MPRA
Munich Personal RePEc Archive

On Quantity Competition With Switching Costs

Langus, Gregor and Lipatov, Vilen

European University Institute, Hannover University

2008

Online at https://mpra.ub.uni-muenchen.de/15457/
MPRA Paper No. 15457, posted 01 Jun 2009 07:09 UTC

On Quantity Competition with Switching Costs *

Gregor Langus
European Commission

Vilen Lipatov †
University of Hannover

May 28, 2009

Abstract

We build a simple model of quantity competition to analyze the effect of switching costs on equilibrium behavior of duopolists. We characterize the industry structure as a function of initial sales of two firms. Contrary to the literature, initial asymmetries persist in our model even though the firms are identical. When the disparity between initial sales is large, the smaller firm may become very aggressive and get more than half of the market in equilibrium. When the firms have similar initial positions, they tend to be locked in them.

JEL Classification: L11, L13

Keywords: quantity competition, switching costs

1 Introduction

Since a series of pioneering work by Klemperer [5, 6, 4] and von Weizsäcker [9] it has been widely accepted by economists that costs incurred by consumers while changing providers of goods and services play an important role in organization of industries. To list just a few aspects, switching costs affect competition intensity, attractiveness of entry, collusion possibilities, and the market structure. The costs themselves originate from different sources. Klemperer [6] identified three types: learning costs, transaction costs and artificial contractual costs.

Learning costs are the effort and time spent to reach an operating level of knowledge of special characteristics of a new product that allows the consumer to use this product with the same ease as an old one. For example,

*The authors wish to thank Jan Boone, Eric van Damme, Massimo Motta, and seminar participants in Bern, Hannover, Tel Aviv and Tilburg for comments and suggestions. The usual disclaimer applies.

†Corresponding author. Königsworther Platz 1, 30167 Hannover, Germany.
Phone: +49 511 762 48 66
Fax: +49 511 762 29 89
Email: **lipatov@mik.uni-hannover.de**

computer operating systems may be (arguably) functionally identical, but require different specific knowledge. Transaction costs arise, for example, while changing a bank account: it takes both time and effort to close one account and to open another. Contractual costs are caused by deliberate actions of firms creating cost of switching away from the current provider. This type of costs is exemplified by frequent-flyer programs. In total, it is hard to find a market in which products do not exhibit any of three types of switching costs. Farell and Klemperer [7] is a comprehensive survey that deals mainly with the effects of switching costs on competition and entry.

The economic literature identifies two effects that switching costs have on entry. On one hand, they facilitate entry, as the incumbent is less interested in new customers. Without discrimination between the old and the new customers the price will have to be lower for the whole customer base, not only for the new customers. On the other hand, switching costs facilitate entry deterrence, as the incumbent can use limit pricing more easily. In particular, in the period of entry the entrant must price significantly below the incumbent to attract new consumers.

The former effect dominates in the model of Farell and Shapiro [3]. Their demand stems from overlapping generations of buyers (in each period a cohort of young buyers enters the market and lives for two periods). On the supply side there are two sellers. In this model the firm with attached customers specializes in serving them and concedes new buyers to its rival. The switching costs lock in consumers and confer a significant market power that results in higher profits. However, these higher profits attract new entrants and may even lead to inefficiently high entry. Klemperer [5] in a two period model with a single consumer generation shows, however, that the incumbent may preempt entry by capturing a large market share or in other circumstances by keeping a small customers base to remain an aggressive competitor.

We do not consider entry explicitly, but we note that (i) switching costs make entry deterrence possible in our model; (ii) the scale of entry depends on the magnitude of switching costs.

Another problem discussed in the literature is the effect of switching costs on the competitiveness of markets. Klemperer [4] builds a two-period differentiated-products duopoly with switching costs and finds that the non-cooperative equilibrium in an oligopoly with switching costs leads to vigorous competition for market share in the early stages of the market's development. This results in the price rise from the first period to the subsequent periods, because the firms compete for market share that is valuable later. However, the prices in this model may be higher than in competition without switching costs.

Padilla [8] shows that switching costs always relax competition compared to the situation with no switching costs. However, he only considers very high switching costs with some consumers uninformed and some being replaced by the new ones. As a result, all the equilibria are mixed pricing strategy equilibria with asymmetric market shares.

Similarly, in our model the firms have asymmetric market shares in equilibrium. However, it is not the fact that the firms use mixed strategies in equilibrium that generates this result. We assume that the firms start the game with exogenously allocated customer bases that need not be the equilibrium ones. Solving the game for all such allocations we characterize the resulting equilibria. In our model the information is complete and perfect and pure strategy asymmetric equilibria exist also for a subset of initially symmetric market shares.

In an attempt to characterize industry dynamics Padilla [8] interprets the mixed strategy random equilibrium realizations of very low prices as sales or stochastic price wars. In his model, when both firms set a low price as a realization of random equilibrium strategies, price wars obtain; when only one of the firms sets a low price, unilateral sales occur. We believe that the resulting fluctuating price series that this model generates do not reflect well the observed stability of industry prices. Moreover, his model cannot explain persistent asymmetric market shares that we observe in many industries. Namely in Padilla [8] there is a persistent tendency to symmetric market shares and asymmetries will only result from randomization over strategies in equilibrium. Our model, on the other hand, captures both these features of reality, relative stability of prices, and persistent asymmetries in market shares.

The focus of our paper, however, is the short-term industry dynamics rather than long-term outcome analyzed in the most of the literature. Switching costs allow for history dependence, which plays a crucial role in the short term. We characterize period-to-period dynamics for any initial level of outputs of two firms.

Despite the focus on short term dynamics we show that convergence to the classical symmetric Cournot equilibrium does not happen even in an infinitely repeated game even if switching costs are small.

We model the industry by a one-shot game where the firms simultaneously decide on the quantities they produce. Demand is given exogenously by a linear function. [1] We are looking for subgame perfect Nash equilibria.

For very small initial output levels, including zero output for both firms, we obtain unique symmetric equilibrium, where the quantity produced increases from the initial one and lower than the quantity produced in equilibrium in the absence of switching costs. This is similar to the result of Klemperer [4], where the competition is most intensive in the initial period.

[1] We can think of this demand as being generated by a continuum of customers whose valuations are uniformly distributed on an interval. Thus, the valuation of qth customer is $p(q)$. A customer incurs fixed costs whenever he did not make a purchase from the same firm in the previous period. Thus, the costs are incurred whenever the customer switches a firm or when he first purchases the good. This is the same demand as [5].

For higher initial quantities of either of the firms we obtain also asymmetric equilibria. To the best of our knowledge, this result is not present in the literature to date.

Quantity increases are relatively less attractive with a higher customer base. That is why the initial allocations in which both firms have high outputs result in the set of equilibria where the quantities remains unchanged. These are situations in which the incentives for both firms to harvest existing customer base are stronger than the incentives for expansion.

An interpretation of these results is in the choice of entry mode when one can opt for an early entry with a limited capacity or for a later entry with a large capacity. The latter may be preferred in industries with switching costs even if after the entry capacity expansion is allowed. Namely, for large enough switching costs and captured market, the incentives for expansion are absent, and the firm might get locked in a less profitable equilibrium.

Interesting asymmetric equilibria obtain when either one or both of the firms have an initially allocated output in the medium range (we shall characterize medium range more precisely later). Each of the firms, given rival's initial output wants to increase the output - future profit increases are attractive. However, if the rival increases the quantity largely enough, lowering the price further, the firm no longer wants to increase the output and prefers to keep a high present price and harvest existing customers. This results in an asymmetric equilibrium where one of the firms ends up bigger, and the other does not change its output.

We consider industry dynamics in telecommunications of 6 European countries to illustrate our findings. The data are supportive of our general prediction that more asymmetric firms tend to symmetry more.

The quantitative results of our analysis survive in a multi-period setting. Thus, our model does not predict convergence to symmetric output allocations over time.

The rest of the paper is structured as follows. In section 2 we formulate and solve the model, in section 3 we discuss comparative statics, in section 4 - entry. Section 5 is devoted to short-term industrial dynamics, section 6 - to the extension of our model to multiple periods, section 7 - to the implications of our results.

2 The Model

We consider a one-shot Cournot game with two firms, demand $p(q)$ and production costs $C(q)$. Switching costs of changing a provider are s. The suppliers cannot discriminate between different consumers according to whether or not they have made the purchase in the previous period. Thus, whenever the sellers want to expand the sales they have to offer a discount to all consumers. The formulation of demand is identical to [5]. We add initial sales to the model, whereby the firms start playing the game with some history, which proxies the customer base of the firm. The maximization problem of a seller is thus

$$\max\ \pi_i = \left(p\left(q_i^2 + q_j^2\right) - s\mathbb{I}_i\right) q_i^2 - C\left(q_i^2\right), \tag{1}$$

where

$$\mathbb{I}_i = \begin{cases} 1, & q_i^2 > q_i^1; \\ 0, & q_i^2 \le q_i^1. \end{cases}$$

$\mathbb{I}_i$ captures the discount when the seller wants to increase sales from that of previous period and q_i^1 denotes the volume of initial sales. The initial sales are treated as exogenous. To be able to obtain analytical results we look at the linear demand $p(q) = a - bq$ and linear costs $C(q) = cq$.

$$\pi_i = \left(a - b\left(q_i^2 + q_j^2\right) - s\mathbb{I}_i\right) q_i^2 - cq_i^2, \tag{2}$$

Denote for convenience $x = \frac{a-c}{3b}$ and $S = \frac{s}{3b}$. Next, fix the strategy of firm A to q_A^2. The best response of the firm B given its initial sales q_B^1 is to maximize

$$\pi_B\left(q_B^2 | q_B^1, q_A^2\right) = \begin{cases} b\left(3x - \left(q_B^2 + q_A^2\right)\right) q_B^2 & q_B^2 \le q_B^1; \\ b\left(3x - \left(q_B^2 + q_A^2\right) - 3S\right) q_B^2 & q_B^2 > q_B^1. \end{cases} \tag{3}$$

The problem is concave in q_B^1 on each interval, so we can find optima separately and then compare them. Differentiation gives

$$\pi'_B\left(q_B^2 | q_B^1, q_A^2\right) = \begin{cases} b\left(3x - \left(2q_B^2 + q_A^2\right)\right), & q_B^2 \le q_B^1; \\ b\left(3x - \left(2q_B^2 + q_A^2\right) - 3S\right), & q_B^2 > q_B^1. \end{cases} \tag{4}$$

From equation above we obtain candidate best responses and rewrite them as

$$q_B^2 = \begin{cases} \frac{1}{2}\left(3x - q_A\right), & q_B^2 \le q_B^1; \\ \frac{1}{2}\left(3x - 3S - q_A\right), & q_B^2 > q_B^1. \end{cases} \tag{5}$$

This condition simply states that in the second period the seller B, when he is expanding the quantity will, given strategy of A, expand to $\frac{1}{2}(3x - 3S - q_A)$. When B is contracting sales, given strategy of A, he will set the quantity to $\frac{1}{2}(3x - q_A)$. After plugging the corresponding expressions for the second period quantities into the conditions and the realizing that in the remaining interval B responds with no change in quantity, $q_B^2 = q_B^1$, we obtain:

$$q_B^2 = \begin{cases} \frac{1}{2}\left(3x - q_A\right), & \frac{1}{2}\left(3x - q_A\right) \le q_B; \\ q_B^1, & \frac{1}{2}\left(3x - 3S - q_A\right) \le q_B < \frac{1}{2}\left(3x - q_A\right); \\ \frac{1}{2}\left(3x - 3S - q_A\right), & \frac{1}{2}\left(3x - 3S - q_A\right) > q_B. \end{cases} \tag{6}$$

We have dropped the superscript for convenience, as now only initial sales are present in the rhs.

The part of the best response function that is relevant for quantity increase is computed under the assumption that it is optimal for a firm to increase the quantity. However, it may not be so for output values close to the best response. The firm would in that case prefer not to raise the quantity, because of the penalized price which it obtains in doing so. Therefore, we compute the set of initial allocations for which the firm is indifferent between increasing the quantity to best response and keeping it as it was before, and then define the global best response for the second period as

$$q_B^2 = \begin{cases} \frac{1}{2}\left(3x - q_A\right), & \frac{1}{2}\left(3x - q_A\right) \leq q_B; \\ q_B, & y_B \leq q_B < \frac{1}{2}\left(3x - q_A\right); \\ \frac{1}{2}\left(3x - 3S - q_A\right), & q_B < y_B. \end{cases} \tag{7}$$

Here

$$y_i = \frac{1}{2}(3x - q_j - \sqrt{3S}\sqrt{-3S + 6x - 2q_j}),$$

is the curve which characterizes the initial q_i for each strategy q_j for which the firm i is indifferent between increasing the quantity and not changing it from the one initially allocated. In Figure 1 these are plotted as dashed convex curves.

It proves useful to take on the following notation. First, define

$$z = \frac{3}{25}(2S + 5x - 2\sqrt{6S}\sqrt{-4S + 5x}), \tag{8}$$

as q_i coordinate of the intersection of the higher best response line for firm j with the indifference set for firm i, y_i. In the figure this is denoted by dashed horizontal line. Moreover,

$$\phi = \frac{1}{2}(S + 2x - \sqrt{3}\sqrt{S}\sqrt{-S + 4x}) \tag{9}$$

is the q_i coordinate of the projection of the intersection of lower best responses $(x - S, x - S)$ on y_i. Finally,

$$\nu = \frac{1}{3}(-7S + 3x + 2\sqrt{S^2 + 3Sx}) \tag{10}$$

is the q_i coordinate of the intersection of lower best response of firm j with i's indifference set y_i.

We proceed to find the equilibria and characterize them in the following propositions. We shall assume, without loss of generality, that firm B never has higher initial sales than firm A.

Proposition 1. The unique Nash equilibrium of the game specified above is characterized by the following strategy profiles:

(i)

$$(x-S, x-S) \quad \text{if} \quad q_{A,B} \le \nu, \tag{11}$$
$$(x, x) \quad \text{if} \quad q_{A,B} \ge x. \tag{12}$$

(ii)

$$(q_A, \frac{3}{2}(x-S) - \frac{1}{2}q_A) \text{ if}$$
$$((q_B \le y_B) \cap (y_A \le q_A \le x+S)) \cup ((\phi \le q_A \le y_A) \cap (q_B \le \nu)), \tag{13}$$

$$(x+S, x-2S) \quad \text{if} \quad (q_A \ge x+S) \cap ((q_B \le z) \cup (q_B \le y_B)), \tag{14}$$
$$(\frac{3}{2}x - \frac{1}{2}q_B, q_B) \quad \text{if} \quad (q_A \ge \frac{3}{2}x - \frac{1}{2}q_B) \cap (x \ge q_B \ge z); \tag{15}$$

(iii)

$$(q_A, q_B) \quad \text{if} \quad (q_B \ge y_B) \cap (q_A \le \frac{3}{2}x - \frac{1}{2}q_B), \tag{16}$$

Proof. We construct the equilibrium from intersection of global best responses, as outlined by (7). ■

Equations (11) and (12) characterize the two symmetric equilibria denoted, respectively, by letters A and C in Figure 1. The first equilibrium results from low, including 0, initial sales for both sellers. In Figure 1 this initial allocation corresponds to the white area under the diagonal close to the origin. Both firms increase the quantity but total sales in the resulting equilibrium are low. Any other equilibrium in our model is characterized by higher total sales. The second symmetric equilibrium (denoted by C in the figure) results from both firms selling large volume in the previous period. This area of initial sales volumes is above the horizontal dashed line through C. In this equilibrium both firms decrease the quantity to the level of the equilibrium without switching costs. This is also the equilibrium where total quantity sold is the highest.

Equations (13)- (15) characterize equilibria in which the initially smaller firm (weakly) increases the quantity and the bigger one (weakly) decreases it. This type of equilibria results when the asymmetry in initial sales is large and the the larger firm A's sales exceed the threshold defined by y_A. Equation (13) thus characterizes the unique asymmetric equilibrium which in Figure 1 corresponds to the area (13) below the curve y_b to the left of $x+S$.

Equation (14) in turn characterizes the equilibrium resulting from the bigger of the firms inheriting large sales (in Figure 1 this means that A has sales beyond $\tilde{q}_A$), whereas the smaller firm had much smaller sales(B

had initial sales below the dashed indifference curve or below the dashed horizontal line denoted by z). In equilibrium, the bigger firm will decrease its sales volume whereas the smaller one will increase it moderately. The equilibrium in the Figure is now at the intersection of best response lines, denoted by B.

Equation (15) gives equilibrium sales volumes for initial allocations which in Figure 1 fall into the region to the right of the higher of the best response lines for firm A and between the horizontal lines through C and z. In this case the large firm, A, will decrease the quantity the other firm will not change sales.

Equation (16) characterizes equilibrium resulting from levels of initial sales in the medium range. In this case none of the firms has an incentive neither to increase nor to decrease its sales from the initial ones. In the figure this set is represented by the grey central area. Clearly for relatively high levels of initial sales the opportunity costs of expansion are high for both firms and none of them has an incentive to increase sales.

As we have shown, at very low, including zero quantities in the first period there is only a symmetric equilibrium where both firms increase sales (the white area below the 45 degrees line close to the origin in the figure).

However, for a set of initial allocations where both firms still have relatively low, but at least one of the firms has initial sales larger than ν, multiple equilibria may obtain. This leads us to the following proposition.

Proposition 2. The multiple Nash equilibria of the game specified above are characterized by the following strategy profiles:

$$\begin{aligned} &(q_A, \frac{3}{2}(x-S) - \frac{1}{2}q_A) \quad \text{together with} \\ &(x-S, x-S) \quad \text{if} \quad (\phi \geq q_A \geq \nu) \cap (q_B \leq \nu), \end{aligned} \tag{17}$$

$$\begin{aligned} &(q_A, \frac{3}{2}(x-S) - \frac{1}{2}q_A) \quad \text{together with} \\ &(x-S, x-S) \quad \text{and} \\ &(\frac{3}{2}(x-S) - \frac{1}{2}q_B, q_B) \quad \text{if} \quad (\phi \geq q_A \geq \nu) \cap (\phi \geq q_B \geq \nu), \end{aligned} \tag{18}$$

$$\begin{aligned} &(q_A, \frac{3}{2}(x-S) - \frac{1}{2}q_A) \quad \text{together with} \\ &(\frac{3}{2}(x-S) - \frac{1}{2}q_B, q_B) \quad \text{if} \quad (y_A \geq q_A \geq \phi) \cap (\nu \leq q_B). \end{aligned} \tag{19}$$

Proof. Analogous to Proposition 1. ■

Multiple equilibria arise because of the interaction between the strategies played by the other player and incentives to increase the sales. A relatively large increase in sales by one of the players may cause the other player to be better off not changing its sales from the initial ones. For the set of initial sales which give multiple equilibria both firms are potentially interested in increasing sales, and at least one does so. If both firms indeed increase sales, this leads to a symmetric equilibrium. The larger, in our case firm A, however, has an incentive to increase sales only as long as B does not choose a large increase in sales. As A's customers base is no longer very small it becomes optimal not to increase the quantity for large increases in B's quantity. In turn, large increase, as a response to a strategy of no change of A for this strategy of B, becomes attractive for B. These strategic interactions imply an additional asymmetric equilibrium in conjunction with the symmetric one.

The first set of multiple equilibria which result from firm B being initially significantly smaller than A is characterized by (17), which can also be seen from the Figure 1. It is obvious that either the firm B will be bigger in equilibrium or both firms will have equal sales volumes at $x - S$.

If we move initial sales of B to the levels close to those of A we have 3 possible equilibria - where either A or B has a higher output and a symmetric equilibrium with both firms having equal outputs. In Figure 1 this set of initial sales volumes is denoted by (18). The resulting equilibria are characterized by the corresponding equation.

There is also a possibility of two asymmetric equilibria when sales volumes of the firms in the initial period are close. In the figure this set is the region (19). The resulting equilibria are characterized by the corresponding equation.

In line with the literature on the switching costs, the firm with a smaller initial market share is relatively more aggressive. The reason is that the larger firm has greater incentives to exploit its customer base and thus lacks incentives for costly expansion. In the present model, however, we can trace the adverse effect of aggressive strategies on the expansive intentions of the other player and obtain asymmetric equilibria, even when the firms are completely symmetric along all dimensions.

As the propositions make clear, equilibrium quantities depend on the initial allocation of output between firms in the presence of switching costs. The outcomes are sometimes sensitive to small changes in the initial conditions. This sensitivity is reflected in the abrupt changes of the equilibrium quantities for small changes in initial sales volumes of one or both firms. Together with possible multiple equilibria, this implies that an attempt at prediction of the industry structure outcomes in reality with switching costs may not be a very fruitful operation. This has been a recurrent, but never satisfactorily explained argument in the literature on switching costs.

We have shown that for a one period model asymmetric equilibria will result for a subset of asymmetric (and a subset of symmetric) initial sales allocations for otherwise identical firms. In the presence of switching costs this is a normal competitive outcome, which need not be a red flag for the

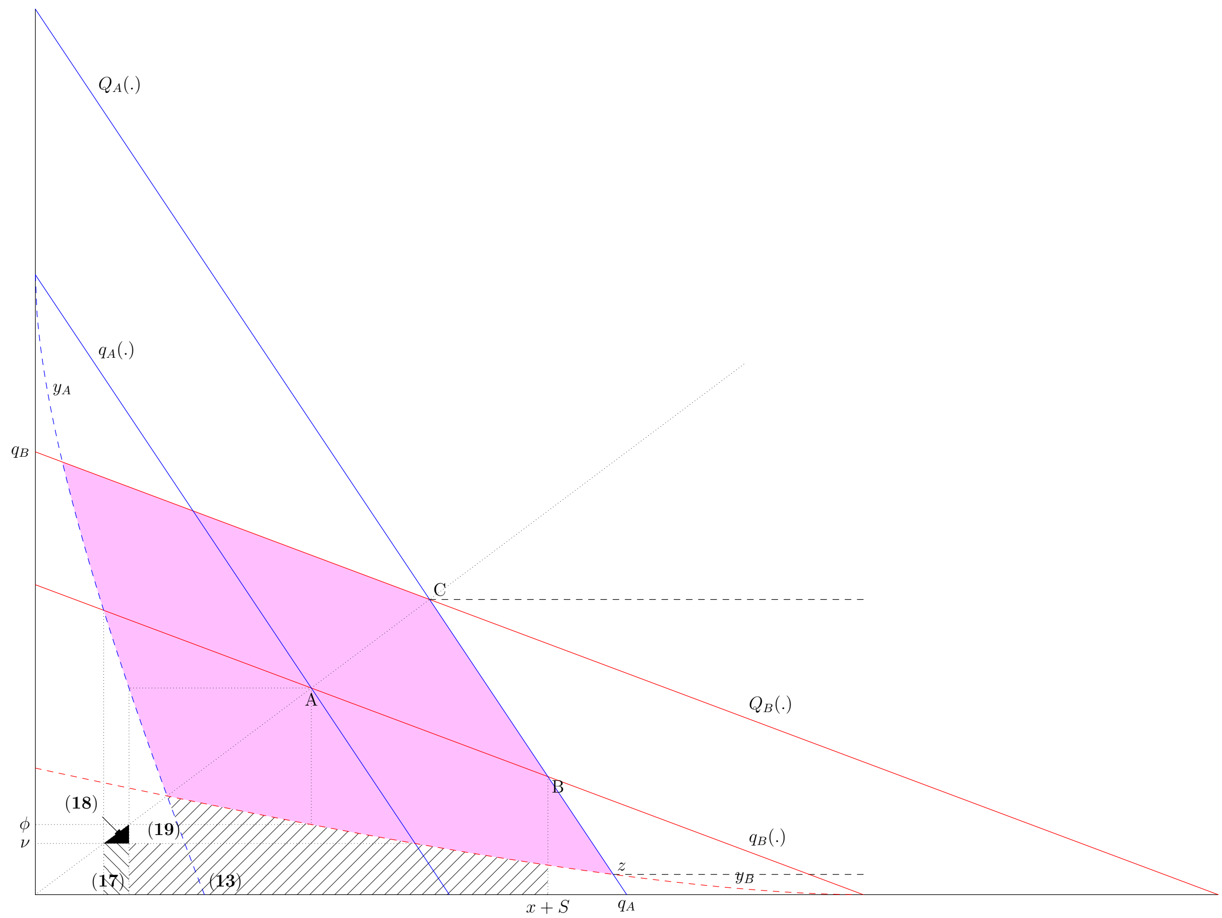

Figure 1: Equilibria

antitrust authority. This is relevant, particularly because we often observe persistent asymmetries in market shares in reality and this seems to often be a great concern for a regulator or a competition authority.

Further, even if a firm has larger sales (larger customers base) initially, this may not be true in equilibrium. In our model often it is the initially smaller firm, which is more aggressive, that has higher equilibrium sales. Taking this result to reality, we should not be surprised if such industries are exhibiting occasional volatile changes in leadership. Moreover, the result should serve as a warning for the regulator from hastily accepting a paternalistic attitude towards the small firms in industries characterized by switching costs.

3 Comparative statics

It is clear that the conventional Cournot duopoly is a limiting case in our model when switching costs tend to zero. The grey area of inaction on the Figure is growing larger with increase of costs s. This is very intuitive: none of the firms wants to adjust its position if the adjustment is costly. Notice that for very high costs there is no initial position that makes firms increase their sells even from zero - in such case entry is successfully blocked.

The size of the market a obviously has the opposite effect on the region where the firms do not change their positions in equilibrium. The slope of demand function b matters for this region in so far as it enters x and S, higher slope thus leading to smaller set of inaction. This also seems intuitive, as more elastic demand is more attractive for price cuts holding costs of switching constant.

Note also that the upper-right border of the grey region have the slope $-\frac{1}{2}$ and -2 regardless of the parameters of the model. Size of the market, elasticity of demand, production and switching costs all change position and size of the area of inaction, but do not change its form. This feature is a result of our assumption that the two firms are identical apart from initial positions.

The size of the region with multiple equilibria depends on how large is $\phi - \nu$. It can be shown that this difference is increasing in S and decreasing in x. Hence, the effect of switching costs and other parameters on this region is similar to that on grey region.

4 Entry

Given that we have solved for all the initial allocations of consumers, we can use the results to examine entry into an industry characterized by switching costs. The entrant that does not face any sunk costs is equivalent to an incumbent which has no initial sales. Thus, Propositions 1 and 2 allow us to characterize the resulting equilibrium for any strictly positive initial sales of the incumbent, firm A. Namely, equations (11), (13), (14), (17) contain all the relevant information for entry analysis.

We start the discussion with monopolistic initial sales of the incumbent. We shall consider as monopolistic the points at zero production of the firm B between lower best response of firm A and most asymmetric equilibrium quantity of firm A, $q_A^m \in \left[\frac{3}{2}(x-S), x+S\right]$. In such case the entry will involve the entrant setting the quantity equal to $\frac{3}{2}(x-S) - \frac{1}{2}q_A^m$, which is the best response to monopolistic quantity q_A^m. This can also be seen in equation (13). The incumbent will not change its sales from the initial ones. The market share of the incumbent in the new equilibrium will be, under the assumptions we made above, somewhat higher than that of the entrant. The difference depends on how steep the best response is.

When we consider the initial sales of the incumbent below monopolistic, the entrant's equilibrium sales are higher. This is according to the same best response defined in (13). When the incumbent's initial sales are below $x-S$, the entrant's equilibrium market share is actually higher than that of the incumbent.

However, if we consider very small initial sales of the incumbent ($q_A \leq y_A$), the industry will exhibit symmetric sales $(x-S, x-S)$. This corresponds to equation (11) of Proposition 1. As a qualification, there is also a small interval $q_A \in [\nu, \phi]$ that results in two equilibria: symmetric $(x-S, x-S)$ and asymmetric $\left(q_A, \frac{3}{2}(x-S) - \frac{1}{2}q_A\right)$. This can be seen from equation (17) of proposition 2.

For larger than monopolistic initial sales of the incumbent, $q_A > x+S$, its equilibrium sales decrease. Despite this, the asymmetry in this case is maximal: the equilibrium is $(x+S, x-2S)$, as can be seen from equation (14).

5 Dynamics

Our results can be applied to get some insights into the adjustment of market structure to demand shocks. Initial sales in our model can be interpreted as the equilibrium sales in the previous period characterized by initial demand. Suppose now between the periods a demand shock (symmetric or asymmetric) is realized, such that the new demand is as in our model. In this manner low initial sales allocations (those close to the origin in Figure 1) correspond to a positive shock in demand and the initial allocations with high sales correspond to negative demand shocks. Thus we can choose any initial state and analyze the adjustment to shocks.

Similarly, evolving industries and growing markets exhibit large potential size, and this corresponds in the model to initial allocations at low sales close to the origin of the graph (increasing the constant term in the demand function would have exactly such an effect). On the other hand, in the model shrinking markets would exhibit small potential size and accordingly initial allocations further from the origin.

With this interpretation we can explore the implications of the model for industry dynamics. As shown in Proposition 1, relatively low initial sales and significant asymmetries in these give rise, in equilibrium, to large changes in sales by at least one of the sellers. Thus, it follows from the model that we should not be surprised to observe sudden shifts in the sales leadership in growing markets or after positive demand shocks. On the other hand, such reversals would be less likely for industries where sellers are operating in stagnating markets. The set of equilibria characterized by the proposition above also imply that these are the situations in which persistent asymmetries in market shares are more likely.

In the markets with small potential size our model predicts that large initial asymmetries will decrease in equilibrium through the smaller firm increasing its sales faster than the bigger one. Once a firm locks in sufficiently high a customers base the incentives for further expansion are low and the model predicts convergence to stable market shares. Note that the model does not predict that small asymmetries will decrease in such markets.

When the initial sales allocations are rather symmetric and shocks to the demand are small, the model implies that neither of the firms will be changing the level of sales in equilibrium. In this region the incentives of sellers to increase market share are weak. These initial allocations can be interpreted as historical customer bases for mature markets along the same lines as before. Thus in industries (markets) which are growing at slower rates the model predicts more stable symmetric or asymmetric market shares.

There are at least two testable hypotheses that come out of our analysis. Firstly, we are more likely to find alternating leadership in the growing industries with switching costs. Secondly, we should see stabilized market shares (symmetric or asymmetric) in mature industries.

We look at the data on the dynamics of market shares in telecommunications industry, where switching costs are substantial. In German mobile telecom [2], the market shares of two leading providers remained stable during 2000-07. They ranged 36.7% - 41.6% for T-Mobile and 34.7% - 40.0% for Vodafone. In terms of our model, initially symmetric distribution of the market shares remains stable.

In Austria 2000-06 [3] the market shares of two leading operators (Mobilkom and T-Mobile) were consistently declining from joint 88.4% to joint 63.9%. Correspondingly, the share of other operators has grown from 11.6% to 36.1%, consistently with our model featuring large initial asymmetries that are reduced over time. The same is true for the Netherlands over 1999

[2] available at http://www.bundesnetzagentur.de/

[3] available at http://www.rtr.at/de/tk/Marktinfos

- 2003, Norway 1993- 2006 [2], the UK 2001 - 05 [1]. A similar story can be told about Italy [4], where TIM had the whole market in 1995, and by 2003 TIM, Vodafone and Wind had respectively 45%, 35% and 15% of the market.

Certainly, these patterns can not serve as a solid evidence in favor of our model results. A comprehensive econometric model is needed to disentangle the effects of switching costs and multiple other factors that affect competition. However, the features observed in the data are indicative of relevance of our conclusions.

6 Extension to multiple periods

In this section we extent our model to multiple (in fact, infinitely many) periods to see if it is robust to such a modification. In general formulation, the optimization problem of the firm A in an infinitely repeated Cournot game with switching costs is

$$V\left(q_A, q_B\right) = \max_{q'_A} \left\{\pi\left(q'_A, q'_B \| q_A, q_B\right) + \delta V\left(q'_A, q'_B\right)\right\},$$

s.t.

$$\begin{aligned} p = P\left(q'_A + q'_B\right) - s \quad &\text{if} \quad q'_A > q_A \\ p = P\left(q'_A + q'_B\right) \quad &\text{if} \quad q'_A \leq q_A \end{aligned}$$

Now our candidate equilibrium is to move to a pair of output (q^*_A, q^*_B)in the first period. Suppose we start with initial vector $q^0 \ll q^*$. We consider a unilateral deviation of first moving to some quantity $q_i \| q^0_i < q_i < q^*_i$. Then the corresponding values are

$$V^* = \frac{P\left(q^*_A + q^*_B\right) q^*_A}{1-\delta} - s q^*_A$$

and

$$V^{dev} = P\left(q_A + q^*_B\right) q_A + \frac{\delta P\left(q^*_A + q^*_B\right) q^*_A}{1-\delta} - s\left(q_A + \delta q^*_A\right).$$

The firm A will prefer not to deviate (and hence change production only once), if

$$V^* - V^{dev} = P\left(q^*_A + q^*_B\right) q^*_A - P\left(q_A + q^*_B\right) q_A + s\left(q_A + (\delta - 1)\, q^*_A\right) \geq 0.$$

Intuitively, the inverse demand should not react too drastically to the reduction of quantity. In case of linear demand we have

$$\left(q^*_A - q_A\right)\left(a - b\left(q_A + q^*_A + q^*_B\right)\right) + s\left(q_A + (\delta - 1)\, q^*_A\right) \geq 0.$$

[4]available at http://www.group.abnamro.com/index.cfm

Clearly, the first term is most likely to be negative, if $q_A \to q_A^*$, in which case we have $a - b\,(2q_A^* + q_B^*)$.

But this is just our demand function, the supremum of the argument is 3 times Cournot quantity, so the infinum of the function is exactly zero. Thus, the first term is always positive. The second term is positive, if $q_A > (1-\delta)\,q_A^*$. Note though that at in the opposite case (small quantities) the first term becomes large: $(\delta q_A^*)\,(a - b\,((2-\delta)\,q_A^* + q_B^*))$. Taken at the extreme, we have $q_A^*\,(a - b\,(q_A^* + q_B^*)) + s\,((\delta - 1)\,q_A^*) \geq 0$ meaning that frictionless Cournot profit should exceed switching costs, which is obviously satisfied if the market is to exhibit any changes in quantities at all.

In effect, with linear demand our candidate equilibrium brings larger value than deviation [5].

The fact that there exists a Markov perfect Nash equilibrium where the firms only move once allows us to compute the regions of initial allocations for which a firm will not change its output in the same fashion as for the one-shot game. In fact, the shape of these regions turns out to be very similar, except that the set is smaller, but not empty for $\delta > 0$. For $\delta = 1$ obviously this set is empty and the only Markov perfect equilibrium is the Cournot equilibrium of the frictionless game.

For any $0 < \delta < 1$ we can thus perform an analysis similar to the one-shot game above to find both symmetric and asymmetric equilibria, analogously as in the one-shot game.

Thus the qualitative results of the model persist when we extend the number of periods (even when we consider infinitely many periods with the discount factor below 1) and restrict ourselves to the simplest equilibrium concept consistent with rational behavior in an infinitely repeated game setting.

7 Discussion of the results

Our simple Cournot model shows that in the presence of switching costs equilibrium allocation depends on the initial allocation. The initial allocation in this model can be interpreted as the firms' market shares relative to potential demand. Thus, initial allocations close to the origin of the graph correspond to situations where the market has significant potential for growth, and the allocations where both firms have large initial sales corresponds to a situation in which market is shrinking. In this view a sudden shock, say increase in expected market size, could induce a change in relative market shares if it is large enough. This response could lead to a reversal in the order of market share sizes. One implication of the model is that the adjustment to shocks in demand is hard to predict and may involve sudden shifts in market

[5] It is standard to show that the same is true for a deviation in any other period and for deviations in multiple periods.

positions of the firms. Industries exhibiting persistent asymmetries in market shares, periods of relatively stable division of market followed by sudden readjustments or longer periods of symmetric market division would all be consistent with the presence of switching costs and imperfect differentiation between old and new customers, as in our model.

Entry decision can also be analyzed in our model. In new industries with a large growth potential the model would predict a relatively symmetric market shares after the entry, as the entrant holds large sales upon entry and the incumbent does not fight aggressively for a market share. At the other extreme an entry to a shrinking monopolized market would result in a relatively asymmetric market allocation, despite the fact that the incumbent is even less aggressive in such a case.

Recently a theory of the stepping stones, or the ladder of investment theory, has become prominent in the literature and among regulators of some industries (telecommunication) where the cost of initial investment into infrastructure are high. The idea of the ladder of investment is that an entrant be given access to the infrastructure of the incumbent so he can build a customer base, which would then justify investment in own infrastructure. If the access to infrastructure is limited so that initially the entrant can not supply the whole market the model predicts that it could easily happen that after the entrant has captured a significant customers base it may lose the incentive to increase sales further and with it the incentive to invest in own infrastructure, thus defeating the purpose of the ladder idea. The entrant would invest in large infrastructure capacity in the absence of the ladder, but after capturing a significant customer base it may no longer be optimal for it to build his own infrastructure.

8 Conclusion

Our analysis in this paper is centered around one basic feature of reality: history dependence. In our simple Cournot setup history matters because the customers have to incur switching costs whenever they buy from a new seller. We are able to characterize equilibrium of Cournot game for any initial allocation of sales. Our main finding is persisting asymmetry in market shares of otherwise identical firms. This result survives extending the model to multiple periods, including an infinitely repeated game.

We also show that when initial asymmetries are small, they tend to remain small, as none of the firms is motivated to behave aggressively. When initial asymmetries are large, the smaller firm has an incentive to expand, and sometimes it does so to the extent that it takes more than half of the market. This gives us empirically testable hypotheses of stable market shares in the markets with uniform distribution of market shares and high volatility in the markets with very uneven distribution of market shares.

Our model also provides rationale for a large-scale one-time entry versus gradual buildup of capacities. The intuition remains intact: with switching costs a fresh entrant is the one who has "nothing to lose" and is relatively more aggressive than a seller with an established customer base.

Linear demand and homogenous good framework are the main limitations of our model. However, different demand functions do not change the nature of competition, so we do not expect our qualitative results to be altered significantly. Heterogenous goods framework would be an interesting extension to our analysis, adding new channels for switching costs to work through. At the same time, the main effects of customer lock-in outlined here will remain on its place.

The analysis presented is general and can be applied to any industries characterized by switching costs. Telecommunications, banking, airlines are among classical examples of such industries.

References

[1] The communications market 2006. page 138, 2006.

[2] The norwegian electronic telecommunications market. page 29, 2007.

[3] Joseph Farrell and Carl Shapiro. Dynamic competition with switching costs. *The RAND Journal of Economics*, 19(1):123–137, 1988.

[4] Paul Klemperer. The competitiveness of markets with switching costs. *The RAND Journal of Economics*, 18(1):138–150, 1987.

[5] Paul Klemperer. Entry deterrence in markets with consumer switching costs. *The Economic Journal*, 97:99–117, 1987.

[6] Paul Klemperer. Markets with consumer switching costs. *The Quarterly Journal of Economics*, 102(2):375–394, may 1987.

[7] Paul Klemperer and Joseph Farrell. Coordination and Lock-In: Competition with Switching Costs and Network Effects. *working paper*, 2006.

[8] A. Jorge Padilla. Mixed pricing in oligopoly with consumer switching costs. *International Journal of Industrial Organisation*, 10:393–411, 1992.

[9] C.C. von Weizsäcker. The Cost of Substitution. *Econometrica*, 52:1085–1116, 1984.