Estimating the lock-in effects of switching costs from firm-level data*

Gábor Kézdi† and Gergely Csorba‡

November 30, 2009

Abstract

This paper proposes a method to identify and quantify the lock-in effects of switching costs using firm-level data. The method compares the behavior of already contracted consumers to the behavior of new consumers, as the latter are not subject to switching costs. In two panel regressions on new and quitting consumers, we look at the differential response to firm-specific price changes, and identify the effect of switching costs from the difference between the two. We apply our method to the Hungarian personal loan market and find significant switching costs: the estimated reaction by a bank's old consumers is around 70 percent weaker than the estimated reaction of new consumers to the market.

Keywords: switching costs, panel data estimations
JEL codes: D12, L14

*The methodology presented in this paper was developed as part of the 2007-2008 banking sector inquiry of the Hungarian Competition Authority (GVH), in which the first author worked as an expert. We would like to thank Dávid Farkas for excellent research assistance, and Jacques Crémer, Timothy Hannan, Gábor Koltay, Surd Kováts, Balázs Muraközy, Konrad Stahl, and seminar participants at the 2008 EARIE Conference (Toulouse), the Mannheim Competition Policy Forum and the MTA-KTI Seminar (Budapest) for their helpful comments. All remaining errors are ours. The views expressed in this paper are not purported to represent those of the GVH.

†Central European University and Institute of Economics at the Hungarian Academy of Sciences. kezdig@ceu.hu

‡Hungarian Competition Authority (GVH) and Institute of Economics at the Hungarian Academy of Sciences. csorba.gergely@gvh.hu

1 Introduction

Switching costs can increase firms' market power and might allow for ex-post price increases. They can also contribute to barriers of expansion for rival firms, which might help formerly dominant firms to conserve their strong market positions, especially if they are able to increase these costs. As a result, switching costs are of important concern for competition authorities. The significant theoretical literature is summarized in the thorough review of Farrell and Klemperer (2007). At the same time, empirical results are scarce, primarily because identifying and quantifying switching costs are difficult. Consumer-level data on potential switchers are rare although they are probably best suited for such analysis. Firm-level aggregates are more frequent to encounter, and most regulatory bodies have the legislative power to acquire them. However, few empirical methods have been developed so far for such aggregates, and they rest on specific model assumptions in order to estimate the magnitude of switching costs (see, e.g. Shy, 2002, Chen and Hitt, 2002, Kim, Kliger and Vale, 2003).

This paper proposes a simple method for estimating the lock-in effects of switching costs from firm-level data. We focus on the behavioral effects that are most important for competition policy. We stay within the framework of demand analysis and avoid structural assumptions on market structure or type of competition. As a result, the lock-in effects of switching costs can be identified in a robust way using firm-level data.

Our central idea is that the behavior of new consumers (those making their first purchase decision on the market) describes the behavior of old consumers (those already engaged in some contractual relationship with a specific firm) in the absence of switching costs. Therefore, the behavior of new consumers can be used as the counterfactual for the behavior of old consumers. The lock-in effect of switching costs is identified by the difference of the effect of the price increase on new consumers versus old consumers.

Estimation requires firm-level data that is only slightly more detailed than usual. The data should include prices and two quantities: the number of consumers who are new to each firm, and the number of consumers who leave each firm but used to buy from the firm. Such data are usually available in markets with long-term contracts such as consumer credits, utilities, telecommunication and media services (these are exactly those liberalized network industries where the competition-hindering effect of switching costs is usually feared). Given the appropriate data, our estimation strategy does not use any explicit assumptions on market structure and equilibrium, which was our practical motivation in developing this kind of methodology. We illustrate that our reduced form-model has strong links to a more structural discrete choice model, as well as the classical multi-period competition model with switching costs by Beggs an Klemperer (1992). Our approach is similar in its spirit to a recent paper by Schiraldi (2009) who uses individual data in order to estimate transaction costs in the Italian car market. His approach is based on a comparison of the share of consumers holding a car to the share of consumers buying the same car. Our model is also in line with a reduced-form empirical method presented by National Economic Research Associates (2003, Appendix B, Section 3): if in a homogenous good industry a firm is able to raise prices without losing significant sales to competitors (i.e. the estimated price-cross elasticity is small), then this fact can be attributed to switching costs.

The estimation model is a system of two panel regressions, estimated in first differences. The left-hand side variable of the first equation is the change in the market share in terms of consumers who are new to the market. It can be interpreted as the change in the probability of new consumer choosing the firm. The left-hand side variable in the second equation is the change in the share of old consumers who remain loyal to their previously chosen firm. This second variable can be interpreted as the change in the probability of staying loyal to the firm. The right-hand side variables are the same in the two regressions, and they contain price changes. The lock-in effects of switching costs are measured by the comparison of the two effects of the same price change. It can be interpreted as the fraction of old consumers who were prevented from switching but would have changed the firm they choose if they were new consumers.

In applications, however, consumers new on the market and old consumers switching to other firms are not observed. Instead, the number of consumers joining or leaving a specific firm can often be observed, and we use these measures to construct proxy variables. We derive the conditions under which using the proxies identify our parameters of interest, and we address the potential bias if those conditions are not satisfied.

As an illustration, we apply the measurement model to the market of personal loans in Hungary. Our results indicate significant switching costs on the Hungarian personal loan market. According to our estimates, one percentage point increase in prices leads to a between 0.43 and 0.61 percentage point decrease in demand among new consumers, but only a 0.13 decrease among the bank's old consumers. Old consumers' responsiveness is therefore between 0.32 and 0.48 percentage points, or between 70 and 79 per cent lower than new consumers' responsiveness. These findings imply substantial market power for banks, especially for increasing some fees ex-post.

Our results are comparable to the more structural estimates of switching costs in markets of household financial services. Shy (2002) derives a static model in which firms' prices are set at given switching costs such that there are nobody has any incentives to undercut their rivals, but this model predicts by definition no switching and stable market shares. As an illustration, he examines the Israeli cellular phone market (estimated switching costs are 35-50% of average price) and the Finnish bank deposit market (estimated switching costs vary between 0 and 11% of the average balance). Kim et al (2003) model the consumers' transitions and banks' intertemporal decision making in a dynamic framework and then apply it to the Norwegian loan market, and estimate switching costs for a bit more than 4% of the average loan's value. Both of these papers trace back switching costs from the fact that their presence affects firms' equilibrium pricing decisions, while we take firms' behavior as given and apply robust fixed-effects panel methods to recover demand responses. Moreover, our results provide direct measures for the lock-in effects of switching costs, which is often the object of interest for competition policy.

2 Underlying modelling framework

2.1 Assumptions on consumer choice

In order to bring our thought experiment closer to empirical applications, assume that consumers are heterogeneous. Consider a randomly taken consumer, N, who is new to the market in period t, and let n_j denote the probability that she buys the product from firm j under existing prices. Consider another consumer, O, who is randomly chosen from the old consumers of firm j (that is she bought from firm j in period $t-1$), and let l_j denote the probability of her consumer to stay loyal to firm j.

Now consider a small increase in p_j, the price of firm j. If consumer N had bought the product from firm j, the price increase may make her change her mind and buy the product from another rim k. The probability n_j that she would buy from firm j is therefore decreased. Responding to the same price increase, consumer O might also change his mind, therefore the probability l_j that he would stay loyal with firm j may decrease. However, if switching costs have a lock-in effect, the same price increase would lead to a smaller decrease of l_j than of n_j. Some old consumers may not switch when switching costs are high even though they would switch if there were no switching costs (in which case they reaction would be the same as if they were new consumers). The fraction of consumers who are prevented from switching in this way may depend on the distribution of switching costs, the Guided buy this intuition, we aim to identify the lock-in effects of switching costs from the difference of the effects of the same price increase on these two probabilities.

2.2 Industry background

We have in mind an industry with J firms offering a contract for a good or a service lasting for period T with required payments p_{jt} in each period.[1] Consumers are heterogenous in their reservation prices and possibly in some taste parameters. The goods or services may be but need not be differentiated across firms. In each period t, some new consumers enter the market (drawn from the same population as consumers in $t-1$), and some old consumers leave the market because of expiring contracts. Consumers maximize their lifetime utility

[1] Technically, there is no problem to allow for the entry of new firms (so J should not be fixed) or the supply of more services by one firm, but then naturally one would need to have data on all these components.

and discount future per-period utilities (or costs) by some discount factor. Firms may also maximize their profit streams and discount future revenues and costs, but we do not model firm behavior in this analysis.

If an old consumer would like to leave firm j for another firm because of a better price offer, she faces switching costs that may vary across individuals. In what follows, we model switching costs as fixed, and we have search costs and one-time transaction costs in mind. Richer switching costs structures would not change the qualitative conclusions so they would yield similar reduced-form relationships between price changes and choice probabilities.

Our stylized competition framework is close to the spirit of the classical theoretical framework of Beggs and Klemperer (1992). They study the main trade-off firms are facing in a similar dynamic setting: whether to charge high prices to rip-off locked-up consumers or low prices to attract new ones. The main difference, however, is that Beggs and Klemperer (1992) assume that there is no switching (presumably because prohibitively large switching costs[2]), while we allow switching costs to take any value and thus prevent only a fraction of old consumers from switching.

A possible way to justify whether our competition framework in the presence of switching costs can be applied to the industry studied, one can check simple observable market facts against some main theoretical results of Beggs and Klemperer (1992): entry is still attractive despite the presence of switching costs, while demand growth and an increase of discount rate (fall in the interest rate) causes prices to fall.

2.3 A discrete choice justification

In this subsection, we show that the reduced form approach we take can be grounded in a more structural discrete choice framework with switching costs. As we shall see, the model is quite complex, and its full analysis is beyond the scope of our paper. Instead, we use the model for confirming the basic intuition behind our more reduced-form estimation strategy and getting additional insight to interpret the reduced-form coefficients.

In the model, individuals are denoted by i, firms by $j = 1$.[3] Buying a product entails entering into a contractual relationship. Time spent since buying the product (i.e. since in a contractual relationship but not necessarily in the same contract with the same firm) is denoted by $s = 1..S$. The individual who starts buying the product from a firm j may stay loyal to firm j or switch to another firm k at some time $s^* > 1$. If switched, we do not re-start s but assume instead that the individual starts buying the product from firm k starting with period $s^* + 1$, from which point she can stay loyal to k or switch again. The

[2] So the customers of firm j also enter in a long-term (although not explicitly contractual) relationship with the firm and paying p_{jt} in subsequent periods.

[3] The outside option may or may not be included among the firms; as we shall see, our empirical implementation handles outside options with the inclusion of period fixed effects.

problem ends at S, which may be finite or infinite. Calendar time is denoted by t; at any given time t, buyers of firm j's product may be at different contract periods s. Price of firm j at time t is denoted by p_{jt}, while y_{ijt} denotes the choice of individual i at calendar time t ($y_{ijt} = 1$ if the individual chooses form j and 0 otherwise).

The goal of this section is to show two things: First, without switching costs, the effect of a price increase $\Delta p_{jt} > 0$ on the choice probability $\Pr(y_{ijt} = 1)$ is the same for individual i if she is new ($s = 1$) and if she is old ($s > 1$). Second, switching costs make the effect smaller on average in absolute value if she is old. What we mean by "on average" will be clarified later.

The per period costs of the product are the sum of the observed price p_{js} and an unobserved component that is specific to the match of firm and individual u_{ij}. We assume that the unobserved component is time-invariant. Unobservables capture individual-specific monetary costs, brand preferences, and other components that are specific to the match of the firm and the individual and affect the utility of the consumer. While obviously an extreme assumption, time-invariance captures the idea that many of those match-specific utility components may be persistent (such as taste heterogeneity, regional differences in availability, brand loyalty, etc.).

In period $s = 1$ individual i is a new consumer, the costs of choosing firm j is the discounted present value of the expected per-period costs. Assume that individual i's best expectations for future p_{js} is its current value so that $E_1[p_{js}] = p_{j1}$.[4] Therefore, the new consumers' expected costs is

$$e_{ij1} = E_1\left[\sum_{s=1}^{S}\frac{p_{js}+u_{ij}}{(1+\rho)^s}\right] = \sum_{s=1}^{S}\frac{E_1[p_{js}]+u_{ij}}{(1+\rho)^s} = \sum_{s=1}^{S}\frac{p_{j1}+u_{ij}}{(1+\rho)^s} \tag{1}$$

The choice problem amounts to picking the firm that has the lowest costs: choose j if $e_{ij1} \leq e_{ik1}\ \forall k \neq j$. The condition simplifies to $u_{ij} - u_{ik} \leq p_{k1} - p_{j1}\ \ \forall k \neq j$, or in vector notation

$$\mathbf{u}_{ij} - \mathbf{u}_{ik} \leq \mathbf{p}_{kt} - \mathbf{p}_{jt} \text{ if } s = 1 \tag{2}$$

where the dimension of the vectors is $(J-1)\times 1$, $\mathbf{u}_{ij}$ is a vector with all elements u_{ij}, $\mathbf{p}_{j1}$ is a vector with all elements p_{j1}, while $\mathbf{u}_{ik}$ and $\mathbf{p}_{k1}$ are the $(J-1)\times 1$ vectors of the different u_{ik} and p_{k1} entries, respectively ($k \neq j$). Intuitively, the individual should choose firm j if the prices of all other firms exceed firm j's price to a degree that the difference is larger than firm j's subjective costs relative to all other firms' subjective costs.

Assume that the vector of unobservables is distributed i.i.d. across individuals. The

[4] We think this assumption is justified. Loan or loyalty contracts usually specificy the same per-period fixed fee, while the variable payments depending on consumption (like minutes called) can usually be best proxied by current consumption.

probability of new consumers choosing firm j at calendar time t is then

$$n_{jt} \equiv \Pr\left(y_{ijt} = 1 | s = 1\right) = \Pr\left(\mathbf{u}_{ij} - \mathbf{u}_{ik} \leq \mathbf{p}_{kt} - \mathbf{p}_{jt}\right) = F\left(\mathbf{p}_{kt} - \mathbf{p}_{jt}\right) \tag{3}$$

where F is the joint c.d.f. of the unobserved cost differentials $\mathbf{u}_{ij} - \mathbf{u}_{ik}$. We assume that F is twice continuously differentiable. This is a familiar discrete choice problem.

Intuitively, an increase in p_{jt} would make some new consumers change their mind and choose another firm instead: these are those for whom at least one element of the threshold (the left-hand side of 2) is high enough exceeding for the corresponding relative price. n_{jt}, the fraction of consumers buying from firm j, is decreased by the fraction of such consumers. The magnitude is determined by the fraction of such marginal individuals, which is determined by the shape of F at $\mathbf{p}_{kt} - \mathbf{p}_{jt}$.

The choice problem for periods $s > 1$ is perhaps easiest understood as an optimal stopping problem. As we shall see, a consumer that chooses firm j a $s = 1$ stays loyal to firm j as long as the relative prices of the other firms $(\mathbf{p}_{kt} - \mathbf{p}_{jt})$ remain above a certain threshold. If at some s^* the threshold is passed, the consumer switches to the cheaper firm, and the stopping problem restarts from $s^* + 1$. Since price differences are likely to be stationary, it is possible that the individual stays loyal to firm j for the entire duration of the contract (until period S), even if S is infinitely large.

Denote the probability of staying loyal to firm j by l_{jt} :

$$l_{jt} = \Pr\left(y_{ijt} = 1 | s > 1, y_{ijt-1} = 1\right) \tag{4}$$

Unlike the choice probability of new consumers (n_{jt}), the probability of staying loyal (l_{jt}) is a conditional probability: it is conditional on the individual's choice of firm in the previous period. But that choice itself was a loyalty decision, too, which was again conditional on the consumer's earlier choice, etc. Therefore, the loyalty probability is the result of a fairly complicated problem that is a function of all past prices not only current ones. Solving and analyzing the model in its full depth is beyond the scope of our paper. Instead, we focus on some intuitive implications from the loyalty conditions themselves.

The expected cost for consumer i if stays loyal to firm j in period $s > 1$ is

$$e_{ijs} = E_s\left[\sum_{\tau=s}^{S} \frac{p_{j\tau} + u_{ij}}{(1+\rho)^{\tau}}\right] = \sum_{\tau=s}^{S} \frac{p_{j\tau} + u_{ij}}{(1+\rho)^{\tau}}, \text{for } s > 1 \text{ if } y_{ij(s-1)} = 1 \tag{5}$$

Choosing another firm would entail a one-time additional switching cost $\mathrm{C}_{ij} \geq 0$. The individual cost of switching to firm k is

$$\begin{aligned} e_{iks} &= \sum_{\tau=s}^{S} E_s\left[\frac{p_{k\tau} + u_{ik\tau}}{(1+\rho)^{\tau}}\right] + C_{ij} = \\ &= \sum_{\tau=s}^{S} \frac{p_{k\tau} + u_{ik\tau}}{(1+\rho)^{\tau}} + C_{ij}, \text{ for } s > 1 \text{ if } y_{ij(s-1)} = 1 \end{aligned} \tag{6}$$

The individual would then stay loyal to firm j if and only if $u_{ijt} - u_{ikt} \leq p_{ks} - p_{js} + C_{ij}/\sum_{\tau=s}^{S} \frac{1}{(1+\rho)^{\tau}} \ \ \forall k \neq j$. Define $c_{ijs} = C_{ij}/\sum_{\tau=s}^{S} \frac{1}{(1+\rho)^{\tau}}$: it measures switching costs scaled by the remaining time. For a given S, c_{ijs} is negatively related to $S - s$ (and positively related to s): In a forward-looking decision, the longer the remaining time the smaller the role of one-time switching costs. Or, in other words, switching costs are expected to be more important the closer the end date S (the larger s).[5]

Using vector notation and rearranging the condition, it can be rewritten in the following way. An individual who in all previous periods 1 through $s-1$ was buying the product from form j should stay loyal to it in period s if

$$\mathbf{u}_{ij} - \mathbf{u}_{ik} - \mathbf{c}_{ijs} \leq \mathbf{p}_{kt} - \mathbf{p}_{jt} \tag{7}$$

where $\mathbf{c}_{ijs}$ is a vector with all elements c_{ijs} (therefore $\mathbf{c}_{ijs} \geq 0$) .

An immediate conclusion from (7) is that the condition of staying loyal to j is the same as the condition of choosing j in the first place (condition 2) if there are no switching costs, (i.e. if $C_{ij} = 0$, which would result in $c_{ijs} = 0$ for all s). This result confirms the intuition behind our reduced-form approach: without switching costs, the choice problem of staying loyal to a firm is identical to the choice problem of first purchases. As a result, the price responsiveness of new consumers can be a valid approximation of the price responsiveness of old consumers in the absence of switching costs.

Switching costs ($C_{ij} > 0$ and thus $c_{ijs} > 0$), on the other hand, decrease the threshold that other firms' relative prices have to exceed in order for consumer i to stay loyal to firm j. One consequence is that for given prices, the loyalty probability is greater than the choice probability of new consumers. The other consequence is that, starting from above the threshold for new consumers (left-hand side of 2), own prices have to increase more (other firms' prices have to decrease more) in order for passing the threshold for old consumers. In particular, for a given increase in own price ($\Delta p_{jt} > 0$), there are always consumers who would switch without switching costs but whose c_{ijs} is high enough to prevent switching. The fraction of such consumers depends on the c.d.f. of switching costs c_{ijs}.

If we look at all individuals who are consumers of firm j at a given calendar time t, the probability of a random individual to stay loyal to firm j is an average of the individual probabilities. The average is taken over c_{ijs}, which means that it takes into account both individual heterogeneity in C_{ij} and heterogeneity in remaining contract time $S-s$. Since c_{ijs} is increasing in s (decreasing in the remaining contract time $S-s$) we expect more people to switch in growing markets than in stationary markets ceteris paribus, because of the fraction of small-s consumers is larger in the first case.

[5] This property is qualitatively but is decreased in magnitude if we add preiod-specific switching costs or costs that are scaled by the remaining time directly (as in cases when original firms make switching consumers pay a sum related to remaining time).

The intuition behind the reduced-form measurement model is thus supported by a more structural discrete-choice framework. In the absence of switching costs, old consumers should react to price changes the same way as new consumers. The presence of switching costs not only makes switching less likely, but it also makes consumers less likely to react to price changes. As a result, the same price change would lead to a smaller decrease in the loyalty probability of old consumers than the choice probability of new consumers. The difference between the two effects is therefore an intuitive measure of the lock-in effects of switching costs. The level of switching costs themselves are, on the other hand, impossible to identify without further assumptions.

3 Empirical strategy

3.1 Parameters of interest

Our first way to measure switching costs is the difference between the effect of an increase in the price charged by firm j on the choice probability of new consumers (n_{jt}; defined in (3)) and the loyalty probability of old consumers (l_{jt}; defined in (4)). This difference we denote it δ:

$$\delta = \left|\frac{\Delta n_{jt}}{\Delta p_{jt}}\right| - \left|\frac{\Delta l_{jt}}{\Delta p_{jt}}\right| \tag{8}$$

Indicator δ is relatively straightforward to interpret: it shows how much more likely it is for a consumer to switch away from firm j in response to a small price increase if she is new to the market than if she is already a consumer of firm j. In a frequentist interpretation, this difference shows the fraction of consumers who are prevented from switching but who would have switched in the absence of switching costs. One can also think of δ as a difference-in-differences estimator.

If no old consumer is constrained by switching costs, the two probabilities respond in the same way so that $\delta = 0$. If switching costs prevent some old consumers from responding to the price change then $|\Delta n_{jt}/\Delta p_{jt}| > |\Delta l_{jt}/\Delta p_{jt}|$, which leads to $\delta > 0$. If switching costs are prohibitive for everyone then no old consumer changes her purchasing behavior, so $\Delta l_{jt}/\Delta p_{jt} = 0$ and thus $\delta = |\Delta n_{jt}/\Delta p_{jt}|$.

The same difference can be measured in a relative way as well. The magnitude of δ depends on the magnitude of the own-price elasticity of the product among new consumers. Different markets may be characterized by different demand elasticities, and therefore the probability responses of new consumers may be different as well. As a result, δ is not directly comparable across markets, but a relative difference is. Let θ denote the normalized version of δ:

$$\theta = \frac{|\Delta n_{jt}/\Delta p_{jt}| - |\Delta l_{jt}/\Delta p_{jt}|}{|\Delta n_{jt}/\Delta p_{jt}|}. \tag{9}$$

θ shows the fraction of consumers who are prevented from switching from among those who would have switched in the absence of switching costs. By this definition, θ might take values between 0 (nobody is constrained by switching costs) and 1 (everybody is constrained by switching costs).

3.2 Measurement of the variables

The goal of our empirical analysis is estimating δ and θ. We argue that estimating them is feasible using panel data on all firms and with information on prices and two quantities: the quantity bought by new consumers and the quantity bought by old consumers. Recall that we index firms by index j and calendar time by t. The first step is estimating the probability responses $\Delta n_{jt}/\Delta p_{jt}$ and $\Delta l_{jt}/\Delta p_{jt}$.

Note that in principle, we should keep the prices of other firms ($\mathbf{p}_{kt}$) constant . In our application, we shall make the simplification of including the price of firm j relative to other prices in a single variable, instead of entering all other prices. Depending on the specific market context, relative prices might be defined as differences, ratios or log differences, with at least two possible benchmark prices: the best possible offer (smallest price) or the market average. Comparing to the smallest price is consistent with perfectly informed and rational consumers. This might be better for markets where prices are relatively easy to acquire and compare, like internet subscriptions. Comparing to the average price is consistent with consumers who cannot collect and process all price information available, so they compare the price of firm j to only a few competitors.[6] This latter might be better for markets where search costs are likely to be significant, like banking or telephone services.

Let S_{jt} denote the stock of all consumers who buy from firm j in period t. We denote the number of incoming consumers to firm j by IN_{jt} and the number of outgoing consumers from firm j by OUT_{jt} . If we can separate the number of consumers whose contract is expiring with firm j (that is they do not face explicit exit costs) from the outgoing consumers, we denote this number by X_{jt} - in this case, OUT_{jt} measures consumers who deliberately terminated their ongoing purchasing relationship with firm j. The evaluation of firm j's stock is therefore:

$$S_{jt} = S_{jt-1} + IN_{jt} - OUT_{jt} - X_{jt}. \tag{10}$$

Incoming consumers can be further separated in two categories: completely new consumers N_{jt} and switchers from other firms F_{jt}. Outgoing consumers also belong to two potential groups: Q_{jt} quit the market for good (likely because of a change in an individual factor like income) and T_{jt} switch to other firms (likely because of a price change). Therefore,

[6] The average may be weighted by previous market shares, but in a regression of market shares on prices such weighting may lead to endogeneity.

we have

$$S_{jt} = S_{jt-1} + (N_{jt} + F_{jt}) - (Q_{jt} + T_{jt}) - X_{jt}. \tag{11}$$

To illustrate these decompositions, let us take an example from the market of banking loans, to which we shall return in our application. The stock S_{jt-1} is the number of consumers having a loan contract with bank j in the beginning of period t. The stock may change in three ways: by IN_{jt} new loans signed, OUT_{jt} loans repaid earlier, and X_{jt} loans expiring in the respective period. Consumers Q_{jt} repay their loan because they wanted to clear their debt regardless of any possible price change, while T_{jt} consumers refinance their loan with another bank. Finally, the bank's incoming consumers consist of consumers who are new to the market (N_{jt}) and refinancing consumers (F_{jt}) arriving from other banks.

Note that in certain market contexts, we might measure all variables in the value of contracts or revenues instead of the number of consumers. If you have both variables, like in our banking example, you might want to work with both in order to check the robustness of your results.

The important measurement problem to deal with is that although we would need decomposition (11) in order to ideally implement our thought experiment, firm-level aggregates usually allow us to track back decomposition (10). We shall address this problem later when we discuss feasible estimation.

A new consumer's realized probability of joining firm j is

$$n_{jt} = \frac{N_{jt}}{\Sigma_j N_{jt}},$$

which is simply firm j's market share from new consumers in period t. At the beginning of period t, firm j has S_{jt-1} old consumers. From among them, $X_{jt} + Q_{jt}$ leave the firm without switching, and an additional T_{jt} leave due to switching. The pool of potential switchers is therefore $S_{jt-1} - X_{jt} - Q_{jt}$, and form them T_{jt} choose to switch. The realized probability of staying loyal to firm j is therefore

$$l_{jt} = 1 - \frac{T_{jt}}{S_{jt-1} - X_{jt} - Q_{jt}}.$$

This probability equals the fraction of consumers loyal to firm j from among all consumers who could have been loyal to it.

3.3 Ideal estimation

Suppose for the moment that we can observe all terms in equation (11) and therefore compute n_{jt} and l_{jt}.We are interested in the changes of these probabilities in reaction to price changes, which we estimate from the following basic system of two equations:

$$\Delta n_{jt} = \alpha_n + \beta^* \Delta p_{jt-1} + u_{njt}, \text{ and} \tag{12}$$
$$\Delta l_{jt} = \alpha_l + \gamma^* \Delta p_{jt-1} + u_{ljt}. \tag{13}$$

where the star superscript denote estimation in an ideal situation in which the n and l variables are observed. Recall that we measure prices p relative to the market average. A more sophisticated way of keeping other prices constant would be to control for each of the prices. The measurement model is easily generalizable to that more sophisticated case. We stay within the simpler framework both because of notational simplicity and because the more sophisticated approach would require long time series, which are not always available.

We argue that in most applications it makes sense to relate changes in consumer decision to lagged price changes. Search for best prices takes price, and in many applications, transactions follow consumer decisions with considerable lag. In such cases, unless frequency of observations is low (i.e. time periods are wide), we can expect price changes in one period to affect measured transactions in the next period. Entering price changes with a lag also helps with endogeneity of price changes. An important source of endogeneity is the reaction of firms to changes in new demand or the stock of their consumers. Lagged prices are clean from this endogeneity because firms cannot change their price retroactively.[7]

Now suppose that the following two conditions hold (besides the maintained assumption of homogenous goods:

Condition 1 *New and old consumers are similar on average in terms of characteristics that matter for demand.*

This first condition is necessary for new consumers to serve as valid counterfactuals for old consumers, that is to adequately describe what the behavior of old consumers were without switching costs. This property is more likely to be satisfied on a stable market.

Condition 2 *Price changes are exogenous to demand.*

The second condition is needed to identify changes in demand.

Under these two conditions, OLS regressions of (12) and (13) consistently estimate the

[7] Of course changes in demand are likely to be serially correlated, which induces endogeneity in lagged prices as well. But such endogeneity can be handled by entering contemporenous price changes as additional controls. Our preferred specification in the application will include such controls.

theoretical β and γ coefficients.[8] As a result,

$$\hat{\delta} = \hat{\beta}^* - \hat{\gamma}^*, \text{ and} \tag{14}$$

$$\hat{\theta} = \frac{\hat{\beta}^* - \hat{\gamma}^*}{\hat{\beta}^*} \tag{15}$$

are consistent estimators of δ and θ as defined in (8) and (9) because of their continuity in the consistent $\hat{\beta}^*$ and $\hat{\gamma}^*$ estimators. Their sampling distribution is, complicated as it involves the covariance of $\hat{\beta}^*$ and $\hat{\gamma}^*$. Moreover, $\hat{\theta}$ is nonlinear in the regression estimators, and its distribution is non-standard. Estimating confidence intervals requires bootstrap or other simulation-based methods.

Firm-specific time-invariant heterogeneity in market share in new contracts (n_{jt}) and loyalty (l_{jt}) are filtered out in the regressions because they are specified in first differences. Similarly, as we estimate the evolutions of shares, the specifications take care of the shocks affecting all firms in the same way (although it is strictly true for n_{jt} only). For this latter reason it may be advisable to include time fixed effects to the regressions, and additional cross-section fixed-effects may be also included in order to control for firm-specific trends.

Note that time fixed-effects control for everything that is common for all banks in a given time period. These controlled factors include the potential benchmark price whether they are the average or the minimum. As a result, the theoretically important distinction of using absolute versus relative prices becomes empirically irrelevant if time fixed-effects are included.[9] Time fixed effects can also control, to some degree, for changes in market structure or the outside option.

3.4 Feasible estimation

In a typical application, the ideal left-hand side variables in (12) and (13) are unobserved: aggregate data are seldom available by the status of the consumer in previous time periods. However, the number of incoming and outgoing consumers from decomposition (10) is usually available in some markets, and we argue that these can be used as proxies in our estimations.

[8] (12) and (13) define a seemingly unrelated regression (SUR) system. Since each equation includes the same right-hand side variables equation-by-equation OLS is identical to GLS and therefore there is no efficiency loss.

[9] Naturally, $\hat{\beta}$ and $\hat{\gamma}$ are estimated from responses to price changes that are observed in the data. Generalization to price changes that are outside the observed range may be problematic. If, for example, switching costs have a common lower bound across consumers and firms keep their price increase below that lower bound, no consumer would switch. As a result, we would estimate $\hat{\theta} = 1$ implying that switching costs are prohibitive for everyone. This is of course true for the observed price changes but would not be true for larger ones. Note that this problem is not unique to our method but *any* regression-based estimation of switching costs, including those using individual data.

We denote the proxy of n by m and the proxy of l by k, and define them as follows:

$$m_{jt} = \frac{IN_{jt}}{\Sigma_k IN_{kt}}, \tag{16}$$

$$k_{jt} = 1 - \frac{OUT_{jt}}{S_{jt-1} - X_{jt}}. \tag{17}$$

By using the proxies, our regressions to estimate are:

$$\Delta m_{jt} = \alpha_m + \beta \Delta p_{jt-1} + u_{mjt} \tag{18}$$

$$\Delta k_{jt} = \alpha_k + \gamma \Delta p_{jt-1} + u_{kjt}. \tag{19}$$

The principal question is how estimators $\hat{\beta}$ and $\hat{\gamma}$ are related to the ideal estimators $\hat{\beta}^*$ and $\hat{\gamma}^*$, respectively. This depends on whether the discrepancies between proxy and ideal variables are correlated to (lagged) price changes, the right-hand side variable of each regression. Formally, a sufficient condition for $\hat{\beta}$ and $\hat{\gamma}$ to be consistent for the same parameters as $\hat{\beta}^*$ and $\hat{\gamma}^*$ is the following.

Condition 3 *The discrepancy between proxy left-hand side variables and ideal ones is uncorrelated to price changes:*

$$\begin{aligned} Cov\left(\Delta d_{mjt}, \Delta p_{jt-1}\right) &= Cov\left(\Delta d_{kjt}, \Delta p_{jt-1}\right) = 0, \textit{ where} \\ \Delta d_{mjt} &\equiv \Delta m_{jt} - \Delta n_{jt}, \textit{ and} \\ \Delta d_{kjt} &\equiv \Delta k_{jt} - \Delta l_{jt}. \end{aligned}$$

Condition 3 adds to the two necessary conditions listed in the previous Subsection. Therefore, applications of the measurement model using proxy variables should address Condition 3 besides Conditions 1 and 2.

In the applied estimation model outlined in this section, the proxy of n_{jt} is m_{jt}, the market share in all new loans issued in period t as defined in (16). This proxy variable errs by potentially including switchers F_{jt} from other banks : $IN_{jt} = N_{jt} + F_{jt}$. Therefore the discrepancy between the ideal variable and the measured one, $d_{mjt} = m_{jt} - n_{jt}$, may include switchers. If price changes induce any switching, an increase in firm j's price may discourage switchers as well as new consumers. As a result, the estimated reaction of new consumers is biased downwards (looks stronger than it is).

Formally, we have that

$$\begin{aligned} \text{p}\lim \hat{\beta} &= \text{p}\lim \hat{\beta}^* + \frac{Cov\left(\Delta d_{mjt}, \Delta p_{jt-1}\right)}{V\left(\Delta p_{jt-1}\right)} = \text{p}\lim \hat{\beta}^* + bias \\ bias &= \frac{Cov\left(\Delta d_{mjt}, \Delta p_{jt-1}\right)}{V\left(\Delta p_{jt-1}\right)} < 0 \end{aligned}$$

The bias is due to changes in switching consumers as a response to price changes, and is therefore related to γ^*. If switching costs prevent everybody to change firms, there is no bias in $\hat{\beta}$. The bias is likely to be larger the stronger the switching response, and therefore the larger γ^* is. In the Appendix, we show that an upper bound to the bias can be approximated as proportional to γ^*, where the proportionality factor is the average of the ration of firm-level stocks (minus exiting consumers) to the sum of all incoming consumers:

$$\begin{aligned} bias &\leq a\gamma^* \\ a &\approx E_{j,t}\left[\frac{S_{jt-1} - X_{jt}}{\Sigma_k IN_{kt}}\right] \end{aligned}$$

The proxy of l_{jt} is k_{jt} as defined in (17), based on contract terminations (loan repayments in our example) before due date. Recall that this variable is meant to proxy the fraction of consumers who did not switch after the price change. It errs on two counts. First, the numerator is $OUT_{jt} = T_{jt} + Q_{jt}$ instead of T_{jt}. It therefore includes consumers Q_{jt} who repay their loans before due date but do not refinance it by other banks. Second, the denominator is $S_{jt-1} - X_{jt}$ instead of $S_{jt-1} - X_{jt} - Q_{jt}$, which again includes Q_{jt}. The discrepancy $d_{kjt} = k_{jt} - m_{jt}$ is due to these two facts: the numerator and the denominator of l are both increased by the same Q_{jt}. The discrepancy is nonzero though because the numerator of l is smaller.

Contrary to the discrepancy for new consumers, this one will not lead to an estimation bias. The question is whether (normalized) non-refinancing terminations Q_{jt} are correlated to price changes in the previous period. We have no reasons to think that they are, because these terminations are typically due to positive income shocks, which are typically unrelated to price movements. As a reason we assume that Condition 3 holds: $Cov\left(\Delta d_{kjt}, \Delta p_{jt-1}\right) = 0$. Therefore, estimates of β_k are consistent for the same parameter as estimates of γ^* would be under ideal circumstances: $p\lim \hat{\beta}_k = p\lim \hat{\beta}_l$.

As a result, if Conditions 1 and 2 are satisfied,

$$\begin{aligned} \beta^* &\leq \operatorname{p}\lim \hat{\beta} \leq \beta^* + a\gamma^* \\ \operatorname{p}\lim \hat{\gamma} &= \gamma^* \end{aligned}$$

Consistency of $\hat{\gamma}$ allows us to get a simple bias-corrected version of $\hat{\beta}$ and thus the switching cost estimators:

$$\hat{\delta}_{corrected} = \left(\hat{\beta} - a\hat{\beta}\right) - \hat{\gamma} \tag{20}$$

$$\hat{\theta}_{corrected} = \frac{\left(\hat{\beta} - a\hat{\beta}\right) - \hat{\gamma}}{\hat{\beta} - a\hat{\beta}} \tag{21}$$

In terms of the absolute value, the corrected estimators are the lower bounds of the true parameters δ and θ, respectively. The effect of the bias correction is larger the stronger the estimated switching response (the larger $\hat{\beta}$). But effect is different for δ and θ, and a smaller effect is expected in terms of the latter.

In order to meet Conditions 1 and 2, the regression models (18) and (19) may in general include other variables. As we noted already, firm and time fixed-effects may be a good idea to include. Note that the model is defined in first differences so firm-specific time-invariant factors in market share and loyalty are automatically controlled for. Including additional firm fixed-effects amount to controlling for firm-specific (possibly stochastic) trends.

Using lagged price changes on the right-hand side makes sure they do not reflect firms' responses to autonomous unobservable changes in demand, at least if unobservables are serially uncorrelated. Controlling for firm-specific trends can significantly decrease serial correlation but may not eliminate it, in which case the contemporary price change Δp_{jt} may be a good idea to include as a proxy for endogeneity.

Comparing switching costs in different regimes is straightforward by comparing $\hat{\delta}$ and $\hat{\theta}$ estimated from separate samples. Such estimation may be more efficient if carried out in a pooled sample with appropriate interactions with Δp_{t-1}. Similar interactions may be useful to assess the role of observable firm-specific switching cost components. By interacting their level with price changes in regressions (18) and (19), one can estimate switching costs $\hat{\delta}$ and $\hat{\theta}$ at different levels of observed cost components.

Note however, that interactions with firms-specific cost components can be problematic as they are choice variables to firms. In our example of banking loans, loan termination fees are potentially observed firm-specific cost components. If banks see an exogenous increase in early repayment of loans (Q_{jt}), they may increase the termination fee in order cover possibly convex costs associated with such repayments. But that would create a correlation between the discrepancy Δd_{kjt} and observed termination fees, resulting in biased estimates. Moreover, termination fees may respond to switching itself (T_{jt}), leading to additional simultaneity bias. These problems need to be addressed in applications.

4 Empirical application

In this section we present an application in order to show how our measurement model can be put to work. The application was part of the banking sector inquiry of the Hungarian Competition Authority (GVH) in 2007.

We look at the market of personal loans, i.e. loans for undetermined use that can be either unsecured or secured with a mortgage. We consider the Hungarian market for such loans between 2002 and 2006, a market that has seen a significant growth during this period. Apart form being secured or unsecured, personal loans can be either in home currency (Hungarian forint, labeled as home currency) or foreign currencies, the latter mostly in Swiss Franc and Euro (combined and labeled as foreign currency). In the beginning of the period observed, the typical personal loan was in home currency, while from 2004 there has been a large switch towards secured loan, mostly in foreign currency, as they entered the market with lucrative interest rates for consumers. Table 1 below shows the shares for the stocks of these four loan groups, and Table 2 shows market segment growth rates.

Table 1. Market shares in percentages (value stocks over all consumers)

Loan type	2002	2003	2004	2005	2006
Home, unsecured	44	56	53	39	28
Foreign, unsecured	0	0	4	10	10
Home, secured	56	44	17	6	4
Foreign, secured	0	0	26	44	58
Percentage Total	100	100	100	100	100

Table 2. Yearly growth rates in percentages (contract values)

Loan type	2003	2004	2005	2006
Home, unsecured	88	173	29	8
Foreign, unsecured	-	-	323	54
Home, secured	19	8	-34	-14
Foreign, secured	17	164000	194	98
Market total	49	187	74	51

The two leading products are home currency unsecured and foreign currency secured loans, they consist at least 80 percent of total stocks for the last 3 years. Of the two of them, our analysis is concentrated to home currency unsecured loans. Although its share has been decreasing, it still has a significant presence on the market and it continues to show moderate growth in absolute terms. Therefore it might be seen as the most mature segment, where our analysis can be more fruitfully applied.

The overall dataset covers the nine largest banks in Hungary, identified as having at least one per cent market share on the personal loan market (they cover over 90 per cent of the market for all personal loans). We use quarterly data on prices and the number and contract value of new contracts and terminated contracts. Of the ten banks, seven provided adequate data for the number of consumers and six to the value of terminated contracts. Nonrespondents were from among the smaller players on the market.[10]

Prices p_{jt} are measured by the annual percentage rate (APR) of the bank's most popular (modal) product within the loan category. Given Hungarian financial regulation, APR includes all entry costs but not the termination costs. Lagging price changes by one period

[10]It is the *OUT* data that are missing for these banks. The two banks with missing both number and value data have a combined market share of 10 per cent, averaged over the sample period. The bank with missing value data has had a market share of an additional 10 per cent over the period.

makes perfect sense in our case. The data covers issued loans, and therefore consumer reaction can be measured by the time lag of loan administration. We use the difference of prices to (unweighted) market averages. Since our most preferred regressions include time fixed-effects, the use of this relative price measure yields the same estimates as raw prices would.

Our most preferred versions the regressions (18) and (19) include the contemporary price change Δp_{jt} as a proxy for endogenous price changes. The also include bank and time period fixed effects. Including time fixed effects ensures that all variables are measured as deviations from market-wide averages within t. In particular, a price change for bank j can to be interpreted as a price change relative either to average or minimum market price.

In the main text we present the OLS estimates of the coefficients β_m and β_k, and based on them, the estimate of δ and θ. Summary statistics and the complete set of parameter estimates and regression statistics are in the Appendix (tables A1 and A2).

We present boostrap confidence intervals on the $\left(5^{th}, 95^{th}\right)$ percentiles. These are estimated by block-bootstrap (re-samplig of complete firm histories as opposed to individual firm-year observations) in order to account for potential serial correlation. While the confidence intervals contain only 90 per cent of the sampling distribution, they are nevertheless rather conservative. This is can be seen by comparing the bootstrap standard errors of the $\hat{\beta}$ coefficients to their analytical standard errors shown in the Appendix: the bootstrap standard errors are significantly larger.

Table 3 shows the estimates of the regression parameters and the switching cost parameters assuming no bias in $\hat{\beta}_m$.

Table 3. Estimates of switching costs in the personal loan market in Hungary, 2002 to 2006. Assuming no bias in $\hat{\beta}_m$ the estimation of switching costs (δ and θ)

	# consumers	value
Response of new consumers (β_m)	-0.61	-0.74
(bootstrap SE)	(0.22)	(0.23)
(confidence interval)	$(-0.93, -0.14)$	$(-0.99, -0.22)$
Response of old consumers (β_k)	-0.13	-0.18
(bootstrap SE)	(0.06)	(0.07)
(confidence interval)	$(-0.18, -0.01)$	$(-0.24, -0.00)$
Switching costs: difference (δ)	0.48	0.56
(confidence interval)	$(0.13, 0.87)$	$(0.22, 0.81)$
Switching costs: normalized (θ)	0.79	0.76
(confidence interval)	$(0.66, 1.00)$	$(0.68, 1.00)$

Block-bootstrap confidence intervals (5^{th} percentile, 95^{th} percentile) based on 2000 runs.

As we have shown, however, $\hat{\beta}_m$ may be a biased estimator of β_n. We derived an upper bound to the bias, with a proportionality factor a, see (20) and (21). Table 4 shows the switching cost estimates allowing for this maximum bias.

Table 4: Corrected estimates of switching costs allowing for maximum bias.

	# consumers	value of contracts
Switching costs: difference (δ)	0.33	0.31
(confidence interval)	$(0.03, 0.80)$	$(0.10, 0.61)$
Switching costs: normalized (θ)	0.70	0.63
(confidence interval)	$(0.35, 1.00)$	$(0.41, 1.00)$

Note: The estimated value of the proportionality factor $a = 1.4$
Block-bootstrap confidence intervals (5^{th} percentile, 95^{th} percentile) based on 2000 runs.

Given the sample size, the estimates are reasonably precise. The confidence interval around parameter of major interest, θ, is especially tight. When maximum bias is allowed for, the estimates of δ become significantly smaller. At the same time, however, the bias-corrected estimates of θ are very close to the non-corrected estimates.

According to the point estimates, one percentage point increase in bank j's APR charges leads to an overall 0.61 percentage point decrease in the probability of new consumers choosing bank j (a bias-corrected point estimate would be $0.43 = 0.61 - 1.4 \times 0.13$). Note that the corresponding measure of demand elasticity is strong: one per cent increase in loan service (i.e. one percentage points increase in APR) leads to 4.7 per cent decrease in market shares evaluated at average market shares (the bias-corrected .estimate would be -3.3).

The same price increase is estimated to induce a mere 0.13 percentage points decrease in the probability of bank j's old consumers to stay loyal. This implies a demand elasticity of -0.13 (evaluated at the average loyalty probability of 0.98). Clearly, these estimates imply that firms have significant market power with respect to they old consumers. The corresponding estimates for the value of contracts are somewhat stringer, in line with the presumption that consumers with larger contracts are more price sensitive.

These estimates imply that the lock-in effects switching costs are substantial, whether measured in terms of choice probabilities or contract values. Estimates of θ indicate that due to switching costs, consumers are 79 per cent more likely to stay in their existing personal loan contract than if they were about to choose a new contract. The corresponding bias-corrected estimates are also high, at 70 per cent. Estimated switching costs are somewhat smaller in terms of contract value (they prevent 76 or 63 per cent of contract value to switch). The difference is small but it may indicate that consumers with larger contracts are somewhat less constrained by switching costs, a result consistent with part of switching costs being fixed.

5 Conclusions

Based on a simple thought experiment, we proposed a measurement model for estimating the lock-in effects of switching costs. The basic idea is to compare demand responses to price changes for consumers who are new to the market and for consumers who already are buyers of some firm, and the difference should attributable to the presence of switching costs. Implementation of the method requires proxies for these two quantities in each period: new transactions on the market and transactions (contracts) terminated by consumers.

We applied our method to the Hungarian personal loan market (we look at the modal and most mature product, unsecured personal loans denominated in home currency). According to our estimates, one percentage point increase in prices leads to a 0.61 percentage point decrease in demand among new consumers but 0.13 among the bank's old consumers. Old consumers' responsiveness is therefore 0.48 percentage points, or 79 per cent lower than new consumers' responsiveness. Even allowing for maximum bias to our estimates due to measurement problems, the price responsiveness of old consumers is 70 per cent lower. If demand responses are estimated in terms of contract value as opposed to number of consumers, old consumers' responsiveness is slightly less depressed implying slightly smaller lock-in effects. The results imply strong lock-in effects of old consumers and thus substantial market power for banks, especially in terms of ex-post fee increases.

Appendices

A Summary statistics and complete results

Table A1. Summary statistics (personal loans denominated in home currency, unsecured)

	# consumers			contract value		
	mean	*std.dev.*	*obs.*	*mean*	*std.dev.*	*obs.*
m_{jt}	0.13	0.18	105	0.14	0.18	87
k_{jt}	0.98	0.02	105	0.97	0.02	87
p_{jt}	0.06	0.03	105	0.05	0.03	87
Δm_{jt}	-0.007	0.049	105	-0.008	0.051	87
Δk_{jt}	-0.001	0.010	105	-0.001	0.012	87
Δp_{jt-1}	-0.001	0.001	105	-0.001	0.020	87

Table A2. Complete regression estimates (personal loans denominated in home currency, unsecured)

	# consumers		contract value	
	m	k	m	k
Δp_{jt-1}	-0.61	-0.13	-0.74	-0.18
SE	$(0.14)^{**}$	$(0.05)^{*}$	$(0.19)^{**}$	$(0.06)^{*}$
Δp_{jt}	-0.62	0.11	-0.83	0.16
SE	$(0.13)^{**}$	$(0.04)^{*}$	$(0.13)^{**}$	$(0.05)^{*}$
Firm FE	yes	yes	yes	yes
Period FE	yes	yes	yes	yes
R^2	0.49	0.43	0.55	0.44
# firms	7	7	6	6
Observations	105	105	87	87

Analytical standard error estimates (clustered at firm level) in parentheses.
R^2 include the explanatory power of fixed-effects.
** significant at 1%, * significant at 5%

B Deriving the bias to $\hat{\beta}_m$

Our goal is to approximate the bias to $\hat{\beta}_m$, which is the following:

$$bias = \frac{Cov\left(\Delta d_{mjt}, \Delta p_{jt-1}\right)}{V\left(\Delta p_{jt-1}\right)}$$

We start from the definition of the discrepancy d_m :

$$\begin{aligned}
n_{jt} &= \frac{N_{jt}}{\Sigma_k N_{kt}} \text{ and} \\
m_{jt} &= \frac{IN_{jt}}{\Sigma_k IN_{kt}} = \frac{N_{jt}+F_{jt}}{\Sigma_k N_{kt}+\Sigma_k F_{kt}}, \text{ so} \\
d_{mjt} &= m_{jt}-n_{jt} = \frac{N_{jt}+F_{jt}}{\Sigma_k N_{kt}+\Sigma_k F_{kt}} - \frac{N_{jt}}{\Sigma_k N_{kt}} \leq \frac{F_{jt}}{\Sigma_k IN_{kt}}.
\end{aligned}$$

The last result follows from the fact that $\frac{N_{jt}+F_{jt}}{\Sigma_k N_{kt}+\Sigma_k F_{kt}} - \frac{N_{jt}}{\Sigma_k N_{kt}} \leq \frac{N_{jt}+F_{jt}}{\Sigma_k N_{kt}+\Sigma_k F_{kt}} - \frac{N_{jt}}{\Sigma_k N_{kt}+\Sigma_k F_{kt}}$, because $\frac{N_{jt}}{\Sigma_k N_{kt}} \geq \frac{N_{jt}}{\Sigma_k N_{kt}+\Sigma_k F_{kt}}$.

Assume that the market is stationary in the sense that the number of new consumers is the same in each time period. Therefore, $\sum_j N_{jt} = \sum_j N_{jt-1}$, and so we have that

$$\Delta d_{mjt} = (m_{jt}-n_{jt}) - (m_{jt-1}-n_{jt-1}) \leq \frac{\Delta F_{jt}}{\sum_k IN_{kt}}.$$

The switching response to a price increase is captured by β_l. Next, we expand the definition of β_l in order to connect it to the measure of switchers in the discrepancy term.

$$\begin{aligned}
\beta_l &= \frac{Cov\left(\Delta l_{jt}, \Delta p_{jt-1}\right)}{V\left(\Delta p_{jt-1}\right)} \text{ and} \\
l_{jt} &= 1 - \frac{T_{jt}}{S_{jt-1} - X_{jt} - Q_{jt}}, \text{ so} \\
\Delta l_{jt} &= 1 - \frac{T_{jt}}{S_{jt-1} - X_{jt} - Q_{jt}} - \left(1 - \frac{T_{jt-1}}{S_{jt-2} - X_{jt-1} - Q_{jt-1}}\right) \\
&= \frac{T_{jt-1}}{S_{jt-2} - X_{jt-1} - Q_{jt-1}} - \frac{T_{jt}}{S_{jt-1} - X_{jt} - Q_{jt}} \\
&\approx \frac{T_{jt-1} - T_{jt}}{S_{jt-1} - X_{jt} - Q_{jt}} = -\frac{\Delta T_{jt}}{S_{jt-1} - X_{jt} - Q_{jt}}.
\end{aligned}$$

If firms are symmetric and consumers homogenous, the change in switching from bank j and to bank j are equal in absolute value (and they always have opposite sign):

$$\Delta T_{jt} \approx -\Delta F_{jt}.$$

As a result,

$$\begin{aligned}
\Delta l_{jt} &\approx -\frac{\Delta T_{jt}}{S_{jt-1} - X_{jt} - Q_{jt}} \approx \frac{\Delta F_{jt}}{S_{jt-1} - X_{jt} - Q_{jt}} \\
&= \frac{\Delta F_{jt}}{\sum_k IN_{kt}} \frac{\sum_k IN_{kt}}{S_{jt-1} - X_{jt} - Q_{jt}}.
\end{aligned}$$

This leads to a bound to the bias for each firm j in each time period t the following way:

$$\begin{aligned}
Cov\left(\Delta l_{jt}, \Delta p_{jt-1}\right) &\approx Cov\left(-\frac{\Delta T_{jt}}{S_{jt-1} - X_{jt} - Q_{jt}}, \Delta p_{jt-1}\right) \approx Cov\left(\frac{\Delta F_{jt}}{S_{jt-1} - X_{jt} - Q_{jt}}, \Delta p_{jt-1}\right) \\
&= \frac{\sum_k IN_{kt}}{S_{jt-1} - X_{jt} - Q_{jt}} Cov\left(\frac{\Delta F_{jt}}{\sum_k IN_{kt}}, \Delta p_{jt-1}\right), \\
Cov\left(\Delta d_{mjt}, \Delta p_{jt-1}\right) &\leq Cov\left(\frac{\Delta F_{jt}}{\sum_k IN_{kt}}, \Delta p_{jt-1}\right) \\
&\approx \frac{S_{jt-1} - X_{jt} - Q_{jt}}{\sum_k IN_{kt}} Cov\left(\Delta l_{jt}, \Delta p_{jt-1}\right) \leq a_{jt} Cov\left(\Delta l_{jt}, \Delta p_{jt-1}\right), \text{ where} \\
a_{jt} &= \frac{S_{jt-1} - X_{jt}}{\sum_k IN_{kt}}
\end{aligned}$$

In the last inequality we replaced $\frac{S_{jt-1} - X_{jt} - Q_{jt}}{\Sigma_k IN_{kt}}$ by $a_{jt} = \frac{S_{jt-1} - X_{jt}}{\Sigma_k IN_{kt}}$ because the latter is estimable, while the former is not.

Based on these results, we can approximate the upper bound to the bias in a panel of firms by the average of the jt bias bounding terms:

$$\begin{aligned} bias &= \frac{Cov\left(\Delta d_{mjt}, \Delta p_{jt-1}\right)}{V\left(\Delta p_{jt-1}\right)} \leq aCov\left(\Delta l_{jt}, \Delta p_{jt-1}\right), \text{ where} \\ a &= E_{j,t}\left[\frac{S_{jt-1} - X_{jt}}{\sum_k IN_{kt}}\right]. \end{aligned}$$

References

[1] BEGGS, ALAN W AND KLEMPERER, PAUL (1992). "Multi-period Competition with Switching Costs," *Econometrica*, 60(3), 651-66.

[2] CHEN, P. S. AND L. M. HITT (2002), "Measuring Switching Costs and the Determinants of Customer Retention in Internet-Enabled Businesses: A Study of the Online Brokerage Industry." *Information Systems Research.* 13(3), 255-274.

[3] FARRELL, J. AND P. KLEMPERER (2007), "Coordination and Lock-In: Competition with Switching Costs and Network Effects." *Handbook of Industrial Organization 3*, Elsevier.

[4] KIM, M., D. KLIGER, AND B. VALE (2003), "Estimating switching costs: the case of banking." *Journal of Financial Intermediation* 12, 25-56.

[5] NATIONAL ECONOMIC RESEARCH ASSOCIATES (2003), "Switching Costs". *Office of Fair Trading Economic Discussion Paper* 5.

[6] SHY, O. (2002), "A quick-and-easy method for estimating switching costs." *International Journal of Industrial Organization* 20, 71-87.