# Consumer heterogeneity and pricing in a duopoly with switching costs*

Tommy Staahl Gabrielsen and Steinar Vagstad

Department of Economics, University of Bergen,

Fosswinckelsgate 6, N-5007 Bergen, Norway

April 13, 2000

**Abstract**

It is well-known that switching costs may facilitate monopoly pricing in a market with price competition between two suppliers of a homogenous good, provided the switching cost is above some critical level. We show that introducing consumer heterogeneity tends to increase the critical switching cost and thereby reduce the stability of the collusive outcome. A testable implication is that widespread price discrimination should go hand in hand with efforts to create switching costs.

## 1 Introduction

A wave of privatization and deregulation has rolled over the world in recent years. The old national monopolies – be it railroad services, airlines, telecommunication

*A preliminary version of this paper was presented at the Nordic Workshop in Industrial Organization (NORIO II) in Bergen, May 1999, and the Conference in Game Theory, Experimental Economics and Applications in Valencia, June 1999. Parts of the paper was written while we were visiting the University of Pompeu Fabra, Barcelona, whose hospitality is gratefully acknowledged. This research has been sponsored by Telenor, through the Foundation for Research in Economics and Business Administration, and the Meltzer Foundation.

or electricity provision or distribution – now typically have been forced to share their markets with one or more entrants. This give rise to numerous interesting issues of competition in general and pricing behavior in particular. A particular problem facing the firms is how to escape the Bertrand paradox: they compete in markets for more or less homogeneous goods, with prices as the main strategic variable. In our view, the most compelling solution to the paradox is the existence of switching costs: the fact that even if consumers don't care about which product they start to buy, there may be costs associated with switching suppliers.[1] These costs dampen competition in mature markets in a variety of settings, as shown by Paul Klemperer in numerous articles (see his 1995 survey). Recent efforts to raise barriers for consumers who might consider to switch supplier must be seen in light of this theory.[2]

Another characteristic of some of the industries in question – telecommunications in particular – is the degree of sophistication in pricing behavior: the typical tariff is non-linear, and normally consumers are offered the choice between a variety of schemes, with the purpose of price discriminating between heterogeneous consumers.[3] The aim of the present paper is to study the interplay between switching costs and pricing behavior in a market where consumers differ in some respect relevant for price discrimination. In addition to the already mentioned literature on switching costs and non-linear pricing, there are also many contributions studying

[1] Other proposed solutions include product differentiation (physically or informationally) and tacit collusion, as laid out in any modern treatments of Industrial Organization, e.g. Tirole (1988).

[2] Examples of such barriers include frequent-flyer's programs used by airlines (for some Norwegian evidence, see Risvold 2000) and subsidizing new mobile phone customers' purchase of the phone if they sign up for a minimum period of one year (for some Norwegian evidence, see Seime 1999).

[3] Such pricing behavior reflects the fact that the products in question are typically non-transferable services, effectively limiting the possibilities for arbitrage. Moreover, the firms have limited information about different consumers' tastes, or they are (explicitly or implicitly) restricted from exploiting the little information they have about tastes in different sub-markets (one could for instance imagine prices that differ according to gender, age and location), leaving second-degree price discrimination as the only viable price discrimination option. (See Wilson, 1993, for a survey of non-linear pricing.)

non-linear pricing in more or less competitive settings. What these contributions have in common, however, is that they model other sources of market power than switching costs. Wilson (1993, part 12.3) consider Cournot competition, while Stole (1995), Armstrong and Vickers (1999) and Rochet and Stole (1999) are examples of studies based on the assumption that products are differentiated.

To our knowledge, none has studied the effects of switching costs in a homogeneous-good duopoly with switching costs and heterogeneous consumers. We conduct the analysis within a model allowing any kind of non-linear pricing. Our main result is that heterogeneity tend to reduce collusive stability, in the sense that collusive pricing can be sustained with homogeneous consumers but not with heterogeneous. One testable implication of this phenomenon is that one should expect consumer heterogeneity and sophisticated pricing to go hand in hand with efforts to raise barriers for consumers who may want to switch supplier, a hypothesis that seems to fit for instance the markets for mobile telephony.

We have also performed essentially the same analysis for some other types of pricing behavior. It turns out that our main result is rather robust with respect to pricing behavior — whatever pricing assumption, heterogeneity tend to make it more difficult to sustain monopoly pricing. Perhaps somewhat surprising is that the result also applies to the case of linear pricing, revealing that the main result does not hinge upon nonlinear pricing but rather on the mere heterogeneity.

The paper proceeds as follows. In the next section we present our basic model, we define the key notion of *critical switching cost*, and we perform some preliminary analysis. In Section 3 we study how consumer heterogeneity affects the critical switching costs under the assumption of non-linear pricing. A brief discussion of how other pricing schedules affect the results is found in Section 4, together with a discussion of different matters left out of the main analysis. Some concluding remarks are gathered in Section 5.

## 2 The model and a benchmark

Consider two firms setting prices in a market with two kinds of consumers — $H$ ("high" demand) and $L$ ("low" demand). The two firms offer functionally identical products, but each consumer has already bought from one of the firms, and if a consumer wants to switch to the other supplier, switching costs are incurred. We assume that all consumers have identical positive switching costs denoted $s$.[4] In particular, the costs of switching does not depend on a consumer's demand volume.[5]

Next, we assume that the firms have access to full non-linear pricing. A contract is a payment-quantity pair $(q_i, T_i)$. (In the general case a menu of contracts is described by a payment function of quantity demanded; $T_i = T(q_i)$.) Here we have only two types of consumers, implying that each supplier offers the consumers a choice between two contracts: $(q_L, T_L)$ intended for the low-demand consumers, and $(q_H, T_H)$ for the high-demand consumers.

Consumer preferences over contracts are described by utility functions that are linear in money and quadratic in quantity:

$$u(\theta, q, T) = \theta q - \frac{1}{2}q^2 - T, \quad \text{for} \quad \theta \in \{L, H\} \tag{1}$$

where $\theta$ is the consumer's "type", $q$ is demand volume and $T$ is monetary payment for the good in question. These preferences give rise to individual demand functions that are linear in prices with no income effects on demand:

$$q = q(p, \theta) = \theta - p, \quad \text{for } \theta \in \{L, H\} \tag{2}$$

Moreover, firms are assumed to be symmetric both as regards costs and customer

[4]This implies that the only candidate for equilibrium in pure strategies entails monopoly pricing (as specified below).

[5]This is obviously not the only way to model switching costs. Consider switching mobile telephone operator. This would entail some fixed costs, for instance the effort of contacting the operators and make them do what you want, possible penalties for terminating the relationship with your existing operator, and costs of opening a new relationship. Typically there are also volume-dependent switching costs, for instance the costs attached to lack of number portability which is presumably a larger problem for a pizza chain than from a typical private consumer, but may be substantial even for private consumers.

bases (from an at the time being unmodelled first period).[6] In particular, each firm has a market share of 50% within each market segment (that is, for each type of consumer). To obtain closed-form solutions to the pricing problem we need marginal costs to be constant, normalized to zero. Finally, to simplify notation we set $H = 1$ while $L \in (0, 1)$. This is without loss of generality as only their relative magnitudes are of importance.

When we in subsequent sections describe a market with heterogeneous consumers we are going to compare results with a benchmark with homogeneous consumers. We will therefore perform some preliminary analysis assuming that all consumers are identical, with demand parameter $\theta$. Then it is well-known that a profit-maximizing monopolist will offer a contract that maximizes social surplus, and by setting an appropriate fixed fee he can convert social surplus into profits. Social surplus ($S$) equals consumers' utility plus profits, and is given by

$$S = u + \pi = \theta q - \frac{1}{2}q^2 - T + T = \theta q - \frac{1}{2}q^2 \tag{3}$$

Social surplus is maximized by setting $q = \theta$, and this surplus is shifted over to the firm by setting $T = \frac{1}{2}\theta^2$. (Note that this solution can be implemented by a two-part tariff $T = F + pq$, where $F = \frac{1}{2}\theta^2$ and $p = 0$.)

Here we do not have a monopoly, though. However, as long as all consumers have positive switching costs, Klemperer (1987) have argued — in a framework of linear pricing — that if there is a pricing equilibrium in pure strategies, this equilibrium must entail monopoly pricing. The argument goes as follows. Let $p_M$ denote the monopoly price. At any lower common price, each firm has an incentive to slightly increase its price, which more fully exploits its own customers without losing any to its competitor. Note that even small switching costs suffices to make the (possible) equilibrium switch from competitive pricing to monopoly pricing.

It should be clear that the logic of small deviations applies equally well to situations involving non-linear pricing: even if firm $A$ uses linear prices, it would pay for firm $B$ to price non-linearly, for instance using two-part tariffs. With linear pricing there is one single instrument — the price — serving two different purposes: effi-

[6] Competition for market shares at an earlier stage is left to Section 4.

ciency and extraction of consumers' surplus. The virtue of two-part tariffs is that they separate these two aims: efficiency is achieved by marginal cost pricing, and consumers' surplus is extracted by the fixed term.

However, the proposed equilibrium may be vulnerable to non-marginal price changes: it is still the case that a sufficiently large price cut will make one firm corner the market, and if the switching costs are too small, cornering the market becomes so attractive that monopoly pricing is not an equilibrium either – implying that there is no equilibrium in pure strategies at all.[7] In this respect the magnitude of the switching cost is important.

To be more precise, to attract one's competitor's customers, one must offer them a price cut that can compensate them for their costs of switching supplier. When all consumers have demand parameter $\theta$, we have seen that monopoly pricing entails $(q, T) = (\theta, \frac{1}{2}\theta^2)$. In order to capture the rival's customers, one will have to undercut by an amount equal to their switching costs, i.e., one will have to set $T \leq \frac{1}{2}\theta^2 - s$, yielding profit of $2T$.[8] Such an undercutting is profitable iff

$$2\left(\frac{1}{2}\theta^2 - s\right) > \frac{1}{2}\theta^2 \tag{4}$$

This is a more precise expression for our statement above that for monopoly pricing to be an equilibrium, the switching cost must be large enough. Solving this inequality for $s$ yields the following Lemma:

**Lemma 1** *(Homogeneous consumers) With homogeneous consumers who have demand parameter $\theta$, an equilibrium in pure strategies exists iff $s \geq s^*(\theta) \equiv \frac{1}{4}\theta^2$. If such an equilibrium exists, it is unique and entails $(q, T) = (\theta, \frac{1}{2}\theta^2)$.*

[7] There is always an equilibrium in mixed strategies, however, see Klemperer (1987). This equilibrium is rather complicated even in a model with linear pricing and homogeneous consumers, and it is beyond the scope of the present paper to analyze mixed-strategy equilibria of the current model.

[8] It is easily checked that the firm can not increase its profit by distorting the quantities, as this will only serve to reduce the value added without helping the suppliers to reap a larger fraction of it.

## 3 Pricing with heterogeneous consumers

Now we turn to situations in which there are both types of consumers ($H$ and $L$). To simplify the exposition, suppose there are four consumers, among which there is one high-demand consumer and one low-demand consumer "belonging" to each of the firms (the results generalize easily to other symmetric structures, and with some effort also to cases of asymmetric customer bases). Depending on the parameters, qualitatively different situations may occur. Demand from $L$-type consumers may be so low that a monopolist may choose to sell only to high-demand consumers, and a firm considering to deviate from monopoly pricing may find it attractive to set his prices to attract the competitor's high-demand consumers; his low-demand consumers; or all of his consumers. All these cases are considered below.

**Monopoly pricing.** Suppose first that the monopolist wants both types of consumers to buy (this is not necessarily the case). Then his problem is to design a pair of contracts to maximize income — $\max_{\{(T_i,q_i)\}} T_H + T_L$ — subject to standard participation and incentive constraints:

$$q_H - \frac{1}{2}q_H^2 - T_H \geq 0 \tag{5}$$

$$Lq_L - \frac{1}{2}q_L^2 - T_L \geq 0 \tag{6}$$

$$q_H - \frac{1}{2}q_H^2 - T_H \geq q_L - \frac{1}{2}q_L^2 - T_L \tag{7}$$

$$Lq_L - \frac{1}{2}q_L^2 - T_L \geq Lq_H - \frac{1}{2}q_H^2 - T_H \tag{8}$$

Suppose that only low-demand consumers' participation constraint and high-demand consumers' incentive constraint bind (it can easily be verified that this is indeed true for the optimal mechanism). We see that the firm has not preferences over quantities — $q_H$ and $q_L$ only affect the objective function indirectly, through the relevant constraints (6) and (7). First we note that $q_H$ enters (7) only, and it should therefore be set to soften this constraint as much as possible, implying $q_H = 1$ (this amounts to setting price equal to marginal costs for high-demand consumers). Assuming the constraints (6) and (7) bind, by inserting them into the objective function of the monopolist the maximization problem is reduced to

$\max_{q_L}\left(\left(\frac{1}{2}-q_L+Lq_L\right)+\left(Lq_L-\frac{1}{2}q_L^2\right)\right)$, with no constraints. Its solution is given by $q_L = 2L-1$ (provided $L \geq \frac{1}{2}$) and the optimal charges satisfy

$$\begin{aligned} T_L &= L(2L-1)-\frac{1}{2}(2L-1)^2 = L-\frac{1}{2} \\ T_H &= \frac{1}{2}-(1-L)(2L-1) \end{aligned}$$

Consequently, maximum profit is given by

$$\pi = T_H + T_L = \left(\frac{1}{2}-(1-L)(2L-1)\right)+\left(L-\frac{1}{2}\right) = 1-2L+2L^2$$

Selling to low-demand consumers is costly in terms of giving up consumers' surplus to high-demand consumers. If their demand is sufficiently low — $L \leq \frac{1}{2}$, to be precise — it pays to neglect them altogether, setting $q_L = T_L = 0$. The following Lemma summarizes the above discussion:

**Lemma 2** *(Monopoly pricing) If $L \leq \frac{1}{2}$, then monopoly pricing entails $(q_L, T_L) = (0,0)$ and $(q_H, T_H) = (1, \frac{1}{2})$, with monopoly profit of $\pi_M = \frac{1}{2}$. If, in contrast, $L > \frac{1}{2}$, then $(q_L, T_L) = (2L-1, L-\frac{1}{2})$ and $(q_H, T_H) = (1, \frac{1}{2}-(2L-1)(1-L))$, yielding monopoly profit of $\pi_M = 1-2L+2L^2$.*

**Optimal undercutting.** To find the critical switching costs needed to sustain monopoly pricing, we now derive the optimal undercutting strategies for the firms. To attract one's competitor's customers, one must offer them a price cut that can compensate them for having to bear their switching costs. Since consumers are heterogeneous, there are two different ways to undercut the rival: one can either go for his high-demand consumers (to be dubbed strategy "High") or for all the competitor's consumers (strategy "All").[9] In what follows we will describe each of these strategies in detail, for different values of $L$. We start with the simpler cases when $L \leq \frac{1}{2}$, implying that only high-demand consumers are served under

[9] In principle there is also a third strategy: going for the rival's low-demand consumers only. It is easily checked that this is never a viable strategy as long as we maintain our assumption of all consumers having the same switching costs. It could change if we allowed low-demand consumers to have substantially lower switching costs.

monopoly pricing, and thereafter we do the more complicated cases when $L > \frac{1}{2}$. For each case we calculate profit of an undercutting firm, and by comparing with profits under monopoly pricing we derive conditions for stability of the pure-strategy monopoly pricing equilibrium. The purpose of this exercise is to compare these stability conditions with those obtained with homogeneous consumers in Lemma 1, in order to find out the implications of heterogeneity for stability.

*Strategy Undercutting "High".* We start with situations in which only high-demand consumers are served under monopoly pricing, i.e., $L \in \left(0, \frac{1}{2}\right]$. Strategy "High" undercutting entails paying the switching costs of the rival's $H$-type. In monopoly, selling to low-demand consumers were unattractive because the big difference in demand between the different types of consumers. When undercutting, however, it may be worthwhile to start to sell to own low demand consumers. Suppose that the undercutting firm does exactly this when undercutting. The undercutting firm offers a menu of contracts $(q_H, T_H)$ to its own and the rival´s high-demand consumers and $(q_L, T_L)$ to its own low-demand consumer to maximize

$$\pi = 2T_H + T_L$$

subject to a new participation constraint for the rival´s high-demand consumer (he must earn at least $s$ to switch):

$$u_H = q_H - \frac{1}{2}q_H^2 - T_H \geq s \tag{9}$$

and (6), (7) and (8).

We then have the following result:

**Lemma 3** *For $L \in \left(0, \frac{1}{2}\right]$, the switching cost needed to block undercutting the rival's high-demand consumers is*

$$s \geq s^H \equiv \frac{1}{4}\left(1 + L^2\right).$$

**Proof.** See the appendix. ■

We see from Lemma 3 that the critical switching cost is increasing in $L$ and it is always higher than the critical switching cost for a homogeneous population of high-demand consumers (as stated in Lemma 1). The reason for this is that when undercutting the rival's high-demand customers it becomes attractive to sell to own low-demand customers, and this makes undercutting more tempting. Hence, switching costs must be higher to make such undercutting unprofitable.

In contrast, when $L \in (\frac{1}{2}, 1)$ both types of consumers buy under monopoly pricing. In this equilibrium the $L$-type earns zero and the $H$-type earns a rent equal to $u_H = (2L-1)(1-L)$. The undercutting firm now maximizes $\pi = 2T_H + T_L$ with respect to (6), (7), (8) and the high-demand consumer's new participation constraint:

$$u_H = q_H - \frac{1}{2}q_H^2 - T_H \geq \frac{1}{2} - \left(\frac{1}{2} - (2L-1)(1-L)\right) + s \tag{10}$$

This yields the following result:

**Lemma 4** *For $L \in \left(\frac{1}{2}, 1\right)$, the switching cost needed to block undercutting the rival's high-demand consumers is*

$$s \geq s^H \equiv \begin{cases} 1 - 2L + \frac{5}{4}L^2 & \text{if} \quad L \in \left(\frac{1}{2}, \frac{2}{3}\right] \\ \left(1 - L + \frac{1}{2}\sqrt{8L - 2L^2 - 4}\right)(1-L) & \text{if} \quad L \in \left(\frac{2}{3}, \frac{13}{11} - \frac{1}{11}\sqrt{26}\right] \\ \frac{17}{12} - \frac{17}{6}L + \frac{19}{12}L^2 & \text{if} \quad L \in \left(\frac{13}{11} - \frac{1}{11}\sqrt{26}, 1\right) \end{cases}$$

**Proof.** See the appendix. ■

We see from Lemma 4 that the critical switching cost for the interval $L \in \left(\frac{1}{2}, 1\right)$ is defined in three pieces, depending on which of the incentive and participations constraints that bind.

Having established the critical switching cost when undercutting the rival's high-demand customers, we must also check whether it might be better to undercut all the rival's customers. This sounds like a sensible way to undercut if $L$ is sufficiently close to one.

*Strategy Undercutting "All".* This strategy involves lowering the payment by $s$ for all customers, without affecting the issue of distorting quantities.[10] This tells us

[10] Technically, the undercutting firm maximizes $2(T_L + T_H)$ – that is, twice the monopoly profit – subject to a set of constraints that is identical to the one facing the monopolist, the only exception being that all consumers' reservation utilities shift upward by an amount $s$.

that it is no point in undercutting the rival's low-demand consumers if $L \leq \frac{1}{2}$. If $L > \frac{1}{2}$, we have the following:

**Lemma 5** *For $L \in \left(\frac{1}{2}, 1\right)$, the switching cost needed to block undercutting all the rival's customers is $s \geq s^A \equiv \frac{1}{4} - \frac{1}{2}L + \frac{1}{2}L^2$.*

**Proof.** See the appendix. ■

**Stability.** By comparing the critical costs for the two respective strategies, it turns out that it requires a larger switching cost to block undercutting of all consumers than to block undercutting of high-demand consumers only, as long as $L > \frac{14}{13} - \frac{1}{13}\sqrt{14}$, in which case the critical switching cost is given by Lemma 5. Since what matters for collusive stability is whether undercutting is profitable or not — exactly which kind of undercutting will take place is of less interest — it is the highest of the critical switching costs that is of our interest: $s^* = \max\{s^H, s^A\}$. Then by summing up the information from Lemmas 3-5, we have that undercutting with either strategy is blocked when $s \geq s^*$, where

$$s^* = \begin{cases} \frac{1}{4} + \frac{1}{4}L^2 & \text{if} \quad 0 \leq L < \frac{1}{2} \\ 1 - 2L + \frac{5}{4}L^2 & \text{if} \quad \frac{1}{2} \leq L \leq \frac{2}{3} \\ \left(1 - L + \frac{1}{2}\sqrt{8L - 2L^2 - 4}\right)(1 - L) & \text{if} \quad \frac{2}{3} < L \leq \frac{13}{11} - \frac{1}{11}\sqrt{26} \\ \frac{17}{12} - \frac{17}{6}L + \frac{19}{12}L^2 & \text{if} \quad \frac{13}{11} - \frac{1}{11}\sqrt{26} < L \leq \frac{14}{13} - \frac{1}{13}\sqrt{14} \\ \frac{1}{4} - \frac{1}{2}L + \frac{1}{2}L^2 & \text{if} \quad \frac{14}{13} - \frac{1}{13}\sqrt{14} < L \leq 1 \end{cases} \tag{11}$$

In Figure 1 below we have plotted $s^*$ together with the critical switching costs for homogeneous populations of high-demand consumers ($s^*(1) = \frac{1}{4}$), low-demand consumers ($s^*(L) = \frac{1}{4}L^2$) and "average" consumers (that is, a hypothetical situation in which all consumers have demand parameter $\frac{1+L}{2}$, yielding $s^*(\frac{1+L}{2})$).

From Figure 1 we see that $s^*$ (the dotted line) is non-monotonic. Starting out from $L = 0$, $s^*$ increases until $L = \frac{1}{2}$, then decreases until $L = \frac{14}{13} - \frac{1}{13}\sqrt{14}$, whereafter it increases again. The turning points are when the monopolist starts selling to low-demand consumers ($L = \frac{1}{2}$) and when the undercutting firm changes undercutting strategy from undercutting only the rival's high-demand consumers to going after all

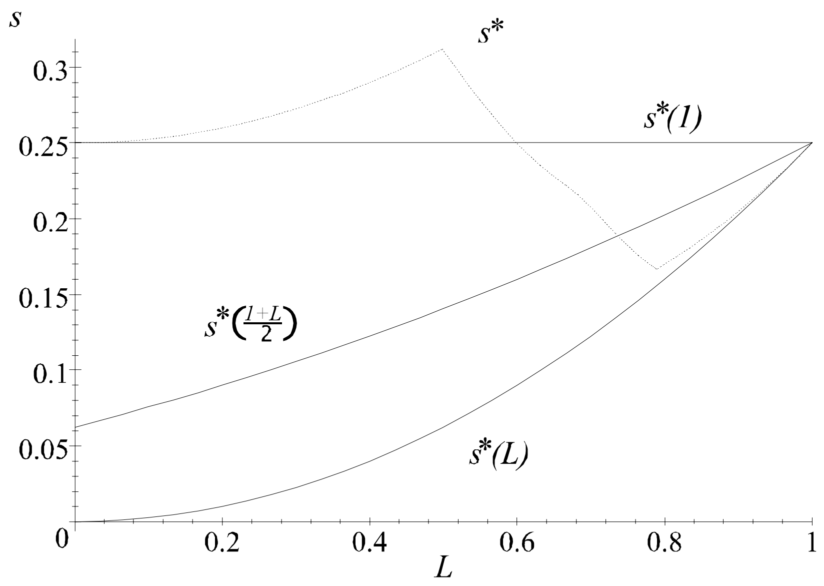


Figure 1: Critical switching costs

consumers. The increase in $s^*$ for $L \in (0, \frac{1}{2})$ is due to the increase in demand from an undercutting firm's own low-demand customers. Since $L$ is low, these consumers are inactive in monopoly, but when undercutting the rival's high-demand customer it is worthwhile to include them. This makes the temptation to undercut bigger, and more so the more surplus that can be extracted from the low-demand customer, hence the increase in $s^*$. When $L > \frac{1}{2}$ the low-demand consumers already buy under monopoly. At first in this interval, the low-demand consumers are given highly distortive contracts to prevent the high-demand consumers from mimicking a low-demand consumer. When undercutting it is possible to offer efficient contracts to all consumers which in itself is a gain, but the gain is lower the more efficient the contract is from the outset, hence $s^*$ decreases. As $L$ increases, the undercutting firm starts worrying about the low-demand consumer mimicking a high-demand consumer. To avoid this, the firm has to offer the high-demand consumer a distortive contract, and at some point also leave rent to the low-demand consumer. Finally, for sufficiently high $L$ the undercutting firm will find it better to undercut all the rival's customers, and then of course the profit from undercutting is increasing in $L$, and so is the critical switching cost needed to block undercutting.

It is noteworthy that for low values of $L$, i.e., $L \in (0, \frac{3}{5})$, consumer heterogeneity implies instability in a strong sense, meaning that it now takes a higher switching cost to sustain monopoly pricing than if all consumers were of the high-demand type. At the other extreme for high $L$ we see that consumer heterogeneity instead implies stability, but only in a weak sense. For high values of $L$ it takes a lower switching cost to sustain monopoly pricing than if all consumers were an average of high and low-demand consumers.[11]

To summarize the first part, we have:

**Proposition 1** *With homogeneous consumers, the critical switching cost* $s^* \leq s^*(1) = \frac{1}{4}$. *With heterogeneous consumers, however, if* $L \in \left(0, \frac{3}{5}\right)$ *then* $s^* > s^*(1)$, *and if* $L \in \left(\frac{3}{5}, 1\right)$ *then* $s^*(L) < s^* < s^*(1)$.

# 4 Discussion

In the analysis above we have assumed that the firms are rather sophisticated price-setters — they use general non-linear pricing schemes. We have also considered less sophisticated pricing rules and investigated how such rules affect the results. In what follows we will present the main results from one particular pricing assumption: linear pricing. Clearly, with linear pricing there is no scope for price discrimination. This in turn means that the result derived with linear pricing cannot stem from price discrimination, but must be interpreted as pure effects of heterogeneity. Intuitively, since more sophisticated pricing presumably extracts more surplus from the consumers, we would expect that more sophisticated pricing creates more instability, and we will see that this intuition is largely correct.

[11] At first glance this might seem surprising. However, the reason is that when undercutting a homogeneous population of an average of high and low-demand consumers all rent apart from the switching costs can be extracted. However, heterogeneity implies that in addition to leaving rent to cover switching costs some information rent has to be left to the high demand consumers. Hence, the profit that can be extracted from a population of heterogeneous consumers is less that the profit that can be extracted from a homogeneous population of average consumers, and therefore the temptation to undercut is lower under heterogeneity than under homogeneity. As a consequence, heterogeneity yields a lower critical switching cost.

As in the previous section, the task is to compare the critical switching costs with homogeneous versus heterogeneous consumers. As in the previous section, with heterogeneous consumers the monopolist will serve all consumers if the low-demand consumers' demand is not too low, and concentrate on the high-demand consumers otherwise (however, the two cases occur for different parameter sets under the two pricing schemes).

Performing the analysis yields the following conclusions:[12] First, also with linear pricing we find that heterogeneity implies loss of stability, but only for parameters for which low-demand consumers are not served under monopoly pricing, that is, for low values of $L$. This result warrants some explanation. Also for higher values of $L$ the undercutting firm earns a benefit on increased sales to his low-demand consumers, but this benefit when undercutting is more than offset by the fact that the monopoly price is set to accommodate an average consumer, with demand parameter $\theta = \frac{1+L}{2}$. Undercutting such a low price is little tempting. Note that this latter effect was not present with non-linear pricing, and appears here only because the firms are restricted to one single instrument — the price — which drops discontinuously when the firms start selling to low-demand consumers.

Second, for most levels of $L$, linear pricing implies lower critical switching costs compared to non-linear pricing, suggesting that sophisticated pricing implies loss of collusive stability. This is illustrated in Figure 2 below.

In the figure the solid line represents the critical switching cost with fully non-linear prices, while the dotted line represents critical switching costs under linear pricing. We see that the solid curve is above the dotted curve for most parameters, indicating that sophisticated pricing tend to reduce stability of monopoly pricing. However, the reverse is also possible, mainly because switching to "undercutting all" occurs for lower values of $L$ when the firms use non-linear pricing than when they use linear pricing.

Up to now we have taken the distribution of market shares for granted. Klemperer (1987, 1995) shows in his model of homogeneous consumers that second period lock-in that is the result of switching costs may intensify competition in the first

[12] Details can be obtained from the authors upon request.

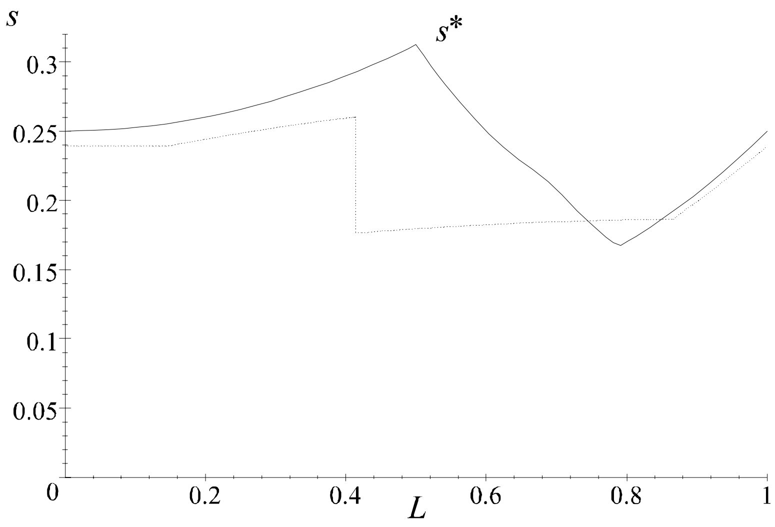


Figure 2: Critical switching costs for linear and nonlinear pricing

period. This effect will clearly be present here as well. But consumer heterogeneity adds another dimension to the first-period problem: in addition to setting low prices, the firms may affect the composition of its clientele by their choice of first-period tariffs. If — as the above analysis suggests — heterogeneity reduces the scope for setting monopoly prices later on, the firms have (at least collectively) a reason to restrict heterogeneity. This can be done in several ways: the firms may abstain from selling to low-demand consumers in the first period, or — perhaps more realistically — they may specialize in the first period: one firm sells to the high-demand consumers while the other sells to the low-demand consumers. However, a full-fledged analysis of these matters is beyond the scope of this paper.

Clearly, collusive stability depends on how the switching costs are distributed, but some alternative formulations would lead to similar conclusions as the present analysis. In particular, letting the switching costs be continuously distributed on an interval not including zero would basically complicate the analysis without generating much new insights. More dramatic changes would follow if the switching costs were distributed on an interval including zero, as this will make the static equilibrium involve some competition.[13] Even more dramatic changes would obtain

[13]E.g. prices lower than the monopoly price but higher than marginal costs in the case of linear

if we allowed a positive fraction of consumers to have zero switching costs, as this will basically take us back to the Bertrand paradox.

If we were to include switching cost heterogeneity we would also have to address how such heterogeneity blends with the already modelled demand parameter heterogeneity. Stochastic independence is clearly the easiest assumption to work with, but it might be more realistic to assume that high-demand consumers have higher switching costs than low-demand consumers, at least in a stochastic sense. Also this is beyond the scope of the present paper.

Our analysis has been conducted using a two-point distribution of demand parameters. This is obviously a simplification that may affect results, but a simplification that seems necessary as a first step. An important issue for further research is to investigate whether similar results can be obtained in a model allowing for continuously distributed consumers. In future work we would also like to investigate whether our choice of utility function affects results. Our analysis is performed with linear and parallel demand for high- and low-demand consumers. The linearity assumption is a simplification that enables us to derive explicit solutions for the critical switching cost and to make a detailed comparison over the different cases considered. But also other utility functions give rise to linear (but not parallel) demand functions, e.g. the following:

$$u(\theta, q, T) = \theta\left(q - \frac{1}{2}q^2\right) - T$$

# 5 Concluding remarks

There are different ways to escape the Bertrand paradox threatening the profit of price-setting firms competing in a market for homogeneous products. We have studied one such possibility — the creation of consumer switching costs — in a market with heterogeneous consumers. We have argued that this market structure as well as this particular strategy to reduce competition fits the telecommunications industry in recent years. We have seen that heterogeneity tend to reduce collusive stabil-

pricing. See Klemperer (1987) for a discussion.

ity, with the immediate implication that the more heterogeneity, the higher efforts to raise barriers for consumers who may want to switch supplier. Heterogeneity is however not immediately observable, but it should be reasonable to assume that heterogeneity is positively correlated to the spread of tariffs offered, and then we have a testable implication: sophisticated pricing should go hand in hand with efforts to create consumers switching costs, a hypothesis that also seems to fit modern telecommunications markets, and the market for mobile telephony in particular.

We have performed some analysis for other types of pricing behavior, essentially confirming our basic result. It should be noted that we still have limited knowledge about the effects of consumer heterogeneity in more general models, for instance models allowing for continuously distributed demand characteristics; models with heterogeneous switching costs in addition to the demand heterogeneity already modeled; or both. However, a full-fledged analysis of these matters is beyond the scope of the present paper and left as an issue for further research.

In future work we would also like to extend our analysis in a more fundamental way, by allowing for dynamics, that is, by allowing for tacit collusion in addition to switching costs. Padilla (1995) has studied the interplay between switching costs and the scope for reaching a collusive agreement in a repeated price game, and it should be possible to extend his analysis to allow for heterogeneous consumers.

# 6 Appendices

## 6.1 Proof of lemmas 3-5

**Proof of Lemma 3:** Undercutting the rival's high-demand customers entails the possibility that none of the incentive constraints will bind. First, suppose that none of the incentive constraints bind. This implies that quantities are set at their efficient levels, yielding $T_H = \frac{1}{2} - s$, $T_L = \frac{1}{2}L^2$ and $\pi = 1 - 2s + \frac{1}{2}L^2$. Hence, undercutting is blocked iff

$$1 - 2s + \frac{1}{2}L^2 \leq \frac{1}{2} \Leftrightarrow s \geq s^H \equiv \frac{1}{4}\left(1 + L^2\right)$$

For these contracts to be incentive compatible, we must have that

$$\begin{aligned} u_H &= s \geq q_L - \frac{1}{2}q_L^2 - T_L = L - L^2 \\ u_L &= 0 \geq Lq_H - \frac{1}{2}q_H^2 - T_H = L - 1 + s \end{aligned}$$

which amounts to the condition that $L - L^2 \leq s \leq 1 - L$, which is always met for the critical switching cost $s^H$ as long as $L \leq \frac{1}{2}$.

But the incentive constraints may of course bind for optimal undercutting if $s \neq s^H$, and then the expression for an undercutting firm's profit would be more complicated. However, we need not analyze these cases in order to know what we need to know about stability, and the argument is as follows. Suppose $s > s^H$. Then if optimal undercutting does not make any incentive constraint bind, this case is already covered above, and we know that undercutting is not profitable. $s$ may be so much higher than $s^H$ that undercutting makes the low-demand consumers' incentive constraint bind. But this just add another constraint to the undercutting firm's profit maximization problem, and then undercutting becomes even less tempting. Next, suppose $s < s^H$. Again the only interesting cases appear when optimal undercutting makes an incentive constraint bind, this time the high-demand consumers'. However, from the analysis above we know that if $s < s^H$ then there is a number $\epsilon > 0$ such that undercutting by $s^H - \epsilon$ does not make any incentive constraint bind. Moreover, such an undercutting makes the rivals' high-demand consumer switch (as long as $\epsilon < s^H - s$), and such undercutting is profitable (remember that the firm would have been indifferent if undercutting by $s^H$ were required, and would therefore strictly prefer undercutting if a smaller amount sufficed). (Such undercutting is not optimal, but that does not matter for the argument.) QED

**Proof of Lemma 4:** Suppose none of the incentive constraints bind (this will be checked below). If so, then quantities should be set at their efficient levels, and $T_H = \frac{1}{2} - (2L-1)(1-L) - s$, $T_L = \frac{1}{2}L^2$ and $\pi = 2\left(\frac{1}{2} - (2L-1)(1-L) - s\right) + \frac{1}{2}L^2$. By comparing this expression for profit with the firm's profit under monopoly pricing, we have that in order to block undercutting,

$$s \geq s^H \equiv 1 - 2L + \frac{5}{4}L^2$$

What remains is to check whether the incentive constraints are actually satisfied for the proposed contracts. Inserting the proposed contracts in the incentive constraints of the two types of consumers yields

$$
\begin{aligned}
u_H &= \frac{1}{2} - \left(\frac{1}{2} - (2L-1)(1-L)\right) + s \geq L - \frac{1}{2}L^2 - \frac{1}{2}L^2 \\
u_L &= 0 \geq L - \frac{1}{2} - \left(\frac{3}{2} - 3L + 2L^2 - s\right)
\end{aligned}
$$

which amounts to the condition that $(1-L)^2 \leq s \leq 2(1-L)^2$. This constraint holds for $s = s^H$ as long as $L \leq \frac{2}{3}$. (Again it is easily verified that we need not check other values of $s$.)

In contrast, when $L \in \left(\frac{2}{3}, 1\right)$, undercutting high-demand consumers will — for the critical switching cost — imply that the low-demand consumers' incentive constraint binds and that at least one of the participation constraints bind. The undercutting firm then maximizes $\pi = 2T_H + T_L$ subject to (6) and the following constraints (it can now be checked that the high-demand consumers' incentive constraint does not bind):

$$
q_H - \frac{1}{2}q_H^2 - T_H \geq \frac{1}{2} - \left(\frac{1}{2} - (2L-1)(1-L)\right) + s \tag{12}
$$

$$
Lq_L - \frac{1}{2}q_L^2 - T_L \geq Lq_H - \frac{1}{2}q_H^2 - T_H \tag{13}
$$

Since quantities do not enter the objective function and $q_L$ enters (6) and (13) only, $q_L$ should be set in order to relax these constraints as much as possible, implying that $q_L = L$. Rewriting these constraints — assuming that (13) binds — yields

$$
q_H - \frac{1}{2}q_H^2 - T_H \geq \frac{1}{2} - \left(\frac{1}{2} - (2L-1)(1-L)\right) + s \tag{14}
$$

$$
\frac{1}{2}L^2 - T_L = Lq_H - \frac{1}{2}q_H^2 - T_H \tag{15}
$$

$$
\frac{1}{2}L^2 - T_L \geq 0 \tag{16}
$$

Suppose first that all three constrains are binding (that is, that $q_H$ is distorted upward to extract all the low-demand consumers' rent). Then, straight forward computation yields

$$
q_H = \frac{1 - s - 3L + 2L^2}{L - 1}
$$

$$
\begin{aligned}
T_H &= \frac{1}{2}\frac{4L-1-4sL+2sL^2-s^2+2s+2L^3-5L^2}{(L-1)^2} \\
T_L &= \frac{1}{2}L^2 \\
\pi_U &= 2T_H+T_L = \frac{1}{2}\frac{8L-2-8sL+4sL^2-2s^2+4s+2L^3-9L^2+L^4}{(L-1)^2}
\end{aligned}
$$

Comparing $\pi_U$ with $\pi_M$ reveals that in order to block undercutting,

$$
s \geq s^H = \left(1-L+\frac{1}{2}\sqrt{8L-2L^2-4}\right)(1-L)
$$

Secondly, for sufficiently high $L$ it might be the case, however, that the low-demand consumers' participation constraint stops binding. Suppose this is the case. Then by solving the two remaining constraints with equality yields

$$
\begin{aligned}
T_H &= -3L+1+2L^2-s+q_H-\frac{1}{2}q_H^2 \\
T_L &= -Lq_H+\frac{1}{2}q_H^2+T_H+\frac{1}{2}L^2 = -Lq_H-3L+1+\frac{5}{2}L^2-s+q_H \\
\pi &= 2\left(-3L+1+2L^2-s+q_H-\frac{1}{2}q_H^2\right)+\left(-Lq_H-3L+1+\frac{5}{2}L^2-s+q_H\right)
\end{aligned}
$$

The undercutting firm sets $q_H$ to maximize profit. Straightforward calculus reveals that the optimal quantity equals $q_H = \frac{3}{2}-\frac{1}{2}L$, with profit given by $\pi_U = -\frac{21}{2}L+\frac{21}{4}+\frac{27}{4}L^2-3s$. Comparing with the monopoly profit reveals that undercutting is blocked iff

$$
s \geq s^H = \frac{17}{12}-\frac{17}{6}L+\frac{19}{12}L^2
$$

Again we must check whether this solution satisfies the participation constraint of the low-demand consumer. That is, whether $q_H = \frac{3}{2}-\frac{1}{2}L$ yields non-negative $u_L$.

$$
\begin{aligned}
u_L &= \frac{1}{2}L^2-T_L = \frac{1}{2}L^2-\left(-Lq_H-3L+1+\frac{5}{2}L^2-s+q_H\right) \\
&= -\frac{11}{12}L^2+\frac{13}{6}L-\frac{13}{12} \geq 0 \Leftrightarrow L \geq \frac{13}{11}-\frac{1}{11}\sqrt{26}
\end{aligned}
$$

QED.

**Proof of Lemma 5**: Undercutting all is unprofitable if

$$
\begin{aligned}
1-2L+2L^2 &\geq \left(2\left(1-2L+2L^2\right)-4s\right) \\
&\Updownarrow \\
s &\geq s^A \equiv \frac{1}{4}-\frac{1}{2}L+\frac{1}{2}L^2
\end{aligned}
$$

# References


[1] Armstrong, M. and J. Vickers (1999): "Competitive price discrimination." Mimeo, Oxford.

[2] Beggs, A. og P.D. Klemperer (1992): "Multiperiod competition with switching costs," Econometrica 60, 651-66.

[3] Deneckre, R. (1983): "Duopoly supergames with product differentiation, Economics Letters 11, 37-42.

[4] Farrel, J. og C. Shapiro (1988): "Dynamic competition with switching costs," RAND Journal of Economics 19, 123-37.

[5] Klemperer, P. (1987): "Markets with consumer switching costs," Quarterly Journal of Economics 102, 375-94.

[6] Klemperer, P. (1995): "Competition when consumers have switching costs: An overview with applications to Industrial Organization, macroeconomics, and international trade," Review of Economic Studies 62, 515-39.

[7] Padilla, A. J. (1995): "Revisiting Dynamic Duopoly wih Consumer Switching Costs," Journal of Economic Theory 67, 520-30.

[8] Risvold, B. (2000): "Bonusprogram og konkurransen i det norske luftfartsmarkedet [Frequent flyer programs and competition in the Norwegian market for airline services]," Thesis submitted for the degree of Cand. Polit., Department of Economics, University of Bergen.

[9] Rochet, J.-C. and L.A. Stole (1999): "Nonlinear pricing with random participation constraints." Mimeo, Chicago.

[10] Seime, G.R. (1999): "Konkurransen i det norske mobiltelefonimarkedet [The competition in the Norwegian market for mobile telephony]," Report no. 49/99, Foundation for Research in Economics and Business Administration, Bergen.

[11] Stole, L.A. (1995): "Nonlinear pricing and oligopoly." Journal of Economics and Management Strategy, 4, 529-62.

[12] Tirole, J. (1988): The theory of Industrial Organization, MIT Press.

[13] Wilson, R. (1993): Nonlinear pricing, Oxford University Press.