

CONSUMER SWITCHING COSTS: A SURVEY

A. J. PADILLA*

Nuffield College, Oxford

En este artículo nuestro propósito es revisar la literatura existente sobre los costes de sustitución en el consumo (consumer switching costs). En particular, nos centraremos en las distintas implicaciones que los diferentes tipos de switching costs (exógenos y endógenos) tienen en el funcionamiento de mercados oligopolistas. Algunos desarrollos recientes de esta literatura, como las implicaciones macroeconómicas de los switching costs o su impacto en el modo de evolución de la estructura industrial, son también analizados. Finalmente, el artículo concluye con algunas sugerencias de cara a futuros desarrollos.

1. Introduction

Consumer Switching Costs can be regarded as the costs that a consumer may have to face when switching suppliers. In other words, these are the costs that result when a consumer switches from the firm he was patronizing to a different one. In principle, the existence of switching costs is not restricted to relationships between firms and their customers, but also applies to any situation in which, once two parties have traded, staying together confers a gain relative to starting new trading relationships¹. In a broad sense thus, they can be understood as an example of idiosyncratic investment. Notice that we do not identify the possible ultimate causes for the existence of these costs nor specify who pays for them. In addition, note that under some circumstances switching costs are actually «start-up costs», i.e. the costs of starting a new trading relationship, a concept that generalizes the previous definition to some extent. Despite their different meanings, their consequences are similar and hence, we will not make any distinction between them in what follows. In conclusion, switching costs generate brand loyalty in many markets even when the products of other brands are functionally identical. We could say that ex-

* This paper is a revised version of the second chapter of my M. Phil Thesis at the University of Oxford. Parts of this paper have drawn heavily on Klemperer (1990b). I am grateful to my supervisor Chris Harris for his helpful suggestions. This work has substantially benefited from the comments of P. Bertoletti, F. Daveri, P. Klemperer, K. U. Kühn, E. Pint, an anonymous referee and, the Editorial Board of this review. The usual caveats apply. Financial support from the Bank of Spain is gratefully acknowledged.

¹ See Williamson (1976).

ante homogeneous products, become ex-post differentiated because of the existence of switching costs².

One form of classification distinguishes three types of switching costs (Klemperer (1987a)): transaction costs, learning costs, and artificial or contractual costs. Examples of the first type are the costs of switching between two banks offering identical facilities; or of returning rented equipment to substitute it for identical equipment from an alternative supplier, and so on. In the case of learning costs a typical example is the costs of learning how to operate a particular brand of software, so that, even when different manufacturers produce functionally identical machines, a consumer prefers to keep buying the original brand provided that different brands of software are not compatible. These two types of switching costs are said to reflect *real* social costs of switching between firms, even though their size can be determined or at least influenced by firms. Finally, it is easy to find examples of the third type of switching costs: frequent flyer plans, landlord's practice of charging high rents to new tenants and lower rents to old tenants, casinos' practice of laying on all sorts of special benefits for regular customers (Banerjee-Summers (1987)). Here, switching costs are created by firms and they are characterized by the absence of any real costs of switching. Similar costs can be created by other sort of contracts³.

Another important classification made in the literature, that is not directly related to the previous one, is between *exogenous and endogenous* switching costs. Sometimes switching costs are completely independent of firms' or consumers' decisions, for instance, they may be created by the existent technology. In this case we speak of exogenous switching costs. However, there are situations where switching costs are a direct consequence of firms' decisions; these are called endogenous switching costs. A particularly important kind of endogenous switching costs are those originated in pricing policies that discriminate between loyal customers and newcomers (what we called artificial costs above). Examples of relationships between buyers and sellers that involve lower transaction prices for loyal customers than those paid by buyers with shorter history of purchases from the same seller are, prime rates on bank loans, frequent flyer programs and supermarket coupons. Pricing practices like these may arise in very different industries with different market structures. They may take different forms, such as precommitment to a price or precommitment to a discount (in absolute or relative terms)⁴. These pricing

² Switching costs have also been called costs of substitution (Von Weiszacker (1984)); lock-in (Farrell (1986)); market inertia (Rosenthal (1982)); etc.

³ Other types of switching costs of less relevance are for instance, psychological costs, that is, when consumers keep buying from the same firm because of habit or loyalty without any substantial economic reason. In addition, switching costs may also arise when there is uncertainty about the product quality so that consumers are reluctant to switch to untested products. Here, some complications such as the possibility of using prices to signal quality arise. See Schmalensee (1982).

⁴ See Caminal-Matutes (1989).

policies are used to generate inertia in market share through artificially created switching costs. They differ fundamentally from quantity discounts or general non-linear pricing schemes where firms commit to a schedule of absolute prices for purchases of different quantities.

The concept of switching costs should be of interest if it satisfies at least two requirements, namely, it helps to create a better understanding of real phenomena and it is consistent with the more general paradigm in which it is inserted, in other words if it is theoretically relevant. We have already provided some examples of the importance of switching costs in the real world⁵, and so we turn now to its theoretical value. Switching costs may explain why managers seem more concerned with market share than with short-run profitability and, further, why this is rational behaviour. Switching costs give firms monopoly power over their customer base and in consequence make current market share an important determinant of future profits. They may also explain phenomena such as price wars and sales. In general, models with switching costs, by stressing the links between time periods and the importance of market share on future profits, help to increase our understanding of competition for market share. In particular, switching costs allow outcomes that look collusive in a repeated non-cooperative framework even when we rule out the possibility of bootstrap equilibria (e.g. equilibria sustained by trigger strategies). Finally, switching costs may have important consequences for macro-economics e.g. explaining why the price-costs margin moves counter-cyclically.

The paper is organized as follows. Section 2 presents a model of the effects of exogenous switching costs on the degree of competitiveness. We attempt to integrate different results in a simple framework which draws from Klemperer (1990 b) and Padilla (1990a, b). Finally, we review the literature on endogenous switching costs. Section 3 deals with some recent developments such as the study of price wars, entry and the evolution of industrial structure in this type of market. The macroeconomic implications of switching costs are also thoroughly reviewed. Section 4 concludes.

2. The competitiveness of markets with consumer switching costs

Most of the existing models on switching costs have focused on how the competitiveness of a market is modified by the presence of such costs (von Weizsäcker (1984), Farrell (1986), Klemperer (1987a, b), Farrell-Shapiro (1988), Beggs-Klemperer (1989), Padilla (1990)). In this section we proceed to describe the main results obtained in this literature. So we concentrate on how competition is relaxed when switching is costly in a two period model of exogenous switching costs. Then, the main results of our two-period analysis are generalized in a multi-period framework. Finally, we look at models of endo-

⁵ This is a qualitative statement since we are not aware of any empirical study of quantitative character. A partial exception could be the evidence provided by Banerjee-Summers (1987).

genous switching costs. The fundamental results found in the analysis of exogenous switching costs are shown to be generally valid in the case of endogenous costs. Let us concentrate first on a two period model where switching costs are exogenous.

At an intuitive level, the effect of switching costs on the degree of competitiveness differs in each of the two periods to be considered. Because switching costs locked consumers into the firms they were patronizing, firms succeeded in exploiting their customer bases in the second period. Second-period competition is then relaxed, and therefore second-period profitability increased. All this makes market share intrinsically valuable. In consequence, firms try to obtain the largest market share possible in the first period. Both firms charge prices below those optimal with no switching costs in order to secure a large number of locked-in consumers in the second period. First period competition becomes thus fiercer when there are consumer switching costs. However, the net effect on overall competition and profitability is, in most cases, ambiguous. Though second period competition is relaxed, this is partially or totally offset by fiercer competition in the first period.

More formally let us denote the profits of the i th firm, $i = 1, 2$, in period t , $t = 1, 2$, by Π_t^i . The first period profit function of firm i is $\Pi_1^i = \Pi_1^i(p_1^i, p_1^j)$ which depends on firm's own price and the price of its rival. Similarly, for the second period, $\Pi_2^i = \Pi_2^i(p_2^i, p_2^j, \sigma_i)$ where σ_i represents the market share of firm i in the previous period. In the second period, firms maximize Π_2^i given their respective market shares σ_i . Since consumers find costly to switch from the firm they were buying in the first period, firms may charge higher prices than in the absence of switching costs without losing its customer base. In period one, each firm wants to maximize its total discounted profits,

$$\pi^i = \Pi_1^i(p_1^i, p_1^j) + \delta \Pi_2^i(p_2^i, p_2^j, \sigma_i) \quad [1]$$

where δ is the discount factor, $0 < \delta < 1$, common to both firms (Π_t^i is assumed to be continuous, twice differentiable and strictly concave). The first order conditions on first period prices are

$$\partial \Pi^i / \partial p_1^i = (\partial \Pi_1^i / \partial p_1^i) + \delta (\partial \Pi_2^i / \partial \sigma_i) (\partial \sigma_i / \partial p_1^i) = 0 \quad [2]$$

Now, provided that $\partial \sigma_i / \partial p_1^i < 0$ and $\partial \Pi_2^i / \partial \sigma_i > 0$, it must be the case that $\partial \Pi_1^i / \partial p_1^i > 0$. That is, provided that an increase in prices leads to a lower market share, and market share is valuable, competition in the first period increases with respect to the no switching costs case (note that p_1^i , the price actually charged by firm i in the first period is lower than the price for which $\partial \Pi_1^i / \partial p_1^i = 0$). However, we can construct situations in which one of the two requirements above fails, and so our previous intuition. This is the origin of the different results existing in the literature. For instance, our previous result may be modified by the presence of rational consumers. Rational consumers recognize that a lower price in the first period lead to an increase in market share and hence implies higher prices tomorrow. Because of their ability to anti-

pate the possibility of exploitation in the second period, rational consumers will not react immediately to a price cut in the first period. Therefore, competition in the first period is mitigated and it may be even the case that the first period results less competitive than in an identical market without switching costs. In addition, it is possible that $\partial \Pi_2 / \partial \sigma_i$ is negative and market share is not valuable, e. g., when there is growing demand from unattached consumers (Klemperer (1987c), Padilla (1990b)).

Given that market share will be in general valuable, it becomes clear that whether switching costs lead to a more competitive outcome or not depends on the sign and magnitude of $\partial \sigma_i / \partial p_i^1$, that is, on the sign and magnitude of the *degree of price sensitivity*. Padilla (1990b) constitutes a natural framework where these effects can be studied. It is a model where products are undifferentiated except by switching costs. It also allows for the existence of dynamic consumer turnover and differently informed consumers (the latter differentiates it from previous papers in this area, e. g. Klemperer (1987a)). Therefore two types of consumers coexist in the market in each period. Firstly, there is a proportion of price sensitive consumers who react to small changes in firms' prices. Secondly, some other consumers are price insensitive either because they are locked into a particular firm or, even if uncommitted to any firm, they are only imperfectly informed about firms' prices. Padilla (1990b) shows that it is precisely the proportion of price sensitive consumers which drives the results on competitiveness. So, when this proportion approaches zero (equivalent to $\partial \sigma_i / \partial p_i^1$, small) collusive outcomes result; when these perfectly sensitive consumers form the whole of the market, then the outcome is highly competitive. In this model, the usual distinction between the cases of myopic and rational consumers can be handled by the same analysis. Myopic consumers are highly sensitive to price cuts and therefore we should expect tough competition for market share. Rational consumers do not react so much to price-cutting since they anticipate future firms' behaviour. Therefore, competition for market share will be less fierce.

All this very much depends on whether the firm can credibly commit to a given path of future prices or not. In fact, our previous analysis assumed that firms cannot credibly commit not to exploit locked-in consumers in the second period. Using a model with differentiated products and with consumers uncertain about their future preferences von Weizsacker (1984) shows that when firms are able to commit to charging the same price in the future, exogenous switching costs make the market more competitive. The reason is that the higher switching costs are, the more consumers' choices are influenced by the future and the less important are their current preferences so that the products are effectively less differentiated. In other words, consumers are more *sensitive* to prices and therefore firms have a higher incentive to undercut. Von Weizsacker calls this a *reputation equilibrium*, where firms commit to maintain the same price in every period. Nevertheless, as we just noticed above, once a consumer has bought from one firm the existence of a cost of switching decreases his sensitiveness to price cuts and therefore, the patronized firm has an incentive to exploit him. Therefore, once the ability of firms

to commit in dispensed with switching costs, cannot lead to an unequivocal increase in competitiveness, and we have to come back to our previous considerations.

In conclusion, two period models have not been able to solve definitively the ambiguity about the effect of switching costs on the degree of competitiveness. The natural way to proceed is then to move to multiperiod models (see Beggs-Klemperer (1989)⁶). This model has been able to clarify this effect by showing that under general circumstances and in different settings, switching costs tend to raise prices.

Following Klemperer (1990b), in a multiperiod setting each firm i wants to maximize,

$$V_t^i = \Pi_t^i(p_t^i, p_t^j) + \delta V_{t+1}^i \quad [3]$$

where V_t^i is the valuation function of the i th firm at t . The first order condition on p_t^i is

$$\partial \Pi_t^i / \partial p_t^i = \delta [(\partial V_{t+1}^i / \partial p_t^i) + (\partial V_{t+1}^i / \partial p_{t+1}^j) (\partial p_{t+1}^j / \partial p_t^i)] \quad [4]$$

Again we have two opposite forces interacting. First, firms want to keep prices low to build up market share. Second, charging higher prices and exploiting the consumer base today has a direct effect on profitability. However, as Klemperer (1990b) notes, there are several reasons why we expect this second effect to dominate. In the first place, today is always more valuable than tomorrow since future is discounted. Second, rational consumers cannot be cheated by undercutting, since they will anticipate future exploitation. Finally, higher prices today may relax price competition tomorrow so that we have what we could call an *intertemporal strategic complementarity*. We should conclude then, that switching costs do relax price competition in a multiperiod framework.

Formally, if $\partial V_{t+1}^i / \partial p_{t+1}^j < 0$, $\partial V_{t+1}^i / \partial p_{t+1}^i > 0$ and $\partial p_{t+1}^j / \partial p_t^i > 0$, then $\partial \Pi_t^i / \partial p_t^i < 0$ for all t which means that prices are larger than those charged in the absence of switching costs.

Before considering the case of endogenous switching costs, it is convenient to note that not only are there differences in the degree of competitiveness but also in the actual path followed by prices. Most two period models derive an increasing sequence of prices, that is second period prices are higher than first period prices. This regardless of whether competition is encouraged in the first period or not. This result arises from the two period structure of the game. However, there are cases e. g., when consumers do not pay the start-up costs or when firms can price discriminate between loyal customers and newcomers, where a decreasing price pattern may obtain.

⁶ Farrell-Shapiro (1988) developed the first many period model of switching costs, but in a context in which the absence of switching costs results in zero profits, so that switching costs must necessarily increase firms profits.

2.1. Endogenous Switching Costs

Endogenous switching costs have received far less attention than exogenous switching costs. Obviously, this is because endogeneity introduces an additional dimension to the strategy spaces of the firms, making its treatment more complex. One case that has been studied is endogenous switching costs caused by discriminatory pricing policies. Two good examples of this literature are Banerjee-Summers (1987) and, Caminal-Matutes (1989).

Banerjee-Summers (1987) consider a two period model where two firms compete in prices⁷. Both firms produce a homogeneous good at zero marginal cost and attempt to maximize total profits. There is a continuum of consumers $[0,1]$, the same in both periods, with unit inelastic demands in each period at constant reservation price. Consumers are rational in the sense that they maximize life-time utility given this period's and next period's expected prices. Firms first announce coupons, that is, discounts for second period prices granted to the first period consumers if they repeat purchase in period two. Then, firms compete in prices. The results of this model support our analysis above claiming that switching costs act as a collusive device. Prices in period one and the (effective) price in period two coincide with the monopoly price. This means that the introduction of coupons may reduce competition in both periods, and hence in overall terms.

In Banerjee-Summers when firms commit with coupons they are able to attain fully collusive outcomes. This should not lead us to conclude that any form of commitment will produce the same results. In fact, Caminal-Matutes (1989) have shown that the specific form of commitment has rather strong implications for the performance of such a conduct. Using a duopolistic model of price competition with differentiated products, they show how two different kinds of commitment i. e., precommitment with coupons and precommitment to second period prices, result in different overall profitability. Both forms of commitment represent different ways of endogenizing switching costs, and both generate a decreasing sequence of prices. However, whereas precommitment with coupons leads to an overall increase in profitability with respect to a situation with no commitment, precommitment to a special second period price for loyal consumers produces the opposite result⁸. Caminal-Matutes suggest that this discrepancy with Banerjee-Summers is due to the lack to product differentiation in this latter model. The absence of product differentiation increases the collusive effect of coupons. Caminal-Matutes then analyse the case where the firms can decide the type of commitment at the same time that they decide whether to commit or not. Under these circumstances it is shown that the equilibrium involves price commitment. Summarising, they argue that the impact of switching costs on the equilibrium price dynamics and profitability depends on whether swit-

⁷ There is sequential price setting, that is, one firm acts as a Stackelberg leader and the other as the follower. The leader is either determined randomly or by size (i.e., the largest firm is the leader).

⁸ Nevertheless, the equilibrium path of prices is decreasing both for loyal customers and newcomers.

ching costs are paid by firms or by consumers. When switching costs are paid by consumers, as it was the case in the typical two period model described at the beginning of this section, the price sequence is increasing. When they are paid by firms like in Caminal-Matutes the price sequence is decreasing. They also argue that the distinction between exogenous and endogenous switching costs does matter regardless of who pays for them. The argument is that overall profitability is different depending upon whether switching costs are endogenized through precommitment to a special second period price or precommitment with coupons.

However, Beggs-Klemperer (1989) argue that the evolution of the price pattern may be independent of who pays for the switching costs. In their model, the time path of the effective price to consumers (that is, of the price plus any start-up or switching costs the consumer must pay himself) is the same whether the consumers must pay the switching (or start-up) costs themselves or whether firms pay these costs for consumers. (The time path of this total effective price is decreasing, as in Caminal-Matutes). In addition, any switching cost that arises endogenously in a model could have been imposed exogenously on the same model structure, yielding the same price path. (And in at least some of the existing models of exogenous switching costs e. g. Klemperer (1987a,b), Padilla (1990b), could have arisen endogenously if an additional stage had been added to the beginning of the game in which firms were able, perhaps through their technology choices, to choose whether or not to create switching costs). In our view the distinction between exogenous and endogenous switching costs cannot, per se, matter for price paths. The derived differences in overall profitability seem to owe more to the possibility of price discriminating between loyal customers and newcomers than to the type of switching costs considered. In particular, note that in Caminal-Matutes (1989) firms commit unilaterally and are able to price discriminate. (For the latter they need to preclude the existence of second hand markets and the possibility of storage).

3. Other topics in consumer switching costs

This section reviews some of the recent developments on switching costs. We concentrate on the relevance of switching costs to consumer choice among competing technologies; on price wars and sales; on the role that this concept may play in the explanation of the evolution of industrial structure and entry decisions; and finally on the macroeconomic implications of switching costs.

3.1. *Network Externalities and Product Compatibility*

By and large, (positive) network externalities arise when a product is more valuable to a user the more users adopt the same good or similar ones⁹. There

⁹ This definition is taken from Tirole (1988) chp. 10. For surveys on this literature see Matutes (1989) and Tirole (1988).

are two types of network externalities, direct and indirect. An example of the first type is computer hardware & software. The more users adopt a brand of computer hardware, the more software will be written for that brand, making the hardware more valuable. Indirect externalities are generated by increasing returns to scale in production. In this case a greater number of compatible products can be supplied at a lower cost if the network expands. Beggs (1989) focuses on a particular example of this kind of externality, created by the joint existence of consumer switching costs and increasing returns. He points out that these externalities may lead to the adoption of inferior technologies and thus, to an inefficient outcome, because the more consumers buy a product the more likely it is to survive and the more attractive it is to other consumers. Which product will survive does not depend exclusively on its technological efficiency but also on the (self-fulfilling) expectations of consumers.

Therefore, provided that the cost differences are not too great, it is possible to construct an equilibrium where the less efficient firm wins. Of course, it is also possible to construct an equilibrium in which the more efficient firm wins. The essential feature of Beggs model is that consumers expect a surplus in the second period because of the need to buy spare parts and the fact that they are completely locked-in.

A somewhat similar result is obtained by Farrell-Shapiro (1988). Their model is one of overlapping generations and switching costs. In each period there are two existing cohorts of consumers, youngsters and oldsters, each of which lives exactly two periods. Suppose that each consumer would pay a premium, z , for a product consumed by both cohorts, in comparison with one consumed only by his own cohort. Then $z > 0$ implies that one firm serving all consumers is the efficient outcome. For small values of z this will not happen, because the firm with the largest market share prefers to exploit its customer base. Therefore, the presence of switching costs also leads to an inefficient outcome in this case.

To increase social welfare, it would be desirable to ensure that the more efficient firm serves the whole market. The validity of tax and subsidy schemes to improve welfare hinges upon the observability of the costs of the two firms in Beggs' model. However, as Beggs (1989) notes, if the government were able to sell licences to produce in this market, it would assign the license to the most efficient firm by conducting a second price auction that would induce truthful revelation of the respective valuations. Roughly speaking, the presence of switching costs and their negative effects on welfare support the imposition of compulsory standards¹⁰, which would have the effect of diluting such costs.

So far, we have dealt with the implications of exogenous switching costs on technology choices. Klemperer-Regibeau (1988) consider endogenous switching costs. In this case, firms' equilibrium policies tend to eliminate incompatibilities among firms' products and therefore, eliminate the extent of such

¹⁰ For an account of the advantages and disadvantages of standardization see Tirole (1988) chp. 10.

costs. This result supports the view that the kind of inefficiencies reviewed above are due more to coordination failures on the demand side than to problems of product compatibility on the supply side.

3.2. Price Wars and Sales

Though price wars and sales are quite similar in appearance, they show clear-cut differences in motivation and context. Price wars appear as a consequence of some kinds of oligopolistic competition and they are clearly incompatible with a monopolistic market in the absence of potential entry. On the contrary, it is always possible to design models in which a monopoly alternates periods of sales with periods of high prices. Nevertheless, both price wars and sales show an identical evolution of the pricing path over time. There is initial price cutting, a period in which low prices are sustained and finally, a return to higher prices.

Gallini-Karp (1989) show that in a monopolistic market, consumer lock-in may give rise to a dynamic price path with alternating periods of high prices and sales (low prices). Sales in this model are due to the impossibility of the monopoly to commit credibly to a future price path as a consequence of the incorporation into the market of new consumers who are both price sensitive, and have rational expectations about future prices. They only buy at low prices because they know in advance that the existence of switching costs will allow the monopoly to exploit them in future periods. Thus, the dynamically consistent path followed by the new product monopolist unable to commit to future prices consists of alternating periods of steady high prices, during which only old or captive customers buy the product, and periods of low prices during which new consumers adopt it. Padilla (1990a,b) undertake the extension of this model to allow for strategic competition in a two period framework. Here, sales are low realizations of a mixed strategy equilibrium in the tradition of Varian (1980). However, this model is not fully dynamic and therefore its conclusions are bound to be partial.

With respect to price wars, Klemperer (1989) formulates a model where price wars are caused by switching costs. This is a four-period model of quantity competition where in each period prices are such that the market clears. Klemperer studies an industry where a dominant firm coexists with a competitive fringe. Consumers have different reservation prices for a single unit of the product. In this model new entry triggers a price war in equilibrium. We can distinguish two types of price wars. Type 1, corresponding to small costs, is characterized by the post-entry price of the dominant firm not being higher than the price charged by the entrants. Type 2, with large switching costs, the incumbent's price remains higher both in the entry period and the post-entry period, though it evolves as a typical price war. In the first type new entrants' price rise in the post-entry period because loyal consumers are to some extent locked-in. Since the incumbent cannot set its entry period price much higher than entrants' price to avoid switching, the incumbent also raises its price in the post-entry period. However, in the second type, the incumbent may be

able to sustain higher prices than the entrants' both in the entry and post-entry period. This is because he acquires a stock of high-reservation-price consumers before entry occurs. Note finally, that in Klemperer's model, markets with switching costs may generate limit pricing behaviour in an equilibrium with complete information. The incumbent may in some circumstances cut its price before entry takes place in order to lock-in consumers and thus reduce the scale of entry.

Another example of consumer switching costs as a possible cause of price wars is Padilla (1990a,b). In this model the occurrence of price wars in markets with switching costs is accidental (stochastic). The probability of a price war is higher the higher the proportion of price sensitive consumers (uncommitted and perfectly informed) in the market. This model differs from Klemperer (1989) in several respects. Whereas Klemperer considers a dominant firm versus a fringe of competitors, Padilla considers a duopoly in which there is no Stackelberg leader. In Klemperer's model price wars are triggered by entry while in Padilla, they are perfectly accidental.

3.3. *The evolution of the industrial structure*

There is a very extensive and dynamic literature on the evolution of industry¹¹. Recently, some duopolistic models have explored the relevance of consumer switching costs to the study of industry evolution. In one of them (Farrell-Shapiro (1988)) market shares remain constant on average, so that we can speak of *action-reaction* in market share. The term action-reaction was firstly used by Vickers (1986) in this context, and it defines a process in which market leadership changes hands constantly. The opposite evolution is denoted by *increasing dominance*. A more interesting example of this literature in Beggs-Klemperer (1989). They obtain action-reaction in the evolution of the industry (represented in these models as the evolution of market share). Set σ_{it} denote the market share of firm i at time t . Then, the following equation describes the evolution of market structure in Beggs-Klemperer,

$$\sigma_{it} = \sigma_i + \mu^t(\sigma_{i0} - \sigma_i) \quad [5]$$

where σ_{i0} is the initial market share σ_i is its long-run value. Beggs-Klemperer show that in their model there is a monotonic, $\mu > 0$, convergence in the direction of decreasing dominance, $\mu < 1$, and this convergence is quite fast, $\mu < 1/3$. Market shares converge in both Farrell-Shapiro and Beggs-Klemperer because the larger firm will charge higher prices, selling to fewer new consumers. Whether this convergence is oscillatory or monotonic depends on the specific assumptions concerning the dynamics of the model. Oscillations are due to the overlapping generations and sequential price setting assumptions of Farrell-Shapiro (1988). Monotonicity seems to be due to the functional forms of the valuation functions of consumers and firms in Beggs-Klemperer (1989).

¹¹ For instance, Reinganum (1985); Vickers (1986); Harris (1988).

This tendency of models with consumer switching costs to support action-reaction equilibria contrasts with previous approaches to this problem, e. g., Harris (1988), Vickers (1986). From these latter models it seems clear that the natural evolution of market structure is characterized by increasing dominance, that is, by a tendency towards more asymmetric market structures. Beggs-Klemperer (1989) provide some possible explanations for that contradiction. First, in their model effort rates (prices) affect current flow profits and thus make it cheaper for the follower (smaller firm) to offer lower prices (work harder). This does not occur in other models because profits are additively separable from the costs of attracting new consumers. In the second place, rational consumers and consumer turnover serve to erode dominance. Rational consumers prefer to be allocated to the smaller firm since they expect it to charge lower prices in the future. Consumer turnover also erodes dominance because locked-in consumers progressively leave the market.

The action-reaction results have been criticised on several grounds. For instance, Farrell-Shapiro (1988) results on action-reaction owe much to the particular overlapping generations framework assumed. Similarly, the Beggs-Klemperer (1989) results are inevitable given that they restrict themselves to looking for equilibria in which firms' equilibrium prices and consumers' value functions are both affine functions of firms' market shares. In any such equilibrium, firms' market shares must evolve according to a linear difference equation as that in equation [5]. In addition, market shares have to satisfy certain boundary constraints, namely $0 \leq \sigma_{it} \leq N \forall_{it}$, where N is the total number of consumers. The combination of linear dynamics and boundary conditions implies necessarily action-reaction as it can be seen in Figure 1.

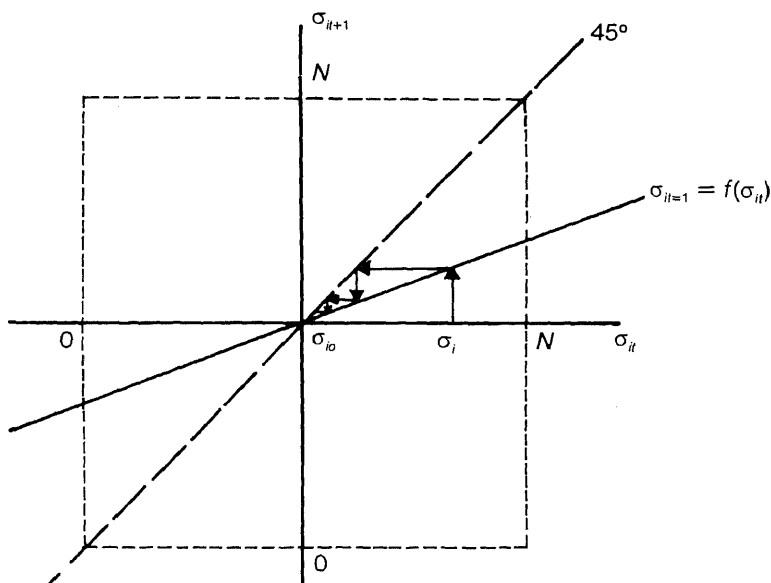


Figure 1

Linear dynamics and boundary conditions imply action-reaction.

In this figure, we draw σ_{it+1} as a function of the previous market share, $\sigma_{it+1} = f(\sigma_{it})$ where f is a linear function representing [5] above. Then f' , equal to μ in [5], has to be smaller than one if the boundary conditions are to hold. Hence, action-reaction obtains. Beggs-Klemperer only show that action-reaction is a possible equilibrium. They do not show, however, that it is the only one in this type of models.

Nevertheless, some support for action-reaction results has been recently provided by Padilla (1990b) in a two period model. In this model, two firms producing an ex-ante homogenous product compete in prices in a market with consumer switching costs and unit inelastic demands. Consider two firms A and B with market shares σ_A, σ_B respectively, $\sigma_A \geq \sigma_B$ without loss of generality. It can be shown that whenever locked-in and uncommitted consumers coexist in the market, the unique equilibrium is characterized by each firm naming prices according to the following distribution function

$$F_i(p) = \frac{(p - \underline{p}) [\sigma_j + \psi N/2]}{\bar{p} \psi N} \quad i, j = A, B. \quad i \neq j. \tag{6}$$

where ψ is the proportion of uncommitted consumers, \underline{p} is the infimum of the support of F_i and N is the total number of consumers in the market. In addition \bar{p} belongs to the common support of F_i, F_j . Firm A names higher prices than firm B in expectation as it is easy to see from Figure 2 below. Hence, action-reaction obtains at least in expected terms. Notice that nothing excludes the occurrence of increasing dominance with positive probability. In conclusion there is still some room for further research on this point.

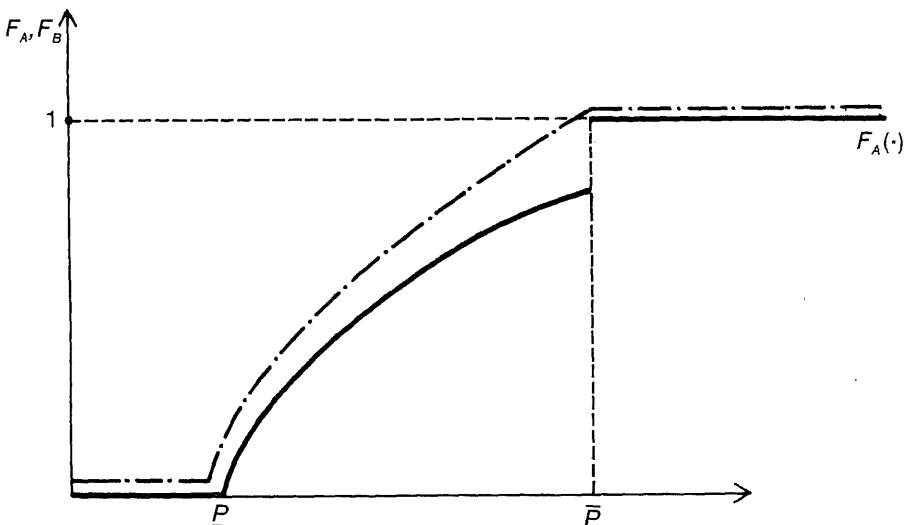


Figure 2
Action-reaction mixed strategy equilibrium (Padilla (1990a,b)).

3.4. Entry in markets with consumer switching costs

We will proceed to review the main results suggested by the existing literature on entry in markets with switching costs. Switching costs tend to make entry more attractive since in most cases competition is relaxed with respect to markets without switching costs. However, we know from section 2 that there are examples where this result is reversed.

In some models, e.g. Farrell-Shapiro (1988), Klemperer (1988), entry into these markets may be socially detrimental. It follows that a public policy oriented to the elimination of barriers to entry may be not necessarily well-founded in industries with switching costs. However, Padilla (1990) finds entry to be socially desirable even when very large switching costs are involved if there are uninformed consumers.

Klemperer (1988) analyzes the case of a monopolist facing a potential entrant when there are no explicit barriers to entry. If the potential entrant decides to enter, a single period of Cournot competition follows in which the incumbent and the follower choose outputs q^i , q^e respectively. Each consumer has switching costs s of buying from a new entrant and the q th consumer has a reservation price $f(q)$ net of switching cost. In any post-entry equilibrium, the incumbent's price is $p^i = f(q^i + q^e)$ and the entrant's price is equal to $p^e = p^i - s$. p^e is such that at least some consumers are willing to switch to the entrant's product, which is modeled as functionally identical to the incumbent's. In this model, entry is welfare reducing for some values of the switching costs s . These values depend on the exact form of the demand curve $f(\cdot)$. Entry causes a welfare loss of $s[q^m - q^i]$, where q^m is monopoly output. This loss is caused by those consumers switching to the new entrant. This negative effect is partially offset by the gain in consumer surplus due to extra production, that is,

$$\int_{q^m}^{q^i+q^e} [f(q) - s - c] dq. \quad [7]$$

where c is the unit cost of production. For small q^e , the social value $f(\cdot) - s - c$ of each unit produced in excess of q^m is also small, because the entering firm sets its marginal revenue (approx. equal to $f(q^i + q^e) - s$) equal to the unit cost of production. Thus any gain is of second order while the loss is always of first order. Therefore, a small amount of entry is welfare-reducing whenever the incumbent's output declines with entry. This result also holds when the entrant's unit costs are lower than the incumbent's, provided the difference is not big. The relationship is depicted graphically in Figure 3

The results on entry deterrence are mixed for these markets. Common sense would suggest that larger switching costs are more effective for entry deterrence but this is not actually true. For example, Klemperer (1987c) notes that entry deterrence is easier when switching costs are either very large or very small. The rationale behind this is the double role played by switching costs.

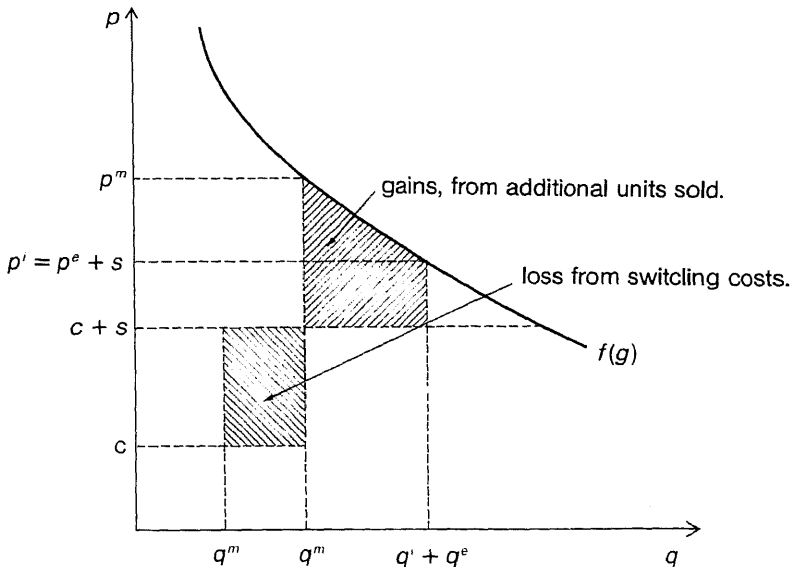


Figure 3

Welfare Effects of Entry into markets with Switching Costs (Klemperer (1987c)).

First, they lock in consumers, but second, by reducing the incentive to fight for new consumers later, they may end up encouraging entry.

Finally, we consider limit pricing. Both limit pricing (or overinvestment in market share) and limit over-pricing (or underinvestment) may be rational strategies for an incumbent facing new entry. A firm may want to cut price before entry takes place, building up a large customer base and so reducing the potential profits of any new entrant. However, if demand is growing and/or there is a strong consumer turnover, the incumbent may prefer higher prices. Underinvestment in market share signals that the incumbent will fight if entry takes place. Some examples of this behaviour (lean and hungry look) can be found in Klemperer (1987c) and Padilla (1990b).

3.5. Macroeconomic Implications

The macroeconomic implications of markets with switching costs are considered to be an important field for future research in this area. In a more general sense, it has not been until very recently that the study of imperfect competition has been incorporated to macroeconomic modelling¹². Focusing on switching costs, two recent contributions are Beggs-Klemperer (1989) and Froot-Klemperer (1989). (Although the latter refers to any situation where market shares matter, it can be regarded as a model of switching costs).

It is not our purpose to analyze these models in detail but simply to highlight the kind of problems they deal with and to evaluate their strengths and weak-

¹² See the excellent survey of Carlton (1989).

nesses. Two main questions have been studied in these models though both are related to the issue of price stickiness. First, because switching costs place value on market shares, they may explain why we sometimes observe countercyclical movements in price-cost margins. So, Beggs-Klemperer (1989) study the effects of a temporary boom in demand. They show that when switching costs are present, price-cost margins may increase during recessions, and conversely, decrease during expansions. They work with a multiperiod model, the ideal setting for the study of temporary shocks. However, the main idea may be illustrated using a simple two period model; The exposition of the remainder of this section follows Klemperer (1990b). We have a duopoly where the total discounted profits of each firm are,

$$\Pi^i = \Pi_1^i(p_1^i, p_1^j) + \delta \Pi_2^i(p_2^i, p_2^j, \sigma_i) \quad [8]$$

where $\sigma_i = \sigma_i(p_1^i, p_1^j)$ is the market share of firm i at 1; $i = 1, 2$. The first order condition for profit maximization with respect to p_1^i implies

$$\partial \Pi_1^i / \partial p_1^i = -\delta (\partial \Pi_2^i / \partial \sigma_i) (\partial \sigma_i / \partial p_1^i), \quad i = 1, 2 \quad [9]$$

A boom in demand during period one can be represented by a decline in the discount factor δ . In this case if consumers are myopic,

$$\partial / \partial \delta [\partial \Pi_1^i / \partial p_1^i] = -(\partial \Pi_2^i / \partial \sigma_i) (\partial \sigma_i / \partial p_1^i)^{13} \quad [10]$$

and therefore, $\partial / \partial \delta (\cdot) > 0$ provided that $\partial \Pi_2^i / \partial \sigma_i > 0$ and $\partial \sigma_i / \partial p_1^i < 0$. That is, a temporary boom in demand in period one leads to an increase in the degree of competitiveness, and consequently to a reduction in prices in the first period (obviously this result hinges upon the concavity of the profit function). With switching costs, a boom leads to smaller increases in price-cost margins than in an otherwise identical market without switching costs. Prices appear to be sticky despite the fact that they are perfectly flexible. This could be an alternative explanation to observed price rigidities, different from menu costs, for example.

Note, however, that a Klemperer (1990b) discusses there is not a unique way of defining a boom in this kind of model. A boom may imply more consumers in the first period and in this case countercyclical behaviour results. A similar result is obtained if we regard a boom in demand as an increase in the permanent income of the consumers in the market, or if firms are credit constrained and apply profit smoothing pricing strategies. On the other hand, if a boom means that all consumers in the market buy more, then price-cost margins will move procyclically. Unfortunately, the empirical evidence here is mixed.

The same two period model can be used to illustrate an unexpected consequence of increases in real interest rates. An increase in interest rates implies a

¹³ If consumers are myopic then $\partial \sigma_i / \partial p_1^i$ is independent of δ . When consumers are rational this is no longer true (Klemperer (1987b)). In particular, $\partial \sigma_i / \partial p_1^i$ is decreasing in δ . However, our analysis remains valid if $\partial \sigma_i / \partial p_1^i < \delta (\partial (\partial \sigma_i / \partial p_1^i) / \partial \delta)$.

reduction in the discount factor $\delta = e^{-r}$, where r denotes the real interest rate. Then $\partial \Pi^i / \partial p^i$ will decrease and in consequence, current prices will rise. Therefore, increases in real interest rates may raise inflation in the short run. Notice that we do not explicitly model money supply so that broader conclusions are difficult.

A second implication of switching costs is related to international oligopoly. Froot-Klemperer (1989) look at the degree of exchange rate pass-through when market share is a relevant variable. In particular, they investigate the pass-through from exchange rates to import prices when firms' future demands depend on current market shares. Foreign firms may either decrease or increase their prices when the home currency depreciates temporarily. They also show that import prices may be more sensitive to expected future than to current exchange rates. Each firm maximizes total discounted profits as before, given the price of the rival. Profits are given by

$$\begin{aligned} \Pi^d &= \Pi_1^d(p^d, p^f) + \delta^d \Pi_2^d(\sigma^d(p^d, p^f) e_2) \\ \text{and } \Pi^f &= \Pi_1^f(p^f, p^d) e_1 + \delta^f \Pi_2^f(\sigma^f(p^f, p^d) e_2) e_2, \end{aligned} \quad [11]$$

where Π^d and Π^f are the total discounted profits of the domestic firm, d , and a foreign firm, f , competing in the domestic market. δ^d, δ^f are the respective discount factors, where δ^i is inversely related to the real interest rate in the i th country, r^i .

Uncovered interest parity implies:

$$\delta^d = (e_2/e_1) \delta^f, \quad [12]$$

where e_t is the exchange rate at t , measured in units of domestic currency per unit of foreign one. It is easy to obtain p^i as a function of the costs of the foreign firm in period one and two, c_1^f, c_2^f , respectively and of the discount factor δ . Finally the effects on prices of a change in the exchange rate are shown by

$$\begin{aligned} dp^i/de_1 &= -c_1^f(\partial p^i/\partial c_1^f) - \delta(\partial p^i/\partial \delta) \\ dp^i/de_2 &= -c_2^f(\partial p^i/\partial c_2^f) + \delta(\partial p^i/\partial \delta) \end{aligned} \quad [13]$$

where $e_1 = \log e_1, e_2 = \log e_2$.

The first term in each of these two equations is called the *cost effect* and it is shown to be negative ($-c_1^f(\partial p^i/\partial c_1^f) < 0$). The second term is called the *real interest rate effect* and it is positive ($-\delta(\partial p^i/\partial \delta) > 0$). Hence, the final effect on import prices, p^f , of a temporary appreciation of the exchange rate, (an increase in e_1), is ambiguous. It is therefore not difficult to find examples where the real interest rate effect dominates, generating a perverse pass-through effect. Notice that a permanent appreciation always produces the expected results, that is, a decrease in import prices. The effect of a permanent change ($de_1 = de_2$) is simply

$$dp^i/de_1 + dp^i/de_2 = -c_1^f(\partial p^i/\partial c_1^f) - c_2^f(\partial p^i/\partial c_2^f) < 0 \quad [14]$$

The reason why a temporary appreciation (depreciation) may induce increases (decreases) in imports prices is because the change in the exchange rate modifies the relative value of future profits with respect to current ones. Therefore, the foreign firm reduces its investment in market share increasing prices. Similar conclusions can be obtained if we use quantity rather than price competition or, in a multiperiod model (Beggs-Klemperer (1989)). There are some points, however that are not that clear in these models. In particular, there is not an explicit mechanism of expectations formation about future exchange rates. In addition, the conclusions of Froot-Klemperer depend on the continuity of the payoff functions and on the sign of $\partial \Pi_2^i / \partial \sigma_i$, which (as we noticed above) is not uniquely determined. Finally, if we consider the case of an expected change instead of a current one, the conventional results are reinforced by the presence of switching costs.

4. Concluding remarks

Switching costs have been shown to be relevant for a better understanding of competition for market share. In particular, they provide alternative explanations for price wars and sales, for the observed behaviour of managers, for macroeconomic phenomena like exchange rate pass-through, etc.

The literature on switching costs is still growing in very different directions. This survey is not a completely comprehensive exercise. Topics such as the design of optimal contracts with switching costs (Farrell-Shapiro (1989)); choices of product lines in multiproduct oligopoly with switching costs (Klemperer (1990a)); auctions of procurement contracts with switching costs; or the study of merger profitability (Padilla (1990c)) have not been included in this survey. They constitute interesting and promising lines for further research.

References

- Banerjee, A and Summers, L. (1987): «On Frequent Flyer Programs and other Loyalty-inducing Economic Arrangements», Harvard Institute of Economic Research, D. P. 1337.
- Beggs, A. (1989): «A Note on Switching costs and Technology Choice», *Journal of Industrial Economics*, no. 37, pp. 437-440.
- Beggs, A. and Klemperer, P. (1989): «Multiperiod Competition with Switching costs», Nuffield College, D.P. 45, Oxford.
- Budd, C.; Harris, C. J. and Vickers, J. (1990): «A Model of the Evolution of Duopoly: Does the asymmetry between firms tend to increase or decrease?», mimeo, Oxford.
- Bulow, J.; Geneakoplos and Klemperer, P. (1985): «Multimarket Oligopoly: Strategic Substitutes and Complements», *Journal of Political Economy*, no. 93, pp. 488-510.
- Caminal, R. and Matutes, C. (1989): «On Precommitment and Competition: Endogenous Switching Costs in a Duopoly Model», Institut d'Anàlisi Econòmica, Discussion Paper UAB, Barcelona.
- Carlton, D. (1989): «The Theory and Facts of How Markets Clear: Is Industrial Organization Valuable for Understanding Macroeconomics», in *Handbook of Industrial Organization*, Scmalense, R. and Willig, R. (eds.) North Holland.

- Dasgupta, P. and Maskin, E. (1986a): «The Existence of Equilibrium in Discontinuous Economic Games I: Theory», *Review of Economic Studies*, pp. 1-27.
- Dasgupta, P. and Maskin, E. (1986b): «The Existence of Equilibrium in Discontinuous Economic Games II: Applications», *Review of Economic Studies*, pp. 28-42.
- Farrell, J. (1986): «A Note on Inertia in Market Share», *Economic Letters*, no. 21, pp. 73-75.
- Farrell, J. and Shapiro, C. (1988): «Dynamic Competition with Switching Costs», *Rand Journal of Economics*, no. 19, pp. 123-127.
- Farrell, J. and Shapiro, C. (1989): «Optimal Contracts with Lock-in», *American Economic Review*, no. 79, 1, pp. 55-68.
- Friedman, J. (1986): *Game Theory with Applications to Economics*, Oxford University Press, Oxford.
- Froot, K. and Klemperer, P. (1989): «Exchange Rate Pass-Through when Market Share Matters», *American Economic Review*, no. 79, 3, pp. 637-654.
- Gallini, N. and Karp, L. (1989): «Sales and Consumer Lock-in», *Económica*, no. 56, pp. 279-294.
- Green, J. and Scotchmer, S. (1986): «Price Competition with a Distribution of Switching Costs and Reservation Prices», Harvard Institute of Economic Research, D.P. 1260.
- Harris, C. J. (1988): «Dynamic Competition for Market Share: an Undiscounted Model», Nuffield College. D.P. 35, Oxford.
- Klemperer, P. (1987a): «Markets with Consumer Switching Costs», *Quarterly Journal of Economics*, pp. 375-394.
- Klemperer, P. (1987b): «The Competitiveness of Markets with Switching Costs», *Rand Journal of Economics*, no. 18, pp. 138-151.
- Klemperer, P. (1987c): «Entry Deterrence in Markets with Consumer Switching Costs», *Economic Journal*, Conference Supplement, pp. 9-117.
- Klemperer, P. (1988): «Welfare Effects of Entry into Markets with Switching Costs», *Journal of Industrial Economics*, no. 37, pp. 99-117.
- Klemperer, P. (1989): «Price Wars Caused by Switching Costs», *Review of Economic Studies*, no. 56, pp. 405-420.
- Klemperer, P. (1990a): «Product Line Competition and Shopping Costs: Why Firms May Choose to Compete Head-to-Head», University of Oxford, mimeo.
- Klemperer, P. (1990b): «Markets with Consumer Switching Costs: an overview», *Review of Economic Studies*, Lecture, delivered at the Royal Economic Society Conference, March 1990.
- Klemperer, P. and Regibeau, P. (1988): «Product Compatibility, Endogenous Switching Costs and Integration», Kellogg School of Management, mimeo.
- Kreps, D. M. and Scheinkman, J. (1983): «Quantity Precommitment and Bertrand Competition yield Cournot Outcomes», *Bell Journal of Economics*, pp. 326-337.
- Maskin, E. and Tirole, J. (1988): «A Theory of Dynamic Oligopoly II: Price Competition, Kinked Curves and Edgeworth Cycles», *Econometrica*, pp. 571-599.
- Matutes, C. (1988): «Teoría de Juegos y Estrategias Tecnológicas», *Cuadernos Económicos ICE*, no. 40, 3, pp. 229-252.
- Padilla, A. J. (1990a): «Mixed Pricing Oligopoly in Markets with Consumer Switching Costs», Unpublished M. Phil Thesis. St. Antony's College, Oxford.
- Padilla, A. J. (1990b): «Mixed Pricing Oligopoly in Markets with Consumer Switching Costs», Nuffield College, Oxford, mimeo.
- Padilla, A. J. (1990c): «Horizontal Mergers: Bertrand Competition and Switching costs yield Cournot Outcomes», Nuffield College, Oxford, mimeo.
- Rasmussen, E. (1989): *Games and Information: An Introduction to Game Theory*, Blackwell, Oxford.
- Reinganum, J. (1979): «A Simple Model of Equilibrium Price Dispersion», *Journal of Political Economy*, no. 87, pp. 851-858.

- Reinganum, J. (1985): «Innovation and Industry Equilibrium», *Quarterly Journal of Economics*, no. 100, pp. 81-99.
- Rosenthal, R. W. (1982): «A Dynamic Model of Duopoly with Customer Loyalties», *Journal of Economic Theory*, no. 27, pp. 69-76.

Abstract

In this paper our purpose is to overview the literature on consumer switching costs focusing on the distinct implications of different types of switching costs, namely exogenous and endogenous costs, on the functioning of oligopolistic markets. In particular, we will analyze how the degree of competitiveness is altered by the existence of switching costs. We will also focus on other recent developments in this literature, such as the macroeconomic implications of markets with consumer switching costs and the evolution of industrial structure in this type of market. We conclude with some suggestions for further research.

Recepción del original, noviembre de 1990
Versión final, junio de 1991