Price negotiation in differentiated product markets: The case of insured mortgages in Canada.*

Jason Allen
Bank of Canada

Robert Clark
HEC Montréal

Jean-François Houde
UW-Madison

PRELIMINARY: PLEASE DO NOT CITE

June 21, 2011

*Correspondence to Jason Allen: Bank of Canada, Ottawa, Ontario, K1A 0G9; Phone: (613) 782-8712; Email: jallen@bankofcanada.ca, Robert Clark: HEC Montréal, CIRANO and CIRPÉE, Montreal, Quebec; Phone: (514) 340-7034; Email: robert.clark@hec.ca, Jean-François Houde: University of Wisconsin-Madison, Madison, Wisconsin and CIRANO; Phone: (608) 262-3805; Email: houdejf@ssc.wisc.edu. This research has benefited from the financial support of the NSF (SES-1024840).

1 Introduction

The goal of this paper is to document and evaluate market power in the Canadian mortgage market. We propose and estimate a model of mortgage-choice and measure the importance of search costs, switching costs and branch network size in generating market power.

A key feature of mortgages markets is that banks are able to price discriminate by setting different rates for different consumers. Like many markets (cars, insurance, etc), although sellers post one price, there is room for discounting. Despite its prevalence, this form of pricing has been largely ignored by researchers studying market power in differentiated products markets. But, ignoring the actual pricing mechanism can lead to an incomplete and biased analysis. Specifically, the researcher has no knowledge of the distribution of rejected prices. To the extent that transaction prices reveal something about the valuation of consumers for the product that they choose, this can lead to a biased estimate of preferences. There have been two main approaches to solving this problem. In their study of the demand for new automobiles, Berry, Levinsohn, and Pakes (2004) ignore transaction prices, abstracting away from the price setting mechanism actually used in the market. In contrast, in their analysis of sub-prime used car loans, Adams, Einav, and Levin (2009) assume monopoly pricing.

In this paper we study the price setting mechanism and the importance of market power in the mortgage market. For most households a mortgage represents its largest debt. Allen, Clark, and Houde (2011) find that the difference in total interest payments on a 5 year fixed-rate mortgage between someone in the 25th percentile discount and the 75th percentile is $12,468. There are substantial gains, therefore, in moving up the distribution. The size of discount a consumer can negotiate should depend on the number and characteristics of lenders present in his/her local market. In practice, however, predicting discount sizes is complicated by the fact that financial institutions are selling services that are differentiated. For instance, the location of retail branches determines the cost of shopping for mortgages (horizontal differentiation), while the quality of complementary services (e.g. convenience of teller/ATM network) affects the value of contracting with each institution (vertical differentiation). Consumers also differ in terms of their preferences for these amenities, as well as their ability to negotiate the best deal.

To shed light on these issues, we analyze detailed transaction-level data on a large set of approved mortgages in Canada between 1999 and 2004 and administered by either the Canadian Mortgage and Housing Corporation or Genworth Financial. These data provide information on features of the mortgage, household characteristics (including place of residence), and market-level characteristics. The richness of these consumer data in combination with lender-level location data (MicroMedia ProQuest) allow us to empirically examine the functioning of this important market.

To estimate the importance of switching costs, search, and branch size in the mortgage market, we propose and estimate a sequential search model where consumers are initially matched with a home bank to obtain a mortgage quote. For a given search cost consumers decide whether or not to gather quotes from the other banks in their neighborhood. The home bank doesn't know the consumers true search cost but tries to screen consumers who are unlikely to search. This is the source of the incumbency advantage. Banks with a larger consumer base are more likely to be initially matched with consumers, and therefore more likely to match with consumers that have worst outside options. Larger network banks also benefit from differentiation. To the extent that there are switching costs from moving a checking account, investment account, etc., consumers are willing to pay a higher price to stay with their home bank. The key parameters of the model are the mean and variance of search costs, the home bank premium, and the marginal utility of branch network size. We find that the average search cost is equal to 5% of the monthly mortgage payment, or approximately $56 per month. There is substantial dispersion in search costs, however, with roughly half the borrowers in the sample with search costs estimated below $38 per month. The home bank premium is slightly less than the search cost parameter, approximately $40 per month. That is, on average consumers are willing to pay $40 a month to stay with their home bank; in other words they are willing to forgo $40 a month to avoid switching banks. The final key parameter is the marginal utility of network size. In this case, the borrowers valuation is relatively small, approximately $3 per month. The results suggest the premium we observe for banks with large networks and the discounts received by switchers comes primarily from search and switching costs.

The paper is organized as follows. Section 2 presents a description of the household-level data, a discussion of the mortgage industry, and descriptive regressions analyzing interest rates. Section 3 presents the model. Section 4 discusses the estimation strategy and the empirical results. Section 5 concludes.

2 Data

This section describes the data as well as some key institutional features of the industry. Section 2.1 gives details on mortgage contracts, as well as the features of the Canadian mortgage market that are relevant for understanding the main results. Section 2.2 describes the structure of the industry. Section 2.3 presents descriptive regressions analyzing the extent of interest rate dispersion and shopping behavior. A detailed description of the data can be found in Allen, Clark, and Houde (2011).

2.1 Mortgage contracts and sample selection

Our main data-set is a sample of insured contracts from Canada Mortgage and Housing Corporation (CMHC) and Genworth Financial between 1999 and 2004. Over this period mortgage insurance was required for households borrowing more than 75% of the cost of the home from a regulated financial institution.[1] Furthermore, mortgage insurance is for the life of the contract. The result is that about 73 per cent of all residential mortgages are insured. We obtained a 10% random sample of contracts of the 12 largest lenders from CMHC and the full set of contracts of the 12 largest lenders from Genworth Financial. We further sample from the Genworth contracts to match their annual market share, which by 2004 was approximately 30%.

In total we have access to 20 household/mortgage characteristics, including all of the financial characteristics of the contract (i.e. rate, loan size, house price, debt-ratio, risk-type), and some demographic characteristics (e.g. income, prior relationship with the bank, residential status, dwelling type). Table 10 in the Appendix lists all of the variables included in data-set. In addition, we observe the location of the purchased house up to the forward sortation area. While the average forward sortation area (FSA) has a radius of 7.6 kilometers, the median is much lower at 2.6 kilometers.[2]

With respect to the lender information, we signed confidentiality agreements with the 12 largest lenders in order to link each contract with a financial-institution. For the remaining contracts, we only know wether the lender is a bank, a credit-union, or a trust/insurance company. We come back to the description of the market structure in the next section.

We restrict our sample to contracts with homogenous terms. In particular, from the original sample we select contracts that have the following characteristics: (i) 25 years amortization period, (ii) 5 year fixed-rate term, (iii) newly issued mortgages (i.e. excluding refinancing and renewal), (iii) contracts that were negotiated individually (i.e. without a broker). A 5 year fixed-rate mortgage contract must be renegotiated every five

[1]Today mortgage insurance is required on all contracts where the amount borrowed is more than 80%.
[2]The FSA is the first half of a postal code. We observe nearly 1,300 FSA in the sample.

Table 1: Summary statistics on contractual characteristics in the full sample

	Distribution observations	
	Number	Fraction
New home purchase	139,488	0.866
25 Years amortization	143,193	0.889
Fixed-rate term	145,770	0.905
5 Years term	134,173	0.833
Big-12 Bank	133,045	0.826
Non-broker transaction	100,467	0.698
Missing values (broker, fico, residential status)	17,074	0.106
Total sample size	53,154	0.33

years, which in effect acts like an adjustable rate mortgage with a fixed time-frame to renegotiate. This contract type has traditionally been the most popular in Canada, although we do observe a slight shift in favor of short-term and variable-rate contracts over the last two years of our sample. In addition, we drop contracts that were initiated by smaller institutions that remained anonymous, as well as contracts with missing values for key attributes (e.g. credit score, broker and residential status). Table 1 illustrates the breakdown of the full sample according to those characteristics. The final sample includes slightly more than fifty thousand observations, or 33% of the initial sample. Most of this drop originates from omitting broker transactions, which represent more than 30% of newly issued mortgages.

Table 2 describes the main financial and demographic characteristics of the borrowers in our sample, where we trim the top and bottom 1% of observations in terms of income, loan-size, and interest-rate premium. The resulting sample corresponds to a fairly symmetric distribution of income and loan size. The average loan size is nearly $\$140,000$ which is twice the average annual household income. The total debt service (or TDS) ratio is capped at 40%, but there are some consumers that are not constrained by this maximum. Figure 1b illustrates the distribution of TDS between 1999 and 2004. From

Table 2: Summary statistics on mortgage contracts

	N	Mean	SD	Min	Median	Max
Loan (X100K)	47,039	1.39	.548	.425	1.31	3.16
Income (X100K)	47,039	.681	.258	.161	.644	2
Other debt (X1000)	47,039	.862	.527	.00143	.761	5.04
LTV	47,039	.91	.0442	.75	.907	.95
FICO (mid-point)	47,039	.672	.0691	.5	.7	.75
Switchers	35,560	.187	.39			
Renters	47,039	.488	.5			
Living with parents	47,039	.0709	.257			

Sample: 5-year fixed-rate contracts issued by one of the Big-12 lenders between 1999 and 2004. Contracts negotiated through brokers are excluded. The sample also excludes top and bottom 1% of the loan size distribution.

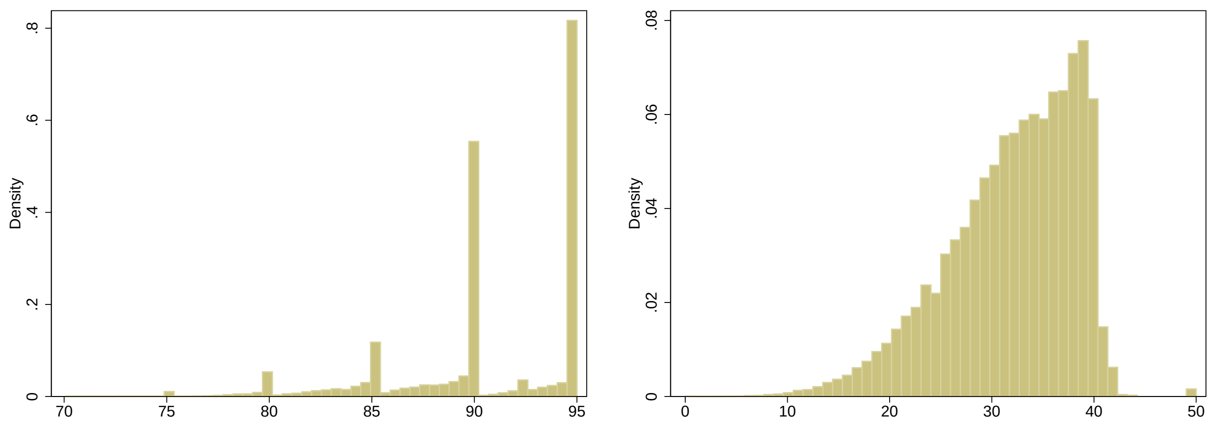
this variable we construct a measure of the total other monthly debt payments subtracting the mortgage payments from the total debt services. On average households monthly debt payments other than the mortgage are $862.

The loan-to-value (LTV) variable shows that many consumers are constrained by the minimum down-payment of 5% imposed by the government. Nearly 40% of households invest the minimum, and the average loan-to-value is 91%. Figure 1a plots the distribution of the LTV ratio. LTV ratios are highly localized around 90 and 95, and to a lesser extent 75, 80, and 85. The clustering comes about because the insurance premium schedule is discrete, and there are only a small number of price-quantity pairs. Moreover, the vast majority of households in our data (i.e. 96%) roll-over the insurance premium into the initial mortgage loan. As a result, those households pay interests on the insurance premium, in addition to the premium itself.

The variable labeled “switchers” is a dummy variable equal to one if the duration of the prior relationship with the mortgage lender is zero. Slightly more than 80% of households choose a lender with which they already have a prior financial relationship. The fraction of switchers is significantly larger for new home-buyers (i.e. formerly renters or living with their parents). The mean borrower has been with his/her financial institution

Figure 1: Loan to Value and Total Debt Service Ratios: 1999-2004

(a) Loan to value (b) Total debt service



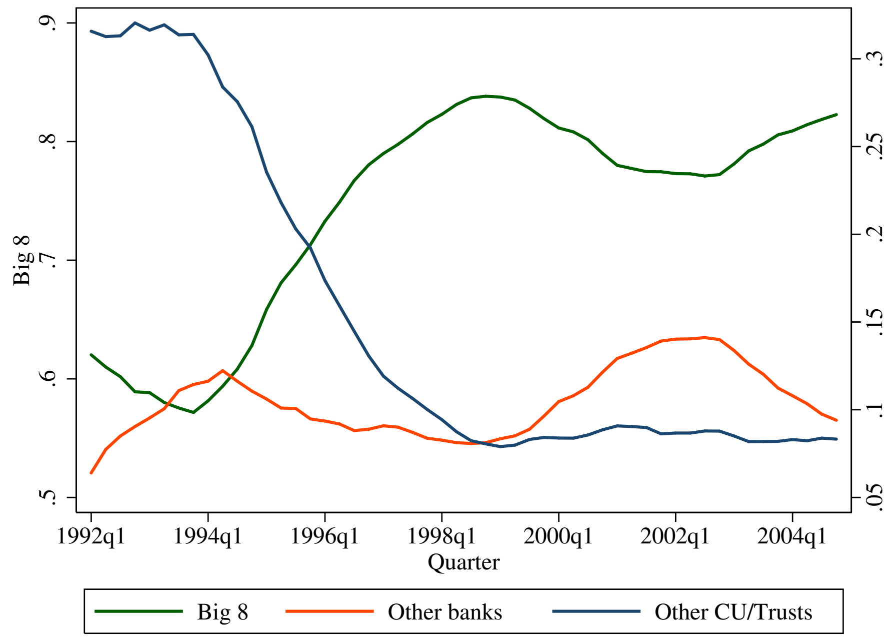
48 months before the contract is signed, about 6 months more than the mean for new home-owners and 20 months less than the mean for previous home-owners (unreported statistics).

2.2 Market Structure

The Canadian mortgage market is dominated by the "Big 6" Canadian banks (Bank of Montreal, Bank of Nova Scotia, Banque Nationale, Canadian Imperial Bank of Commerce, Royal Bank Financial Group, and TD Bank Financial Group), a regional cooperative network – Desjardins – and a provincially owned deposit-taking institution – Alberta's ATB Financial. Collectively, they control 90 per cent of assets in the banking industry and are called the "Big 8". We also includes three other regional institutions in our sample: HSBC, Vancity, Canada-Trust (acquired by TD in 2000), and Banque Laurentienne. We label this set the Big 12.

The dominance of the Big-8 stems from the period of consolidation that occurred throughout the 1990s, when the large banks acquired nearly all of the trust companies, who until that point had played an important role in the mortgage market. Poor loans

Figure 2: Market shares of newly issued insured mortgages between 1992 and 2004



in the 1980s left the trust companies or their holding companies in financial distress. As a response to these troubles, and to the fact that trust companies had an unfair legislative advantage when it came to making loans (having to do with reserve requirements), legislative changes took place in 1992 to allow banks to enter the trust business. Figure 2 illustrates the changes in the distribution of new mortgages market shares following this reform. After the last merger in 2000, between Toronto-Dominion and Canada Trust, the market remained relatively stable. The 8 largest lenders jointly control 80% of the mortgage market. This statistics is higher when we excludes broker transactions, which predominantly deal with smaller financial institutions and trust companies (unreported). Allen, Clark, and Houde (2011) provide evidence on how the evolution of the Canadian banking sector led Canadians to treat their primary bank as a “one-stop shop” (universal bank) where they purchase the majority of their financial services.

We characterize the market structure facing each consumer by matching the house location with the postal code associated with each financial institution's branches between 1999 and 2004. The branch location data comes from Micromedia-ProQuest; a provider of commercial address information in Canada. The information relative to the location of each house is coarser than the location of branches. Therefore, we assume that each house is located in the center of its FSA, and calculate a somewhat large euclidian distance radius of 10 KM around it to define the borrower's choice-set. This radius is defined such that the vast majority of contracts are signed with a bank located within each consumer's choice-set.

Figures 3 illustrate the distribution of minimum distances between each house's FSA centroid and the closest branch of each lender. On average consumers transact with banks that tend to be located close to their house. The average minimum Euclidian distance nearly 2 KM for the chosen institution, and above three for the other lenders. In fact the distributions indicate that 80% of consumers transact with a bank that has a branch within 2 KM of their new house, while only 30% of consumers have an average distance to competing lenders lower or equal to 2 KM.

This feature reflects the fact consumers tend to choose lenders with large networks of branches. In Table 3 we measure the average network size of the chosen institution relative to the average size of others present in the same neighborhood (i.e. relative network size). On average consumers transact with lenders that are nearly 60% larger than their competitors in terms of branches; the median is smaller at 28%.

The remaining variables in Table 3 measure the level of concentration aggregated at the census-division level. On average each consumer faces six lenders within 10 KM. Most of these banks have a relatively small presence, as indicated by the large Herfindahl-Hirschman index, calculated using the distribution of branches within 10 KM of each contract (i.e. both the mean and median are above two thousands). The C1 and HHI-contract measures also suggest a lack of competition. On average, the top lender in each

Figure 3: Distribution of minimum distances between banks and consumers

(a) Distance to chosen lender (b) Distance to other lenders

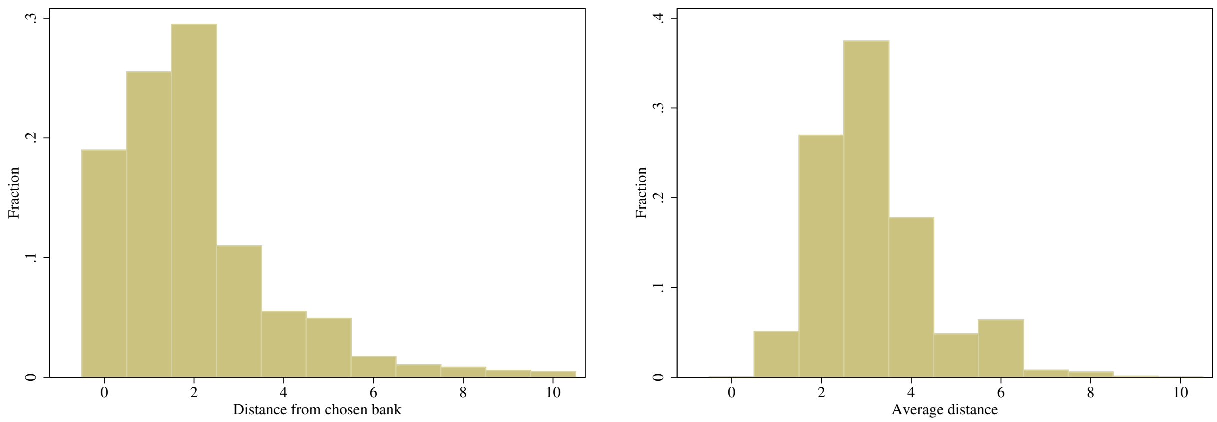


Table 3: Descriptive statistics on local market structure

	Mean	**Min**	**P25**	**P50**	**P75**	**Max**
Nb. contracts	455	11	29	169	410	4288
Nb. FIs (in 10 KM)	6.09	2	5.18	6.12	7.03	8.12
HHI-Branch (in 10 KM)	2240	1527	1874	2089	2325	5370
C1-Contract	41.4	21.6	29.2	36.8	48.5	90
HHI-Contract	1304	338	517	762	1424	7300
Relative network size	1.58	.831	1.11	1.28	1.52	10.6

Markets are defined as census-divisions (130 obs.). Sample excludes market with less than 10 contracts between 1999 and 2004, and only includes contracts with Big-12 lenders.

region controls 41.4% of contracts. The HHI-contract variable suggests a somewhat lower level of concentration, although this variable is subject to measurement error due to the small sample in some region. This difference nonetheless suggests that, although the top lender in each region has a disproportionately large share, the remaining contracts are distributed more uniformly across other banks.

2.3 Descriptive evidence on discounting

Most Canadian banks operate nationally and post prices that are common across the country. Lenders typically post the mortgage rate for their different products on a weekly basis in both national and local newspapers, as well as online. Moreover, there is little dispersion in posted prices, especially among the Big six financial institutions. In fact, the coefficient of variation on posted rates for the Big six during the early part of our sample period is always around zero. Allen and McVanel (2009) provide a detailed analysis of movements in Canadian banks' posted rates. Allen, Clark, and Houde (2011) provide a detailed description of mortgage discounting in Canada. Here we focus on a substantially smaller set of contracts in than in Allen, Clark, and Houde (2011).

When shopping for a mortgage contract one option for consumers is to pay the posted price of their home bank or of some rival bank. However, in Canada this is not their only option. Local branch managers have the authority to offer borrowers discounts below the posted price under general guidelines from headquarters. Rather than settle for the posted price consumers can instead try to obtain mortgage contracts with lower rates. There are in general two ways for them to do this: (i) negotiating directly with branch managers by gathering, or threatening to gather, additional quotes, or (ii) hire a broker. In this paper we focus on the first option, and discard all contracts initiated by a broker.

Our data do not provide direct information on the number of quotes gathered by borrowers. However, survey evidence from CAAMP reveals that on average borrowers negotiate with between one and two financial institutions when searching for a rate, and between 46% and 61% of first-time home buyers gather multiple quotes. Table 4 reproduces these statistics from an annual survey conducted by CAAMP.

This survey also reaffirms the leading role played by the main institution of consumers; defined as the one with which borrowers conduct day-to-day banking activities. In 2004, 80% of new borrowers reveal that they contact their main financial institution

Table 4: Summary statistics on shopping habits

	1999	2000	2001	2002	2003	2004
Contact main FI						80%
Contact other FI						32%
Number of FI contacts						2-4
Number of quotes						1-2
Several rate offers	61%	56%	46%	57%	51%	
Arranged via Broker	18%	26%	38%	22%	30%	32%
Loyalty to main FI	57%	57%	48%	63%	63%	54%

Source: Canadian Association of Accredited Mortgage Professionals (CAAMP). Each entry estimates the average answer of new home-buyers.

when shopping for their mortgage. Moreover, depending on the year, nearly 60% of new home-buyers remained loyal to their main institution. This statistic is smaller than what is suggested by our sample, in which 19% of borrowers contracted with a financial institution with whom they have no prior experience. The two numbers are not directly comparable however, since the CAAMP survey focusses only on new-home buyers. In our sample, 44% of contracts originate from borrowers who already own a house, and first-time home buyers are more likely to switch institutions.

The fact that transaction interest rates are negotiated rather than posted induces a substantial amount of dispersion. Table 5 measures the dispersion in transaction interest rates (in logs). Between 1997 and 2004, which includes the sample that we study in this paper, the standard-deviation of log-rates was 0.0924 after removing the contribution of aggregate trends in the level of interest rates (see table footnote). The residual dispersion of log-rates, which conditions on observable financial attributes of the contract, is very similar (0.087), suggesting that most of the dispersion is idiosyncratic and driven by non-financial attributes.

The table also illustrates an increase in the amount of dispersion over time. This trend is mainly due to the fact that the fraction of consumers paying the posted-rate went from 40% in the first half of the 1990s, to 16% after 1997.

Table 5: Evolution of interest rate dispersion and variance decomposition

	1992-1996	1997-2004
Fraction paying posted rate	0.4019	0.1655
Std.-Deviation: log-rate	0.0616	0.0924
Std.-Deviation: residual log-rate	0.0610	0.0876
Variance decomposition (fraction between):		
Lender	0.0232	0.0289
Neighborhood (fsa)	0.0238	0.0702
Contractual characteristics (37)	0.0193	0.0588
HHI-Branch (10 km)	0.0037	0.0118
Number Lenders (10 km)	0.0025	0.0079

The log-rate is expressed in deviation from month/year fixed-effects in order to remove trends in the level of interest rates. The "residual log-rate" variable is obtained by projecting the natural log of transaction interest rate onto month/year fixed-effects, and financial characteristics of the contract (i.e. loan-size, income, fico score, ltv). The "contractual characteristics" is a set of 37 discrete categories of contracts based on the income, loan-size, and loan to value. The "fraction between" measures the ratio of the between group variance over the the total variance.

The bottom half of Table 5 presents the contribution to the variance of systematic differences across borrowers and local markets. The results suggest again that most of the observed dispersion is idiosyncratic. The ratio of the variance between groups over the total variance ranges from 1% to 7%. The location of houses is the most disaggregate category and explains only 7% of the variance. It captures systematic regional differences across market structure and consumers. Over time however, the contribution of all categories increased significantly, suggesting that negotiated rates reflect more closely the characteristics of consumers.

Next we analyze the relationship between transaction rates and observed consumer and market attributes. To remove aggregate trends in interest rates, we measure the margin of lenders by the 5 year bond rate from the transaction interest rate. The average margin in the data is 1.20 percentage point (standard-deviation of 0.67), which is slightly higher than the average discount based on the posted price (i.e. mean and standard deviation of $0.91/0.62$). Table 7 presents the results of five linear regressions of transaction margins on financial and demographic characteristics of borrowers, as well as market structure controls. The first three columns analyze the relationship between income, loan size, and transaction rates. Without conditioning on loan size, the results suggest a negative relationship between rates and income, which in part is associated with the wealth of the residential neighborhood (i.e. column (2) versus (1)). The sign of this relationship is reversed in columns (3) to (5), suggesting that larger loans are associated with lower transaction rates. In fact the marginal effect of income in specification (5) and (6) is positive and statistically different from zero for most observed contracts, except for richer households with relatively small loans. The remaining coefficients associated with financial characteristics reveal that consumers with high risk scores and that are not financially constrained tend to pay lower rates.

These results suggest that consumers who are more "profitable" for the lenders, tend to receive higher discounts. The profitability of the transaction here does not only reflect the risk of default (where the cost of default is negligible given that the contracts are insured), but rather the overall value of serving consumers, which include complementary services offered by the banks and the compensation of branch managers (which could be increasing in the size of loans). There is also some risk of prepayment by the borrower which would reduce the revenue earned directly from the mortgage contract. More specifically, in some cases borrowers will contribute over and above their monthly payment to pay down the mortgage more quickly than the lender expected. Although the amount of prepayment is typically capped at 15 per cent of the value per year, we know from other sources that richer households are more likely to use weekly or bi-monthly payments, and therefore might also be more likely to prepay in advance a fraction of their mortgage. This would explain in part the positive relationship between income and interest rate, conditional on loan-size.

The next four coefficients are related with consumers' search effort and valuation of banks attributes. First, we learn that new home-buyers, and especially consumers who used to live with their parents, receive larger discounts. Similarly, consumers switching to a financial institution with whom they have no prior relationship pay significantly lower rates than loyal consumers. In both cases, the discounts represent nearly 8 basis points, or 6.6% of the average margin. The coefficient associated with the relative network size also shows that consumers banking with relatively large banks in their neighborhood pay higher rates.

These correlations are consistent with two interpretations. On the one hand, the fact that a larger proportion of consumers combine all of their financial services under one bank means that consumers value the complementary services offer by banks, which is increasing in the network size. This extra premium in the willingness to pay is also enhanced by the presence of switching costs, which could explains in part the switching coefficient. On the other hand, if consumers differ in their ability to search or gather multiple quotes, loyal consumers and consumers matched with large network institutions could represent a larger proportion of non-searchers. The model we present in the next section accommodates both interpretations, and tries to measure their relative importance.

The last four variables in Table 7 illustrate the negative relationship between the number of lenders and the transaction rates. As seen earlier, most consumers have at least 6 lenders in their choice-set, and very few have less than 5. As a result the competitive effect of concentration can be seen by comparing local markets with 7 or more lenders to markets with less than 7. Pass this threshold the relationship is significant. Consumers in local markets with 7 lenders pay on average 5 basis points less than markets in the reference group, and nearly 15 points less in local markets with 9 or more lenders. The magnitude of this correlation is reduced when we control for census variables at the FSA-level, and even further when we add location fixed-effects. The relationship between the number of lenders and rates, however, remain statistically significant in most cases.

In Table 7 we study the relationship between borrower characteristics and the probability of switching institutions. As we alluded to earlier new home buyers are more likely to switch, especially consumers formerly renting an apartment. Moreover, consumers who transact with larger network institutions are less likely to switch, suggesting as before that consumers matched with dominant banks are more likely to be loyal. The number of available options also influences the decision to switch. Consumers located in less competitive markets (i.e. fewer than seven lenders) are less likely to switch.

Finally, the loan to income ratio indicates that consumers shopping for a larger loan are less likely to remain loyal to their home institutions, while richer household are more likely. This relationship remains significant when we include loan and income as linear terms as well. It is consistent with the previous results relating income and loan size with transaction rates. Since switching in our context proxies for the search effort of consumers, this result importantly suggests that high income consumers are less likely to gather multiple quotes, while the opposite is true for consumers with larger loans.

3 Model

We propose a sequential search model in which consumers with heterogeneous search costs are initially matched with their "home" bank to obtain an initial quote, and then decide to keep searching by gathering multiple quotes from the remaining lenders in their local market. The initial stage is analogous to a bargaining model with incomplete information. The home bank makes an initial offer without knowing the cost for the consumer of gathering multiple quotes, and tries to screen consumers who are unlikely to search. This is the first source of the incumbency advantage: banks with a large consumer base are more likely to be initially matched with potential clients, and therefore more likely to transact with consumers with low outside option.

Large network banks benefit from a second advantage because of differentiation. All else being equal, the model allows consumer to value their home bank more, and/or lenders with a large network of retail branches. Quality differentiation in this market arises because banks are multi-product firms, and a large fraction of consumers combine their day-to-day banking and lending transactions with the same institution. Moreover, to the extent that consumers face a switching cost to transact with a different bank than their home bank, the average willingness to pay for the home will be higher.

We describe the model in details in the next three subsections. First, we describe additional notation, and formally define the timing of the model. Then we solve the model backward starting with the competition stage after consumers reject the initial offer. The last subsection describes the distributional assumptions that we use to estimate the model.

3.1 Timing and some notation

The timing of the model is as followed. First, consumers obtain an initial quote p^0 from their home bank h. Consumers can accept or reject this offer. If the offer is rejected, consumers organize a multi-lateral negotiation between the $\mathcal{N}$ lenders located in their neighborhood. At this point, the initial offer serves as a ceiling, since consumers have the option of recalling p^0 in the event that none of the competing offers dominate the initial offer (which will not happen in equilibrium because of our ll information assumption). Consumers evaluate the offers by trading-off the value of banking with each lender, denoted by θ_{ij}, and the cost of the contract offered (i.e. p_j). In this version of the model, we assume that firms and consumers have symmetric information with respect to the profitability of serving each consumer; a feature that is suggested by the previous regression results.

We assume the following payoff functional forms:

$$\text{Consumers:} \quad U_{ij} = \theta_{ij} - p_{ij} \tag{1}$$

$$\text{Firms:} \quad \pi_{ij} = \underbrace{L_i(r_j - c_i) + \omega_{ij} + u_{ij}}_{\text{Profits from lending + complementary services}} \tag{2}$$

where $p_{ij} = L_i \times r_j$ is the cost of the contract for consumers, and L_i is the fixed loan size.

The cost of the contract p_{ij} measures the monthly payment made by household i. We assume that households borrow the amount of the insurance premium at the same interest rate. Therefore L_i incorporates the lump-sum insurance premium paid by the banks to the insurance company. This insurance premium is an increasing function of the loan-to-value ratio.

The marginal cost of lending c_i is common across banks and is equal to the 5 years bond rate at the starting date of the contract. The scale of r_j and c_i measure the monthly payment that consumers and banks have to incur on a loan of size 1 over a common amortization period of 25 years.

The lender's profits include an additional component measuring the indirect profit earned through complementary services offered by the bank. It is further decomposed into a function of consumers and bank characteristics $\omega_{ij} = Z_{ij}\gamma + \epsilon_i$, and a random variable u_{ij} unobserved to the econometrician. The random variable ϵ_i is an unobserved attribute of consumer i that affects banks' profits symmetrically.

The value of banking with bank j for consumer i is a function of the "quality" of bank j's services in the neighborhood of consumer i, and a premium earned by consumers to accept a contract from a bank with which they have prior experience. We measure the quality of banks services by the density of their branch network, denoted by q_{ij}. Furthermore, the experience of of consumer i with bank j, denoted by E_{ij}. We assume that consumers have positive months of experienced with at most one bank. The willingness to pay function is expressed as:

$$\theta_{ij} = q_{ij}\alpha + \lambda 1(E_{ij} > 0) \tag{3}$$

This willingness to pay function implies that firms are vertically differentiated. While we assume that the marginal utility for quality is common across consumers, the size of the loan affects the ranking of offers. This is because the scale of p_i in function of L_i, while θ_{ij} is common across consumers facing the same choice-set and having the same home bank. As a result, everything else being equal, consumers with larger loans are more likely to search and choose the lowest rate (as opposed to the highest quality option).

3.2 Equilibrium prices

The model is solved backward. Conditional on rejecting p^0, all lenders in the choice-set $\mathcal{N}_i$ compete for the contract. We assume that firms have full-information: firms observed θ_{ij}, ϵ_i and u_{ij} for all j. They simultaneously offer a price r_j and consumers pick the best offer, based on U_{ij}. At this stage the lender generating the highest surplus wins the contract, and the transaction surplus is defined by:

$$V_{ij} = \theta_{ij} - L_i c_i + \omega_{ij} + u_{ij}. \tag{4}$$

If bank j is the highest surplus lender, the Bertrand-Nash equilibrium price is defined by the zero profit condition associated with the second-highest surplus bank. That is, bank j offers a rate r_j^* such that the consumer is indifferent between bank j and the zero-profit offer from the second-highest surplus bank:

$$\theta_{ij} - L_i r_j^* = V_{i,(2)} = \max_{k \neq j} V_{ik} \Leftrightarrow r_j^* = \frac{1}{L_i} \left(\theta_{ij} - V_{i,(2)}\right) \tag{5}$$

Therefore, the value of gathering multiple quotes for consumer i is equal to $V_{i,(2)}$ minus the sunk cost of shopping, denoted by κ_i. Notice that only the top two offers matter to determine the transaction price and surplus. As a result, our search assumption implies that consumers need to visit at most three banks to maximize their surplus.

In the initial stage, consumers and banks observe $V_{i,(2)}$, but consumers privately observe κ. The search cost is distributed in the population according to an exponential distribution with parameters $(\sigma_i, \bar{\kappa})$.[3] The initial offer maximizes bank h expected profit:

$$\begin{aligned} \max_{p^0} \quad & \left(p^0 - L_i c_i + \omega_{ih} + u_{ih}\right) \Pr\left(\theta_{ih} - p^0 > V_{i,(2)} - \kappa\right) \\ & + 1\left(V_{ih} > V_{i,(2)}\right) \Pr\left(\theta_{ih} - p^0 < V_{i,(2)} - \kappa\right) \underbrace{\left(p_h^* - L_i c_i + \omega_{ih} + u_{ih}\right)}_{V_{ih} - V_{i,(2)}}. \end{aligned} \tag{6}$$

The optimal initial offer takes the following form:

$$p_h^0 = \begin{cases} L_i c_i - \omega_{ih} - u_{ih} + \sigma_i & \text{If } V_{ih} \leq V_{i,(2)}, \\ \theta_{ih} - V_{i,(2)} + \sigma_i & \text{Otherwise.} \end{cases} \tag{7}$$

Plugging in these two prices into the probability of searching, we obtain the following expression:

$$H_{ih} = \begin{cases} 1 - \exp\left(-\frac{1}{\sigma_i} \max\left\{V_{i,(2)} - V_{ih} + \sigma_i - \bar{\kappa}, 0\right\}\right) & \text{If } V_{ih} \leq V_{i,(2)}, \\ 1 - \exp\left(-\frac{1}{\sigma_i} \max\left\{\sigma_i - \bar{\kappa}_i, 0\right\}\right) & \text{Otherwise.} \end{cases} \tag{8}$$

[3]More precisely $\Pr(\kappa < u) = 1 - \exp\left(-\frac{1}{\sigma_i} \max\{0, u - \bar{\kappa}\}\right)$, where σ_i is function of observed consumer covariates.

Where $\Pr\left(\theta_{ih} - p^0 < V_{i,(2)} - \kappa\right) = H_{ih}$.

3.3 Distribution assumptions

The model has three sources of randomness: (i) identity of banks with prior experience, (ii) common unobserved profit shock ϵ_i, and (iii) idiosyncratic match values u_{ij}.

The identity of home banks is partially observed when consumers transact with a bank which they have at least one month of experience. We assume that $1(E_{ij} > 0)$ is a binomial random variable with probability distribution ψ_{ij} that is function of the location of consumers (i.e. region), and income group. This probability distribution is estimated separately using data on bank affiliation. We come back to the estimation of this distribution below.

The common unobserved lending cost ϵ_i is normally distributed with mean zero and variance σ_ϵ^2.

The bank-specific idiosyncratic match values $\{u_{ij}\}_{j=1...N}$ are independently distributed according to a type-1 extreme-value (EV) distribution with location and scale parameters $(0, \sigma_u)$. As a result, conditional on $\{\theta_{ij}, \omega_{ij}, L_i, c_i\}$, the surplus V_{ij} is also distributed according to a type-1 extreme-value distribution with location $\xi_{ij} = \theta_{ij} - L_i c_i + \omega_{ij}$ and scale σ. Let $F(u; \xi_{ij}, \sigma_u)$ and $f(u; \xi_{ij}, \sigma_u)$ denote the CDF and PDF of V_{ij}.

The EV distribution leads to analytical expressions for the density functions of the first and second-order statistics $V_{i,(1)}$ and $V_{i,(2)}$. For instance, the distribution of highest surplus in consumer i's choice-set takes the form of an EV density with location parameter

$$\xi_{i,\max} = \sigma_u \log\left(\sum_{j \in \mathcal{N}_i} \exp\left(\xi_{ij}/\sigma_u\right)\right)$$

and scale σ_u. This lead to the familiar multinomial logit form for the probability that bank

j offers the highest surplus:

$$\rho_{ij} = \Pr\left(V_{ij} = \max_{k \in \mathcal{N}_i}\{V_{i,k}\}\right) = \frac{\exp(\sigma_u \xi_{ij})}{\sum_{k \in \mathcal{N}_i} \exp(\sigma_u \xi_{ik})} = \frac{\partial \xi_{i,\max}}{\partial \xi_{ij}}. \tag{9}$$

The second-order statistics of the V's distribution can also be derived analytically:

$$\Pr(V_{i,(2)} < v | V_{ij} = V_{i,(1)}) = G_{ij}(v|\mathcal{N}_i) = \frac{1}{\rho_{ij}}\Big(F(v; \xi_{i,-j}, \sigma_v) + (\rho_{ij} - 1) F(v; \xi_{i,\max}, \sigma_u)\Big) \tag{10}$$

$$\begin{aligned} \Pr(V_{i,(2)} < v) = G_i(v|\mathcal{N}_i) &= \sum_{j \in \mathcal{N}_i} F(v; \xi_{i,-j}, \sigma_u) + F(u; \xi_{i,\max}, \sigma_u) \sum_{j \in \mathcal{N}_i} (\rho_{ij} - 1) \\ &= \sum_{j \in \mathcal{N}_i} F(v; \xi_{i,-j}, \sigma_u) + (1 - N_i) F(u; \xi_{i,\max}, \sigma_u) \end{aligned} \tag{11}$$

where $N_i = |\mathcal{N}_i|$, and $\xi_{i,-j} = \sigma_u \log\left(\sum_{k \neq j} \exp(\xi_{ik}/\sigma_u)\right)$. The densities $g_{ij}(v|\mathcal{N}_i)$ and $g_i(v|\mathcal{N}_i)$ are defined analogously.

4 Estimation

4.1 Likelihood function

Consider the likelihood contribution of an individual i. We first condition on Z_i which groups all the relevant information to calculate θ_{ij}, c_i and ω_{ij}, as well as the model parameter vector β and the identity of the home bank giving the first quote. The observed outcomes are the chosen lender and transaction price: (b_i, p_i). The observed prices are either generated from consumers accepting the initial quote p_b^0, or accepting the competitive offer p_b^*. Only the later case is feasible if $b_i \neq h_i$, while both cases have positive likelihood if $b_i = h_i$.

Case 1: $b_i = h_i$

With probability ρ_{i,b_i} the initial quote is drawn from the highest-surplus bank. If that is the case, the transaction price is equal to $p_i = \theta_{ih} - V_{i,(2)} + \sigma$ with probability $1 - H_{ih}$, and to $p_i = \theta_{ih} - V_{i,(2)}$ with probability H_{ih}. The only random variable affecting outcomes, $V_{i,(2)}$, can thus be recovered directly from p_i and the parameters. The likelihood contribution is thus given by:

$$L(b_i = h_i, p_i|Z_i, \beta, V_{ih} = V_{i,(1)}) = H_{i,b_i} g_{i,b_i}(\theta_{i,b_i} - p_i|\mathcal{N}_i) + (1 - H_{i,b_i}) g_{i,b_i}(\theta_{i,b_i} + \sigma_i - p_i|\mathcal{N}_i),$$

where $H_{i,b_i} = 1 - \exp\left(-\frac{1}{\sigma_i}\max\{\bar{\kappa} - \sigma_i, 0\}\right)$.

If bank b is not the highest-surplus option, the transaction price is equal to the initial offer: $p_i = L_i c_i - \omega_{i,b_i} - u_{i,b_i} + \sigma_i$. The probability that consumer i accepts this quote is equal to $1 - H_{i,b_i}$, which is function of the unobserved second-highest surplus option $V_{(2)}$. The likelihood contribution must therefore integrate this variable out to calculate the observed choice probability:

$$\begin{aligned} L(b_i = h_i, p_i|Z_i, \beta, V_{i,b_i} < V_{i,(1)}) &= f\left(V_{i,b_i}; \xi_{i,b_i}, \sigma_u\right) \\ &\times \int \exp\left(-\frac{1}{\sigma_i}\max\left\{\max\{V_{i,b_i}, v_{(2)}\} - V_{i,b_i} + \sigma_i - \bar{\kappa}_i, 0\right\}\right) g_i(v_{(2)}|\mathcal{N}_i \backslash b_i) dv_{(2)} \\ &= f\left(V_{i,b_i}; \xi_{i,b_i}, \sigma_u\right) \times \Bigg[\exp\left(-\frac{1}{\sigma_i}\max\left\{\sigma_i - \bar{\kappa}_i, 0\right\}\right) G_i\left(V_{i,b_i}|\mathcal{N}_i \backslash b_i\right) \\ &+ \int_{V_{i,b_i}}^{\infty} \exp\left(-\frac{1}{\sigma_i}\max\left\{v_{(2)} - V_{i,b_i} + \sigma_i - \bar{\kappa}_i, 0\right\}\right) g_i(v_{(2)}|\mathcal{N}_i \backslash b_i) dv_{(2)}\Bigg], \end{aligned}$$

where $V_{i,b_i} = \theta_{i,b_i} - p_i + \sigma_i$. The previous expression integrates $v_{(2)}$ first over the region where $V_{i,b_i} > v_{(2)}$ (i.e. bank b_i is the second-highest surplus), and over the region where b_i is dominated by other options.[4]

[4]Recall that $g_i(v|\mathcal{N}_i \backslash j)$ denotes the unconditional density of the second-highest surplus option among

Joining these two events, the likelihood contribution of individual i's outcomes when $h_i = b_i$ is:

$$L(b_i = h_i, p_i|Z_i, \beta) = \rho_{i,b_i} L(b_i = h_i, p_i|Z_i, V_{ih} = V_{i,(1)}) + (1 - \rho_{i,b_i}) L(b_i = h_i, p_i|Z_i, V_{ih} < V_{i,(1)}) \tag{12}$$

Case 2: $b_i \neq h_i$

In this case, lender b_i is automatically the highest surplus option in consumer i's choice-set. The transaction price is therefore equal to $p_i = \theta_{i,b_i} - V_{i,(2)}$ and $V_{i,(2)} = \theta_{i,b_i} - p_i$. Moreover, before choosing b_i the initial quote must have been rejected with probability H_{i,h_i}, which is function of the unobserved value of the home bank, V_h. This random variable must thus be integrated-out to calculate the choice-probability. In that case the distribution of transaction price is given by the conditional distribution of second-order statistics g_{i,b_i}, conditional on option b_i offering the highest surplus. The likelihood contribution integrates out the value of the initial option taking into accounts this fact.

$$\begin{aligned} L(b_i \neq h_i, p_i|Z_i, \beta) &= \rho_{i,b_i} \int_{-\infty}^{V_{i,(2)}} 1 - \exp\left(-\frac{1}{\sigma_i} \max\left\{V_{i,(2)} - V_h + \sigma_i - \bar{\kappa}, 0\right\}\right) P\left(p_i|V_h\right) \frac{f(V_h; \xi_{i,h_i}, \sigma_u)}{F(V_{i,(2)}; \xi_{i,h_i}, \sigma_u)} dV_h \\ &= \rho_{i,b_i} g_{i,b_i}(V_{i,(2)}|\mathcal{N}_i) \int_{-\infty}^{V_{i,(2)}} 1 - \exp\left(-\frac{1}{\sigma_i} \max\left\{V_{i,(2)} - v_h + \sigma_i - \bar{\kappa}, 0\right\}\right) \frac{f(v_h; \xi_{i,h_i}, \sigma_u)}{F(V_{i,(2)}; \xi_{i,h_i}, \sigma_u)} dv_h \end{aligned}$$

where $V_{i,(2)} = \theta_{i,b_i} - p_i$ and $H_{i,h_i} = 1 - \exp\left(-\frac{1}{\sigma_i}\left(\sigma_i - \bar{\kappa}\right)\right)$.

The likelihood function is evaluated by integrated out two other unobservables: h_i and ϵ_i. The common lending profit shock ϵ_i is distributed according to a normal distribution with common variance σ_ϵ. We integrate it out using quadrature methods.

In the first case, the observed loyalty of consumers fully identifies the identity of the initial offer. For the contracts that are switching institutions, the likelihood must integrate

the $N_i - 1$ lenders excluding bank j.

the identify of the bank with prior experience. Moreover, this variable is absent for the contracts insured by Genworth. We get around this problem by separately estimating the distribution of the main financial institution from a survey of consumer finances performed by Epsos-Reid. This data-set surveys nearly $12,000$ households per year in all the regions of the country. We group the data into six years, ten regions, and four income categories. Within these subsamples we estimate the probability of choosing one of the twelve largest lender as their main financial institution. We denote this estimated probability by ψ_{ij}, where i indexes the contract identifier. This probability corresponds to the density of positive experience level $1(E_{ij} > 0)$ given the income and location of borrower i.

In addition, consumers go first shop at their home bank if it is present in their neighborhood, which is non-zero for some consumers. For instance, a bank might not be present in the new residential neighborhood of consumers, or they might be affiliated with one of the smaller institution outside of the Big-12. As a result, the identity of the first offer (i.e. h_i) is not always equal to the "home" bank, which means that we must integrate out two possibilities when evaluating the likelihood contribution of an individual: (i) receiving an initial quote from the home bank (i.e. $E_{ih} > 0$), and (ii) receiving an initial quote from a bank with no prior experience (i.e. $E_{ih} = 0$). In the latter case, we assume that consumers affiliated with a bank that is not in their choice-set are matched with banks randomly as function of the branch network size, denoted by s_i. Formally the probability of pairs (h_i, E_{ij}) is:

$$\Pr\left(h_i = j, E_{ij}\right) = \begin{cases} 1(j \in \mathcal{N}_i)\hat{\psi}_{ij} & \text{If } E_{ij} > 0 \\ \sum_{k \notin \mathcal{N}_i} \hat{\psi}_{ik} s_{ij} & \text{If } E_{ij} = 0 \end{cases} \tag{13}$$

In words, the initial comes from the home bank of consumer i if possible, and is randomly sampled from the set of available options otherwise.

The likelihood contribution of a contract i therefore can be written as:

$$L(b_i, p_i|X_i, \beta) = \int \left(\sum_{j \in \mathcal{N}_i, E \in \{0,1\}} \Pr\left(h_i = j, E_{ij}\right) L(b_i, p_i|X_i, h_i, \epsilon_i, \beta \right) f(\epsilon_i; \sigma_\epsilon) d\epsilon_i, \quad (14)$$

where X_i is a vector of exogenous covariates characterizing the payoff functions.

4.2 Results

Table 8 presents the maximum likelihood estimates for the key model parameters. The price coefficient is normalized to one and monthly payments are measured in hundreds of dollars. The scale of the parameters translates into $\$100$ of monthly expenses for the life of the contract (i.e. 5 years).

The two parameters entering the search cost distribution suggest that search frictions are economically important, and heterogeneous in the population. The baseline cost is equal to $\$23$ while the average is $\$56$ per month; roughly 5% of the average monthly payment. Under the exponential distribution assumption, half of the population of mortgage clients face a search cost lower than $\$38$ per month, implying substantial dispersion. According to the model, the marginal consumer accepting the initial quote from a winning bank is indifferent between searching and reducing his monthly payment by $\$56$, or accepting p^0.

The home bank premium λ is equal to $\$40$, while the marginal utility of network size α is equal to $\$3$. Therefore, on average consumers are willing to pay $\$40$ every month to stay with the bank with which they have prior experience. Assuming that this utility gain originates from avoid the cost of switching bank affiliation, our result suggests that switching costs are large, and of similar order of magnitude as the cost of gathering multiple quotes. In comparison, the value of network size is much smaller. The difference in the willingness to pay for a lender with a branch network that is half the size of the average network, and a lender with a branch network that is twice the size of the average network is slightly more than $\$5$ per month; significantly smaller than the cost of switching.

These results suggest that the premium observed for large network banks and the discount received by switchers are mainly caused by search and switching costs. In other words, banks with a large consumer base have important control of prices because they receive a larger fraction of "first-visits." This allows them to exploit the search cost of consumers and serve a larger proportion of non-searchers. The second source of market power originates from brand loyalty, and implies that even conditional on facing competition, the home bank is more likely to retain consumers. In comparison, the role of the complementarity service quality in generating market power is much less important.

The remaining parameters associated with the firms' profit function suggest that firms are nearly symmetric when it comes to the cost of lending. Most of the unobserved heterogeneity between consumers is common across firms, since the variance of unobserved match value (i.e. σ_u) is 8 times smaller than the common shock (i.e. σ_ϵ). The standard deviation across lenders in idiosyncratic profits is estimated at $\$7$, while the standard deviation across consumers is equal to $\$56$.

This has important implications for competition. Abstracting from loyalty issues, the average difference between the first and second highest surplus is close to zero in the average market with 8 lenders. As a result, the market for "non-loyal" consumers is very competitive: banks are nearly homogeneous and have similar cost structure, which lead to a Bertrand-type equilibrium. This is not to say that the identity of lenders is irrelevant. Bank fixed-effects are important to explain the data, and the systematic difference between the most and the least efficient lender is nearly $\$40$.

Table 9 evaluates the goodness of fit of the model by comparing the reduced-form relationship between retail margins and switching probability, and the characteristics of consumers. For this we simulated 500 realizations of the market outcomes for each consumer, and compare the average regression coefficients with the observed ones. This comparison confirms that the model replicates most of the observed correlations. The model fits well the observed relationships between consumers financial characteristics and transaction rates, as well as the relationship between the number of lenders and rates. However, the model tends to under-estimate the magnitude of the discount that switchers receive, while over-estimating the premium paid by consumers dealing with large network institutions.

5 Conclusion

References

Adams, W., L. Einav, and J. Levin (2009). Liquidity constraints and imperfect information in subprime lending. *American Economic Review 99*, 49–84.

Allen, J., R. Clark, and J. F. Houde (2011). Discounting in mortgage markets. Bank of Canada Working Paper No. 2011-3.

Allen, J. and D. McVanel (2009). Price movements in the Canadian residential mortgage market. Bank of Canada Working Paper 2009-13.

Beggs, A. and P. Klemperer (1992). Multi-period competition with switching costs. *Econometrica: Journal of the Econometric Society 60*(3), 651–666.

Berger, A. N. and A. Dick (2007). Entry into banking markets and the early-mover advantage. *Journal of Money, Credit and Banking*.

Berry, S., J. Levinsohn, and A. Pakes (2004, Feb). Differentiated products demand systems from a combination of micro and macro data: The new car market. *The Journal of Political Economy 112*(1), 68.

Kim, M., D. Kliger, and B. Vale (2003). Estimating switching costs: The case of banking. *Journal of Financial Intermediation 12*, 25–56.

Marquez, R. (2002). Competition, adverse selection, and information dispersion in the banking industry. *Review of Financial Studies 15*, 901–926.

Table 6: Margin regression results

VARIABLES	(1) Margin	(2) Margin	(3) Margin	(4) Margin	(5) Margin
Annual income (X 100K)	-0.14^a	-0.076^a	0.15^a	-0.22^a	-0.19^a
	(0.011)	(0.011)	(0.014)	(0.036)	(0.036)
Loan size (X 100K)				0.035^b	0.050^a
				(0.017)	(0.018)
Loan/Income				-0.20^a	-0.18^a
				(0.012)	(0.013)
Other debt (X 100K)				-0.086^a	-0.085^a
				(0.0076)	(0.0076)
$0.85 \leq LTV < 90$				0.065^a	0.061^a
				(0.0088)	(0.0089)
$0.90 \leq LTV < 0.95$				0.10^a	0.097^a
				(0.011)	(0.011)
$LTV = 0.95$				0.19^a	0.18^a
				(0.0092)	(0.0093)
FICO (mid-point)				-0.75^a	-0.76^a
				(0.038)	(0.038)
Renter	0.00022	0.0023	-0.00077	-0.035^a	-0.029^a
	(0.0071)	(0.0071)	(0.0070)	(0.0071)	(0.0072)
Living w/ parents	-0.058^a	-0.054^a	-0.066^a	-0.078^a	-0.069^a
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Switcher	-0.080^a	-0.077^a	-0.072^a	-0.075^a	-0.069^a
	(0.0089)	(0.0089)	(0.0088)	(0.0087)	(0.0088)
Relative network size	0.057^a	0.055^a	0.053^a	0.053^a	0.048^a
	(0.0044)	(0.0044)	(0.0043)	(0.0043)	(0.0049)
Nb. FIs=7	-0.094^a	-0.072^a	-0.050^a	-0.046^a	-0.020
	(0.0080)	(0.0080)	(0.0080)	(0.0079)	(0.014)
Nb. FIs=8	-0.16^a	-0.13^a	-0.11^a	-0.097^a	-0.068^a
	(0.0088)	(0.0091)	(0.0091)	(0.0090)	(0.020)
Nb. FIs=9	-0.22^a	-0.17^a	-0.14^a	-0.13^a	-0.073^b
	(0.014)	(0.015)	(0.015)	(0.015)	(0.029)
Nb. FIs>9	-0.27^a	-0.21^a	-0.15^a	-0.15^a	-0.079
	(0.037)	(0.037)	(0.037)	(0.037)	(0.054)
Constant	1.30^a	1.43^a	1.46^a	2.26^a	2.28^a
	(0.040)	(0.046)	(0.045)	(0.059)	(0.16)
Observations	47,039	47,039	47,039	47,039	47,039
R-squared	0.263	0.272	0.284	0.305	0.333
Census variables	N	Y	Y	Y	Y
FSA FE	N	N	N	N	Y

Robust standard errors in parenthesis. ^{a} p<0.01, p<0.05, ^{a} p<0.1 Controls variables: Bank+Prov. FE, Year+Prov. FE, Month dummies.

Table 7: Switching probability linear regression

VARIABLES	(1) Switching	(2) Switching
Loan/Income	0.051[a]	0.043[a]
	(0.0086)	(0.0087)
Renter	0.091[a]	0.087[a]
	(0.0043)	(0.0044)
Living w/ parents	0.056[a]	0.053[a]
	(0.0063)	(0.0064)
Relative network	-0.021[a]	-0.022[a]
	(0.0034)	(0.0035)
Nb. FIs in $[1, 7)$	-0.028[a]	-0.018[a]
	(0.0049)	(0.0057)
Constant	0.55[a]	0.71[a]
	(0.042)	(0.094)
Observations	35,560	35,560
R-squared	0.252	0.257
City FE	N	Y

Robust standard errors in parentheses. [a] p<0.01, [b] p<0.05, [a] p<0.1 Control variables: Bank+Prov. FE, Year+Prov. FE, Month dummies.

Table 8: Maximum likelihood estimation results

Variables	Parameters
Negotiation cost	
Intercept $\bar{\kappa}$	0.23274
	(0.008)
Mean private-value ($\sigma_\kappa - \bar{\kappa}$)	0.328
	(0.007)
Differentiation	
Quality (α)	0.030
	(0.012)
Home bank premium (λ)	0.429
	(0.007)
Idiosyncratic preference shock (σ_u)	0.101
	(0.001)
Cost function	
Residual (σ_ϵ)	0.564
	(0.003)
Bank FE (range)	0.389
Average payment: $r_{ij} \times L_i$ ($\times$ 100 $)	9.359
	(3.911)
Average transaction rate r_{ij}	6.7037
	(0.645)
LLF	-2.279
N	5,000

Asymptotic standard-errors in parenthesis. Control variables in the profit function: Loan size, income, FICO score, previous owner, and group fixed-effects (year, bank and regions). The utility and profit functions are expressed in 100 dollars units. Sample size: $5,000$.

Table 9: Comparison between observed and predicted margin and switching regressions

	Margin		Switching prob.	
Variables	Observed	Simulated	Observed	Simulated
Switcher	-0.0630	-0.0024		
RSize	0.0103	0.0344	-0.0061	-0.0618
Loan/Income	-0.1022	-0.0663	0.0213	0.0009
Previous owner	0.0205	0.0526	-0.0672	-0.0022
Loan size	1.4664	1.5601		
Income	-0.2204	-0.0261		
Bond rate	-0.2430	-0.2571		
FICO	-0.4226	-0.9139		
Maximum LTV	0.0814	0.1088		
N>8	-0.0243	-0.0268		
N≤6			-0.0201	-0.0404
Intercept	1.2841	1.1892		
Average outcome (sd)	0.7339	0.6825	0.1875	0.3808
	(0.418)	(0.657)		
Switch probability			0.500	0.516

Note: Margin is calculated as the difference between the monthly payments received by the lender (i.e. r_{ij}) and the monthly cost evaluated at the 5 year bond-rate (i.e. c_i) on a $\$1000$ loan. The predicted probability of searching is equal to 0.538. Omitted FEs: banks, year and region. Number of simulations: 500.

Table 10: Definition of Household / Mortgage Characteristics

Name	Description
FI	Type of lender
Source	Identifies how lender generated the loan (branch, online, broker, etc)
Income	Total amount of the borrower(s) salary, wages, and income from other sources
TSD	Ratio of total debt service to income
Duration	Length of the relationship between the borrower and FI
R-status	Borrowers residential status upon insurance application
FSA	Forward sortation area of the mortgaged property
Market value	Selling price or estimated market price if refinancing
Applicant type	Quartile of the borrowers risk of default
Dwelling type	10 options that define the physical structure
Close	Closing date of purchase or date of refinance
Loan amount	Dollar amount of the loan excluding the loan insurance premium
Premium	Loan insurance premium
Purpose	Purpose of the loan (purchase, port, refinance, etc.)
LTV	Loan amount divided by lending value
Price	Interest rate of the mortgage
Term	Represents the term over which the interest rate applies to the loan
Amortization	Represents the period the loan will be paid off
Interest type	Fixed or adjustable rate
$CREDIT$	Summarized application credit score (minimum borrower credit score).

Some variables were only included by one of the mortgage insurers.