DEPARTMENT OF ECONOMICS
WORKING PAPERS

economics.ceu.hu

Estimating consumer lock-in effects from firm-level data

by

Gábor Kézdi[1] and Gergely Csorba[2]

2012/17

[1] Department of Economics, Central European University, IEHAS, kezdig@ceu.hu
[2] Institute of Economics, IEHAS

Abstract

This paper proposes a practical method for estimating consumer lock-in effects from firm-level data. The method compares the behavior of already contracted consumers to the behavior of new consumers, the latter serving as a counterfactual to the former. In panel regressions on firms' incoming and quitting consumers, we look at the differential response to price changes and identify the lock-in effect from the difference between the two. We discuss the potential econometric issues and measurement problems and offer solutions to them. We illustrate our method by analyzing the market for personal loans in Hungary and find strong lock-in effects.

Keywords: lock-in, switching costs, demand analysis, difference-in-differences, personal loans

JEL codes: C33, D12, L13

Acknowledgements

The empirical methodology presented in this paper was developed as part of the 2007-2009 retail banking sector inquiry of the Hungarian Competition Authority (GVH), in which the first author worked as an expert and the second author worked as Chief Economist. The views expressed in this paper are not purported to represent those of the GVH.

We thank Dávid Farkas for excellent research assistance, and Jacques Crémer, Timothy Hannan, Gábor Koltay, Surd Kováts, Aliz McLean, Balázs Muraközy, Konrad Stahl, Ádám Szentpétery, Chris Wilson and seminar participants at CEU, GVH, IE-HAS Budapest, Mannheim and Toulouse for their valuable comments on previous versions of the paper circulated under the title "Estimating the Lock-in Effects of Switching Costs from Firm-Level Data". All remaining errors are ours.

1 Introduction

Consumer lock-in is a critical factor in evaluating firms' market power in a dynamic framework. If consumers' ability to change their service provider is limited, due to the presence of switching costs for instance, then it decreases the elasticity of demand, which can lead to price increases in later periods. Consumer lock-in may also contribute to raising barriers to expansion for competitors, enabling incumbent firms to conserve their strong market position. These theories of harm imply that lock-in is an important concern for competition authorities and sectoral regulators alike,[1] and regulatory policies were specifically designed to increase consumer mobility in most network industries.[2]

This paper proposes a practical and intuitive approach for estimating lock-in effects in a direct way, which has two practical advantages. First, it stays within a demand analysis framework and thus there is no need to make strong assumptions on market structure. Second, it requires only firm-level data, which are less expensive and easier to collect than consumer-level data, and most regulatory bodies have the legislative power to acquire them.

We develop an empirical method that compares the reactions of new consumers to the reactions of old consumers with respect to price changes.[3] The difference between these two reactions will give an indication on firms' market power over old consumers. The basis of our identification is the idea that the behavior of new consumers can describe the behavior of old consumers in the absence of lock-in. In other words, the group of new consumers can provide the counterfactual outcomes to the group of old consumers, in the spirit of the program evaluation literature (see, for example, Imbens and Wooldridge, 2009).[4]

[1] The theoretical literature is thoroughly reviewed by Farrell and Klemperer (2007).

[2] An important example is mobile telephone number portability. See Maicas et al. (2009), and the references therein, on estimating the effect of number portability on switching costs and consumer mobility.

[3] A new consumer is defined as someone making her first purchase decision on the market, while an old consumer is already a customer of a firm.

[4] A simple thought experiment can be given with two identical consumers *New* and *Old* who differ only in that *Old* has been the customer of firm j for some time. Suppose that at current prices *New* would also buy from firm j, but there is a change in the relative price of firm j that is large enough to make *New* choose another firm. If there were no lock-in effects, *Old* would react in the exact same way to this price change and would switch. However, if

When we have ideal data at hand, our empirical model consists of a system of two panel regressions estimated in first differences, which measures the effect of a change in the relative price of a given firm on the respective firm's "market share" between two groups of consumers. The dependent variable in the first equation is the market share of the respective firm in terms of consumers who are new to the market. The dependent variable in the second equation is the retention rate, i.e. the "market share" of the firm among its own customers from the previous time period. The lock-in effects are measured by comparing the effects of the same price change on these two market shares. The difference of the two responses is the fraction of old consumers who would have switched if they had been new consumers but were prevented from doing so because of lock-in. This identification strategy is very similar to a "difference-in-differences" approach, in which one compares the behavioral response of a group that may be locked-in to the behavioral response of another group that is not.

For the counterfactual approach to successfully identify the lock-in effects, the two groups of old and new consumers have to be similar in terms of their price elasticity and switching behavior. This assumption does allow for differences among consumers entering the market in different time periods, as long as such differences are controlled for or are unrelated to the behavior we examine.[5] We believe that this condition is more likely to be satisfied in developed markets and/or markets with relatively homogenous goods, in which market entry may be due to exogenous shocks. Examples for such markets include loan contracts - which we analyze in our application -, consumer utilities and standard telecommunication services. Incidentally, these are also the typical industries where competition policy and regulation are more concerned with problems of market power.

With appropriate data, this counterfactual approach can be implemented in a relatively straightforward fashion by panel data methods with firm-specific and time fixed effects. Unfortunately, however, information on consumers who are new to the market and old consumers who switch to other firms is typically not present in firm-level data.[6] We therefore implement the proposed

switching costs are sufficiently large, *Old* might stay locked-in with firm j.

[5] For exampe, more informed or more patient consumers might enter the market in different time periods, but this heterogeneity does not necessarily mean that their switching reactions will be different depending on whether they are locked-in or not.

[6] While the development of information technologies now allows most firms to collect rich data on their customers,

method by using the number of consumers joining firms and the number of consumers leaving firms and construct proxy variables for changes in the fraction of new and old consumers. We address the potential biases due to the use of such proxy variables, and develop an easy-to-implement formula that corrects for the biases under reasonable assumptions. Data on prices and the number of consumers joining and leaving firms are usually available in markets with long-term contracts.

As an illustration, we apply the estimation method to the market of personal loans in Hungary. According to our estimates, a one percentage point increase in interest rates leads to a 0.61 percentage points decrease in demand among new consumers, compared to a 0.13 percentage points decrease among the banks' old consumers. This means that old consumers' responsiveness is four fifths lower than the responsiveness of new consumers. Our bias-corrected estimates indicate substantial lock-in as well. We can reject the hypothesis of perfect consumer mobility, while the hypothesis of complete consumer lock-in (as assumed in many theoretical models) cannot be rejected. Our method was applied in the banking sector inquiry of the Hungarian Competition Authority, and as a result, several regulatory recommendations were made to facilitate consumers' switching between firms.

2 Economic framework and parameters of interest

We do not explicitly model firms' behavior because our empirical model does not require an explicit structure, but we nevertheless outline the economic setup in the background of our measurement strategy. There are J firms offering a contract for a good (or service) lasting for T periods with required payments p_{jt} in each period. Each consumer demands at most one good, which might be homogenous or differentiated. In each period t, some new consumers, who are drawn from the same population as previous consumers, enter the market; and some old consumers leave the market because of expiring contracts. Consumers can be heterogenous in their reservation prices and in other taste parameters. We assume that both new and old customers of a given firm face the same

these databases are rarely merged at the industry level for a general analysis.

price p_{jt}.[7]

If in any period t a consumer buying from firm j in period $t-1$ (an "old" consumer of firm j) wants to leave firm j for firm k because of a better price offer, she faces some switching costs.[8] We allow switching costs to vary across individuals, so there is no loss of generality in assuming that switching costs are fixed (i.e. they do not depend on the value of the transaction).

If a consumer i is new to the market in period t, let n_{ijt} denote the probability that she buys the product from firm j under existing prices. Similarly, if consumer i is an old customer of firm j (that is she bought from firm j in period $t-1$), let r_{ijt} denote the probability that firm j retains this consumer (so she continues to buy from firm j). The realized choice probabilities will be the share of firm j from the total set of new consumers and from firm j's old consumers, respectively.

We are interested in how an increase in the price of firm j affects these choice probabilities. The responses to price changes are likely non-positive, so $\Delta n_{jt}/\Delta p_{jt} \leq 0$ and $\Delta r_{jt}/\Delta p_{jt} \leq 0$. Our most important assumption is that the effect of a price increase on the behavior of new and old customers would be the same in the absence of lock-in. If switching costs are high enough, they will increase the threshold value of a price increase that induces a reaction for some consumers. Such consumers would not switch if the actual price increase is lower than this threshold even though they would have chosen another firm in the absence of switching costs. These consumers are locked in because of switching costs. As a result, the average effect of the same price increase is likely to be smaller for old consumers than for new consumers, i.e. $|\Delta n_{jt}/\Delta p_{jt}| \geq |\Delta r_{jt}/\Delta p_{jt}|$. Ideally, these properties on the changes to be measured could be derived from a more structural framework of discrete choice with switching costs.[9]

Guided by the above properties, we aim to identify the lock-in effects from the difference of the effects of the same price increase on the choice probability of new consumers and the retention

[7] While this assumption is correct in the context of our empirical application, it is common industry practice to charge different prices to old and new consumers in some other markets. In such cases one should measure two prices or address the potential bias resulting from price differences if only one data point on price is available.

[8] See the next section for a more detailed discussion on what factors can be taken as switching costs.

[9] Our Online Appendix C specifies a simple theoretical choice model that delivers the intuitive results of our counterfactual approach.

probability:

$$\delta_{jt} = \left|\frac{\Delta n_{jt}}{\Delta p_{jt}}\right| - \left|\frac{\Delta r_{jt}}{\Delta p_{jt}}\right| = \frac{\Delta r_{jt}}{\Delta p_{jt}} - \frac{\Delta n_{jt}}{\Delta p_{jt}}. \tag{1}$$

Indicator δ_{jt} shows how much more likely it is that a consumer switches away from firm j in response to a small increase in p_{jt} if she is new to the market than if she is already a customer of firm j. In a frequentist interpretation, this difference shows the fraction of old consumers who are prevented from leaving firm j in period t but would have switched in the absence of lock-in. If no old consumer is constrained by lock-in then $\delta_{jt} = 0$, while if lock-in is complete for all customers of firm j in period t, then $\delta_{jt} = |\Delta n_{jt}/\Delta p_{jt}|$. Naturally, the value of δ_{jt} depends on the distribution of demand parameters and switching costs, as well as on the period-specific market position of firm j. In empirical applications, an average value of δ is likely to be the best absolute indicator of industry lock-in effect.

Since different markets can be characterized by different demand elasticities and market structures, δ is not necessarily comparable across markets. For comparisons, it is more convenient to use a normalized version of δ_{jt}:

$$\theta_{jt} = \frac{|\Delta n_{jt}/\Delta p_{jt}| - |\Delta r_{jt}/\Delta p_{jt}|}{|\Delta n_{jt}/\Delta p_{jt}|} = \frac{\Delta n_{jt}/\Delta p_{jt} - \Delta r_{jt}/\Delta p_{jt}}{\Delta n_{jt}/\Delta p_{jt}}. \tag{2}$$

This relative indicator of the lock-in effect shows the fraction of consumers prevented from switching from among the consumers who would have switched in the absence of any lock-in. By this definition, θ_{jt} might takes on values between 0 (nobody is constrained by lock-in) and 1 (all of those who would have switched without switching costs are constrained by lock-in).

Using the estimated θ and its sampling distribution, it is possible to test the hypothesis of two polar cases. The hypothesis of $\theta = 0$ corresponds to perfect consumer mobility (no lock-in of any degree), while the hypothesis of $\theta = 1$ corresponds to complete lock-in.

3 Links between lock-in effects and switching costs

Lock-in and switching costs are very much interrelated, as discussed in length by Farrell and Klemperer (2007). The presence of significant switching costs will necessarily lead to some locked-in consumers, but lock-in might actually result from factors other than switching costs in the strict sense. It is therefore useful to clarify what we mean by switching costs and how this connects to our central definition of lock-in.

As Farrell and Klemperer (2007, p. 1972) define, "a product has classic switching costs if a buyer will purchase it repeatedly and will find it costly to switch from one seller to another". Switching costs can include all transaction costs related to the process of switching: exit costs from the previous seller, the costs of searching for new sellers and the costs of entry related to the new seller. Switching costs should not be restricted only to monetary costs: search costs are usually opportunity costs of time, while exit costs could entail psychological transaction costs or loss of utility when the consumer purchased complementary products from the specific firm or has a specific preference for variety.

The approach taken by our paper defines consumer lock-in as customer inertia, a tendency to remain in an ongoing relationship with a specific seller. This outcome can especially manifest as a reduced responsiveness to the given seller's (relative) price increases, but also as an increased likelihood to repurchase the original or complementary goods from the supplier. The state dependency of consumer choice is a crucial factor to consider, and we should be careful not to interpret all situations with small price responsiveness as lock-in.[10] For example, strong consumer preference for a given firm might also lead to lower own-price elasticities, but this kind of loyalty does not result from the previous relationship with the seller itself.[11] It is therefore crucial in our empirical applications that we can identify the old consumers (installed base) of each firm.

By using the broad definition of switching costs presented above, there are many cases where we can attribute lock-in to switching costs. For example, if a consumer settles for a specific service provider after some searching and then does not reevaluate the available options until she has bad experiences with that firm, then this mental lock-in can be seen as a result of psychological switching

[10] See Dubé et al (2010) for a detailed discussion of state dependence in consumer choices and the various factors that could explain it.

[11] It should be noted though that the preference towards a given firm might evolve because of the experience of using that specific firm's product, especially if the consumer was not fully informed before the transaction. This process, however, changes the reservation price (demand) of the given consumer for the product, not the costs of switching.

costs. However, if a firm targets its old customers with personalized marketing offers (information leaflets, promotions), these specialized actions of the firm (and not only the relationship itself) are the ones affecting the price responsiveness of consumers, so switching costs have a much smaller (maybe negligible) role to play.

In conclusion, the existence of switching costs is maybe the most significant, but not the only contributing factor to lock-in. It is hard to empirically assess the magnitude of switching costs' contribution, but this is not strictly necessary either as we focus on estimating the lock-in effects. Still, our empirical approach can benefit from the insights of the switching costs literature.

3.1 Connections to the existing literature

The Industrial Organization model closest to our setup is the theoretical framework of Beggs and Klemperer (1992). This classic paper analyzes dynamic competition in the presence of switching costs, and studies the main trade-off between charging high prices to rip off locked-in consumers versus low prices to attract new ones. A substantive difference to our setting is, however, that Beggs and Klemperer assume that switching costs are so large that they prohibit consumers from switching in equilibrium (so there is complete lock-in), while we allow switching costs to take on any value, so some fraction of old consumers may switch.[12] When solving for equilibrium, Beggs and Klemperer analyze affine strategies in which each firm's price is a linear function of its market share plus some firm-specific constants. Our basic equations are of a similar form, but we estimate the effects of price changes on changes in choice probabilities and corresponding market shares. There are also several theoretical predictions of the Beggs-Klemperer model that can be checked in empirical applications: entry should be attractive despite the presence of switching costs; growth in demand should cause prices to fall; and larger firms should initially set higher prices and therefore lose market shares.

From a practical point of view, three issues are of principal empirical importance considering lock-in: 1) the presence and magnitude of switching costs or other factors causing lock-in, 2) the

[12] In most markets characterized by switching costs some switching does occurs, although it may be of small magnitude. This is the case in our application as well, which we study later.

magnitude of consumer lock-in and 3) the resulting effects on prices. Most empirical papers have focused exclusively on switching costs as the underlying factor behind lock-in and analyzed the third question without the intermediate step,[13] and some reviewed in this section look at the first. However, empirical evidence on lock-in effects as such is scarce in Industrial Organization applications, because its identification and quantification present multiple methodological challenges.[14]

There are a few examples where switching costs were estimated from individual-level data: for example, in the online brokerage industry by Chen and Hitt (2002), for breakfast cereals by Shum (2004), for Internet portals by Goldfarb (2006), for mobile telephone subscriptions by Grzybowski (2008) and for television subscriptions by Shcherbakov (2009). These papers usually base their estimation approach on a discrete choice model, but they are estimating switching costs, and not the resulting lock-in effect.

We can also find examples where the state dependency of consumer choices was empirically examined, typically by studying individual-level data.[15].These papers, however, focus on separating the key factors behind state dependency (which is possible by observing a large sample of subsequent purchases), while our paper focuses on the simple testing of state-dependency based on aggregated data.

Some papers use firm-level data to estimate switching costs. These papers focus on the magnitude of switching costs compared to prices as opposed to the lock-in effects. We know of two structural studies, which derive a specific model of competition in the presence of switching costs and then estimate equilibrium conditions for prices or market shares. Shy (2002) builds a static model with given switching costs in which firms' prices are set such that nobody has any incentives to undercut their rivals. By construction, his model predicts no switching and stable market shares and it is used as a benchmark for identifying the existence of switching costs. As an illustration, Shy (2002) estimates that switching costs are 35-50% of average price on the Israeli cellular phone

[13] See Farrell and Klemperer (2007, Chapter 2.2) for a detailed overview.

[14] At the same time, lock-in effects have been analyzed in the labor literature. For example, Madrian (1994) tests for the "job-lock" effect of employer-provided health insurance plans by comparing the effect of medical expenditures on the job switching behavior of those with insurance plans to those without such plans.

[15] See Chintagunta (1998), Dubé et al (2010), Seetharaman (2004), Seetharaman et al. (1999).

market, and vary between 0 and 11% of the average balance on the Finnish bank deposit market. Kim et al. (2003) model consumers' transitions and banks' intertemporal decision-making in a dynamic framework and apply it to the Norwegian loan market: their estimated switching costs are 4% of the average loan's value. Both papers measure switching costs in terms of prices, but they do not provide direct estimates for the lock-in effect of switching costs. Because of the structural approach, they also need correctly specified models of competition, in contrast to our counterfactual approach.

The idea that a reduced-form model can capture how the presence of switching costs alters consumers' price responsiveness is of course not completely new. In a homogenous good industry, small cross-price elasticity estimates across firms may indicate large switching costs because price increases do not result in significant losses to competitors.[16] Our method requires more data (two measures of quantity as opposed to one), but it has the additional advantage of identifying the magnitude of the lock-in effects of switching costs, and is also applicable to differentiated goods industries. There are also a few empirical papers on switching costs that use proxies in their empirical implementations, but they use proxies directly for the unobservable switching costs,[17] while our paper uses variables to approximate the magnitude of terminating and switching consumers.

Schiraldi (2011) uses a counterfactual approach in order to estimate transaction costs (relative to prices) in the Italian car market and finds large variation in transaction costs. This paper is similar to our approach in that he compares the share of consumers holding a car to the share of consumers buying the same car in the same period, but it uses individual data to estimate a structural dynamic model of consumer demand.

4 Empirical strategy

4.1 Ideal estimation

We are interested in the effect of a change in the price of the product on two probabilities. The first is the probability of a new consumer choosing firm j, which we denote by n_{jt}. The second is

[16] This method is discussed for example by NERA (2003, Appendix B).

[17] Sharpe (1997) is a notable example, with an application to switching costs in retail banking deposit markets.

the probability of an old consumer of firm j to remaining with firm j, and we denote this by r_{jt}. The empirical counterparts of these probabilities are firm j's market shares among new consumers and its own old consumers, respectively. We estimate the effects of price changes on these two probabilities from the following two equations:

$$\Delta n_{jt} = \alpha_n + \beta^* \Delta p_{jt-1} + u_{njt}, \text{ and} \tag{3}$$

$$\Delta r_{jt} = \alpha_r + \gamma^* \Delta p_{jt-1} + u_{rjt}. \tag{4}$$

where star superscripts denote estimations in the ideal situation where the n and r variables are observed. Recall that we measure prices p relative to the market average. The regressions are easily generalizable to alternative measures of relative prices.

In most applications, it makes sense to relate changes in consumer decisions to lagged price changes. Searching for the best prices takes time, and in many applications, transactions follow consumer decisions with a considerable lag. In such cases, unless the frequency of observations is low (i.e. time periods are long), we can expect price changes in one period to affect measured transactions in the next period. Entering price changes with a lag also alleviates the problem of the endogeneity of price changes (see the next section for more details).

Now suppose that the following two conditions hold:

Condition 1 *Characteristics that matter for demand changes have a similar distribution among new and old consumers.*

This first condition is necessary for new consumers to serve as valid counterfactuals for old consumers, that is to adequately describe what the reactions of old consumers would be without switching costs. This property is more likely to be satisfied in a stable market with relatively homogenous goods. Note, however, that the fact that more (or less) informed, sophisticated or impatient consumers enter the market in earlier periods does not necessarily violate the condition as long as these different consumer cohorts' behavioral reactions to price changes are similar regarding their choice of firm.

The similarity of new and old consumers is required in terms of the price changes they face as well. This is obviously satisfied if firms cannot charge different prices to new and old consumers.

It may also be satisfied, however, if such price discrimination is feasible as long Δp is the same for new and old consumers. Examples for the latter include fixed discounts or free complementary items for new consumers if prices are entered in levels in the regressions, or proportional discounts if prices are entered in logarithmic form.

Condition 2 *Price changes are exogenous to changes in demand.*

The second condition is needed to identify changes in demand.

Under these two conditions, OLS regressions of (3) and (4) consistently estimate the theoretical β and γ coefficients.[18] As a result,

$$\hat{\delta} = \hat{\beta}^* - \hat{\gamma}^*, \text{ and} \tag{5}$$

$$\hat{\theta} = \frac{\hat{\beta}^* - \hat{\gamma}^*}{\hat{\beta}^*} \tag{6}$$

are consistent estimators of δ and θ as defined in (1) and (2) because they are continuous in the consistent $\hat{\beta}^*$ and $\hat{\gamma}^*$ estimators. Their sampling distribution involves the joint sampling distribution of $\hat{\beta}^*$ and $\hat{\gamma}^*$. $\hat{\theta}$ is also nonlinear in the regression estimators. Therefore, estimating confidence intervals is probably best done by bootstrapping or other simulation-based methods.

The regression coefficients β^* and γ^* do not vary across firms. As a result, the lock-in measures δ and θ do not vary across firms, either. This restriction is consistent with the assumption of homogenous lock-in effects, but that assumption is unlikely to be true. Firms may be very different in their costumer base, loyalty programs etc. In principle, our framework can be generalized to allow for firm-specific coefficients on consumer response $\left(\beta_j^* \text{ and } \gamma_j^*\right)$ and firm-specific measures of lock-in (δ_j and θ_j). However, because of sample size issues, applications using firm-level data are likely to produce very imprecise estimates of such firm-specific parameters. The restricted parameters are meaningful without making the homogeneity assumption, too: β^* and γ^* measure average responses across all firms, and thus δ and θ measure the average degree of lock-in.

Firm-specific time-invariant heterogeneity in market share in new contracts (n_{jt}) and the fraction of remaining consumers (r_{jt}) are filtered out in the regressions because they are specified in

[18] Equations (3) and (4) 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.

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 this is strictly true only for n_{jt}).

It is advisable to include time fixed effects and additional cross-section fixed effects in order to control for aggregate changes and firm-specific trends. Note that time fixed effects control for everything that is common to all firms in a given time period, including the potential benchmark price, whether it is the average or the minimum. As a result, the theoretically important distinction between using absolute versus relative prices becomes less relevant empirically if time fixed effects are included.[19] Time fixed effects can also control, to some degree, for changes in market structure or the outside option.

4.2 Addressing potential econometric problems

In order to meet Conditions 1 and 2, regression models (3) and (4) may in general include other variables. As we noted previously, it may be a good idea to include firm and time fixed effects. Note that the model is defined in first differences so firm-specific time-invariant factors are controlled for automatically. Additional firm fixed effects in the first-difference specification control for firm-specific (possibly stochastic) trends. Additional year fixed effects control for nonlinearities in aggregate trends.

Condition 2 requires exogenous variation in prices. Such exogeneity is best ensured by natural experiments or the use of valid instrumental variables. Note, however, that finding valid instruments is difficult in these applications, even more so than in demand analysis in general. It is standard in the Empirical Industrial Organization literature to use the competitors' characteristics as instruments. That is obviously ruled out here as the competitors' behavior is likely to affect switching

[19] Naturally, $\hat{\beta}$ and $\hat{\gamma}$ are identified 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, consumers do not switch to other firms below a lower bound of price changes and firms keep their price increases below that lower bound, no consumer would switch. As a result, we would estimate $\hat{\theta} = 1$, implying that all consumers are locked in. This conclusion is true within observed price changes but would not be true for larger ones. Note that this problem is not unique to our method but applies to *any* regression-based estimation of lock-in, including those using individual data.

(and thus Δr) directly. Another set of usual variables are "cost shifters." Since our application looks for variation in prices within the same market, cost shifters are likely to be extremely weak instruments because they are likely to affect competitors in similar ways. In fact, any instrument that is likely to affect all firms within the market in a similar way would be a bad candidate.

An alternative to instrumental variables (although typically imperfect) is the use of proxy variables for endogenous price changes. Note that in our model the behavioral effects are captured by lagged price changes on the right-hand side (Δp_{jt-1}) in order to allow for delays in the responses. An important potential source of endogeneity is the reaction of firms to changes in new demand or the stock of their consumers. Lagged prices are free of this endogeneity since firms cannot change their prices retroactively. As a result, Δp_{jt-1} the u_{jt} variables are uncorrelated in the absence of serial correlation. On the other hand, serial correlation may lead to endogeneity if it affects both unobservables (u) and price changes (Δp). Including contemporaneous price changes Δp_{jt} can capture serial correlation in the right-hand side variables and therefore control for endogenous price changes.[20]

Comparing lock-in in different regimes is straightforward by comparing $\hat{\delta}$ and $\hat{\theta}$ estimated from separate samples. Such estimation can be more efficient if carried out in a pooled sample with appropriate interactions with Δp_{t-1}. Interactions may be helpful in assessing the role of observable firm-specific switching cost components, too. By interacting their level with price changes in regressions (3) and (4), one can estimate switching costs $\hat{\delta}$ and $\hat{\theta}$ at different levels of observed cost components. In our example of banking loans, loan termination fees are potentially observed firm-specific switching cost components. Note however, that interactions with firms-specific switching cost components can be problematic as they are choice variables to firms. Termination fees may respond to switching itself, leading to additional simultaneity bias. As a result, observed variation in termination fees is far from ideal to address the role of monetary switching costs in consumer lock-in.

The problems listed above may or may not occur in specific applications, and they need to be

[20] This approach is sometimes called as the proxy variable solution to endogeneity (see, e.g., Wooldridge, 2002, chapter 9.2).

assessed on a case by case basis.

4.3 Measurement of the main variables

It is feasible to estimate the lock-in effects by using panel data on all firms and with information on prices and two quantities: the number of consumers joining and leaving each firm. In order to see the relationship of these quantities to probabilities n_{jt} and r_{jt}, we need to understand in detail how they are measured.

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. Firm j's stock can be therefore decomposed as:

$$S_{jt} = S_{jt-1} + IN_{jt} - OUT_{jt} - X_{jt}. \tag{7}$$

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

$$S_{jt} = S_{jt-1} + (N_{jt} + T_{jt}) - (Q_{jt} + F_{jt}) - X_{jt}. \tag{8}$$

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 who have a loan contract with bank j at the beginning of period t. The stock may change in three ways: by IN_{jt} new loans are signed, OUT_{jt} loans are repaid before their expiration date, and X_{jt} loans expire in the respective period. Q_{jt} consumers repay their loans before their expiration and quit the market, while F_{jt} consumers refinance their loans and leave bank j for another bank. Finally, the bank's incoming consumers consist of consumers who are new to the market (N_{jt}) and consumers who refinance (T_{jt}), arriving to bank j from other banks.

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

$$n_{jt} = \frac{N_{jt}}{\Sigma_k N_{kt}}, \tag{9}$$

which is simply firm j's market share from among 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 F_{jt} leave due to switching. The pool of potential switchers is therefore $S_{jt-1} - X_{jt} - Q_{jt}$, and F_{jt} of them choose to switch. The realized retention rate of firm j is therefore

$$r_{jt} = 1 - \frac{F_{jt}}{S_{jt-1} - X_{jt} - Q_{jt}}. \tag{10}$$

This probability equals the fraction of consumers staying with firm j from among all consumers who could have been with that firm.

In order to implement the estimation model specified in (3) and (4) all variables in decomposition (8) should be measured. However, firm-level aggregates usually allow us to measure the variables in decomposition (7) only. We address this problem in the next section.

In certain market contexts, all stock and flow variables can be measured in the value of contracts or revenues instead of the number of consumers. If one can measure both variables, like in our banking example, one might want to work with both in order to check the robustness of the results.

The measurement of prices is more straightforward but it is not without problems. In principle, we should keep the prices of all other firms 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 may 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 may 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 not being able to collect and process all available price information, and thus comparing the price of firm j to that of only a few competitors.[21] This latter

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

approach may serve better for markets where search costs are likely to be significant, like banking or telecommunications services.

4.4 Addressing data problems in firm-level analysis

In a typical application on firm-level data, the ideal left-hand side variables in (3) and (4) are unobserved: aggregate data on the status of the consumer in previous time periods are seldom available. However, the number of incoming and outgoing consumers from decomposition (7) is often available, and we argue that these can be used as proxies in our estimations.

We denote the proxy of n by m and the proxy of r by q, and define them as follows:

$$m_{jt} = \frac{IN_{jt}}{\Sigma_k IN_{kt}}, \tag{11}$$

$$q_{jt} = 1 - \frac{OUT_{jt}}{S_{jt-1} - X_{jt}}. \tag{12}$$

By using these proxies, our regressions to estimate become:

$$\Delta m_{jt} = \alpha_m + \beta \Delta p_{jt-1} + u_{mjt} \tag{13}$$

$$\Delta q_{jt} = \alpha_q + \gamma \Delta p_{jt-1} + u_{qjt}. \tag{14}$$

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 with (lagged) price changes, the right-hand side variable of each regression. Formally, we would need $Cov\left(\Delta d_{mjt}, \Delta p_{jt-1}\right) = 0$ and $Cov\left(\Delta d_{qjt}, \Delta p_{jt-1}\right) = 0$ to hold, where $d_{mjt} = m_{jt} - n_{jt}$ and $d_{qjt} = q_{jt} - r_{jt}$. We argue that the second covariance condition is likely to be satisfied, but the first is not.

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 (11). This proxy variable errs by potentially including switching consumers T_{jt} to firm j: $IN_{jt} = N_{jt} + T_{jt}$. Therefore the discrepancy between the ideal variable and the measured one, $d_{mjt} = m_{jt} - n_{jt}$, includes switching consumers. If price changes induce any switching, an increase in firm j's price may discourage switching consumers 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

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

The bias is due to the fact that a price increase decreases switching to the firm from other firms as well as the market share in terms of new consumers. The magnitude of the bias is related to the decrease in the probability of switching to the firm in response to the price increase. It is therefore related to the effect of the price increase on the retention probability of the other firms. The stronger the effect of a relative price change on switching to a firm from other firms is, the larger the bias in measuring the effect on new consumers.

In case of perfect lock-in, there are no consumers switching from other firms and thus there is no bias in $\hat{\beta}$. An immediate consequence is that the bias has no effect on the consistency of a test for complete lock-in because there is no bias under its null hypothesis ($H_0 : \theta = 1$). This is an important result as, in many cases, testing for perfect lock-in is one of the most important questions of the analysis.

In the absence of perfect lock-in, the bias is related to the effect of price changes on the retention probability in the market. In Appendix A, we show that under reasonable assumptions, one can approximate the upper bound to the bias by the product of the average effect of relative price changes on the retention probability in the market (γ^*) and the average relative size of the two "markets" of a firm, its old consumers and the pool of new 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 measurement problem in the retention probability is likely to cause less problems. The proxy of r_{jt} is q_{jt} as defined in (12), based on contract terminations (loan repayments in our example) before expiration 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 $OUT_{jt} = F_{jt} + Q_{jt}$ instead of F_{jt}. It therefore includes consumers Q_{jt} who repay their loans before expiration date but do not refinance at 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 $d_{qjt} = q_{jt} - r_{jt}$ is due to these two facts: the numerator and the denominator of r_{jt} are both increased by the same Q_{jt}. The discrepancy is positive, since the numerator of r_{jt} is smaller than the denominator. But this discrepancy is unlikely to lead to biased estimates in the regression. The question is whether non-refinancing terminations Q_{jt} are correlated with price changes in the previous period. Most likely they are not, because these terminations are typically due to positive income shocks, which are typically unrelated to price movements. Therefore, we can assume that $Cov\left(\Delta d_{qjt}, \Delta p_{jt-1}\right) = 0$. So estimates of γ are consistent for the same parameter as estimates of γ^* would be under ideal circumstances: $p\lim \hat{\gamma} = p\lim \hat{\gamma}^*$.As a result, if Conditions 1 and 2 are satisfied,

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

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

$$\hat{\delta}_{corrected} = \left(\hat{\beta} - a\hat{\gamma}\right) - \hat{\gamma} \tag{15}$$

$$\hat{\theta}_{corrected} = \frac{\left(\hat{\beta} - a\hat{\gamma}\right) - \hat{\gamma}}{\hat{\beta} - a\hat{\gamma}} \tag{16}$$

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

5 Illustrative application

In this section we present an application in order to show how our measurement model can be put to work. This application was part of the retail banking sector inquiry of the Hungarian Competition Authority (GVH), conducted between 2007 and 2009. The inquiry explored consumer lock-in and switching costs in relation to current accounts and bank loans, and it made recommendations to improve the effective functioning of competition in this sector. Many of the recommendations aimed at facilitating consumers' switching and constraining the market power of banks in terms of ex-post price changes by unilateral contract modifications.[22]

We focus on the market of personal loans between 2002 and 2006, i.e. loans for undetermined use that can in principle be either unsecured or secured by a mortgage. In Hungary, personal loans can be denominated either in the home currency or in foreign currencies, the latter mostly in Swiss Francs and Euros. Our application focuses on unsecured loans denominated in the home currency. This segment was the largest and most developed of the personal loan products in the sample period (the volume of foreign currency denominated loans only started to grow after 2005).

Old consumers are clients who already have a loan contract at one of the banks. In any period (quarter) they can stay at their original bank or terminate the loan contract and refinance it with a loan from another bank at a more favorable rate. The latter behavior is the switching we are interested in. There is no price discrimination between the old and the new consumers of a certain bank concerning personal loans.

The overall dataset covers the nine largest banks in Hungary that hold at least a one per cent market share on the personal loan market. Together, 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 nine banks, seven provided adequate data on the number of consumers and six on the value of terminated contracts. The unbalanced panel used in our analysis covers all or almost all observations for at least half of the banks, and at least half of the quarters for the remaining ones, leading to a total number of observations of 117 for consumer numbers and 97 for contract values. Nonrespondents were among the smaller banks. Prices p_{jt} are measured by the annual percentage rate (APR) of the banks' most popular (modal) product in terms of loan value and duration, quoted by the bank in that specific quarter. Most banks had

[22] The full report including the estimation results of this Section is avalaible in Hungarian at the GVH's homepage: http://www.gvh.hu/domain2/files/modules/module25/777170A574AD8E91.pdf

An English executive summary can be also found at

http://www.gvh.hu/domain2/files/modules/module25/8801AA394BE9C1EF.pdf a single product in the personal loan market during the period. Hungarian financial regulations require banks to calculate and publish APR for all products. By regulation, the APR includes all entry costs but not the termination costs. Table 1 shows the most important data for the market using our sample.

Table 1. Features of the personal loan market in Hungary (unsecured loans)

	2002	2003	2004	2005	2006
Number of consumers ('000)	50	78	222	413	506
Value of contracts (billion HUF)	23	42	100	167	187
Number of firms	7	7	9	9	9
Average interest rate (APR), per cent	26.8	25.8	28.4	25.9	24.3

Note. Number of consumers and value of contracts are measured as yearly averages.

The market grew dynamically during the observed time period, its growth rate slowing down somewhat after 2005. We also observe decreasing average prices,[23] some entry in the middle of the period, and larger firms with higher prices and declining market shares, facts that may be consistent with the previously discussed Beggs-Klemperer (1992) framework of dynamic competition with switching costs. Of course, observed changes in APR may be due to various reasons. In order to filter out the aggregate component of such changes, year fixed effects are included in the regressions.

The main explanatory variable is the APR relative to the market average, lagged by one period. Lagging makes sense because changes are advertised only after they are made and thus consumers are likely to react with some time lag. We estimate regressions (13) and (14) and include bank and time period fixed effects. Including bank fixed effects ensures that potential bank-specific trends do not interfere with the identification. Including time period (quarterly) fixed effects ensures that common effects on all banks do not interfere with the identification. In particular, the effects of potential changes in the outside option, business cycle or seasonality are filtered out (to the extent of a linear approximation). We do not have credible instruments for price changes nor clean natural experiments that would ensure that variation in prices is exogenous to consumer

[23] There was also a decrease in the base rate in the period examined (with a big jump upwards at the end of 2003), but the changes in the APR of the products studied was more significant.

demand. Instead, in the spirit of our discussion of the econometric problems in section 4.2, we include contemporaneous price changes Δp_{jt} next to the main variable Δp_{jt-1} as a proxy variable for potential endogeneity.

In the main text we present the OLS estimates of the coefficients β and γ and the estimates of δ and θ based on those OLS coefficients. We return to the bias-corrected estimates later. Summary statistics and the complete set of parameter estimates and regression statistics are in Appendix B (Tables B1 and B2). We present bootstrap confidence intervals on the 5^{th} and 95^{th} percentiles. These are estimated by block-bootstrap (i.e. re-sampling of complete firm histories as opposed to individual firm-year observations) in order to account for serial correlation. While the confidence intervals contain only 90 per cent of the sampling distribution, they are nevertheless rather conservative. This can be seen by comparing the bootstrap standard errors of the $\hat{\beta}$ coefficients to their analytical standard errors shown in Appendix B: the bootstrap standard errors are significantly larger.

Table 2 shows the estimates of the regression parameters and the switching cost parameters assuming no bias in $\hat{\beta}$.

Table 2. Estimates of lock-in for unsecured personal loans assuming no bias in $\hat{\beta}$

	# consumers	value
Response of new consumers (β)	-0.61	-0.74
(bootstrap SE)	(0.22)	(0.23)
(confidence interval)	$(-0.93, -0.14)$	$(-0.99, -0.22)$
Response of old consumers (γ)	-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} and 95^{th} percentiles) based on 2000 iterations.

Given the sample size, the estimates are reasonably precise. The confidence interval around the parameter of major interest, θ, is especially tight. Based on those confidence intervals, we cannot reject the null hypothesis of complete lock-in. Note that while the estimated consumer response coefficients $\left(\hat{\beta} \text{ and } \hat{\gamma}\right)$ have confidence intervals that overlap to some extent, their difference is significantly different from zero because the error terms in the two market share equations are not independent of each other.

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. In order to have a better feel for the magnitudes, we can transform the estimates into elasticities that measure the effect of a one per cent change in the APR (as opposed to a one percentage point change). The corresponding elasticities are quite strong: one per cent increase in the APR is estimated to lead to a 4.7 per cent decrease in market shares, evaluated at the average market share The same price increase is estimated to induce a mere 0.13 percentage points decrease in the probability of bank j's old consumers staying with the bank. This implies an elasticity of -0.13 (evaluated at the average retention rate of 0.98). The corresponding estimates for the value of contracts are somewhat stronger, in line with the presumption that consumers with larger contracts are more price sensitive.

As we have shown, $\hat{\beta}$ may be a biased estimator of β^*. Since $\hat{\gamma}$ is small, the bias is likely to be relatively small. In fact, the bias may be zero because we cannot reject the null hypothesis of complete lock-in. Table 3 shows the estimated lock-in effects corrected for the bias using the formulae in (15) and (16). The bias-corrected point estimates are somewhat smaller, but close to the non-corrected point estimates. The bias corrections do not change the qualitative conclusions from the previous results.

Table 3: 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 is 1.4

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

Taken together, our estimates imply that the lock-in effects of switching costs are substantial, whether measured in terms of choice probabilities or contract values.[24] According to our point estimates, old consumers with ongoing loan contracts are 70 to 80 per cent less likely to react to price changes in the market than new consumers who choose a loan product for the first time. The lock-in may be complete: we cannot reject the hypothesis that old consumers are 100 per cent prevented from switching to other banks if they face a relative price increase. Estimated lock-in effects are somewhat smaller in terms of contract value. The difference is small but it may indicate that consumers with larger contracts are somewhat less constrained by switching costs.

[24] The significance of the lock-in effects in this market is supported by other, more circumstantial evidence. From 2006, the demand for loans denominated in local currency started to decrease so it was not in the banks' interests anymore to vigorously compete for new consumers. In line with such incentives, many banks increased their interest rates significantly, especially after the financial crisis of 2008 that led to a further drop in the number potential new consumers.

We have carried out robustness checks using various specifications and various sub-periods.[25] The results indicate that the inclusion of time fixed effects is important, bank fixed effects are less important (recall that these are bank fixed effects added to panel regressions in first differences), and the contemporaneous and leaded proxy variables do not make a difference. We have also estimated regressions with the value of the termination fee included. Termination fees are one-time fees incurred when repaying a loan before expiration date and are therefore potentially important elements of switching costs. The estimated lock-in effects are stronger at higher levels of termination fees, indicating that monetary costs are indeed an important element of switching costs.

6 Conclusions

Based on a simple thought experiment, we proposed a practical method for estimating the lock-in effects in a direct way by using firm-level data. The basic idea was to compare demand responses to price changes for consumers who are new to the market and for consumers who are already customers of a given firm, and the difference should be attributed to lock-in. Implementation of the method required proxies for the following two quantities in each period: new transactions on the market and transactions (contracts) terminated by consumers. Using these proxy variables may lead to biased estimates, but we derived a way to correct for these biases.

We illustrated our method with an application to the Hungarian market of unsecured personal loans and found substantial switching costs. Old consumers' responsiveness to price changes is estimated four fifths lower than new consumers' responsiveness (or somewhat smaller but still more than two thirds lower if allowing for the bias due to measurement problems). The results indicate the existence of strong lock-in effects, to the extent that they might be consistent with complete consumer lock-in.

The Hungarian Competition Authority used this method in a detailed sector inquiry of the banking sector, and the robust conclusion was that switching costs are substantial enough to significantly raise the market power of banks, especially by ex-post fee increases. Therefore, the sector

[25] The results of the robustness checks and the estimates with termination fees are included in the Online Appendix D.

inquiry made several recommendations to the lawmakers and financial regulators that were partly aimed at facilitating the switching of consumers (developing a price comparison website in order to increase transparency, foreign currency risks should be also represented in the APR, early repayment charges should be maximized) and at constraining the market power of banks in terms of ex-post price changes by unilateral contract modifications.

References

[1] Beggs A, Klemperer P (1992), Multi-period Competition with Switching Costs,*Econometrica*, 60, 651-66.

[2] Chen P,.Hitt L (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, 255-274.

[3] Chintagunta, P (1998), Inertia and Variety Seeking in a Model of Brand-Purchase Timing, *Marketing Science*, 17 (3), 253-70.

[4] Dubé JP, Hitsch J, Rossi P (2010), State Dependence and Alternative Explanations for Consumer Inertia, *RAND Journal of Economics* 41, 417-445.

[5] Farrell J, Klemperer P (2007), Coordination and Lock-In: Competition with Switching Costs and Network Effects. In: Armstrong M, Porter R (editors), *Handbook of Industrial Organization 3* Elsevier, 1967-2072.

[6] Grzybowski L (2008), Estimating Switching Costs in Mobile Telephony in the UK, *Journal of Industry, Competition and Trade* 8, 113-132.

[7] Goldfarb A (2006), State Dependence at Internet Portals, *Journal of Economics & Management Strategy* 15, 317-352.

[8] Imbens G and Wooldridge J (2009), Recent Developments in the Econometrics of Program Evaluation, *Journal of Economic Literature*, 47, 5–86.

[9] Kim M, Kliger D, Vale B (2003), Estimating Switching Costs: the Case of Banking, *Journal of Financial Intermediation* 12, 25-56.

[10] Madrian B (1994), Employment-Based Health Insurance and Job Mobility: Is There Evidence of Job-Lock?*The Quarterly Journal of Economics,* 109(1), 27-54.

[11] Maicas J, Polo Y, Sese J (2009), Reducing the Level of Switching Costs in Mobile Communications: The Case of Mobile Number Portability, *Telecommunications Policy* 33, 544-554.

[12] National Economic Research Associates (2003), Switching Costs, *Office of Fair Trading Economic Discussion Paper* 5.

[13] Shcherbakov O (2009), Measuring consumer switching costs in the television industry, mimeo, University of Mannheim.

[14] Schiraldi P (2011), Automobile Replacement: a Dynamic Structural Approach, *RAND Journal of Economics* 42, 266-291.

[15] Seetharaman P (2004), Modeling Multiple Sources of State Dependence in Random Utility Models: A Distributed Lag Approach, *Marketing Science* 23 (2), 263-71.

[16] Seetharaman P, Ainslie A, Chintagunta P (1999), Investigating Household State Dependence Effects across Categories, *Journal of Marketing Research* 36, 488-500.

[17] Sharpe S (1997), The Effect of Consumer Switching Costs on Prices: a Theory and its Application to the Bank Deposit Market, *Review of Industrial Organization* 2, 79-94.

[18] Shum M (2004), Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast-Cereals Market, *Journal of Economics & Management Strategy* 13, 241-272.

[19] Shy O (2002), A Quick-and-easy Method for Estimating Switching Costs, *International Journal of Industrial Organization* 20, 71-87.

[20] Wooldridge J (2002), *Introductory Econometrics: A Modern Approach,* South-Western.

Appendices

A Deriving an approximation to the bias to $\hat{\beta}$

Our goal is to derive an approximation of the bias to $\hat{\beta}$, defined as $bias = Cov\left(\Delta d_{mjt}, \Delta p_{jt-1}\right)/V\left(\Delta p_{jt-1}\right)$. We want to show that, under reasonable assumptions, the following inequality holds (providing an upper limit to the bias):

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

We start from the definition of the discrepancy d_m :

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

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

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

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

The switching response to a price increase is captured by γ^{*}. Here we expand the definition of γ^{*} in order to connect it to the measure of switching consumers in the discrepancy term.

$$\begin{aligned}
\gamma^{*} &= \frac{Cov\left(\Delta r_{jt}, \Delta p_{jt-1}\right)}{V\left(\Delta p_{jt-1}\right)} \text{ and} \\
r_{jt} &= 1 - \frac{F_{jt}}{S_{jt-1} - X_{jt} - Q_{jt}}, \text{ so}
\end{aligned}$$

$$\begin{aligned}
\Delta r_{jt} &= 1 - \frac{F_{jt}}{S_{jt-1} - X_{jt} - Q_{jt}} - \left(1 - \frac{F_{jt-1}}{S_{jt-2} - X_{jt-1} - Q_{jt-1}}\right) \\
&= \frac{F_{jt-1}}{S_{jt-2} - X_{jt-1} - Q_{jt-1}} - \frac{F_{jt}}{S_{jt-1} - X_{jt} - Q_{jt}} \\
&\approx \frac{F_{jt-1} - F_{jt}}{S_{jt-1} - X_{jt} - Q_{jt}} = -\frac{\Delta F_{jt}}{S_{jt-1} - X_{jt} - Q_{jt}}.
\end{aligned}$$

Note that the magnitude of the change in switching from bank j (ΔF) is likely to have the opposite sign to the change in switching to bank j (ΔT). Assume furthermore that the two changes are similar in magnitude:[26]

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

As a result,

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

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

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

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

This completes our derivation.

---

[26] The two magnitudes are the same in the aggregate. If lock-in is sufficiently strong, the two magnitudes are likely to be small and close to each other, too. If old consumers switch to other firms primarily because their own firm changes its price (and not because of other firms' price changes), a price change by firm j may lead to ΔF smaller than ΔT in magnitude, leading to smaller bias than derived below. If lock-in is week and old consumers switch to other firms as a response to price changes at other firms, too, a price change by firm j may lead to ΔF larger than ΔT in magnitude, leading to larger bias than derived below.

B Summary statistics and complete results

Table B1. Summary statistics for unsecured personal loans denominated in home currency

	# 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

Notes.

m_{jt}: market share of firm j in quarter t in terms of newly contracted consumers

k_{jt}: fraction of consumers of of firm j in quarter $t-1$ that are retained in t

p_{jt}: price of loan charged by firm j in t, measured as annual percentage rate (APR)

Δx_{jt} : change of variable x $(m,\ k,\ p)$ from quarter $t-1$ to quarter t.

Table B2. Complete regression estimates (unsecured personal loans in home currency)

	# 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 includes the explanatory power of fixed effects.

** significant at 1%, * significant at 5%

C Details of the discrete choice background

The model describes the decision problem of consumer i who enters into a contractual relationship lasting for S periods with one of the J firms.[27] The consumer who starts buying the product from firm j may stay with this firm till the end or switch to another firm k at some period $s^* > 1$. If she switched her service provider, we continue the decision problem from period $s+1$. The problem ends at S (which may be finite or infinite). Note that since consumers arrive at different periods in terms of calendar time, customers of firm j's product may be at different contract periods s at a given calendar period t.

Let $y_{ijs} = 1$ mean that consumer i chooses firm j in period s and $y_{ijs} = 0$ otherwise. The cost of being the customer of firm j in period s will be the observed price p_{js} and an unobserved component u_{ij} that is specific to the match of individual i and firm j. We assume that the unobserved component is time-invariant, which 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.). Furthermore, if consumer i switches from firm $\dot{j}$ to firm k in period $s \in [2, S]$(that is if $y_{ij(s-1)} = 1$ and $y_{ijs} = 0$) she faces additional switching costs C_{ij} to be paid at the time of switching. Additionally, we assume that consumers cannot predict future price changes so that $E_s[p_{jr}] = p_{js}$ for $r > s$.[28]

If consumer i is a new consumer at $s = 1$, she will minimize her discounted present value of the expected per-period costs denoted by $e_{ij1} = E_1\left[\sum_{r=1}^{S}(p_{jr} + u_{ij})/(1+\rho)^r\right]$. She then chooses firm j if $e_{ij1} \leq e_{ik1}$ for $\forall k \neq j$, which condition simplifies to $u_{ij} - u_{ik} \leq p_{k1} - p_{j1}$ for $\forall k \neq j$, or in vector notation[29]

$$\mathbf{u}_{ij} - \mathbf{u}_{ik} \leq \mathbf{p}_{k1} - \mathbf{p}_{j1}. \tag{17}$$

[27] 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.

[28] We think this assumption is justified in many applications. Loan contracts provided to individuals or subscription fees usually specificy the same per-period fixed fee, while future consumption affecting variable payments (like minutes called) can usually be proxied best by current consumption.

[29] 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.

If we assume that the vector of unobservables is i.i.d. across individuals, the probability of new consumers choosing firm j at calendar time t is

$$n_{jt} \equiv \Pr\left(y_{ijt} = 1 | s = 1\right) = \Pr\left(\mathbf{u}_{ij} - \mathbf{u}_{ik} \leq \mathbf{p}_{kt} - \mathbf{p}_{jt}\right) = F\left(\mathbf{p}_{kt} - \mathbf{p}_{jt}\right) \tag{18}$$

where F is the joint c.d.f. of the unobserved cost differentials $\mathbf{u}_{ij} - \mathbf{u}_{ik}$. 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 (17)) is high enough to exceed 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 $\mathbf{p}_{kt} - \mathbf{p}_{jt}$.

Now suppose that consumer i is an old consumer of firm j in period $s \in [2, S]$, so $y_{ij(s-1)} = 1$. The expected costs of staying with firm j is $e_{ijs} = E_s\left[\sum_{r=s}^{S}(p_{jr} + u_{ij}) / \left(1+\rho\right)^r\right]$, while choosing another firm k would mean expected costs $e_{iks} = E_s\left[\sum_{r=s}^{S}(p_{kr} + u_{ik} + c_{ij}) / \left(1+\rho\right)^r\right]$, where c_{ijs} is the discounted switching cost distributed equally for all subsequent periods so that $C_{ij} = \sum_{r=s}^{S} c_{ijs} / \left(1+\rho\right)^r$ (the s subscript in c_{ijs} denotes the time period of switching). Consequently, consumer j would stay with firm j if and only if $u_{ijt} - u_{ikt} \leq p_{ks} - p_{js} + c_{ijs}$ for $\forall k \neq j$, or in vector notation if

$$\mathbf{u}_{ij} - \mathbf{u}_{ik} - \mathbf{c}_{ijs} \leq \mathbf{p}_{ks} - \mathbf{p}_{js} \tag{19}$$

where $\mathbf{c}_{ijs}$ is a vector with all elements $c_{ijs} \geq 0$.Note that 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 prohibitive the closer the end date S (the larger s).[30]

[30] The magnitude of this effect is decreased if we add period-specific switching costs or costs that are scaled directly to the remaining time (as in cases when incumbent firms make switching consumers pay a sum related to time remaining).

In this way we can write down the retention rate of firm j in $s > 1$ by

$$r_{jt} = \Pr\left(y_{ijt} = 1 | y_{ijt-1} = 1\right). \tag{20}$$

This choice probability is conditional on the individual's choice in the previous period. That choice itself was a decision to stay with firm j, too, which was again conditional on the consumer's earlier choice, etc. As a result, the retention probability is a fairly complicated function of all past prices, and solving the retention problem is beyond the scope of our paper. Instead, we focus on some intuitive implications of the retention conditions themselves.

The first immediate consequence of condition (19) is that if there are no switching costs, the condition of staying with firm j is the same as condition (17) for choosing j in the first place. This confirms the intuition behind our reduced-form approach: the price responsiveness of new consumers can be a valid approximation of the price responsiveness of old consumers in the absence of switching costs.

On the other hand, condition (19) shows that switching costs decrease the threshold that other firms' relative prices have to exceed in order for consumer i to with firm j. One consequence is that for given prices, the retention 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 (17)), own prices have to increase more (other firms' prices have to decrease more) in order to pass 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 in the absence of switching costs but whose c_{ijs} is high enough to prevent switching. As a result, the same price increase leads to a weaker average reaction of old consumers. The fraction of consumers who are prevented from switching depends on the c.d.f. of switching costs c_{ijs}, which in turn depends on the distribution of C_{ij} and heterogeneity in the 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.

Estimating consumer lock-in effects from firm-level data
Appendix D
Additional regression results

Table D1. Regression results for the entire sample period (standard errors are clustered at the bank level)

	Simplest specification				Time Fixed Effects				Time and Bank Fixed Effects			
	# Consumers		Contract Value		# Consumers		Contract Value		# Consumers		Contract Value	
	M	k	m	k	m	k	m	k	m	k	m	k
Δp_{tj-1}	-1.06	-0.06	-1.17	-0.10	-0.73	-0.12	-0.89	-0.17	-0.62	-0.12	-0.77	-0.17
	(0.25)	-0.03	(0.28)	-0.06	(0.21)	(0.04)	(0.23)	(0.06)	(0.14)	(0.04)	(0.16)	(0.06)
Obs	105	105	87	87	105	105	87	87	105	105	87	87
R-sq	0.16	0.01	0.21	0.03	0.33	0.4	0.31	0.39	0.46	0.41	0.49	0.4
Uncorrected point estimates of the structural parameters												
δ	1.00		1.07		0.61		0.72		0.50		0.60	
θ	0.94		0.91		0.84		0.81		0.81		0.78	

	Time and Bank Fixed Effects, Period t proxy				Time and Bank Fixed Effects, Period t and t+1 Proxies			
	# Consumers		Contract Value		# Consumers		Contract Value	
	m	k	m	k	m	k	m	k
Δp_{tj-1}	-0.61	-0.13	-0.74	-0.18	-0.60	-0.11	-0.70	-0.17
	(0.13)	(0.04)	(0.17)	(0.06)	(0.13)	(0.05)	(0.17)	(0.05)
Δp_{tj}	-0.62	0.11	-0.83	0.16	-0.6	0.09	-0.83	0.13
	(0.12)	(0.04)	(0.12)	(0.04)	(0.12)	(0.03)	(0.12)	(0.05)
Δp_{tj+1}					0.07	0.05	0.12	0.04
					(0.11)	(0.04)	(0.07)	(0.03)
Obs	105	105	87	87	98	98	81	81
R-sq	0.49	0.43	0.55	0.44	0.48	0.44	0.54	0.48
Uncorrected point estimates of the structural parameters								
δ	0.48		0.56		0.49		0.53	
θ	0.79		0.76		0.82		0.76	

Table D2. Regression results for the period of 2003 to 2006 (standard errors are clustered at the bank level)

	Simplest specification				Time Fixed Effects				Time and Bank Fixed Effects			
	# Consumers		Contract Value		# Consumers		Contract Value		# Consumers		Contract Value	
	m	k	m	k	m	k	m	k	m	k	m	k
Δp_{tj-1}	-1.07	-0.08	-1.16	-0.13	-0.68	-0.13	-0.88	-0.18	-0.57	-0.13	-0.76	-0.19
	(0.26)	(0.04	(0.30)	(0.07	(0.22)	(0.04)	(0.25)	(0.06)	(0.13)	(0.05)	(0.17)	(0.06)
Obs	89	89	75	75	89	89	75	75	89	89	75	75
R-sq	0.16	0.01	0.21	0.03	0.33	0.4	0.31	0.39	0.46	0.41	0.49	0.4
Uncorrected point estimates of the structural parameters												
δ	0.99		1.03		0.55		0.70		0.44		0.57	
θ	0.93		0.89		0.81		0.80		0.77		0.75	

	Time and Bank Fixed Effects, Period t proxy				Time and Bank Fixed Effects, Period t and t+1 Proxies			
	# Consumers		Contract Value		# Consumers		Contract Value	
	m	k	m	k	m	k	m	k
Δp_{tj-1}	-0.57	-0.13	-0.76	-0.19	-0.55	-0.11	-0.72	-0.17
	(0.13)	(0.04)	(0.17)	(0.06)	(0.12)	(0.06)	(0.15)	(0.06)
Δp_{tj}	-0.41	0.1	-0.55	0.14	-0.38	0.08	-0.55	0.11
	(0.10)	(0.03)	(0.12)	(0.05)	(0.10)	(0.02)	(0.12)	(0.05)
Δp_{tj+1}					0.13	0.06	0.14	0.07
					(0.15)	(0.06)	(0.13)	(0.01)
Obs	89	89	75	75	82	82	69	69
R-sq	0.61	0.44	0.7	0.41	0.6	0.46	0.69	0.46
Uncorrected point estimates of the structural parameters								
δ	0.44		0.57		0.44		0.55	
θ	0.77		0.75		0.80		0.76	

Table D3. Regression results for the period of 2004 to 2006 (point estimates and standard errors clustered at the bank level)

	Simplest specification				Time Fixed Effects				Time and Bank Fixed Effects			
	# Consumers		Contract Value		# Consumers		Contract Value		# Consumers		Contract Value	
	m	k	m	k	m	k	m	k	m	k	m	k
Δp_{tj-1}	-0.49	0.06	-0.66	0.04	-0.57	-0.04	-0.72	-0.04	-0.18	-0.06	-0.24	-0.04
	(0.38)	(0.1)	(0.42)	(0.11)	(0.43)	(0.05)	(0.5)	(0.05)	(0.27)	(0.05)	(0.22)	(0.07)
Obs	69	69	59	59	69	69	59	59	69	69	59	59
R-sq	0.03	0	0.06	0	0.04	0.42	0.07	0.33	0.46	0.43	0.66	0.35
Uncorrected point estimates of the structural parameters												
δ	0.55		0.70		0.53		0.68		0.12		0.20	
θ	1.12		1.06		0.93		0.94		0.67		0.83	

	Time and Bank Fixed Effects, Period t proxy				Time and Bank Fixed Effects, Period t and t+1 Proxies			
	# Consumers		Contract Value		# Consumers		Contract Value	
	m	k	m	k	m	k	m	k
Δp_{tj-1}	-0.17	-0.05	-0.23	-0.05	-0.15	-0.05	-0.19	-0.06
	(0.27)	(0.06)	(0.2)	(0.08)	(0.24)	(0.04)	(0.18)	(0.08)
Δp_{tj}	-0.19	-0.02	-0.11	0.07	-0.11	-0.09	-0.08	-0.01
	(0.12)	(0.1)	(0.23)	(0.15)	(0.14)	(0.13)	(0.25)	(0.2)
Δp_{tj+1}					0.25	0.11	0.17	0.08
					(0.18)	(0.07)	(0.13)	(0.03)
Obs	69	69	59	59	62	62	53	53
R-sq	0.46	0.43	0.67	0.35	0.44	0.47	0.65	0.41
Uncorrected point estimates of the structural parameters								
δ	0.12		0.18		0.10		0.13	
Θ	0.71		0.78		0.67		0.68	

Table D4. Regression results for the period of 2002 to 2005 (standard errors are clustered at the bank level)

	Simplest specification				Time Fixed Effects				Time and Bank Fixed Effects			
	# Consumers		Contract Value		# Consumers		Contract Value		# Consumers		Contract Value	
	m	k	m	k	m	k	m	k	m	k	m	k
Δp_{tj-1}	-1.11	-0.06	-1.23	-0.11	-0.77	-0.13	-0.94	-0.18	-0.67	-0.13	-0.81	-0.19
	(0.22)	(0.03	(0.25)	(0.06	(0.20)	(0.05)	(0.22)	(0.06)	(0.12)	(0.04)	(0.14)	(0.05)
Obs	77	77	63	63	77	77	63	63	77	77	63	63
R-sq	0.19	0.02	0.25	0.06	0.37	0.38	0.34	0.43	0.47	0.42	0.5	0.46
Uncorrected point estimates of the structural parameters												
δ	1.05		1.12		0.64		0.76		0.54		0.62	
θ	0.95		0.91		0.83		0.81		0.81		0.77	

	Time and Bank Fixed Effects, Period t proxy				Time and Bank Fixed Effects, Period t and t+1 Proxies			
	# Consumers		Contract Value		# Consumers		Contract Value	
	m	k	m	k	m	k	m	k
Δp_{tj-1}	-0.67	-0.13	-0.80	-0.19	-0.65	-0.12	-0.77	-0.18
	(0.12)	(0.04)	(0.15)	(0.05)	(0.12)	(0.05)	(0.15)	(0.05)
Δp_{tj}	-0.64	0.09	-0.9	0.15	-0.74	0.11	-0.96	0.15
	(0.16)	(0.04)	(0.13)	(0.06)	(0.17)	(0.03)	(0.13)	(0.06)
Δp_{tj+1}					-0.02	0.05	0.07	0.06
					(0.15)	(0.04)	(0.15)	(0.03)
Obs	77	77	63	63	70	70	57	57
R-sq	0.51	0.44	0.57	0.51	0.52	0.44	0.55	0.54
Uncorrected point estimates of the structural parameters								
δ	0.54		0.61		0.53		0.59	81
θ	0.81		0.76		0.82		0.77	0.48

Table D5. Regression results with the value of termination fee included as additional regresor (F). (Measured in HUF '00,000, min=0.02, max=0.35. Entire sample period; standard errors are clustered at the bank level)

	Simplest specification				Time Fixed Effects				Time and Bank Fixed Effects			
	# Consumers		Contract Value		# Consumers		Contract Value		# Consumers		Contract Value	
	m	k	m	k	m	k	m	k	m	k	m	k
Δp_{tj-1}	-1.06	-0.06	-1.17	-0.10	-0.73	-0.12	-0.89	-0.17	-0.62	-0.12	-0.77	-0.17
	(0.25)	-0.03	(0.28)	-0.06	(0.21)	(0.04)	(0.23)	(0.06)	(0.14)	(0.04)	(0.16)	(0.06)
$\Delta p_{tj-1} \times F_{tj-1}$	-2.26	-0.51	-3.28	-0.94	-0.87	-0.41	-2.69	-0.83	-1.42	-0.37	-3.45	-0.73
	(2.25	(0.20)	(1.99	(0.26)	(2.37	(0.16)	(2.26	(0.18)	(2.79	(0.17)	(2.52	(0.15)
Obs	69	69	59	59	69	69	59	59	69	69	59	59
R-sq	0.03	0	0.06	0	0.04	0.42	0.07	0.33	0.46	0.43	0.66	0.35
Uncorrected point estimates of the structural parameters at various levels of the termination fee												
	δ	θ	δ	θ	δ	θ	δ	θ	δ	θ	δ	θ
Minimum	0.60	1.11	0.54	1.26	0.50	0.96	0.29	1.13	0.25	0.87	-0.06	0.91
25th per cent	0.82	0.99	0.84	0.99	0.56	0.89	0.53	0.88	0.39	0.82	0.30	0.77
Median	0.91	0.97	0.96	0.94	0.58	0.86	0.62	0.84	0.44	0.81	0.43	0.78
75th per cent	1.17	0.92	1.31	0.87	0.65	0.81	0.90	0.79	0.60	0.79	0.84	0.78
Maximum	1.17	0.92	1.31	0.87	0.65	0.81	0.90	0.79	0.60	0.79	0.84	0.78

	Time and Bank Fixed Effects, Period t proxy				Time and Bank Fixed Effects, Period t and t+1 Proxies			
	# Consumers		Contract Value		# Consumers		Contract Value	
	m	k	m	k	m	k	m	k
Δp_{tj-1}	-0.20	-0.04	0.21	0.01	-0.12	-0.01	0.32	0.00
	(0.7	(0.06	(0.58	(0.06	(0.72	(0.05	(0.64	(0.06
Δp_{tj}	-0.62	0.11	-0.84	0.16	-0.98	-0.04	-1.22	0.18
	(0.13)	(0.04)	(0.14)	(0.04)	(0.32)	(0.13	(0.47)	(0.13
Δp_{tj+1}	-0.11	0.00	-0.03	0.00	0.01	0.02	0.07	0.00
	(0.04)	(0.01	(0.08	(0.01	(0.26	(0.06	(0.39	(0.01
$\Delta p_{tj-1} \times F_{tj-1}$	-1.64	-0.33	-3.65	-0.70	-1.86	-0.42	-4.01	-0.68
	(2.33	(0.14)	(1.93	(0.11)	(2.39	(0.12)	(2.14	(0.13)
F_{tj}					-0.13	-0.03	-0.12	0
					(0.27	(0.06	(0.37	(0.02
$\Delta p_{tj} \times F_{tj}$					1.27	0.53	1.37	-0.08
					(1.01	(0.43	(1.42	(0.4
Obs	69	69	59	59	62	62	53	53
R-sq	0.46	0.43	0.67	0.35	0.44	0.47	0.65	0.41
Uncorrected point estimates of the structural parameters at various levels of the termination fee								
	δ	θ	δ	θ	δ	θ	δ	θ
minimum	0.19	0.80	-0.14	1.03	0.14	0.88	-0.25	1.06
25th per cent	0.36	0.80	0.24	0.72	0.33	0.82	0.18	0.64
median	0.42	0.80	0.39	0.75	0.40	0.81	0.35	0.72
75th per cent	0.62	0.80	0.83	0.78	0.61	0.80	0.85	0.78
maximum	0.62	0.80	0.83	0.78	0.61	0.80	0.85	0.78

Table D6. Regression results with the log of the termination fee included as additional regresor (log(F+!)). (min=7.6, max=10.5. Entire sample period; standard errors are clustered at the bank level)

	Simplest specification				Time Fixed Effects				Time and Bank Fixed Effects			
	# Consumers		Contract Value		# Consumers		Contract Value		# Consumers		Contract Value	
	m	k	m	k	m	k	m	k	m	k	m	k
Δp_{tj-1}	1.85	0.99	3.41	1.84	-0.80	0.88	2.11	1.72	1.31	0.77	4.89	1.48
	(5.93	(0.55	(5.73	(0.72)	(6.01	(0.37)	(6.02	(0.32)	(6.83	(0.37)	(6.51	(0.26)
$\Delta p_{tj-1} \times F_{tj-1}$	-0.29	-0.10	-0.46	-0.19	0.01	-0.10	-0.30	-0.19	-0.19	-0.09	-0.56	-0.16
	(0.58	(0.05	(0.57	(0.07)	(0.59	(0.03)	(0.59	(0.03)	(0.67	(0.04)	(0.64	(0.02)
Obs	105	105	87	87	105	105	87	87	105	105	87	87
R-sq	0.17	0.02	0.23	0.05	0.33	0.4	0.32	0.4	0.48	0.41	0.5	0.41
Uncorrected point estimates of the structural parameters at various levels of the termination fee												
	δ	θ	δ	θ	δ	θ	δ	θ	δ	θ	δ	θ
minimum	0.58	1.65	0.48	5.57	0.84	1.17	0.45	2.62	0.22	1.64	-0.37	0.58
25^{th} per cent	0.97	1.03	1.03	1.01	0.62	0.88	0.67	0.86	0.42	0.82	0.44	0.88
median	1.02	1.00	1.10	0.96	0.59	0.84	0.70	0.81	0.45	0.79	0.55	0.84
75^{th} per cent	1.13	0.95	1.26	0.89	0.53	0.76	0.76	0.74	0.51	0.75	0.78	0.80
maximum	1.13	0.95	1.26	0.89	0.53	0.76	0.76	0.74	0.51	0.75	0.78	0.80

	Time and Bank Fixed Effects, Period t proxy				Time and Bank Fixed Effects, Period t and t+1 Proxies			
	# Consumers		Contract Value		# Consumers		Contract Value	
	m	k	m	k	m	k	m	k
Δp_{tj-1}	2.29	0.59	5.99	1.28	2.56	0.87	6.37	1.29
	(5.56	(0.34	(4.86	(0.26)	(6.07	(0.33)	(5.67	(0.38)
Δp_{tj}	-0.61	0.11	-0.84	0.15	-1.92	-1.21	-2.79	0.11
	(0.11)	(0.04)	(0.13)	(0.04)	(2.66	(1.02	(3.94	(0.89
Δp_{tj+1}	-0.02	0.00	-0.01	0.00	-0.02	0.00	-0.01	0.00
	(0.01)	0.00	(0.02	0.00	(0.04	(0.01	(0.06	0.00
$\Delta p_{tj-1} \times F_{tj-1}$	-0.29	-0.07	-0.67	-0.14	-0.32	-0.10	-0.70	-0.15
	(0.54	(0.03)	(0.48	(0.02)	(0.59	(0.03)	(0.56	(0.04)
F_{tj}					0	0	0	0
					(0.04	(0.01	(0.06	0
$\Delta p_{tj} \times F_{tj}$					0.13	0.13	0.19	0
					(0.26	(0.1	(0.38	(0.09
Obs	105	105	87	87	105	105	87	87
R-sq	0.51	0.43	0.56	0.44	0.51	0.44	0.57	0.44
Uncorrected point estimates of the structural parameters at various levels of the termination fee								
	δ	θ	δ	θ	δ	θ	δ	θ
minimum	-0.03	0.32	-0.68	0.76	-0.02	0.14	-0.90	0.86
25^{th} per cent	0.42	0.83	0.39	0.85	0.43	0.82	0.21	0.58
median	0.48	0.82	0.54	0.83	0.49	0.80	0.37	0.65
75^{th} per cent	0.60	0.81	0.84	0.82	0.61	0.78	0.67	0.71
maximum	0.60	0.81	0.84	0.82	0.61	0.78	0.67	0.71