

Search Frictions and Market Power in Negotiated-Price Markets

Jason Allen

Bank of Canada

Robert Clark

Queen's University

Jean-François Houde

University of Wisconsin–Madison and National Bureau of Economic Research

We provide a framework for empirical analysis of negotiated-price markets. Using mortgage market data and a search and negotiation model, we characterize the welfare impact of search frictions and quantify the role of search costs and brand loyalty for market power. Search frictions reduce consumer surplus by \$12/month/consumer, 28 percent of which can be associated with discrimination, 22 percent with inefficient matching, and 50 percent with search costs. Banks with large consumer bases have margins 70 percent higher than those with small consumer bases. The main source of this incumbency advantage is brand loyalty; however, price discrimination based on search frictions accounts for almost a third.

I. Introduction

In a large number of markets, sellers post prices, but actual transaction prices are achieved via bilateral bargaining. This is the case, for instance,

This research has benefited from the financial support of the National Science Foundation (SES-1024840). We thank the Canada Mortgage and Housing Corporation and Gen-

Electronically published June 24, 2019

[*Journal of Political Economy*, 2019, vol. 127, no. 4]

© 2019 by The University of Chicago. All rights reserved. 0022-3808/2019/12704-0002\$10.00

in the markets for new or used cars (Goldberg 1996; Scott-Morton, Zettelmeyer, and Silva-Risso 2001; Busse, Silva-Risso, and Zettelmeyer 2006), health insurance (Dafny 2010), capital assets (Gavazza 2016), and financial products (Woodward and Hall 2012; Allen, Clark, and Houde 2014a), as well as for most business-to-business transactions (e.g., Joskow 1987; Town and Vistnes 2001; Salz 2015).

In this paper, we are interested in two key features characterizing many of these markets. First, since buyers incur a cost to gather price quotes, these markets are characterized by important search frictions. Second, the repeated relationship that develops between a buyer and a seller creates a loyalty advantage, which increases the value of transacting with the same seller. The reasons can be switching costs associated with changing suppliers, cost advantages of the incumbent sellers, or complementarities from the sale of related products.

Search frictions and brand loyalty have implications for market power. Search costs open the door to price discrimination: the seller offering the first quote is in a quasi monopoly position and can make relatively high offers to consumers with poor outside options and/or high expected search costs. Brand loyalty reduces the bargaining leverage of consumers, because incumbent sellers provide higher value, which creates a form of lock-in. Together, these features imply that a firm with an extensive consumer base has an incumbency advantage over rival firms in the same market.

We study one particular negotiated-price setting, the Canadian mortgage market, for which we have access to an administrative data set on a large number of individually negotiated mortgage contracts, which we use to estimate a model of search and price negotiation. In this market, in contrast to the United States, the vast majority of mortgages are originated by national lenders that are vertically integrated along the supply chain: from origination to servicing and financing. These lenders post common interest rates, but in-branch loan officers have considerable freedom to negotiate directly with borrowers. Importantly, there is evidence of search frictions and brand loyalty in this setting. According to market research firm Ipsos-Reid, about 70 percent of consumers in this market combine day-to-day banking and mortgage services at their main financial institution. Our transaction-level data suggest that when originating new mortgages, roughly 80 percent of consumers get a rate quote from this

worth 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 Michael Grubb, Ken Hendricks, Ali Hortaçsu, Matt Lewis, Francesca Molinari, Jack Porter, Dan Quint, Jacques Robert, Marc Rysman, Nicolas Sahuguet, 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. This paper has previously circulated under the title "Price Negotiation in Differentiated Product Markets: The Case of Insured Mortgages in Canada." Instructions for accessing the data are provided as supplementary material online.

lender. Moreover, despite the fact that approximately 60 percent of consumers admit in surveys to searching for additional quotes, less than 30 percent obtain a mortgage from a lender other than their main institution.

In this setting, we consider two questions. First, what is the impact of search frictions on consumer welfare? Second, what are the sources and magnitude of market power? We focus on quantifying the incumbency advantage that stems from having a large consumer base and decomposing it into two parts: (i) a first-mover advantage arising from price discrimination and search frictions and (ii) a loyalty advantage originating from long-term relationships.

To address these questions, we estimate a structural model of demand and supply applied to negotiated-price markets. Our contribution is to develop a framework that accounts for the fact that in these settings buyers negotiate prices with potentially many differentiated sellers but often sign exclusive contracts with just one. As a result, researchers typically observe only transaction prices and the identity of sellers, including whether or not buyers remain loyal to their current supplier. This poses a serious challenge for empirical work, since buyers' outside options are unobserved.¹ In our case study, we do not observe rejected offers or whether consumers search for more than one lender. This is not unique to mortgage lending; most data sets used in previous empirical work on price negotiation in consumer goods and business-to-business markets share these same features.²

To overcome these challenges, we develop a two-stage game of bargaining and search related to the models advanced by Wolinsky (1987), Bester (1988), and Chatterjee and Lee (1998). Like Chatterjee and Lee's model, our setup involves one-sided incomplete information in which the uninformed party (the home bank) makes a take-it-or-leave-it offer to the consumer, who then decides, on the basis of expected net gain from searching, whether or not to gather additional quotes. The home bank uses its initial quote to price discriminate by screening high-search cost consumers. If the initial quote is rejected, consumers pay a search cost, and local lenders compete via an English auction for the contract. Using an auction to characterize the competition stage represents a tractable approach to address the missing prices problem. In particular, it accounts for the

¹ As a result we cannot use recent approaches based on the simultaneous complete-information multilateral negotiation game proposed by Horn and Wolinsky (1988) that have modeled the outside option as observed prices paid by a given buyer to alternative suppliers (Crawford and Yurukoglu 2012; Grennan 2013; Gowrisankaran, Nevo, and Town 2015; Lewis and Pflum 2015; Ho and Lee 2017).

² For instance, the data set used by Goldberg (1996) contains information on the price the consumer paid, the brand of the purchased vehicle, and whether the consumer previously bought the same brand. Cicala (2015) has data on coal deliveries to power plants and transaction prices, while Salz (2015) has information on the contract terms between businesses and waste carters. Their data sets allow them to measure the duration of contractual relationships with incumbent suppliers. See also Jindal and Newberry (2018) for home appliances.

fact that sellers can counter rivals' offers by lowering prices, a process that mimics an English auction. The expected outcome of the auction, net of the search cost, determines the bargaining leverage of the consumer. Since consumers' outside options are privately observed, the model implies that search occurs in equilibrium, a feature that we observe in the data.

The tractability of the model also allows us to analyze identification of the parameters in a transparent way. We describe conditions under which the search and lending cost distributions are nonparametrically identified, using insights from the labor, discrete-choice, and empirical auctions literatures. The identification argument is generalizable to other negotiated-price settings in which researchers have access to data on transaction prices and switching decisions (but not necessarily search). Although the search and lending cost distributions in theory could be nonparametrically estimated, we instead estimate a parametric version of the model using maximum likelihood. This allows us to more easily incorporate observable differences between consumers and firms.

The results can be summarized as follows. We find that firms face relatively homogeneous lending costs for the same borrower. In contrast, we find that borrowers face significant search costs and a brand loyalty advantage. On average, consumers in our sample face an up-front search cost of \$1,150. In addition, the incumbent bank has an average cost advantage of \$17.10 per month (for a \$100,000 loan) generating a sizable loyalty advantage.

We use the model estimates to characterize the impact of search frictions on consumer welfare and to measure market power. To quantify the welfare cost of search frictions, we perform a set of counterfactual experiments in which we eliminate the search costs of consumers. The surplus loss from search frictions originates from three sources: (i) misallocation of buyers and sellers, (ii) price discrimination, and (iii) the direct cost of gathering multiple quotes. Our results suggest that search frictions reduce average consumer surplus by almost \$12 per month, over a 5-year period. Approximately 28 percent of the loss in consumer surplus comes from the ability of incumbent banks to price discriminate with their initial quote. A further 22 percent is associated with the misallocation of contracts and 50 percent with the direct cost of searching. We also find that the presence of a posted rate limits the ability of firms to price discriminate, thereby reducing the welfare cost of search frictions. Competition also amplifies the adverse effects of search frictions on consumer welfare.

Our results also suggest that the market is fairly competitive. The average profit margin is estimated to be just over 20 basis points (bps), which corresponds to a Lerner index of 3.2 percent. However, margins vary considerably depending on whether consumers search and/or switch. On average, firms charge a markup that is 90 percent higher for consumers who are not searching. Banks' profits from switching consumers are

\$14.99 per month (17.1 bps) compared to \$20.22 per month from loyal consumers (24.6 bps).

The increased profits earned from loyal consumers correspond to the incumbency advantage and are directly related to the size of the bank's consumer base. To measure the source and magnitude of the advantage we use the simulated model to evaluate the correlation between consumer base and profitability. We find that banks with the largest consumer bases earn, on average, 62 percent of the profits generated in their markets, compared to only 2 percent for those with the smallest. This difference is driven by the fact that lenders with large consumer bases control a large share of the mortgage market and earn significantly more profit per contract than smaller banks.

We measure the incumbency advantage as the increased market power of banks with large consumer bases relative to those with the smallest. Our estimates suggest that banks with large consumer bases have margins that are 70 percent higher than those with small consumer bases. To identify the importance of the two sources of the incumbency advantage we simulate a series of counterfactual experiments aimed at varying the first-mover advantage and the differentiation component independently. Our results suggest that about 50 percent of the incumbency advantage can be directly attributed to brand loyalty, 30 percent to search frictions, and the remaining 20 percent to their interaction.

Our paper is related to three strands of literature. First are the recent empirical papers based on the complete-information multilateral negotiation game proposed by Horn and Wolinsky (1988) mentioned above (see, e.g., Grennan 2013; Gowrisankaran et al. 2015). This method for measuring the buyers' outside options is suitable for the case of bargaining between buyers and their network of suppliers but is not applicable when buyers transact with a single seller. Our paper is also related to the industrial organization search literature (see Sorensen 2001; Hortaçsu and Syverson 2004; Hong and Shum 2006; Wildenbeest 2011; De los Santos, Hortaçsu, and Wildenbeest 2012; Honka 2014; Marshall 2016; Alexandrov and Koulayev 2017). Although these papers take into account concentration and differentiation, they have mostly focused on cases in which firms offer random posted prices to consumers irrespective of their characteristics (as opposed to targeted negotiated-price offers). Finally, our findings contribute to the literature on the advantages accruing to incumbent firms from demand inertia and brand loyalty. Bronnenberg and Dubé (2017) provide an extensive survey of this literature in industrial organization and marketing. Our model allows us to quantify the relative importance of two sources of state dependence and market power associated with the incumbency advantage.

The paper is organized as follows. Section II presents details on the Canadian mortgage market and introduces our data sets. Section III pre-

sents the model, and Section IV discusses conditions for nonparametric identification of the primitives. Section V discusses the estimation strategy, and Section VI describes the empirical results. Section VII analyzes the impact of search friction and brand loyalty on consumer welfare and market power. Finally, Section VIII presents conclusions. Additional information on the data, proofs, and results can be found in the online appendixes.

II. Institutional Details and Data

A. *Institutional Details*

The Canadian mortgage market is 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 percent of banking industry assets. For convenience we label these institutions the "Big 8."

Canada features two types of mortgage contracts: conventional, which are uninsured since they have a low loan-to-value ratio, and high loan-to-value, which require insurance (for the lifetime of the mortgage). Most 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.³ CMHC's market share during our sample period averages around 80 percent. Both insurers use the same insurance guidelines and charge lenders an insurance premium, ranging from 1.75 percent to 3.75 percent of the value of the loan, which is passed on to borrowers.⁴

The large Canadian banks operate nationally, and their head offices post prices that are common across the country on a weekly basis in both national and local newspapers, as well as online. Throughout our entire sample period the posted rate is nearly always common across lenders and represents a ceiling in the negotiation with borrowers.⁵

According to data collected by marketing firm Ipsos-Reid, the majority of Canadians have a main financial institution in which they combine checking and mortgage accounts. Therefore, potential borrowers can accept to pay the rate posted by their home bank or search for and negotiate over rates. Borrowers bargain directly with local branch managers or

³ A private firm, Genworth Financial, also provided insurance in this period and had a 90 percent government guarantee.

⁴ Appendix A describes the insurance rules and defines all of the variables included in the data set.

⁵ In Canada pricing over the posted rate is illegal. A similar setup is implied in other retail markets featuring negotiation in the presence of manufacturers' suggested retail prices.

hire a broker to search on their behalf.⁶ Our model excludes broker transactions and focuses only on branch-level transactions.

B. Mortgage Data

Our main data set is a 10 percent random sample of insured contracts from the CMHC, from January 1999 to October 2002. The data set contains information on the terms of the contract (transaction rate, loan size, and house price), as well as financial and demographic characteristics of the borrower. In the empirical analysis we focus on borrower income, FICO risk score, the loan-to-value ratio, and the 5-year bond rate valid at the time of negotiation. In addition, we observe the closing date of the contract and the location of the purchased house up to the forward sortation area (FSA).⁷

The data set contains lender information for the Big 8, a large trust company (Canada Trust), and three small regional lenders (Vancity, Manulife, and Canada Western Trust). Mortgage contracts for which we do not have a lender name but only a type are coded as “other credit union,” “other trusts,” and “other bank.” “Other bank” includes mostly two institutions: Laurentian Bank and HSBC. The former is present only in Québec and Eastern Ontario and the latter mostly in British Columbia and Ontario. We exploit this geographic segmentation and assign “other banks” customers to HSBC or Laurentian on the basis of their relative presence in the local market around each home location. The credit union and trust categories are fragmented and contain mostly regional financial institutions. We therefore combine both along with the three smaller regional lenders into a single “other lender” category. Overall, therefore, consumers face 12 lending options.

We restrict our sample to contracts with homogeneous terms. 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), (iv) contracts that were negotiated individually (i.e., without a broker), and (v) contracts without missing values for key attributes (e.g., credit score, broker, and residential status).

The final sample includes around 26,000 observations, or about one-third of the initial sample. Approximately 18 percent of the initial sample contained missing characteristics: either risk type or business originator (i.e., branch or broker). The reason is that CMHC started collecting these

⁶ 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 are compensated by lenders.

⁷ The FSA is the first half of a postal code. We observe nearly 1,300 FSAs in the sample. While the average FSA has a radius of 7.6 km, the median is 2.6 km.

transaction characteristics systematically only in the second half of 1999. We also drop broker transactions (28 percent) as well as short-term, variable rate, and refinanced contracts (40 percent).

We use the data to construct three main outcome variables: (i) monthly payment, (ii) negotiated discounts, and (iii) loyalty. The monthly payment, denoted by p_i , is constructed using the transaction interest rate, the loan size, and the amortization period (60 months) specified in borrower i 's contract. To construct negotiated discounts, we must first identify the posted rate valid at the time of negotiation. Since our contract data include only the closing date, to pin down the appropriate posted rate we must infer the negotiation week. To do so we identify the length of time ahead of closing that minimizes the aggregate fraction of consumers paying above the posted rate. This turns out to be 33 days prior to closing. Finally, the loyalty variable is a dummy variable equal to one if a consumer has prior experience dealing with the chosen lender. Since 75 percent of consumers are new home buyers, this most likely identifies the bank with which the borrowers have a savings or checking account. Note that this variable is not available for one lender, and we therefore treat the loyalty outcome as partly missing when constructing the likelihood function.

Finally, since the main data set does not provide direct information on the number of quotes gathered by borrowers, we supplement it with survey evidence from the Altus Group (FIRM survey). The survey asks 841 people who purchased a house during our sample period about their shopping habits. We use the aggregate results of this survey to construct auxiliary moments characterizing the fraction of consumers who report searching for more than one lender, by demographic groups. We focus in particular on city size, regions, and income groups.

C. Market Structure Data

The market structure is described by the consumer base of each bank and the number of lenders available in consumers' choice sets. The consumer base of a lender is defined by its share of the market for day-to-day banking services. In the model, this is used to approximate the fraction of consumers in a given market that have prior experience with each potential lender. To construct this variable, we use microdata from a representative survey conducted by Ipsos-Reid (Consumer Finance Monitor, 1999–2002). Each year, Ipsos-Reid surveys nearly 12,000 households in all regions of the country. We group the data into year (four), region (10), and income (four) categories. Within these subsamples we estimate the probability of consumers choosing one of the 12 largest lenders as their main financial institution or home bank denoted by h . We use $\psi_h(x_i)$ to denote the probability that a consumer with characteristics x_i has prior experience with bank h .

A consumer's choice set is defined by the location of the house being purchased. We assume that consumers have access to lenders that have a branch located within 10 km of the centroid of their FSAs.⁸ This choice is justified by the data: over 90 percent of loans are originated by a lender present within 10 km of each FSA. In addition, the fact that rates are negotiated directly with loan officers limits the ability of consumers to perform the transaction online. Indeed, CMHC reports that less than 2 percent of mortgages are originated through the internet or phone.

The location of each financial institution's branches is available annually from Micromedia-ProQuest. We use this data set to match the new house location with branch locations and construct each consumer's choice set. Formally, a lender is part of consumer i 's choice set if it has a branch located within 10 km of the house location. We use \mathcal{N}_i to denote the set of rival lenders available to consumer i (excluding the home bank), while n_i is the number of banks in \mathcal{N}_i .

D. Market Features

Before introducing the model, we provide descriptive evidence outlining the key features of the Canadian mortgage market that we want to capture. Table 1 describes the main financial and demographic characteristics of the borrowers in our sample. Table 2 reports a subset of the coefficients of two reduced-form regressions describing the relationship between transaction characteristics and negotiation rates, as well as the probability of remaining loyal to the home bank.

The estimation sample corresponds to a fairly symmetric distribution of income and loan size. The average loan size is about \$136,000, which is twice the average annual household income. The loan-to-value (LTV) variable shows that many consumers are constrained by the minimum down payment of 5 percent imposed by the government guidelines. Nearly 40 percent of households invest the minimum. Our focus is on the monthly payment made by a borrower, and so when we talk about quotes and rates, they will be based on a given monthly payment. The average monthly payment made by borrowers in our sample is \$925.

In what follows we present five key features that characterize shopping behavior and outcomes in the Canadian mortgage market and most negotiated-price markets.

Feature 1: Mortgage transaction rates are dispersed.

⁸ Our results are robust to alternative neighborhood size definitions. We also considered a 5 km neighborhood, since this captures the fact that the average distance to chosen lenders is about 2 km, compared to slightly less than 4 km for the average distance to other financial institutions. Results using this definition are available on request.

TABLE 1
SUMMARY STATISTICS ON MORTGAGE CONTRACTS AND LOYALTY IN THE SELECTED SAMPLE

Variable	Mean (1)	Standard Deviation (2)	P25 (3)	P50 (4)	P75 (5)
Interest rate spread	120	59.3	81	115	161
Positive discounts	95.3	45.4	70	95	125
1(Discount = 0)	.127	.333	0	0	0
Monthly payment	925	385	619	858	1,169
Total loan (\$100,000)	136	57.6	90.4	126	174
Income (\$100,000)	68.4	27.9	48.5	64.1	82.1
FICO score	669	74	650	700	750
LTV	91	4.38	89.7	90	95
1(LTV = max)	.385	.487	0	0	1
1(Previous owner)	.251	.433	0	0	1
1(Loyal)	.737	.440	0	1	1
Number of lenders	8.65	1.44	8	9	10
Relative network	1.6	1.02	.989	1.37	1.93

NOTE.—Sample size = 26,218. Number of missing loyal observations = 5,599. The sample covers the period from January 1999 to October 2002. We trim the top and bottom 0.5 percent of observations in terms of income and loan size. Interest rates and discounts are expressed in percentage basis points (bps). The number of lenders is within 10 km of the borrowers' new home (neighborhood). Relative network is defined as the average network size of the chosen institution relative to the average size of others present in the same neighborhood.

There is little within-week dispersion in posted prices, especially among the big banks, where the coefficient of variation on posted rates is very close to zero. In contrast, the coefficient of variation on transaction rates is 50 percent, and there is substantial residual dispersion as illustrated by the R^2 of .61 in table 2. See Allen, Clark, and Houde (2014b) for more details.

Feature 2: Consumers who are loyal and are located in concentrated markets tend to pay higher rates.

The rate regression shows that clients who remain loyal to their home bank receive discounts that are about 9.7 bps smaller than do new clients. It also shows that discounts are increasing in the number of local lenders and decreasing in relative network size.

Feature 3: Consumers search more than they switch.

The search and negotiation process typically begins with the consumer's main financial institution: about 80 percent of consumers get a quote from their main institution (see Allen et al. 2014a). A little over 60 percent of consumers search, but only about 26 percent switch away from their main institution.

Feature 4: Consumers are more loyal in concentrated markets and to banks with larger branch networks.

The loyalty regression shows that the likelihood of remaining loyal is decreasing in the number of lenders present in the market and increasing in relative network size.

TABLE 2
MORTGAGE RATES AND LOYALTY IN THE SELECTED SAMPLE:
REDUCED-FORM REGRESSION

Variable	Rate (1)	1 (Loyal) (2)
1 (Loyal)	.097*** (.0079)	
Previous owner	.025*** (.0084)	.11*** (.0072)
Relative network	.023*** (.0046)	.026*** (.0045)
# lenders (log)	-.13*** (.022)	-.076*** (.019)
Observations	20,619	20,619
R^2	.612	.095
Marginal effect: income	.29	.18
Marginal effect: loan	.47	-.19

NOTE.—Number of missing loyal observations = 5,599. The sample covers the period from January 1999 to October 2002. We trim the top and bottom 0.5 percent of observations in terms of income and loan size. Interest rates and discounts are expressed in percentage basis points (bps). The number of lenders is within 10 km of the borrowers' new home (neighborhood). Relative network is defined as the average network size of the chosen institution relative to the average size of others present in the same neighborhood. Each regression also includes market and quarter/year fixed effects and other financial characteristics, i.e., posted rate, bond rate, FICO score, LTV, 1 (LTV = max), loan size, income, and loan/income. Robust standard errors are in parentheses.

* $p < .1$.

** $p < .05$.

*** $p < .01$.

Feature 5: Lenders with strong retail presence have larger market shares.

On average, consumers face 8.6 lenders within their neighborhood. Consumers tend to choose lenders with large branch networks, transacting with lenders that are nearly 60 percent larger than their competitors in terms of branches. Lenders with larger branch networks also tend to have a bigger share of the day-to-day banking market, generating a link between day-to-day market share and mortgage market share that provides large banks with an incumbency advantage.

III. Model

In this section we build a model that captures the five key features just mentioned. Consumers receive an initial quote from their main financial institution and then decide whether to accept or reject this quote on the basis of their heterogeneous search costs and their expected gain from gathering multiple quotes, which depends, among other things, on how competitive is their local market.

In addition to capturing these features, the model takes into account the fact that, during negotiation, loan officers can lower previously made

offers in an effort to attract or retain potential clients. Furthermore, competition takes place locally between managers of competing banks, since consumers must contact loan officers directly to obtain discounts. We also suppose that branches that are part of the same network do not compete for the same borrowers, a feature of the Canadian mortgage market and of some, but not all, negotiated-price markets.

The next three subsections describe the model. First, we present preferences and cost functions and the bargaining protocol. Then we solve the model backward, starting with the second stage of the game in which banks compete for consumers. Finally, we describe the consumer search decision and the process generating the initial quote. All variables introduced in the model vary at the consumer level, i , on the basis of observed or unobserved characteristics. To simplify notation we omit the borrower's index i and will add it back in the next section for random variables and consumer characteristics.

A. Preferences and Cost Functions

Consumers solve a discrete-choice problem over which lender to use to finance their mortgage:

$$\max_{j \in \mathcal{J}} v_j - p_j, \quad (1)$$

where \mathcal{J} is the set of lenders offering a quote, p_j denotes the monthly payment offered by lender j , and v_j denotes the maximum willingness to pay (WTP) associated with bank j .

The choice set \mathcal{J} is defined both by where consumers live and by their search decision. Consumers can obtain a quote from their home bank (h) and from the n lenders in \mathcal{N} . We assume that the cost of obtaining a quote from the home bank is zero, while the cost of getting additional quotes is $\kappa > 0$. This search cost does not depend on the number of quotes and is distributed in the population according to the cumulative distribution function (CDF) $H(\cdot)$.

The WTP of consumers is a combination of differentiation and mortgage valuation:

$$v_j = \begin{cases} \bar{v} + \lambda & \text{if } j = h \\ \bar{v} & \text{else.} \end{cases}$$

The valuation for a mortgage, \bar{v} , is common across all lenders. Throughout we assume that it is large enough not to affect the set of consumers present in our sample. The parameter $\lambda \geq 0$ measures consumers' WTP for their home bank relative to other lenders.

We also assume that banks have a constant borrower-specific marginal cost of lending. This measures the direct lending costs for the bank (i.e.,

default and prepayment risks), net of the future benefits associated with selling complementary services to the borrower.⁹ Since we do not observe the performance of the contract along the risk and complementarity dimensions, we use a reduced-form function to approximate the net present value of the contract. The monthly cost for bank j to lend to the consumer is

$$c_j = \begin{cases} c - \Delta & \text{if } j = h \\ c + \omega_j & \text{if } j \neq h, \end{cases} \quad (2)$$

where c is the common cost of lending to the consumer; ω_j is the cost differential of lender j relative to the home bank (or its match value); and Δ is the home bank's *cost advantage*. This advantage arises because of the multi-product nature of financial institutions and the fact that the home bank is potentially already selling profitable products to the consumer.¹⁰ It could come from real complementarities generated by bundling products (economies of scope) or from the fact that costs include not just the direct cost of mortgage lending but also revenues/costs derived from the sale of additional products.¹¹ In contrast to the home bank, competing lenders may need to offer discounts on these products to overcome the switching costs or may not earn any revenues at all from them if consumers do not switch.

As we will see below, the importance of brand loyalty in the market is driven by the sum of the cost and WTP advantages of the home bank: $\gamma = \Delta + \lambda$. We refer to γ as the home bank *loyalty advantage*.

The idiosyncratic component, ω_j , is distributed according to $G(\cdot)$, with $E(\omega_j) = 0$. We use subscript (k) to denote the k th-lowest cost match value among the non-home bank lenders. The CDF of the k th-order statistic among the n lenders is given by $G_{(k)}(w|n) = \Pr(\omega_{(k)} < w|n)$.

Finally, lenders' quotes are constrained by a common posted price \bar{p} .¹² The posted price determines both the reservation price of consumers (i.e., $\bar{v} > \bar{p}$) and whether or not consumers qualify for a loan at a given lender (i.e., $\bar{p} > c_j$).

⁹ While lenders are fully insured against default risk, the event of default implies additional transaction costs to lenders that lower the value of lending to risky borrowers. Prepayment risk is perhaps more relevant in our context, since consumers are allowed to reimburse up to 20 percent of their mortgage every year without penalty.

¹⁰ For instance, banks also offer credit cards, which have a 50–60 percent return on equity (ROE), compared to Canadian banks' overall ROE of 16 percent.

¹¹ Note that we rule out the possibility that the incumbent bank has more information than other lenders, since otherwise the problem would involve adverse selection, and the initial quote would be much more complicated. For a discussion about competition when one firm has more information about consumers learned from their past purchases, see Fudenberg and Villas-Boas (2007). A subset of this literature has focused on credit markets and the extent to which lenders can learn about the ability of their borrowers to repay loans and use this information in their future credit decisions and pricing. See, e.g., Dell'Ariccia, Friedman, and Marquez (1999) and Petriconi (2015).

¹² Since there is almost no dispersion in posted prices, we assume that every lender has the same posted rate.

B. *Bargaining Protocol, Information, and Timing of the Game*

In an initial period outside the model, consumers choose the type of house they want to buy, the loan size, L , and the timing of the home purchase (including closing date). Our focus is on the negotiation process, which we model as a two-stage game. In the first stage, the home bank makes an initial offer p^0 . At this point, the borrower can accept the offer or search for additional quotes by paying the search cost κ . If the initial quote is rejected, the borrower organizes an English auction among the home bank and the n other banks present in his or her neighborhood. The lender choice maximizes the utility of consumers, as in equation (1).

Information about costs and preferences is revealed sequentially. At the initial stage, all parties observe the posted price \bar{p} , the number of rival banks n , the common component of the lending cost c , and the home bank cost and WTP advantages (λ, Δ) . These variables define the observed state vector: $s = (c, \lambda, \Delta, \bar{p}, n)$. This information is common to all players. The search cost is privately observed by consumers. The home bank knows only the distribution, which can vary across consumers on the basis of observed demographic attributes. Finally, in the second stage of the game, each lender learns its idiosyncratic lending cost, ω_j .

Before solving the game, two remarks are in order. First, consumers are price takers in the model, and so lenders have full bargaining power. This does not mean, however, that consumers have no bargaining leverage, since they have an informational advantage from knowing their search cost. This prevents the home bank from extracting the entire surplus of consumers, as in Allen et al. (2014a).¹³ Second, consumers are assumed to pay the cost of generating offers at the auction stage (rather than firms). Therefore, banks that are not competitive relative to the home bank are, in theory, indifferent between submitting and not submitting a quote. In these cases we assume that banks always submit a truthful offer that is consistent with their realized match values.

Next, we describe the solution of the negotiation by backward induction, starting with the competition stage.

C. *Competition Stage*

Conditional on rejecting p^0 , the home bank competes with lenders in the borrower's choice set. We model competition as an English auction with

¹³ Beckert, Smith, and Takahashi (2016) take a different approach by assuming that consumers and firms split the known surplus from the auction using a Nash bargaining protocol. In this context, the relative bargaining power of consumers, instead of the search cost distribution, determines the split of the surplus.

heterogeneous firms and a cost advantage for the home bank.¹⁴ Since the initial quote can be recalled, firms face a reservation price: $p^0 \leq \bar{p}$.

We can distinguish between two cases leading to a transaction: (i) $\bar{p} < c - \Delta$ and (ii) $c < p^0 + \Delta \leq \bar{p} + \Delta$. In the first case the borrower does not qualify at the home bank. Borrowers not qualifying at their home bank must search, and their reservation price is \bar{p} . This borrower may qualify at other banks because of differences in ω_j . The lowest-cost qualifying bank wins by offering a price equal to the lending cost of the second most efficient qualifying lender:

$$p^* = \min\{c + \omega_{(2)}, \bar{p}\}. \quad (3)$$

This occurs if and only if $0 < \bar{p} - c - \omega_{(1)}$.

If the borrower qualifies at the home bank, the highest-surplus bank wins and offers a quote that provides the same utility as the second-best option. The equilibrium pricing function is

$$p^* = \begin{cases} p^0 & \text{if } \bar{v} + \lambda - p^0 \geq \bar{v} - c - \omega_{(1)} \\ c + \omega_{(1)} + \lambda & \text{if } \bar{v} + \lambda - p^0 < \bar{v} - c - \omega_{(1)} < \bar{v} - c + \gamma \\ c - \gamma & \text{if } \bar{v} - c - \omega_{(1)} > \bar{v} - c + \gamma > \bar{v} - c - \omega_{(2)} \\ c + \omega_{(2)} & \text{if } \bar{v} - c - \omega_{(2)} > \bar{v} - c + \gamma. \end{cases} \quad (4)$$

This equation highlights the fact that, at the competition stage, lenders directly competing with the home bank will, on average, have to offer a discount equal to the loyalty advantage in order to attract new customers.¹⁵ In cases 1 and 2 the home bank provides the highest utility and so wins the auction. In case 1, the initial quote provides higher utility than does the next-best lender's quote, and so the consumer pays p^0 . In case 2, the reverse is true, and so the consumer pays $c + \omega_{(1)} + \lambda$ and gets utility of $\bar{v} - c - \omega_{(1)}$. In cases 3 and 4, the home bank is not the highest-surplus lender, and the consumer pays $c - \gamma$ or $c + \omega_{(2)}$, depending on whether the home bank is the second- or third-highest surplus lender.

¹⁴ Our approach is close in spirit to the literature in both labor economics and finance studying search and matching frictions in markets with bargaining (see, e.g., Postel-Vinay and Robin [2002] for on-the-job search and Duffie, Garleanu, and Pedersen [2005] for over-the-counter markets). More recently, the application of auction-like models to price negotiation settings has been used in the context of business-to-business transactions (e.g., Salz 2015; Beckett et al. 2016) and consumer markets (e.g., Woodward and Hall 2012; Allen et al. 2014a).

¹⁵ Equations (3) and (4) also highlight the fact that the transaction price is determined by three lenders: the home bank and the two most cost-efficient lenders. Therefore, while we assume that consumers search the entire choice set, an implication of the model is that consumers need to obtain formal quotes from at most three lenders. This is in line with a Bertrand-Nash interpretation of the game, in which consumers learn lenders' cost ranking after paying the search cost, for instance, through advertising, by calling banks directly, or indirectly through a real estate agent.

D. Search Decision and Initial Quote

The borrower chooses to search by weighing the value of accepting p^0 or paying a sunk cost κ to search in order to lower his expected monthly payment. The utility gain from search is

$$\begin{aligned} \bar{\kappa}(p^0, s) &= \underbrace{\bar{v} + \lambda[1 - G_{(1)}(-\gamma)] - E[p^*|p^0, s]}_{\text{2nd-stage expected utility}} - \underbrace{[\bar{v} + \lambda - p^0]}_{\text{1st-stage utility}} \\ &= p^0 - E[p^*|p^0, s] - \lambda G_{(1)}(-\gamma), \end{aligned}$$

where $1 - G_{(1)}(-\gamma)$ is the retention probability of the home bank in the competition stage. A consumer will reject p^0 if and only if the gain from search is larger than the search cost. Therefore, the search probability is

$$\Pr(\kappa < p^0 - E[p^*|p^0, s] - \lambda G_{(1)}(-\gamma|n)) \equiv H(\bar{\kappa}(p^0, s)). \tag{5}$$

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

$$\max_{p^0 \leq \bar{p}} (p^0 - c + \Delta)[1 - H(\bar{\kappa}(p^0, s))] + H(\bar{\kappa}(p^0, s))E(\pi_h^*|p^0, s),$$

where

$$E(\pi_h^*|p^0, s) = (p^0 - c + \Delta)[1 - G_{(1)}(p^0 - \lambda - c)] + \int_{-c}^{p^0 - c - \lambda} [\omega_{(1)} + \gamma] dG_{(1)}$$

are the expected profits from the auction for the home bank. The first term represents the case in which the initial quote provides higher utility than the next-highest surplus lender, while the second is the reverse.

Importantly, the home bank will offer a quote only if it makes positive profit at the posted price: $0 < \bar{p} - c + \Delta$. In the interior solution, the optimal initial quote is implicitly defined by the following first-order condition:

$$\begin{aligned} p^0 - c + \Delta &= \underbrace{\frac{1 - H(\bar{\kappa}(p^0, s))}{H'(\bar{\kappa}(p^0, s))\bar{\kappa}_{p^0}(p^0, s)}}_{\text{Search cost distribution}} + \underbrace{\frac{E(\pi_h^*|p^0, s)}{\partial p^0}}_{\text{Cost and quality differentiation}} \\ &\quad + \underbrace{\frac{H(\bar{\kappa}(p^0, s))}{H'(\bar{\kappa}(p^0, s))\bar{\kappa}_{p^0}(p^0, s)} \frac{\partial E(\pi_h^*|p^0, s)}{\partial p^0}}_{\text{Reserve price effect}}, \end{aligned} \tag{6}$$

where

$$\bar{\kappa}_{p^0}(p^0, s) = \frac{\partial \bar{\kappa}(p^0, s)}{\partial p^0}.$$

Equation (6) implicitly defines the home bank's profit margins from price discrimination. It highlights three sources of profits: (i) positive average search costs, (ii) market power from differentiation in cost and quality (i.e., match value differences and home bank cost advantage), and (iii) the reserve price effect. If firms are homogeneous, 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, in the following proposition (proven in app. B) we claim that, in the interior, it is additive in the common cost shock. This simplifies the problem, since we need to numerically solve the first-order condition for only one value of c per consumer.

PROPOSITION 1. The optimal initial quote, p^0 , is additive in c in the interior: $p^0 = c + \mu(\Delta, \lambda, n)$.

From this proposition, we can characterize the initial quote as follows:

$$p^0(s) = \begin{cases} \bar{p} & \text{if } c > \bar{p} - \mu(\Delta, \lambda, n) \\ c + \mu(\Delta, \lambda, n) & \text{else.} \end{cases}$$

To summarize, the model predicts three equilibrium functions: (i) the initial quote $p^0(s)$, (ii) the search cost threshold $\bar{\kappa}(s)$, and (iii) the competitive price $p^*(\omega, s)$. Although it is difficult to characterize these functions analytically, the separability of the initial quote in the interior leads to a series of useful predictions that are summarized in corollary 1. We use these implications in the identification section below.

COROLLARY 1. The following predictions about the distribution of prices and search probability in the interior, when $p^0(s) < \bar{p}$, follow from proposition 1:

- i. The equilibrium search probability is independent of c .
- ii. The equilibrium search probability is affected symmetrically by λ and Δ .
- iii. The distribution of p^* for switchers is a function only of $\gamma = \lambda + \Delta$.
- iv. The average transaction price paid by loyal consumers is affected asymmetrically by λ and Δ , and the effect of λ is stronger.

IV. Identification

The model contains four primitives: (i) the distribution of the common lending cost conditional on observed attributes of borrower i and re-

gion and period fixed effects (x_i) , $F(c_i|x_i)$; (ii) the distribution of idiosyncratic cost differences, $G(\omega_{ij})$; (iii) the search cost distribution, $H(\kappa_i)$; and (iv) the loyalty advantage parameters, (λ, Δ) . From the model description, we maintain the assumptions that the lending cost function is additively separable in c_i and ω_{ij} and that κ_i and ω_{ij} are independent and identically distributed. We also assume that loyalty parameters are common across consumers and that (κ_i, ω_{ij}) are independent of observed borrower characteristics x_i (this last assumption is partially relaxed in the empirical analysis).

In this section we discuss nonparametric identification of the search and cost distributions, as well as identification of the loyalty parameters. In the next, we estimate a parametric version of the model, which allows us to more easily incorporate observable differences between consumers and firms.

The data correspond to a cross section of transaction prices, borrower characteristics, closing weeks ($t(i)$), and lender choices (including whether or not the lender is the home bank). The closing week allows us to infer the posted price valid at the time of the transaction ($\bar{p}_{t(i)}$), while the location determines the number of available rival options in the borrower's neighborhood (n_i). From the data, we can therefore characterize the probability of switching lenders conditional on $(x_i, \bar{p}_{t(i)}, n_i)$, as well as the distribution of transaction prices given $(x_i, \bar{p}_{t(i)}, n_i)$ separately for switching and loyal consumers. These three distributions correspond to the reduced form of the model.

We face two challenges when discussing the mapping from the reduced form to the primitives of the model. First, since we observe only accepted offers, we must infer the distributions of the two unobserved heterogeneity components (c_i and ω_i) from a single price. Second, since we do not observe search and switch decisions separately, we need to distinguish between two sources of attachment to the home bank—search costs and the loyalty advantage—solely using the conditional probability of remaining loyal to the home bank.

To overcome these challenges, in addition to the assumptions listed above, we rely on two exclusion restrictions. We assume that the number of lenders and the posted price are independent of the c_i , conditional on the observed attributes of the borrower, x_i . Furthermore, we require that both variables exhibit enough variation across borrowers. We formally introduce these assumptions in appendix C and propose a sequential approach to show that they are sufficient to guarantee identification of the model. The argument can be summarized as follows.

1. Consider first the distribution of prices for switching borrowers facing very high posted prices: $\bar{p} \rightarrow \infty$. These transactions are generated from the auction and reflect the cost of the second most efficient lender (including potentially the home bank). Furthermore, since the posted price constraint is not binding, selection into the competition stage is indepen-

dent of the realization of c_i (from corollary 1 (i)). This eliminates the selection bias that arises from looking separately at switching consumers.¹⁶

In this subsample, the distribution of transaction prices across markets with different n 's can be used to separately identify $F(c_i|x_i)$, $G(\omega_i)$, and the sum of the two loyalty parameters ($\gamma = \Delta + \lambda$). To see this, note that when $n_i = 1$, the transaction price is equal to $p_i^* = c_i - \gamma$, which can be used to identify $F(c_i|x_i)$ given γ . Next, consider markets with a small number of lenders $n_i > 1$. In such markets, the presence of a positive loyalty advantage implies that prices paid by switchers mostly reflect the common cost component, which is independent of n_i . As the number of lenders increases, the probability that a rival lender, and not the home bank, is the next-best alternative converges to one. For large n_b the distribution of idiosyncratic cost differences, $G(\omega_{ij})$, is identified using standard English auction arguments.

In between, the correlation among the number of rivals and the price paid by switchers depends on the magnitude of the loyalty advantage parameter: the larger is γ , the smaller is the effect of n on prices. Therefore, γ is identified from the strength of the correlation between the number of rivals and p^* , as the number of competitors becomes large.

2. Consider next data on the probability of remaining loyal to the home bank, conditional on $(x_i, \bar{p}_{l(i)}, n_i)$. In the model, this probability corresponds to the product of the search probability and the probability that the home bank retains the consumer at the auction stage (i.e., $G_{(n)}(-\gamma)$). The previous argument suggests that the gain from search and the retention probability, which are functions of $G(\omega)$ and γ , can be computed directly from the distribution of prices for unconstrained switching consumers. However, in the absence of the constraint imposed by the posted price, the switching probability takes only discrete values in equilibrium, one for each $n_i \in \{1, 2, \dots, \bar{n}\}$ because c_i does not affect the search probability. These moments would be sufficient to test the null hypothesis that search costs are zero, but not to identify the distribution $H(\kappa_i)$ nonparametrically.¹⁷

The presence of a binding posted price constraint breaks this independence and creates dispersion in the search cost thresholds across consumers within the same market. In particular, for consumers receiving $p^0 = \bar{p}$, the search probability is monotonically increasing in \bar{p} . Therefore, exogenous variation in \bar{p} can be used to nonparametrically identify the distribution of search costs by varying the search cost threshold across consumers with similar x_i and n_i .

¹⁶ This is similar to the identification at infinity arguments used in labor economics to study "Roy-type" models (e.g., French and Taber 2011).

¹⁷ Under the null hypothesis that search costs are zero, the observed probability of remaining loyal to the home bank is equal to the retention probability, $G_{(n)}(-\gamma)$. This probability can be calculated from data on the price paid by switching consumers.

3. Finally, the observed distribution of prices among loyal consumers can be used to separate the effect of loyalty on cost (i.e., Δ) and willingness to pay (i.e., λ). This distribution is a mixture of initial quote offers and auction prices. We know from corollary 1(iv) that λ and Δ have different impacts on the average transaction price of loyal consumers. In contrast, λ and Δ affect symmetrically the equilibrium search probability in the interior (corollary 1(ii)) and the distribution of prices for switchers (corollary 1(iii)). Therefore, while both parameters influence in the same way the observed retention probability, they have different effects on the average price difference between loyal and switching consumers. This moment can thus be used to identify λ separately from Δ .

This identification argument relies on the existence of important variation in the number of lenders and the posted price across consumers. This is particularly relevant for the identification of the search cost distribution (step 2). In practice, we observe fairly limited time-series variation in the posted rate and very few consumers with fewer than four lenders in their neighborhoods. Given these shortcomings of the data, we incorporate additional aggregate moments measuring the fraction of borrowers gathering more than one quote, conditional on (limited) demographic characteristics (from the FIRM survey). With this additional information, the separate identification of the search and loyalty parameters becomes even more transparent. We now have two measures of state dependence: the average switching probability (\bar{S}) and the average search probability (\bar{H}). Using these measures, one can use the predicted switching probability to estimate the aggregate retention probability of the home bank at the auction stage:

$$\bar{S} = \bar{H} \times G_{(1)}(-\gamma), \quad G_{(1)}(-\gamma) = \frac{\bar{S}}{\bar{H}}.$$

For instance, in our sample the average switching probability is a little less than 30 percent, while the aggregate search probability from the FIRM survey is just over 60 percent. On average, the home bank therefore wins the auction with probability 46 percent. Since, on average, the number of lenders per neighborhood is eight, this implies that the loyalty advantage is positive and large relative to the dispersion of idiosyncratic cost differences.

V. 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. We then derive the likelihood function induced by the model and discuss the sources of identification.

A. *Distributional Assumptions and Functional Forms*

The lending cost function differs slightly from the model presentation. Specifically, we account for loan size differences across borrowers, and we allow observed bank characteristics to affect the distribution of cost differences across lenders (i.e., ω_{ij} and Δ_i).

We model the monthly cost of lending $\$L_i$ over a 25-year amortization period using a linear function of borrower and lender characteristics:

$$c_{ij} = L_i \times (c_i + \omega_{ij}), \quad (7)$$

where the common cost component is normally distributed, $c_i \sim N(x_i\beta, \sigma_c^2)$, and the idiosyncratic cost differences are distributed according to a lender-specific type 1 extreme-value distribution, $\omega_{ij} \sim \text{T1EV}(\xi_{ij} - e\sigma_\omega, \sigma_\omega)$.¹⁸

The location parameter of the idiosyncratic cost difference distribution, ξ_{ij} , varies across lenders because of the presence of bank fixed effects and the size of the branch network in the neighborhood of the consumer (normalized by the average network size of rivals). The type 1 extreme-value distribution assumption leads to analytical expressions for the distribution functions of the first- and second-order statistics and is often used to model asymmetric value distributions in auction settings (see, e.g., Brannan and Froeb 2000).

The loan size is normalized so that the per-unit lending cost in equation (7) measures the monthly cost of a \$100,000 loan. The vector x_i controls for observed financial characteristics of the borrower (income, loan size, FICO score, LTV, etc.), the bond rate, as well as period and location 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 code district). The period fixed effects are defined at the quarter-year level.

The lending cost of the home bank is expressed slightly differently because of the home bank cost advantage parameter:

$$c_{ih} = L_i \times [c_i + \Delta_{ih}],$$

where h is the home bank index of borrower i , and $\Delta_{ih} = \xi_{ih} - \Delta(z_i^2)$ is consumer i 's home bank deterministic cost differential. In the application, we allow the cost advantage parameter to depend on the borrower's income and home ownership status:

$$\Delta(z_i^2) = L_i \times (\Delta_0 + \Delta_{\text{inc}}\text{Income}_i + \Delta_{\text{owner}}\text{Previous Owner}_i).$$

The WTP component of the loyalty advantage is defined analogously as a linear function of income and home ownership status:

¹⁸ The location parameter of the type 1 extreme-value distribution is adjusted by a factor $e\sigma_\omega$ to guarantee that the error is mean zero (i.e., e is the Euler constant).

$$\lambda(z_i^2) = L_i \times (\lambda_0 + \lambda_{\text{inc}} \text{Income}_i + \lambda_{\text{owner}} \text{Previous Owner}_i).$$

Finally, we assume that the search cost is exponentially distributed with a consumer-specific mean that depends on income and home ownership status:

$$H(\kappa|z_i^1) = 1 - \exp\left(-\frac{1}{\alpha(z_i^1)} \kappa\right),$$

$$\log \alpha(z_i^1) = \alpha_0 + \alpha_{\text{inc}} \log \text{Income}_i + \alpha_{\text{owner}} \text{Previous Owner}_i.$$

B. Likelihood Function

We estimate the model by maximum likelihood. The endogenous outcomes of the model are the chosen lender and monthly payment $\{b(i), p_i\}$, as well as whether consumers remain loyal to their home bank or switch. The observed prices are generated from consumers either accepting the initial quote (i.e., $p_i = p^0(s)$) or accepting the competitive offer (i.e., $p_i = p^*(\omega, s)$). Importantly, only the latter case is feasible if consumers switch financial institutions, while both cases have a positive likelihood for loyal consumers.

Moreover, the identity of the home bank is known for loyal consumers, while it is unobserved for switching consumers. To construct the likelihood function, we first condition on the identity of the home bank for both types of transactions and then integrate out h using the empirical distribution of h defined in Section II.

In order to derive the likelihood contribution of each individual, we first condition on the choice set \mathcal{N}_i ,¹⁹ the observed characteristics x_i , the identity of home bank h , the posted price valid at the time consumer i negotiated the contract $\bar{p}_{l(i)}$, and the model parameter vector $\theta = \{\beta, \xi, \sigma_\omega, \sigma_c, \alpha, \Delta, \lambda\}$. Let $\mathcal{I}_i = \{\mathcal{N}_i, x_i, \bar{p}_{l(i)}\}$ summarize the information known by the econometrician about consumer i .

In order to simplify notation, we use individual subscripts i for the borrower characteristics and random variables, with the understanding that all functions and variables are consumer specific and depend on \mathcal{I}_i and the parameter vector θ . For instance, $\Delta_{i,h} = \xi_{i,h} - \Delta(z_i^2)$ and $\lambda_i = \lambda(z_i^2)$ denote the home bank cost and WTP advantages, and $\mu_i \equiv \mu(\mathcal{N}_i, \Delta_{i,h}, \lambda_i)$ is used to denote the initial quote markup (interior solution). In addition, we use c_i to summarize the state variable in the initial stage of the game instead of $s_i = \{c_i, \bar{p}_{l(i)}, \mathcal{N}_i, \Delta_{i,h}, \lambda_i\}$. For instance, $\bar{\kappa}(c_i) \equiv \bar{\kappa}(s_i)$ and

¹⁹ We use \mathcal{N}_i rather than n_i to characterize the choice set of consumers, since the identities of banks present in each neighborhood (not just the number) enter the distribution of lending costs.

$p^0(c_i) \equiv p^0(s_i)$ correspond to the equilibrium search cost threshold and initial quote, respectively.

Next we summarize the likelihood contribution for loyal and switching consumers. Appendix D describes in greater details the derivation of the likelihood function.

Likelihood Contribution for Loyal Consumers

The main obstacle in evaluating the likelihood function is that we do not observe whether or not consumers search. The unconditional likelihood contribution of loyal consumers is therefore

$$\begin{aligned} L(p_i, b(i)) &= h|\mathcal{I}_i, h, \theta \\ &= L(p_i = p^0(c_i), b(i) = h|\mathcal{I}_i, h, \theta) \\ &\quad + L(p_i = p^*(\omega_i, c_i), b(i) = h|\mathcal{I}_i, h, \theta). \end{aligned} \quad (8)$$

The first term is a function of the solution to the optimal initial quote: $p^0(c_i) = \min\{\bar{p}_{l(i)}, c_i + \mu_i\}$. Since the markup is independent of c_i in the interior, the distribution of p_i takes the form of a truncated distribution:

$$\begin{aligned} L(p_i = p^0(c_i), b(i) = h|\mathcal{I}_i, h, \theta) &= \begin{cases} f(p_i - \mu_i|x_i)[1 - H(\bar{\kappa}(p_i - \mu_i))] & \text{if } p_i < \bar{p}_{l(i)} \\ \int_{\bar{p}_{l(i)} - \mu_i}^{\bar{p}_{l(i)} + \Delta_{i,h}} [1 - H(\bar{\kappa}(c_i))] dF(c_i|x_i) & \text{if } p_i = \bar{p}_{l(i)}. \end{cases} \end{aligned} \quad (9)$$

The second element measures the probability of observing a constrained initial quote. This event occurs if $c_i > \bar{p}_{l(i)} - \mu_i$ and the consumer qualifies for a loan at its home bank (i.e., $c_i < \bar{p}_{l(i)} - \Delta_{i,h}$).

In addition to the search cost and the common lending cost, the likelihood contribution from searching consumers reflects the realization of the lowest-cost differential in \mathcal{N}_i (i.e., $\omega_{i(1)}$). The transaction price is given by $p_i = p^0(c_i)$ if $\omega_{i(1)} > p^0(c_i) - c_i - \lambda_i$ or by $p_i = c_i + \omega_{i(1)} + \lambda_i$ otherwise:

$$\begin{aligned} L(p_i = p^*(\omega_i, c_i), b(i) = h|\mathcal{I}_i, h, \theta) &= \begin{cases} [1 - G_{(1)}(\mu_i - \lambda_i|\mathcal{N}_i)]H(\bar{\kappa}(p_i - \mu_i))f(p_i - \mu_i|x_i) \\ \quad + \int_{p_i - \mu_i}^{p_i + \Delta_{i,h}} g_{(1)}(p_i - c_i - \lambda_i)H(\bar{\kappa}(c_i))dF(c_i|x_i) & \text{if } p_i < \bar{p}_{l(i)} \\ \int_{\bar{p}_{l(i)} - \mu_i}^{\bar{p}_{l(i)} + \Delta_{i,h}} [1 - G_{(1)}(\bar{p}_{l(i)} - c_i - \lambda_i|\mathcal{N}_i)]H(\bar{\kappa}(c_i))dF(c_i|x_i) & \text{if } p_i = \bar{p}_{l(i)}. \end{cases} \end{aligned} \quad (10)$$

Likelihood Contribution for Switching Consumers

For switching consumers, the likelihood contribution depends on the relative position of the home bank in the surplus distribution of lenders belonging to \mathcal{N}_i . We use $g_b(\omega)$ to denote the density of the cost differential of the chosen lender and $g_{-b}(\omega|\mathcal{N}_i)$ to denote the density of the most efficient lender in \mathcal{N}_i other than b .²⁰

If the observed price is unconstrained, the transaction price is equal to the minimum of $c_i - (\Delta_{i,h} + \lambda_i)$ and $c_i + \omega_{i,-b}$. If the consumer does not qualify for a loan at his home bank, the transaction price is the minimum of the posted price and the second-lowest cost. This occurs if $c_i > \bar{p}_{l(i)} + \Delta_{i,h}$. Therefore, the transaction price for switching consumers is equal to \bar{p} if and only if the chosen lender is the only qualifying bank. This leads to the following likelihood contribution:

$$\begin{aligned}
 &L(p_i, b(i) \neq h | \mathcal{I}_i, h, \theta) \\
 &= \begin{cases} 1(\bar{p}_{l(i)} > p_i + \lambda_i) \left\{ [1 - G_{-b}(-\Delta_{i,h} - \lambda_i | \mathcal{N}_i)] G_b(-\Delta_{i,h} - \lambda_i) \right. \\ \quad \left. \times H(\bar{\kappa}(p_i + \Delta_{i,h} + \lambda_i)) f(p_i + \Delta_{i,h} + \lambda_i | x_i) \right\} \\ \quad + \int_{p_i + \Delta_{i,h} + \lambda_i}^{\infty} G_b(p_i - c_i) H(\bar{\kappa}(c_i)) g_{-b}(p_i - c_i | \mathcal{N}_i) dF(c_i | x_i) & \text{if } p_i < \bar{p}_{l(i)} \\ \quad \int_{\bar{p}_{l(i)} + \Delta_{i,h}}^{\infty} G_b(\bar{p} - c_i) [1 - G_{-b}(\bar{p}_{l(i)} - c_i | \mathcal{N}_i)] dF(c_i | x_i) & \text{if } p_i = \bar{p}_{l(i)}. \end{cases} \tag{11}
 \end{aligned}$$

Note that the first term is equal to zero if $\bar{p}_{l(i)} < p_i + \lambda_i$.²¹ This condition ensures that the home bank's lending cost is below $\bar{p}_{l(i)}$ at the observed transaction price.

Integration of the Home Bank Identity and Selection

The unconditional likelihood contribution of each individual is evaluated by integrating out the identity of the home bank. Recall that h is missing for a sample of contracts and is unobserved for switchers. In the former case we use the unconditional distribution of home banks, while in the latter case we condition on the fact that the chosen lender is not the home bank. This leads to the following unconditional likelihood:

²⁰ The density $g_{-b}(\omega|\mathcal{N}_i)$ is $g_{(1)}(\omega|\mathcal{N}\setminus b)$.

²¹ This reduces the smoothness of the likelihood, affecting primarily the parameters determining λ . To remedy this problem we smooth the likelihood by multiplying the second term in eq. (11) by $[1 + \exp((\lambda_i - \bar{p}_{l(i)} + p_i)/s)]^{-1}$, where s is a smoothing parameter set to 0.01.

$$L(p_i, b(i)|\mathcal{I}_i, \theta) = \begin{cases} L(p_i, b(i)|\mathcal{I}_i, h = b(i), \theta) & \text{if } 1(\text{Loyal}_i) = 1 \\ \sum_{h \neq b(i)} \frac{\psi_h(x_i)}{\sum_{j \neq b(i)} \psi_j(x_i)} L(p_i, b(i)|\mathcal{I}_i, h, \theta) & \text{if } 1(\text{Loyal}_i) = 0 \\ \sum_h \psi_h(x_i) L(p_i, b(i)|\mathcal{I}_i, h, \theta) & \text{if } 1(\text{Loyal}_i) = M/V, \end{cases} \quad (12)$$

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

In addition, the fact that we observe only accepted offers implies that the unconditional likelihood suffers from a sample selection problem. The probability that consumer i is in our sample is given by the probability of qualifying for a loan from at least one bank in i 's choice set. This is given by the probability that the minimum of $c_i - \Delta_{i,h}$ and $c_i + \omega_{i,(1)}$ is less than $\bar{p}_{i(i)}$:

$$\Pr(\text{Qualify}|\mathcal{I}_i, \theta) = \sum_h \psi_h(x_i) \int F(\bar{p}_{i(i)} - \min\{\omega_{i,(1)}, -\Delta_{i,h}\} | x_i) dG_{(1)}(\omega_{i,(1)} | \mathcal{N}_i). \quad (13)$$

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

$$L^c(p_i, b(i)|\mathcal{I}_i, \theta) = L(p_i, b(i)|\mathcal{I}_i, \theta) / \Pr(\text{Qualify}|\mathcal{I}_i, \theta). \quad (14)$$

Aggregate Likelihood Function

To construct the likelihood function we need to aggregate the information contained in equation (14) across the N observed contracts while incorporating additional external aggregate information on search effort. We use the results of the annual FIRM survey (described in Sec. II) to match the probability of gathering more than one quote along three dimensions: city size, region, and income group.

Using the model and the observed new home buyer characteristics, we calculate the probability of rejecting the initial quote, integrating over the model shocks and the identity of the incumbent bank. Let $\bar{H}_g(\theta)$ denote this function for demographic group g . Similarly, let \hat{H}_g denote the analogue probability calculated from the survey. The difference between the two, $m_g(\theta) = \bar{H}_g(\theta) - \hat{H}_g$, is a mean-zero error under the null hypothesis that the model is correctly specified. We use $G = 8$ aggregate moments.

Several econometric approaches have been proposed in the literature for combining data from multiple surveys. When individual data from independent surveys are available, a standard approach is to maximize a joint likelihood defined as the product of density functions calculated from separate data sets (e.g., van den Berg and van der Klaauw 2001). This

approach is not feasible in our case, since we observe only aggregate moments and do not have access to the microdata from the search probability survey. Alternatively, we could use a constrained maximum likelihood estimator (MLE) that maximizes the sum of individual likelihood contributions subject to the constraint that the aggregate moment conditions are satisfied exactly (see Ridder and Moffitt 2007). The disadvantage of this approach is that it ignores the fact that the aggregate moments are themselves measured with error. In our application the number of observations used to measure the aggregate moments is less than 500, compared to close to 30,000 in the contract data. A third approach, which takes the relative sample sizes of the two data sets into account, is the generalized method of moments (GMM) estimator proposed by Imbens and Lancaster (1994). This approach combines moment restrictions obtained from the score of the log likelihood function, with the vector of aggregate errors obtained by matching moments from the survey.

Although this option would be a natural choice, it can be difficult to implement in practice and does not perform well numerically for our specific problem. The reason is that to evaluate the GMM objective function we must rely on numerical derivatives to compute the score function. This is challenging since the likelihood function involves repeatedly solving a nested fixed point and numerically approximating several integrals. With over 60 parameters this represents a nontrivial increase in computation time relative to evaluating the likelihood function once. Furthermore, the numerical score function is less smooth than the likelihood function, making optimization of the GMM problem numerically more prone to convergence problems. We experimented with different optimization routines without success and decided to use an alternative estimating procedure.

We use a quasi likelihood estimator that relies on a normal approximation to the density of the aggregate residuals. Let σ_g^2 denote the predicted variance in the search probability across consumers in group g (calculated from the model). From the central-limit theorem, $\sqrt{M_g} m_g(\theta) / \sigma_g$ is a sample average that is normally distributed when M_g is large enough (i.e., the number of consumers surveyed in group g). In our case, the number of households surveyed by the Altus Group in each group ranges between 265 and 441.

Under this assumption, the combined quasi likelihood is the product of the conditional likelihood function obtained from the contract data (product of eq. [14] across N) and the normal densities associated with each of the aggregate moments. This leads to the following aggregate log likelihood function:²²

²² The parameters are estimated by maximizing the aggregate log likelihood function using the Broyden-Fletcher-Goldfarb-Shanno numerical optimization algorithm within the Ox programming language (Doornik 2007).

$$\max_{\theta} \sum_i \log L(p_i, b_i | \mathcal{I}_i, \theta) - m(\theta)^T \hat{W}_2^{-1} m(\theta), \quad (15)$$

where $m(\theta)$ is a $K \times 1$ vector of errors from the auxiliary moments, and \hat{W}_2 is a diagonal matrix with the estimated asymptotic variance of the moments.²³

Note that the constrained MLE problem takes a similar form:

$$\max_{\theta, \rho} \sum_i \log L(p_i, b_i | \mathcal{I}_i, \theta) - \rho m(\theta)^T \hat{W}_2^{-1} m(\theta), \quad (16)$$

where $\rho \geq 0$ is a Lagrangian multiplier. Intuitively, as the number of observations in the auxiliary survey goes to infinity (holding fixed N), \hat{W}_2^{-1} goes to infinity (in eq. [15]), and our quasi likelihood estimator forces the aggregate moments to be satisfied with equality almost surely (just as with constrained MLE).

By setting $\rho = 1$, the weight that the quasi likelihood puts on the auxiliary moments depends on the sample size.²⁴ In that sense, our approach is similar to the GMM estimator proposed by Imbens and Lancaster (1994). However, the two estimators cannot be nested in any sense. The moment conditions in Imbens and Lancaster's study are not the same as the score of the quasi likelihood defined in equation (15). When using a block-diagonal weighting matrix for each set of moment conditions, the GMM estimator minimizes the sum of the square of the scores minus a penalty function to account for the sum of square of the moment residual, while our estimator maximizes the sum the log likelihood function minus the same quadratic penalty function. We have conducted a series of Monte Carlo simulations to analyze the small-sample performance of both estimators and found that our quasi likelihood estimator performs equally as well as or better than GMM. These results are available in appendix E. The appendix also provides additional details as to the differences between GMM and our quasi likelihood approach.

Computation

In order to evaluate the aggregate likelihood function, we must first solve the optimal initial offer defined implicitly by equation (6). This nonlinear equation needs to be solved separately for every consumer/home bank combination. We perform this operation numerically using a New-

²³ We estimate σ_g by calculating the within-group variance in search probability using the sample of individual contracts. Since this variance depends on the model parameter values, we follow a two-step approach: (i) calculate σ_g using an initial estimate of θ (e.g., starting with $\sigma_g = 1$), and (ii) hold σ_g fixed to estimate $\hat{\theta}$. The term \hat{W}_2 is a diagonal matrix with element (g, g) given by $2\hat{\sigma}_g^2/M_g$. The multiple 2 is coming from the fact that the log of the normal density is proportional to $-0.5(x - \mu)' \Sigma^{-1}(x - \mu)$.

²⁴ In the empirical application, we compare the results for two values of the Lagrangian multiplier: $\rho = 1$ and $\rho = 100$.

ton algorithm that uses the first and second derivatives of firms' expected profits. We use starting values defined as the expected initial quote from the complete information problem, for which we have an analytical expression. This procedure is robust and converges in a small number of steps. Notice that since the interior solution is additive in c , this equation needs to be solved only once for each evaluation of the likelihood contribution of each household, $L(l_i, b(i)|\mathcal{I}_i, h, \theta)$. In addition, the integrals are evaluated numerically using a quadrature approximation.

VI. Estimation Results

A. Parameter Estimates

Table 3 summarizes the maximum likelihood estimates from three specifications, each one varying the source of the loyalty advantage. In specification 1, the loyalty advantage takes the form of a WTP term, λ , for the home bank. In specification 2, the home bank has a cost advantage, Δ , over competing lenders. Specification 3 nests both models.

Each specification implies that the home bank is more likely to "win" against rival banks at the competition stage but have different implications for the price differences between loyal and switching borrowers. Holding fixed the magnitude of the idiosyncratic cost differences between lenders (σ_ω), the WTP model implies a larger average price difference between loyal and switching borrowers, relative to the cost advantage model. This difference is relatively small in the data: loyal borrowers pay about 10 bps more than switching borrowers, or about 10 percent of the standard deviation of residual rates. In specification 1, the model reconciles these two features with small estimates of σ_ω and λ_0 . In contrast, the cost advantage model leads to larger estimates of the differentiation parameters, Δ and σ_ω . Also, the cost advantage model fits the data significantly better.

We formally assess the performance of the two modeling choices by estimating specification 3. The last row reports the results of two likelihood ratio tests testing the null hypothesis that $\lambda_i = 0$ and $\Delta_i = 0$. We can easily reject the null hypothesis that the cost advantage parameters are zero; the test statistic is more than 40 times larger than the 1 percent critical value (i.e., 660.7 vs. 16.3). In contrast, the null hypothesis of zero home bank WTP parameters is much more modestly rejected (i.e., 45.7 vs. 16.3).

A closer look at the estimates of λ in specification 3 reveals that the intercept and owner parameters are not significantly different from zero statistically or economically, while the estimated cost advantage parameters are large and precisely estimated. The reverse is true for the interaction of income and loyalty. This suggests that the relationship between loyalty and income is better explained by the WTP model. Still, the effect

TABLE 3
MAXIMUM LIKELIHOOD ESTIMATION RESULTS

	Specification 1	Specification 2: Baseline	Specification 3
Heterogeneity and preferences:			
Common shock (σ_ϵ)	.356 (.003)	.358 (.003)	.358 (.003)
Idiosyncratic shock (σ_ω)	.047 (.002)	.102 (.002)	.094 (.002)
Average search cost (log):			
α_0	-1.539 (.042)	-1.506 (.026)	-1.592 (.034)
α_{inc}	.458 (.052)	.401 (.038)	.356 (.045)
α_{owner}	.184 (.054)	.086 (.059)	.143 (.059)
Home bank WTP:			
λ_0	.064 (.003)		.010 (.007)
λ_{owner}	.032 (.002)		-.016 (.008)
λ_{inc}	.002 (.003)		.023 (.01)
Home bank cost advantage:			
Δ_0		.146 (.005)	.126 (.008)
Δ_{owner}		.066 (.004)	.075 (.008)
Δ_{inc}		.012 (.006)	-.010 (.01)
Cost function:			
Intercept	5.332 (.229)	5.495 (.229)	5.479 (.23)
Bond rate	.307 (.026)	.306 (.026)	.306 (.026)
Range posted rate	-.147 (.017)	-.145 (.017)	-.145 (.017)
Total loan	-.220 (.073)	-.208 (.073)	-.208 (.073)
Income	-.228 (.026)	-.214 (.026)	-.228 (.027)
Loan/income	-.100 (.01)	-.102 (.01)	-.102 (.01)
Previous owner	-.003 (.007)	.047 (.007)	.051 (.008)
House price	.222 (.066)	.211 (.066)	.211 (.066)
FICO	-.662 (.038)	-.656 (.038)	-.660 (.038)
LTV	1.111 (.157)	1.092 (.158)	1.093 (.157)
1 (LTV = 95%)	.029 (.008)	.029 (.008)	.029 (.008)
Relative network size	-.019 (.001)	-.039 (.002)	-.036 (.002)
Range of bank fixed effects	[-.041, .038]	[-.088, .063]	[-.08, .059]
Quarter-year fixed effects	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes
Sample size	26,218	26,218	26,218
Log likelihood function/ N	-2.059	-2.048	-2.047
Search moments weight	1	1	1
Likelihood ratio test ($\chi^2(3)$)	660.678	45.660	

NOTE.—Units: \$100s per month. Average search cost function: $\log \alpha(z_i^1) = \alpha_0 + \alpha_{inc} \log(\text{Income}_i) + \alpha_{owner} \text{Previous Owner}_i$. Home bank willingness to pay: $\lambda(z_i^2) = L_i \times (\lambda_0 + \lambda_{inc} \text{Income}_i + \lambda_{owner} \text{Previous Owner}_i)$. Home bank cost advantage: $\Delta(z_i^3) = L_i \times (\Delta_0 + \Delta_{inc} \text{Income}_i + \Delta_{owner} \text{Previous Owner}_i)$. Cost function: $c_{ij} = L_i \times (c_i + \omega_{ij})$, where $c_i \sim N(x_i \beta, \sigma_c^2)$ and $\omega_{ij} \sim \text{TIEV}(\xi_j + \xi_{branch}, \text{Rel. network size}_{ij} - e\sigma_\omega, \sigma_\omega)$. The likelihood ratio test compares models 1 and 2 against model 3 (alternative hypothesis). The 1 percent significance level critical value is 16.266. Specification 2 is our baseline model. Robust standard errors are in parentheses (White 1982).

of income on the loyalty advantage is economically small and imprecise in all three specifications. Since the data do not support the WTP model, we use the cost advantage model as our baseline specification.

Table G.1 in appendix G evaluates the robustness of the results to the weight assigned to the auxiliary search moments. Specifically, we reestimated the model with weights of zero and 100 on the auxiliary search moments.

A weight of 100 is analogous to increasing the sample size of the search survey to be roughly on par with the number of observations in the mortgage contract data. Doing so tends to increase the magnitude and heterogeneity of the loyalty advantage parameters (i.e., λ and Δ) and changes the sign of the income coefficient in the search cost function. This allows the model to better match the observed heterogeneity in the search probability across market size and income groups (see the goodness of fit discussion below).

By setting a zero weight, the parameters are identified solely using the mortgage contract data. The results from specifications 4 and 5 are similar to the results presented in table 3, which is not surprising given the fact that the sample size in the contract data is much larger than in the search survey. The most noticeable differences between the two estimates are that the average search cost is lower with a weight of zero (by about 15–20 percent) and that the dispersion of costs across lenders is larger (e.g., $\sigma_\omega = 0.12$ instead of $\sigma_\omega = 0.1$). Both features imply a larger predicted search probability in specifications 4 and 5 relative to 2 and 3 (approximately 3 percentage points). The fact that these differences are fairly minor confirms that the model's key parameters can be identified without using direct information on search behavior.

Next, we discuss the economic magnitude of the parameter estimates, focusing on the lending cost function and the search cost distribution. To better understand the magnitude of the estimates, recall that consumers choose a lender by minimizing their monthly payment net of the search cost. The monthly cost of supplying a \$100,000 loan is a linear function of borrowers' observed and unobserved characteristics, and the parameters are expressed in \$100 per month. For instance, in table 3 the variance parameter of the common shock, $\sigma_c = 0.358$, implies that the common lending cost standard deviation for a \$100,000 loan with fixed attributes is equal to \$35.80 per month.

Lending Cost Function

The first two parameters, σ_c and σ_ω , measure the relative importance of consumer unobserved heterogeneity with respect to the cost of lending. The standard deviation of the common component is 64 percent larger than the standard deviation of idiosyncratic shock (i.e., 0.358 vs. 0.128), suggesting that most of the residual price dispersion is due to consumer-level unobserved heterogeneity rather than to idiosyncratic differences across lenders.²⁵

²⁵ The standard deviation of an extreme-value random variable is equal to $\sigma_c\pi/\sqrt{6}$, or 0.102 in our case.

The estimate of σ_ω has key implications for our understanding of the importance of market power in this context. Abstracting from systematic differences across banks, the average cost difference between the first- and second-lowest-cost lender, $c_{(1)}$ and $c_{(2)}$, is equal to \$20 in duopoly markets, is \$17 with three lenders, and approaches \$14 when the number of lenders is 11.

In the model, market power also arises because of systematic cost differences across banks: (i) bank fixed effects, (ii) network size, and (iii) home bank cost advantage. The estimates of the fixed effects reveal relatively small differences across banks. Three of the 11 coefficients are not statistically different from zero (relative to the reference bank), and the range of fixed effects is equal to \$15 per month in our baseline specification, or about the same scale as the standard deviation of the idiosyncratic components.

We incorporate network size in the model by allowing the lending cost to depend on the relative branch network size of lenders in the same neighborhood. The estimates reveal that a lender with three times more branches than the average would experience a cost advantage of about \$12 per month (compared to a single-branch institution). This is consistent with our interpretation of the lending cost function, as capturing elements of profits from complementary banking services that are increasing in branch network size.

Turning to the estimate of Δ_i , we find that the presence of the loyalty advantage corresponds to an average cost advantage of \$17.10 per month (for a loan size of \$100,000). This cost advantage is substantial, given the fact that σ_ω is relatively small. At the estimated parameters, the probability that the home bank has a cost lower than the most efficient lender in \mathcal{N}_i (i.e., the retention probability at the auction) is equal to $G_{(1)}(\omega_h) = 51$ percent, substantially more than the uniform probability of choosing a lender at random in the average choice set (i.e., $1/8 = 12$ percent).

As mentioned, this cost advantage arises from the presence of switching costs, or complementarities between mortgage lending and other financial services, since the home bank enjoys a cost advantage relative to rival lenders due to its profits from other services. To capture these gains, rival lenders must offer (costly) discounts on other products to get consumers to switch institutions.

Panel A of table 4 summarizes the distribution of Δ_i across borrowers. Recall that the loyalty advantage is a deterministic function of income and prior ownership status. We find that the home bank cost advantage is particularly important for previous owners, suggesting that underlying switching costs are more important for older borrowers with longer prior experience. In comparison, the effect of income on the loyalty advantage is positive but much smaller (less than \$0.50 per month).

TABLE 4
SUMMARY STATISTICS ON THE HOME BANK COST ADVANTAGE, SEARCH, AND INTEREST COSTS

Variable	Mean	Standard Deviation	P25	P50	P75
A. Home Bank Cost Advantage					
New home buyers:					
Income < \$60,000	15.1	.111	15.1	15.1	15.2
Income ≥ \$60,000	15.6	.273	15.4	15.5	15.7
Previous owners:					
Income < \$60,000	21.7	.107	21.6	21.7	21.8
Income ≥ \$60,000	22.3	.336	22	22.2	22.4
B. Search and Interest Cost					
Nonsearchers:					
Total search cost	2.3	1.3	1.47	2	2.82
Interest cost	44.2	18.3	30.1	40.9	55.3
Searchers:					
Total search cost	.549	.443	.203	.461	.809
Interest cost	45.6	19	30.6	42.6	57.6
Total:					
Total search cost	1.15	1.19	.323	.784	1.58
Interest cost	45.1	18.8	30.4	42	56.9

NOTE.—Units in panel A are \$/month. The cost advantage is measured for a \$100,000 loan. Units in panel B are \$1,000s. The search and interest costs correspond to the total over the term of the mortgage contract (60 months).

Search Cost Distribution

Table 3 reports the parameters of the average search costs. Recall that we use an exponential distribution and model the mean as a log-linear function of income and prior ownership status. We find that search costs are increasing in income and ownership experience. New home buyers are estimated to have lower search costs on average (8.6 percent), and a 1 percent increase in income leads to a 0.4 percent 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.

Since search costs are not paid on a monthly basis, panel B of table 4 summarizes the simulated distribution of search costs expressed over the 5-year term of the mortgage contract.²⁶ The bottom rows report the unconditional distribution, and the top two rows illustrate the selection effect of consumers' search decisions. On average, we estimate that the cost

²⁶ Most mortgage contracts in Canada involve substantial financial penalties for borrowers who decide to prepay their mortgage before the end of the 5-year term period. Borrowers are free to switch lenders after this period. It is therefore reasonable to use the term period length as the planning horizon.

of searching for multiple offers is equal to \$1,150 (with a median of \$784). The difference between searchers and nonsearchers is substantial. We estimate that the search cost of “searchers” is \$549, while “nonsearchers” decided to accept the initial offer in order to avoid paying, on average, \$2,300 in search costs.

To put these numbers in perspective, we also report in panel B of table 4 the total interest cost over 5 years. While the search cost estimates are nominally very important, they represent, on average, only 2.5 percent of the overall cost of the contracts (i.e., $2.5\% = 1.15/45.1$).

An important feature of the model is that consumers financing larger loans are more likely to search. The reason is that the gains from search are increasing in loan size, while the search cost is fixed. As a result, in panel B of table 4 we find that searchers incur 3 percent larger total interest costs. The reason is that they finance loans that are, on average, \$11,000 larger than those of nonsearchers, despite paying, on average, 20 bps lower rates.

Are these number realistic? Woodward and Hall (2012) calculate that a US 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 distribution is consistent with this measure. Our results are also comparable to those in Allen et al. (2014a), where, using a simpler complete-information analogue to the bargaining model employed here, results suggest that for the Canadian mortgage market search costs represent about 4 percent of the overall cost.

How do our results compare to existing estimates of search costs in the literature? Perhaps the closest point of comparison comes from Honka's (2014) analysis of the insurance market. She estimates the cost of searching for policies to be \$35 per online search and a little over \$100 per off-line search. These numbers represent roughly 6 percent and 20 percent of annual insurance premia, respectively, and are therefore somewhat larger than the 2.5 percent reported above.

We can also compare our findings to those of Salz (2015), Hortaçsu and Syverson (2004), and Hong and Shum (2006). Salz (2015) studies the New York City trade waste market in which businesses contract with waste carters for waste disposal and finds that search costs represent between 30 percent and 50 percent of total expenses. Hortaçsu and Syverson (2004) estimate a median search cost of 5 bps, yielding a ratio of 8 percent. The average search cost across the four books considered by Hong and Shum (2006) is \$1.58 (for nonsequential search), yielding a ratio of 33 percent.

Although somewhat lower, our search cost estimates are comparable with those found in the literature, despite the fact that, because of the negotiation process, it is more complicated to obtain information about mortgage prices than about most products studied until now.

B. Goodness of Fit

We next provide a number of tests for the goodness of fit of the baseline model. To do so, we construct 300 random samples of 1,000 individuals, drawn with replacement from the main data set. The final simulated data set includes 300,000 contracts obtained using the following steps:

1. Sample individual shocks from the estimated distributions: $(c_i, w_{i1}, \dots, w_{im}, \kappa_i)$.
2. Sample borrower characteristics from the empirical distribution: $(L_i, \bar{p}_{l(i)}, x_i, h(i))$.
3. Solve the model and compute the endogenous outcomes: $(p_i^0, p_i^*, 1(\kappa_i < \bar{\kappa}_i(p^0)), b_i)$.
4. Drop consumers who failed to qualify for a loan at any bank—about 5.5 percent of consumers.

Panel A of table 5 presents summary statistics for the key endogenous outcomes of the model. The top rows summarize the observed sample, while the bottom rows summarize the simulated data. Overall, the baseline model is able to match well the interest rate spread (transaction rate minus bond rate) and monthly payments. Predicted and observed discounts are also fairly similar, but the model tends to underpredict the fraction of borrowers paying the posted rate (9.2 percent vs. 12.7 percent).

The last two columns of panel A of table 5 highlight how well the model matches aggregate lender choices. The model slightly overpredicts the fraction of loyal consumers (67 percent vs. 65 percent) and the fact that borrowers tend to choose lenders with larger than average branch networks (1.678 vs. 1.599).

Next we measure the ability of the model to fit the aggregate search moments measured from the national survey of new home buyers. On average, the baseline specification predicts that 65.7 percent of consumers reject the initial offer and search, compared to 62.5 percent in the survey. This difference is significantly different from zero only at a 10 percent significance level. We can also contrast the survey results with the predicted search probabilities from the two alternative specifications in table G.1 in appendix G, which vary the weight placed on the search moments (i.e., $\rho = 0$ vs. $\rho = 100$ in eq. [15]). When the search moments are not used in the estimation, the model tends to predict a larger search probability (69.4 percent). In contrast, by assigning a weight of $\rho = 100$ to the search moments, the model is able to reproduce almost perfectly the survey predictions (63.5 percent). We also find that the model reproduces the general patterns of the survey across regions and city sizes but tends to underestimate the amount of heterogeneity across demographics groups.

TABLE 5
SUMMARY STATISTICS FOR SIMULATED AND OBSERVED DATA

		A. NEGOTIATED PRICE AND BANK CHOICE						
		Spread (bps)	Discounts (bps)	1 (Discount = 0)	Payment (\$/Month)	1 (Loyal)	Relative Network	
Observed:								
Mean		119.5	95.3	.127	924.6	.651	1.599	
Standard deviation		59.3	45.4	.333	385.0	.477	1.015	
Simulated:								
Mean		119.4	92.2	.092	962.8	.670	1.678	
Standard deviation		62.0	53.4	.289	397.3	.470	1.136	
		B. SEARCH PROBABILITIES						
		Baseline ($\rho = 1$)	Zero Moment Weight ($\rho = 0$)	Large Moment Weight ($\rho = 100$)	Survey Data			
					Frequency	M_g		
City:								
Population > 1 million		.673	.717**	.661	.660	338		
1 million > population > 100,000		.657	.695	.639	.654	268		
Population \leq 100,000		.628**	.655***	.584	.560	275		
Regions:								
East		.626***	.656***	.582	.557	289		
West		.651	.688*	.628	.643	327		
Ontario		.673	.713	.659	.668	265		
Income:								
> \$60,000		.639***	.670***	.586	.579	400		
\leq \$60,000		.669	.712**	.670	.666	441		
Total		.657*	.694***	.635	.625	841		

NOTE.—The simulated sample is obtained by simulating 300,000 contracts from the baseline model and dropping consumers who fail to qualify for a loan (5.5 percent). Spread is defined as transaction rate minus bond rate.

* $p < .1$.

** $p < .05$.

*** $p < .01$.

The main takeaway of this simulation exercise is that the model estimated from the contract data alone tends to predict slightly more search than what the aggregate survey suggests. To understand the source of this discrepancy, it is useful to look at the ability of the model to explain the rate difference between loyal and switching consumers. As in the data, the model predicts that loyal consumers obtain lower discounts than do switching consumers, but the model predicts an even greater difference (16 vs. 9 bps). The reason is that in the model, “switching” consumers must have rejected an initial offer and must pay a competitive price. This timing restriction is probably too restrictive. In practice, the timing of moves is likely heterogeneous across consumers, in ways that we cannot identify in our data. Honka, Hortaçsu, and Vitorino (2017), for instance, consider a richer search/matching model that exploits data on search and consideration set formation.

In appendix F we provide more information on the fit of the model (including the two results discussed above). We show that the model reproduces very well the lenders’ aggregate market shares. We also evaluate the ability of the model to reproduce the reduced-form relationships observed in the data between rates, loyalty, and transaction characteristics. In general, the model does a good job of predicting the relationship between discounts and financial attributes.

VII. Search Frictions and Market Power

In this section, we use the model to quantify the effect of search frictions and market power on consumer surplus and firms’ profits. In the model, market power and search frictions are tightly linked, since lenders are able to use the initial quote to screen consumers with high search costs. We start by quantifying the welfare impact of search frictions by computing the equilibrium allocation of contracts in the absence of search costs. We then quantify the importance of market power in the industry by focusing on the incumbency advantage.

A. *Quantifying the Effect of Search Frictions on Welfare*

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

$$\begin{aligned}
\Delta CS_i &= \underbrace{\bar{v} - p_i - 1(\kappa_i < \bar{\kappa}(p_i^0))\kappa_i}_{CS_i} - \underbrace{(\bar{v} - \tilde{p}_i)}_{\tilde{CS}_i} \\
&= [\tilde{c}_{i,b(i)} - c_{i,b(i)}] - [m_{i,b(i)} - \tilde{m}_{i,b(i)}] - 1(\kappa_i < \bar{\kappa}(p_i^0))\kappa_i \\
&= \Delta V_{i,b(i)} - \Delta m_{i,b(i)} - 1(\kappa_i < \bar{\kappa}(p_i^0))\kappa_i,
\end{aligned} \tag{17}$$

where the tilde indicates the equilibrium outcomes without search cost, \bar{v} is the WTP for mortgages (policy invariant), $V_{i,b(i)} = \bar{v} - c_{i,b(i)}$ is the transaction surplus (excluding the search cost), $m_{i,b(i)} = p_i - c_{i,b(i)}$ is the profit margin, and $1(\kappa_i < \bar{\kappa}(p_i^0))$ is an indicator variable equal to one if the consumer rejects the initial offer. As before, we assume that the WTP for mortgages is large enough that the same group of consumers would enter the housing market with or without search frictions.

We label the three components *misallocation*, *discrimination*, and *search cost*, respectively. The sum of the first and third components measures the change in total welfare caused by search frictions. The discrimination component is related to the surplus split between firms and consumers.

We simulate the counterfactual experiments as before. The only difference between the baseline and the zero search cost environments is that, without search frictions, consumers do not obtain an initial quote. As a result, the posted rate becomes the reservation price in the competition stage. Table 6 presents the main simulation results. Columns 1–3 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 reports the fraction of simulated consumers experiencing zero changes, and the next four describe the conditional distribution of nonzero changes. To calculate the cumulative changes, we average the changes across all qualifying consumers. The percentage shares of each component are expressed relative to the cumulative changes.

We estimate that the cumulative reduction in consumer surplus associated with search frictions is equal to \$12.80 per month, or 2 percent of the total interest cost of mortgages in our data set. The largest component (50 percent) is attributed to the sunk cost of searching, followed by the increase in margins associated with price discrimination (28 percent) and misallocation (21 percent). Over 98 percent of consumers are affected. The sum of the misallocation and discrimination components corresponds to the effect of search frictions on monthly payments alone: \$6.37 per month, on average, per borrower. This leads to an increase in interest payments of \$503 over 5 years (col. 5), or \$1,569 for consumers who are directly affected by the price change.

The sum of the misallocation and search cost components corresponds to the total welfare cost of search frictions: \$9.15 per month per borrower.

TABLE 6
DECOMPOSING THE EFFECT OF SEARCH FRICTIONS ON WELFARE

	CONSUMER SURPLUS CHANGE: ZERO SEARCH COST			INTEREST COST CHANGE (\$) (5)	CONSUMER SURPLUS CHANGE FROM MARKET POWER (\$/Month) (6)
	Misallocation (\$/Month) (1)	Discrimination (\$/Month) (2)	Search (\$/Month) (3)		
Zero changes (%)	.83	.68	.32	.68	.00
Nonzero changes:					
Mean	-16.32	11.37	9.49	1,569	-15.12
P10	-33.44	-7.59	1.46	344	-34.24
P50	-11.70	12.43	7.86	1,591	-9.89
P90	-2.19	28.48	19.01	2,697	-1.53
Cumulative:					
\$	-2.73	3.64	6.42	503	-15.12
%	.21	.28	.50	1.00	

NOTE.—Each entry corresponds to an average over 300,000 simulated contracts. Statistics in lines 2–5 are calculated using the samples of consumers facing nonzero changes. Cumulative changes are the sum of all changes divided by the total number of qualifying consumers. The welfare decomposition in cols. 1–3 corresponds to $\Delta CS_i = \Delta V_i - \Delta m_i - \Delta \kappa_i \delta_i$. The last row reports the contribution of each component, in percentage, to the cumulative change. Col. 5 summarizes the effect of search frictions on the total interest payment over 5 years: total interest cost ($\kappa_i > 0$) – total interest cost ($\kappa_i = 0$). Col. 6 reports the further reduction in consumer surplus arising from the presence of market power in the second stage of the game: $CS(\kappa_i = 0, m_i > 0) - CS(\kappa_i = 0, m_i = 0)$.

For these two components, the fraction of zero changes measures the percentage of buyers and sellers that are matched efficiently and the fraction of nonsearchers in the presence of search frictions, respectively. Search frictions cause 17 percent of transactions to be misallocated, despite the fact that more than 32 percent of consumers do not search. Note that the difference between these two fractions would be close to zero if the loyalty advantage were null. Since banks' fixed effects are not highly dispersed, this occurs mostly because consumers visit the highest expected surplus seller first, which reduces the fraction of inefficient matches.

Focusing directly on the change in profit margins, column 2 shows that the relatively small contribution of the discrimination component is explained by the fact that some consumers pay higher markups in the frictionless market. The median change in profit margins is equal to \$12.43 per month, significantly more than the median increase in search costs (i.e., 12.43 vs. 7.86). However, the 10th percentile consumer benefits from a \$7.59 reduction in profit margins, which brings the cumulative effect down to \$3.64.

To understand this heterogeneity, recall that the initial quote is used both as a screening tool and as a price ceiling in the competition stage. The home bank is in a monopoly position in the first stage and can set individual prices based on consumers' expected outside options. This is analogous to first-degree price discrimination and strictly increases the expected profit of the home bank. This adverse effect is weighed against the fact that the initial offer can be recalled and so protects consumers against excessive market power in the competition stage. In the zero search cost environment, the price ceiling is, on average, higher (i.e., it is the posted rate), which explains why some consumers experience an increase in profit margins after eliminating search frictions.

To put these numbers in perspective, column 6 summarizes the distribution of consumer surplus changes arising from eliminating market power entirely, relative to the zero search cost environment. We calculate the difference in surplus between the zero search cost environment and one with no search frictions and zero profit margins. This is equivalent to shifting the bargaining power entirely to consumers in the competition stage and therefore maximizing the surplus of consumers. Relative to the baseline environment, eliminating market power and search frictions would increase consumer surplus by \$27.92 per month on average (i.e., $12.80 + 15.12$). Therefore, eliminating search frictions would allow consumers to reach 46 percent of their maximum surplus.

These results can be compared to those of Gavazza (2016), who performs a similar decomposition of the effect of search frictions on welfare in decentralized asset markets. Using data from the business aircraft market, he finds that, relative to his estimated model, when search costs are set to zero, welfare falls slightly (by roughly \$1 million per quarter). This

small decrease is the result of a reduction in direct search costs (of about \$6 million), a reduction of misallocation (\$3 million), and an offsetting increase in dealer costs (\$11 million).

B. Quantifying the Importance of Market Power

Overall, we find that the market is competitive. Figure 1A plots the distribution profit margins. The average profit margin is 22.1 bps, which corresponds to a Lerner index of 3.2 percent. This is consistent with our earlier findings that mortgage contracts are fairly homogeneous across lenders, and search costs represent a small share of consumers' overall mortgage spending. It is also fairly consistent with the findings in Allen et al. (2014a), which suggest margins of around 35 bps before the merger and 40 bps afterward.

This implies that a large fraction of the observed spread between negotiated rates and the 5-year bond rate corresponds to transaction costs. We estimate that each contract costs roughly 100 bps to originate, beyond the financing cost, which is proxied by the bond rate. This cost stems from a variety of sources: the compensation of loan officers (bonuses and commissions), the advantage associated with prepayment risks, transaction costs associated with the securitization of contracts, and upstream profit margins from financing.

The distribution of profit margins is also very dispersed. The coefficient of dispersion of profit margins is equal to 72 percent, and the range exceeds 100 bps. Figure 1B shows that part of this dispersion is caused by heterogeneous search efforts. On average, firms charge a markup that is 90 percent larger on consumers who are not searching (i.e., 32.1 vs. 16.9). The margin distribution for searchers also exhibits an important mass between 0 and 20 bps, and the median margin among searchers is only 13 bps (compared to 32 bps in the nonsearcher sample).

The dispersion in profit margins also reflects the fact that market power arises from a variety of sources: (i) price discrimination, (ii) loyalty advantage, (iii) observed cost differences (i.e., bank fixed effects and network size), and (iv) idiosyncratic cost differences (i.e., ω_{ij}).

The last two components ensure positive profit margins in the competition stage. On average, the difference between the lowest and second-lowest cost among rival lenders is equal to \$15.70 per month. This is the profit margin that would be realized if the home bank were not present and there were no posted rate (i.e., ceiling) and therefore can be thought of as an upper bound on the market power of rival banks. In practice, rival lenders earn slightly less: banks' average profits from switching consumers are \$14.99 per month (or 17.1 bps), compared to \$20.22 per month for loyal consumers (or 24.6 bps).

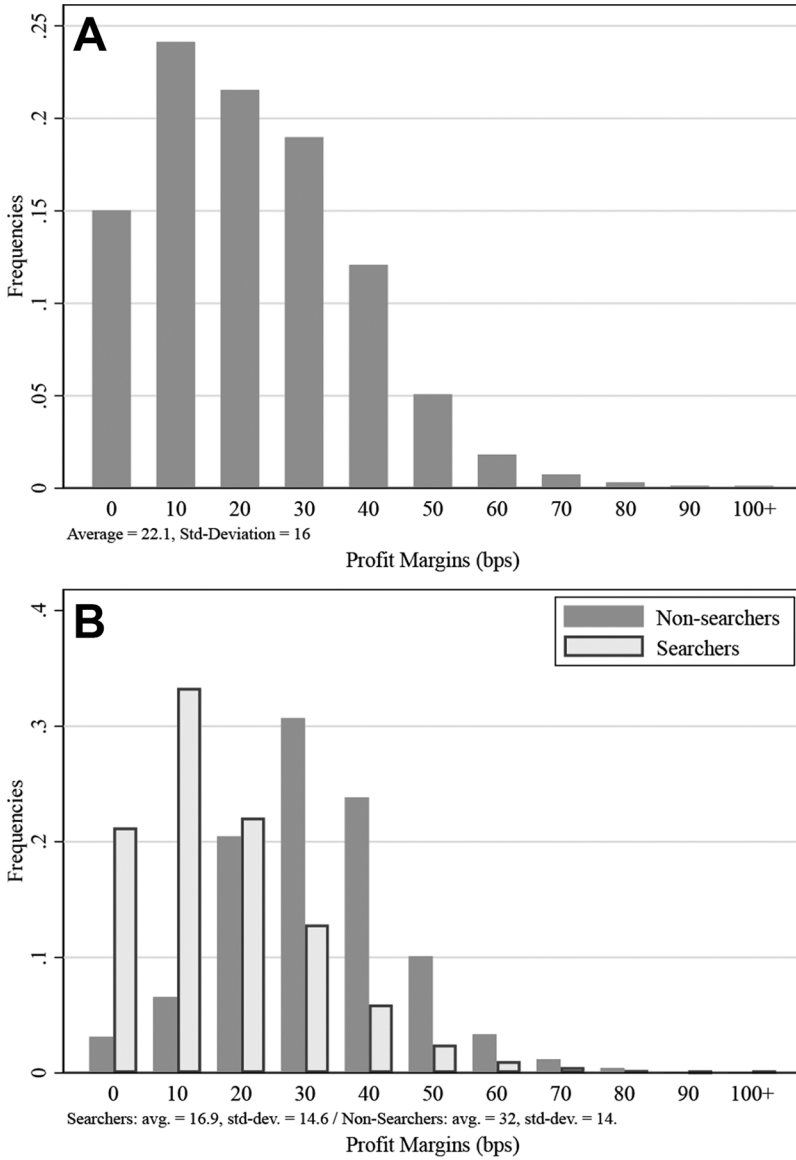


FIG. 1.—Distribution of profit margins. A, Full sample. B, Searchers and nonsearchers.

The profit gain from loyalty corresponds to an *incumbency advantage*: Banks with a large consumer base have more market power because of a first-mover advantage and loyalty advantage (or differentiation). We find that the loyalty advantage is substantial: the average home bank cost ad-

vantage is 33 percent larger than the standard deviation of idiosyncratic cost differences. As a result the home bank is able to retain, on average, 51 percent of searching consumers. The first-mover advantage arises because the home bank is in a quasi monopoly position in the first stage of the game and can price discriminate between consumers on the basis of heterogeneity in their expected reservation prices. The ability to make the first quote allows the home bank to charge a higher markup and retain a larger fraction of consumers who, without search costs, would choose another lender.

To measure the source and magnitude of the incumbency advantage, we use the simulated model to evaluate the correlation between the size of a lender's consumer base and its profitability. In the model, the consumer base of a given bank is defined as the share of consumers with whom it has an existing day-to-day banking relationship, and this base determines the fraction of consumers in a given market who start their search with the bank (i.e., ψ_{ij}). Recall that this matching probability is defined at the level of a neighborhood (FSA), income group, and year. We use this definition to construct markets that each have a homogeneous consumer base distribution, and we construct measures of profits and concentration at this level of aggregation. Doing so yields slightly more than 8,000 unique markets.

To construct a measure of consumer base that is comparable across markets, we compute, for each market i , the ratio of the matching probability of lender j over the average matching probability among banks in the market:

$$\text{Matching probability ratio} = \frac{\psi_{ij}}{\bar{\psi}_i} = \frac{\psi_{ij}}{n_i + 1}.$$

Panel A of table 7 summarizes the distribution of contracts and profits across different types of lenders. The table ranks banks from the smallest consumer base (i.e., between 0 and 25 percent of the average size in the same market) to the largest (i.e., between four and seven times the average size). As we saw earlier, most consumers choose a mortgage lender with a large branch presence. This is reflected in the distribution of contracts shown in column 1: 46 percent of contracts are issued by banks with a consumer base between one and two times larger than the average bank in their market.

Columns 2 and 3 report the weighted average share of profits and contracts generated by each bank type. To get this number, for each market, we calculate the average share of profits and contracts generated by lenders with consumer bases belonging to one of the six categories. We then aggregate these shares across markets, using the total number of contracts originated in each market as weight.

If there were no relationship between banks' consumer bases and mortgages, contracts and profits would be uniformly distributed across catego-

TABLE 7
INCUMBENT ADVANTAGE AND MARKET POWER

MATCHING PROBABILITY RATIO	SAMPLE FREQUENCY (1)	WITHIN MARKET SHARES			SECOND-STAGE PROFITS (%) (4)	MARGINS (bps) (5)
		Profits (2)	Contracts (3)			
0-1/4 (small consumer base)	.05	.02	.02	.90	16.6	
1/4-1/2	.04	.04	.05	.62	19.1	
1/2-1	.17	.07	.08	.51	18.9	
1-2	.46	.16	.16	.50	20.4	
2-4	.25	.34	.30	.51	24.0	
4-7 (large consumer base)	.04	.62	.54	.40	30.7	

B. DISTRIBUTION OF BANK PROFITABILITY IN THE BASELINE AND COUNTERFACTUAL ENVIRONMENTS

STATISTICS AND VARIABLES	Baseline (1)	CF-1: $\Delta_i = 0$	CF-2: $\psi_i = 1/(n+1)$	CF-3: $\psi_i = 1/(n+1)$ and $\Delta_i = 0$
		(2)	(3)	(4)
Ratio: large base/small base:				
Margins (bps)	1.851	1.369	1.485	1.145
Profit shares	35.717	11.652	17.159	6.699
Contract shares	24.582	9.769	13.204	6.243
Full sample averages:				
Search probability	.656	.774	.838	.822
2nd-stage profits (%)	.531	.727	.809	.784
Margins (bps)	22.07	18.56	21.34	18.60
Matching probability ratio	1.709	1.546	1.605	1.439

C. DECOMPOSITION OF THE INCUMBENCY ADVANTAGE

LARGE/SMALL RATIO	Incumbency Advantage:			
	Loyalty Premium: Base - CF-3 (1)	CF-2 - CF-3 (2)	Price Discrimination: CF-1 - CF-3 (3)	Interaction (4)
Margins	.707	.340 (.48)	.224 (.32)	.142 (.2)
Profit share	29.018	10.460 (.36)	4.954 (.17)	13.605 (.47)
Contract share	18.339	6.961 (.38)	3.526 (.19)	7.852 (.43)

NOTE.—Each entry in panel A is the weighted average outcome of lenders belonging to each category (rows). The weights are proportional to the number of contracts originated in each market (i.e., neighborhood/year/income). Variable definitions: matching probability ratio = consumer base of bank j /average consumer base; sample frequency = market share of lenders in each category; second-stage profit (%) = average share of profits originating from the searching consumers; within market share = average share of profits or contracts generated by lenders in each category; margins are presented in percentage basis points; ratio: large base/small base = ratio of the average outcomes of lenders in the large group over those in the small group. Counterfactual environments: (1) zero home bank cost advantage; (2) uniform matching probability; (3) combination of 1 and 2.

ries (i.e., would be about 12 percent on average). The resulting distributions are significantly more concentrated. Banks in the top category (4–7) earn, on average, 62 percent of the profits generated in their respective markets, compared to only 2 percent, on average, for the smallest banks. Note that the average profit share increases very quickly with the size of the consumer base.

In addition, the distribution of profits is more concentrated than the distribution of contracts. On average, the top lenders originate 54 percent of contracts but earn 62 percent of the profits. This pattern reflects the fact that banks with a large consumer base charge, on average, higher markups. Column 5 shows that the average profit margin for banks in the top category is equal to 30.7 bps, compared to only 16.6 bps for banks in the bottom category. This discrepancy is largely explained by the difference in markups between searchers and nonsearchers. Banks in the smallest consumer base category earn, on average, 90 percent of their profits from consumers reaching the second stage of the game, compared to 40 percent for banks in the largest category. This confirms the importance of the first-mover advantage as a source of market power for large consumer base lenders.

Identifying the relative importance of the first-mover advantage and differentiation is not an easy task, however, since the two interact to generate a correlation between profitability and size of consumer base. For instance, the profit gain from being able to make the first offer depends on the amount of differentiation, since lower-cost banks have more leverage in the initial negotiation stage. Similarly, the presence of a cost advantage reduces the incentive for consumers to search and increases the fraction of profits generated from price discrimination.

To measure each of the components that generate the incumbency advantage, we simulate a series of counterfactual experiments aimed at varying the first-mover advantage and the differentiation component independently. In particular, to eliminate the differentiation component, CF-1 simulates a model in which the cost advantage of the home bank is set to zero, which is analogous to separating the provision of mortgages from other banking services. We eliminate the ability of firms to screen high-search cost consumers by imposing a uniform matching probability and breaking the link between the ability to make the first offer and the size of the consumer base (CF-2). Finally, CF-3 combines the previous two environments by assuming a uniform matching probability and zero loyalty advantage.²⁷

²⁷ An alternative approach for eliminating the first-mover advantage is to set consumer search costs to zero. We chose instead to modify the order of moves by setting $\psi_{ij} = 1/(n + 1)$, since doing so does not fundamentally change the degree of competition in the market. The zero search cost counterfactual yields very similar conclusions. Results are available from the authors on request.

Results are displayed in panel B of table 7. The top rows summarize concentration in the industry as well as the dispersion in profit margins between large and small banks. The bottom rows describe some of the key equilibrium outcomes in the baseline and counterfactual environments. The ratio of the profit margin of large and small banks is a measure of the incumbency advantage: how much more market power banks with large consumer bases have relative to banks with small consumer bases. In the baseline environment, we estimate that large banks' profit margins are 85.1 percent larger. Eliminating the first-mover and the loyalty advantage shrinks the margin difference to 14.5 percent (CF-3), and so this is a measure of the market power of large banks that stems solely from brand and branch network differences.²⁸ The difference, or $0.707 = 1.851 - 1.145$, is explained by the incumbency advantage.

Column 1 of panel C of table 7 summarizes the incumbency advantage in terms of profit margins, profit shares, and market shares (or contract). Columns 2–4 use the uniform matching probability (CF-2) and the zero loyal advantage (CF-1) counterfactual environments to decompose the incumbency advantage into three terms:

$$\underbrace{\text{Incumbency advantage}}_{0.707} = \underbrace{\text{Loyalty advantage}}_{0.34} + \underbrace{\text{Price discrimination}}_{0.22} + \underbrace{\text{Interaction}}_{0.14}.$$

Therefore, relative to CF-3, almost 50 percent of the market power of large banks is caused by the home bank cost advantage, just over 30 percent by the first-mover advantage, and the remaining 20 percent by the interaction of both elements.

The interaction term originates from the joint equilibrium effect of differentiation and the order of moves on the search probability. As the middle row indicates, the combined effect of the home bank cost and first-mover advantage is to lower the search probability from 82.2 percent to 65.6 percent, which increases the profit margin ratio by an extra 14 percentage points through a change in the composition of loyal borrowers. Independently, the two elements have little or no effect on the search probability relative to the CF-3 environment.

The concentration of profits and contracts is similarly affected. Eliminating both the loyalty advantage and the first-mover advantage substantially reduces the concentration of profits: large banks' share of profits is 35.72 times larger than that of small banks in the baseline, compared to only 6.70 times in CF-3. As with margins, the loyalty advantage alone explains a bigger share of the drop in concentration (36 percent) than the first-mover advantage (17 percent). However, in contrast to the case with

²⁸ All ratios would be equal to one if the difference between lenders were caused only by idiosyncratic cost differences.

margins, a larger portion of the profit share ratio is explained by the interaction of differentiation and discrimination: 47 percent of the profit share difference between large and small banks is explained by the interaction term. The reason is that the increase in the search probability from letting the most efficient lender make the first offer has a very large effect on banks' retention probability and therefore on their overall profitability.

VIII. Conclusion

The paper makes three main contributions. The first is to provide an empirical framework for studying markets with negotiated prices. 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. Finally, the paper also demonstrates the importance of having a large consumer base for market power and decomposes the effect into a first-mover advantage and brand loyalty. We find that brand loyalty is the most important source of market power, but search frictions play an important role through the first-mover advantage.

A few caveats should be mentioned. First, the assumption that monthly payment has no effect on the loan size could imply (depending on the elasticity of loan demand) that the distortions arising from search and market power are larger than the ones we calculate. Second, although the overall fit of our model is good, it predicts that loyal consumers pay more than they are observed to in the data. This difference is directly related to 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.

The model also overestimates the impact of competition on rates, likely because market structure is assumed to be independent of consumers' unobserved attributes, up to regional fixed effects. Otherwise, our estimates of firms' cost differences would suffer from an attenuation bias, and our results would correspond to a lower bound on margins. A related interpretation of the small reduced-form effect of competition on rates is that consumers face heterogeneous consideration sets, conditional on being located in the same area. This would create measurement error in consumer choice sets, which is computationally prohibitive to incorporate, since lenders are *ex ante* heterogeneous. Moreover, we do not have data on the set, or identity, of lenders considered by borrowers.

References

- Alexandrov, A., and S. Koulayev. 2017. "No Shopping in the U.S. Mortgage Market: Direct and Strategic Effects of Providing Information." Working Paper no. 2017-01 (April), Consumer Financial Protection Bur., Washington, DC.

- Allen, J., R. Clark, and J. F. Houde. 2014a. "The Effect of Mergers in Search Markets: Evidence from the Canadian Mortgage Industry." *A.E.R.* 104:3365–96.
- . 2014b. "Price Dispersion in Mortgage Markets." *J. Indus. Econ.* 62:377–416.
- Beckert, W., H. Smith, and Y. Takahashi. 2016. "Competition and Price Discrimination in a Spatially Differentiated Product Market with Negotiated Prices." Working paper, Johns Hopkins Univ.
- Bester, H. 1988. "Bargaining, Search Costs and Equilibrium Price Distributions." *Rev. Econ. Studies* 55 (1): 201–14.
- Brannan, L., and L. M. Froeb. 2000. "Mergers, Cartels, Set-Asides, and Bidding Preferences in Asymmetric Oral Auctions." *Rev. Econ. and Statis.* 82 (May): 283–90.
- Bronnenberg, B. J., and J. Dubé. 2017. "The Formation of Consumer Brand Preferences." *Ann. Rev. Econ.* 9:353–82.
- Busse, M., J. Silva-Risso, and F. Zettelmeyer. 2006. "\$1,000 Cash Back: The Pass-Through of Auto Manufacturer Promotions." *A.E.R.* 96 (4): 1253–70.
- Chatterjee, K., and C. C. Lee. 1998. "Bargaining and Search with Incomplete Information about Outside Options." *Games and Econ. Behavior* 22:203–37.
- Cicala, S. 2015. "When Does Regulation Distort Costs? Lessons from Fuel Procurement in US Electricity Generation." *A.E.R.* 105:411–44.
- Crawford, G., and A. Yurukoglu. 2012. "The Welfare Effects of Bundling in Multichannel Television." *A.E.R.* 102 (2): 643–85.
- Dafny, L. 2010. "Are Health Insurance Markets Competitive?" *A.E.R.* 100 (4): 1399–1431.
- Dell’Ariccia, G., E. Friedman, and R. Marquez. 1999. "Adverse Selection as a Barrier to Entry in the Banking Industry." *RAND J. Econ.* 30:515–34.
- De los Santos, B., A. Hortaçsu, and M. Wildenbeest. 2012. "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior." *A.E.R.* 102:2955–80.
- Doornik, J. A. 2007. *Ox—an Object-Oriented Matrix Programming Language*. Richmond, UK: Timberlake Consultants Press and Oxford Univ.
- Duffie, D., N. Garleanu, and L. H. Pedersen. 2005. "Over-the-Counter Markets." *Econometrica* 73 (6): 1815–47.
- French, E., and C. Taber. 2011. "Identification of Models of the Labor Market." In *Handbook of Labor Economics*, vol. 4A, edited by O. Ashenfelter and D. Card, chap. 6. Amsterdam: Elsevier.
- Fudenberg, D., and M. Villas-Boas. 2007. "Behavior-Based Price Discrimination and Customer Recognition." In *Economics and Information Systems*, vol. 1, edited by T. Hendershott. Oxford: Elsevier Sci.
- Gavazza, A. 2016. "An Empirical Equilibrium Model of a Decentralized Asset Market." *Econometrica* 84 (5): 1755–98.
- Goldberg, P. K. 1996. "Dealer Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey." *J.P.E.* 104 (3): 622–54.
- Gowrisankaran, G., A. Nevo, and R. Town. 2015. "Mergers When Prices Are Negotiated: Evidence from the Hospital Industry." *A.E.R.* 105:172–203.
- Grennan, M. 2013. "Price Discrimination and Bargaining: Empirical Evidence from Medical Devices." *A.E.R.* 103 (1): 145–77.
- Ho, K., and R. Lee. 2017. "Insurer Competition in Health Care Markets." *Econometrica* 85:379–417.
- Hong, H., and M. Shum. 2006. "Using Price Distributions to Estimate Search Costs." *RAND J. Econ.* 37 (2): 257–75.
- Honka, E. 2014. "Quantifying Search and Switching Costs in the U.S. Auto Insurance Industry." *RAND J. Econ.* 45:847–84.
- Honka, E., A. Hortaçsu, and M. A. Vitorino. 2017. "Advertising, Consumer Awareness, and Choice: Evidence from the U.S. Banking Industry." *RAND J. Econ.* 48:611–46.

- Horn, H., and A. Wolinsky. 1988. "Bilateral Monopolies and Incentives for Merger." *RAND J. Econ.* 19 (3): 408–19.
- Hortaçsu, A., and C. Syverson. 2004. "Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds." *Q.J.E.* 119:403–56.
- Imbens, G., and T. Lancaster. 1994. "Combining Micro and Macro Data in Micro-econometric Models." *Rev. Econ. Studies* 61 (October): 655–80.
- Jindal, P., and P. Newberry. 2018. "To Bargain or Not to Bargain: The Role of Fixed Costs in Price Negotiations." *J. Marketing Res.* 55 (6): 832–51.
- Joskow, P. L. 1987. "Contract Duration and Relationship-Specific Investments: The Case of Coal." *A.E.R.* 77:168–85.
- Lewis, M., and K. Pflum. 2015. "Diagnosing Hospital System Bargaining Power in Managed Care Networks." *American Econ. J.: Econ. Policy* 7 (February): 243–74.
- Marshall, G. 2016. "Search and Wholesale Price Discrimination." Working paper, Univ. Illinois.
- Petriconi, S. 2015. "Bank Competition, Information Choice and Inefficient Lending Booms." Working paper. <https://ssrn.com/abstract=2705043>.
- Postel-Vinay, F., and J.-M. Robin. 2002. "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica* 70 (6): 2295–2350.
- Ridder, G., and R. Moffitt. 2007. "The Econometrics of Data Combination." In *Handbook of Econometrics*, vol. 6B, edited by J. J. Heckman and E. E. Leamer, 5469–5547. Amsterdam: Elsevier.
- Salz, T. 2015. "Intermediation and Competition in Search Markets: An Empirical Case Study." Working paper, New York Univ.
- Scott-Morton, F., F. Zettelmeyer, and J. Silva-Risso. 2001. "Internet Car Retailing." *J. Indus. Econ.* 49 (4): 501–19.
- Sorensen, A. 2001. "An Empirical Model of Heterogeneous Consumer Search for Retail Prescription Drugs." Working Paper no. 8548 (October), NBER, Cambridge, MA.
- Town, R., and G. Vistnes. 2001. "Competition in Networks: An Analysis of Hospital Pricing Behavior." *J. Health Econ.* 20:719–53.
- van den Berg, G., and B. van der Klaauw. 2001. "Combining Micro and Macro Unemployment Duration Data." *J. Econometrics* 102:271–309.
- White, H. 1982. "Maximum Likelihood Estimation of Misspecified Models." *Econometrica* 50:1–25.
- Wildenbeest, M. R. 2011. "An Empirical Model of Search with Vertically Differentiated Products." *RAND J. Econ.* 42 (4): 729–57.
- Wolinsky, A. 1987. "Matching, Search, and Bargaining." *J. Econ. Theory* 42:311–33.
- Woodward, S., and R. E. Hall. 2012. "Diagnosing Consumer Confusion and Sub-optimal Shopping Effort: Theory and Mortgage-Market Evidence." *A.E.R.* 102 (December): 3249–76.