Consumer heterogeneity, incomplete information and pricing in a duopoly with switching costs

Tommy Staahl Gabrielsen, Steinar Vagstad*

Department of Economics, University of Bergen, Fossweinckelsgate 6, N-5007 Bergen, Norway

Received 20 August 2002; accepted 9 December 2002

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. With heterogeneous consumers and incomplete information about individual consumers’ types, monopoly pricing entails second-degree price discrimination with inefficient contracts for low demand types. We show that introducing consumer heterogeneity may increase the critical switching cost needed to sustain a pure-strategy equilibrium involving monopoly pricing.

© 2003 Elsevier B.V. All rights reserved.

Keywords: Incomplete information; Switching costs; Duopoly; Price discrimination

JEL Classification: D43; D82; L10; L13

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 or electricity provision or distribution—now typically have been forced to share their markets with one or more entrants. This gives rise to numerous interesting issues of competition in general and pricing behavior in particular. A particular problem facing firms is how to escape the Bertrand paradox: they compete in markets for more or less homogeneous goods, with prices as the main

*Corresponding author.
E-mail address: steinar.vagstad@econ.uib.no (S. Vagstad).

0167-6245/03/$ – see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0167-6245(03)00004-0
strategic variable. In our view, the most compelling solution to the paradox is the existence of switching costs: the fact that even if consumers do not care about which product they start to buy, there may be costs associated with switching suppliers. 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.

Despite the large theoretical literature on the importance of switching costs on pricing in a variety of markets, there are only a limited number of empirical analyses on the measurement of switching costs. Notable exceptions are Elzinga and Mills (1998), who study the wholesale distribution of cigarettes; Kim et al. (2001), who study bank loans; and Chen and Hitt (2002), whose focus is on what they call internet-enabled businesses. They all agree that switching costs are an important source of market power.

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 discrimination between heterogeneous consumers. 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, and where the firms lack information about these consumer differences. In addition to the already mentioned literature on switching costs and non-linear pricing, there are also many contributions studying non-linear pricing in more or less competitive settings. What these contributions have in common, however, is that they model sources of market power other than switching costs. Wilson (1993, Part 12.3) considers Cournot

---

1 An early contribution to the analysis of markets with switching costs is von Weizsäcker (1984), who also provides a thorough discussion of how switching costs relate to incomplete information. Other proposed solutions to the Bertrand paradox 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). See also Nilssen (1992) for some important distinctions between different types of switching costs.

3 In addition, there are many attempts to measure the related notion of ‘consumer loyalty’ (for a discussion, see Chen and Hitt, 2002, and references therein).

4 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 incomplete 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.)
competition, while Stole (1995), Armstrong and Vickers (2001) and Rochet and Stole (2002) are examples of studies based on the assumption that products are differentiated.\(^5\) Whereas the central focus in all these papers is how market power enables non-linear pricing when firms compete, our focus is different. In a model of imperfect competition based on consumer switching costs we are interested in how consumer heterogeneity affects the critical switching costs needed to sustain non-competitive prices as an equilibrium outcome.\(^6\)

To our knowledge, none has studied the effects of consumer heterogeneity in a homogeneous-good duopoly with switching costs. With homogeneous consumers and linear pricing, Klemperer (1987) shows that monopoly pricing can be achieved if all consumers have (i) strictly positive switching costs and (ii) the switching costs are sufficiently high. We conduct the analysis within a model allowing consumer heterogeneity and non-linear pricing. Our main result is that consumer heterogeneity may increase the critical switching cost, implying that monopoly pricing can be sustained with homogeneous but not with heterogeneous consumers.

To illustrate this result consider a switching cost just high enough to sustain monopoly pricing when all firms have homogeneous consumers of the same type. The critical switching cost is derived from the temptation to offer non-marginal price cuts to the rival’s customers. Suppose then that half of all consumers belonging to each firm experience a drastic reduction in their willingness to pay for the good in question (in fact, so drastic that a monopolist would choose not to serve those consumers). Intuitively, one should think that since the value of each firm’s customer base has been reduced, the temptation to undercut to attract these customers should be reduced and hence the needed switching costs to keep customers from switching could be lower. However, our results show that the opposite may be true. Specifically, in this heterogeneous market, monopoly pricing will no longer be sustainable, in other words the switching cost is no longer sufficient to sustain monopoly pricing as an equilibrium outcome.

The reason for this result is that monopoly pricing with heterogeneous consumers entails inefficient contracts to low-demand types, and that these inefficiencies can be reduced when undercutting your rival, thereby making undercutting more tempting. With homogeneous consumers contracts are efficient, and the temptation to undercut stems solely from the desire to increase the number of customers at a lower price. With heterogeneous consumers, contracts are inefficient to low-demand types to prevent high-demand mimicking low-demand types. When undercutting the rival’s high-demand customers a competitor must offer the high-demand types a more advantageous contract. By doing this it becomes less tempting for high-demand types to mimic low-demand types, and

\(^5\) See also Deneckere (1983).
\(^6\) Farrell and Shapiro (1988) and Wang and Wen (1998) study switching costs models, but their focus is on other aspects of the pricing game.
therefore contracts to own low-demand types can be made more efficient. Thus, with heterogeneous consumers the temptation to undercut stems from two sources: (i) increasing the number of customers as before, and (ii) the efficiency gain from less distortion of low-demand customers. When low-demand contracts would be very inefficient in monopoly, the potential for an efficiency gain is larger, and due to this the temptation to undercut may be larger than when consumers are all of the same type.

The present analysis focuses on the existence of a pure strategy equilibrium involving monopoly pricing. There are at least two reasons why firms should be concerned about the existence of such an equilibrium. First, the alternative to the proposed equilibrium is an equilibrium in which each firm draws tariffs from a distribution ranging from competitive pricing to monopoly pricing, with lower profits as an inevitable result. Second, heterogeneity is not entirely exogenous to the firms, but depends on actions taken in earlier periods. Since we advocate homogeneity, we also provide guidelines for firms seeking to affect the prospects of ending up in an equilibrium involving monopoly pricing.7

The relevance of the present paper to information economics is two-fold. First, there is the relation between the intertemporal nature of information problems and the existence of switching costs as discussed by von Weizsäcker (1984), von Weizsäcker (1984) argues that problems of incomplete information "intrinsically are problems of intertemporal choice", where the benefits of acquiring information today may accrue at a later point in time. He uses reputation as an example. Consumers would be unwilling to be 'locked-in' to a specific supplier unless the supplier has built a reputation for not fully exploiting his advantage in the future. Thus, incomplete information is a prerequisite for the existence of switching costs. Second, there is the issue of firms having incomplete information about consumers' types. With perfect information of consumers' types, firms may offer efficient contracts and fully extract consumers' surplus. When there is private information about consumers' types, some kind of inefficiency is needed to induce self-selection.

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. Some concluding remarks are gathered in Section 4.

---

7 In the present article we abstract from both the issue of equilibrium in mixed strategies and the issue of earlier-period competition. Concentrating on second-period pricing is in line with the approach in Shilony (1977) and Shaffer and Zhang (2000). Moreover, abstracting from mixed-strategy equilibria is in line with Klemperer (1987). See Padilla (1992) for more on mixed-strategy equilibria and switching costs. We believe both issues are important, but beyond the scope of this article. Instead, here we have chosen to focus on the relationship between heterogeneity and the critical switching cost.
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 $s$. In particular, the costs of switching does not depend on a consumer’s demand volume.\(^8\)

In the general case a menu of contracts is described by a payment that is a function of quantity demanded: $T_i = T(q_i)$. With only two types of consumers we can limit attention to two contracts, one for each type, and where each contract is a point contract $(q_i, T_i)$.\(^9\) Consequently, 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:

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

(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:

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

(2)

Moreover, firms are assumed to be symmetric both as regards costs and customer bases from an unmodelled first period. In particular, each firm has a

---

\(^8\) This implies that the only candidate for equilibrium in pure strategies entails monopoly pricing (as specified below). Note that, even if some consumers do not buy in a (proposed) equilibrium, they would still have to pay the switching cost if they want to switch supplier. This may happen because prices are typically low in the (unmodelled) initial period in which firms build customer bases, while the locked-in consumers are exposed to high prices in the present period.

\(^9\) 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 for a typical private consumer, but may be substantial even for private consumers.

\(^10\) We have also performed the analysis assuming simpler pricing schemes, even simple linear pricing, without altering our basic results. This illustrates that our results do not hinge on the ability to price discriminate, but stem solely from the heterogeneity in consumers’ demand. For more details, see Gabrielsen and Vagstad (2000).
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, for simplicity 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 describe a market with heterogeneous consumers in subsequent sections 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 charging an appropriate fixed payment 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. 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. At any lower common price than the monopoly 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 suffice 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: efficiency 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.\(^{11}\) In this respect the magnitude of the switching cost is important.

\(^{11}\) 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.
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 $2T$. Such an undercutting is not profitable iff

$$2(\frac{1}{2} \theta^2 - s) \leq \frac{1}{2} \theta^2.$$  

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 $s \geq s^\ast(\theta) = \frac{1}{4} \theta^2$. If such an equilibrium exists, it is unique and entails $(q, T) = (\theta, \frac{1}{2} \theta^2)$.

### 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 whom 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.

#### 3.1. 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—$\max_{(T, q)} T_H + T_L$—subject to standard participation and incentive constraints:

\[ 2(\frac{1}{2} \theta^2 - s) \leq \frac{1}{2} \theta^2. \]

Here we have assumed that each firm has one consumer of type $\theta$. (What matters is that each firm has a 50% share of the market.) It is easily checked that the firm cannot 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.
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 no 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 under two-part pricing schemes). Assuming 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

$$T_L = L(2L - 1) - \frac{1}{2}(2L - 1)^2 = L - \frac{1}{2},$$

$$T_H = \frac{1}{2} - (1 - L)(2L - 1).$$

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 $\pi_M = 1 - 2L + 2L^2$. 
3.2. 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 (dubbed strategy ‘High’) or for all the competitor’s consumers (strategy ‘All’). 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 monopoly pricing, and thereafter we do the more complicated cases when \( L > \frac{1}{2} \). For each case we calculate the profit of an undercutting firm, and by comparing with profits under monopoly pricing we derive conditions for the existence of the pure-strategy monopoly pricing equilibrium. The purpose of this exercise is to compare the conditions for existence under heterogeneity with those obtained with homogeneous consumers in Lemma 1.

3.2.1. Strategy undercutting ‘High’

We start with situations in which only high-demand consumers are served under monopoly pricing, i.e., \( L \in (0, \frac{1}{2}] \). Strategy ‘High’ entails paying the switching costs of the rival’s \( H \)-type. In monopoly, selling to low-demand consumers was unattractive because of the large 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

\[
\pi = 2T_H + T_L, \tag{12}
\]

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

\[
u_H = q_H - \frac{1}{2} q_H^2 - T_H \geq s, \tag{13}\]

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

We then have the following result:

\[\text{In principle, there is also a third strategy: going for the rival’s low-demand consumers only. It is easily checked that this is never a viable strategy as long as we maintain our assumption of all consumers having the same switching costs. It could change if we allowed low-demand consumers to have substantially lower switching costs.}\]
Lemma 3. For $L \in (0, \frac{1}{2}]$, the switching cost needed to block undercutting the rival’s high-demand consumers is
\[ s \geq s^H(L) = \frac{1}{2} (1 + L^2). \tag{14} \]

Proof. See Appendix A.

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 $L \in (\frac{1}{3}, 1)$ 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:
\[ u_H = q_H - \frac{1}{2} q_H^2 - T_H \geq \frac{1}{2} - \left[ \frac{1}{2} - (2L - 1)(1 - L) \right] + s. \tag{15} \]
This yields the following result:

Lemma 4. For $L \in (\frac{1}{3}, 1)$, the switching cost needed to block undercutting the rival’s high-demand consumers is
\[ s \geq s^H(L) \begin{cases} 1 - 2L + \frac{3}{4}L^2, & \text{if } L \in \left(\frac{1}{3}, \frac{3}{5}\right], \\ \left[1 - L + \frac{1}{2} \sqrt{8L - 2L^2 - 4}\right](1 - L), & \text{if } L \in \left(\frac{7}{11}, \frac{13}{11} - \frac{1}{11} \sqrt{26}\right], \\ \frac{17}{12} - \frac{17}{6} L + \frac{19}{12} L^2, & \text{if } L \in \left(\frac{13}{11} - \frac{1}{11} \sqrt{26}, 1\right). \end{cases} \tag{16} \]

Proof. See Appendix A.

We see from Lemma 4 that the critical switching cost for the interval $L \in (\frac{1}{3}, 1)$ is defined in three parts, depending on which of the incentive and participation constraints that bind. In the first interval, $L \in (\frac{1}{3}, \frac{3}{5}]$, none of the incentive constraints bind when undercutting, hence efficient contracts are offered to both types. For $L \in (\frac{3}{5}, 1)$ undercutting implies that the incentive constraints start to bind the ’wrong way’. Now the undercutting firm needs to worry about the low-demand type mimicking the high-demand type. This interval is again divided into two depending on which participation constraints are binding.\(^\dagger\)

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.

\(^\dagger\) For more details, see the proof of Lemma 4 in Appendix A.
3.2.2. Strategy undercutting ‘All’

This strategy involves lowering the payment by \( s \) for all customers, without affecting the issue of distorting quantities.\(^{15}\) This tells us that it is no point in undercutting the rival’s low-demand consumers if \( L \geq \frac{1}{2} \). If \( L > \frac{1}{2} \), we have the following:

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

**Proof.** See Appendix A.

3.3. Existence of a pure-strategy equilibrium

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 \( L > \frac{14}{15} - \frac{1}{15} \sqrt{13} \), in which case the critical switching cost is given by Lemma 5. Since what matters for existence 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 interest: \( \max(s^M, s^H) \). Then by summing up the information from Lemmas 3–5, we have that undercutting with either strategy is blocked when \( s \geq s^{**}(L) \), where

\[
s^{**}(L) = \begin{cases} 
\frac{1}{4} + \frac{1}{2}L^2, & \text{if } 0 \leq L \leq \frac{1}{2}, \\
1 - 2L + \frac{1}{2}L^2, & \text{if } \frac{1}{2} \leq L \leq \frac{3}{2}, \\
(1 - L + \frac{1}{2} \sqrt{8L - 2L^2 - 4})(1 - L), & \text{if } \frac{3}{5} < L \leq \frac{13}{15} - \frac{1}{15} \sqrt{26}, \\
\frac{17}{12} - \frac{17}{6}L + \frac{12}{15}L^2, & \text{if } \frac{13}{15} - \frac{1}{15} \sqrt{26} < L \leq \frac{14}{15} - \frac{1}{15} \sqrt{14}, \\
\frac{1}{4} - \frac{1}{2}L + \frac{1}{2}L^2, & \text{if } \frac{14}{15} - \frac{1}{15} \sqrt{14} < L \leq 1. 
\end{cases}
\]

(17)

\(^{15}\) Technically, the undercutting firm maximizes \( 2(T_u + T_m) \)—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 \( s \).
In Fig. 1 we have plotted \( s^*(L) \) together with the critical switching costs for homogeneous populations of high-demand consumers \( (s^*(1) = \frac{1}{2}) \), low-demand consumers \( (s^*(L) = \frac{1}{2}L^2) \) and ‘average’ consumers (that is, a hypothetical situation in which all consumers have demand parameter \( (1 + L)/2 \), yielding \( s^*((1 + L)/2) \)).

From Fig. 1 we see that \( s^*(L) \) (the bold line) is non-monotonic. Starting out from \( L = 0 \), \( s^*(L) \) 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 the undercutting strategy from undercutting only the rival’s high-demand consumers to going after all consumers. The increase in \( s^*(L) \) 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 greater, and more so the more surplus that can be extracted from the low-demand customer, hence the increase in \( s^*(L) \). 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 \( s^*(L) \) decreases in \( L \). 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., $L \in (0, \frac{1}{2})$, consumer heterogeneity threatens the existence of the pure-strategy equilibrium 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 increases the possibility of existence, 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.\footnote{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.}

To summarize, we have the following result:

**Proposition 1.** With homogeneous consumers, the critical switching cost $s^*(\theta) \leq s^*(1) = \frac{1}{4}$. With heterogeneous consumers, however, if $L \in (0, \frac{1}{2})$ then $s^{**}(L) > s^*(1)$, and if $L \in (\frac{1}{2}, 1)$ then $s^*(L) < s^{**}(L) < s^*(1)$.

4. 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 and incomplete information about consumers’ types. 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 may create problems with the existence of an equilibrium involving monopoly pricing, and especially so if heterogeneity is large. This result has the immediate implication that the more heterogeneity, the greater the efforts to raise barriers for consumers who may want to switch supplier. Heterogeneity, however, is 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: spread of tariffs offered 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.
In an earlier version of this article (Gabrielsen and Vagstad, 2000), we reported results 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 is left as an issue for further research.

In future work we would also like to extend our analysis in two other ways. First, we wish to study the interplay between the second-period effects on the critical switching cost studied in the present paper, and the competition for (different types of) customers in the first period. Such an extension could build heavily on Paul Klemperer’s earlier work (see the survey Klemperer, 1995). Second, we would also like to determine how things change if we allow for more dynamics, e.g. by allowing for tacit collusion in addition to the switching costs studied here. Beggs and Klemperer (1992) and Padilla (1995) have studied the interplay between switching costs and the scope for reaching a collusive agreement in a repeated version of simpler pricing games, and an extension should take these contributions as starting points.

Acknowledgements

Earlier versions of this paper were presented at conferences in Bergen (NORIO II 1999), Valencia (Game Theory, Experimental Economics and Applications 1999) and Seattle (ESWC 2000). We thank Paul Klemperer, Don Lamberton, Trond Olsen, and seminar participants at the University of Bergen and the University of Helsinki for helpful comments. Parts of this paper were written while we were visiting the University of Pompeu Fabra, Barcelona, whose hospitality is gratefully acknowledged. This research was sponsored by Telenor, through the Foundation for Research in Economics and Business Administration, and by the Meltzer Foundation.

Appendix A

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
For these contracts to be incentive compatible, we must have that
\[ 1 - 2s + \frac{1}{2}L^2 \leq \frac{1}{2} \iff s \geq s^H = \frac{1}{4}(1 + L^2). \]  
(A.1)

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 existence, 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 adds 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.) \( \square \)

**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
\[
T_H = \frac{1}{2} - (2L - 1)(1 - L) - s,
\]
\[
T_L = \frac{1}{2}L^2 \quad \text{and} \quad \pi = 2[\frac{1}{2} - (2L - 1)(1 - L) - s] + \frac{1}{2}L^2.
\]

By comparing this expression for profit with the firm’s profit under monopoly pricing, we have that, in order to block undercutting,
\[ s \geq s^H = 1 - 2L + \frac{5}{4}L^2. \]  
(A.4)

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
\[ u_H = \frac{1}{2} - \left[ \frac{1}{2} - (2L - 1)(1 - L) \right] + s \geq L - \frac{1}{2}L^2 - \frac{1}{2}L^2, \]  
(A.5)

\[ u_L = 0 \geq L - \frac{1}{2} - \left( \frac{3}{2} - 3L + 2L^2 - s \right), \]  
(A.6)

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 \(L \in (\frac{2}{3}, 1)\), 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):

\[ q_H - \frac{1}{2}q_H^2 - T_H \geq \frac{1}{2} - \left( \frac{1}{2} - (2L - 1)(1 - L) \right) + s, \]  
(A.7)

\[ Lq_L - \frac{1}{2}q_L^2 - T_L \geq Lq_H - \frac{1}{2}q_H^2 - T_H. \]  
(A.8)

Since quantities do not enter the objective function and \(q_L\) enters (6) and (A.8) 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 (A.8) binds—yields

\[ q_H - \frac{1}{2}q_H^2 - T_H \geq \frac{1}{2} - \left( \frac{1}{2} - (2L - 1)(1 - L) \right) + s, \]  
(A.9)

\[ \frac{1}{2}L^2 - T_L = Lq_H - \frac{1}{2}q_H^2 - T_H, \]  
(A.10)

\[ \frac{1}{2}L^2 - T_L \geq 0. \]  
(A.11)

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, straightforward computation yields

\[ q_H = \frac{1 - s - 3L + 2L^2}{L - 1}, \]  
(A.12)

\[ T_H = \frac{1}{2} \frac{4L - 1 - 4sL + 2sL^2 - s^2 + 2s + 2L^3 - 5L^2}{(L - 1)^2}, \]  
(A.13)

\[ T_L = \frac{1}{2}L^2, \]  
(A.14)

\[ \pi_U = 2T_H + T_L \]

\[ = \frac{1}{2} \frac{8L - 2 - 8sL + 4sL^2 - 2s^2 + 4s + 2L^3 - 9L^2 + L^4}{(L - 1)^2}. \]  
(A.15)
Comparing $\pi_u$ with $\pi_M$ reveals that, in order to block undercutting,

$$s \geq s^H = (1 - L + \frac{1}{2}\sqrt{8L - 2L^2 - 4})(1 - L).$$ \hspace{1cm} (A.16)

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

$$T_h = -3L + 1 + 2L^2 - s + q_h - \frac{1}{2}q_h^2,$$ \hspace{1cm} (A.17)

$$T_l = -Lq_h + \frac{1}{2}q_h^2 + T_h + \frac{1}{2}L^2 = -Lq_h - 3L + 1 + \frac{5}{2}L^2 - s + q_h,$$ \hspace{1cm} (A.18)

$$\pi = 2(-3L + 1 + 2L^2 - s + q_h - \frac{1}{2}q_h^2) + (-Lq_h - 3L + 1 + \frac{5}{2}L^2 - s + q_h).$$ \hspace{1cm} (A.19)

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

$$s \geq s^H = \frac{17}{12} - \frac{11}{14}L + \frac{13}{12}L^2.$$ \hspace{1cm} (A.20)

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

$$u_L = \frac{1}{2}L^2 - T_L = \frac{1}{2}L^2 - (-Lq_h - 3L + 1 + \frac{5}{2}L^2 - s + q_h),$$ \hspace{1cm} (A.21)

$$= -\frac{11}{12}L^2 + \frac{13}{6}L - \frac{13}{12} \geq 0 \iff L \geq \frac{13}{11} - \frac{11}{11} \sqrt{26}. \hspace{1cm} (A.22)$$

**Proof of Lemma 5.** Undercutting ‘All’ is unprofitable if

$$1 - 2L + 2L^2 \equiv [2(1 - 2L + 2L^2) - 4s],$$ \hspace{1cm} (A.23)

$$\Downarrow$$

$$s \geq s^A = \frac{1}{4} - \frac{1}{2}L + \frac{1}{2}L^2.$$ \hspace{1cm} (A.24)

**References**


