumers' trades with sellers, so there may then be less switching than is socially desirable.

(iii) *switching costs are not real social costs*

If switching costs are contractual, and not social costs, consumers will *ceteris paribus* switch less than is efficient. But if real (social) switching costs exist, then contractual switching costs may prevent socially inefficient switches of the types discussed above.[45]

2.5.3 Multiproduct Firms

A consumer who buys several products in a single period may incur additional "shopping costs" for each additional supplier used. These shopping costs may be the same as the switching costs incurred by consumers who change suppliers between periods. However, the dynamic issues that switching-cost models usually emphasise no longer arise. In particular, firms and consumers can contract on all prices, so the analogy with economies of scope in production is particularly strong.[46] Thus shopping costs provide an efficiency reason for multiproduct firms just as economies of scope in production do.[47]

[44]The literature has largely assumed that consumers have no commitment power (see Section 2.8 for exceptions).

[45]In Fudenberg and Tirole (2000) firms endogenously offer long term contracts that create contractual switching costs that reduce inefficient switching to less preferred products and increase social welfare, conditional on firms being permitted to price discriminate between old and new customers.

[46]But some superficially single-period contexts are better understood as dynamic. For instance, supermarkets advertise just a few "loss leaders"; unadvertised prices are chosen to be attractive once the consumer is in the shop ("locked in") but might not have drawn him in. (See note LM $\approx$ 5.)

[47]Dranove and White (1996) models hospitals as multi-product providers with switching costs between providers. Several studies document that travelers stongly prefer to

The analogy is not perfect, because switching costs and shopping costs are based on specific consumer-firm matches, whereas the production-side economies of scope emphasised by Panzar and Willig (1981) and others depend only on a firm's total sales of each product and not on whether the same consumers buy the firm's different products or whether some consumers use multiple suppliers.[48]

However, the analogy is particularly good if firms' product lines are sufficiently broad that most consumers use just one supplier. For example, Klemperer and Padilla (1997) demonstrate that selling an additional product can provide strategic benefits for a firm in the markets for its current products if consumers have shopping costs of using additional suppliers (because selling an extra variety can attract demand away from rival suppliers for this firm's existing varieties). This parallels Bulow et al (1985)'s demonstration of the same result if consumers' shopping costs are replaced by production-side economies of scope (because selling an additional variety lowers the firm's marginal costs of its existing products). In both cases each firm, and therefore the market, may therefore provide too many different products. More obviously, mergers can be explained either by consumer switching costs (Klemperer and Padilla (1997)) or by production economies of scope.

Some results about *single*-product competition over *many* periods with switching costs carry over to *multi*-product competition in a *single* period with shopping costs. For example, we suggested earlier in this section that when switching costs are learning costs, oligopolists might benefit by synchronizing their sales to minimize switching and so reduce the pool of highly price-sensitive (no-switching cost) customers. Likewise multiproduct firms competing in a single period may have a joint incentive to minimize the number of consumers who buy from more than one firm. Indeed Klemperer (1992, 1995 ex.4) shows that firms may inefficiently offer similar products to

use a single airline for a multi-segment trip, and the importance of these demand-side complementarities in air travel (e.g. Carlton, Landes, and Posner (1980)).

[48] As we noted in Section 2.1, if firms can discriminate between consumers, then each consumer becomes an independent market which, in the presence of switching costs, is closely analogous to a market with production economies of scope.

each other, or similar product lines to each other, for this reason. Taken together with the previous paragraph's result, this suggests that each firm may produce too many products, but that there may nevertheless be too little variety produced by the industry as a whole.

An important set of shopping-cost models are the "mix-and-match" models pioneered by Matutes and Regibeau (1988), Economides (1989) and Einhorn (1992). Most of this literature takes each firm's product-line as given, and asks whether firms prefer to be compatible (no shopping costs) or incompatible (effectively infinite shopping costs); see Sections 2.7.3 and 2.8.

Similarly, when firms "bundle" products (see, e.g., Whinston (1990), Nalebuff (1999, 2000)) they are creating contractual shopping costs between their products; we discuss bundling briefly in Sections 2.7.3 and 2.8.[49]

"Shopping costs" models are distinguished from other "switching costs" models in that consumers can observe and contract on all prices at the same time in the "shopping costs" models. We will henceforth use the term switching costs to cover all these costs, but continue to focus mainly on dynamic switching costs.

2.6 Battles for Market Share

2.6.1 The Value of Market Share

We have seen that with switching costs (or indeed proprietary network effects—see Section 3.7), a firm's current customer base is an important determinant of its future profits.

We can therefore write a firm's current-period value function, (i.e., total discounted future profits), V_t, as the sum of its current profits, π_t, and its discounted next-period value function $\delta V_{t+1}(\sigma_t)$, in which δ is the discount factor and the next-period value function, $V_{t+1}(\cdot)$, is a function of the size of its current-period customer base, σ_t.

$$V_t = \pi_t + \delta V_{t+1}(\sigma_t) \tag{2}$$

For example, in our core model with free entry, $V_{t+1} = s\sigma_t$, and Biglaiser,

[49] Varian's (1989) and Stole's (forthcoming) surveys describe models of quantity discounts and bundling in Volume 1 and the current volume of this Series, respectively.

Crémer and Dobos (2003) have explored various cases in which this simple formula holds. More generally, however, even (2) is a simplification. In general, the firm's future profits depend on its customers' types and their full histories, how market share is distributed among competing firms, how many consumers in the market make no purchase, etc. However, V_{t+1} depends only on current-period market share in models such as Klemperer (1987b, 1995), Farrell and Shapiro (1988), Beggs and Klemperer (1992), Padilla (1992, 1995), and Chen and Rosenthal (1996), which all model just two firms and a fixed set of consumers whose reservation prices are high enough that they always purchase. (For example, equation (1) shows for Beggs and Klemperer's model how prices, and therefore also quantities, and hence value functions, in a period depend on the firm's previous-period market share.) So σ_t is often interpreted as "market share", and this explains firms' very strong concern with market shares in markets with switching costs and/or (we shall see) network effects.[50]

2.6.2 Penetration Pricing

The firm's first-order condition for the optimal choice of a period-t price is then

$$0 = \frac{\partial V_t}{\partial p_t} = \frac{\partial \pi_t}{\partial p_t} + \delta \frac{\partial V_{t+1}}{\partial \sigma_t} \frac{\partial \sigma_t}{\partial p_t} \tag{3}$$

Provided that the firm's value function is increasing in its market share, $\partial V_{t+1}/\partial \sigma_t > 0$,[51] therefore, the firm charges a lower price or sets a higher quantity[52] than would maximize short-run profits, in order to raise its cus-

[50]Because switching costs make current market share such an important determinant of a manufacturer's future profits, Valletti (2000) suggests they may provide a motive for vertical integration with retailers to ensure sufficient investment in a base of repeat subscribers.

[51]This seems the usual case, although in principle, stealing customers from rival(s) may make the rival(s) so much more aggressive that the firm is worse off. See Banerjee and Summers (1987), Klemperer (1987c).

In Beggs and Klemperer (1992), V_{t+1} is quadratic in σ_t. (The fact that the sum of the duopolists' value functions is therefore maximised at the boundaries is consistent with stable dynamics because lowering current price is less costly in current profits for the firm with the smaller market share. See Budd et al (1993).)

[52]We can perform a similar analysis with similar results for a quantity-setting firm. The analysis is also unaffected by whether each firm sets a single price to all consumers or

tomer base and hence its future profits. That is, $\partial\pi_t/\partial p_t > 0$ (since we assume $\partial\sigma_t/\partial p_t < 0$).

In the early stages of a market, therefore, when few consumers are locked in, so even short-run profit-maximising prices are not high relative to costs, equation (3) implies low penetration pricing, just as in the core two-period model.[53] [54] Equation (3) also suggests that the larger the value of the future market, V_{t+1}, the deeper the penetration pricing will be. For example, a more rapidly growing market will have lower prices.[55]

2.6.3 Harvesting vs Investing: Macroeconomic and International Trade Applications

As equations (2) and (3) illustrate, the firm must balance the incentive to charge high prices to "harvest" greater current profits ((3) showed π_t is increasing in p_t) against the incentive for low prices that "invest" in market share and hence increase future profits (V_{t+1} is increasing in σ_t, which is decreasing in p_t).

Anything that increases the marginal value of market share will make the firm lower price further to invest more in market share. Thus, for example, a lower δ, that is, a higher real interest rate, reduces the present value of future market share (see (2)) so leads to higher current prices (see (3): lower δ implies lower $\partial\pi_t/\partial p_t$ implies higher p_t[56]).

Chevalier and Scharfstein (1996) develop this logic in a switching-cost

whether, as in section 2.5, each firm sets different prices to different groups of consumers in any period.

[53] It is unclear whether we should expect "penetration pricing" patterns from a monopolist, since $\partial V_{t+1}/\partial\sigma_t$ may be smaller in monopoly—where consumers have nowhere else to go—than in oligopoly, and (if goods are durable) durable-goods effects imply falling prices in monopoly absent switching-cost effects (equation (3) only implies that early period prices are lower than in the absence of switching-costs, not that prices necessarily rise). Cabral et al (1999) show it is hard to obtain penetration pricing in a network-effects monopoly model (see Section 3.6).

[54] Of course, as noted in Section 2.3.2, in a more general model the "penetration" might be through advertising or other marketing activities rather than just low prices.

[55] Strictly, (3) tells us prices are lower if $\partial V_{t+1}/\partial\sigma_t$ is larger, but this is often true for a more rapidly growing market. See, for example, Beggs and Klemperer (1992), Borenstein, Mackie-Mason and Netz (2000) and also Holmes' (1990) steady-state model of a monopolist selling a single product to overlapping generations of consumers who incur set-up costs before buying the product.

[56] See Klemperer (1995). We assume stable, symmetric, oligopoly and that the dominant effect of lowering δ is the direct effect.

model based on Klemperer (1995). They argue that liquidity-constrained firms perceive very high real interest rates and therefore set high prices, sacrificing future profits in order to raise cash in the short term. They provide evidence that the most financially-constrained supermarket chains indeed raise their prices relative to other chains during recessions, and Campello and Fluck (2004)'s subsequent empirical work shows that these effects are larger in industries where consumers face higher switching costs.[57]

Fitoussi and Phelps (1988) use a similar logic (emphasising search costs rather than switching costs) to argue that high interest rates contributed to the high rates of inflation in Europe in the early 1980s.

Froot and Klemperer (1989) also apply the same logic to international trade in a model of competition for market share motivated by switching costs and network effects. A current appreciation of the domestic currency lowers a foreign firm's costs (expressed in domestic currency) so tends to lower prices, but if the appreciation is expected to be only temporary then the fact that the domestic currency will be worth less tomorrow is equivalent to an increase in the real interest rates which raises prices. So exchange-rate changes that are expected to be temporary may have very little impact on import prices. But if the currency is anticipated to appreciate in the future, both the "cost effect" and "interest-rate effect" are in the same direction—market share tomorrow is probably worth more if future costs are lower, and tomorrow's profits are worth more than today's profits, so for both reasons today is a good time to invest in market share rather than harvest current profits. So import prices may be very sensitive to anticipated exchange-rate changes. Froot and Klemperer (1989) and Sapir and Sekkat (1995) provide empirical support for these theories.[58]

[57] See also Campello (2003). Beggs and Klemperer (1989, Section 5.3) and Klemperer (1995) provide further discussion of how "booms" and "busts" affect the trade-offs embodied in equation (3) and hence affect price-cost margins.

[58] For other applications of switching-costs theory to international trade, see Tivig (1996) who develops "J-curves" (since sales quantities respond only slowly to price changes if there are switching costs), Gottfries (2002), To (1994), and Hartigan (1995).

2.7 Entry

Switching costs may have important effects on entry: with real, exogenous switching costs, small-scale entry to win new, unattached, consumers is often easy and indeed often too easy, but attracting even some of the old "locked-in" customers may not just be hard, but also be too hard from the social standpoint.

Furthermore, these results take the switching costs as given. Firms may also create unnecessary switching costs in order to discourage entry.

2.7.1 Small-Scale Entry is (Too) Easy

We saw in Section 2.4 that if firms cannot discriminate between old and new consumers, then the "fat cat" effect may make small scale entry very easy: incumbent firms' desire to extract profits from their old customers creates a price umbrella under which entrants can profitably win new unattached (or low switching cost) customers. And even after entry has occurred, the erstwhile incumbent(s) will continue to charge higher prices than the entrant, and lose market share to the entrant, so long as they remain "fatter" firms with more old consumers to exploit.

So if there are no economies of scale, even an entrant that is somewhat less efficient than the incumbent(s) can enter successfully at a small scale that attracts only unattached buyers.[59] (See Klemperer (1987c), Farrell and Shapiro (1988), Gabszewicz, Pepall and Thisse (1992), Wang and Wen (1998), etc.)

Of course, the flip-side of this is that the same switching costs that en-

[59] [NAPP] This result depends on there being (sufficient) new customers in each period (which is a natural assumption). For an analogous result that entry was easy into just one product in a shopping-cost market, there would have to be sufficient buyers without shopping costs, or who wished to purchase just that product (this may be a less natural assumption). Failing that, "small scale" entry in a shopping cost market is not easy.

Our assumption of no discrimination between old and new consumers means the easy-entry result also does not apply to aftermarkets. Entry may be hard in this case if first-period prices cannot fall too low, and the incumbent has a reputational or similar advantage. For example, the UK Office of Fair Trading found in 2001 that new entry was very hard into the hospital segment of the market served by NAPP Pharmaceutical Holdings Ltd where prices were less than one-tenth of those in the "follow-on" community market.

courage new entry also encourage the new entrants to remain at a relatively small scale unless there are many unattached buyers.

2.7.2 Large Scale Entry is (Too) Hard

While the fat-cat effect gives new entrants an advantage in competing for new customers, it is very hard for them to compete for customers who are already attached to an incumbent. There is also adverse selection: consumers who switch are likely to be less loyal, hence less valuable, ones.[60] So entry may be hard if small-scale entry is impractical, due perhaps to economies of scale, or to network effects. Furthermore, even new consumers may be wary of buying from a new supplier if they know that it can only survive at a large scale, since with switching costs consumers care about the future prospects of the firms they deal with.

Of course, this does not imply that there is *too* little large-scale entry. If switching costs are social costs, then large-scale entry may not be efficient even if the entrant's production costs are modestly lower than an incumbent's. That is, to some extent these obstacles to profitable large-scale entry reflect social costs of such entry.

However, this reflection is imperfect. If the entrant cannot discriminate between consumers, then large-scale entry requires charging all consumers a price equal to the incumbent's price less the marginal old buyer's switching cost. But socially the switching cost applies only to the old switching buyers, not to the new consumers, and only applies to switching buyers at the average level of their switching cost, not at the marginal switching cost. So efficient large-scale entry may be blocked.

Furthermore, entry can sometimes be strategically blockaded. In par-

[60] Some work on the credit card market emphasises this adverse-selection problem: creditworthy borrowers may have been granted high credit limits by their current card issuers so have higher switching costs. Furthermore, low-default risk customers may be less willing to switch (or even search) because they do not intend to borrow—but they often do borrow nevertheless (Ausubel, 1991). Calem and Mester (1995) provide empirical evidence that this adverse selection is important, Ausubel provides evidence that the US bank credit card issuing market earns positive economic profit and attributes this, at least in part, to switching costs or search costs, and Stango (2002) also argues that switching costs are an important influence on pricing.

ticular, an incumbent may "limit price", that is, cut price to lock in more customers and make entry unprofitable at the necessary scale, when entry at the same scale would have been profitable, and perhaps efficient, if the additional customers had not been "locked-up" prior to entry (see Klemperer (1987c)).

Of course, entry can be too easy or too hard for more standard reasons. Entry can be too hard if it expands market output, and consumers rather than the entrant capture the surplus generated. And entry is too easy if its main effect is to shift profits from the incumbent to the entrant.[61] But these caveats apply whether or not there are switching costs; the arguments specific to switching costs suggest that entry that depends for its success on consumers switching is not just hard, but too hard.

2.7.3 Single-Product Entry May Be (Too) Hard

If switching costs (or shopping costs) "tie" sales together so consumers prefer not to patronise more than one firm (see section 2.5.3), then an entrant may be forced to offer a full range of products to attract new customers (let alone any old consumers). If offering a full range is impractical, entry can effectively be foreclosed. Thus in Whinston (1990), Nalebuff (1999), and Klemperer and Padilla (1997), tying can foreclose firms that can only sell single products. In Whinston and Nalebuff the "switching costs" are contractual, while in Klemperer and Padilla the products are "tied" by real shopping costs.[62] If the switching/shopping costs are real, entry need not be too hard *given* the switching costs, but the arguments of the previous subsection suggest it often may be.

[61] Klemperer (1988) illustrates the latter case, showing that new entry into a mature market with switching costs can sometimes be socially undesirable. The point is that just as entry of a firm whose costs exceed the incumbent's is often inefficient in a standard Cournot model without switching costs (Bulow et al, 1985, section VI E, Mankiw and Whinston, 1986) so entry of a firm whose production cost *plus* consumers' switching cost exceeds the incumbent's production cost is often inefficient in a quantity-setting model with switching costs (see note K88).

[62] Choi (1996a) shows that tying in markets where R&D is critical can allow a firm with an R&D lead in just one market to pre-empt both. The welfare effects are ambiguous.

2.7.4 Artificial Switching Costs Make Entry (Too) Hard

The previous discussion addressed whether entry is too easy or too hard, taking the switching costs as given: we observed that switching costs make certain kinds of entry hard, but that this is at least in part because they also make entry socially costly, so entry may not be very much *too* hard given the switching costs. A larger issue is whether the switching costs are inevitable real social costs. They may instead be contractual,[63] or may be real but caused by an unnecessary technological choice that an entrant cannot copy. In these cases, it is the incumbent's ability to choose incompatibility that is the crucial entry barrier.

2.8 Endogenous Switching Costs: Choosing How to Compete

Market participants may seek to either raise or to lower switching costs in order to reduce inefficiencies (including the switching cost itself), to enhance market power, to deter new entry, or to extract returns from a new entrant:

2.8.1 Reducing Switching Costs to Enhance Efficiency

As we have seen, a firm that cannot commit not to exploit its ex-post monopoly power must charge a lower introductory price. If the price-path (or quality-path) is very inefficient for the firm and consumers jointly, the firm's surplus as well as joint surplus may be increased by nullifying the switching costs. Thus, for example, a company may license a second source to create a future competitor to which consumers can costlessly switch (Farrell and Gallini (1988)).[64]

Likewise, firms producing differentiated products (or product lines) may deliberately make them compatible (i.e., choose zero switching costs). This increases the variety of options available to consumers who can then "mix-and-match" products from more than one firm without paying a switching

[63]This includes those created by "loyalty contracts", "exclusive contracts" and "bundling" or "tying" etc.

[64]In Gilbert and Klemperer (2000) a firm commits to low prices that will result in rationing but will not fully exploit the consumers ex-post, to induce them to pay the start-up costs of switching to the firm.

cost. So eliminating switching costs can raise all firms' demands, and hence all firms' profits.[65]

Where suppliers are unwilling to reduce switching costs (see below), third parties may supply converters,[66] or regulators may intervene.

We have also already noted that customers may incur the switching (or start-up) cost of using more than one supplier, or may pre-commit to ignoring the switching costs in deciding whether to switch,[67] in order to force suppliers to behave more competitively.[68]

Finally, firms may be able to mitigate the inefficiencies of distorted prices and/or qualities by developing reputations for behaving as if there were no switching costs.[69]

2.8.2 Increasing Switching Costs to Enhance Efficiency

Firms may also mitigate the inefficiencies of distorted prices and qualities by contracting, or even vertically integrating, with their customers.[70] [71] Likewise Taylor (2003) finds firms might set lower prices to loyal consumers to reduce inefficient switching. Of course, a downside of these strategies of increasing switching costs is that they also limit the variety available to consumers unless they pay the switching costs.

[65] See Matutes and Regibeau (1988), Economides (1989), Garcia Mariñoso (2001), Stahl (1982), etc. But the mix-and-match models reveal other effects too; see Section 2.8.4.

[66] See Section 3.8.3 for more on converters.

[67] See Cabral and Greenstein (1990).

[68] Greenstein (1993) discusses the procurement strategies used by U.S. federal agencies in the late 1970s to force suppliers of mainframe computers to make their systems compatible with those of their rivals.

[69] See Eber (1999). Perhaps more plausibly firms may develop reputations for, or otherwise commit to, treating old and new customers alike (since this behaviour is easy for consumers to understand and monitor); this behavior may also mitigate the inefficiencies due to the distorted prices (though see note FS [˜2 below]) – it is most likely to be profitable if bargain-then-ripoff pricing is particularly inefficient.

[70] See Williamson (1975) and Klein, Crawford, and Alchian (1978).

[71] [Note FS] However *incomplete* contracts to protect against suppliers' opportunism may be less desirable than none at all. Farrell and Shapiro (1989) call this the Principle of Negative Protection. The point is that it is better (ex ante) for customers to be exploited efficiently than inefficiently ex-post. So if contracts cannot set all future variables (e.g. can set prices but not qualities), so customers anyway expect to be exploited ex-post, it may be better that there are no contracts.

2.8.3 Increasing Switching Costs to Enhance Oligopoly Power

Although switching costs typically reduce social surplus, we saw in Sections 2.3-2.5 that they nevertheless often increase firms' profits. If so, firms jointly prefer to commit (before they compete) to real social switching costs than to no switching costs. Thus, firms may artificially create or increase switching costs.

Of course, a firm may prefer switching costs *from* but not *to* its product if it can achieve this, especially where the switching costs are real social costs. Adams (1978) describes how Gillette and its rivals tried to make their razor blades (the profitable follow-on product) fit one another's razors but their razors accept only their own blades. However, Koh (1993) analyses a model in which each duopolist chooses a real social cost of switching to it, and shows the possibility that each chooses a positive switching cost in order to relax competition.[72]

In Banerjee and Summers (1987) and Caminal and Matutes (1990) firms have the option to generate contractual switching costs by committing in period zero to offering repeat-purchase coupons in a two-period duopoly, and both firms (independently) take this option.[73] Similarly Fudenberg and Tirole (2000) explore a two-period model in which firms can price discriminate between consumers based on their past demands; if firms can also offer long term contracts—that is, generate contractual switching costs—then firms do offer such contracts in equilibrium, in addition to spot contracts.[74]

2.8.4 Reducing Switching Costs to Enhance Oligopoly Power

An important class of models which suggests that firms may often be biased towards too much compatibility from the social viewpoint is the "mix-and-match" models (see Section 2.5) in which different firms have different abilities in producing the different components of a "system". Consumers'

[72]Similarly Bouckaert and Degryse (2004) show in a two-period credit market model that each bank may reduce switching costs *from* itself, in order to relax competition.

[73]However, Kim and Koh (2002) find that a firm with a small market share may reduce contractual switching costs by choosing to honour repeat-purchase coupons that its rivals have offered to their old customers.

[74]These papers are discussed in more detail elsewhere in this Volume, in Stole (forthcoming).

ability to mix-and-match the best product(s) offered by each firm is an efficiency gain from compatibility (that is, from zero rather than infinite shopping costs), but firms' private gains from compatibility may be even greater because—perhaps surprisingly—compatibility can increase prices.

In the simplest such model, Einhorn (1992) assumed that a single consumer wants one each of a list of components produced by firms A, B, with production costs a_i and b_i respectively for component i. In compatible competition the price for each component is $\max\{a_i, b_i\}$, so the consumer pays a total price $\sum_i \max\{a_i, b_i\}$ for the system. But if the firms are incompatible, the Bertrand price for a system is $\max\{\sum_i a_i, \sum_i b_i\}$ which is lower unless the same firm is best at everything: if different firms are best at providing different components, then the winning seller on each component appropriates its full efficiency margin in compatible competition, but in *in*compatible competition the winner's margin is its efficiency advantage where it is best, *minus* its rival's advantage where its rival is best. Firms thus (jointly) more than appropriate the efficiency gain from compatibility, and consumers actually prefer incompatibility.

This result depends on (among other assumptions) duopoly at each level. If more than two firms produce each component, the sum of the second-lowest cost of each component (which the consumer pays under compatibility) may easily be lower than the second-lowest system cost when firms are incompatible, so consumers often prefer compatibility and firms' incentives may be biased either way (see Farrell, Monroe and Saloner (1998)).[75]

The "order-statistic" effect emphasised in these models is not the only force, however. Matutes and Regibeau (1988) stressed that under compatibility a price cut by one firm in one component increases demand for the *other* firms' complements, whereas under incompatibility all of this boost in complementary demand accrues to the firm, so compatibility reduces incentives to cut prices.[76] Economides (1989) argued that, unlike the Einhorn

[75]Einhorn's results, but not those of Farrell, Monroe and Saloner, are qualitatively unaffected by whether or not firms know their own efficiencies in each component. The analysis of these two papers is related to Palfrey (1983).

[76]Matutes and Regibeau (1992) allowed firms to set separate prices for bundles (not necessarily the sum of the component prices) and found that the force toward compatibility weakens. Furthermore, compatibility also changes the structure of demand, so

result, this logic does not depend on duopoly, so provides a clear argument why firms may try too hard to reduce switching costs and shopping costs.[77]

2.8.5 Increasing Switching Costs to Prevent or Exploit Entry

The mix-and-match literature of the previous subsection ignores the fact that entry provides a much greater discipline on prices when compatability means a new firm can enter offering just one component of a system than when any entrant needs to offer a whole system.

More generally, we have seen (Section 2.7) that an incumbent firm may protect a monopoly position against entry by writing exclusionary contracts, or by artificially creating real switching costs through technological incompatibility with potential entrants.[78] Imposing contractual switching costs (but not real social switching costs) can also enable an incumbent to extract rents from an entrant without preventing its entry—the entrant is forced to pay a fee (the "liquidated damages") to break the contracts.[79]

2.9 Switching Costs and Policy

As we have seen, with (large) switching costs firms compete over streams of goods and services rather than over single transactions. So one must

even Matutes and Regibeau (1988) found that firms are sometimes biased towards incompatibility. And Klemperer (1992) also shows that firms may prefer incompatibility to compatibility when the latter is socially preferred, and that the firms may even distort their product choices to sustain incompatibility. Garcia Mariñoso (2001) examines a mix-and-match model in which purchase takes place over two periods, and finds that firms are biased towards compatibility because it reduces the intensity of competition in the first period – see also Haucap (2003) and Garcia Mariñoso (2003). (All these models, unlike Einhorn and Farrell, Monroe, and Saloner, assume some product differentiation between firms' components even under compatibility). See also Anderson and Leruth (1993).

[77]Most of the "mix-and-match" literature assumes that each firm offers a full line of products, but DeNicolo (2000) analyzes competition with one full-line and a pair of specialist firms. In our terminology, there are then no additional shopping costs of buying from an additional specialist firm after having bought from one of the specialist firms, but the specialist firms do not internalize the complementarities between them.

[78]Imposing switching costs would not be worthwhile for the incumbent if they reduced consumers' willingnesses to pay by more than the gains from excluding entry. In models such as Rasmusen, Ramseyer, and Wiley (1989), and Segal and Whinston (2000), it is unprofitable to enter and serve only one customer, so no customer loses by signing an exclusive contract if other customers have already done so; in equilibrium this can mean that no customer needs to be compensated for signing an exclusive contract.

Deterring entry is also profitable if it can transfer rents from an entrant to the incumbent.

[79]See Aghion and Bolton (1987) and Diamond and Maskin (1979).

not jump from the fact that buyers become locked in to the conclusion that there is an overall competitive problem. Nor should one draw naive inferences from individual transaction prices, as if each transaction were the locus of ordinary competition. Some individual transactions may be priced well above cost even when no firm has (ex-ante) market power; others may be priced below cost without being in the least predatory.[80] [81] Thus switching-cost markets can be more competitive than they look, and switching costs need not generate supernormal profits, even in a closed oligopoly. These points emerge clearly from the core two-period model with which we began.

But, as our further discussion shows, while switching costs need not cause competitive problems, they probably do make competition more fragile, especially when they coexist with ordinary scale economies (or, as we will see in section 3, with network effects). Because large-scale entry into switching-cost markets is hard (whether or not inefficiently so), there may be much more incentive for monopolizing strategies such as predation or merger than there is in markets in which easy entry limits any market power. Thus switching costs, in combination with other factors, could justify heightened antitrust scrutiny.[82]

Furthermore, while sometimes (as in our core model) firms must give all their ex post rents to consumers in ex ante competition, that is not always

[80]For instance, in an aftermarket context such as the Kodak case, the fact that repair services are priced well above cost does not by itself prove that there is a serious competitive problem.

[81]Another naïve argument is that if one observes little or no switching, then firms do not constrain one another's prices: firms that compete on a life-cycle basis (rather than on an individual transaction basis) constrain one another's life-cycle prices and, of course, firms may be constrained even ex post by the threat of customer switching even when that threat is not carried out in equilibrium.

[82]For example, the UK Competition Commission in July 2001 blocked the proposed merger of two banks, Lloyds TSB and Abbey National, even though Abbey National accounted for only 5 per cent of the market for personal banking. An important part of the Commission's reasoning was that consumer switching costs, combined with some scale economies, make new entry very hard, and that existing firms with low market shares tend to compete more aggressively than larger firms in markets with switching costs, so smaller firms are particularly valuable competitors to retain. (Klemperer is a UK Competition Commissioner, but was not involved in this decision.) See also Lofaro and Ridyard (2003).

Note ˜59 (NAPP) gives another example where policy makers were concerned that entry was very hard in a market with switching costs. In this case the UK regulator (the Director of the Office of Fair Trading) limited NAPP's aftermarket price to no more than five times the foremarket price in order to ameliorate the bargains-then-ripoffs price pattern. (He also limited the absolute level of the aftermarket price.)

true. The ex post rents may be less than fully competed away, as in most of the oligopoly models we discussed. Or, if the ex post rents are dissipated in unproductive activities such as excessive marketing or advertising, then consumers are harmed by switching costs even if firms are no better off. So switching costs often do raise average prices. Moreover, as in our core model, switching costs often cause a bargain-then-ripoff pattern of prices, and (going beyond the core model) this can be inefficient even when the average level of prices remains competitive; they make matching less efficient by discouraging re-matching or the use of multiple suppliers; and, of course, they result in direct costs when consumers do switch.

For these reasons, despite the warnings in our first paragraph, markets may indeed perform less well with switching costs than without, so policy intervention to reduce switching costs may be appropriate.[83] For example, policy might cautiously override intellectual property rights, especially of copyright-like intellectual property that may have little inherent novelty, if those rights are used only as a tool to enforce incompatibility and so create private rewards that bear no relationship to the innovation's incremental value.[84]

In general firms may be biased either towards or against compatibility relative to the social standpoint. But switching costs seem more likely to lower than to raise efficiency, so when firms favor switching costs the reason is often because they enhance monopoly or oligopoly power by directly raising prices or by inhibiting new entry.[85] This suggests that policy-makers should

[83]Gans and King (2001) examine the regulatory trade-offs in intervening to reduce switching costs and show that who is required to bear the costs of ameliorating switching costs can importantly affect the efficiency of the outcome. See also Galbi (2001).

Viard (2003) found that the introduction of number portability for U.S. toll-free telephone services substantially reduced switching costs and led to the largest firm substantially reducing prices; the U.S. wireless industry strongly resisted the introduction of number portability in the wireless market. Aoki and Small (2000) and Gans, King, and Woodbridge (2001) also analyse number portability in the telecoms market.

The UK government is currently considering recommendations to reduce switching costs in the mortgage market, see Miles (2004).

[84]Thus, for example, the European Commission in 2004 ruled that Microsoft had abused its market power by, *inter alia*, refusing to supply interface infrastructure to competitors, thus making entry hard by products that could form part of a "mix-and-match" system with Microsoft's dominant Windows PC operating system. Microsoft was ordered to provide this information even if it was protected by intellectual property.

[85]A caveat is that firms often do not make a coordinated joint choice of whether to

take a close look when firms with market power choose to have switching costs (through contract form or product design) when choosing compatibility would be no more costly.[86] [87]

3 Network Effects and Competition

4 Conclusion

References

compete with switching costs or without, and different firms may be able to control the costs of different switches. See Section 2.8.

[86]For example, the Swedish competition authority argued that Scandinavian Airlines' "frequent-flyer" program blocked new entry on just one or a few routes in the Swedish domestic air-travel market in which entry on the whole range of routes was impractical (see Section 2.7.3), and the airline was ordered to alter the program from October 2001. A similar decision was made by the Norwegian competition authority with effect from April 2002. Fernandes (2001) provides some support for these decisions by studying alliances formed by U.S. airlines, and showing that "frequent-flyer" programs that cover more routes are more attractive to consumers and confer greater market power on the airlines operating the programs. See also Klemperer and Png (1986).

[87]A caveat is that the policy debate is often held ex-post of some lock-in. At this point incumbents' preference to maintain high switching costs is unsurprising and does not prove that switching costs raise prices overall (nor do the switching costs necessarily cause inefficiencies). Reducing switching costs ex-post also expropriates the incumbents' ex-ante investments, which may be thought objectionable, though the fear of expropriation of this kind of ex-ante investment seems unlikely to harm dynamic efficiency (and may in fact improve efficiency).