Switching costs, consumers’ heterogeneity and price discrimination in the mobile communications service industry

Nicoletta Corrocher     Lorenzo Zirulia

June 4, 2008

Abstract

In this paper we develop a model that captures some basic features of competition in the infant period of the mobile communications service industry. In a duopoly with price discrimination and switching costs, we analyze the evolution of market structure, when an incumbent and a new entrant compete, and a new class of users with lower willingness to pay appears in the market. We find that the market share of the new entrant depends on the degree of heterogeneity and the level of switching costs. In particular, if the degree of heterogeneity is intermediate, the evolution of market structure is similar for high and null switching costs. Since consumer surplus and social welfare are unambiguously lower under high switching costs, this result points at the risk of inferring the degree of competitiveness from the convergence in market shares.

Key words: switching costs, price discrimination, mobile communications.

JEL Codes: L13, L96.

1 Introduction

When analysing the evolution of the mobile communications service industry in European countries and in other countries as well, such as Australia, Japan or South Korea, some recurrent features emerge.¹ In the initial phase of mobile telephony diffusion, most governments granted one licence to the former monopolist in the fixed telephony market, which became monopolist also in the mobile telephony market. In some countries, such as in the UK, a duopoly was established. In any case, prices tended to be high and penetration rates low. Later on, following technological progress, more operators were allowed to enter the market: competition became more intense, with significant effect on prices and ultimately on diffusion (Gruber, 2005). In all markets, we observe a progressive convergence of market shares (in terms of total subscribers) between

¹The evolution of the US market is peculiar, since local duopolies were established.
the incumbent(s) and the new entrant(s). However, the speed and degree of convergence display differences across countries. In Italy, for instance, the market share of TIM (the incumbent) went from 100% in 1995 to 43.5% in 2005, while the market share of Vodafone Omnitel (the second entrant in the market) increased up to 35.7% in 2005. In the UK, Vodafone, that together with Cellnet was operating as duopolist in the initial phase, saw its market share decreasing from 38% in 1997 to 23.8% in 2005, while, for a later entrant such as Orange, the market share increased from 14% to 22.1% over the same period.

This paper aims at analyzing the dynamics of the market structure in the mobile communications service industry, by considering a model of oligopolistic competition which includes two distinctive features of the industry, i.e. switching costs and price discrimination.

In the mobile communication service industry, switching costs can arise from different sources. There are transaction costs, related to the lack of number portability, which has characterized the industry for a long time. If a customer has to change his number when he changes provider, he bears a cost in communicating the new number to his habitual contacts (or potential costs exist if he does not to do so). However, it is important to underline that the lack of number portability is not the only possible source of switching costs in the industry. Contracts (and "extras"/"seasonal promotions") are usually designed so that on-net tariffs (i.e., tariffs for consumers served by the same operator) are lower than off-net tariffs (i.e., tariffs for consumers served by different operators). Also with number portability, consumers have to communicate their habitual contacts that they have changed operator, since this change affects the cost of calls for those who call them. In some countries, switching costs can also arise because in some countries operators can lock handsets to be used exclusively within their own networks, and unlocking comes at a cost. Furthermore, there are search costs because consumers have to gather information about other operators’ characteristics, in particular about tariff plans. Finally, switching costs can be also psychological, due to consumers’ inertia. Evidence for the existence and relevance of switching costs in mobile communications can be found, for instance, in Lee et al. (2006) and Grzybowski (2008). Based on a survey of 466 consumers, Lee et al. (2006) shows that in the Korean mobile communications industry number portability has reduced switching costs, but these are still significant: the mean value of switching costs is 34$. Grzybowski (2008) shows that switching costs are significant in the UK mobile communications industry, even with number portability.

Consumers’ heterogeneity and price discrimination are other distinctive features of this industry. Consumers vary greatly in terms of willingness to pay and modes of service usage. Consumers assigning a higher value to the service (e.g. businessmen) adopted earlier, since they were the only ones that could afford the initial high costs of handsets (Gruber, 2005). At the same time, second-degree price discrimination is commonly observed: firms offer multiple tariff

\[\text{For examples of models that capture other relevant aspects of this industry, see Valletti (1999) and the survey by Gans et al. (2004).}\]
plans, each designed for a specific class of consumers. For instance, the Italian mobile operators (TIM, Vodafone, Wind and 3 Italy) collectively introduced 121 tariff plans (for pre-paid cards) from 2000 to 2005. Similarly, in 1997 the UK operators (Cellnet, Vodafone, Mercury One-2-One, Orange) were offering 20 basic tariff plans (Valletti and Cave, 1998).

In the model we consider a situation where an incumbent firm, which served consumers with high willingness to pay in an unmodeled first period, compete with a new entrant when facing a new cohort of consumers with lower willingness to pay. We model explicitly the "second period" competition, and we compare two cases. In the first case, switching costs are absent, so that both firms compete for both types of consumers (with high and low willingness to pay). In the second case, switching costs are high, so that old consumers are "locked-in" to the incumbent firm. When serving new consumers, however, the incumbent must apply second-degree price discrimination, since contracts that are specifically designed for consumers with lower willingness to pay are attractive also for old users.

In comparing the two cases, we first look at the new entrant market share, which can be interpreted as a measure of convergence in market shares between the leader and the follower (notably, convergence may even lead to the new entrant becoming the new leader in the market, i.e. "leapfrogging" can occur). A-priori, it is not clear if we should expect the new entrant market share to be higher under positive or under null switching costs. With null switching costs and symmetric firms, we show that in equilibrium firms, as expected, equally share the market between them. With positive (and high) switching costs, there are two opposite forces at work. On one side, the incumbent can benefit from an installed base of locked-in consumers. On the other side, one can expect the incumbent firm to be less aggressive than the new entrant on the new consumers, which is what we actually find in our set-up. This is due to a "fat cat" effect which is commonly found in the literature on switching costs: large firms (in this case the incumbent firm) are less aggressive on new consumers, because they want to exploit their installed base.

The comparison provides the following results. If consumers' heterogeneity is absent or low, under null switching costs the new entrant market share is higher and the two firms equally share the market between them. However, under high switching costs the incumbent firm has sufficient incentives to serve also new consumers and it remains the market leader. If instead consumers' heterogeneity is high, under high switching costs the new entrant becomes the leader, because the incumbent prefers exploiting old consumers. In the intermediate case, the result depends on the share of consumers with high willingness to pay: if this is high, the new entrant market share is higher under null switching costs. From another point of view, this means that, for intermediate degrees of heterogeneity, ex post market structure can be similar for high or zero switching costs, which implies that the rate of convergence is not sufficient to determine the degree of "frictions" in the market. This results has relevant policy implications, since we are able to show that the two cases significantly differ in terms of firms' profits (under high switching costs, the incumbent increases its profit and the
new entrant decreases it, so that the net effect is unambiguously negative) and in terms of social welfare, which is unambiguously lower under high switching costs, with a negative effect both on total profits and consumer surplus. While our model refers to the past evolution of the industry (which is now mature in most countries), we believe that other sectors as well may present similar characteristics, so that our results may be also relevant to them.

From a theoretical point of view, this paper makes a contribution to two different streams of literature. First, with reference to the literature on competition in presence of switching costs (see Farrell and Klemperer, 2007, for a survey), this paper investigates pricing strategies when new cohorts of consumers arrive over time and the implications in terms of industrial dynamics. We add consumers’ heterogeneity with second-degree price discrimination to the framework usually adopted in this literature. Second, we contribute to the literature on second-degree price discrimination in oligopoly (Stole, 1995; Valletti, 2000; see also Stole, 2007, for a survey of price discrimination in competitive environments), by focusing on the role of the installed base of consumers in providing the incentives to offer new contracts.

This paper is structured as follows. Section 2 describes the set-up of the model, which is solved in Section 3. In Section 4 the results under null and high switching costs are compared in terms of ex post market structure and social welfare. Finally, Section 5 concludes.

2 The model

The model builds on the recent literature that extends second-degree price discrimination to competitive environments (Stole, 1995; Valletti, 2000; Gabrielsen and Vagstad, 2003). We consider a market with two firms, an incumbent (firm I) and a new entrant (firm E). The two firms compete offering two horizontally differentiated products. Consumers are heterogenous along two dimensions: they are uniformly located on a segment $[0, 1]$, so that they are identified by their position $d_i \in [0, 1]$, and they have a different willingness to pay. In particular, consumers can attribute a high value to the good ($\theta = \bar{\theta}$) with probability $x$ - from now on these will be high type consumers - or a low value to the good ($\theta = \underline{\theta}$), with probability $(1 - x)$ - from now on these will be low type consumers. We assume that $x \leq \frac{\bar{\theta}}{\underline{\theta}}$. This guarantees that a monopolist, if switching costs were null, would have incentives to serve both types of consumers. Finally, we assume that consumers’ position is observable and firms can offer contracts contingent to $d$. This is in line with Stole (1995) and Valletti (2000).

In the first, unmodeled period, only firm I and high type consumers are in the market. We assume that the market is fully covered, i.e. all high type consumers buy from firm I. What we model explicitly is the second period competition, which is in line with Gabrielsen and Vagstad (2003 and 2004) and Shaffer and Zhang (2000). The link between the first and the second period lies

---

3 For a model with heterogenous consumers, but without price discrimination, see Klemperer (1989).
in the switching costs \( \sigma \geq 0 \) that high type consumers bear if they buy from firm \( E \) in the second period.

On the supply side, firms have the same cost structure, with zero fixed and marginal costs. Firms are located at the extremes of the interval \([0, 1]\), with firm \( I \) localized at 0 and firm \( E \) localized at 1. Consequently, consumers' position corresponds to the relative preference for one firm or the other. Since \( d \) is observable, firms offer contracts contingent to \( d \). A contract offered by a firm is \( \{q, T\} \), where \( q \) is the quantity of the good and \( T \) is a money transfer to the firm. We allow firms to use non-linear tariffs, which are commonly observed in the mobile communications service industry.

On the demand side, consumer's utility is specified as follows. When consumer's type is \( \{d_i, \theta\} \) and she buys a contract \( \{q, T\} \) from a firm located at \( d \), her utility is given by:

\[
U(d_i, \theta, d) = \theta(1 - |d - d_i|)q - \frac{1}{2}q^2 - T - \sigma d
\]

(1)

while for low type consumers, the utility is given by:

\[
U(d_i, \theta, d) = \theta(1 - |d - d_i|)q - \frac{1}{2}q^2 - T
\]

(2)

The assumed utility function results in a linear inverse demand function for the individual consumer of type \( \{d_i, \theta\} \) who buys from a firm located at \( d \): \( p = \theta(1 - |d - d_i|) - q \). If consumers do not buy any contract, they obtain a utility level normalized to 0.

In the second period competition, we consider a simultaneous move game, in which the two firms choose a menu of contracts \( M_j(d) \), \( j = I, E \) in each location \( d \). Since \( \theta \) can assume two values, in equilibrium firms can offer zero, one or two contracts.

As a solution concept for each location \( d \), we adapt the notion of undercut-proof equilibrium (UPE), introduced by Morgan and Shy (2000), and further discussed in Shy (2001) and Shy (2002). In our game, as in the games discussed by Morgan and Shy, a Nash equilibrium in pure strategies does not exist when \( \sigma > 0 \), because Edgeworth cycles in tariffs emerge.\(^4\) UPE considers a specific type of conjectural variations behaviour in which each firm assumes that the rival firm will modify its menu of contracts only if such an action satisfies two properties: (a) the firm will appropriate the customers of the firm it "undercuts"\(^5\), and (b) such undercutting is profitable.

**Definition 1** A pair of weekly undominated\(^6\) menu \((M^*_I(d), M^*_E(d))\) constitutes a UPE at \( d \) if:

\(^4\)The nature of this cycle is discussed in the proof of Proposition 2.

\(^5\)It is worth noting that "undercutting" here does not occur in the space of prices, but in the space of consumers' utility. In general, a firm "undercuts" the other offering a contract which yields a slightly higher utility for the consumers. This is line with Armstrong and Vickers (2001).

\(^6\)Restriction to weekly undominated strategies guarantees the uniqueness of equilibrium. The same restriction is in Valletti (2000).
1. for any $M'_I(d)$ such that $\Pi_I(M'_I(d), M^*_E(d)) > \Pi_I(M^*_I(d), M^*_E(d))$, there exists a contract $M''_E(d)$ such that $\Pi_I(M'_I(d), M''_E(d)) < \Pi_I(M^*_I(d), M^*_E(d))$ and $\Pi_E(M'_I(d), M'_E(d)) > \Pi_E(M^*_I(d), M'_E(d))$.

2. for any $M'_E(d)$ such that $\Pi_E(M'_I(d), M'_E(d)) > \Pi_E(M^*_I(d), M'_E(d))$, there exists a contract $M''_I(d)$ such that $\Pi_E(M'_I(d), M''_I(d)) < \Pi_E(M^*_I(d), M'_E(d))$ and $\Pi_I(M'_I(d), M_E(d)) > \Pi_I(M^*_I(d), M_E(d))$.

### 3 Solution

As we mentioned in the Introduction, we solve the model for two specific values of $\sigma$: 1) null switching costs ($\sigma = 0$); 2) high switching costs ($\sigma = \sigma > \frac{22}{2}$). The analysis of the intermediate cases is significantly more complex and does not provide further economic insights.

In the case of $\sigma = 0$, firms compete for both high type and low type consumers. In the second case, firms compete for low types only. However, such competition affects high type consumers (and the profits firm $I$ can extract from them), through their incentive compatibility constraint.

#### 3.1 No switching costs

If $\sigma = 0$, the model is isomorphic to Valletti (2000). Given the symmetry of the set-up, the equilibrium is also symmetric, with firm $I$ serving the first half of the market (where it has a competitive advantage with respect to the rival) and firm $E$ serving the remaining half. To see why this must be the case, consider the interval $[0, \frac{1}{2}]$. Firm $E$ cannot serve consumers in an UPE at $d$ in this interval. Otherwise, if firm $E$ makes non-negative profits, firm $I$ can offer the same contract as $E$, obtaining higher profits. However, competition does have an impact, because firm $I$ must provide to each type of consumers a utility level which is at least as high as the maximum utility the other firm can provide. If not, this will give room for profitable undercutting for firm $E$.

Given symmetry, we focus on $[0, \frac{1}{2}]$, where in equilibrium consumers are served by firm $I$. The hypothesis of null marginal and fixed costs implies that the contract by firm $E$ which maximizes the utility of a consumer of type $\{d_i, \theta\}$ is $T^\text{max} = 0$ and $q^\text{max} = \arg \max q \theta d_i q - \frac{1}{2} q^2 \equiv d_i q$, which entails a maximum utility $u^\text{max}_L = \frac{\theta^2 d_i^2}{2}$ for low type consumers and $u^\text{max}_H = \frac{\theta^2 d_i^2}{2}$ for high type consumers. The optimization problem for firm $I$ is equivalent to a monopolistic second-degree price discrimination problem with type-dependent reservation utility. In particular, firm $I$ solves the program $P1$:
Next proposition summarizes the optimal contract.

**Proposition 2** The solution of program $P_1$ is characterized as follows:

1. For $d_i \in [0, d_A]$, where $d_A$ is the unique solution of $(\frac{\theta + \bar{q}}{2}) (1 - d_A)^2 = 0$, the optimal menu of contracts implies that $IC_H$ and $IR_L$ are binding, with:

   \[
   \begin{align*}
   T_H^* &= \left( \frac{\bar{q}^2}{2} - \frac{(\theta - \bar{q})(\theta - x\bar{q})}{(1 - x)} \right) (1 - d_i)^2 - \frac{\bar{q}^2 d_i^2}{2} \\
   T_L^* &= \left( \bar{q}(1 - d_i) \right) (1 - d_i) \\
   q_H^* &= \frac{(\theta - x\bar{q})}{(1 - x)} (1 - d_i) \\
   q_L^* &= \frac{(\theta - x\bar{q})}{(1 - x)} (1 - d_i) \\
   \Pi^* &= \left( x\bar{q}^2 + \frac{1}{2} \frac{(\theta - x\bar{q})^2}{(1 - x)} \right) (1 - d_i)^2 - \frac{\bar{q}^2 d_i^2}{2}
   \end{align*}
   \]

2. For $d_i \in [d_A, d_B]$, where $d_B$ is the unique solution of $(\frac{\bar{q} + \theta}{2}) d_B^2 - \theta (1 - d_B)^2 = 0$, the optimal menu of contracts implies that $IC_H$, $IR_H$, $IR_L$ are binding, with:
3. For \( d_i \in [d_B, \frac{1}{2}] \), the optimal menu of contracts implies that \( IR_H \) and \( IR_L \) are binding, with:

\[
T_H^* = \frac{\bar{\theta}^2}{2} (1 - d_i)^2 - \frac{\theta^2 d_i^2}{2} \\
T_L^* = \frac{\theta^2 (1 - d_i)^2}{2} - \frac{\theta^2 d_i^2}{2} \\
q_H^* = \bar{\theta}(1 - d_i) \\
q_L^* = \frac{(\bar{\theta} + \theta) d_i^2}{2 (1 - d_i)^2} \\
\Pi^* = x \left( \frac{\bar{\theta}^2}{2} (1 - d_i)^2 - \frac{\theta^2 d_i^2}{2} \right) + (1 - x) \left( \frac{\bar{\theta}^2}{2} (1 - d_i)^2 - \frac{\theta^2 d_i^2}{2} \right)
\]


3.2 High switching costs

If \( \sigma = \pi \), firms compete for low types only, since high types, independently from their location, do not switch to \( E \) even if this firm guarantees to them the highest utility level. However, competition for low types is affected by the existence of an installed base for firm \( I \), since this firm must preserve the incentive compatibility for high type consumers, if it wants to serve both types of consumers.

For the same argument as before, firm \( I \) cannot serve low type consumers in \((\frac{1}{2}, 1]\). Therefore, we focus here on the interval \([0, \frac{1}{2}]\). Firm \( I \) has two alternatives: serving both types of consumers or serving high type only. We consider these two alternatives in turn.

Suppose that firm \( I \) serves both types of consumers at \( d \). In this case, it solves a monopolistic screening problem with type-dependent reservation utility, similar to the one described in the previous section. The difference stands in the reservation utility for high type consumers. For high type consumers, the only
feasible alternative is not buying (i.e. utility equal to 0). Low type consumers, instead, have the alternative of buying from $E$.

Suppose that firm $I$ is willing to serve both types of consumers at $d$. Since firm $E$ can provide low type consumers a maximum utility $u_L^\text{max} = \frac{\theta^2 d^2}{2}$, the optimal contract solves the problem $P2$:

$$\max_{(q_H,q_L,T_H, T_L)} \Pi \equiv xT_H + (1-x)T_L \quad \text{s.t.}$$

$$\bar{\theta}(1-d_i) - \frac{1}{2}q_H^2 - T_H \geq 0 \quad \text{(IR}_H)$$

$$\bar{\theta}(1-d_i) - \frac{1}{2}q_L^2 - T_L \geq \frac{\theta^2 d^2}{2} \quad \text{(IR}_L)$$

$$\bar{\theta}(1-d_i) - \frac{1}{2}q_H^2 - T_H \geq \bar{\theta}(1-d_i) - \frac{1}{2}q_L^2 - T_L \quad \text{(IC}_H)$$

$$\bar{\theta}(1-d_i) - \frac{1}{2}q_L^2 - T_L \geq \theta(1-d_i) - \frac{1}{2}q_H^2 - T_H \quad \text{(IC}_L)$$

Next proposition summarizes the results. We denote with $\Pi_{HL}^*$ the equilibrium profits obtained by firm $I$ if it serves both types of consumers.

**Proposition 3** The solution of the program $P1$ implies that, in the optimal menu of contracts, $IC_H$ and $IR_L$ are binding, with:

$$T_H^* = \left(\frac{\theta^2}{2} - \frac{(\bar{\theta} - \theta)(\bar{\theta} - x\bar{q})}{(1-x)}\right)(1-d)^2 - \frac{\theta^2 d^2}{2} \quad (19)$$

$$T_L^* = \left(\frac{\theta}{2} - \frac{1}{2} \frac{(\bar{\theta} - x\bar{q})}{(1-x)}\right)(1-d)^2 - \frac{\theta^2 d^2}{2} \quad (20)$$

$$q_H^* = \bar{\theta}(1-d_i) \quad (21)$$

$$q_L^* = \frac{\theta - x\bar{q}}{(1-x)}(1-d_i) \quad (22)$$

$$\Pi_{HL}^* = \left(\frac{\theta^2}{2} + \frac{1}{2} \frac{(\bar{\theta} - x\bar{q})^2}{(1-x)}\right)(1-d)^2 - \frac{\theta^2 d^2}{2} \quad (23)$$

**Proof.** We first note that, at the optimum, $IR_L$ must be binding. From $IC_H$ and $\bar{\theta} > \theta$ we obtain:

$$\bar{\theta}(1-d_i) - \frac{1}{2}q_H^2 - T_H \geq \bar{\theta}(1-d_i) - \frac{1}{2}q_L^2 - T_L > \theta(1-d_i) - \frac{1}{2}q_L^2 - T_L - \frac{\theta^2 d^2}{2} \quad (24)$$

If $\theta(1-d_i) - \frac{1}{2}q_L^2 - T_L > \frac{\theta^2 d^2}{2}$, firm $I$ can increase its profits by increasing $T_H$ and $T_L$ by the same amount, with all the constraints still satisfied. Then, $T_L = \theta(1-d_i) - \frac{1}{2}q_L^2 - \frac{\theta^2 d^2}{2}$.
Second, IC\textsubscript{H} must be binding at the optimum. Ab absurdo, if:

\[
\overline{u}(1-d_i) - \frac{1}{2} q_i^2 - T_H > \overline{u}(1-d_i) - \frac{1}{2} q_i^2 - T_L \geq \overline{u}(1-d_i) - \frac{1}{2} q_i^2 + T_L = 0 \tag{25}
\]

firm E can increase its profits by increasing \( T_H \), while IC\textsubscript{H} and IC\textsubscript{L} are still satisfied. This implies \( T_H = \overline{u}(1-d_i) - \frac{1}{2} q_i^2 - \overline{u}(1-d_i) + \frac{1}{2} q_i^2 + T_L \).

If we plug \( T_H \) and \( T_L \) into the profit function, and we derive with respect to \( q_H \) and \( q_L \), we obtain \( q_i^* \) and \( q_i^* \), and then \( T_H \) and \( q_L \) reported in the Proposition. It is then immediate to verify that these values satisfy the ignored constraints IC\textsubscript{H} and IC\textsubscript{L}.

Firm I has the alternative to serve high type consumers only, by offering a single contract. In this case, firm I maximizes \( T_H \) under the individual rationality constraint. This yields \( T_H^* = \overline{u}(1-d) \) and profits are \( \Pi_H(d) = x \overline{u}(1-d)^2 \). Given the profits associated to each alternative, firm I will choose the menu of contracts which maximizes its profits. If we define \( G(d) \equiv \Pi_H(d) - \Pi_L(d) = \frac{1}{2} \left( \frac{\theta - x \overline{u}}{(1-x)} \right)^2 (1-d)^2 - \frac{\theta^2 d^2}{2} \), firm I will serve both types of consumers if \( G(d) > 0 \); otherwise, it will serve high type consumers only. Next proposition summarizes the equilibrium menu of contracts offered by firm I in each \( d \in [0, \frac{1}{2}] \).

We find that firm I serves consumers at locations which are sufficiently close to 0, i.e. for \( d \leq d^* \), with \( d^* \leq \frac{1}{2} \).

**Proposition 4** There exists a unique \( d^* \in [0, \frac{1}{2}] \) such that, for \( d \leq d^* \), in the UPE firm E serves both types of consumers, while for \( d^* < d \leq \frac{1}{2} \), firm I serves only high type consumers and firm E serves low type consumers. \( d^* \) is the unique solution to \( G(d) = 0 \). In \( [0, d^*] \), firm I offers the menu of contracts obtained solving \( P2 \). In \( (d^*, \frac{1}{2}] \), firm I offers \( \{q_H^* = \overline{u}(1-d), T_H^* = \overline{u}(1-d)^2\} \) and firm E offers the contract \( \{q_L^* = \theta d, T_L^* = \theta^2 d^2 - \frac{1}{2} \left( \frac{\theta - x \overline{u}}{(1-x)} \right)^2 (1-d)^2\} \).

**Proof.** We first prove that \( d^* \) exists and it is unique. We have \( G(0) = \frac{1}{2} \left( \frac{\theta - x \overline{u}}{(1-x)} \right)^2 > 0 \), \( G(\frac{1}{2}) = \frac{1}{8} \left[ \left( \frac{\theta - x \overline{u}}{(1-x)} \right)^2 - \overline{u}^2 \right] < 0 \) (since it is decreasing in \( \overline{u} \) and negative for the smallest conceivable value of \( \overline{u}, \overline{u} \)) and \( G'(d) = \left[ -\frac{(\theta - x \overline{u})^2}{(1-x)} (1-d) - \theta^2 d \right] < 0 \). Then, for the Weierstrass intermediate value theorem, existence and uniqueness follows. In order to prove that firms share the market between them as indicated in the proposition, note the following. If \( d \leq d^* \) (which implies \( G(d) \geq 0 \)), firm I finds it convenient to serve both types of consumers, even for the maximum value of utility firm E can guarantee to low type consumers. The contract firm I offers is the solution to \( P2 \). If \( d > d^* \) (which implies \( G(d) < 0 \)), firm I prefers to serve high types only rather than serving both types with a contract that makes undercutting by E unprofitable. Firm I serves high types, while firm E serves low types with the contract specified in the proposition. Notice that firm E must guarantee a utility equal to \( \frac{1}{2} \left( \frac{\theta - x \overline{u}}{(1-x)} \right)^2 (1-d)^2 \); otherwise, firm I would find it convenient to "undercut", serving both type of consumers.
3.2.1 Equilibrium market shares

On the basis of the proposition above, we can derive the equilibrium in the overall segment. In \( \frac{1}{2}, 1 \), firm I will serve high type consumers (because of switching costs), while firm E will serve low type consumers (since \( G(d) < 0 \) in this interval). Firm I will offer \( \{ q^*_H = \tilde{b}(1 - d), T^*_H = \frac{\tilde{b}^2(1-d)^2}{2} \} \), while firm E will offer \( \{ q^*_L = 2d, T^*_L = \frac{\tilde{b}^2}{2} - \frac{1}{2} \frac{(1 - x)^2}{1 - \frac{d}{x}} (1 - d)^2 \} \).

Firm I market share in the low type segment is simply \( d \) and consequently firm E market share is \( 1 - d \). Since \( d < \frac{1}{2} \), firm E is the leader in the low type segment. This result comes from a "fat cat" effect (in the terminology introduced by Fudenberg and Tirole, 1987), which is common in the literature on the dynamic effects of switching costs on market competition (see for instance Beggs and Klemperer, 1992). The incumbent faces a trade-off between the exploitation of the existing installed base of consumers and the acquisition of new consumers, and for this reason it is less aggressive than the new entrant. Even if the firm can discriminate between cohorts, it can do it by self-selection only, and this guarantees locked-in consumers an informational rent. It is important to underline that the result does not depend on the existence of consumers’ heterogeneity, but it is reinforced when new consumers have lower willingness to pay. With respect to the existing literature, an element of novelty is that the lower aggressiveness of the incumbent is reflected in the possibility that it does not offer a contract for new consumers at a particular location.

Given firms’ market shares in the high type and in the low type segment, we can finally derive firms’ overall market shares. Firm I market share is \( x + (1 - x)d^* \), while firm E market share is \( (1 - x)(1 - d^*) \).

4 Market shares’ evolution and social welfare

In this section, we compare the results obtained under \( \sigma = 0 \) (no switching costs) with the results obtained with \( \sigma = \sigma \) (high switching costs) The comparison is carried out along two dimensions: first, we look at ex-post market structure, for different degree of consumers’ heterogeneity; second, we analyze social welfare in the two cases, disentangling the effects on profits and consumer surplus.

4.1 Market shares’ evolution

The first aspect we investigate concerns the evolution of market shares and market structure. In particular, we use firm E market share in the overall market as a measure of "convergence". Notice that convergence may also lead to "leapfrogging", when firm E becomes the leader in the overall market.

In the case of \( \sigma = 0 \), firms have equal market shares. In case of \( \sigma = \sigma \), firm E market share is \( (1 - x)(1 - d^*) \). "Convergence" is higher under high switching costs if \( (1 - x)(1 - d^*) > \frac{1}{2} \), which is clearly also the condition under which firm E "leapfrogs" firm I. Next proposition summarizes the results.
Proposition 5  If \( \bar{\theta} > \frac{1}{3\bar{\sigma}} \), firm E market share is higher under \( \sigma = 0 \) than under \( \sigma = \bar{\sigma} \). If \( \bar{\theta} < \frac{1}{3\bar{\sigma}} \), firm E market share is higher under \( \sigma = \bar{\sigma} \) than under \( \sigma = 0 \). For \( \frac{1}{3\bar{\sigma}} < \bar{\theta} < \frac{3}{2\bar{\sigma}} \), firm E market share is higher under \( \sigma = \bar{\sigma} \) if \( x > x^*(\bar{\theta}, \bar{\sigma}) \).

Proof. Since \( d^* = \frac{1}{\sqrt{1-x}} \frac{(\theta-x\sigma)}{(2\theta-x\sigma)} \), firm E market share is \((1-x)(1-d^*) - \frac{1}{2} = (1-x) - \sqrt{1-x} \frac{(\theta-x\sigma)}{(2\theta-x\sigma)} - \frac{1}{2} \). We define \( H(x) \equiv (1-x) - \sqrt{1-x} \frac{(\theta-x\sigma)}{(2\theta-x\sigma)} - \frac{1}{2} \) (then firm E market share is higher under \( \sigma = \bar{\sigma} \) if \( H(x) > 0 \)). We note the following:

1. \( \lim_{x \to 0} H(x) = 0 \) and \( \lim_{x \to \frac{\theta}{\bar{\theta}}} H(x) = 1 - 2\frac{\theta}{\bar{\theta}} \). Then, \( H(x) \) takes a positive value in \( x = \frac{\theta}{\bar{\theta}} \) if \( \bar{\theta} > \frac{1}{2\bar{\sigma}} \).

2. The derivative of \( H(x) \) is \( H'(x) = -1 + \frac{\theta}{2\bar{\theta}} + \frac{1}{4} \). Then, the derivative is positive for \( \bar{\theta} > \frac{1}{2\bar{\sigma}} \).

3. In any \( x^* > 0 \) such that \( H(x^*) = 0 \), \( H'(x) \) is positive, which implies that such \( x^* \) (if it exists), it is unique. Given the definition of \( x^* \), we have \( \frac{(\theta-x\sigma)}{(2\theta-x\sigma)} = \sqrt{1-x^*} + \frac{1}{2\sqrt{1-x}} \). If we substitute into \( H'(x) \), we obtain \( H'(x^*) = -1 + \frac{\theta}{2\bar{\theta}} + \frac{1}{4} + \frac{1}{4} > 0 \). This function is increasing in \( x^* \). Since \( \lim_{x^* \to 0} H'(x^*) = -1 + \frac{\theta}{2\bar{\theta}} + \frac{1}{4} + \frac{1}{4} > 0 \), we have the claim.

Results (1-3) together imply that there are three possible situations: i) if \( \bar{\theta} < \frac{1}{3\bar{\sigma}} \), \( H(x) \) is always negative; ii) if \( \bar{\theta} > \frac{3}{2\bar{\sigma}} \), \( H(x) \) is always positive; iii) if \( \frac{1}{3\bar{\sigma}} < \bar{\theta} < \frac{3}{2\bar{\sigma}} \), \( H(x) \) is negative for \( x < x^* \), and positive for \( x > x^* \). The Proposition follows. \( \blacksquare \)

Proposition 5 can be interpreted as follows. While firm E market share is independent of \( \bar{\theta} \) when \( \sigma = 0 \), it is decreasing in \( \bar{\theta} \) when \( \sigma = \bar{\sigma} \). The intuition is that the higher is consumers’ heterogeneity, the lower is the incentive for firm I to serve low type consumers, since the higher is the cost of serving them while preserving incentive compatibility for high type consumers. For an intermediate degree of heterogeneity, the case with high and no switching costs are similar in terms of market shares (being equal for \( x = x^* \)). What is most interesting here is that a certain degree of heterogeneity is a necessary condition for "observational equivalence" to occur. In other terms, if consumers are heterogenous and firms apply second-degree price discrimination, observing the rate of convergence in market shares is not sufficient to evaluate the extent to which switching costs exist in the market.
4.2 Social welfare

In our view, the most interesting results of the previous section concern the fact that the ex-post market structure can be similar in the two cases of null or high switching costs. Having a similar market structure in terms of firms’ market shares, however, does not clearly imply similarity over other dimensions. In particular, the level and distribution of firm profits are likely to be affected by the presence of switching costs, and the same is for consumer surplus. For this reason, we now compare the two cases \( \sigma = 0 \) and \( \sigma = \bar{\sigma} \) under these dimensions.

First of all, we compare firm \( I \) and firm \( E \) for a given \( \sigma \). If \( \sigma = 0 \), firm \( I \) and firm \( E \) makes equal profits, due to symmetry. In Proposition 6 we show that firm \( I \) makes always higher profits in presence of high switching costs. This occurs independently of firms’ market shares. When \( E \) becomes the market leader, the negative effect on firm \( I \) profits is more than compensated by the larger surplus extracted from high type consumers.

**Proposition 6** If \( \sigma = \bar{\sigma} \), firm \( I \) makes always higher profits than firm \( E \).

**Proof.** Firm \( I \) profits are equal to:

\[
\int_0^{1-d^*} \left[ \frac{x^2}{2} + \frac{1}{(1-x)} \left( (1-d)^2 - x \bar{\theta}^2 \right) \right] \, dd + \int_{1-d^*}^1 \left[ \frac{x^2}{2} (1-d)^2 \right] \, dd
\]

Firm \( E \) profits are equal to:

\[
\int_{1-d^*}^1 \left[ (1-x) \frac{\bar{\theta}^2}{2} d^2 - \frac{1}{(1-x)} \left( (1-d)^2 - x \bar{\theta}^2 \right) \right] \, dd
\]

Solving the integrals, firm \( I \) profits are higher if

\[
\frac{1}{6} \left( \frac{\theta - x \bar{\theta}}{1-x} \right)^2 - \frac{1}{6} (1-x) \bar{\theta}^2 + \frac{1}{6} x \bar{\theta}^2 \geq 0
\]

and simplyfying we get

\[
\bar{\theta}^2 + x^2 \bar{\theta}^2 - 2x \theta \bar{\theta} - \bar{\theta}^2 - x^2 \bar{\theta}^2 + 2x \theta^2 + x \bar{\theta}^2 - x^2 \bar{\theta}^2 = x(\bar{\theta} - \theta)^2 \geq 0
\]

The second result concerns the comparison between firms’ total profits under under \( \sigma = 0 \) and under \( \sigma = \bar{\sigma} \). We first prove that that firm \( I \) profits are always higher in presence of switching costs.

**Proposition 7** Firm \( I \)’s profits are higher with \( \sigma = \bar{\sigma} \) than with \( \sigma = 0 \).

**Proof.** Consider first the range \([0; \frac{1}{2}]\). Suppose that firm \( I \) serves both types of consumers when \( \sigma = \bar{\sigma} \). Firm \( I \) would obtain higher profits than in the case of...
\( \sigma = 0 \), since high type consumers have a lower reservation utility in this case. In the interval \((d^*; \frac{1}{2})\), where firm \( I \) choose to serve high type consumers only, its profits are a fortiori higher. Finally, if \( \sigma = \sigma \), firm \( I \) serves also high types in the interval \([\frac{1}{2}; 1]\), obtaining positive profits. \( \blacksquare \)

We were not able to prove a similar analytical result for firm \( E \). In this case, two opposing forces are at work: while switching costs impede the competition for high type consumers, at the same time they relax the competition for new, low type consumers. For this reason, we fix \( \theta = 1 \) (without loss of generality, since only the ratio \( \frac{\sigma}{\theta} \) matters for the comparison) and we plot the difference between total profits under \( \sigma = 0 \) and under \( \sigma = \sigma \) in a three-dimensional graph, as a function of \( x \) and \( \theta = 1 \) (Figure 1 below). It turns out that the total profits are higher without switching costs. In particular, this means that it is not in the interest of firm \( E \) to maintain high switching cost to relax competition.

**Proposition 8** Total profits are higher when \( \sigma = 0 \) than when \( \sigma = \sigma \).

**Proposition 9** Consumer surplus is higher when \( \sigma = 0 \) than when \( \sigma = \sigma \).

**Proof.** Let us consider high type consumers and low type consumers in turn. Since \( q_H \) is the same for all high type consumers, independently from their location and from the level of switching costs, their surplus depends only on the transfer they pay. If \( d^* \leq d_A \), high type consumers obtain the same level of utility in \( \sigma = 0 \) and \( \sigma = \sigma \) if they are located at \( d \leq d^* \), and a lower level (equal to 0) for \( \sigma = \sigma \) otherwise. If \( d^* > d_A \), high type consumers obtain the same level of utility in \( \sigma = 0 \) and \( \sigma = \sigma \) if they are located at \( d \leq d_A \), and a lower level otherwise. As far as low type consumers are concerned, in \([0; d^*]\) they obtain in equilibrium the same utility with \( \sigma = \sigma \) and \( \sigma = 0 \). In \([d^*; \frac{1}{2}]\), low type consumers get a rent \( \frac{\sigma^2d^2}{2} \) under \( \sigma = 0 \), while they get \( \frac{1}{2}(\frac{\sigma-x}{1-x})^2(1-d)^2 \) under \( \sigma = \sigma \). Consumers are worse off with high switching costs since in \([d^*; \frac{1}{2}]\) it is \( G(d) = \frac{1}{2}(\frac{\sigma-x}{1-x})^2(1-d)^2 - \frac{\sigma^2d^2}{2} < 0 \). Finally, in \((\frac{1}{2}; 1]\), low type consumers would get a rent \( \frac{\sigma^2(1-d)^2}{2} \) under \( \sigma = 0 \), while they get \( \frac{1}{2}(\frac{\sigma-x}{1-x})^2(1-d)^2 \) under \( \sigma = \sigma \). It is \( \frac{\sigma^2}{2} > \frac{1}{2}(\frac{\sigma-x}{1-x})^2(1-d)^2 \). To see this, consider the smallest value of \( \theta \), i.e. \( \bar{\theta} \). Since \( \frac{1}{2}(\frac{\sigma-x}{1-x})^2(1-d)^2 \) is decreasing in \( \theta \), if the inequality holds in \( \bar{\theta} = \bar{\theta} \), it would hold for any \( \bar{\theta} \) in \([\bar{\theta}; \frac{\sigma}{\theta}]\). In \( \bar{\theta} = \bar{\theta} \), the inequality collapses to \( \frac{\sigma^2}{2} > \frac{1}{2}(\frac{\sigma-x}{1-x})^2(1-d)^2 \), which is satisfied for any \( x < 1 \). \( \blacksquare \)
4.3 Comment

The main message of our results is that the presence of switching costs unambiguously benefit the incumbent firm, although this may not reflect in the market structure. More than that, the new entrant market share may be even higher in presence of switching costs, when the heterogeneity of consumers’ willingness to pay is sufficiently high.

The gain of profitability for the incumbent firm occurs at the expenses of both consumers and the new entrant. While the result concerning consumers is expected, the loss in profitability for the new entrant is the net effect of two opposing forces. On the one hand, the presence of switching costs makes it impossible, for the new entrant, to attract high type consumers; on the other hand, it reduces the intensity of competition for the new users. This is a general issue in the switching costs literature (see Farrell and Klemperer, 2007). We show that, in our set-up, the first effect always prevails. If entry costs are present, then switching costs can even deter entry although the scale at which it would occur is high. This occurs because the incumbent firm leaves to the new firm only consumers with low willingness to pay.

This result seems consistent with the available empirical evidence. In the UK, for instance, market shares are substantially equal for four main operators. However, Vodafone and (to a lesser extent) O2, the first firms to enter the market, are disproportionately serving consumers who use more intensively the mobile phone (as shown by the data on call volumes) and who have a higher willingness to pay (as suggested by data on revenues per user) (Of tel, 2008).

From a policy perspective, our model suggests that the evolution of market shares cannot be used to infer the magnitude of switching costs. Per se, the result is not surprising, if one considers that, although for different reasons, one could expect convergence in market shares both with and without switching costs. What is important is that i) the comparison is affected by the degree of consumers’ heterogeneity; ii) the new entrant market share can be higher under high switching costs (with the new entrant becoming the leader in the overall market) only if heterogeneity is present. On the last point, one can correctly claim that “leapfrogging” is rarely observed in the real world\(^7\). However, it must be noticed that our model does not include features as coverage and financial resources which constitute a further advantage for the incumbent. In a sense, we are underestimating the incumbent ability to maintain its market share, imposing symmetry on all the dimensions except the existence of an installed base. What our model predicts is that, ceteris paribus, the incentives and ability of new firms to attract customers may be either increased or decreased by the presence of switching costs.

When applied to the mobile communications service industry, our model refers to the past evolution of this industry. However, several other markets present similar characteristics, in particular the joint occurrence of switching costs, entry of new users and new firms in the market, and consumers’ heterogeneity with (imperfect) price discrimination. Typewriters, food proces-

\(^7\)One exception is Greece, where Cosmote became the market leader three years after entry.
sors (from restaurants to the consumer market), video cassette recorders (from broadcasting studios to expensive home application and then finally to the mass market) and hard disk drives are all examples of markets in which new entrants offer lower cost systems to attract consumers with lower willingness to pay and compete with incumbents that usually focus on the high end of the market.

An interesting case where we might observe new entrants in the market serving low type consumers and different degrees of switching costs is that of mini-computers considered as ‘front ends’ to mainframe systems, where ‘stand alone’ applications (with almost no switching costs) coexist with some ‘integrated’ applications which involve higher switching costs. Other similar examples can be found in the field of medical devices, machine tools and automotive applications.

5 Conclusions

In this paper we considered a model with some features characterizing the infant period of the mobile communication industry. We analyze a growing market where new consumers with a willingness to pay which is lower than old consumers’ willingness to pay appear together with a new firm. We look at the outcome of competition in presence or in absence of switching costs for the old users. Our results show that switching costs can increase or decrease the new entrant market share. The comparison crucially depends on the degree of consumers’ heterogeneity, and the new firm market share can be higher under positive switching costs only if a sufficient degree of heterogeneity is present. Independently of market shares, however, the incumbent firm benefits from switching costs in terms of profitability, with a reduction in new entrant profits and consumer surplus.

Our model presents several limitations which can be a starting point for fruitful and interesting extensions. One can think of developing a full fledged dynamic model, in which the market grows following the continuous appearance of class of consumers with lower willingness to pay, and see if the results of the model are robust to such a richer environment. Alternatively, we could model more in detail firms’ incentives to introduce new contracts, by analyzing a framework in which firms can actively invest resources to obtain information about new types of consumers. Finally, another possible venue of research is one in which switching costs are made endogenous, so that firms strategically decide how to defend their installed base of customers, while attracting new users.

References


