NBER WORKING PAPER SERIES

SEARCH FRICTIONS AND MARKET POWER IN NEGOTIATED PRICE MARKETS

Jason Allen
Robert Clark
Jean-François Houde

Working Paper 19883
http://www.nber.org/papers/w19883

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2014

This research has benefited from the financial support of the NSF (SES-1024840). We thank the Canada Mortgage and Housing Corporation and Genworth Financial for providing us with the data. We also thank the Altus-Group. We thank the many seminar participants who have provided feedback on this paper. We have greatly benefited from discussions with Ken Hendricks, Ali Hortaçsu, Matt Lewis, Alan Sorensen, and Andrew Sweeting. The views in this paper are those of the authors and do not necessarily reflect those of the Bank of Canada. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Search Frictions and Market Power in Negotiated Price Markets
Jason Allen, Robert Clark, and Jean-François Houde
NBER Working Paper No. 19883
February 2014
JEL No. L13,L41,L81

ABSTRACT

This paper develops and estimates a search and bargaining model designed to measure the welfare loss associated with frictions in oligopoly markets with negotiated prices. We use the model to quantify the consumer surplus loss induced by the presence of search frictions in the Canadian mortgage market, and evaluate the relative importance of market power, inefficient allocation, and direct search costs in explaining the loss. Our results suggest that search frictions reduce consumer surplus by almost \$20 per month on a \$100, 000 loan, and that 17% of this reduction can be associated with discrimination, 30% with inefficient matching, and the remainder with the search cost. In addition, we find that product differentiation attenuates the effect of search frictions by reducing the cost of gathering quotes and improving efficiency, while posted prices do so through the ability of the first-mover to price discriminate. In contrast, competition amplifies the welfare effect of search frictions. Despite this, the overall effect of competition is to increase aggregate consumer surplus and drive prices down, but these effects are not spread equally across consumers: those with low search costs benefit more from competition.

Jason Allen
Bank of Canada
234 Wellington Street
Ottawa, ON K1A 0G9
jallen@bankofcanada.ca

Robert Clark
Institute of Applied Economics
HEC Montreal
3000, chemin de la Cote-Sainte-Catherine
Montreal (Quebec) H3T 2A7
Canada
Robert.Clark@hec.ca

Jean-François Houde
Wharton School
University of Pennsylvania
3620 Locust Walk
3008 Steinberg Hall - Dietrich Hall
Philadelphia, PA 19104
and NBER
houde@wharton.upenn.edu

1 Introduction

What is the welfare cost of search frictions in oligopoly markets with negotiated prices? In price haggling environments, the surplus loss associated with these frictions can originate from three sources. First, search frictions can hinder the ability of consumers to match with the most efficient firms, generating a misallocation of buyers and sellers. Second, they can generate market power by allowing first movers to price discriminate against consumers with poor outside options and high search costs. Finally, there is a direct cost imposed on consumers searching for multiple quotes. In this paper we develop and estimate a structural model of search and price negotiation to quantify the contribution of each of these components to the welfare loss associated with search frictions.

Our case study is the Canadian mortgage market. In most mortgage markets, lenders post interest rates, but contract terms for each borrower are determined through a search and negotiation process, with borrowers searching across different lender options, and then bargaining over rates. There is important heterogeneity in the ability of consumers to understand the subtleties of financial contracts, and in their willingness to search for multiple quotes. These same features are present in many markets, such as those for other financial products, insurance, new and used automobiles, and housing, and are known to generate important dispersion in the prices paid for observationally identical products.

Quantifying the importance of market frictions when prices are individually negotiated requires the imposition of some structure on how prices are determined. This is because researchers typically observe only accepted offers (not the rejected ones), and often have limited information on search behavior. To circumvent this problem, we develop and estimate a sequential search model that explicitly takes into account the heterogeneous outside options of consumers when they start their search process. In particular, individual borrowers are initially matched with their main financial institution (home bank) to obtain a mortgage quote, and can then decide, based on their expected net gain from searching, whether or not to gather additional quotes. If they reject the initial offer and choose to search, lenders compete via an English auction for the mortgage contract.

This modeling strategy is related to the search and bargaining models developed by Chatterjee and Lee (1998), Bester (1993), and Armstrong and Zhou (2011), in which consumers negotiate with one firm, but can search across stores for better prices. Our modeling assumption is also analogous to the on-the-job-search models of Postel-Vinay and Robin (2002) and Dey and Flynn (2005), in which wages are determined via bilateral bargaining, rather than wage posting.

In this framework, market power arises from three sources. First, banks with a large retail presence enjoy market power because of the timing of the search process. Like in most sequential search models, the initial lender is in a quasi-monopoly position, and can tailor individual offers to discriminate across consumers based on differences in expected search probabilities. This gives a first-mover advantage to banks that are dominant in a given local market. In addition, all lenders in our data-set offer multiple complementary financial services. To the extent that the cost of switching banks is non-negligible, this implies that consumers are willing to pay more to remain loyal to their main financial institution. The home bank of consumers therefore offers a quality differentiated service, which increases the market power of large banks. Finally, the small number of lenders available in the market, combined with the presence of idiosyncratic cost differences, creates market power even without search frictions.

To estimate the model we use detailed administrative data on a large set of approved mortgages in Canada between 1999 and 2002. Our analysis focuses on individually negotiated contracts, thereby excluding transactions generated through intermediaries (e.g. mortgage brokers), which account for about 25% of total transactions. These data provide information on features of the mortgage, household characteristics (including place of residence), and market-level characteristics. An advantage of our setting is that all of the mortgage contracts in our sample are insured by the government. This allows us to abstract from concerns related to the risk of default, and focus on homogenous contracts.

In order to quantify the magnitude of search costs, relative to other forms of brand loyalty, we supplement these mortgage-contract data with aggregate moments obtained from a survey of the shopping behavior of new mortgage buyers. We use this auxiliary source of information to calculate the search probability across different demographic groups: city size, income group, region, and new/previous home buyers.

We use the estimated model to provide three sets of empirical results. First, the model allows us to decompose observed price differences across consumers into cost-related factors, and consumer-specific markups consistent with price discrimination. For instance, in Allen et al. (2013b), we show that richer households pay higher rates (conditional on loan size), while new home buyers and consumers financing larger loans obtain larger discounts. Our results show that less than half of these reduced-form differences are explained by cost differences across borrowers; the remainder being caused by observable differences in the expected cost and benefits of search across borrowers.

Second, we quantify the importance of market power, by measuring the level and dispersion of profit margins across lenders and borrowers. The key parameters generating market power are those related to search costs and the loyalty premium, and the relative importance of cost differences across lenders. In the latter case, we estimate that most of the unobserved heterogeneity across borrowers is common to all lenders, which implies that firms face relatively homogenous lending costs. As a result, absent search frictions or product differentiation, the market would be very competitive, with profit margins averaging near $7/month for a loan size of $100,000 (i.e. about 1% markup).

In contrast, we estimate that borrowers face significant search costs and loyalty-premium. On average, search costs are \$29 per month. When discounted over the length of each contract (i.e. 5 years), these estimates correspond to an average upfront search cost of \$1,657, and a median of \$1,028. In addition, on average, consumers are willing to forego \$22 a month to stay with their home bank, and potentially avoid having to switch banks.

The presence of large search costs and product differentiation are responsible for generating nearly 60% of the average positive profit margins. The average profit margin above the marginal cost of lenders is estimated to be slightly lower than 30 basis points (bps), or about a 4.31% markup. In terms of monthly payment for a \$100,000 loans, this corresponds to about \$18/month.

We therefore conclude that the market is fairly competitive, since conditional on searching, consumers are able to extract most of the transaction surplus.[1] Indeed, the median margin is 33 bps for non-searchers, compared to 16 bps for searchers. In addition, more than 15% of transactions among consumers who search for more than one quote lead to zero profits. However, not every borrower is able to benefit from this intense competition between lenders. The dispersion of profit margins is therefore very large: the inter-decile range of profit margins is equal to 55 bps, or about twice the median margin.

Finally, we quantify the welfare cost of search frictions and market power. To do so we perform a set of counter-factual experiments in which we eliminate the search costs of consumers, and vary the amount of competition in the market. In particular we decompose the welfare cost of search frictions into three components (i.e. misallocation, price discrimination, and search costs), and analyze the role played by market structure (i.e. competition, differentiation, and price ceilings) in amplifying or attenuating these adverse effects.

Our results suggest that, overall, search frictions reduce consumer surplus by almost \$20 per month. Approximately 17% of the loss in consumer surplus comes from the ability of home banks to price discriminate with their initial quote. A further 30% loss is associated with the misallocation of contracts, and 55% is associated with the direct cost of searching for multiple quotes.

Next, we study the importance of product differentiation and price ceilings in attenuating or amplifying the adverse effect of search frictions. We find that differentiation, captured by the loyalty premium, attenuates the effect of search frictions mostly by reducing direct search costs and improving allocation: by starting their search process with the highest quality firm, consumers reduce the extent of misallocation in the market. This improvement in consumer surplus is present despite the fact that differentiation between lenders leads to higher profits from price discrimination.

In the case of the posted-rate, we find that the presence of a price ceiling reduces the welfare cost of search frictions, by limiting the ability of firms to price discriminate. However, reducing the amount of price haggling comes at the cost of limiting the overall supply of loans, which we cannot quantify in the framework of the model. We estimate that eliminating the posted-rate

[1] In our context the transaction surplus is the difference between the borrower's willingness to pay for a contract, or loyalty premium, and the marginal cost of the contract.

would increase by 5% the total number of loans issued in our sample, and reduce the search cost of consumers who are not qualifying for a loan from their initial lender.

To study the role of competition we simulate counter-factual mergers in increasingly competitive markets. In contrast to product differentiation and the posted price, competition amplifies the welfare effect of search frictions. As the number of firms in the market increases, the welfare loss from price discrimination shrinks, but the welfare loss from misallocation and direct search costs increases.

While the welfare cost of search frictions is larger the more banks are present in a market, we show that competition improves consumer welfare by reducing profit margins, and by creating more profitable search opportunities. The impact on consumer surplus of moving from a duopoly market, to a market with twelve lenders is similar in magnitude to the impact of removing entirely search frictions.

Importantly, because of the presence of heterogeneous search costs, these benefits are not spread equally across consumers. We investigate this heterogeneity by analyzing the effect of losing a potential bargaining partner on the distribution of transaction rates. We find that consumers who are more likely to search benefit more from competition. Therefore, eliminating a lender significantly raise rates paid by consumers at the bottom and middle of the rate distribution, but has little effect on consumers at the top. As a result, we show that the dispersion of transaction prices is increasing in the number of competitors in the market. This prediction is confirmed empirically by Allen et al. (2013a). In that paper, we conduct a retrospective analysis of a merger that took place in the same market, and find that the loss of a lender is associated with a 10% to 15% reduction in the residual dispersion of transaction rates.

The paper contributes to four distinct literatures. First, it develops an empirical framework for analyzing market power in differentiated product markets with negotiated prices. So far, studies that attempted to model pricing behavior in these markets have either ignored the dispersion in transaction prices across consumers, or abstracted away from the price-setting mechanism actually used in the market. For instance Berry et al. (2004) in their study of the demand for new automobiles use the median transaction price, while others, like Adams et al. (2009), ignore the possible competition between firms by assuming a monopoly pricing model.

Second, it contributes to a large body of literature studying the dispersion of transaction prices in negotiated price markets. Most of this research has been focused on testing for the presence of price discrimination (e.g. Goldberg (1996), Scott-Morton et al. (2001, 2003), Dafny (2010), Allen et al. (2013b)), measuring the pass-through of promotions (e.g. Busse et al. (2006), and Sallee (2011)), and studying the impact of asymmetric information (e.g. Busse et al. (2013)). Similarly, researchers have used data on housing transactions to test the predictions of alternative search and bargaining models, see of instance Hendel et al. (2010) and Merlo and Ortalo-Magne (2004). In contrast, we are interested in quantifying the consequences of market power and search frictions, and we therefore follow a more structural econometrics approach.

Third, the paper also contributes to a growing literature measuring search costs in retail markets. In particular, recent papers have estimated equilibrium models of price of competition, in which firms randomize their prices in order exploit the fact that consumers incur search costs (see for instance Hortaçsu and Syverson (2004), Hong and Shum (2006), Wildenbeest (2011)). There is also a large literature in economics and marketing, devoted to measuring the size of consumer search costs, using exogenous price distributions. This includes for instance Sorensen (2001), De Los Santos et al. (2011), and Honka (2012). This modeling strategy is appropriate for posted-price markets, in which firms offer random price menus to consumers, irrespectively of their characteristics.

The closest paper to our's in the search literature is Gavazza (2013), which estimates an equilibrium search and bargaining model applied to the secondary business-aircraft market. In this market, a large number of buyers and sellers engage in costly search for a trading partner, which leads to a misallocation of resources, and wasteful search efforts. The main difference with our paper is that we study a concentrated market in which sellers (lenders) enjoy market power. As a result, we evaluate the welfare cost of search frictions relative to an environment in which consumers are able to freely obtain a larger number of quotes, while still facing an imperfectly competitive market. In contrast, Gavazza compares the allocation of products relative to a perfectly competitive Walrasian market.

Finally, there is also a growing empirical literature evaluating market power in markets with bargaining. This literature has mostly concentrated on health care markets (e.g., Town and Vistnes (2001), Capps et al. (2003), Dranove et al. (2008), Grennan (2013), and Gowrisankaran et al. (2013)), although more recently Crawford and Yurukoglu (2011) study the cable market. A limitation of this literature is that it focuses on bilateral bargaining models with perfect information. Specifically, a buyer's outside option is not determined as an equilibrium object dependent on offers they could expect to get from other sellers. Consequently, negotiations never fail and matches are efficient.

The paper is organized as follows. Section 2 presents details on the Canadian mortgage market, including market structure, contract types, and pricing strategies, and introduces our data sets. Section 3 presents the model. Section 4 discusses the estimation strategy and Section 5 describes the empirical results of the model. Section 6 presents the counterfactuals. Finally, Section 7 concludes.

2 Data

2.1 Mortgage contracts and sample selection

There are two types of mortgage contracts in Canada – conventional mortgages, which are uninsured since they have a low loan-to-value ratio, and high loan-to-value mortgages, which require insurance (for the lifetime of the mortgage). Today, 80% of new home-buyers require mortgage insurance. The primary insurer is the Canada Mortgage and Housing Corporation (CMHC), a crown corporation with an explicit guarantee from the federal government. During our sample period a private firm, Genworth Financial, also provided mortgage insurance, and had a government guarantee, although for only 90%. CMHC's market share during our sample period averages around 80%.

All insurers use the same guidelines for insuring mortgages. First, borrowers with less than 25% equity must purchase insurance.[2] Second, borrowers with monthly gross debt payments that are more than 32% of gross income or a total debt service ratio of more than 40% will almost certainly be rejected.[3] The mortgage insurers charge the lenders an insurance premium, ranging from 1.75 to 3.75% of the value of the loan – lenders pass this premium onto borrowers. Insurance qualifications (and premiums) are common across lenders and based on the posted rate. Borrowers qualifying at one bank, therefore, know that they can qualify at other institutions, given that the lender is protected in case of default.

Our main data-set is a sample of insured contracts from the Canada Mortgage and Housing Corporation (CMHC), from January 1999 and October 2002.[4] We obtained a 10% random sample of all contracts from CMHC. The data-sets contain information on 20 household/mortgage characteristics, including 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 13 in the Appendix lists all of the variables included in the data-set. In addition, we observe the location of the purchased house up to the

[2]This is, in fact, not a guideline, but a legal requirement for regulated lenders. After our sample period, the requirement was adjusted and today borrowers with less than 20% equity must purchase insurance.

[3]Gross debt service (GDS) is defined as principal and interest payments on the home, property taxes, heating costs, annual site lease in case of leasehold, and 50% of condominium fees. Total debt service (TDS) is defined as all payments for housing and other debt. Both measures are as a percentage of gross income. These guidelines have been updated post our sample period to also be based on credit scores; borrowers with lower credit scores now face higher GDS requirements. Crucial to the guidelines is that the TDS and GDS calculations are based on the posted rate and not the discounted price. Otherwise, given mortgages are insured, lenders might provide larger discounts to borrowers above a TDS of 40 in order to lower their TDS below the cut-off. The guidelines are based on the posted rate to discourage this behavior.

[4]Although we have data from 1992 to 2004, there are a number of reasons to restrict the sample to 1999-2001. See Allen et al. (2013b) for a discussion of the complete data-set. First, between 1992 and 1999, the market transited from markets with a larger fraction of posted-price transactions and loans originated by trust companies, to a decentralized market dominated by large multi-product lenders. Our model is a better description of the latter period. Second, between November 2002 and September 2003, TD-Canada Trust experienced with a new pricing scheme based on a "no-haggle" principle. Understanding the consequences of this experiment is beyond the scope of this paper.

Table 1: Summary statistics on mortgage contracts in the selected sample

VARIABLES	N	Mean	SD	P25	P50	P75
Interest rate spread (bps)	29,000	129	61.4	86.5	123	171
Residual spread (bps)	29,000	0	49.7	-32.1	-2.96	34.7
Positive discounts (bps)	22,240	77.7	40	50	75	95
1(Discount=0)	29,000	23.3	42.3			
Monthly payment ($)	29,000	966	393	654	906	1219
Total loan ($/100K)	29,000	138	57.2	92.2	129	176
Income ($/100K)	29,000	69.1	27.9	49.2	64.8	82.8
FICO score	29,000	669	73.6	650	700	750
Switcher	22,875	26.7	44.2			
1(Max. LTV)	29,000	38.2	48.6			
1(Previous owner)	29,000	24.3	42.9			
Number of FIs (5 KM)	29,000	7.82	1.73	7	8	9
HHI (5 KM)	29,000	1800	509	1493	1679	1918
Relative branch network	29,000	1.46	.945	.84	1.22	1.83

forward sortation area (FSA).[5]

We also have access to data from Genworth Financial, but use these only for robustness, since we are missing some key information for these contracts. We obtained the full set of contracts originated by the 12 largest lenders and further sampled from these contracts to match Genworth's annual market share.

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 year 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), (iv) contracts without missing values for key attributes (e.g. credit score, broker, and residential status).

The final sample includes 29,000 observations, or about 33% of the initial sample. 18% of the initial sample contained missing characteristics; either risk type or business originator (i.e. branch or broker). This is because CMHC started collecting these transaction characteristics systematically only in the second half of 1999. We also dropped broker transactions, (28%), as well as short-term, variable rate and mortgage renewal contracts (40%). Finally, we drop 10% of transactions for which the lender is located more than 5 KM away from the centroid of FSA the new house (see discussion below).

Table 1 describes the main financial and demographic characteristics of the borrowers in our sample, where we trim the top and bottom 0.5% of observations in terms of income, and loan-size. The resulting sample corresponds to a fairly symmetric distribution of income and loan size. The

[5]The FSA is the first half of a postal code. We observe nearly 1,300 FSA in the sample. While the average forward sortation area (FSA) has a radius of 7.6 kilometers, the median is much lower at 2.6 kilometers.

average loan size is about $\$138,000$ which is twice the average annual household income. The average monthly payment is $966, and the average interest rate spread is 129 basis points.

Importantly, only about 27% of households switch banks when negotiating a new mortgage loan. This large loyalty rate suggests that most consumers combine multiple financial services with the same bank. This is consistent with the fact the large Canadian banks are increasingly offering bundles of services to their clients, helped in part by the deregulation of the industry in the early 1990s. For instance, a representative survey of Canadian finances from Ipsos-Reid shows that 67% of Canadian households have their mortgage at the same financial institution as their main checking account.[6] In addition, 55% of household loans, 78% of credit cards, 73% of term deposits, 45% of bonds/guaranteed investments and 39% of mutual funds are held at the same financial institution as the households main checking account.

The loan-to-value (LTV) variable shows that many consumers are constrained by the minimum down-payment of 5% imposed by the government guidelines. Nearly 40% of households invest the minimum, and the average loan-to-value ratio is 91%. LTV ratios are highly localized around 90 and 95. Moreover, the vast majority of households in our data (i.e. 96%) roll-over the insurance premium into the initial mortgage loan. The loan size measure that we use includes the insurance premium for those households.

2.2 Pricing and negotiation

The Canadian mortgage market is currently dominated by six national 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 in Québec), and a provincially owned deposit-taking institution (Alberta's ATB Financial). Collectively, they control 90% of assets in the banking industry. For convenience we label these institutions the "Big 8."

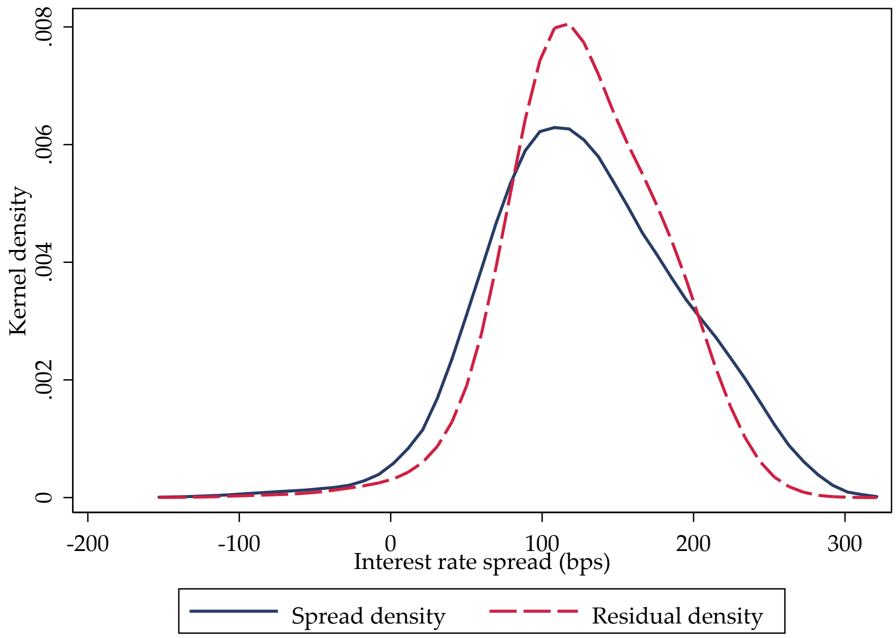
The large Canadian banks operate nationally and post prices that are common across the country on a weekly basis in both national and local newspapers, as well as online. There is little dispersion in posted prices, especially at the big banks where the coefficient of variation on posted rates is close to zero. In contrast, there is a significant amount of dispersion in transaction rates. Approximately 25% of borrowers pay the posted rate.[7] The remainder receive a discount.

Figure 1 illustrates this dispersion by plotting the distribution of retail interest rates in the sample. We measure spreads using the 5-year bond-rate as a proxy for marginal cost. The transaction rate is on average 1.3 percentage points above the 5-year bond rate, and exhibits substan-

[6] This figure is slightly lower than the 73% reported in Table 1 because we excluded broker-negotiated transactions. Consumers dealing with brokers are significantly more likely to switch bank (75%).

[7] The 25% is based on the posted price being defined as the posted rate within 90 days from the closing date minus the negotiated rate. The majority of lenders offer 90-day rate guarantees, which is why we use this definition. Some lenders have occasionally offered 120-day rate guarantees.

Figure 1: Dispersion of interest rate spreads between 1999-2001



tial dispersion. Importantly, a large share of the dispersion is left unexplained when we control for a rich set of covariates: financial characteristics, week fixed effects, lender/province fixed-effects, lender/year fixed-effects, and location fixed-effects. These covariates explain 44% of the total variance of observed spreads. The figure also plots the residual dispersion in spreads. The standard-deviation of retail spreads is equal to 61 basis points, while the residual spread has a standard-deviation of 50 basis points.

This dispersion comes about because potential borrowers can search for and negotiate over rates. Borrowers bargain directly with local branch managers or hire a broker to search on their behalf.[8] Our model focuses only on branch-level transactions, and therefore we exclude broker transactions.

Our data do not provide direct information on the number of quotes gathered by borrowers. However, survey evidence from the Altus Group (FIRM), reveals a number of facts regarding the way Canadians shop for mortgages. First, 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. Conditioning on home-buyers with mortgage insurance between 1999 and 2002, we present the search probability (P(s)), and the number of respondents (N), broken down by region, city population, income, home-buyer type (new or renewer), and loyalty. From

[8] Local branch managers compete against rival banks, but not against other branches of the same bank. Brokers are "hired" by borrowers to gather the best quotes from multiple lenders but compensated by lenders.

Table 2: Summary statistics on shopping habits

Category	Fraction (%)	N	Category	Fraction (%)	N
Pop. size				Home buyers	
Pop $\leq$ 100K	44.3	79	NHB	67.3	153
100K<Pop<1M	66.7	114	Renewer	50.9	106
Pop>1M	64	75			
Income				Loyal/switch	
Inc<60K	56	141	Switch	76	25
Inc.$\geq$60K	61.9	126	Loyal	46	70
Regions					
East	51.5	103			
Ontario	71.6	102			
West	53.4	73	**Total**	59	278

Source: Altus-Group FIRM Survey: 1999-2002. NHB is new home-buyer.

this we see that the search probability is higher in more populated areas. Also, new home-buyers are more likely to search than those renewing a mortgage. The search probability for high and low income individuals is about the same, and individuals who switch are much more likely to have searched. We use these statistics on the fraction of consumers gathering more than one quote by demographic group in the estimation of the search and negotiation model (see section 4.2).

Not only do almost half of all consumers not search, but some of those who search end up nonetheless transacting with their home bank. The variable labeled "switchers" is a dummy variable equal to one if the duration of the prior relationship with the mortgage lender is zero. 79% of households choose a lender with which they already have a prior financial relationship.[9] The fraction of switchers is significantly larger for new home-buyers, and for contracts negotiated through a broker.

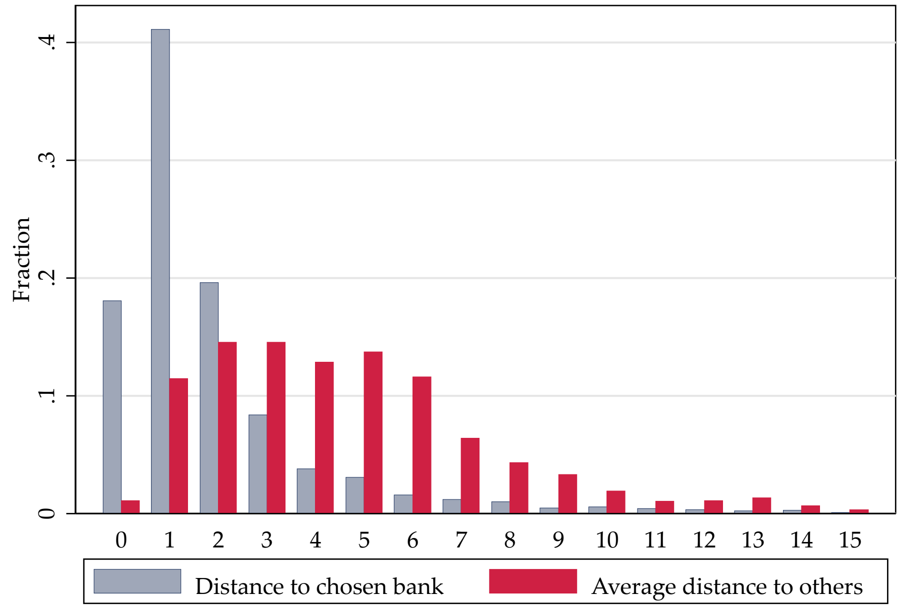
2.3 Local markets and lender information

Our main data-set contains the lender information for 10 lenders during our sample period (the big 8 plus Canada Trust and Vancity). For mortgage contracts where we do not have a lender name but only a lender type, these are coded as "Other Bank", "Other credit union", and "Other trusts". The credit-union and trust categories are very fragmented, and contain mostly regional financial institutions. We therefore combine both into a single "Other Lender" category.

The "Other Bank" category includes mostly two institutions: Laurentian Bank and HSBC. The former is only present in Québec and Eastern Ontario, while the latter is present mostly in British Colombia and Ontario. We exploit this geographic segmentation and assign the "Other banks"

[9]Note that due to data limitations we do not measure the switcher variable for contracts issued by Genworth, and for one financial institution. The fraction of switchers is measures using only the remaining contracts.

Figure 2: Distribution of minimum distances between banks and consumers



Note: Distances are calculated over bank located within at least 50 Km.

customers to HSBC or Laurentian based on their relative presence in the local market around each home location. After performing this imputation, consumers face at most 13 lending options: the Big 8, Canada Trust, Laurentian Bank, Vancity, HSBC, and Other Lender.

Not all consumers have access to every option, because of the uneven distribution of branches across local markets. We exploit this variation by assuming that consumers shop for their mortgage locally, in a neighborhood around the location of their new house (e.g. municipality). We define this as a consumer's choice set, which is their home bank h plus all other banks in their neighborhood, $\mathcal{N}_i$. To implement this, we match the new house location with the postal code associated with each financial institution's branches (available annually from Micromedia-ProQuest). 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 5KM around it to define the borrower's maximum choice-set. Formally, a lender is part of consumer i's maximum choice-set if it has a branch located within less than 5KM of the house location. We use this definition to measure the relative presence of each lender (i.e. number of branches in a choice-set), and the number of lenders within each choice-set (i.e. number of lenders with at least one branch).

Figure 2 illustrates 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 is nearly 1.5KM for the chosen institution, and 2.4KM for the other lenders. In fact the distributions indicate that 80% of consumers transact with a bank that has a branch within 2KM of their new house, while only 40% of consumers have an average distance to competing lenders lower or equal to 2KM.

This feature reflects the fact that consumers tend to choose lenders with large networks of branches. Table 1 reports 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%. Table 1 also presents measures on the level of concentration in a consumers choice-set. On average each consumer faces 7.8 lenders within 5 KM. Most of these banks have a relatively small presence, indicated by the large Herfindahl-Hirschman index, calculated using the distribution of branches within 5KM of each contract.

3 Model

We propose a sequential model in which consumers with heterogeneous search costs initially obtain a quote from their home bank and then decide whether or not to keep searching by gathering multiple quotes from the remaining lenders in their choice-sets. We describe the model in detail in the next three subsections. First, we present the notation, and formally define the timing of the model. Then, we solve the model backwards, starting with the second stage of the model in which banks are allowed to compete for consumers. Finally, we describe the search decision of consumers, and the process generating the initial quote.

3.1 Timing and payoffs

The timing of the model is as follows. In an initial period outside the model, consumers choose the type of house they want to buy, the loan size L_i, and the timing of home purchase. Buyers also observe the posted price of lenders, which for simplicity we assume is the same across banks. Empirically this is nearly always true throughout our entire sample period. In addition, for a particular bank, the posted price is common for all regions in the country.

Taking these characteristics as given, consumers then visit their home bank h, and receive an initial quote, p_i^0, measured in dollars per month. At this point information about the home bank's cost is publicly revealed, and consumers privately observe their cost of gathering additional quotes (denoted κ_i).[10]

If the initial offer is rejected, consumers organize a multilateral negotiation game between banks in their choice-set, denoted by $\mathcal{N}_i$. The full choice-set is given by $\mathcal{N}_i + h$ since we allow the home bank to participate in the second-stage negotiation.

[10]Note that information about lenders other than the home bank are not revealed at this stage. This would be a "full information" version of the model, and is available upon request.

We model the multilateral negotiation process as an English auction game among lenders in $\mathcal{N}_i + h$, with a bid-preference advantage for the home-bank. The simultaneous assumption in the second stage allows us to abstract from considerations related to the order of arrival of competing offers. We believe it is a more accurate description of the market than a model with sequential offers without recall. In practice, banks are able to lower their initial offer if consumers receive a lower price quote from a competing bank.

We assume the following payoff structure for consumers and firms, respectively:

$$\text{Consumers:} \qquad U_{ij} = \lambda_i 1(j = h) - p_{ij}, \tag{1}$$
$$\text{Firms:} \qquad \pi_{ij} = p_{ij} - (c_{ij} + \varepsilon_i), \tag{2}$$

where p_{ij} is the monthly payment offered by bank j.

The parameter λ_i measures consumer i's willingness to pay for their home-bank. Throughout we refer to λ using the terminology loyalty premium and switching cost interchangeably. Consumers are assumed to be associated with at most one lender, and therefore $1(j = h)$ is a dummy variable equal to one if consumer i has prior experience dealing with bank j, and zero otherwise.

The cost term measures the direct lending costs for the bank, net of the future benefits associated with selling complementary services to consumer i. Both components are related to variables affecting the risk of default, and the risk of pre-payment over the length of the contract. While lenders are fully insured against default risk, the event of default implies additional transaction costs to lenders that lowers the value of lending to risky borrowers. Pre-payment risk is perhaps more relevant in our context, since consumers are allowed to reimburse up to 20% of their mortgage every year without penalty.[11]

Since we do not observe the performance of the contract along these two dimensions, $c_{ij} + \varepsilon_i$ approximates the net present value of the contract. This cost function includes a common unobserved attribute ε_i that symmetrically affects all lenders. The realization of the idiosyncratic comment c_{ij} is privately observed by lenders in the game's second stage.

In both pricing stages, banks are constrained by their posted rate, essentially a price ceiling on the negotiation.[12] We assume that each consumer faces a posted price given by the monthly payment associated with the posted rate valid at the time of negotiation, denoted by $\bar{p}_i$. The presence of the posted rate forces some lenders to reject loans that would lead to negative profits. The loan qualifying condition for all banks is given by: $c_{ij} + \varepsilon_i < \bar{p}_i$.

The value of shopping, net of the search cost κ_i, is function of the equilibrium price vector

[11] In practice, however, borrowers pre-pay, on average, an additional 1% of their mortgage every year.

[12] In Canada overage, i.e. pricing over the posted rate, is illegal.

offered by banks. Consumers choose the lender that generates the highest indirect utility:

$$W_i = \begin{cases} \lambda_i - p_h & \text{If } \lambda_i - p_h > -\min_{j\in\mathcal{N}} p_j, \\ -\min_{j\in\mathcal{N}_i} p_j & \text{Else.} \end{cases} \tag{3}$$

The value of shopping is a random variable determined by the realization of firms' costs, and the mode of competition. The search cost κ_i therefore measures both the time cost of generating competition between firms, and the cost of obtaining additional information about lenders.

The gross transaction surplus from an (ij) match is equal to:[13]

$$V_{ij} = \lambda_i 1(j = h) - (c_{ij} + \varepsilon_i). \tag{4}$$

Finally, we use $c_{(k)}$ to denote the k^{th} lowest cost option in $\mathcal{N}_i$. The distribution of costs is given by N CDF's, $G_j(x) = \Pr(c_{ij} < x)$, and we use $G_{(k)}(x) = \Pr(c_{(k)} < x)$ to denote the CDF of the k^{th} order statistic of the cost distribution.

3.2 Competition stage

Conditional on rejecting p^0, the home bank h competes with lenders in the choice-set $\mathcal{N}_i$. We model this competition as an English auction with heterogeneous firms, and bid-preference favoring the home bank. Since the initial quote can be recalled, firms face a reservation price equal to: $p^0 \leq \bar{p}$.

We can distinguish between two cases leading to a transaction: (i) $\bar{p}_i - c_h < \varepsilon_i$, and (ii) $\varepsilon_i < p_i^0 - c_h \leq \bar{p}_i - c_h$. In the first case the borrower does not qualify at the home bank. As such, the lowest cost bank will win by offering a price equal to the lending cost of the second most efficient qualifying lender:

$$p_i^* = \min\{c_{(2)} + \varepsilon_i, \bar{p}_i\}. \tag{5}$$

This occurs if and only if, $\varepsilon_i < \bar{p}_i - c_{(1)}$.

If the borrower qualifies at the home bank, the highest surplus bank will win, and offer a quote that provides the same utility as the second best option. The equilibrium pricing function is:

$$p_i^* = \begin{cases} p_i^0 & \text{If } c_{(1)} > p_i^0 - \varepsilon_i - \lambda_i \\ \lambda_i + c_{(1)} + \varepsilon_i & \text{If } c_h - \lambda_i < c_{(1)} < p_i^0 - \varepsilon_i \\ \min\{c_h - \lambda_i, c_{(2)}\} + \varepsilon_i & \text{Otherwise.} \end{cases} \tag{6}$$

This equation highlights the fact that at the competition stage loyal consumers will on average pay

[13]It should be noted that most of the model's predictions are the same whether or not we assume that the match value enters firms' profits, or consumers' willingness to pay. While we believe that it is more reasonable to think that most of the randomness across consumers arises from differences in lending opportunity costs across banks, as we will see below the choice of lender and the transaction price depend only on the distribution of total surplus.

a premium, while lenders directly competing with the home-bank will on average have to offer a discount by a margin equal to the switching cost in order to attract new customers.

Finally, we assume that consumers and lenders have rational expectations over the outcome of the competition stage, which leads to the following expression for the expected value of shopping:

$$
\begin{aligned}
E[W_i|p_i^0,\varepsilon_i] &= (\lambda_i - p_i^0)(1 - G_{(1)}(p_i^0 - \varepsilon_i - \lambda_i)) + \int_{c_h-\lambda_i}^{p_i^0-\varepsilon_i} -(c_{(1)}+\varepsilon_i)g_{(1)}(c_{(1)})dc_{(1)} \\
&+(\lambda_i - c_h - \varepsilon_i)\left[G_{(1)}(c_h-\lambda_i) - G_{(2)}(c_h-\lambda_i)\right] + \int_{-\infty}^{c_h-\lambda_i} -(c_{(2)}+\varepsilon_i)g_{(2)}(c_{(2)})dc_{(2)}.
\end{aligned} \tag{7}
$$

3.3 Search decision and initial quote

Consumers choose to search for additional quotes by weighing the value of accepting p_i^0, or paying a sunk cost κ_i in order to lower their monthly payment. The search decision of consumers is defined by a threshold function, which yields a search probability that is increasing in the outside option of consumers and decreasing in the loyalty premium:

$$
\Pr\left(\text{Reject}|p_i^0,\varepsilon_i\right) = \Pr\left(\lambda_i - p_i^0 < E\left[W_i|p_i^0,\varepsilon_i\right] - \kappa_i\right) = H_i(p_i^0,\varepsilon_i). \tag{8}
$$

Note that we index the search probability by i to highlight the fact that consumers face different expected value of shopping, and different search cost distributions (e.g. increasing in income).

Lenders do not commit to a fixed interest rate, and are open to haggling with consumers based on their outside options. This practice allows the home bank to price discriminate by offering up to two quotes to the same consumer: (i) an initial quote p^0, and (ii) a competitive quote p^* if the first one is rejected.

The price discrimination problem is based on the expected value of shopping and the distribution of search costs. More specifically, anticipating the second-stage outcome, the home bank chooses p^0 to maximize its expected profit:

$$
\max_{p^0 \leq \bar{p}} \quad (p^0 - c_h - \varepsilon_i)[1 - H_i(p^0,\varepsilon_i)] + H_i(p^0,\varepsilon_i)E(\pi^*|p^0,\varepsilon_i),
$$

where, $E(\pi^*|p^0,\varepsilon_i) = [1 - G_{(1)}(c_h - \lambda_i)]E(p^* - c_h - \varepsilon_i | c_{(1)} > c_h - \lambda_i)$.

Importantly, the home bank will offer a quote only if it makes positive profit: $\epsilon < \overline{p} - c_h$. The optimal initial quote first order condition is:

$$
p^0 - c_h - \varepsilon_i = \underbrace{\frac{1 - H_i(p^0,\varepsilon_i)}{h_i(p^0,\varepsilon_i)}}_{\substack{\text{Search cost}\\\text{distribution}}} + \underbrace{E(\pi^*|p^0,\varepsilon_i)}_{\substack{\text{Cost+Quality}\\\text{Differentiation}}} + \underbrace{\frac{H_i(p^0,\varepsilon_i)}{h_i(p^0,\varepsilon_i)}\frac{\partial E(\pi^*|p^0,\varepsilon_i)}{\partial p^0}}_{\text{Reserve price effect}} \tag{9}
$$

where $h_i(p^0, \varepsilon_i) = \partial H_i(p^0, \varepsilon_i)/\partial p^0$ is the marginal effect of p^0 on the search probability, which is analogous to the slope of the demand curve for the initial lender.

The previous expression implicitly defines firms' profit margins from price discrimination. It highlights three sources of profits for the home bank: (i) positive average search costs, (ii) market power from differentiation in cost and quality (i.e. match value differences and loyalty premium), and (iii) the reserve price effect. If firms are homogenous, the only source of profits will stem from the ability of the home bank to offer higher quotes to high search cost consumers.

Although the initial quote does not have a closed-form solution, it is additive in the common cost shock in the interior: $p_i^0 = \bar{p}_i^0 + \varepsilon_i$. To see this, note that if p_i^0 is additive, since the equilibrium second-stage price is additive in ε_i, the expected value of shopping is also additive in ε_i. As a result, in the interior, the search probability is independent of ε_i, since only the difference between p^0 and $E(W_i|p_i^0)$ matters for determining the threshold of consumers. We use $\bar{H}_i$ to denote the equilibrium search probability when $p_i^0 < \bar{p}_i$. Similarly, the expected second-stage profit is independent of ε_i, since p^* is additive in ϵ_i. Therefore, the right-hand-side of equation 9 is independent of ε_i, which confirms that $p_i^0 = \bar{p}_i^0 + \varepsilon_i$ is a solution to the first-order condition of the initial lender.

4 Estimation method

In this section we describe the steps taken to estimate the model parameters. We begin by describing the functional form assumptions imposed on consumers and lenders' unobserved attributes. Then we derive the likelihood function induced by the model, and discuss the sources of identification in the final subsection.

4.1 Distributional assumptions

Our baseline model has three sources of randomness beyond observed financial and demographic characteristics: (i) the identity of banks with prior experience and origin of the first quote, (ii) the common unobserved profit shock ϵ_i, and (iii) idiosyncratic cost differences between lenders. We describe each in turn.

Distribution of main financial institutions The identity of home banks is partially observed when consumers transact with a bank with which they have at least one month of experience, and consumers are assumed to have experience with at most one bank. For the consumers who switch institutions, the identity of the bank with prior experience is unknown (i.e. we only know that it is not chosen). Moreover, this variable is absent for the 20% of contracts insured by Genworth, and is missing entirely for one bank.

We assume that $1(j = h)$ is a multinomial random variable with probability distribution $\psi_{ij}(X_i)$. This distribution is a function of consumers' locations and income group. We estimate this probability distribution separately using a survey of consumer finances conducted by Ipsos-Reid, which identifies the main financial institution of consumers.[14] This data-set surveys nearly $12,000$ households per year in all regions of the country. We group the data into six years, ten regions, and four income categories. Within these sub-samples we estimate the probability of a consumer choosing one of the twelve largest lenders as their main financial institution. This probability corresponds to the density of positive experience level given the year, income, and location of borrower i.

We use the distribution of main financial institutions to integrate over the identity of the home-bank for switching consumers or for consumers with missing data. Formally, we let Status$_i \in \{\text{Loyal}, \text{Switching}, \text{M/V}\}$ denote the switching status of consumer i. Then the conditional probability that bank h is the first mover is:

$$\Pr\left(h|b_i, \text{Status}_i, X_i, \mathcal{N}_i\right) = \begin{cases} 1(h = b_i) & \text{If Status}_i = \text{Loyal}, \\ 1(h \neq b_i)\psi_h(X_i) / \sum_{j \neq b_i} \psi_j(X_i) & \text{If Status}_i = \text{Switching}, \\ \phi_h(X_i) & \text{If Status}_i = \text{M/V}. \end{cases} \tag{10}$$

An additional problem is that the experience duration variable might be measured with error. For instance, some loyal consumers who obtained a pre-qualifying offer might be considered loyal because they received an offer more than a month before closing. We take this feature into account by incorporating a binomial IID measurement error. With probability ρ the identity of the home bank is drawn from the conditional probability described in equation 10, and with probability $1 - \rho$ the identity of the home bank is drawn from the unconditional distribution $\phi_h(X_i)$. Let $\Pr(h|b_i, \rho, \text{Status}_i, X_i)$ denote the measurement-error adjusted probability distribution function.

Cost function We parametrize the cost of lending to consumer i using the following reduced-form function:

$$c_{ij} = \begin{cases} L_i \times (Z_i\beta + \varepsilon_i - u_{ij}) & \text{If } j \neq h \\ L_i \times (Z_i\beta + \varepsilon_i) & \text{If } j = h. \end{cases} \tag{11}$$

Lending costs are measured on a monthly basis, using a 25 years amortization period. The function in parenthesis parametrizes the monthly cost of a \$100,000 loan. The vector Z_i controls for observed financial characteristics of the borrower (e.g. income, loan size, FICO score, LTV, etc), the bond-rate, as well as period, location and bank fixed-effects. The location fixed-effects identify the region of the country where the house is located, defined using the first digit of the postal code (i.e. postal district). Because of the small number of observations in the Maritimes and Prairies, we group those provinces in two regions. This leaves us with 10 unique markets.

The common shock ϵ_i is normally distributed with mean zero and variance σ_ϵ^2, and the vector

[14] Source: Consumer finance monitor (CFM), Ipsos-Reid, 1999-2002.

of bank-specific idiosyncratic cost shocks $\{u_{ij}\}$ are independently distributed according to a type-1 extreme-value (EV) distribution with location and scale parameters $(-\sigma_u\gamma, \sigma_u)$.[15] We interpret u_{ij} as a mean-zero deviation from the lending cost of the home-bank.

As a result, conditional on ϵ_i, the lending cost is also distributed according to a type-1 extreme-value distribution. The EV distribution assumption leads to analytical expressions for the distribution functions of the first and second-order statistics, and has often been used to model asymmetric value distributions in auction settings (see for instance Brannan and Froeb (2000)).

We use $g_{(k)}(x)$ to denote the density of the k^{th} order statistic of the lending cost distribution, and $f(x)$ to denote the density of the common component ε_i.

Other functional forms Our main empirical specification allows for heterogeneous expected search-cost and loyalty premium. In particular, we allow $\bar{\kappa}$ and λ to vary across new and experienced home buyers, and income categories:

$$\log(\bar{\kappa}_i) = \bar{\kappa}_0 + \bar{\kappa}_{\text{inc}}\text{Income}_i + \bar{\kappa}_{\text{owner}}1(\text{Previous owner}_i),$$
$$\log(\lambda_i) = \lambda_0 + \lambda_{\text{inc}}\text{Income}_i + \lambda_{\text{owner}}1(\text{Previous owner}_i),$$

where $1(\text{Previous owner})$ is an indicator variable equal to 1 if the borrower previously owned a home and equal to 0 if they previously rented or lived with their parents.

4.2 Likelihood function

We estimate the model by maximum likelihood. The endogenous outcomes of the model are: the chosen lender and transaction price (B_i, P_i), as well as the selling mechanism $M_i = \{\mathbf{A}, \mathbf{N}\}$ (i.e. **A**uction versus **N**egotiation). The observed prices are either generated from consumers accepting the initial quote (i.e. $M_i = \mathbf{N}$), or accepting the competitive offer (i.e. $M_i = \mathbf{A}$). Importantly, only the latter case is feasible if $B_i \neq h$, while both cases have positive likelihood if $B_i = h$. We first derive the likelihood contribution for the loyal case followed by the case of switchers.

In order to derive the likelihood contribution of each individual, we first condition on the choice-set $\mathcal{N}_i$, the observed characteristics Z_i, the identity of home-bank h, and the model parameter vector θ. After describing the likelihood contribution conditional on $\mathcal{I}_i = (\mathcal{N}_i, Z_i, h)$, we discuss the integration of h.

Moreover, since we only observed accepted offers, we must adjust the likelihood to control for endogenous selection. In particular, because of the posted-rate, some consumers fail to qualify for a loan at every bank in their choice-set. To control for this possibility, we maximize a conditional likelihood function, adjusted by the probability of qualifying for a loan given observed characteristics Z_i and choice-set $\mathcal{N}_i$.

[15] The location parameter of u_{ij} is normalized to $-\sigma_u\gamma$ so that u_{ij} is mean-zero.

Finally, in the last subsection we describe how we incorporate aggregate moments on the probability of search.

We use the following notation. We use cap-letters to refer to random outcome variables, and small-case letters to refer to the realizations of consumer i. We remove the conditioning $(\mathcal{I}_i, \theta)$ whenever necessary, since it is common to all probabilities. In order to simplify the notation, we also use individual subscripts i only for the outcomes variables and random shocks, with the understanding that all functions and variables are consumer-specific and depend on $\mathcal{I}_i$.

Likelihood contribution for loyal consumers The main obstacle in evaluating the likelihood function is that we do not observe the selling mechanism, S_i. The unconditional likelihood contribution of loyal consumers is therefore:

$$L_i(p_i, B_i = h|\mathcal{I}_i) = \underbrace{L_i(p_i, B_i = h, M_i = \mathbf{n}|\mathcal{I}_i)}_{L_i^N(p_i,h|\mathcal{I}_i)} + \underbrace{L_i(p_i, B_i = h, M_i = \mathbf{a}|\mathcal{I}_i)}_{L_i^A(p_i,h|\mathcal{I}_i)}. \tag{12}$$

Recall that the interior solution of the home-bank first-order condition is additive in ε_i: $p_i^0 = \bar{p}^0 + \varepsilon_i$. Therefore, if $\varepsilon_i < \bar{p} - \bar{p}^0$ we have $p_i = \bar{p}^0 + \epsilon_i$ and the search probability is constant: $H(\varepsilon_i) = \bar{H}$. Otherwise we do not have an interior solution and the price is equal to $\bar{p}$. The likelihood of observing p_i thus has a truncated form:

$$L_i^N(p_i, h|\mathcal{I}_i) = \begin{cases} f(p_i - \bar{p}^0)(1 - \bar{H}) & \text{If } p_i < \bar{p}, \\ \int_{p_i - \bar{p}}^{\bar{p} - c_h} (1 - H(\varepsilon_i)) f(\varepsilon_i) d\varepsilon_i & \text{If } p_i = \bar{p}, \end{cases} \tag{13}$$

where the search probability in the constrained case is equal to $H(\varepsilon_i) = 1 - \exp\big(-(E[W|\bar{p}, \varepsilon_i) - \lambda + \bar{p})/\bar{\kappa}_i\big)$.

The likelihood contribution from the auction mechanism involves the distribution of lowest-cost lender among competing options, denoted by $g_{(1)}(x)$. If the observed price is unconstrained, the transaction price is either equal to the competitive price $\lambda + c_{(1)} + \varepsilon_i$, or the reserve price $\bar{p}^0 + \varepsilon_i$. The latter outcome is realized if the initial quote is preferred to the price offered by the most efficient lender: $\bar{p}^0 + \varepsilon_i < \lambda + c_{(1)} + \varepsilon_i$. In contrast, the observed price is equal to $\bar{p}$ if the competitive price is larger than the posted price, *and* the initial quote is constrained: $\lambda + c_{(1)} + \varepsilon_i > \bar{p} > \bar{p}^0 + \varepsilon_i$. The likelihood of observing p_i from loyal consumers with the auction mechanism is given by:

$$L_i^A(p_i, h|\mathcal{I}_i) = \begin{cases} \begin{array}{l} \int_{-\infty}^{\bar{p} - c_h} g_{(1)}(p_i - \lambda - \varepsilon_i) H(\varepsilon_i) f(\varepsilon_i) d\varepsilon_i \\ \quad + \left[1 - G_{(1)}(\bar{p}^0 - \lambda)\right] \bar{H} f(p_i - \bar{p}^0) \end{array} & \text{If } p_i < \bar{p}, \\ \int_{\bar{p} - \bar{p}^0}^{\bar{p} - c_h} \left[1 - G_{(1)}(\bar{p} - \lambda - \varepsilon_i)\right] H(\varepsilon_i) f(\varepsilon_i) d\varepsilon_i & \text{If } p_i = \bar{p}. \end{cases} \tag{14}$$

Likelihood contribution for switching consumers If the observed price is unconstrained and the home bank offers a quote (i.e. $c_h + \varepsilon_i < \bar{p}$), the transaction price is equal to the minimum of $c_h - \lambda + \varepsilon_i$ and $c_{(2)} + \varepsilon_i$. If the consumer does not qualify for a loan at his/her home bank, the transaction price is the minimum of the posted-price, and the second-lowest cost. This occurs if $\epsilon_i > \bar{p} - c_h$. Therefore, the transaction price for switching consumers is equal to $\bar{p}$ if and only if the chosen lender is the only qualifying bank.

In the two cases where the transaction price is equal to $c_{(2)} + \varepsilon_i$, the consumer's choice reveals the most efficient lender (i.e. $c_{(1)} = c_{b_i}$), and the value of $c_{(2)}$ is the minimum cost among other lenders. We use $g_{-b_i}(x)$ to denote the density of lowest cost among $\mathcal{N}_i \backslash b_i$ lenders. Using this notation, we can write the likelihood contribution in the unconstrained case as the sum of three parts:

$$
\begin{aligned}
&L_i(p_i, b_i|\mathcal{I}_i) = \\
&\quad \int_{\bar{p}-c_h}^{\infty} g_{-b}(p_i - \varepsilon_i) G_{b_i}(p_i - \varepsilon) f(\varepsilon_i) d\varepsilon_i + \int_{p_i - c_h + \lambda}^{\bar{p}-c_h} g_{-b}(p_i - \varepsilon_i) G_{b_i}(p_i - \varepsilon) H(\varepsilon_i) f(\varepsilon_i) d\varepsilon_i \\
&\quad + (1 - G_{-b_i}(c_h - \lambda)) G_{b_i}(c_h - \lambda) f(p_i - c_h + \lambda) H(p_i - c_h + \lambda), \quad \text{If } p_i < \bar{p}.
\end{aligned} \tag{15}
$$

Note that the search probability is set to one in the first term, since the home-bank does not offer a quote (i.e. $c_h + \varepsilon_i > \bar{p}$). Also, the second term is equal to zero if $\bar{p} < p_i + \lambda$.[16]

In the constrained case, the likelihood contribution is given by:

$$
L_i(p_i, b_i|\mathcal{I}_i) = \int_{\bar{p}-c_h}^{\infty} \left(1 - G_{-b_i}(\bar{p} - \varepsilon_i)\right) G_{b_i}(\bar{p}_i - \varepsilon_i) f(\varepsilon_i) d\varepsilon_i, \quad \text{If } p_i = \bar{p}. \tag{16}
$$

Integration of other unobservables and selection The unconditional likelihood contribution of each individual is evaluated by integrating out the identity of the home bank h. Recall, that h is missing for a sample of contracts, and is unobserved for switchers. We therefore express the unconditional likelihood by summing over all possible combinations:

$$
L_i(p_i, b_i|X_i, \theta) = \sum_h \Pr\left(h|b_i, \rho, X_i\right) L_i(p_i, b_i|X_i, h, \beta),
$$

where $\Pr\left(h|b_i, \rho, X_i\right)$ is the conditional probability distribution for the identity of the home bank, and incorporates measurement error (ρ). Note that we condition on b_i when evaluating the home-bank probability since for switchers the probability that $h = b_i$ is zero.

In order to correct for selection, we calculate the probability of qualifying for a loan from at least one bank in consumer i's choice-set. This is given by the probability that the minimum of

[16]This creates a discontinuity in the likelihood, affecting primarily the parameters determining λ. To remedy this problem we smooth the likelihood by multiplying the second term in equation 15 by $(1 + \exp((\lambda - \bar{p} + p_i)/s))^{-1}$, where s is a smoothing parameter set to 0.01.

$c_{(1)} + \varepsilon_i$ and $c_h + \varepsilon_i$ is lower than $\bar{p}$:

$$\Pr(\text{Qualify}|X_i, \theta) = \sum_h \psi_h(X_i) \int_{-\infty}^{\infty} F(\bar{p} - \min\{c_{(1)}, c_h\}) g_{(1)}(c_{(1)}) dc_{(1)}, \tag{17}$$

where $\psi_h(X_i)$ is the unconditional probability distribution for the identity of the home bank.

Using this probability, we can evaluate the conditional likelihood contribution of individual i:

$$L_i^c(p_i, b_i|X_i, \theta) = L_i(p_i, b_i|X_i, \theta) / \Pr(\text{Qualify}|X_i, \theta). \tag{18}$$

Aggregate likelihood function The aggregate likelihood function sums over the n observed contracts, and incorporates additional external survey information on search effort. We use the results of the annual FIRM survey conducted by the Altus Group and presented in Table 2 to match the probability of gathering more than one quote along four dimensions: new-home buyers, city-size, region, and income group.

Using the model and the observed new-home buyers characteristics we calculate the probability of rejecting the initial quote; integrating over the model shocks and the identity of the home bank. Let $\bar{H}_g(\theta)$ denote this function for demographic group G.[17] Similarly, let $\hat{H}_g$ denote the analog probability calculated from the survey.

We use the central-limit theorem to evaluate the likelihood of observing $\hat{H}_g$ under the null hypothesis that the model is correctly specified. That is, under the model specification, $\hat{H}_g - \bar{H}_g(\theta)$ is normally distributed with mean zero and variance σ_g^2/N_g, where σ_g^2 is the model predicted variance in the search probability across consumers in group g, and N_g is the number of households surveyed by the Altus Group.[18] The likelihood of the auxiliary data is therefore given by:

$$Q(\hat{H}|\theta) = \prod_g \phi\left(\sqrt{N_g}(\hat{H}_g - \bar{H}_g(\theta))/\sigma_g\right), \tag{19}$$

where $\phi(x)$ is the standard normal density.

Finally, we combine $Q(\hat{H}|\theta)$ and $L_i^c(p_i, b_i|X_i, \theta)$ to form the aggregate log-likelihood function that is maximized when estimating θ:

$$\mathcal{L}(\mathbf{p}, \mathbf{b}|\mathbf{X}, \theta) = \sum_i \log L_i^c(p_i, b_i|X_i, \theta) + \log Q(\hat{H}|\theta). \tag{20}$$

[17]In order to reduce the computational burden associated with the calculation the average search probability, we simulate 10 realizations of the model shocks for each observed consumers. The results are not sensitive to this choice because we average over a large number of borrowers to calculate $\hat{H}_g$.

[18]We estimate σ_g by calculating the within group variance in search probability using the sample of individual contracts. Heterogeneity in search probability comes from dispersion in the number of options, the timing of house purchase, as well as financial characteristics of households in our data. Since this variance depends on the model parameter values, we follow a sequential approach: (i) calculate σ_g using an initial estimate of θ (e.g. starting with $\sigma_a = 1$), and (ii) hold σ_g fixed to estimate $\hat{\theta}$.

Notice that the two likelihood components are not on the same scale, since the FIRM survey contains fewer observations than the mortgage contract data-set. Therefore, we also test the robustness of our main estimates to the addition of an extra weight ω that penalizes the likelihood for violating the aggregate search moments:

$$\mathcal{L}^{\omega}(\mathbf{p},\mathbf{b}|\mathbf{X},\theta) = \sum_i \log L_i^c(p_i, b_i|X_i,\theta) + \omega \log Q(\hat{H}|\theta). \tag{21}$$

Computational steps In order to evaluate the aggregate likelihood function, we must first solve the optimal initial offer defined implicitly by equation 9. This non-linear equation needs to be solved separately for every consumer/home bank combination. We perform this operation numerically using a Newton algorithm that uses for the first and second derivatives of firms' expected profits. We also use starting values defined as the expected initial quote from the complete information problem, for which we have an analytical expression. This procedure is very robust and converges in a small number of steps. Notice that since the interior solution is additive in ε_i, this non-linear equation needs to be solved only once for each evaluation of the likelihood contribution of each household, $L_i(l_i, b_i|X_i, h, \beta)$.

In addition, the integrals are evaluated numerically using a quadrature approximation. To reduce the computational cost of numerical integration, we use the "sparse grid integration" technique discussed in Skrainka and Judd (2011) when the dimension of integration is larger than one. In addition, we Monte-Carlo integration to calculate the predicted search probability used in equation 19. The average search probability is calculated over a large number of households within each category, and is therefore less sensitive to approximation errors even with a small number of simulated draws (we use 10 draws for each household).

Finally, the parameters are estimated by maximizing the aggregate log-likelihood function defined in equation 20, using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) numerical optimization algorithm within the Ox matrix programming language (Doornik 2007).

4.3 Identification

The model includes four groups of parameters: (i) consumer observed heterogeneity (β), (ii) unobserved cost heterogeneity (σ_u and σ_ϵ), (iii) search cost ($\bar{\kappa}$), and (iv) switching cost (λ).

Although we estimate the model by maximum likelihood, it is useful to consider the empirical moments contained in the data. The contract data include information on market share, and conditional price distributions. For instance, we can measure the reduced-form relationship between average prices and the number of lenders in consumers' choice-sets, or other borrower-specific attributes. Similarly, we measure the fraction of switchers, along with the premium that loyal consumers pay above switchers. Finally, we augment the contract data with the fraction of consumers who gather more than one quote along four key borrower characteristics.

Intuitively, the cost parameters can be identified from the sample of switchers. Under the timing assumption of the model, most switchers are consumers who reject the initial quote, and initiate the competitive stage. The transaction price therefore reflects the second-order statistic of the cost distribution. This conditional price distribution can therefore be used to identify the contribution of observed consumer characteristics.

The residual dispersion can be explained by u or ϵ: the idiosyncratic and common unobserved cost shocks. To differentiate between the two, we exploit variation in the size of consumers' choice-sets. Indeed, the number of lenders directly affects the distribution of the second-order statistic through the value of σ_u. The reduced-form relationship between transaction rates and number of lenders, and the importance of residual price dispersion, therefore identify the relative importance of σ_u and σ_ϵ.

The data exhibit three sources of variation in the choice-set of consumers. First, consumers living in urban areas tend to face a richer choice-set than do consumers living in small cities. We exploit this cross-sectional variation, conditional on postal-code district fixed-effects.[19] Second, nearly 50% of consumers were directly affected by the merger between Canada Trust and Toronto Dominion Bank in 2000, and effectively lost one lender. The third source of variation comes from changes in the distribution of branches across markets.

The two remaining groups of parameters are identified from differences in the price distribution across switching and loyal consumers, as well as from the relative fraction of switchers and searchers. Intuitively the task is to tell the difference between two competing interpretations for the observed consumer loyalty: high switching cost (or loyalty premium), and/or high search cost.

In the model, the search and switching probabilities are functions of the search-cost and loyalty premium parameters. Intuitively, any differences between these two probabilities reveal the presence of positive switching cost. Indeed, we observe that 59% of consumers search in the population, while more than 75% of consumers remain loyal. This suggests a sizable loyalty premium. In addition, the level of the premium is separately identified from the observed price difference between loyal and switching consumers. Therefore, we have at least three moments to identify three parameters.

The model also implies strong restrictions on the relationship between search/switching, and observed characteristics of markets and loans. For instance, the value of shopping is increasing in the loan size and the number of competitors; both features that we observe in the survey data. Therefore, in practice the search cost and loyalty premium parameters are identified from more than three sources of variation.

Finally, the fact that we observe search and switching outcomes by income and new-home

[19] Postal-code districts are defined as the first letter of each postal-code. Ontario and Quebec have five and three districts respectively, and the rest of Canada have one district per province. We observe 16 districts in our data-set.

Table 3: Maximum likelihood estimation results

Heterogeneity and preferences	Est.	S.E.	Cost function	Coef.	S.E.
Common shock (σ_ε)	0.291	0.002	Intercept	3.590	0.054
Idiosyncratic shock (σ_u)	0.146	0.001	Bond rate	0.624	0.007
Avg. search cost			Loan size	0.089	0.016
$\bar{\kappa}_0$	-1.680	0.027	Income	-0.209	0.032
$\bar{\kappa}_{\text{inc}}$	0.603	0.037	Loan/Income	-0.111	0.011
$\bar{\kappa}_{\text{owner}}$	0.289	0.032	Other debt	-0.055	0.006
Home premium			FICO score	-0.510	0.028
λ_0	-2.040	0.006	Max. LTV	0.060	0.005
λ_{inc}	0.715	0.004	Previous owner	-0.012	0.005
λ_{owner}	0.036	0.003			
Measurement error	0.948	0.005	Market FE	X	
			Year FE	X	
Number of parameters	47		Quarter FE	X	
Sample Size	29,000		Bank FE	X	
Log-likelihood/10,000	-4.015		Bank FE Std-Dev	0.101	

Average search cost function: $\log(\bar{\kappa}_i) = \kappa_0 + \kappa_{\text{inc}}\text{Income}_i + \kappa_{\text{owner}}\text{Previous owner}_i$. Home bank premium function: $\log(\lambda_i) = \lambda_0 + \lambda_{\text{inc}}\text{Income}_i + \lambda_{\text{owner}}\text{Previous owner}_i$. Cost function: $C_i = L_i \times (Z_i\beta + \varepsilon_i - u_i)$. Units: $\$/100$

buyers status allows us to parametrize $\bar{\kappa}_i$ and λ_i as a function of these two variables.

5 Estimation results

5.1 Preference and cost function parameter estimates

Table 3 presents the maximum likelihood estimates of the key parameters of the model. The model is estimated on the full sample of 29,000 CMHC-insured contracts. The consumer preference and heterogeneity parameters are presented on the left-hand side, and the cost function parameters (β) on the right. The price coefficient is normalized to one and monthly payments are measured in hundreds of dollars. In order to better illustrate the magnitude of the estimates, we also present in Table 4 a series of marginal effects obtained by simulating contract terms using the estimated model.[20] We use this simulated sample in the goodness of fit analysis presented in the next subsection.

Unobserved heterogeneity and profit margins The first two parameters, σ_ε and σ_u, measure the relative importance of consumer unobserved heterogeneity with respect to the cost of lending.

[20] To obtain a simulated sample of contracts, we sample the random shocks of the model for every household in our main data-set, and compute the equilibrium outcomes. We repeat this process 11 times for each borrower.

The standard-deviation of the common component is 62% larger than the standard-deviation of idiosyncratic shock (i.e. 0.291 versus 0.187), suggesting that most of the residual price dispersion is due to consumer unobserved heterogeneity rather than to idiosyncratic differences across lenders.[21] Similarly, the estimates of the bank fixed-effects reveal relatively small systematic differences across lenders. Three of the eleven coefficients are not statistically different from zero (relative to the reference bank), and the standard deviation across the fixed-effects is equal to 0.106, or about half of the dispersion of the idiosyncratic shock.

Our estimate of σ_u has key implications for our understanding of the importance of market power in this market. Abstracting from bank fixed-effects, the average difference between the first and second lowest cost lender, $c_{(1)}$ and $c_{(2)}$, is equal to \$20 in duopoly settings, \$12 with three lenders, and approaches \$5 when N goes to 12. These differences imply that in an environment without quality differentiation, the competitive stage would lead to profit margins of about \$7 per month for the average market and a loan size of \$100,000.

In the model, market power also exists because of price discrimination motives (i.e. first-stage quote), and product differentiation associated with the loyalty premium. The first two rows of Table 4 show the distribution of monthly payments and lending costs for a homogenous loan size of \$100,000. The difference between the two leads to an average profit margin of \$17; more than twice the profits generated by idiosyncratic cost differences between lenders.

Importantly, profit margins are highly dispersed across consumers. In Figure 3 we plot the distribution of profits, expressed in basis points, for two groups of borrowers: searchers and non-searchers. Consistent with the previous discussion, margins for searchers are significantly lower, and mostly concentrated between 0 and 25 bps (the median is 16 bps). In contrast, the median profit margin is 33 bps for non-searchers. In both cases, the distribution has coverage from 0 to more than 100 bps, and the inter-decile range is equal to 54 bps.

Despite this large amount of dispersion, the average profit margins confirm that the market is fairly competitive. This is consistent with the idea that mortgage contracts are nearly homogenous across lenders, and represent a large share of consumers' budgets.

In addition, profit margins estimates should be contrasted with the average spread between transaction rates and the 5-year bond-rate: 27 bps versus 130 bps. Recall that the average borrower is able to negotiate 75 bps off the posted-rate. The marginal cost of lending is therefore roughly 100 bps below the posted-rate, and 100 bps above the 5-year bond-rate. This implies that each transaction involves significant transaction costs over the cost of funds. These costs can originate from a variety of sources: the compensation of loan officers (bonuses and commissions), the premium associated with pre-payment risks, and transaction costs associated with the securitization of contracts.

Search cost and loyalty premium The bottom two panels of Table 4 report the predicted distribu-

[21]The standard deviation of an extreme-value random variable is equal to $\sigma_u \pi/\sqrt{6}$, or 0.18 in our case.

Table 4: Model predictions and marginal effects

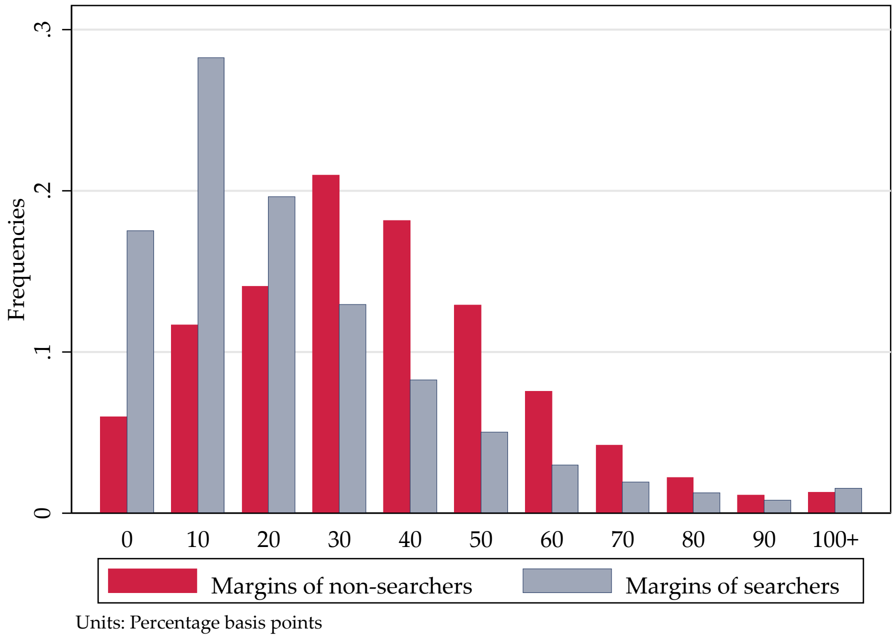
VARIABLES	Mean (1)	Std-Dev (2)	P-25 (3)	Median (4)	P-75 (5)
Monthly payment	705.99	49.55	672.09	703.59	739.94
Lending cost	688.49	50.12	653.80	686.87	724.43
Non-qualifying probability	0.06	0.23			
Payment marginal effects:					
Δ^{sd} Income	4.59	2.70	2.55	4.33	6.36
Δ^{sd} Loan size	-10.83	3.51	-12.70	-10.11	-8.33
Lending cost marginal effects:					
Δ^{sd} Income	1.40	3.19	-1.01	1.10	3.50
Δ^{sd} Loan size	-5.44	4.16	-7.65	-4.58	-2.49
Search cost – κ_i	29.52	33.01	6.86	19.15	40.70
Δ^{sd} Income	5.31	1.42	4.34	4.90	5.88
Δ Previous owner	11.07	7.56	6.19	9.51	13.73
Home bank premium – λ_i	21.99	5.42	18.60	20.82	23.73
Δ^{sd} Income	4.42	1.09	3.74	4.19	4.77
Δ Previous owner	0.80	0.19	0.68	0.76	0.86

Monthly payment and Lending costs are normalized to represent a $100,000 loan. Δ^{sd} corresponds to the effect of a one standard deviation increase in income or loan size. Δ *Previous owner* measures the marginal effect of being a previous owner borrowers relative to a new home buyers. Search costs and home-bank premiums are measured on a per-month basis.

tion of search costs and loyalty premiums, as well as the effect of loan-size and income on these two parameters. The parameters entering the search cost distribution suggest that search frictions are economically important. The average search cost is $\$29$, and is increasing in income and ownership experience. In particular new home-buyers are estimated to have significantly lower search costs on average ($\$11.07$). The effect of income is somewhat smaller. A one standard-deviation increase in income leads to a $5 increase in the average search cost of consumers. This is consistent with an interpretation of search costs as being proportional to the time cost of collecting multiple quotes.

The fact that new home-buyers face lower search costs is somewhat counter-intuitive, since previous owners are, in principle, more experienced at negotiating mortgage contracts. In the data, this difference is identified from the fact that new-home buyers are significantly more likely to switch, and are less likely to gather more than one quote according to the national survey. However, despite these differences, conditional on other financial characteristics, previous owners

Figure 3: Distribution of profit margins



are observed to pay only slightly more than new-home buyers (about 3 bps). Therefore, the model reconciles these facts by inferring that new home buyers face relatively low search costs, but are associated with a higher lending cost of about $1.5/month for a $100,000 loan.

To understand the magnitude of these estimates, it is useful to aggregate the monthly search cost over the length of the contract. According to the model, the marginal consumer accepting the initial quote is indifferent between searching, and reducing his expected monthly payment by $\$\kappa_i$. Over a five year period, assuming an annual discount factor of 0.96, these estimates correspond to an average upfront search cost of $1,657, and a median of $1,028.[22] Are these number realistic? Hall and Woodward (2010) calculate that a U.S. home buyer could save an average of $983 on origination fees by requesting quotes from two brokers rather than one. Our estimate of the search cost is consistent with this measure.

Turning to the estimate of λ_i, we find that the average loyalty premium is equal to $22 per month. Like with search costs, new home-buyers enjoy a smaller premium, but the difference is small ($0.80 per month). In comparison, the effect of income on the loyalty premium is much larger; a one standard deviation increase in income raises λ_i by $4.42 per month.

Over five years, the discounted value of the loyalty premium corresponds to an upfront value of approximately $1,028. Assuming that this utility gain originates from avoiding the cost of

[22]The search cost is measured in terms of monthly payment units. Since the contract is written over a 60 month period, the discounted value of the search cost is equal to $\sum_{t=0}^{60} \frac{\kappa_i}{(1+r)^{60}}$. With an annual discount factor of 0.96 the monthly interest rate is 0.3%.

switching bank affiliations, our results suggest that switching costs are large, and of a similar order of magnitude to the cost of gathering multiple quotes.

Another interpretation, of course, is that the loyalty premium is caused by complementarities between mortgage lending and other financial services. For instance, consumers could perceive that combining multiple accounts under one bank improves the convenience of the services, which would lead to direct utility gains. In addition, ithe home bank can compete with other mortgage lenders by offering discounts on other services, such as checking/saving accounts or preferential terms on other loans or lines of credits. This interpretation is valid only if other multi-product lenders cannot make similar offers, because, for instance, switching main financial institutions is too costly.

Recent surveys of Canadian households' banking activities are consistent with this interpretation. Statistics Canada reports that, on average, Canadians spend about $16 a month on banking fees, and that approximately 29% do not pay any banking fees due to discounting.[23] Moreover, the Canadian Finance Monitor (CFM) survey indicates that transaction fees are increasing in income, which is consistent with our result that higher-income households have a larger loyalty premium.[24]

Loan size and income effects In order to better understand the role played by loan size and income, we report in Table 4 the marginal effect of both variables on monthly payments and marginal costs. The monthly payment marginal effects are obtained by regressing predicted monthly payments on all the state variables of the model (i.e. financial characteristics, market structure, and fixed-effects), while the lending cost marginal effects are obtained directly from the cost function reported in Table 3. Note that monthly payments are measured using a common loan of $100,000, in order to eliminate any mechanical relationships between loan size or income and monthly payments.

Consistent with previous findings in Allen et al. (2013b), the model predicts that, after conditioning on financial and demographic characteristics of borrowers, richer households pay higher rates, and consumers financing bigger loans are more likely to obtain large discounts. For a loan size of $100,000, Table 4 shows that a one standard deviation increase in income increases payments by $4.59/month, while a one standard-deviation increase in loan size decrease payments by $10.83/month.

The estimated lending cost function reveals that only about thirty percent of the income effect on payments is due to cost differences; the rest is explained by larger search costs and a loyalty premium. Similarly the lending cost function is non-monotonic in income: the effect of increasing income by one standard-deviation is negative at the top of the income distribution (i.e. from the 75% percentile).

[23]Source: Statistics Canada, Selected Household expenditures items (2009).
[24]Those results are available upon request.

Figure 4: Distribution of loan-size for searchers and non-searchers

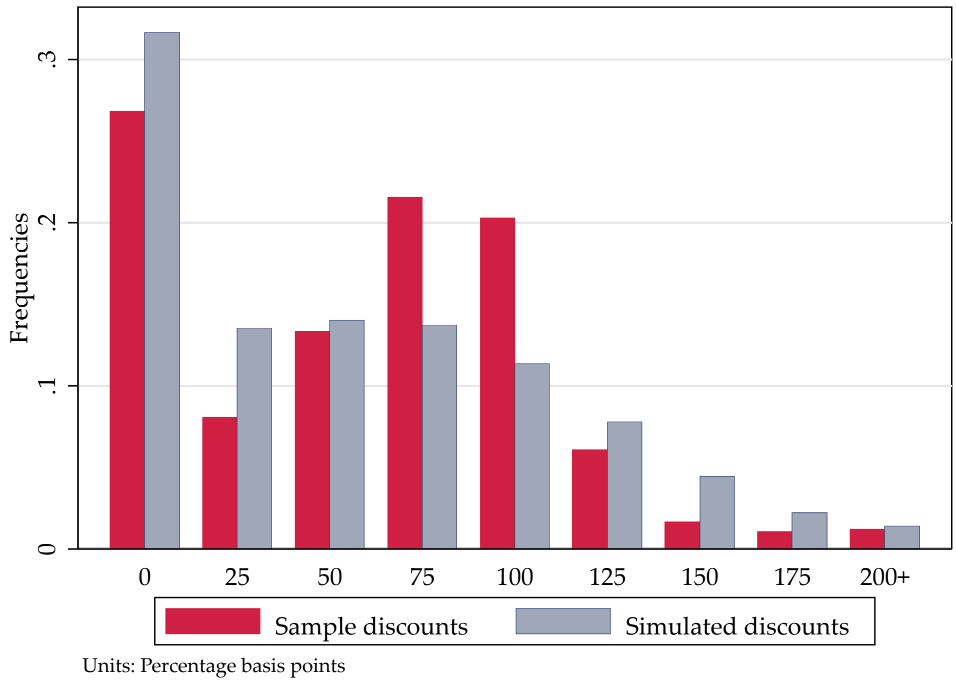
The positive relationship between lending cost and income is consistent with the fact that banks mostly face pre-payment risks, given the insurance coverage provided by the government against default risks. The fact that the sign of the income effect is reversed at the top highlights the value of attracting wealthier customers, both because of lower default risks and larger revenues from complementary services. Given the prevalence of one-stop-shopping in banking, this increases the opportunity cost of not serving wealthier households.

Looking at the loan-size marginal effects, roughly half of the reduced-form relationship is explained by cost differences. A one standard-deviation increase in loan size reduces the cost of lending by $5.44 per month. The remainder is explained by the search decision of consumers. As Figure 4 shows, consumers financing larger loans are more likely to search. This is because the gains from search are increasing in loan size, while the search cost is fixed. Note that this relationship is also true in the FIRM survey. Households earning more than $60,000 (a proxy for loan size) are 10.5% more likely to search multiple lenders than those earning less than $60,000.

Additional specifications Table 12 in the Appendix presents the results of three alternative specifications. The first specification uses a homogenous search-cost distribution and common loyalty premium, the second incorporates data from CMHC and Genworth contracts, and the third increases the weight on the aggregate search moments by setting $\omega = 100$ in equation 19.

The first specification is nested in the baseline specification presented in Table 12, which allows us to formally test the restrictions. The likelihood ratio test shows that incorporating observable

Figure 5: Predicted and observed distribution of negotiated discounts



differences in the search cost and loyalty premium improves significantly the fit of the model, as the null hypothesis represented by columns (1) is easily rejected.

We cannot provide the same statistical interpretations to the likelihood ratio in the second specification, but it is clear that the model fit is better within CMHC data than in the combined sample.[25] This is in part due to the fact the Genworth excludes contracts from the "Other bank" category, while CMHC does not.

The third specification reveals that it is necessary to increase average search cost and loyalty premium in order to match the aggregate search moments. As we will discuss further below, matching these moments also requires larger idiosyncratic differences across lenders. This is reflected by the ratio of σ_ε over σ_u, which is much smaller than in the baseline specification: $\sigma_\varepsilon/\sigma_u = 2.03$ without the penalty weight, versus 1.59 with $\omega = 100$.

5.2 Goodness of fit

We next provide a number of tests for the goodness of fit of our search and price negotiation model. Figure 5 shows that the estimated model reproduces fairly well the overall shape of the discount distribution. There are two main takeaways. First, the data show a large mass of consumers

[25]The log-likelihood in the sample with Genworth is re-weighted so that the two statistics are on the same scale, despite the fact that the Genworth sample has more observations.

receiving 75 and 100 bps discounts. This would appear to be the result of bunching by loan officers around a common discount size, which is not something that the model can predict.

Second, a related implication of this behavior is that few consumers receive small discounts, and the density of discounts is sharply increasing past zero in the data. The model predicts a similar pattern, but is much less pronounced. This prediction from the model is mostly caused by the distribution of discounts among non-searchers, which is strictly decreasing. In contrast, the model implies a discount distribution for searchers that has a similar dip at 25 bps, because few consumers gathering multiple quotes receive small discounts.

Table 5 looks at how well the model matches the search probabilities of different demographic groups. The first column corresponds to the model prediction using our baseline specification, and the last two reproduce the aggregate moments from the national survey of new home buyers.

Overall, the model tends to over predict the amount of search in the market. The unconditional average search probability predicted by the model is 64%, compared with 59% according to the national survey. Similarly, while the model matches reasonably well the qualitative predictions of the survey, it has a hard time matching the magnitude of the differences across groups. This is especially true for the differences across small and large cities, which are nearly 20 percentage points in the survey data, and 10 percentage points in the model. Also, the model cannot rationalize the non-monotonicity in the relationship between city size and search probability.

Note that most of the differences between the model predicted probabilities and survey results are not statistically significant, given the relatively small number of observations in the survey. In the baseline specification, three out ten mean differences are statistically different from zero using a 10% significance level.

Importantly, the middle column shows that the model can rationalize most of the observed search patterns, by imposing a larger weight on the aggregate moments (i.e. specification 3 in Table 12 presented in the Appendix). Across all the groups, the model matches well the survey results, and the predicted search probability is exactly equal to 59%. Only one mean difference is statically different from zero; the one corresponding to the non-monotonicity of the search probability with respect to city size.

The fact that the baseline specification does not as accurately match the aggregate moments highlights a tension between the price and search moments. As hinted by specification (3) in Table 12, the model requires a relatively large search cost and loyalty premium to bring the search probability down to less than 60%. In turn, this increases the predicted average discount that switching consumers obtain, much beyond what we observe. In addition, the model requires larger idiosyncratic differences across lenders to match the observed relationship between market size and search. This is because σ_u determines the rate at which the gain from search increases with competition. However, increasing σ_u also leads to a steeper reduced-form relationship between price and market structure than the one we observe in the data. Since the number of observations

Table 5: Observed and predicted search probability by demographic groups

	Baseline Specification	Penalty Specification	Survey data	
			Avg.	Nb. Obs.
	(1)	(2)	(3)	(4)
Income				
$>$ \$60K	0.657	0.623	0.619	126
$\leq$ \$60K	0.614	0.540	0.560	141
Ownership status				
New home buyers	0.650	0.673	0.673	153
Previous owners	0.606^b	0.509	0.509	106
City size				
Pop. $>$ 1M	0.673	0.645	0.640	75
1M $\geq$ Pop. $>$ 100K	0.627	0.565^b	0.667	114
Pop. $\leq$ 100K	0.584^a	0.506	0.443	79
Regions				
East	0.586	0.492	0.515	103
Ontario	0.669	0.655	0.716	102
West	0.638^c	0.564	0.534	73

Null hypothesis: Survey average = Model average. Significance levels: a = 1%, b = 5%, c = 10%. P-values are calculated using the asymptotic standard-errors of the survey.

in the contract data is much larger than the number of households in the survey, the un-penalized likelihood resolves this conflict by assigning relatively more weight to the price relationships.

Finally, in Table 6 we evaluate the ability of the model to reproduce the observed reduced-form relationships between transaction rates and observed characteristics of borrowers. To highlight the ability of the model to explain the cross-sectional distribution of rates, we regress the interest-rate spread, simulated and observed, on financial and market characteristics of the borrowers.

The comparison between columns (1) and (2) clearly shows that the model does a good job at predicting most reduced-form relationships associated with financial characteristics. The R^2 reported at the bottom also shows that the model predicts a similar amount of residual price dispersion: 0.345 versus 0.407. Similarly, the average marginal effects of loan size and income on transaction rate are well explained by the model (bottom).

The model also predicts well the relationship between the relative size of branch networks and rates. We measure the network size by taking the ratio of the number of branches over the average number of branches of competing banks. The estimated model reproduces well the fact that consumers dealing with large-network banks pay higher rates on average. In the model, this is entirely due to the fact that consumers start their search with their home bank, and face large search costs on average. Column (4) shows that after controlling for the search decision of consumers, the effect of network size on rates is zero.

The model tends to over-predict the impact of the log number of competitors on rates. One

Table 6: Reduced-form interest rate spread regressions

VARIABLES	Sample	Simulations		
	(1)	(2)	(3)	(4)
Prior relationship	-0.0792^{a}	-0.368^{a}	-0.453^{a}	-0.218^{a}
	(0.00866)	(0.00254)	(0.00203)	(0.00248)
Search indicator				-0.353^{a}
				(0.00217)
Previous owner	0.0305^{a}	0.0183^{a}	0.0128^{a}	0.00796^{a}
	(0.00720)	(0.00245)	(0.00225)	(0.00218)
Relative network size	0.0174^{a}	0.0154^{a}	0.00382^{a}	0.00193
	(0.00405)	(0.00145)	(0.00124)	(0.00118)
Number of competitors (log)	-0.0426^{a}	-0.0867^{a}	-0.0630^{a}	-0.0762^{a}
	(0.0128)	(0.00499)	(0.00400)	(0.00395)
Bond rate	-0.445^{a}	-0.391^{a}	-0.392^{a}	-0.391^{a}
	(0.00841)	(0.00270)	(0.00258)	(0.00252)
Loan size (/100,000)	-0.00217	-0.000214	0.0252^{a}	0.0225^{a}
	(0.0165)	(0.00769)	(0.00664)	(0.00610)
Income (/100,000)	-0.181^{a}	-0.145^{a}	-0.170^{a}	-0.160^{a}
	(0.0337)	(0.0154)	(0.0137)	(0.0124)
Loan/income	-0.185^{a}	-0.147^{a}	-0.144^{a}	-0.138^{a}
	(0.0125)	(0.00512)	(0.00478)	(0.00444)
Other debts	-0.0777^{a}	-0.0725^{a}	-0.0737^{a}	-0.0704^{a}
	(0.00852)	(0.00307)	(0.00275)	(0.00262)
FICO score (/1000)	-0.767^{a}	-0.637^{a}	-0.644^{a}	-0.624^{a}
	(0.0402)	(0.0132)	(0.0124)	(0.0122)
Maximum LTV	0.0895^{a}	0.0732^{a}	0.0739^{a}	0.0712^{a}
	(0.00627)	(0.00199)	(0.00194)	(0.00190)
Constant	4.733^{a}	4.362^{a}	4.339^{a}	4.439^{a}
	(0.0804)	(0.0259)	(0.0237)	(0.0229)
Average marginal effects:				
Income effect	0.487	0.384	0.347	0.335
Loan size effect	-0.313	-0.246	-0.215	-0.207
Prior relationship	W/ Error	W/ Error	True	True
Observations	29,000	301,136	301,136	301,136
R-squared	0.345	0.407	0.450	0.493

Robust standard errors in parentheses

^{a} p<0.01, ^{b} p<0.05, ^{c} p<0.1

Dependent variable: Negotiated rate - 5 year bond rate.

Additional controls: Market, year and quarter fixed-effects

likely cause is the correlation between consumers' unobserved characteristics and the number of competitors in each local market. Indeed, controlling for location (FSA) fixed-effects in column

(1) lowers the competition coefficient by a factor of three. This suggests that our estimates suffer from a simultaneity bias, which biases downward our estimate of σ_u; the parameter measuring the strength of competition across lenders. Since this parameter determines the size of profit margins, it implies that our estimates provide a lower bound of the amount of market power.[26]

The reduced-form regressions reveal that the model also over-estimates by a significant margin the premium that loyal consumers pay (i.e. 8 bps versus 36.8 bps). This is despite the fact that the model incorporates measurement error in the switching variable. In the baseline specification, we estimate that about 5% of borrowers report this variable with error. Column (3) shows that incorporating measurement error helps bring this coefficient closer to the data. Using the true switch outcome, the model predicts that loyal borrowers pay on average 45 bps more than switchers. Finally, column (4) shows that this is mostly due to the fact that searchers are more likely to switch. Conditional on searching, loyal consumers pay about 22 bps more than switchers.

5.3 Summary of results

There are three main takeaways from the estimation of our model. The first is that markups are, on average, quite small, but vary substantially across borrowers. The average markup is estimated to be 4.31%, or just below 30 bps. The profit margins of non-searchers are 50% smaller the profit margins of non-searchers, and the within group dispersion is between 50 and 60 bps.

The second important result is that markups are mostly generated by search and switching costs. We find that almost 60% of the average profit margins of lenders can be explained by search costs and quality differences. We estimate an average search cost of $29 per month (median of $19). In addition, on average, consumers are willing to forego $22 a month to stay with their home bank and avoid having to switch banks.

How do these results compare to existing estimates in the literature? Perhaps the closest point of comparison comes from Honka (2012) analysis of the insurance market. She estimates cost of searching for policies to be $28 per online search and $100 per offline search, and switching costs of $115. To compare these numbers to ours we calculate the ratio of the search cost to the standard-deviation of monthly payments for the average loan size (i.e. $70/Month). In our context, this ratio is 27% ($= 19/70$).[27] In comparison, the estimates from Honka (2012) range between 10% for online transaction, and 35% offline.

We can also compare our findings to those of Hortaçsu and Syverson (2004) and Hong and Shum (2006). Hortaçsu and Syverson (2004) estimate a median search cost of 5 basis points (or $5 per $10000 invested), yielding a ratio of 8%. The average search cost across the four books considered by Hong and Shum (2006) is $1.58 (for non sequential search), yielding a ratio of 33%.

[26] Incorporating location fixed-effects in the structural model would likely solve this problem, but make the estimation intractable by increasing the number of parameters to nearly 1,000.

[27] Note that the median search cost of $19 corresponds to roughly 22 bps. Given rate spread in basis points of 61 bps, this corresponds to a ratio of about 35%.

Overall our results are consistent with those in the literature. We should also point out that searching for mortgages is more complicated than searching for many of the products studied in the literature because of the negotiation process. Also, our model implicitly assume that consumers search over all options. However, once they have identified the two most efficient lenders, consumers in practice need to haggle only with two firms. Therefore, we can think of the search process as one with two steps: first, identify the most efficient lenders, second, haggle.

Finally, consistent with reduced-form evidence, our model predicts that, after conditioning on financial and demographic characteristics of borrowers, richer households pay higher rates, and consumers financing bigger loans are more likely to obtain large discounts. Using the model we can decompose this effect into cost-difference and search components. We find that only about thirty percent of the income effect is due to cost differences with the rest coming from larger search costs and the loyalty premium. Similarly, our model suggests that around half of the reduced-from relationship between loan-size and rates is explained by cost differences, with the remainder coming different incentives to search.

6 Counter-Factual Analysis

In this section, we use the estimated model to simulate a counter-factual equilibrium with zero search costs in order to quantify the effect of search frictions on consumer welfare. We also analyze the extent to which quality differentiation and the presence of a price ceiling can amplify or attenuate the welfare cost of search frictions.

We then, in the following subsection, explore the channels through which competition impacts the adverse effects of search frictions. We perform this analysis by calculating the welfare cost of search frictions in hypothetical markets of different sizes, and by simulating the effect of counter-factual mergers across increasingly competitive markets.

6.1 Quantifying the effect of search frictions

The presence of search costs lowers the welfare of consumers for three distinct reasons. First, it imposes a direct burden on consumers searching for multiple quotes. Second, it can prevent non-searching consumers from matching with the most efficient lender in their choice set (adjusting for quality differences), thereby creating a misallocation of buyers and sellers. Lastly, it opens the door to price discrimination, by allowing the initial lender to make relatively high offers to consumers with poor outside options and/or high expected search costs. These factors can be identified by decomposing the change in consumer surplus caused by the presence of search frictions:

$$\begin{aligned}
\Delta\text{CS}_i &= \underbrace{\lambda_i 1(b_i = h) - c_i - p_i - S_i\kappa_i}_{\text{CS}_i} - \underbrace{\lambda_i 1(b_i^0 = h) - c_i^0 - p_i^0}_{\text{CS}_i^0} \\
&= \left[(\lambda_i 1\{b_i = h\} - c_i) - \left(\lambda_i 1\{b_i^0 = h\} - c_i^0\right)\right] - \left(m_i - m_i^0\right) - \kappa S_i \\
&= \Delta V_i - \Delta m_i - S_i\kappa_i, \qquad (22)
\end{aligned}$$

where the 0 superscript indicates the equilibrium outcomes without search cost, $V_i = \lambda_i 1\{b_i = h\} - c_i$ is the transaction surplus (excluding of the search cost), $m_i = p_i - c_i$ is the profit margin, and S_i is an indicator variable equal to one if the consumer rejects the initial offer.

We label the three components *misallocation*, *discrimination*, and *search cost*, respectively. The first and third components are standard frictions in the search literature, and lead to a decrease in total welfare. The elimination of the price discrimination incentive, on the other hand, does not need to be welfare improving for all consumers.

We simulate the counter-factual experiments as follows. First, we randomly select a sample of 5,000 households from the main data set. Second, for each simulated household, we randomly sample the realization of their idiosyncratic shocks: (i) the identity of their home bank h_i, (ii) the common lending cost ε_i, (iii) the vector of idiosyncratic match values u_{ij}, and (iv) the private-value search cost κ_i. Third, at each realization of these shocks, we solve the optimal initial offer and search decision, and then store the endogenous outcomes. Fourth, we simulate the equilibrium without search costs by using solely the competitive auction stage with a reserve price equal to the posted-rate, holding fixed the realized shocks. Finally, we repeat steps two through four 100 times, and average the equilibrium outcomes across the realization of the idiosyncratic shocks.

Table 7 presents the main simulation results. Panel (A) corresponds to the baseline environment, and uses the observed distributions of posted-rates and loyalty premiums. The other two panels report the simulation results in two counter-factual environments: (B) no quality differentiation between lenders (i.e. $\lambda_i = 0$ for all i), and (C) no posted-rates.

Columns 1 through 3 show the change in the misallocation, discrimination and search cost components respectively, while column 4 presents the total change in consumer surplus. To illustrate the heterogeneity across consumers, the first line of each panel reports the fraction of simulated consumers experiencing zero changes, and the next three describe the conditional distribution of non-zero changes. To calculate the cumulative changes, we sum the changes across all consumers qualifying for a loan, and divide by the number of simulated consumers (5,000).[28] The percentage shares of each component is expressed relative to the cumulative changes.

In the baseline environment, we estimate that the cumulative reduction in consumers surplus

[28] Note that the cumulative effect differs from the average across qualifying consumers, since a fraction of consumers fail to qualify for a loan when the posted-rate is present. On average we estimate that 5.4% of consumers in the baseline environment are not able to qualify for a loan because of the posted rate (same in Panel B).

Table 7: Decomposing the effect of search frictions on welfare

		Consumer surplus change decomposition			Δ CS
		Misallocation	Discrimination	Search cost	
		(1)	(2)	(3)	(4)
Panel A: Baseline environment					
Zero changes (%)		79.7	67.5	35.8	3.6
Non-zero changes ($)	P10	-63.18	-15.27	2.01	-42.94
	P50	-19.46	13.22	12.46	-16.24
	P90	-3.17	33.75	35.19	-2.57
Cumulative changes	$	-5.36	3.21	10.34	-18.91
	%	28.4	17.0	54.7	100
Panel B: No differentiation					
Zero changes (%)		77.95	75.38	24.85	0.29
Non-zero changes ($)	P10	-72.24	-22.30	2.34	-49.13
	P50	-26.85	9.39	14.35	-18.60
	P90	-6.56	29.20	40.65	-2.99
Cumulative changes	$	-7.23	1.32	13.60	-22.15
	%	32.7	6.0	61.4	100
Panel C: No price-ceiling					
Zero changes (%)		81.51	68.51	31.12	0.00
Non-zero changes ($)	P10	-65.24	-5.02	2.13	-42.26
	P50	-20.38	22.01	12.67	-18.90
	P90	-3.33	38.30	30.78	-3.03
Cumulative changes	$	-5.36	5.80	10.31	-21.46
	%	25.0	27.0	48.0	100

Each entry corresponds to an average over 100 simulated samples. Each sample is equal to 5,000 consumers. Averages in the first four lines are calculated using the samples of consumers facing non-zero changes. The fraction of zero changes is computed over the samples of qualifying consumers. Cumulative changes are the sum of all changes divided by the total number of simulated consumers, including non-qualifying consumers. The welfare decomposition corresponds to: $\Delta \text{CS}_i = \Delta V_i - \Delta m_i - \Delta \kappa_i S_i = \text{Misallocation}_i - \text{Discrimination}_i - \text{Search cost}_i$. The last row reports the contribution of each component, in percentage of the cumulative change.

associated with search frictions is equal to $18.91 per month. The largest component (55%) is attributed directly to the sunk cost of searching, followed by the misallocation of contracts (28%), and the increase in profit margins associated with price discrimination (17%). Interestingly, over 96% of consumers are adversely affected, with the median consumer experiencing a $16 reduction in surplus.

Looking first at misallocation and search, the first row measures the fraction of buyers and sellers that are matched efficiently and the fraction of non-searchers in the presence of search frictions. In the baseline environment, search costs cause 20% of transactions to be misallocated, despite the fact that more than 35% of consumers are not searching (column (3)). Note that this difference would be zero if firms had homogenous costs and willingness to pay. Given that banks' fixed effects are not highly dispersed, this difference is caused mostly by the fact that consumers visit the highest expected surplus seller first, which reduces the fraction of inefficient matches.

Panel B illustrates this point by removing differentiation between firms. In this case, the home bank makes the first offer, but does not provide the highest expected transaction surplus. This inefficient sequencing of quotes leads to an increase in search, from 64% to 75%, and therefore to an increase in search costs, from \$10.34/month to \$13.60/month. Despite this increase in search effort, the fraction of inefficient matches is larger without quality differentiation by about two percentage points, and the cumulative losses associated with misallocation increase by 35% (i.e. from -5.36 to -7.23). This reflects mostly the direct loss in surplus associated with the reduction in willingness-to-pay for loyal consumers.

We next consider the change in consumer surplus associated with profit margins. Column (2) shows that the relatively small contribution of the price discrimination component (17% in the baseline), is due in large part to the fact that many consumers pay *larger* markups in the frictionless market. In the baseline environment, the median change in profit margins is equal to \$13.22 per month; slightly more than the median increase in search cost. However, the tenth percentile consumer benefits from a \$15.27 reduction in profit margins, which brings the cumulative effect down to \$3.21.

To understand this heterogeneity, it is important to note that the initial quote is used both as a price discrimination tool, and as a price ceiling in the competition stage. The home-bank is in monopoly position in the first stage, and can set individual prices based on consumers' expected outside options. This is a form of first-degree price discrimination that leads it to strictly higher expected profit margins.[29]

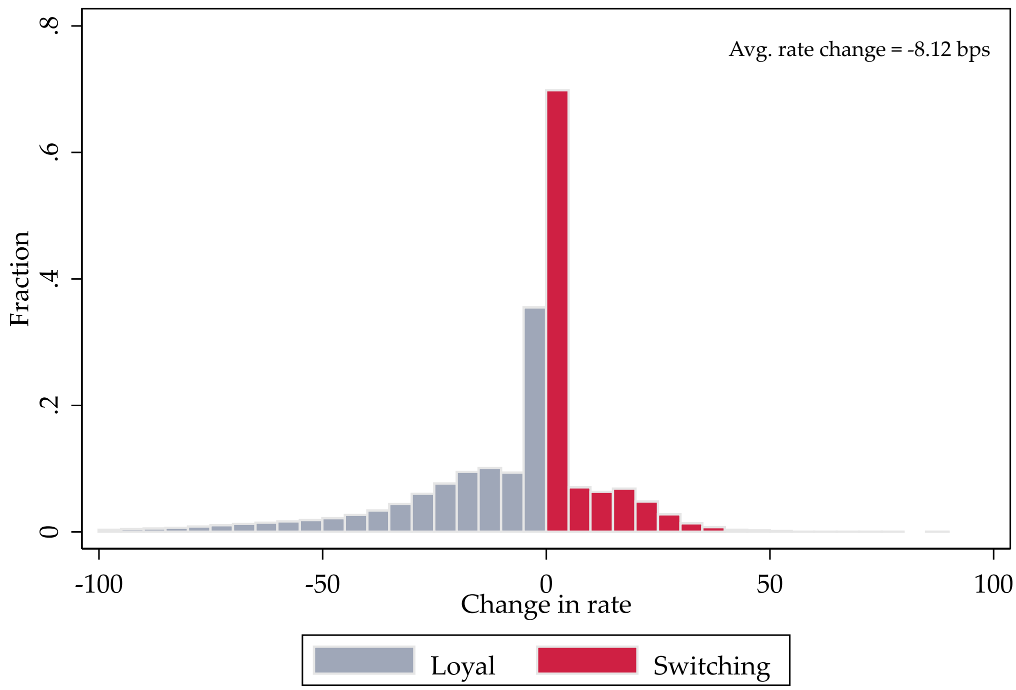
This adverse effect is weighed against the fact that the initial offer can be recalled, and therefore protects consumers against excessive market power at the auction stage. Notice that this factor is present only because consumers and firms have imperfect information about the gains from search. If firms had perfect information about the outcome of the auction, the initial offer would not distort prices in the second stage.

Removing the loyalty premium forces the home bank to behave more competitively in the first stage, which reduces the price discrimination contribution to the welfare cost of search frictions by close to 60%. This change is the result of a reduction in the home bank's markup, and an increase in the relative market power of competing lenders in the second stage. The latter effect implies that a larger fraction of consumers experience a decrease in profit margins following the elimination of search costs, relative to the baseline.

This asymmetry can be seen in Figure 6. This figure plots the distribution of the change in

[29] We use *expected* profits here because the outcome of the second stage is random, and the realized profit margins of the home bank at the auction can be larger than the known margin in the first stage.

Figure 6: Change in rates after removing the loyalty premium



Bar width = 5 bps. Loyal/Switching: Bank choice before removing the home-bank premium.

transaction rates, expressed in bps, following the elimination of the loyalty premium in the baseline environment with search frictions. The market becomes more competitive, and transaction rates decrease by 8.12 bps on average. However, this average effect hides the fact that the distribution of rate changes is strictly positive within the subsample of switching consumers.

Overall, the comparison between the baseline and no-loyalty environments can help us understand the role of product differentiation in search markets. When one seller offers a strictly higher quality service, eliminating differentiation reduces its ability to price discriminate. Because this increases the relative market power of competing firms, it does not offset the increases in misallocation and search costs that result from marking products ex-ante homogenous. Therefore, the cumulative welfare cost of search frictions increases by 17.5% when we remove differentiation.

The last panel simulates an environment with no price ceiling. Relative to the baseline, the cumulative effect on surplus increases by 13.5%, which implies that the presence of the posted rate attenuates the welfare cost of search frictions. As column (2) clearly shows, this is the result of a large increase (80%) in the welfare contribution of price discrimination. This highlights the fact that the presence of a price ceiling greatly limits the ability of the initial lender to price discriminate. In the baseline simulation, we estimate that nearly 50% of initial quotes are constrained by the posted rate.

Notice also that the cumulative search and misallocation components are nearly identical with or without a posted-rate. This is the result of two offsetting forces. On the one hand, eliminating the posted rate raises substantially the initial quote, which in turn increases the search probability among consumers qualifying for a loan from their home bank. On the other hand, by eliminating the price ceiling, more consumers with relatively poor financial characteristics are able to get a quote from their home bank, and therefore do not necessarily have to search for additional quotes (especially those with high search costs). These opposing forces cancel each other out almost exactly.

6.2 Market structure and search frictions

As pointed out by Janssen and Moraga-Gonzalez (2004), the effect of market concentration on consumer welfare and on the distribution of prices is ambiguous when search frictions are important. These relationships have so far only been studied in the context of oligopoly models in which firms randomize posted prices, but a similar ambiguity exists when prices are negotiated individually and search is costly.

In terms of welfare, as markets become more competitive the quality of matching opportunities improves, and the ability of firms to price discriminate is attenuated. Both factors lead to lower prices, and greater surplus for consumers. However, if consumers are more likely to search as the number of firms increases, these gains come at the expense of larger search costs. Moreover, because search opportunities and costs are heterogenous, the overal benefits of competition are not spread equally across consumers. As a result, the impact of an increase in competition on the distribution of prices is ambiguous.

To dissect the effect of market structure on consumer surplus and rates, we simulate a series of counter-factual experiments in which we vary the number of competitors, holding fixed the identity and characteristics of lenders. In particular, we eliminate all systematic cost differences between lenders, and assume that consumers start their search with the same home bank. Then, for each simulated consumer and each realization of the idiosyncratic shocks, we solve the equilibrium outcomes by incrementally changing the number of competitors in $\mathcal{N}_i$, from 1 to 12. Importantly, at each step, we solve the game holding fixed the match values of all existing lenders. Therefore, in the zero search-cost environment, the surplus of consumers is strictly increasing in N.

We use the results of these experiments to answer two questions. First, following on the analysis performed above, we evaluate to what extent competition attenuates the welfare cost of search frictions. Second, we study directly the effect of competition on welfare and prices, by comparing the equilibrium outcomes with search frictions across alternative market structures. This effectively creates twelve counter-factual “mergers”, which we use to study the effect of competition of consumers welfare, search effort, and price dispersion.

Table 8: Welfare effects of search frictions in competitive and non-competitive markets

Number of competitors	Weflare change decomposition (%) Misallocation (1)	Discrimination (2)	Search cost (3)	Δ CS ($/month) (4)
2	12.97	42.03	45.00	-12.13
3	18.89	32.77	48.35	-14.33
4	23.34	26.28	50.38	-15.72
5	26.50	21.53	51.97	-16.76
6	28.72	17.83	53.45	-17.62
7	30.17	15.04	54.79	-18.36
8	30.98	12.94	56.08	-19.01
9	31.48	11.29	57.23	-19.59
10	31.72	9.96	58.32	-20.11
11	31.71	8.96	59.33	-20.57
12	31.48	8.22	60.30	-21.00
13	31.22	7.60	61.19	-21.39

Each entry corresponds to an average over 100 simulated samples. Each sample is equal to 5,000 consumers. The last column reports the cumulate change in consumer welfare divided by the total number of simulated consumers, including non-qualifying consumers. The welfare decomposition corresponds to: $\Delta \text{CS}_i = \Delta V_i - \Delta m_i - \Delta \kappa_i S_i = \text{Misallocation}_i - \text{Discrimination}_i - \text{Search cost}_i$. Columns (2)-(4) report the percentage contribution of each component to the cumulative change.

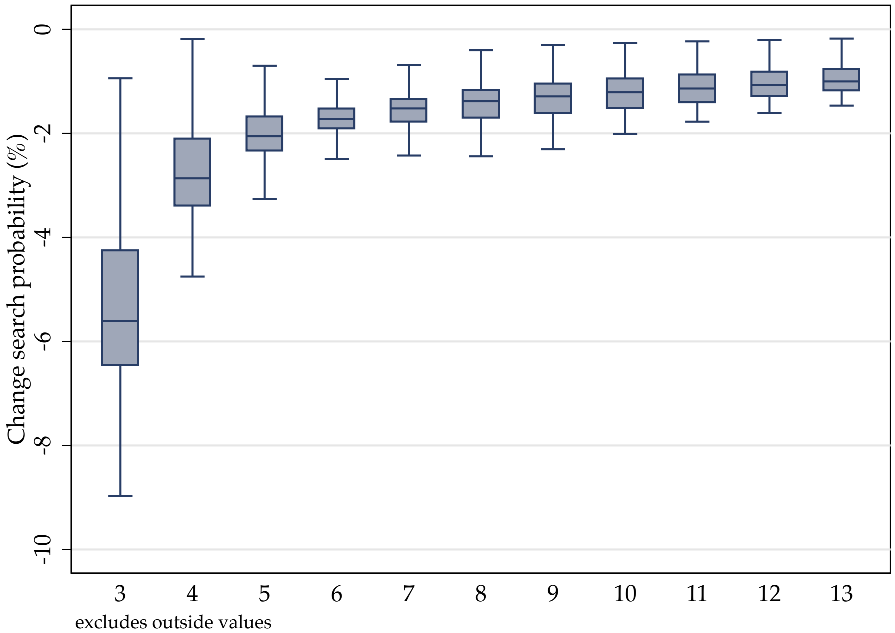
Table 8 decomposes the effect of search frictions on consumer surplus in different market structures. Each row represents a different market structure, and the columns summarize the welfare impact of introducing search frictions. Columns (1) to (3) present the percentage contribution of each component, and the last column presents the cumulative change in consumers surplus, measured as before in $/month per simulated consumer.

Looking at the cumulative surplus changes, the results suggest that competition *amplifies* the adverse effect of search frictions on welfare. In other words, by setting search costs to zero, we can improve the welfare of consumers by a larger margin in competitive markets, than in concentrated markets. For instance, the surplus decrease is 76% larger (in absolute value) in markets with 13 lenders than in markets with two.

The decomposition exercise shows that this might not always be true, and depends on the importance of market power. Indeed, as competition increases, the welfare loss contribution from price discrimination shrinks very sharply from 42% to less than 8%. This is a direct consequence of the improvement in consumer bargaining leverage, caused by the increase in the number of competing lenders. However, in our simulations, this improvement comes at the cost of more search effort and a greater misallocation of contracts.

The contribution of the search cost component increases because, in equilibrium, consumers are more likely to search in competitive markets. In principle, the effect of competition on search

Figure 7: Simulated merger effects on search probabilities



is ambiguous. On the one hand, competition increases the benefit of search, and therefore the probability of rejecting the initial offer (holding fixed the initial offer). On the other hand, the initial quote is decreasing in N, which leads to a decrease in the relative gain from search. Our simulation results show that the former effect dominates for most consumers.

Figure 7 illustrates the distribution of the change in search probability associated with each simulated merger. While the effects are heterogenous across consumers, especially in more concentrated market, the vast majority end up searching less after losing a potential lending option. The median effect goes from nearly -6% in duopoly markets, to about -1% in the most competitive markets.[30]

The contribution of the misallocation component also increases sharply with the number of competitors, from 12% to more than 30%. While the fraction of misallocated contracts is not necessarily larger in more competitive markets (since the fraction of searchers is also larger), the potential improvement in transaction surplus is much larger in competitive markets (i.e. the minimum lending cost decreases sharply with N). Therefore, eliminating search frictions corrects a larger market imperfection in markets with 13 lenders, relative to markets with only 2 lenders.

Next, we study directly the effect of concentration on welfare and prices in markets where search costs are present. As before, we can decompose the effect of losing a potential bargaining

[30] In the figure, the boxes correspond to the inter-quartile range of each variable, and whisker lines represent the maximum and minimum observations excluding outliers.

Table 9: Decomposing the effect of mergers on consumer welfare

Mergers	Δ Fraction non-qualify (1)	Welfare change decomposition ($/Month) Misallocation (2)	Market power (3)	Search cost (4)	Δ CS ($/Month) (5)
3 to 2	0.03	-2.18	5.11	-0.72	-6.57
4 to 3	0.02	-2.06	2.19	-0.40	-3.84
5 to 4	0.02	-2.01	1.05	-0.36	-2.70
6 to 5	0.01	-1.95	0.54	-0.38	-2.11
7 to 6	0.01	-1.87	0.28	-0.39	-1.77
8 to 7	0.01	-1.81	0.13	-0.40	-1.54
9 to 8	0.01	-1.73	0.04	-0.39	-1.39
10 to 9	0.01	-1.65	0.01	-0.38	-1.29
11 to 10	0.00	-1.60	-0.03	-0.37	-1.20
12 to 11	0.00	-1.53	-0.04	-0.36	-1.13
13 to 12	0.00	-1.45	-0.03	-0.34	-1.07
Cumulative effect	0.12	-19.83	9.26	-4.49	-24.61

Each entry corresponds to an average over 100 simulated samples. Each sample is equal to 5,000 consumers. The last column reports the cumulate changes in consumer welfare divided by the total number of simulated consumers, including non-qualifying consumers (5,000). The welfare decomposition corresponds to: $\Delta \text{CS}_i = \Delta V_i - \Delta m_i - \Delta \kappa_i S_i = \text{Misallocation}_i - \text{Market power}_i - \text{Search cost}_i$. Columns (2)-(4) report the contribution value of each component (in $/month).

partner on consumer surplus into three terms:

$$\begin{aligned} \Delta \text{CS}_i &= [V_i(N-1) - V_i(N)] - [m_i(N-1) - m_i(N)] - \kappa_i\,[S_i(N-1) - S_i(N)] \\ &= \Delta V_i - \Delta m_i - \Delta \kappa_i S_i. \end{aligned} \tag{23}$$

The change in transaction surplus, ΔV_i, captures both the change in the allocation of consumers across lenders, and the cost increases associated with each merger. The second component measures the change in market power caused by reducing the number of competitors, rather than solely the elimination of the price discrimination opportunity. The last component measures the cost saving that results from consumers searching less in more concentrated markets.

Table 9 shows the results of this decomposition exercise. The first column shows the evolution of the change in the fraction of non-qualifying consumers as we increase the number of competitors. Similar to what we observed when the posted rate was eliminated, increasing the number of competitors leads to a 12% expansion in the number of loans issued, from 84% in duopoly markets to about 96% in the most competitive market. The adverse effect of mergers on the supply of loans is mostly felt in highly concentrated markets.

In terms of welfare, the merger simulation exercise shows that competition increases aggregate consumer surplus. Although the effect of each merger is modest, the difference in consumer surplus between the most competitive market structure (13 lenders) and the least competitive (2) is larger in magnitude than the effect of eliminating search frictions altogether ($24.6 vs $18.9 per month).

The welfare cost of each merger falls as the number of lenders in the market increases. The largest change comes from the market power component: the effect of losing a competitor on profit margins is decreasing in N. The same pattern exists, but is less pronounced, for misallocation. Finally, the offsetting effect of the reduction in search costs is not large enough to compensate for the misallocation and market-power effects of mergers. The reduction in search effort following each merger reduces the cumulative search costs by less than $0.75/month in duopoly markets, and by about $0.35/month in the most competitive markets.

Lastly, we focus our attention on the effect of mergers on the distribution of prices. While Table 9 clearly shows that the market power effect is rapidly decreasing in the number of competitors, it hides the fact that this competitive effect differs enormously across consumers.

The effect of mergers on negotiated prices is heterogenous in part because it is passed through differently to searchers and non-searchers. To see this, note that we can decompose the average change in rates into a change in the competitive quote, a change in the initial quote, and a change in rates for consumers adjusting their search effort post-merger:

$$\begin{aligned} E\left[\Delta r_i\right] &= \underbrace{E\Big[\big(r_i^*(N-1)-r_i^*(N)\big)S_i(N)\Big]}_{\text{Competitive quote (+)}} + \underbrace{E\Big[\big(r_i^0(N-1)-r_i^0(N)\big)(1-S_i(N))\Big]}_{\text{Initial quote (+)}} \\ &\quad + \underbrace{E\Big[\big(r_i^*(N-1)-r_i^0(N-1)\big)\big(S_i(N-1)-S_i(N)\big)\Big]}_{\text{Search (+/–)}}. \end{aligned} \tag{24}$$

The first two terms are positive, while the third can, in theory, be positive or negative depending on how the probability of search changes with N. Although our model predicts that the probability of search can increase or decrease following a merger, our simulations suggest that removing a lender option lowers the average probability of search (see Figure 7).

Table 10 summarizes the results of this decomposition. The average effect of losing a competitor ranges from 9.26 bps in market with three lenders, to 2 bps in markets with 13 lenders. The decomposition reported in columns (4)-(6) shows that the bulk of this price increase (70%) is associated with consumers facing higher prices in the competition stage. In contrast, the contribution of the initial quote is much smaller, and decreases towards zero for mergers in more competitive markets. The fact that consumers adjust their search behavior to engage in less search following the merger implies that the effect on rates is amplified by the presence of search frictions, but I by a small margin.

The small contribution from the initial stage arises mostly because the initial quote is less responsive to changes in market structure. Columns (2) and (3) compare the changes in the competitive and initial quotes for all simulated consumers, irrespective of their search decisions. Sys-

Table 10: Decomposition of merger effects on negotiation rates

Mergers	Change in rates (bps)			Decomposition (%)			Change
	Average	2nd-stage	1st-stage	Comp. quote	Initial quote	Search	dispersion
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
3 to 2	9.26	11.89	5.04	71.71	17.84	10.44	-10.02
4 to 3	6.01	7.73	2.93	72.68	16.18	11.14	-5.81
5 to 4	4.62	5.83	1.86	73.48	13.30	13.22	-4.23
6 to 5	3.84	4.70	1.24	73.31	10.48	16.21	-3.32
7 to 6	3.35	3.98	0.86	73.33	8.10	18.57	-2.86
8 to 7	2.98	3.42	0.60	72.48	6.26	21.26	-2.46
9 to 8	2.72	3.02	0.44	72.64	4.85	22.52	-2.08
10 to 9	2.52	2.71	0.32	71.97	3.77	24.26	-1.99
11 to 10	2.34	2.46	0.24	71.12	2.99	25.89	-1.71
12 to 11	2.18	2.25	0.18	70.95	2.42	26.64	-1.55
13 to 12	2.04	2.08	0.14	70.67	1.97	27.35	-1.29

Each row corresponds to the simulation of a merger from N to $N-1$. Columns (1) to (3) report the average change in transaction rate, and in the second-stage and initial quotes. Columns (4) to (6) decompose the average rate change into the contribution of searchers and non-searchers, and the contribution of consumers changing their search decisions. The last column reports the change residual rate dispersion, measured using the inter-decile range.

tematically, the change in the second stage offer is two to three times larger than the change in the initial quote.

This difference is explained in part by the presence of the posted rate, which constrains roughly 50% of initial quotes, and is independent of market structure. However, even absent the posted rate, the model predicts larger changes in transaction prices at the auction stage following each merger. Therefore, consumers with low search costs benefit more from increased competition, than do consumers with high search costs.

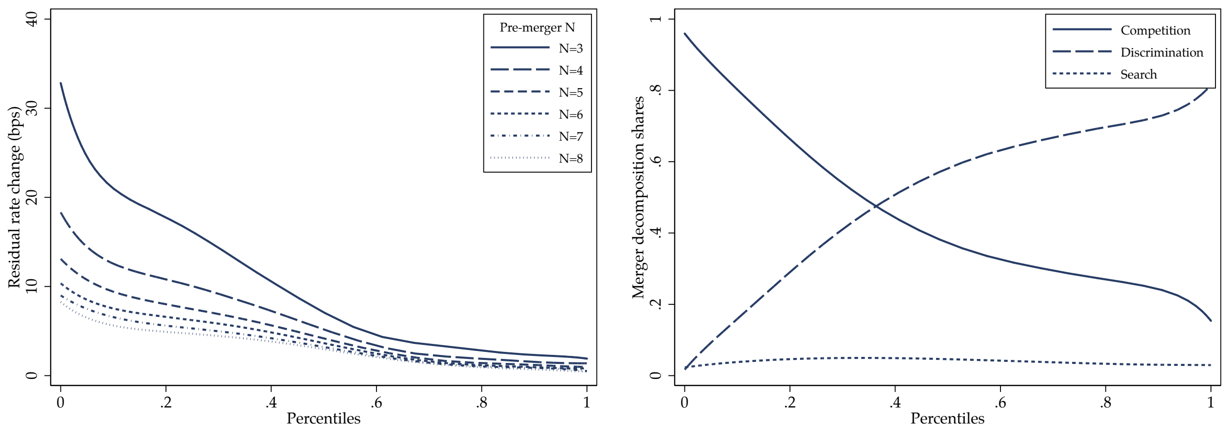
This difference between searchers and non-searchers has implications for the relationship between market structure and price dispersion. In the last column of Table 10 we show that a decrease in the number of competitor systematically leads to a decrease in residual price dispersion. This is because as we increase the number of competitors in a market, the bottom of the price distribution decreases faster than the top, which increases the dispersion of prices.

Figure 8a illustrates this point by plotting the conditional average rate increases following a merger, against the percentiles of the residual rate distribution. Each line represents this relationship for market structures ranging from 3 to 8 lenders. It is clear that it is consumers at the bottom of the rate distribution that are most affected by the loss of a competitor. With 3 lenders, a merger leads to a substantial increase in rates, over 20 bps, at the 10th percentile. Similarly, with 5 competitors, the effect of the merger at the 10th percentile is smaller, about 10 bps, but much larger than at the 90th percentile.

Moreover, as Figure 8b shows, there exists a clear relationship between the percentiles of the

Figure 8: Effect of mergers across the pre-merger distribution of residual rates

(a) Quantile merger effects (b) Decomposition at N=5



residual rate distribution and the contribution of each component. For individuals at the bottom of the residual distribution, i.e. those paying the lowest price, any merger effect is coming from the second-stage. At the top of the residual distribution are individuals who are less likely to search, and therefore most of the merger effect comes through the initial quote. A similar relationship exists for simulated mergers.

The model's predictions about the relationship between market structure and both rates and dispersion are corroborated by findings in Allen et al. (2013a). In it we study the effect of an actual merger between lenders that occurred in the Canadian banking industry. Our results suggest that following the merger there was a small but positive rate increase, but that only consumers at the bottom and middle of the rate distribution were affected. As a consequence, price dispersion decreased following the merger.

6.3 Summary of counter-factual results

Our counter-factual results show that search frictions reduce consumer surplus by almost $20 per month, with the biggest part of this loss associated with the direct cost of searching for multiple quotes, and the remainder with price discrimination and inefficient matching.

Product differentiation increases market power, with loyal consumers paying higher rates. However, overall differentiation attenuates the effect of search frictions by reducing direct search costs and improving allocation: there is a loyalty premium attached to the initial lender, and it makes the first offer. As discussed formally in Weitzman (1979), the sequence of search that we use is optimal only when the home bank offers the higher quality product.

The posted rate also attenuates the welfare cost of search frictions. Its impact comes mostly through its effect on the ability of the home bank to discriminate: in its absence, lenders can increase the initial quote, which can increase the search probability of consumers. This is not to say that the presence of a price ceiling is good overall for consumers.

While the posted price reduces the adverse effect of price discrimination, it also reduces access to credit. We estimate that eliminating the posted rate would increase the number of mortgages issued by 5.4% on average. Furthermore, lowering the posted rate to reflect the average discount in the market would increase the fraction of non-qualifying loan applications to nearly 20%.[31] Therefore, while our model cannot put a number on the value of increasing access to credit, it is important to note that a uniform pricing policy in this market would not necessarily improve consumer welfare.

Our merger simulations reveal that, in contrast to product differentiation and the posted price, competition amplifies the welfare effect of search frictions. As the number of firms in the market increases, the welfare loss from price discrimination shrinks, but the welfare loss from misallocation and direct search costs increases.

We also show that mergers lead to lower search on the part of consumers, and to higher rates. In each case the impact is stronger when the number of banks is larger. We also show that, in terms of welfare change, the impact of moving from duopoly to a market with twelve lenders is similar in magnitude to the impact of removing entirely search frictions. Our findings also show that the effect of competition is not spread equally across all consumers. Specifically, we find consumers with low search costs benefit more from competition, and so eliminating a lender impacts rates paid by consumers at the bottom and middle of the rate distribution, but has no effect on consumers at the top.

These results suggest that price dispersion falls following the mergers, and more so for mergers in less competitive markets. This is consistent with the predictions in Borenstein (1985), and the empirical findings of Borenstein and Rose (1994), which suggest that an increase in competition lowers the prices paid by price-sensitive consumers, while leaving unchanged prices paid by loyal consumers at the top of the price distribution.

7 Conclusion

The paper makes three main contributions. The first is to provide an empirical framework for studying markets in which prices are negotiated. The second is to show that search frictions are important and generate significant welfare losses for consumers that can be decomposed into misallocation, price discrimination, and direct search cost components. We also show that the welfare loss is mitigated by switching costs (loyalty premium) and posted prices, but amplified by competition. The final contribution is to show that the role of competition is also important in

[31] This result is available upon request.

markets with search frictions, but that this effect is not spread equally across consumers.

Although the overall fit of our model is good, the goodness of fit analysis highlights several caveats. First, reduced-form estimates using the data show that loyal consumers pay around 8 bps more, while the model predicts more than 35 bps. This difference is directly related with our modeling assumptions: the timing and order of search are the same for all consumers, and all consumers have a single home bank. These are simplifying assumptions that closely link search and switching in the model.

Similarly, the model tends to over-estimate the impact of competition on rates. This discrepancy likely reflects the fact that we assume that market structure is independent of consumers' unobserved attributes, up to regional fixed-effects. If this assumption is not valid, it would imply that our estimates of firms' cost differences, which determine markup levels, suffer from a attenuation bias, and therefore that our results correspond to a lower bound on the size of profit margins in this market.

A related interpretation of the small reduced-form effect of competition on rates and discounts, is that consumers face heterogenous consideration sets, conditional on being located in the same postal-code area. This would create measurement error in the choice-set of consumers. Because lenders are ex-ante heterogeneous, it is computationally prohibitive to incorporate this type of measurement error in the model. Moreover, we do not have access to data on the set, or identity of lenders considered by borrowers.

Finally, in order to keep the model tractable, we decided to focus only on branch-level transactions, and ignore contracts that are negotiated through brokers. Brokers, act as intermediaries and can potentially lower the search cost of individuals by searching over a larger set of lenders. Since brokers are used by approximately 25% of borrowers it would be important to understand better the role they play in this environment. In an ongoing project we are working on modeling the behavior of these intermediaries.

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 (2013a). The effect of mergers in search markets: Evidence from the canadian mortgage industry. NBER working paper 19126.

Allen, J., R. Clark, and J. F. Houde (2013b). Price dispersion in mortgage markets. Forthcoming *Journal of Industrial Economics*.

Armstrong, M. and J. Zhou (2011). Exploding offers and buy-now discounts.

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.

Bester, H. (1993). Bargaining versus price competition in markets with quality uncertainty. *The American Economic Review 83*(1), 278–288.

Borenstein, S. (1985, Autumn). Price discrimination in free-entry markets. *The Rand Journal of Economics 16*(3), 380–397.

Borenstein, S. and N. Rose (1994). Competition and price dispersion in the u.s. airline industry. *Journal of Political Economy 102*(4), 653–683.

Brannan, L. and L. M. Froeb (2000, May). Mergers, cartels, set-asides, and bidding preferences in asymmetric oral auctions. *The Review of Economic and Statistics 82*(2), 283–290.

Busse, M., A. Israeli, and F. Zettelmeyer (2013, June). Repairing the damage: The effect of price expectations on auto-repair price quotes.

Busse, M., J. Silva-Risso, and F. Zettelmeyer (2006). $1,000 cash back: The pass-through of auto manufacturer promotions. *The American Economic Review 96*(4).

Capps, C., D. Dranove, and M. Satterthwaite (2003). Competition and market power in option demand markets. *RAND Journal of Economics 34*(4), 737–763.

Chatterjee, K. and C. C. Lee (1998). Bargaining and search with incomplete information about outside options. *Games and Economic Bahavior 22*, 203–237.

Crawford, G. and A. Yurukoglu (2011). The welfare effects of bundling in multichannel television.

Dafny, L. (2010). Are health insurance markets competitive? *American Economic Review 100*(4), 1399–1431.

De Los Santos, B., A. Hortacsu, and M. Wildenbeest (2011). Testing models of consumer search using data on web browsing and purchasing behavior. *American Economic Review (Forthcoming)*.

Dey, M. and C. Flynn (2005). An equilibrium model of health insurance provision and wage determination. *Econometrica 73*, 571–627.

Doornik, J. A. (2007). *Ox - An Object-Oriented Matrix Programming Language*. Timberlake Consultants Press and Oxford.

Dranove, D., M. Satterthwaite, and A. Sfekas (2008). Boundedly rational bargaining in option demand markets: An empirical application.

Gavazza, A. (2013, December). An empirical equilibrium model of a decentralized asset market. working paper, LSE.

Goldberg, P. K. (1996). Dealer price discrimination in new car purchases: Evidence from the consumer expenditure survey. *Journal of Political Economy 104*(3), 622–654.

Gowrisankaran, G., A. Nevo, and R. Town (2013). Mergers when prices are negotiated: Evidence from the hospital industry. Mimeo.

Grennan, M. (2013). Price discrimination and bargaining: Empirical evidence from medical devices. *American Economic Review 103*(1), 145–177.

Hall, R. E. and S. Woodward (2010). Diagnosing consumer confusion and sub-optimal shopping effort: Theory and mortgage-market evidence. NBER working paper 16007.

Hendel, I., A. Nevo, and F. Ortalo-Magne (2010). The relative performance of real estate marketing platforms: Mls versus fsbomadison.com. *The American Economic Review*.

Hong, H. and M. Shum (2006). Using price distributions to estimate search costs. *RAND Journal of Economics 37*(2), 257–275.

Honka, E. (2012, January). Quantifying search and switching costs in the u.s. auto insurance industry. Working paper, University of Texas-Dallas.

Hortaçsu, A. and C. Syverson (2004). Product differentiation, search costs, and competition in the mututal fund industry: A case study of S&P 500 index funds. *Quarterly Journal of Economics*, 403–456.

Janssen, M. and J. Moraga-Gonzalez (2004). Strategic pricing, consumer search and the number of firms. *Review of Economic Studies 71*, 1089–1118.

Merlo, A. and F. Ortalo-Magne (2004). Bargaining over residential real estate: evidence from england. *Journal of Urban Economics 56*(2), 192–216.

Postel-Vinay, F. and J.-M. Robin (2002). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica 70*(2), 2295–2350.

Sallee, J. (2011, May). The surprising incidence of tax credits for the toyota prius. *American Economic Journal: Economic Policy 3*.

Scott-Morton, F., F. Zettelmeyer, and J. Silva-Risso (2001). Internet car retailing. *Journal of Industrial Economics 49*(4), 501–519.

Scott Morton, F., F. Zettelmeyer, and J. Silva-Risso (2003). Consumer information and discrimination: Does the internet affect the pricing of new cars to women and minorities? *Quantitative Marketing and Economics 1*, 65–92.

Skrainka, B. S. and K. Judd (2011). High performance quadrature rules: How numerical integration affects a popular model of product differentiation. working paper, Stanford university.

Sorensen, A. (2001, October). An empirical model of heterogeneous consumer search for retail prescription drugs. Working Paper 8548, NBER.

Town, R. and G. Vistnes (2001). Competition in networks: An analysis of hospital pricing behavior. *Journal of Health Economics 20*, 719–753.

Weitzman, M. (1979, May). Optimal searh for the best alternative. *Econometrica 47*(3), 641–654.

Wildenbeest, M. R. (2011). An empirical model of search with vertically differentiated products. *Rand Journal of Economics 42*(4). manuscript.

Appendix A: Robustness

Table 11: Summary statistics on mortgage contracts in the selected sample

variable	N	mean	sd	p25	p50	p75
loan	35,457	140,015	56,606	94,257	131,846	177,548
income	35,457	69,535	27,630	49,946	65,292	83,232
payment	35,457	974	387	665	920	1223
spread	35,457	1.26	.63	.82	1.22	1.69
I(no discount)	35,457	22.6	41.8	0	0	0
switch	22,815	26.7	24.2	0	1	1
credit score	35,457	668	72.1	650	700	750
I(LTV=95)	35,457	36.9	48.2	0	0	1
previous owner	35,457	24.3	42.9	0	0	0

Table 1 provides summary statistics for the main data-set, which is based only on contracts insured by CMHC. For robustness we also include estimate the model using contracts insured by Genworth Financial, even though there are more missing observations. This table provides summary statistics of the full sample.

Table 12: MLE estimation results for alternative specifications[32]

Parameters	(1) No Heterogeneity	(2) W/ Genworth	(3) $\omega = 100$
Common shock (σ_ε)	0.288	0.290	0.247
	(0.002)	(0.002)	(0.002)
Idiosyncratic shock (σ_u)	0.124	0.156	0.155
	(0.002)	(0.002)	(0.003)
Avg. search cost:			
$\bar{\kappa}_0$	-1.080	-1.660	-1.275
	(0.013)	(0.028)	(0.016)
$\bar{\kappa}_{\text{inc}}$		0.576	0.143
		(0.039)	(0.018)
$\bar{\kappa}_{\text{owner}}$		0.326	0.820
		(0.043)	(0.013)
Loyalty premium:			
λ_0	-1.780	-1.973	-1.822
	(0.011)	(0.008)	(0.005)
λ_{inc}		0.692	0.670
		(0.004)	(0.002)
λ_{owner}		0.020	0.260
		(0.003)	(0.002)
Measurement error:	0.936	0.941	0.886
	(0.004)	(0.005)	(0.004)
Cost function:			

Intercept	3.510	3.871	3.430
	(0.063)	(0.247)	(0.043)
Bond rate	0.610	0.580	0.629
	(0.009)	(0.039)	(0.006)
Loan size	0.035	0.083	0.077
	(0.012)	(0.015)	(0.015)
Income	-0.024	-0.214	-0.098
	(0.025)	(0.030)	(0.028)
Loan/Income	-0.078	-0.109	-0.077
	(0.009)	(0.012)	(0.010)
Other debt	-0.054	-0.046	-0.043
	(0.006)	(0.007)	(0.005)
FICO score	-0.501	-0.518	-0.463
	(0.029)	(0.033)	(0.029)
Max. LTV	0.060	0.060	0.053
	(0.004)	(0.005)	(0.004)
Previous owner	0.017	-0.008	-0.093
	(0.005)	(0.006)	(0.005)
Number of parameters	43	47	47
LLF/10,000	-4.062	-4.279	-5.037
Likelihood-ratio test: $2 \times (\mathcal{L}_{\text{base}} - \mathcal{L}_0)$	943.371	5274.540	20437.486
Sample size	29,000	35,457	29,000

[32] Average search cost function: $\log(\bar{\kappa}_i) = \kappa_0 + \kappa_{\text{inc}}\text{Income}_i + \kappa_{\text{owner}}\text{Previous owner}_i$. Home bank premium function: $\log(\lambda_i) = \lambda_0 + \lambda_{\text{inc}}\text{Income}_i + \lambda_{\text{owner}}\text{Previous owner}_i$. Cost function: $C_i = L_i \times (Z_i\beta + \varepsilon_i - u_i)$. Units: \$/100. All specifications include year, market and bank fixed-effects. The likelihood ration test is calculated relative to the baseline specification presented in Table 3.

Appendix B: Data description

Table 13: 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.