The role of intermediaries in selection markets: Evidence from mortgage lending*

Jason Allen[a], Robert Clark[b], Jean-François Houde[c], Shaoteng Li[d], & Anna Trubnikova[e]

Abstract

In this paper we study the role of intermediaries (brokers) in the Canadian mortgage market. In this market consumers can search for quotes in one of two ways: on their own, or via a broker. We provide descriptive evidence that borrowers who transact through brokers are different from those who do not. Broker-clients: (i) finance larger loans, (ii) are more leveraged, and (iii) are less creditworthy. We investigate two explanations for these observations: (i) brokers steer borrowers towards products that are more profitable for them and (ii) borrowers have (unobserved) preferences for riskier loans. We build and estimate a model of mortgage demand that accounts for, and disentangles the explanations for, the differences in product choices across origination channels: selection on observables, selection on unobservables, and broker steering. We find that brokers steer about 15% of borrowers to mortgages with longer amortization, while a borrower's own (unobservable) characteristics drive their decision for smaller down payments.

*This version: February 1, 2023. Correspondence to [a]Jason Allen: Bank of Canada, Ottawa, Ontario, K1A 0G9, email: jallen@bankofcanada.ca; [b]Robert Clark: Queen's University, 94 University Avenue, Kingston, Ontario, K7L 3N6, email: clarkr@econ.queensu.ca; [c]Jean-François Houde: University of Wisconsin, 1180 Observatory Dr. Madison WI, 53706, e-mail: houdejf@icloud.com; [d]Shaoteng Li: Institute for Economic and Social Research, Jinan University, 601 West Huangpu Road, Guangzhou, Guangdong, 510632, email: lis@econ.queensu.ca; [e]Anna Trunikova: Cornerstone, email: anna.trubnikova@cornerstone.com. This research has benefited from the financial support of the NSF (SES-1024840). We thank Claudia Robles-Garcia for excellent discussion and seminar participants at SITE Financial Regulation 2020 and the Bank of Canada. We are also grateful to Michael Boutros, Kim Huynh, and Alex Scherbakov. We also thank the Canada Mortgage and Housing Corporation for providing us with the data. 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.

1 Introduction

Intermediaries serve an important role in the economy by facilitating trade between firms and consumers. In decentralized markets, where search and negotiation are required for exchange, intermediaries can benefit consumers by effectively lowering their costs of searching for a firm with which to trade (Rubinstein and Wolinsky (1987), Gavazza (2016), Biglaiser et al. (2020), Donna et al. (2021), and Salz (2022)). At the same time, however, outsourcing search and negotiation creates an agency problem between the consumer and intermediary. For instance, Egan (2019) describes how intermediaries distort households' investment decisions for convertible bonds. Similar concerns arise for other investment products (Bergstresser et al. (2008), Christoffersen et al. (2013), Chalmers and Reuter (2020), and Egan et al. (2022)), as well as insurance (Schneider (2012) and Anagol et al. (2017)) and real estate (Jia-Barwick et al. (2017)).

This paper concentrates on the mortgage market, where a number of papers have documented conflicts of interest between borrowers and intermediaries—mortgage brokers. LaCour-Little (2009), Robles-Garcia (2020), and Guiso et al. (2022) point out agency problems that exist between brokers and borrowers in various settings. In the Canadian mortgage market that is the focus of this paper, borrowers transacting through brokers are observed to take on riskier mortgages—mortgages with higher loan-to-value (LTV) ratios and longer amortization. Such loans are costlier for the borrower over the lifetime of the mortgage, but, if broker compensation is tied to loan size and amortization length, they may have a financial incentive to steer borrowers towards these types of products.[1]

Importantly, in credit markets there may be an additional reason for borrowers to end up with riskier loans of the sort just described. Unlike in consumer-goods markets, suppliers in credit markets screen borrowers for profitability and often reject applicants (see Marquez (2002), Grigsby et al. (2020), and Argyle et al. (2020)). Generating quotes can therefore be difficult for unprofitable consumers—those, for example, with high prepayment risk and low liquidity.[2] As a result, some

[1] A typical broker earns between 0.5% and 1.2% of the value of the mortgage at origination, and future compensation can include trailer fees, which are proportional to outstanding balances and reward brokers periodically for the duration that a borrower stays with the same lender. All compensation is paid by the lender.

[2] When borrowers prepay, the lender loses out on the interest payments on the part of the principal that is prepaid, and therefore borrowers with high prepayment risk are less profitable. Borrowers with low levels of liquidity are also less profitable, but for different reasons. Individuals with low savings are less likely to be able to invest in high-margin investment products that banks try to cross-sell.

may need to search over a broad set of lenders in order to find one willing to offer a product that allows them to qualify for their desired credit level. To the extent that consumers anticipate the cost of searching and negotiating, less profitable consumers are more likely to use a broker to help them with this more extensive search, thereby rationalizing the observation that broker-clients wind up with riskier products.

The objective of this paper is to study the role that mortgage brokers play in helping borrowers generate quotes and qualify for credit, while being simultaneously incentivized to steer them to riskier products. Determining whether brokers are steering borrowers or simply helping them generate quotes and qualify is crucial for policy makers. For instance, partial responsibility for the U.S. housing bubble and subsequent collapse during the Global Financial Crisis was attributed to the conduct of brokers (c.f. Berndt et al. (2010) and Jiang et al. (2014)).[3] Establishing the appropriate policy response to ensure there is not a repeat of these events depends on what exactly was the function of brokers. If they are steering borrowers to risky products, this suggests a role for regulation. However, if brokers help with search, negotiation, and approval, then they are providing a valuable service to borrowers who might not otherwise be able to purchase a home.

To assess these different explanations we take advantage of a comprehensive dataset containing detailed information on the universe of insured contracts from the Canada Mortgage and Housing Corporation (CMHC), which is the primary mortgage insurer in the country. The dataset comprises information about household/mortgage characteristics, including the rate and chosen lender, about demographic characteristics such as income and source of downpayment, and, most importantly, about which channel the borrower used to arrive at the chosen lender—own search or broker. We complement this dataset with data from Payments Canada and the Canadian Association of Accredited Mortgage Professionals (CAAMP), which provide information on the geographic locations of bank branches and brokers, respectively. Our focus is on the 2005 to 2007 period during which there were several macroprudential policy changes that affected the set of available mortgage products. These regulatory changes helps to identify demand and the extent of steering in Section 5.[4]

[3] For an example of the popular press narrative see https://www.wsj.com/articles/SB117997159688112929.

[4] Macroprudential policies, whereby mortgage-lending guidelines are set by a regulator, have been more widely

We start by providing evidence that consumers transacting through brokers choose different products than consumers transacting directly through banks. Broker-clients choose riskier mortgages—those with higher LTV ratios and longer amortization. We also document that brokers, on average, negotiate lower interest rates than direct negotiation, conditional on consumer characteristics. At the same time, the market appears to exhibit adverse selection into the broker channel. Consumers who use brokers tend to have lower income and credit scores relative to consumers who negotiate directly with a lender. The observable consumer characteristics, however, cannot fully explain the differences in product choices. We provide two possible explanations. First, brokers might steer consumers towards products that are ultimately more profitable to them—*broker-steering channel.* Alternatively, consumers that use a broker may be riskier or less profitable to lenders than other consumers (which is unobserved by the econometrician) because they have a high likelihood of early repayment or have fewer liquid assets for which banks can cross-sell investment opportunities—*selection on unobservables.*

To separately identify selection on observables, selection on unobservables, and broker steering, we develop and estimate a model of mortgage-product choice. In the model, consumers hold savings and carry non-mortgage debt and must make a discrete choice over mortgage products characterized by two features: amortization length and LTV. Consumers maximize indirect utility, which consists of expected monthly consumption and other non-monetary payoffs. We allow for random coefficients on lump-sum payments and total interest costs to capture heterogeneity in consumer savings and prepayment risks. The choice of LTV captures the trade-off between lump-sum payment and monthly consumption, while the amortization choice reflects a trade-off between total interest cost and monthly consumption. Mortgage qualification is incorporated through a total-debt-service (TDS) constraint.

Estimation proceeds in two steps. Our model of product choice takes as given the decision of whether or not to use a broker and the interest rate paid by the borrower. However, unobserved heterogeneity in consumer savings and prepayment risks can endogenously affect mortgage rates,

implemented since the Global Financial Crisis. These include restricting the maximum allowable loan-to-value (LTV), payment-to-income (PTI) ratio, total-debt-service (TDS) ratio, loan-to-income (LTI) ratio, etc. For Canada, see Allen et al. (2020). For a broad review of macroprudential policies, see Claessens (2015). We focus on the short period of loosening since there are many fewer potentially confounding factors than there were during and post-GFC.

origination channel, and product choices. Therefore, we first use a control function approach, similar to Adams et al. (2009), Crawford et al. (2018), and Ioannidou et al. (2022), to take into account this selection issue.[5] Specifically, we construct control residuals from first-stage regressions of mortgage rates and broker usage using instrumental variables related to local market structure (number of brokers and lenders, share of bank branches that do not accept broker-intermediated transactions, etc.).[6] In the second step, we estimate the parameters in the product choice model via maximum likelihood using as input the control residuals generated in the first stage. We maximize the joint likelihood of observing a particular product choice and the level of non-mortgage debt held by the borrower. The main parameters of interest are the random coefficients reflecting the heterogeneity in marginal disutility from lump-sum payments, the marginal disutility from total interest costs, and channel-specific product fixed effects.

Using our model estimates we investigate whether brokers steer consumers towards riskier products. The channel-specific product fixed effects represent consumer preferences other than their tastes over observable mortgage characteristics (i.e., monthly payment, lump-sum payments, and total interest costs). We test the null hypothesis that these product fixed effects are not different across channels. We reject the null, suggesting that brokers do indeed influence borrowers' preferences over mortgage products and hence their choices.

Having established that brokers engage in steering, we next quantify its importance relative to selection on observables and unobservables in explaining the differences in product choices across channels. For this we turn to counterfactual analysis in which we gradually shut down different mechanisms influencing product choice. Setting the product fixed effects equal across channels simulates the elimination of steering. In this scenario, consumers are 5.73 p.p. (39.3% of the total difference) less likely to choose products with amortizations longer than 25 years, and 2.89 p.p. (40.6% of the total difference) less likely to choose a zero down-payment product. Next, we remove the control function variables in the random coefficients so that broker borrowers' unobserved

[5]In Appendix B we propose a model based off of Allen et al. (2019) that rationalizes the control variables constructed to deal with endogeneity of origination-channel choice and the corresponding rates. See Wooldridge (2015) for a discussion of control functions and how they relate to IV estimation.

[6]Some lenders do not originate mortgages through the broker channel. Lenders face a trade-off between a higher volume generated by accepting broker business and an increase in costs, since they have to compensate brokers for these loans.

savings and prepayment risks are closer to those of branch borrowers. Doing so allows us to quantify the importance of selection on unobservables. We find that selection on unobservables explains about 23% of the difference in the share of longer-amortization products and 66% for zero down-payment products.

Overall, we find that, holding observable characteristics fixed, a significant fraction (28.21%) of broker-clients choose a different, and on average costlier, product than they would if they instead used a branch. However, only about half of this is due to steering. We find that 14.57% of broker-clients make different choices once steering is eliminated. In particular, broker steering plays an important role in amortization choice. Focusing on the affected borrowers, we find that eliminating steering would, on average, shorten their amortization length by 4.5 years, lower the outstanding balance by $4,505 after five years of origination, and reduce the interest cost expenses over the entire amortization period by $32,449.[7]

In contrast, steering is not the main reason that broker-clients are more likely to select high-LTV products, as the difference in LTV choices across channels is mainly driven by selection on unobservables. Our findings suggest that, if there were no selection on unobservables, 19.54% of broker-clients would choose different products, on average lowering their loan amount by $9,048, amortization length by 2.8 years, and total interest cost by $36,021. These clients, however, are not steered to risky products, but prefer them. They choose to use a broker to generate quotes and qualify for larger loans, consistent with behavior described in Grigsby et al. (2020).

Our findings have important policy implications. In a number of jurisdictions, including in Canada, regulations have been put in place affecting intermediary incentives. Our findings suggest that, although some movement towards costlier products is the result of steering, not all is, and so regulating broker activity could in fact harm consumers. That being said, our evidence on broker steering, especially as it relates to amortization, implies that borrowers would likely benefit from increased transparency. In Canada, several provincial regulators have introduced policies aimed at increasing transparency. For example, Alberta and Ontario have regulation requiring that brokers

[7]On the other hand, these borrowers benefit from the more competitive rates in the broker channel, which is on average 10.6 bps lower than branch rates. If we fix the product choices, the lower rates translate to an average savings of $4,564 in total interest costs.

disclose information on compensation.[8] Regulation in British Columbia requires brokers to disclose any direct or indirect conflict of interest (e.g., upfront commissions, volume bonuses, reward points, trailer fees, etc.) related to mortgage transactions.[9]

Literature Review In addition to LaCour-Little (2009), Robles-Garcia (2020), and Guiso et al. (2022) mentioned above, Hall and Woodward (2012), Allen et al. (2014b), Jiang et al. (2014), and Myśliwski and Rostom (2022) study the role of brokers in mortgage markets. However, none focus on the role of brokers in helping borrowers generate quotes and qualifying for a loan. The role of intermediaries in other financial markets has been studied—for example, Bar-Isaac and Gavazza (2015) in the Manhattan rental market and Egan (2019) in the market for financial advisors. There is also a growing literature structurally examining the impact of intermediaries on outcomes (in addition to Gavazza (2016) and Salz (2022), see for instance Biglaiser et al. (2020), Donna et al. (2021), and Grunewald et al. (2021)).

We are also related to a series of papers on financial advice, namely Foerster et al. (2017), Egan et al. (2019), Charoenwong et al. (2019), Foà et al. (2019), and Bhattacharya et al. (2021), and to a large and growing literature using structural methods to study credit markets.[10] Allen et al. (2014a), Allen et al. (2019), and Allen and Li (2020) focus on the mortgage-search and negotiation process in order to explain the observed dispersion in transaction rates. Benetton (2021) and Benetton et al. (2019) study the UK market. Considerable attention has been paid to the US mortgage market (c.f. Alexandrov and Koulayev (2017), Aguirregabiria et al. (2020), and Buchak et al. (2020)). Clark et al. (2021) provide a review of this research, as well as related work on other credit products, such as sub-prime auto loans (Adams et al. (2009)), auto insurance (Honka (2014) and Braido and Ledo (2018)), and small business lending (Crawford et al. (2018)).

The rest of the paper is organized as follows. Section 2 describes the Canadian banking market,

[8]See Section 21 of Ontario Regulation 188/08 (`https://www.ontario.ca/laws/regulation/080188/v3`) and Section 65 of the Alberta Real Estate Act Rules (`https://www.reca.ca/wp-content/uploads/2022/08/Rules-2022-08-19.pdf`)

[9]See Section 17.3 of Mortgage Broker Act (`https://www.bclaws.gov.bc.ca/civix/document/id/complete/statreg/96313_01`) and the related guidelines (`https://www.bcfsa.ca/media/2613/download`).

[10]Related to financial advice, researchers have also studied the role of advertising in mortgage markets. For example, Gurun et al. (2016) find that sub-prime lenders that advertise more within a region sell more expensive mortgages; Grundl and Kim (2019) document heterogeneous effects of advertising on borrowers' decisions to refinance.

including the presence of mortgage brokers. Section 3 introduces our datasets and presents a descriptive analysis of the data. Section 4 presents the model. Section 5 discusses the estimation strategy. Sections 6 and 7 describe the empirical results. Section 8 concludes.

2 The Canadian mortgage market and the role of intermediaries

2.1 Insurance, qualification, and macroprudential changes

The Canadian mortgage market features two types of contracts—conventional, which have a low LTV ratio and, hence, are uninsured, and high-ratio, which are high LTV and require insurance (for the full amortization period of the mortgage). During our sample period—2005 to 2007—roughly 80% of new home buyers required mortgage insurance. The primary insurer is Canada Mortgage Housing Corporation (CMHC), a Crown corporation with an explicit guarantee from the federal government. During our sample period a private firm, Genworth Financial also provided mortgage insurance and had a 90% government guarantee. The market share of CMHC during our sample period averages around 70%. Both insurers use the same guidelines for insuring mortgages and charge the lenders an insurance premium, ranging from 1% to 3.7% of the value of the loan, which is passed on by lenders to borrowers.

The government sets the rules for mortgage-insurance qualification. These rules include a maximum LTV and debt-service ratio, a minimum credit score, as well as a maximum amortization. Importantly, some of these rules changed during our sample period to facilitate homeownership. For example, prior to 2006 the maximum allowable amortization was 25 years. In 2006 the federal government sequentially expanded the insurance to mortgages with longer maturities. The maximum amortization period was increased to 30 years in March, then to 35 years in June, reaching its peak of 40 years in December 2006.[11] Near the time of the last change, the maximum LTV was increased from 95% to 100%, thus allowing for zero down-payment loans. Insurance premiums depend on LTV and amortization.[12] These rules were tightened starting in 2008 in response to the

[11] The maximal amortization was subsequently decreased to 35 years in October 2008 in the midst of the financial crisis.

[12] For mortgages with a 25-year amortization, depending on LTV (80%, 85%, 90%, 95%, 100%), the insurance premium is 1%, 1.75%, 2%, 2.75%, and 3.1%, respectively. Each 5-year amortization extension increases the premium

U.S. house-price correction and concerns that the Canadian market was too lax.

Finally, in addition to the mortgage-insurance qualification criteria, lenders can also screen borrowers for profitability. Borrowers are unprofitable if they are likely to prepay their mortgage (high prepayment risk) or have few liquid assets. The former case has been well documented—lenders lose out on the interest payments related to the part of the principal that is prepaid, and so borrowers with high prepayment risk are less profitable.[13] Borrowers with low levels of liquidity are also less profitable because the opportunity to cross-sell high-margin investment products is diminished. For example, Canada has some of the highest mutual fund fees in the world, c.f. Ruckman (2003) and Khorana et al. (2008), and these are sold by the same bank branches that originate mortgages. In contrast, because mortgage brokers only deal in mortgages, cross-selling is rare in the broker channel.

2.2 Market structure

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, operating in Quebec), and a provincially owned deposit-taking institution (Alberta's ATB Financial). Collectively, they control 90% of banking industry assets. A number of smaller banks (ING, HSBC, Laurentian, etc.) and credit unions also serve the market. Mortgage originators also service the loan—they do not sell the service rights to a third party, which is common in the U.S.

In the last twenty years a new set of players has emerged—Mortgage Finance Companies (MFCs). The first of these, MCAP Financial Corporation, was incorporated in 1997. MFCs are non-depository monoline institutions that mostly administer mortgages through brokers and fund their lending through securitization or sales to third parties. By 2010, MFCs controlled about 10% of the mortgage market. Coletti et al. (2016) point out that the rise of the MFCs can be linked to the existence of government policies designed to increase competition in the mortgage

by 0.2 percentage points. Premiums are typically rolled into the loan, and therefore consumers only pay a small monthly fee; they are not required to pay the insurance upfront, which makes them price-insensitive to the insurance.

[13] Prepayment risk includes default risk—since all mortgages are insured, the lender is made whole for the amount outstanding if the borrower defaults.

market, mortgage insurance and securitization, and advances in information technology, consistent with the findings in Buchak et al. (2018). Unlike the big banks, these lenders are not directly subject to prudential regulation and supervision. As a result, they typically have lower levels of capital and contingent liquidity, holding roughly 40–90 cents of capital for every $100 of mortgages underwritten (see Coletti et al. 2016). That being said, since most of their mortgages are insured and then either securitized or sold off,[14] they are subject to the same mortgage-insurance rules as traditional lenders.

2.3 Mortgage brokers

Brokers have been present in the Canadian mortgage market going back to the 1970s, but really only penetrated the market starting in the mid 1990s, establishing a national broker association (CIMBL) in 1994. By 2005, brokers were responsible for negotiating roughly 40% of all insured contracts, and nearly 50% of contracts for first-time buyers.

Brokers in Canada have a fiduciary duty to their clients. They are compensated by lenders, but "hired" free of charge by borrowers to gather the best quotes from multiple lenders. Surveys by Maritz Canada and CAAMP (Maritz 2012 and Dunning 2011) suggest that mortgage brokers contact on average 4.5 lenders for each contract.[15] The broker arrangement in Canada differs from other jurisdictions. In the U.S., brokers receive both a cash-fee from the borrower and a yield-spread premium from the lender. The yield-spread premium is an increasing function of both the loan size and the interest rate, therefore brokers in the U.S. do not have an incentive to find borrowers the lowest rate (e.g., Hall and Woodward 2012). In the U.K., Robles-Garcia (2020) documents that over 70% of first-time buyers use a broker and compensation is split between the lender and borrower (and there is no fiduciary duty to the borrower).

Normally, for every mortgage they arrange, the broker receives an upfront commission from lenders (50–120 bps of the origination loan amount), such that brokers clearly benefit when borrowers take out larger loans. Brokers can also benefit from borrowers choosing longer amortization.

[14] This can be through National Housing Act mortgage-backed securities, via Canada Mortgage Bonds commercial mortgage-backed securities programs, or to federally regulated lenders. See Bédard-Pagé (2019).

[15] In principal, borrowers could potentially hire several brokers; this is something we observe.

Although the maximum amortization length is between 25 and 40 years, loan terms are much shorter, typically just five years. As a result, borrowers must regularly renew their mortgage to receive a new rate for the outstanding balance. At renewal, borrowers can choose to either stay with their current lender or switch to a different mortgage provider. Typically, at renewal brokers receive relatively low commissions if borrowers remain with their current lender. In contrast, if borrowers switch lenders, brokers would obtain higher commissions from the new lender (similar to the upfront commission at origination). This creates incentives for brokers to recommend switching lenders at renewal. Lenders have in part responded to this agency problem by introducing a trailer-fee compensation structure—small upfront commission at origination but annual payments for the duration that the borrower stays with the same lender.[16] This structure implies a number of reasons that brokers prefer products with longer amortization. First, if they receive trailer fees for every year that borrowers stay with the same lender, longer amortization potentially extends the period during which brokers get paid. Second, note that outstanding balances decline more slowly for mortgages with longer amortization due to smaller monthly payments. Since trailer fees are a fixed fraction of outstanding balances, the annual commissions are larger if amortization is longer. Lastly, Allen and Li (2020) find that Canadian borrowers with longer amortizing mortgages are more likely to switch lenders at renewal. Therefore, even under the upfront compensation structure, brokers have an incentive to recommend products with longer amortization since they expect more switching and hence higher commissions at renewal.

3 Data and descriptive analysis

3.1 Data

Our main dataset is a 25% random sample of insured mortgage contracts from the CMHC, from January 2005 to December 2007. The dataset contains information on over a dozen household/mortgage characteristics, including the financial characteristics of the contract (i.e., lender, rate, application

[16]For example, Merix Financial, one of the largest lenders operating in the broker channel, offers two compensation structures. The upfront compensation structure pays 100 bps at origination and nothing in following years even if borrowers stay at renewal. The trailer-fee structure pays 75 bps at origination, and then 8–20 bps of the outstanding balance annually for the duration of the borrower-lender relationship.

date, closing date, loan size, house price, debt ratio, risk type), and some demographic characteristics such as income, prior relationship with the lender, source of down-payment (cash, savings, loan, gift), residential status (renter, living with parents, previous home owner), and dwelling type (detached, semi-detached, condo, etc.). In addition, we observe the location of the purchased house up to the forward sortation area (FSA).[17] Importantly, we know which channel the borrowers use to originate their mortgages—bank branch or broker.[18]

The dataset contains lender identity information for 13 mortgage providers during our sample period. Mortgage contracts for which we do not have a lender name but only a lender type are coded as "Other credit union," "Other trusts," and "Other bank." These three categories are fragmented and contain mostly regional financial institutions. We therefore combine them into a single "Other Lender" category. All together, consumers face 14 lending options.

We restrict our sample to contracts with relatively homogeneous terms. In particular, we restrict our sample to contracts with the following characteristics: (i) 5-year fixed-rate term, (ii) 25-year or longer amortization, (iii) LTV $\geq$ 80%, and (iv) first-time home buyers' newly issued mortgages (i.e., excluding refinancing, renewal, and repeated purchase).[19] We trim the top and bottom 0.5% of borrowers in terms of income and house price. The final sample includes 48,398 observations.

For each borrower, we define the relevant market as a 10 KM circle around the centroid of her FSA. We obtain annual data on bank branch locations from Payments Canada. We also collected information on broker locations from a directory of brokers gathered annually by CAAMP. The directory has information on the broker and his/her associated firm. For the province of Ontario we also confirm the accuracy of this directory by comparing it to administrative data collected by the Financial Services Regulatory Authority of Ontario. We then count the number of banks and brokers in each borrower's local market.

[17] 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 kilometers, the median is much lower at 2.6 kilometers.

[18] We also observe, in less than 1% of cases, loans intermediated by either a real estate agent or construction company. We drop these contracts from our analysis.

[19] In analysis not reported here we examined whether repeat-buyers were different than first-time buyers. Conditional on observable borrower characteristics, repeat-buyers are 7.5 percentage points less likely to use a broker than first-time buyers. Conditional on using a broker, however, they have similar outcomes to first-time buyers.

3.2 Descriptive evidence

Table 1: Summary statistics by origination channel

	Branch	Broker	Mean Diff	SE	T Stat
Borrower characteristics					
Income ($1,000)	70.462	69.622	0.840	0.258	3.25
Borrower age (year)	33.151	33.849	-0.697	0.087	-8.01
I(Low Risk)	0.807	0.777	0.031	0.004	8.10
House price ($1,000)	181.280	202.970	-21.691	0.784	-27.68
Monthly payment (other debt)	931.44	849.07	82.38	5.47	15.05
Mortgage features					
Interest rate (%)	5.422	5.277	0.145	0.006	26.04
Monthly payment (mortgage)	1003.69	1102.32	-98.63	4.12	-23.97
Bond rate	4.241	4.231	0.010	0.004	2.25
Rate premium	1.440	1.436	0.004	0.001	3.49
Loan ($1,000)	168.589	189.486	-20.897	0.724	-28.87
Amortization (year)	27.77	28.8	-1.03	0.050	-20.74
LTV	93.35	93.64	-0.281	0.035	-7.95
TDS	33.36	34.07	-0.704	0.054	-13.05
I(Max Amortization)	0.492	0.569	-0.077	0.005	-17.11
I(Max LTV)	0.481	0.525	-0.044	0.005	-9.67
I(Max TDS)	0.093	0.131	-0.038	0.003	-13.21
Market structures					
Nb. banks	6.006	6.405	-0.398	0.023	-17.34
Broker presence	0.807	0.901	-0.094	0.003	-29.59
Nb. brokers	15.576	20.399	-4.823	0.215	-22.41
Share excluding brokers	0.206	0.211	-0.005	0.001	-3.84
Obs	23,611	24,787			

Note: This table presents the mean of transaction characteristics in both branch and broker channels. We also present the mean differences and standard errors (SE). The t-statistics indicate that most of the mean differences are statistically significant at the 0.1% level. Income is household gross annual income. The indicator variable, I(low risk), is equal to 1 if the credit score is higher than 720 and 0 otherwise. We observe granular FICO credit score buckets in the data, which we use in estimation. Rate premium is the difference between the medium posted rate of the biggest six national banks and the "no-haggle rate" set by broker-lenders (such as ING and First National). The indicator variable, I(Max amortization), is equal to 1 if the borrower chose the maximum allowable amortization at the time of origination. The indicator variable, I(Max LTV), equals 1 if the borrower chose the maximum allowable loan-to-value ratio at the time of origination. The indicator variable, I(Max TDS), equals 1 if the borrower has a total-debt-service ratio greater than 40%. All maximum allowable amounts vary over time. For each borrower, the number of banks and brokers are obtained for her local market, defined as a 10 KM circle around the FSA centroid. We also calculate the share of branches belonging to the banks that exclude brokers.

In this section we provide descriptive evidence that consumers transacting through mortgage brokers are observationally different from consumers transacting directly with financial institutions.

Table 1 describes the main financial and demographic characteristics of the borrowers in our sample, broken down by broker and branch transactions. Notice that slightly more than half of the borrowers in our sample obtain their mortgages via brokers. The top part of the table shows that broker-clients are on average riskier than branch-clients. Specifically, they have lower income and lower FICOs, but buy more expensive houses using longer amortization periods to smooth-out monthly loan payments. Broker-clients are also more likely to be constrained by the maximum allowable TDS ratio.[20] This is the case despite broker-clients having lower monthly payments towards non-mortgage debt such as credit cards, auto loans, etc.

Table 1 also presents summary statistics for the set of instrumental variables that we employ in Section 5. On average, there are 6 banks within a 10 KM radius of a branch-client and 6.4 banks for broker-clients. Overall, there is a high level of broker presence in Canada. For borrowers who transact with a branch, 80.7% have a broker in their neighborhood. For borrowers who transact with a broker the percentage is much higher—90.1%. Furthermore, there are many brokers in an average neighborhood, over 15. Finally, we calculate the share of bank branches owned by lenders that do not intermediate with brokers during our sample. That is, some banks prefer to forgo broker-business in order to save the commissions, even though this likely means less overall business. For branch-clients the share of bank branches in their local market that excludes brokers is 20.6% and for broker-clients the share is 21.1%.

Figures 1(a) and 1(b) display the distributions of LTV and amortization for the branch and broker channels over the sample period. The distributions of LTV and amortization are shifted to the right for broker-clients relative to branch-clients. Figures 1(c) and 1(d) present the same information, but restrict attention to the period during which the most lax guidelines were in place (effective November 19, 2006). The differences in the contract characteristics are even more pronounced in this window—a larger fraction of broker-clients choose the maximum LTV allowed (0 down-payment) and the maximum amortization length (40 years).

In Table 2, we report results from linear probability regressions of broker-use on LTV, amorti-

[20] In the data, we observe relatively few consumers with TDS constraint binding at 45%. In the model, consumers can adjust debt exactly to match the constraint, but in reality this adjustment can be discrete or random, and consumers might want to "over" adjust in order to avoid hitting the constraint at closing. Therefore, we assume that constrained consumers are those with TDS $\in (40\%, 45\%]$.

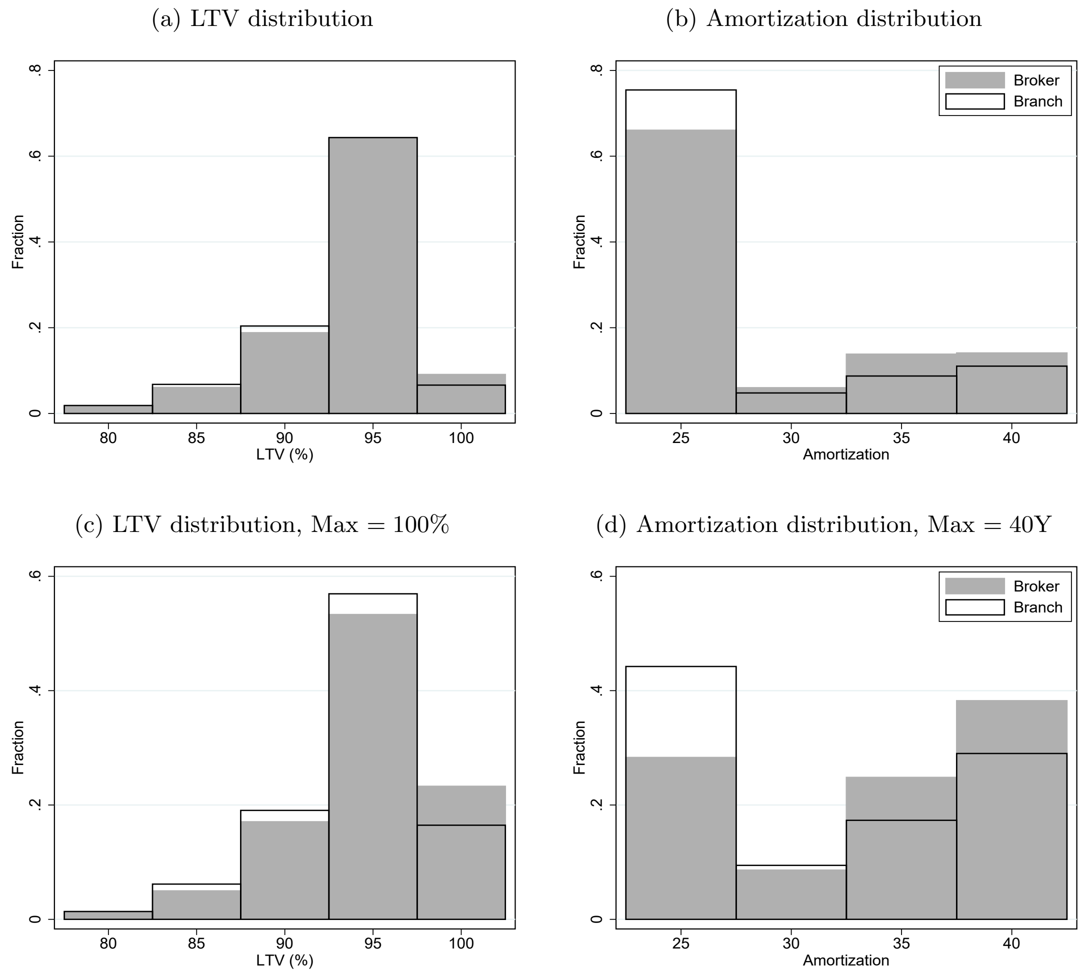
Figure 1: LTV and amortization distributions by origination channels

(a) LTV distribution

(b) Amortization distribution

(c) LTV distribution, Max = 100%

(d) Amortization distribution, Max = 40Y



zation, and a host of controls. The estimates from all specifications confirm the patterns observed in Figure 1—the broker channel is associated with mortgage products featuring higher LTV and longer amortization. We also find that borrowers using brokers have lower income and credit scores relative to borrowers using a branch. Turning to rates, Table 3 highlights that brokers, on average, negotiate better rates (about 10 bps). This difference is robust to controlling for other mortgage and consumer characteristics (column 3) as well as product (column 4) and lender (column 5) fixed effects.[21]

In Table 4, we investigate whether heterogeneity in observed characteristics (FICO score, in-

[21] We only know the lender category rather than the lender identity for 31% of the mortgages. As a result, these observations are dropped in column (5). The lender fixed effects do not explain much of the variation in rates.

Table 2: Broker channel is associated with riskier products

	I(broker)	I(broker)	I(broker)
LTV	0.002***	0.009***	0.009***
	(0.001)	(0.001)	(0.001)
Amortization (year)	0.008***	0.004***	0.008***
	(0.001)	(0.001)	(0.001)
FICO $\in$ [680,717]		-0.021***	-0.021***
		(0.007)	(0.007)
FICO $\in$ [718,759]		-0.036***	-0.035***
		(0.006)	(0.006)
FICO $\geq$760		-0.046***	-0.041***
		(0.007)	(0.007)
Income (log)		-0.214***	-0.193***
		(0.009)	(0.009)
Other controls	N	Y	Y
Year FE	N	N	Y
Region FE	N	N	Y
R^2	0.009	0.068	0.089

Note: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors clustered at the FSA level. Other controls include borrower age, house price (log), bond rate, TDS ratio, I(house age$\leq$5 year), and source of down-payment. There are 48,398 observations.

Table 3: Broker-clients pay less than branch-clients

I(broker)	-0.101***	-0.099***	-0.100***	-0.095***
	(0.006)	(0.006)	(0.006)	(0.007)
Region FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Other controls		Yes	Yes	Yes
Product FE			Yes	Yes
Lender FE				Yes
R^2	0.248	0.372	0.383	0.397
Obs	48398	48398	48398	33536

Note: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors clustered at the FSA level. Other controls include FICO score, income (log), borrower age, house price (log), bond rate, TDS ratio, I(house age$\leq$5 year), and source of down-payment.

come, borrower age, house price, property age, bond rate, TDS ratio, and source of down-payment) can fully explain the correlation between broker usage and product choice. We restrict attention to borrowers who have access to products with 100% LTV or products with amortization longer than 25 years. We find that, even after controlling for a rich set of observable characteristics together with year and region fixed effects, broker borrowers are more likely to choose products with higher LTV and longer amortization when they are available. The results hold even if we include lender fixed effects and focus on big lenders that originate in both the branch and broker channels, indicating that it is unlikely that the results are driven by heterogeneity in lenders across channels.

Table 4: Broker-clients are more likely to choose LTV>95% and amortization>25Y

	Panel A: I(LTV>95%)			
I(broker)	0.075***	0.072***	0.057***	0.052***
	(0.005)	(0.005)	(0.007)	(0.007)
R^2	0.369	0.375	0.413	0.349
Obs	19167	19167	13800	7977
	Panel B: I(AMT>25Y)			
I(broker)	0.084***	0.101***	0.049***	0.053***
	(0.006)	(0.005)	(0.008)	(0.008)
R^2	0.170	0.300	0.309	0.308
Obs	30847	30847	21830	13619
Other controls	Y	Y	Y	Y
Year FE	N	Y	Y	Y
Region FE	N	Y	Y	Y
Lender FE	N	N	Y	Y

Note: Panel A focuses on the subsample of borrowers who have access to products of 100% LTV at time of mortgage origination; while panel B focus on the subsample of borrowers who have access to products of more than 25-year amortization. The last column considers the biggest five lenders who originate in both branch and broker channels. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors clustered at FSA-level. Other controls include FICO score, income (log), borrower age, house price (log), bond rate, TDS ratio, I(house age$\leq$5 year), and source of down-payment.

These findings suggest that there is either selection on unobservable characteristics, or that brokers steer borrowers towards these products. For example, borrowers selecting into the broker channel might have different characteristics unobserved by the econometrician but observed by lenders and brokers (e.g., savings, checking account activity, gender, etc.), which could lead them to prefer higher LTV and longer amortizing products. On the other hand, brokers might steer borrowers towards products that are more profitable to them. In the next section, we build a structural model to account for these possible explanations for the observed differences in product choice across origination channels: selection on observables, selection on unobservables, and broker steering.

4 A model of mortgage-product choice

In this section we develop a model of mortgage-product choice. Taking as given the decision of whether or not to use a broker and the mortgage rate, the consumer makes a discrete choice over products characterized by two features: amortization length and LTV. In Section 5, we use a control function approach to handle the endogeneity in mortgage rates and choices of origination channel.

Given the best interest rate offer p in the chosen origination channel (branch or broker), the consumer makes a discrete choice among mortgage products with different LTV and amortization. The total set of products available is $\mathcal{J} \equiv \{80\%, 85\%, 90\%, 95\%, 100\%\} \times \{25Y, 30Y, 35Y, 40Y\}$. As mentioned in Section 2, our sample period featured a number of changes to macroprudential policies (relaxing the maximal amortization to 40 years and allowing for zero down-payment loans), and so certain choices in $\mathcal{J}$ might not be available when the borrower originates her mortgage. The borrower-specific choice set $\mathcal{J}_i$, therefore, can be smaller than $\mathcal{J}$.[22]

The borrower chooses the product that maximizes her indirect utility, which consists of her expected consumption as well as payoffs that depend on the characteristics of the mortgage. Given a rate of p, the associated monthly mortgage payment per dollar of mortgage loan is P_{AMT_j}, which is decreasing in amortization. We focus on the market for insured mortgages, hence the total loan size is $L_{ij} = (1+\Delta_j) \cdot LTV_j \cdot h_i$, where Δ_j is the insurance premium and h_i is the house price.[23] Then $R_j = P_{AMT_j} \cdot (1+\Delta_j) \cdot LTV_j$ denotes the monthly payment per dollar of house price. Borrowers also have access to non-mortgage debt, D_i (e.g., auto loans and credit cards) at a fixed rate of w, such

[22] The choice set does not depend on the origination channel. That is, every lender and broker offers the same menu of products in $\mathcal{J}_i$. This assumption seems reasonable given our focus on a set of relatively homogeneous products.

[23] We assume that the value of a consumer's housing choice is predetermined, i.e., consumers have a rough idea about the neighborhood they wish to purchase in and therefore the approximate price of the house.

that the monthly payment on non-mortgage debt is given by $d_i = D_i \cdot w$. Hence, we characterize consumption as:

$$c_{ij} = y_i - R_j h_i + D_i - d_i = y_i - R_j h_i + d_i \left(\tfrac{1}{w} - 1\right).$$

Loan qualification: In order to qualify for mortgage insurance, and therefore a loan, a borrower has to meet a TDS requirement—the total monthly payment on all debt should be less than 45% of gross income. The constraint limits the maximum monthly payment on non-mortgage debt:

$$\frac{R_j h_i + d_i}{y_i} \leq 0.45 \Rightarrow d_i \leq (0.45 \times y_i - R_j h_i) \equiv d_{ij}^{\max}. \tag{1}$$

We assume that d_i is drawn from a distribution $F_d(\cdot)$ after the contract is signed. The post-contractual monthly payment is $\tilde{d}_i = \min\{d_i, d_{ij}^{\max}\}$. Therefore, the TDS constraint affects product choices via its impact on consumption.

The expected consumption of consumer i who has chosen product j is as follows:

$$\begin{aligned} Ec_{ij} &= y_i - R_j h_i - \mathbb{E}[\min\{d_i, d_{ij}^{\max}\}] \cdot (1 - \tfrac{1}{w}) \\ &= y_i - R_j h_i - \left[\left(1 - F(d_{ij}^{\max})\right) \cdot d_{ij}^{\max} - \int_{-\infty}^{d_{ij}^{\max}} d_i \cdot dF(d_i)\right] \cdot \left(1 - \tfrac{1}{w}\right) \\ &= y_i - D_{ij}(p), \end{aligned}$$

where $D_{ij}(p) = R_j h_i + \mathbb{E}[\min\{d_i, d_{ij}^{\max}\}] \cdot (1 - \frac{1}{w})$ is the expected net debt payment conditional on the mortgage rate, p.

In addition to the utility derived from expected consumption, consumer payoffs also depend on other mortgage characteristics. We parameterize the indirect utility from choosing mortgage product $j \in \mathcal{J}_i$ (conditional on the origination channel *broker* and the rate offer p) as follows:

$$U_{ij}(p, broker) = Ec_{ij} - \pi_i \left(\frac{(1 - LTV_j) \cdot h_i + 0.03 \cdot h_i}{s_i}\right)^2 - \delta_i I_j + \mu_{j,broker}. \tag{2}$$

The second component represents the consumer's disutility of making an upfront payment from her savings, s_i. This lump-sum payment consists of a down-payment, $(1 - LTV_j) \cdot h_i$, as well as closing costs, assumed to be 3% of the house value.[24] Hence, consumers face a trade-off between the size of the down-payment and consumption: increasing the LTV leads to a smaller down-payment but higher monthly mortgage payments. Borrowers with very high savings prefer low-LTV contracts. As s_i falls, the cost of raising the required lump-sum payment increases, and consumers are more likely to choose options with a higher LTV. Since we do not observe s_i, we denote $\theta_i = \frac{\pi_i}{s_i^2}$ as a borrower-specific coefficient reflecting the heterogeneity in marginal disutility from the lump-sum payment. Besides savings, this coefficient also captures the opportunity cost of making a down-payment. Borrowers with better outside investment opportunities will have a higher disutility of making a large lump-sum payment.[25]

The third component in equation (2), $\delta_i I_j$, captures disutility from high interest rates beyond the direct effect of lowering monthly consumption. Specifically, I_j is the total interest cost associated with product j over the amortization period. This creates a natural trade-off, as products with higher amortization have a smaller monthly payment, R_j, but larger total interest costs, I_j. We allow for borrower-specific marginal disutility from interest costs, given by δ_i. This coefficient captures, in a reduced-form way, differences in borrowers' prepayment probabilities.[26]

Finally, $\mu_{j,broker}$ captures other non-monetary payoffs from choosing product j in origination channel $broker \in \{0, 1\}$. We allow these product fixed effects to differ across the origination channel. For instance, $\mu_{j,broker=0} > \mu_{j,broker=1}$ means that the consumer derives higher utility from product j if she obtains it from the branch channels. We interpret differences in consumer preferences across channels as *broker steering*.

Given the mortgage rate p and origination channel $broker$, the consumer's problem is to choose the mortgage product that maximizes the indirect utility, which generates the value function below:

$$U_i^*(p, broker) = \max_{j \in \mathcal{J}_i} U_{ij}(p, broker).$$

[24] See for example `https://www.canada.ca/en/financial-consumer-agency/services/buying-home.html` and `https://www.rbcroyalbank.com/mortgages/budgeting-for-closing-costs.html`.

[25] More generally, the marginal disutility of lump-sum payments can vary in consumer discount factors.

[26] Borrowers need not strictly follow the amortization schedule. For example, a borrower choosing a mortgage with a 40-year amortization might not be sensitive to the total interest cost over the full amortization period if she plans to sell the house after 5 years and exit the mortgage market.

5 Empirical specification

In this section, we parameterize the mortgage-product choice model to incorporate heterogeneity in observable characteristics. Conditional on the origination channel and mortgage rate, our model predicts the product choice. However, since unobserved consumer heterogeneity in the taste parameters (capturing savings and prepayment risks) can simultaneously affect product choices, origination channels, and mortgage rates, ignoring the endogeneity in origination channel and mortgage rate could bias the estimates of the product choice model. In order to control for the selection bias, we adopt a control function approach similar to Adams et al. (2009), Crawford et al. (2018), and Ioannidou et al. (2022). Specifically, we use first-stage regressions for mortgage rates and broker usage to construct control variables that capture the selection due to unobservables. Then in the second stage, we add these control variables to the taste parameter specifications and estimate the product choice model via maximum likelihood. We discuss identification at the end of this section.

5.1 Empirical specification and two-stage estimation strategy

We estimate the model in two stages, and start by describing the second. Our model predicts that a consumer's choice of product j is a function of observable and unobservable consumer characteristics. In particular, the consumer-specific taste parameters, (θ_i, δ_i), represent the disutility from making lump-sum payments (capturing consumer savings or outside investment opportunities) and the disutility from total interest costs (capturing prepayment risks), respectively. We parameterize them as functions of observable characteristics. Specifically, we assume that these parameters are random coefficients distributed log-normal, conditional on observable consumer characteristics $\boldsymbol{x}_i$ (FICO score, borrower age, income (log), I(house age$\leq$5 year), house price (log), bond rate, rate premium, region, year, and quarter fixed effects):

$$\ln \theta_i = \boldsymbol{x}_i' \boldsymbol{\theta} + u_{1i}, \quad \mathbb{E}(u_{1i}|\boldsymbol{x_i}) = 0, \tag{3}$$

$$\ln \delta_i = \boldsymbol{x}_i' \boldsymbol{\delta} + u_{2i}, \quad \mathbb{E}(u_{2i}|\boldsymbol{x_i}) = 0. \tag{4}$$

We denote $\boldsymbol{u}_i = (u_{1i}, u_{2i})$, and it captures the variation in the random coefficients that cannot be explained by the observable consumer characteristics.

Next, we specify that the monthly payment on non-mortgage debt, d_i, is log-normally distributed with location parameter $\bar{d}_i$ and variance σ_d^2. The location parameter is modeled as a function of observed characteristics $\boldsymbol{x_i}$, as well as income (y_i) and unobserved savings (through θ_i):

$$\bar{d}_i = \boldsymbol{x}_i'\boldsymbol{\lambda} + \lambda_0 y_i \theta_i.$$

Let $\boldsymbol{\beta}_2 = (\boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\delta}, \boldsymbol{\lambda}, \lambda_0)$. Depending on whether the TDS constraint is binding, the probability (density) of observing a monthly payment on non-mortgage debt $\tilde{d}_i \equiv \min\{d_i, d_{ij}^{\max}\}$ is given by

$$\begin{aligned} P_i(\tilde{d}_i) &\equiv P(\tilde{d}_i | p_i, broker_i, \boldsymbol{x_i}, \theta_i, \delta_i; \boldsymbol{\beta}_2) \\ &= P(\tilde{d}_i | p_i, broker_i, \boldsymbol{x_i}, \boldsymbol{u}_i; \boldsymbol{\beta}_2) = \begin{cases} 1 - F_d(d_{ij}^{\max}), & \text{if } TDS_i = T\bar{D}S \\ f_d(d_i), & \text{if } TDS_i < T\bar{D}S. \end{cases} \end{aligned}$$

Given the observed characteristics $\boldsymbol{x_i}$, the origination channel, the mortgage rate, the model parameters, and random coefficients, the model deterministically predicts product choices. To smooth the likelihood function for estimation, we introduce an IID preference shock for each product choice, ε_{ij}, drawing from a Type-1 extreme value distribution with scale parameter of σ_ε and mean of 0. Then for each consumer i, given her origination channel and mortgage rate, the probability of choosing product j can be expressed in a standard logit form:

$$\begin{aligned} P_i(j) &\equiv P_i(j | p_i, broker_i, \boldsymbol{x_i}, \theta_i, \delta_i; \boldsymbol{\beta}_2) \\ &= P_i(j | p_i, broker_i, \boldsymbol{x_i}, \boldsymbol{u}_i; \boldsymbol{\beta}_2) = \frac{\exp(U_{ij}/\sigma_\varepsilon)}{\sum_{k \in \mathcal{J}_i} \exp(U_{ik}/\sigma_\varepsilon)}. \end{aligned}$$

We choose a small scale parameter $\sigma_\varepsilon = 0.02$ to ensure that the idiosyncratic taste shocks have little effect on consumer choices.

Therefore, conditional on the observed characteristics $\boldsymbol{x}_i$, the origination channel $broker_i$, and interest rate p_i, the likelihood contribution of consumer i, who has non-mortgage debt payment $\tilde{d}_i$

and chooses mortgage product j, is given by

$$\begin{aligned}\ell_i(j,\tilde{d}_i,p_i,broker_i,\boldsymbol{x}_i;\boldsymbol{\beta}_2) &= \int P_i(j,\tilde{d}_i|p_i,broker_i,\boldsymbol{x}_i,\boldsymbol{u}_i;\boldsymbol{\beta}_2)\cdot f(\boldsymbol{u}_i|p_i,broker_i,\boldsymbol{x}_i)\,d\boldsymbol{u}_i\\ &= \int P_i(j)P_i(\tilde{d}_i)\cdot f(\boldsymbol{u}_i|p_i,broker_i,\boldsymbol{x}_i)\,d\boldsymbol{u}_i\end{aligned}$$

where $f(\boldsymbol{u}_i|p_i,broker_i,\boldsymbol{x}_i)$ is the conditional probability density function of $\boldsymbol{u}_i$. Recall that $\boldsymbol{u}_i$ captures the variation in the random coefficients (θ_i,δ_i) that cannot be explained by the observable consumer characteristics $\boldsymbol{x}_i$. More intuitively, it contains the consumer's private information regarding her savings and prepayment risk, which may be observable to lenders/brokers but not to us. Naturally, such private information also affects the borrower's profitability and hence her choice of origination channel as well as the mortgage rate, what we call *selection on unobservables*. Ignoring this could lead to biased estimates for parameters $\boldsymbol{\beta}_2$ in the product choice model, and especially for $\boldsymbol{\mu}$, which captures *broker steering*.

For example, a borrower with higher prepayment risk would seem less profitable from a lender's perspective and hence is more likely to use a broker to obtain a mortgage, since it might be too costly for her to search for approval on her own. Meanwhile, the borrower is more likely to choose mortgage products with longer amortization, since she worries less about the total interest cost over the full amortization period. In this case, a model ignoring self-selection into the broker channel due to unobserved prepayment risk would attribute the choice of longer amortization to broker-steering effects, overestimating the influence of brokers on consumer choices.

How can we control for the potential selection bias introduced by the endogeneity in mortgage rates and origination channels? We now turn to our first stage to construct control variables that capture the selection on unobservables.

We estimate the reduced-form pricing equation (5) and broker choice equation (6). Both rates and broker choice are a function of observable characteristics, $\boldsymbol{z}$, which include instrumental variables not in $\boldsymbol{x}$. We explain how these exclusion restrictions help identification in Section 5.2.

$$p_i = \begin{cases} \boldsymbol{z}_i' \boldsymbol{\gamma}_1^{broker=0} + e_{1i}, & \text{if } broker_i = 0 \\ \boldsymbol{z}_i' \boldsymbol{\gamma}_1^{broker=1} + e_{1i}, & \text{if } broker_i = 1 \end{cases}, \quad \mathbb{E}(e_{1i}|\boldsymbol{z_i}) = 0. \tag{5}$$

$$broker_i = \begin{cases} 0, & \text{if } \boldsymbol{z}_i' \boldsymbol{\gamma}_2 + e_{2i} < 0 \\ 1, & \text{if } \boldsymbol{z}_i' \boldsymbol{\gamma}_2 + e_{2i} \geq 0. \end{cases}, \quad e_{2i} \sim \mathcal{N}(0, 1). \tag{6}$$

We estimate the pricing equation (5) by ordinary least squares using the subsamples of branch-clients and broker-clients separately. Given the estimates $(\boldsymbol{\gamma}_1^{broker=0}, \boldsymbol{\gamma}_1^{broker=1})$, we obtain the residuals $\hat{e}_{1i} \equiv e_1(\boldsymbol{z_i}, p_i, broker_i) = p_i - \boldsymbol{z}_i' \hat{\gamma}_1^{broker=broker_i}$.

Equation (6), which determines the use of the broker channel, is estimated using a Probit model. We then obtain the generalized residuals as the inverse Mills ratio:

$$\hat{e}_{2i} \equiv e_2(\boldsymbol{z}_i, broker_i) = \begin{cases} -\phi(-\boldsymbol{z}_i' \hat{\gamma}_2)/\Phi(-\boldsymbol{z}_i' \hat{\gamma}_2), & \text{if } broker_i = 0 \\ \phi(-\boldsymbol{z}_i' \hat{\gamma}_2)/(1 - \Phi(-\boldsymbol{z}_i' \hat{\gamma}_2)) & \text{if } broker_i = 1, \end{cases}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the probability density function and cumulative distribution function of a standard Normal distribution $\mathcal{N}(0, 1)$, respectively.

The residuals $(\hat{e}_{1i}, \hat{e}_{2i})$ contain information on the unobserved heterogeneity in consumer savings and prepayment risks, which endogenously affect borrowers' mortgage rates, origination channel, and product choices. Therefore, we use these residuals as control variables in the random coefficients (θ_i, δ_i). Specifically, we assume that the error terms in equations (3) and (4) are linear in $\hat{e}_{1i}, \hat{e}_{2i}$ and their interaction term. We can rewrite the random coefficient specification as follows:

$$\ln \theta_i = \boldsymbol{x}_i' \boldsymbol{\theta} + \underbrace{\rho_1 \hat{e}_{1i} + \rho_2 \hat{e}_{2i} + \rho_3 \hat{e}_{1i} \hat{e}_{2i} + \tilde{u}_{1i}}_{u_{1i}}, \tag{7}$$

$$\ln \delta_i = \boldsymbol{x}_i' \boldsymbol{\delta} + \underbrace{\tau_1 \hat{e}_{1i} + \tau_2 \hat{e}_{2i} + \tau_3 \hat{e}_{1i} \hat{e}_{2i} + \tilde{u}_{2i}}_{u_{2i}}, \tag{8}$$

where $\tilde{u}_{1i} \sim \mathcal{N}(0, \sigma_\theta^2)$ and $\tilde{u}_{2i} \sim \mathcal{N}(0, \sigma_\delta^2)$ are independent of mortgage rates and the origination channels.

Denote $\boldsymbol{\beta} \equiv (\boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\delta}, \boldsymbol{\lambda}, \lambda_0, \boldsymbol{\rho}, \boldsymbol{\tau}, \sigma_\theta, \sigma_\delta, \sigma_d)$ the full vector of parameters for the second-stage product choice model. With the constructed residuals $(\hat{e}_{1i}, \hat{e}_{2i})$ from the first stage, we can rewrite the likelihood contribution of a borrower who has non-mortgage debt payment $\tilde{d}_i$ and chooses mortgage product j:

$$
\begin{aligned}
&\ell_i(j, \tilde{d}_i, p_i, broker_i, \boldsymbol{x}_i, \hat{e}_{1i}, \hat{e}_{2i}; \boldsymbol{\beta}) \\
&= \int P_i(j, \tilde{d}_i | p_i, broker_i, \boldsymbol{x}_i, \hat{e}_{1i}, \hat{e}_{2i}, \tilde{\boldsymbol{u}}_i; \boldsymbol{\beta}) f(\tilde{\boldsymbol{u}}_i | p_i, broker_i, \boldsymbol{x}_i, \hat{e}_{1i}, \hat{e}_{2i})\, d\tilde{\boldsymbol{u}}_i \\
&= \iint P_i(j | p_i, broker_i, \boldsymbol{x}_i, \theta_i, \delta_i; \boldsymbol{\beta}) P_i(\tilde{d}_i | p_i, broker_i, \boldsymbol{x}_i, \theta_i, \delta_i; \boldsymbol{\beta}) \phi_{\tilde{u}_{1i}}(\tilde{u}_{1i}) \phi_{\tilde{u}_{2i}}(\tilde{u}_{2i})\, d\tilde{u}_{1i} d\tilde{u}_{2i}
\end{aligned}
$$

We maximize the log-likelihood function $\mathcal{L} = \sum_i \ln \ell_i(j, \tilde{d}_i, p_i, broker_i, \boldsymbol{x}_i, \hat{e}_{1i}, \hat{e}_{2i}; \boldsymbol{\beta})$ for all the observations in our sample to obtain the estimate for $\boldsymbol{\beta}$.

5.2 Identification

We now provide a heuristic discussion of identification of the model parameters.

Recall that both rates and broker choices are a function of observable characteristics $\boldsymbol{z}$, which include all control variables in $\boldsymbol{x}$ (FICO score, borrower age, income (log), I(house age$\leq$5 year), house price (log), bond rate, rate premium, region, year, and quarter fixed effects) and hence allow us to control for consumer selection into brokers due to observable heterogeneity, as well as consumer-specific pricing algorithms. The region, year, and quarter fixed effects control for time-invariant mean differences across regions as well as year trends and within-year seasonality.

In addition, $\boldsymbol{z}$ includes instrumental variables that are not in $\boldsymbol{x}$. These variables are related to local market structure: (i) number of banks in a local market, (ii) an indicator variable that equals 1 if at least one broker is available in the local market, (iii) number of brokers in log, (iv) share of bank branches excluding the broker channel, (v) the interaction between (iii) and (iv), as well as (vi) the interaction between logged number of brokers and rate premium. Our exclusion restriction assumption requires that these instrumental variables affect the random coefficients

(θ_i, δ_i) through the constructed residuals in the first-stage regressions ($\tilde{\boldsymbol{u}}$ is independent of $\boldsymbol{z}$). For instance, conditional on $\boldsymbol{x}$, the number of banks in a local market does not directly affect consumer savings or prepayment risks. However, consumers with access to more banks are more likely to use a branch (e.g., due to the lower search costs) and have better rates (due to more competitive market structure), which creates variation in the constructed residuals $(\hat{e}_{1i}, \hat{e}_{2i})$ that are independent of $\boldsymbol{x}_i$. Therefore, the exclusion restrictions allow us to separately identify the effect of observable characteristics $\boldsymbol{x}$ on the random coefficients and the effect of the constructed residuals (selection on unobservables).

We briefly discuss the identification of the remaining parameters. The distribution of $\tilde{d}_i$ is identified from the observed non-mortgage debt payment and TDS constraint, as well as correlations of these variables with observed characteristics. The parameters in the random coefficient θ_i are identified from the correlation between the lump-sum payment of the chosen product, observed consumer characteristics $\boldsymbol{x}_i$, and constructed residuals $(\hat{e}_{1i}, \hat{e}_{2i})$ from the first-stage regressions. For example, if we observe that borrowers with higher credit scores are more likely to choose products with lower LTV (larger down-payment), then we can infer that a higher credit score is associated with lower disutility from a lump-sum payment (lower θ_i and higher savings). Similarly, correlation between total interest costs associated with the chosen product and $\boldsymbol{x}_i$ as well as $(\hat{e}_{1i}, \hat{e}_{2i})$ identifies the parameters in δ_i. The main parameter of interest is $\boldsymbol{\mu}$. It is identified from the differences between the probabilities of choosing products across different channels that are not explained by consumer preferences over expected consumption, lump-sum payments, and total interest costs.

6 Estimation results

In this section we present our model estimates (benchmark model). We then estimate restricted versions of the benchmark model to test for whether or not brokers steer borrowers and whether or not pricing and broker-use are endogenous to product choice.

Table 5: First-stage regression results

	Broker (OLS)	Broker (Probit)	Rate: Branch	Rate: Broker
Instrumental Variables				
Nb. banks	-0.017***	-0.045***	-0.009***	-0.003
	(0.003)	(0.009)	(0.003)	(0.003)
Broker presence	0.064***	0.174***	0.006	0.017
	(0.017)	(0.047)	(0.019)	(0.018)
Nb. brokers (log)	0.101***	0.269***	-0.065**	-0.110***
	(0.022)	(0.059)	(0.032)	(0.031)
Share excluding brokers	0.099***	0.266***	-0.015	-0.023
	(0.037)	(0.099)	(0.041)	(0.053)
Share excluding brokers	-0.114***	-0.301***	0.015	0.046
× Nb. brokers (log)	(0.021)	(0.057)	(0.028)	(0.030)
Rate premium	-0.013	-0.036	0.030	0.055**
× Nb. brokers (log)	(0.014)	(0.039)	(0.022)	(0.021)
Control Variables				
Rate premium	0.032	0.091	0.266***	0.191***
	(0.039)	(0.106)	(0.063)	(0.064)
Bond rate	0.042***	0.110***	0.529***	0.555***
	(0.010)	(0.025)	(0.014)	(0.012)
Income (log)	-0.188***	-0.503***	0.122***	0.135***
	(0.008)	(0.022)	(0.011)	(0.011)
FICO ∈ [680,717]	-0.029***	-0.079***	-0.053***	-0.042***
	(0.007)	(0.018)	(0.012)	(0.011)
FICO ∈ [718,759]	-0.052***	-0.138***	-0.108***	-0.117***
	(0.006)	(0.017)	(0.011)	(0.010)
FICO ≥760	-0.072***	-0.193***	-0.173***	-0.162***
	(0.007)	(0.017)	(0.010)	(0.009)
Borrower age	0.001***	0.002***	0.001***	0.002***
	(0.000)	(0.001)	(0.000)	(0.000)
House price (log)	0.182***	0.488***	-0.288***	-0.233***
	(0.010)	(0.027)	(0.014)	(0.012)
New property	-0.050***	-0.133***	-0.042***	-0.020***
	(0.006)	(0.017)	(0.008)	(0.007)
Constant	0.229*	-0.749**	4.874***	3.932***
	(0.118)	(0.316)	(0.174)	(0.153)
R^2	0.071		0.362	0.373
Obs	48398	48398	23611	24787

Note: The variable 'rate premium' is defined as the difference between the big 6 banks' median posted rate and the no-haggle rate. It has a min, median, and max of 1.2%, 1.45%, and 1.75%, respectively. The variable 'new property' is an indicator variable that equals 1 if the property is less than 5 years old. Region and quarter fixed effects are included. Standard errors clustered at the FSA level. $^{*}p < 0.1$, $^{**}p < 0.05$, *** $p < 0.01$. Marginal effects can be found in Table A1 in the Appendix.

6.1 Benchmark estimates

Table 5 displays the first-stage regression results. We find that local market structure affects the choice of origination channel. Each additional bank in a borrower's neighbourhood decreases the probability of using a broker by 1.7 p.p. The total negative marginal effect of the neighborhood share of lenders excluding brokers suggests that consumers are less likely to use a broker when there are more branches of the excluding banks. Consumers who have a broker in their neighborhood are 6.5 p.p. more likely to use a broker, and the number of brokers is positively correlated with broker choice. Additional controls are also statistically significant. An increase in the bond rate represents an increase in the cost of funding, and this is correlated with an increase in broker-use. Borrowers are also more likely to use a broker if they are purchasing a more expensive house. Variables that are negatively correlated with broker-use include income and credit score.

Focusing on columns (3) and (4), we find that regardless of origination channel, local market structure variables have a smaller effect on rates than they do on channel choice. In addition, we find that the presence of more banks in a neighbourhood is correlated with lower mortgage rates for branches, but not brokers. We also find that the number of brokers is correlated with lower rates, and this relationship is stronger when rate premiums are low. Finally, we find that higher income borrowers pay higher rates and higher credit score borrowers pay lower rates, and the differences between branch and broker are economically small. In general, our regression results are consistent with consumer-based pricing.

Table 6 presents the Maximum Likelihood estimates of the mortgage-product choice model. Recall that θ represents a consumer's marginal disutility of upfront payment—which is decreasing in the consumer's savings and increasing in their outside investment opportunities. We estimate lower disutility from making a lump-sum payment for borrowers who are buying more expensive houses, are more creditworthy (FICO ≥ 760), and are older. These estimates can be explained by the fact that these characteristics could be positively correlated with having higher savings. We estimate higher disutility for higher income borrowers and for borrowers who enter the market when the bond rate is higher, since these borrowers likely have better outside investment opportunities. From the coefficients on the control residuals ($\hat{e}_1$, $\hat{e}_2$), we find that high θ (lower savings or better

Table 6: Product choice model estimates

	(1)	(2)	(3)
	lnθ	$\bar{d}$	lnδ
Constant	19.2972 (0.1737)	-8.9499 (0.0867)	3.9164 (0.313)
FICO∈[680,717]	0.2846 (0.013)	0.1013 (0.0072)	0.2827 (0.023)
FICO∈[718,759]	0.1553 (0.0113)	0.044 (0.0064)	0.2479 (0.0211)
FICO≥760	-0.1999 (0.0111)	-0.1001 (0.006)	0.0592 (0.021)
Borrower age	-0.0051 (0.0004)	0 (0.0002)	-0.0038 (0.0007)
Income (log)	0.1289 (0.0124)	1.4705 (0.0073)	0.1442 (0.0245)
New property	-0.0115 (0.009)	0.0526 (0.0051)	0.1799 (0.0163)
House price (log)	-1.9198 (0.0149)	-0.6395 (0.0078)	-0.4059 (0.0278)
Bond rate	0.1944 (0.0155)	0.0103 (0.0087)	-0.1662 (0.0301)
Rate premium	0.0243 (0.0407)	-0.0143 (0.0241)	-0.3185 (0.0795)
$y\theta$		-0.0345 (0.0013)	
$\hat{e}_1$	0.1913 (0.006)		-0.2067 (0.0081)
$\hat{e}_2$	0.0673 (0.0098)		-0.0555 (0.0086)
$\hat{e}_1\hat{e}_2$	-0.0312 (0.0067)		-0.0592 (0.01)
σ_θ	0.6143 (0.0061)		
σ_d		0.5109 (0.0013)	
σ_δ			0.8355 (0.0094)
Log likelihood		-95,633.69	

Note: Estimates on year, quarter, and region dummies are not reported. Standard errors in parentheses. $\hat{e}_1$ is the residuals obtained from rate regressions reported in columns (3) and (4) of Table 5. $\hat{e}_2$ is the generalized residuals obtained from the Probit regression of broker choices reported in column (2) of Table 5. These residuals and their interaction are used to control for the endogeneity of mortgage rates and broker choices.

outside investment opportunities) are associated with higher mortgage rates and broker-use. In addition, the estimate on the interaction term $\hat{e}_1\hat{e}_2$ suggests that at the broker channel ($\hat{e}_2$ is high), the positive relationship between θ and rate is weaker. A possible reason for this is that cross-selling is rare in the broker channel. Therefore, even clients with low θ (high savings or bad outside investment opportunities), may not enjoy significantly better rates since lenders and brokers cannot earn extra profits from cross-selling other products (e.g., mutual funds).

The estimates in column (2), which are for the location parameter for non-mortgage debt, mostly have the same sign as those in column (1). This is consistent with borrowers with higher savings tending to have lower non-mortgage debt. Finally, the random coefficient in column (3), δ, captures borrower prepayment risks (or marginal disutility from total interest costs). The most important coefficients are for the control residuals, which show that higher rates and broker-use are associated with higher prepayment risks (and hence lower disutility from total interest costs).

6.2 Model fit

Using our model estimates we simulate the product choices for 100,000 broker-clients and 100,000 branch-clients. We then compare the resulting distributions of amortization, LTV, and TDS to those observed in the data. Amortization and LTV are product choices while TDS is an outcome. We present goodness of fit for a version of the model without the logit (taste) shocks. These shocks are needed to smooth the likelihood function in estimation, however, we do not rely on these shocks in the model simulation. Results are presented in Table 7. Consistent with the data, the share of riskier products (longer amortization and higher LTV) originated via brokers is higher than through branches. Quantitatively, the shares in each cell are similar for the data and model. In addition, the model is able to reproduce the difference in TDS distribution across channels. Overall, the model fits the data well.

6.3 Endogeneity in pricing and broker usage

As described in Section 5.1, our empirical model accounts for both the endogeneity in pricing and choice of origination channel using a control function approach. To quantify the importance of selection bias created by this endogeneity problem, we estimate a restricted version of the model in which the coefficients on the residuals obtained from the first-stage regressions ($\hat{e}_1$, $\hat{e}_2$) are set to 0, i.e., assuming that there is no selection on unobservables in the data generating process. See Appendix C for estimation results. The likelihood ratio test-statistic for the null hypothesis that the coefficients on our control residuals are equal to zero is 944.88. The critical value of $\chi^2(6)$ distribution associated with the 0.1% significance level is 22.46, and so we reject the restricted model. This result implies that controlling for endogeneity of prices and origination channel is

Table 7: Benchmark model fit

	Data		Model	
	Branch	Broker	Branch	Broker
AMT=25Y	44.59	28.33	45.82	28.13
AMT=30Y	9.48	8.57	10.04	10.57
AMT=35Y	17.08	24.95	17.22	23.79
AMT=40Y	28.84	38.16	26.92	37.51
LTV=80%	1.36	1.32	1.72	1.73
LTV=85%	6.24	4.77	6.34	5.65
LTV=90%	19.03	17.02	18.52	16.89
LTV=95%	56.39	52.65	56.11	51.75
LTV=100%	16.98	24.24	17.31	23.98
TDS∈(5%,15%]	0.48	0.18	0.18	0.13
TDS∈(15%,25%]	8.73	6.66	10.57	8.42
TDS∈(25%,35%]	39.89	38.50	42.46	41.73
TDS∈(35%,45%]	50.90	54.66	46.79	49.72

Note: the first two columns show the distribution of products in the data. The next two columns present the simulated distribution from the benchmark model for broker and branch channels.

important.

In Section 7 we focus on quantifying the importance of our three channels for product choice in our benchmark model. In Appendix C we do the same using this restricted model. The exercise further highlights that, if one does not account for endogeneity of prices and origination channel, the estimated broker-steering effects could be biased.

6.4 Do brokers steer borrowers?

Next we use our model to investigate whether there is evidence that brokers steer borrowers. Table 8 presents product-specific fixed effects in the branch channel ($\mu_{j,broker=0}$) and the estimated difference between origination channels: $\mu_{j,broker=1} - \mu_{j,broker=0}$. All coefficients are relative to a reference product with amortization of 25 years and LTV of 80%. For example, consider the bottom-right entry (-79.08, 0.57). The negative sign on the first number implies that for branch borrowers, holding all things equal across product choices,[27] the product (100%,40Y) is preferred less than the baseline product (80%,25Y). In terms of monetary values, branch borrowers on average require a monthly compensation of $|$-79.08$\times\$1,000/50|$ =\$1,581.6 to choose the product (100%,40Y) over the baseline product.[28] The second number in the cell captures the difference in product preferences across channels. In this example, it implies that borrowers using the broker channel on average require a monthly compensation of $|(-79.08 + 0.57) \times \$1,000/50|$ =\$1,570.2 to choose the product (100%,40Y) over the baseline product. Since this number is smaller than the compensation required in the branch channel (\$1,581.6), brokers affect borrower preferences relative to the baseline product.

Table 8: Estimates of product fixed effects across branch and broker channels

	LTV=80%	LTV=85%	LTV=90%	LTV=95%	LTV=100%
AMT=25Y	(0,0)	(-6.34,-3.03)	(-20.12,-5.27)	(-37,-7.02)	(-78.32,-2.57)
AMT=30Y	(-2.29,1.06)	(-11.8,-1.07)	(-27.91,-2.65)	(-44.94,-4.42)	(-81.16,-0.74)
AMT=35Y	(-1.69,2.66)	(-12.27,0.62)	(-29.12,-0.98)	(-46.67,-2.1)	(-80.35,0.63)
AMT=40Y	(-3.16,4.71)	(-12.32,0.53)	(-29.83,-0.86)	(-47.75,-1.84)	(-79.08,0.57)

Note: In each cell, the first number is the estimated product preference at branch channel ($\mu_{j,broker=0}$), while the second number is the estimated difference between origination channels ($\mu_{j,broker=1} - \mu_{j,broker=0}$). In order to test the null hypothesis, $\mu_{j,broker=1} = \mu_{j,broker=0}, \forall j$, we estimate the model under the null and obtain a likelihood ratio test statistic of 257. The critical value of $\chi^2(19)$ distribution associated with the 0.1% significance level is 43.82.

Our main objective is to investigate whether, after controlling for heterogeneity in consumer characteristics, preferences, and heterogeneous interest rate offers, brokers exert any influence on consumers' product choices. Specifically, we test the null hypothesis: $\mu_{j,broker=1} = \mu_{j,broker=0}, \forall\ j$. To do this we estimate a restricted version of the benchmark model under the null hypothesis that all product-specific fixed effects are the same in both channels. The likelihood ratio test-statistic for the null hypothesis that the product fixed effects are equal across origination channels is 257.0. The critical value of $\chi^2(19)$ distribution associated with a 0.1% significance level is 43.82; and so we reject the null. This result implies that brokers do exert influence on consumers as to the mortgage product they purchase. In other words, borrowers end up with a product that might not be their

[27] All things being expected consumption, lump-sum payment, and total interest cost.

[28] Since we use a small scale parameter $\sigma_\varepsilon = 0.02$ for the T1EV idiosyncratic preference shock distribution, after normalization the marginal disutility of \$1,000 monthly payment is 50.

optimal choice if they were instead using the bank-branch channel. However, simply from the entries in Table 8, it is difficult to infer the direction of broker steering since brokers affect the relative preferences across all products. As a solution, in the next section we investigate how consumer choices would be affected if we remove the broker-steering effects (i.e., set $\mu_{j,broker=1} = \mu_{j,broker=0}$).

7 Decomposition

We now formally decompose the contribution of broker steering, selection on observables, and selection on unobservables for explaining the different distributions of product choice across origination channels.

Table 9 presents the main results for product shares, and Table 10 reports summary statistics for key variables. Panel (A) of Table 9 describes product choices arising from the benchmark model. Each cell is a product share in the broker channel for (LTV, amortization) combinations. The most popular product has a 95% LTV and 40-year amortization (18.3% share). Choices of products with an LTV below 90% are relatively rare. Panels (B)-(E) also present product shares, but they are generated from different model simulations. In each case, we use our model estimates to simulate the choices of 100,000 broker-clients under different scenarios. Each scenario builds on the previous panel. More specifically, panel (B) removes only broker steering, panel (C) removes broker steering and selection on unobservables, panel (D) removes these first two channels and also rate differences across branches and brokers, and panel (E) removes the first three channels as well as differences in observable characteristics across branches and brokers.[29]

In panel (B) we remove broker steering, reporting product shares from a counterfactual in which we set the differences in product dummies across channels to zero. The key takeaway from this panel is that, in the absence of steering, the share of the products with 100% LTV and long amortization (defined as amortization of at least 35 years) falls. The share of products with 100% LTV falls from 23.98% to 21.09% and the share of products with 35- or 40-year amortization falls from 61.30% to 55.07%. In Table 10 we see that the average amortization length and LTV also fall,

[29] One might ask whether the sequence of removing differences across broker- and branch-clients matters—it does not. Results are available upon request.

Table 9: Product shares in broker channel (%)

	LTV=80%	LTV=85%	LTV=90%	LTV=95%	LTV=100%	**Total**
	Panel A: Model					
AMT=25Y	0.78	2.09	6.22	14.97	4.06	28.13
AMT=30Y	0.25	0.71	1.96	5.18	2.48	10.57
AMT=35Y	0.33	1.27	3.69	13.27	5.22	23.79
AMT=40Y	0.36	1.58	5.01	18.34	12.22	37.51
Total	1.73	5.65	16.89	51.75	23.98	100.00
	Panel B: Model w/o broker steering					
AMT=25Y	0.62	2.19	7.10	19.89	4.06	33.86
AMT=30Y	0.25	0.67	1.80	6.16	2.19	11.07
AMT=35Y	0.26	0.87	2.93	10.82	3.72	18.60
AMT=40Y	0.17	1.52	4.84	18.83	11.12	36.47
Total	1.30	5.24	16.67	55.70	21.09	100.00
	Panel C: Model w/o broker steering, branch-like control function variables					
AMT=25Y	1.04	3.16	9.03	20.69	3.33	37.24
AMT=30Y	0.35	0.88	2.19	6.20	1.86	11.48
AMT=35Y	0.37	1.13	3.45	10.78	2.95	18.67
AMT=40Y	0.21	1.79	5.11	17.27	8.22	32.60
Total	1.97	6.96	19.78	54.93	16.36	100.00
	Panel D: Model w/o broker steering, branch-like control function variables, rate setting same as branch					
AMT=25Y	1.11	3.30	9.20	20.59	3.29	37.49
AMT=30Y	0.37	0.93	2.25	6.24	1.81	11.60
AMT=35Y	0.41	1.16	3.60	10.70	2.85	18.72
AMT=40Y	0.26	1.91	5.32	16.92	7.78	32.18
Total	2.16	7.30	20.36	54.45	15.73	100.00
	Panel E: Model w/o broker steering, branch-like control function variables, rate setting, and observable characteristics same as branch					
AMT=25Y	1.08	3.33	9.74	24.18	4.40	42.73
AMT=30Y	0.29	0.82	1.94	5.77	1.83	10.65
AMT=35Y	0.34	1.05	3.17	10.29	2.93	17.77
AMT=40Y	0.16	1.53	4.41	15.04	7.71	28.85
Total	1.87	6.73	19.25	55.28	16.86	100.00

Note: In panel A we present the predicted product shares in our benchmark model. Panel B presents product shares under the restriction that the product dummies across channels are equal. Panel C presents product shares in a version of the model where we set the control function variables in the random coefficients to mimic those of the branch borrowers. In panel D we simulate broker borrowers' mortgage rates using the first stage rate regression estimates from branch borrowers (column (3) in Table 5). For each broker borrower, we draw the residual rate from the corresponding distribution in the branch channel, preserving the percentile ranking. Panel E draws observations from the branch borrower sample. The sampling scheme is weighted to mimic the share for each market-quarter.

while TDS increases. The last rows of Table 10 document the share of broker-clients that switch products across different panels. Moving from the benchmark to panel (B), we see that 14.57% of broker-clients choose a different product once we remove steering. For these borrowers, the average amortization falls by 4.46 years and their monthly mortgage payment increases by \$53.44 once we remove steering. In contrast, their LTV only decreases by 0.2 p.p. on average, suggesting that broker steering mainly influences consumer choices over amortization.

Table 10: Summary statistics: Panel A to E

	Data	Panel A	Panel B	Panel C	Panel D	Panel E	Branch
AMT (years)	33.65	33.53	32.88	32.33	32.28	31.64	31.26
LTV (%)	94.69	94.53	94.50	93.84	93.71	93.93	94.05
TDS (%)	34.77	35.02	35.15	35.18	35.35	34.59	34.44
GDS (%)	20.60	20.57	20.71	20.67	20.88	19.39	18.76
Rate (%)	5.53	5.53	5.53	5.53	5.64	5.64	5.67
Loan (\$1,000)	213.57	213.01	212.92	211.15	210.82	204.27	191.24
Mortgage payment	1167.89	1166.45	1174.24	1171.99	1183.26	1152.22	1085.57
θ		0.30	0.30	0.27	0.27	0.32	0.38
δ		0.47	0.47	0.52	0.52	0.53	0.56
Debt obligation	846.28	905.76	905.76	912.56	912.56	1003.72	1006.37
Change choice (%)			14.57	19.54	4.20		

Note: This table calculates the mean of variables of interest for broker borrowers using the data or simulated samples from Table 9. The last column ('Branch') presents summary statistics for simulated branch borrowers in panel B of Table 11. In the last row, we show the fraction of borrowers who change their product choices when moving from a previous panel to the current one. θ is the random coefficient reflecting the marginal disutility of lump-sum payments. δ is the marginal disutility of total interest costs.

In panel (C) we simulate a version of the model where we set the control function variables in the random coefficients to mimic those of the branch borrowers. This allows us to capture any selection on unobservables: broker borrowers could be unobservably different in their disutility from lump-sum payments (θ) and from interest costs (δ), and these unobserved differences are captured by $\hat{e}_1$ and $\hat{e}_2$, the (generalized) residuals obtained from Table 5. Specifically, we obtain the generalized residuals $\hat{e}_2$ for broker-clients assuming that they had instead chosen to use a branch. For $\hat{e}_1$, we draw from the residual distribution following the rate regression in the branch channel (column (3) in Table 5), preserving the percentile ranking. This effectively pushes the broker-clients' taste preferences towards those of branch-clients. This leads to a further decrease in the share of products with high LTV and long amortization—suggesting that these unobservable characteristics are also an important determinant of product choice. In Table 10 we can observe the corresponding decrease in average amortization length and LTV—moving from panel (B) to (C), we observe that 19.54% of broker-clients choose a different product once selection on unobservables is removed in addition to steering. For these borrowers, the average amortization falls by 2.8 years and the average LTV by 3.4 p.p.; selection on unobservables is much more important than broker steering in explaining the LTV difference across channels. Their monthly mortgage payment falls by $11.50.

Up to this point we have allowed broker- and branch-clients to have different interest rates. Recall that in the data we observe that broker-clients pay less. To generate the product shares in panel (D) we simulate the model by imposing that broker-clients face the same interest rates as branch-clients (using the first-stage rate regression estimates from branch-clients—column (3) in Table 5—and drawing the residual rate from the corresponding distribution in the branch channel, preserving the percentile ranking). We plot the change in mortgage rates faced by broker-clients in Figure 4 of Appendix A. Despite our counterfactual broker-clients paying about 11 basis points more, the product shares barely change in this counterfactual. That is, even if broker-clients faced the same rate distribution as branch-clients, they would choose their preferred (riskier) product.

Finally, panel (E) reports product shares in a counterfactual where we eliminate differences in observable borrower characteristics (other than the market-quarter that borrowers belong to).[30] The changes in product shares moving from panel (D) to (E) can therefore be attributed to differences in observed characteristics between branch borrowers and broker borrowers, rather than to differences in market composition between the two groups. The shares for different products chosen among broker-clients reported in panel (E) should closely match the product shares in the branch sample. For comparison these are reported in Table 11.

We summarize the change in product shares across panels in two ways. First, Table 12 compresses the 20 available products into four: (i) low-LTV/low-amortization, (ii) low-LTV/high-amortization, (iii) high-LTV/low-amortization, and (iv) high-LTV/high-amortization. We report

[30]Practically what this means is that we draw observations from the branch sample using a weighted sample scheme to ensure that the composition of mortgage-markets remains the same. For example, suppose we only have two markets: Toronto and Montreal, each with 100 mortgages. In Toronto, 20 mortgages are originated by banks and 80 by brokers, while Montreal has 20 from brokers and 80 from banks. In the data, therefore, our broker-client sample has 80 from Toronto and 20 from Montreal. We wish to keep this market composition in the simulation, which is achieved by assigning higher weight to Toronto branch borrowers. Note that if we draw 100 borrowers from the branch sample without weighting, we would have roughly 20 from Toronto and 80 from Montreal.

Table 11: Product shares in branch channel (%)

	LTV=80%	LTV=85%	LTV=90%	LTV=95%	LTV=100%	**Total**
	Panel A: Data					
AMT=25Y	0.82	3.09	10.13	25.43	5.12	44.59
AMT=30Y	0.13	0.73	1.69	5.32	1.61	9.48
AMT=35Y	0.30	1.06	2.95	10.14	2.64	17.08
AMT=40Y	0.11	1.36	4.27	15.50	7.61	28.84
Total	1.36	6.24	19.03	56.39	16.98	100.00
	Panel B: Model					
AMT=25Y	1.01	3.26	9.88	26.63	5.05	45.82
AMT=30Y	0.27	0.69	1.78	5.49	1.80	10.04
AMT=35Y	0.31	0.98	2.97	10.01	2.96	17.22
AMT=40Y	0.14	1.41	3.90	13.98	7.50	26.92
Total	1.72	6.34	18.52	56.11	17.31	100.00

the level change in product shares and in parenthesis we report the contribution from (a) broker steering, (b) selection on unobservables, and (c) selection on observables, respectively. A negative number implies that product shares fall as we move from broker to branch. Overall, we see that if broker-clients used a branch instead, the share of high-amortization products would fall. The largest gain in share would be low-LTV/low-amortization products. The contribution of broker steering towards high-amortization products is similar across LTV-bins, accounting for about 40% of the change in product shares. In contrast, selection on unobservables can explain over 50% of the change in the share of high-LTV/high-amortization products, but the contribution for the change in share of low-LTV/high-amortization products is negative. The choice of these products is mostly driven by selection on observables.

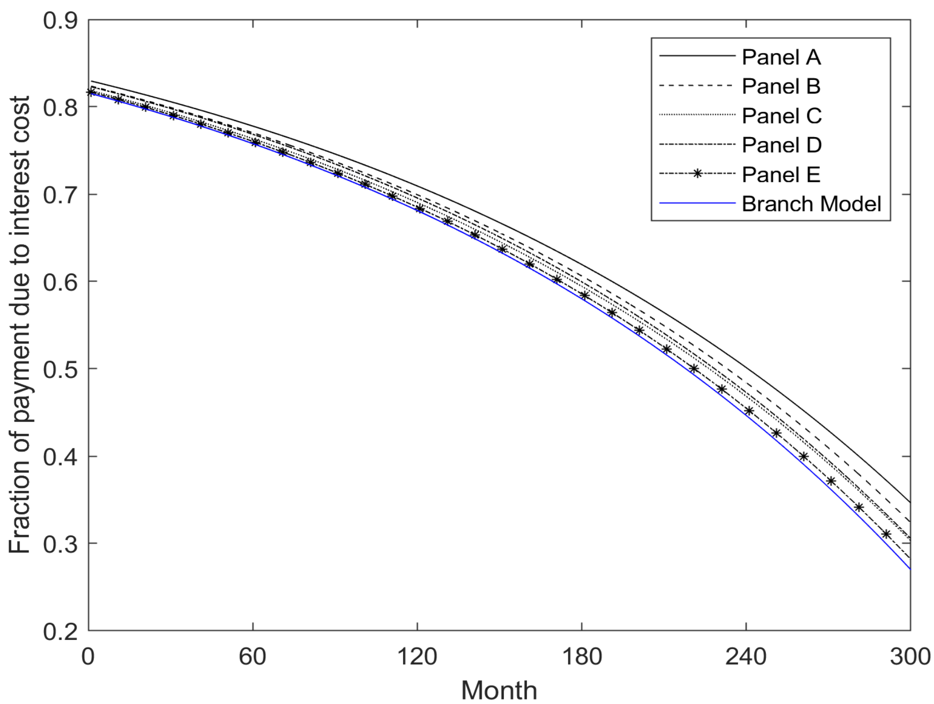
Table 12: Changes in product shares from Panel A to E

	LTV≤95%	LTV=100%	**Total**
AMT=25Y	14.26 (40.16%,28.88%,30.97%)	0.34 (1.17%,-215.25%,314.08%)	14.61 (39.25%,23.18%,37.58%)
AMT>25Y	-7.14 (39.69%,-8.58%,68.9%)	-7.46 (38.83%,53.56%,7.61%)	-14.61 (39.25%,23.18%,37.58%)
Total	7.12 (40.63%,66.43%,-7.06%)	-7.12 (40.63%,66.43%,-7.06%)	0

Note: In each cell, the first number shows the percentage point change in product share from Panel A to Panel E in Table 9. The fraction of the change that is due to broker steering (Panel A to Panel B), unobservables (Panel B to C), and observables (Panel C to E) is shown in order in parentheses.

To further highlight the role of selection on unobservables, we can compare the benchmark model predictions in Table 12 to the restricted model predictions presented in Table A5 of Appendix D. Recall that this restricted model (discussed in Section 6.3) assumes away selection on unobservables by excluding the control residuals in the random coefficient specifications. Compared to the benchmark model predictions in Table 12, the restricted model attributes, at least partially, the effect of selection on unobservables on product choices to broker steering. For instance, the contribution of broker steering in the low-LTV/low-amortization and the high-LTV/high-amortization bins are close to the sum of the contribution from both broker steering and selection on unobservables reported in Table 12. Therefore, the restricted model overestimates the contribution of broker steering away from low-LTV/low-amortization products and the contribution of broker steering towards high-LTV/high-amortization products.

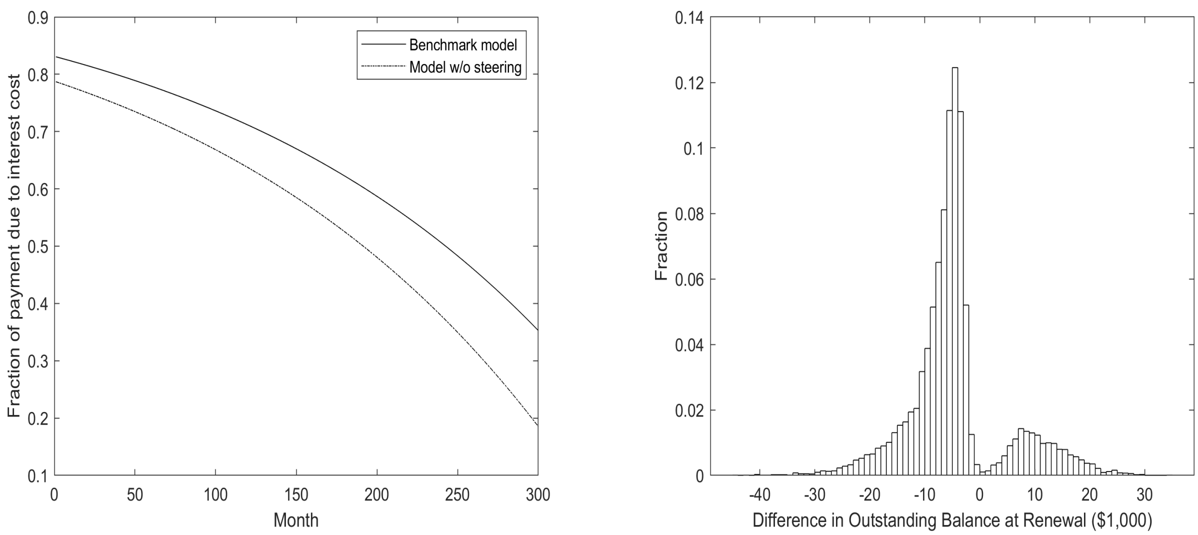
Figure 2: Amortization schedules



Note: This figure plots the fraction of borrowers' monthly mortgage payments that goes towards their interest costs. The calculations are based on simulated broker borrower samples reported in Panels A to E in Table 9. The solid blue line ('Branch Model') shows the amortization schedule for simulated branch borrowers in panel B of Table 11.

Given that broker steering seems to largely impact the average amortization length, our second approach for highlighting our findings is to graphically illustrate the impact on borrowers' total interest costs from being steered. Figure 2 shows the amortization schedule (fraction of the mortgage payment that is interest relative to principal) in the benchmark model versus the simulated models defined above. Moving from 'Panel A' (broker model) to 'Branch Model' we see that borrowers pay down their mortgage faster and as a result pay less in terms of interest.

Figure 3: Amortization schedules and outstanding balances for steered borrowers



(a) (b)

Note: This figure focuses on the 14.57% borrowers who change their product choices without broker steering. Panel (a) plots the fraction of a borrower's monthly mortgage payment that is interest cost (amortization schedule). The solid line is in the benchmark model and the dashed line is in the simulated model where we set the difference in product dummies across origination channels to be zero. Panel (b) plots the difference in outstanding balances five years after origination between the baseline model and the model without steering.

Next, we explore in more detail the 14.57% of borrowers who are steered by brokers. Among them, 74% choose longer amortization due to broker steering, 20% increase their LTV, and 16% decrease their LTV. Our results are presented in Figure 3, which contains two panels. Panel (A) reproduces the amortization schedule shown in Figure 2 for steered borrowers. Removing broker steering leads borrowers to take out shorter-term mortgages. As a result, the share of their monthly payment that is interest relative to principal changes quite drastically. After 25 years, 65% of a steered borrower's payment goes towards principal, versus over 80% in the counterfactual. Over the lifetime of the mortgage, these borrowers can save $32,449 in interest costs if there were no steering. In panel (B) we plot the difference in outstanding balance at renewal for a borrower in the benchmark model versus the simulated model without steering. On average, if borrowers were not steered, they would on average have \$4,505 less to refinance after five years of paying down their mortgage. There is, however, substantial dispersion, with some borrowers having over \$30,000 less to refinance at renewal and some over \$20,000 more.

Why might brokers steer borrowers towards longer amortizing mortgages? The combination of 5-year term contracts and the structure of broker compensation likely explains this phenomenon. The short-term nature of the contract means that borrowers have to renegotiate a new rate every 5 years ('renewal'). From the schedules plotted in Figures 2 and 3, we see that longer amortizing mortgages have larger outstanding balances at renewal. Although brokers receive trailer fees for re-signing their client with the same lender after 5 years (say 20 bps on the outstanding balance), they can also be paid the full compensation (say, 100 bps on the outstanding balance) if they are able to find a new lender. In both cases brokers benefit from the larger outstanding balance, but of course they benefit more if consumers switch at renewal. Focusing solely on branch borrowers, Allen and Li (2020) find that Canadian mortgagors with longer amortizing mortgages are more likely to switch lenders at renewal. Therefore, brokers have an incentive to recommend products with longer amortization since they expect more switching and hence higher commissions at renewal.

Given that broker compensation is proportional to mortgage loan size, why do we observe relatively little steering towards higher LTV products? One reason is that broker-clients tend to have lower θ, either due to higher savings or worse outside investment opportunities, than branch borrowers. However, unlike bank branches, brokers have little opportunity to cross-sell investment products. Therefore, it is hard for brokers to steer a borrower to a higher-than-ideal LTV (i.e., investing less in housing). The observed choice of high LTV is therefore driven mainly by broker-clients' preferences.

Lastly, why would brokers steer a small fraction (2.37%) of borrowers towards lower LTV products? This can hardly be explained by profit maximization. We find that these borrowers on average have very low θ (i.e., high savings or bad outside investment opportunities). If they were to visit a bank branch, a high LTV mortgage bundled with an investment product might be recommended to them. Since brokers have little opportunity to cross-sell, they likely provide a different recommendation: invest more in housing with a low LTV mortgage. Such a recommendation reflects brokers' fiduciary duty and does not necessarily create higher profits.

8 Conclusion

In this paper we study the role of intermediaries in credit markets. Using data from the Canadian mortgage market, we document that consumers who transact through brokers are riskier (have lower income and credit scores) and obtain riskier products (with higher amortization and LTV). We propose two possible explanations for these observations. First, if broker compensation is tied to loan size and amortization length, brokers may steer consumers towards specific products. At the same time, less profitable consumers may be more likely to use a broker to help them search for a lender willing to offer a product that allows them to qualify for their desired credit level. To disentangle the importance of selection on unobservable consumer characteristics (such as available savings) and broker steering, we develop and estimate a model of mortgage-product choice. Our decomposition exercise suggests that broker steering mainly affects consumer choices of amortization length, while the difference in LTV distributions across channels is largely driven by selection on unobservables.

References

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

Aguirregabiria, V., R. Clark, and H. Wang (2020). The geographic flow of bank funding and access to credit: Branch networks and local-market competition. Working paper.

Alexandrov, A. and S. Koulayev (2017). No shopping in the U.S. mortgage market: Direct and strategic effects of providing information. Working paper, CFPB No. 2017-01.

Allen, J., R. Clark, and J. Houde (2019). Search frictions and market power in negotiated-price markets. *Joural of Political Economy 127*, 1550–1598.

Allen, J., R. Clark, and J. F. Houde (2014a). The effect of mergers in search markets: Evidence from the Canadian mortgage industry. *American Economic Review 104*, 3365–3396.

Allen, J., R. Clark, and J. F. Houde (2014b). Price dispersion in mortgage markets. *Journal of Industrial Economics 62*, 377–416.

Allen, J., T. Grieder, B. Peterson, and T. Roberts (2020). The impact of macroprudential housing finance tools in Canada. *Journal of Financial Intermediation 42*, 100761. Macroprudential policies in the Americas.

Allen, J. and S. Li (2020). Dynamic competition in negotiated price markets. Bank of Canada working paper 2020-22.

Anagol, S., S. Cole, and S. Sarkar (2017). Understanding the advice of commissions-motivated agents: Evidence from the Indian life insurance market. *Review of Economics and Statistics 99*, 1–15.

Argyle, B., T. Nadauld, and C. Palmer (2020). Real effects of search frictions in consumer credit markets. Working paper.

Bar-Isaac, H. and A. Gavazza (2015). Brokers' contractual arrangements in Manhattan rental markets. *Journal of Urban Economics 86*, 73–82.

Bédard-Pagé, G. (2019). Non-bank financial intermediation in Canada: An update. Bank of Canada Staff Discussion Paper.

Benetton, M. (2021). Leverage regulation and market structure: An empirical model of the UK mortgage market. *Journal of Finance 76*, 2997–3053.

Benetton, M., A. Gavazza, and P. Surico (2019). Mortgage pricing and monetary policy. Working paper.

Bergstresser, D., J. M. Chalmers, and P. Tufano (2008). Assessing the costs and benefits of brokers in the mutual fund industry. *The Review of Financial Studies 22*(10), 4129–4156.

Berndt, A., B. Hollifield, and P. Sandås (2010, July). The role of mortgage brokers in the subprime crisis. Working Paper 16175, National Bureau of Economic Research.

Bhattacharya, V., G. Illanes, and M. Padi (2021). Fiduciary duty and the market for financial advice. Working paper.

Biglaiser, G., F. Li, C. Murry, and Y. Zhou (2020). Intermediaries and product quality: Theory and evidence from used car dealers. *RAND Journal of Economics 51*, 905–933.

Braido, L. H. and B. C. Ledo (2018). Dynamic price competition in auto insurance brokerage. *The RAND Journal of Economics 49*(4), 914–935.

Buchak, G., G. Matvos, T. Piskorski, and A. Seru (2018). Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics 130*, 453–483.

Buchak, G., G. Matvos, T. Piskorski, and A. Seru (2020). Beyond the balance sheet model of banking: Implications for bank regulation and monetary policy. Forthcoming, *Journal of Political Economy*.

Chalmers, J. and J. Reuter (2020). Is conflicted investment advice better than no advice? *Journal of Financial Economics 138*(2), 366–387.

Charoenwong, B., A. Kwan, and T. Umar (2019). Does regulatory jurisdiction affect the quality of investment-adviser regulation? *American Economic Review 109*, 3681–3712.

Christoffersen, S., R. Evans, and D. K. Musto (2013). What do consumers' fund flows maximize? Evidence from their brokers' incentives. *Journal of Finance 68*, 201–35.

Claessens, S. (2015). An overview of macroprudential policy tools. *Annual Review of Financial Economics 7*(1), 397–422.

Clark, R., J. F. Houde, and J. Kastl (2021). The industrial organization of financial markets. In K. Ho, A. Hortaçsu, and A. Lizzeri (Eds.), *Handbook of Industrial Organization, vol. V*. Elsevier.

Coletti, D., M.-A. Gosselin, and C. MacDonald (2016, December). The rise of mortgage finance companies in Canada: Benefits and vulnerabilities. *Bank of Canada Financial System Review*.

Crawford, G., N. Pavani, and F. Schivardi (2018). Asymmetric information and imperfect competition in lending markets. *American Economics Review 108*(7), 1659–1701.

Donna, J., P. Pereira, T. Pires, and A. Trindade (2021). Measuring the welfare of intermediation in vertical markets. Forthcoming, *Management Science*.

Dunning, W. (2011). Annual state of the residential mortgage market in Canada.

Egan, M. (2019). Brokers vs. retail investors: Conflicting interests and dominated products. *Journal of Finance 74*, 1217–1260.

Egan, M., S. Ge, and J. Tang (2022). Conflicting interests and the effect of fiduciary duty: Evidence from variable annuities. *The Review of Financial Studies 35*(12), 5334–5386.

Egan, M., G. Matvos, and A. Seru (2019). The market for financial adviser misconduct. *Journal of Political Economy 127*, 233–295.

Foà, G., L. Gambacorta, L. Guiso, and P. E. Mistrulli (2019). The supply side of household finance. *The Review of Financial Studies 32*(10), 3762–3798.

Foerster, S., J. Linnainmaa, B. Melzer, and A. Previtero (2017). Retail financial advice: Does one size fit all? *The Journal of Finance 72*(4), 1441–1482.

Gavazza, A. (2016). An empirical equilibrium model of a decentralized asset market. *Econometrica 84*, 1755–1798.

Grigsby, J., A. Hortacsu, G. Matvos, A. Seru, and V. Yao (2020). Searching for approval. Working paper.

Grundl, S. and Y. Kim (2019). Consumer mistakes and advertising: The case of mortgage refinancing. *Quantitative Marketing and Economics 17*, 161–213.

Grunewald, A., J. Lanning, D. Low, and T. Salz (2021). Auto dealer loan intermediation: Consumer behavior and competitive effectsy.

Guiso, L., A. Pozzi, A. Tsoy, L. Gambacorta, and P. E. Mistrulli (2022). The cost of steering in financial markets: Evidence from the mortgage market. *Journal of Financial Economics 143*(3), 1209–1226.

Gurun, U., G. Matvos, and A. Seru (2016). Advertising expensive mortgages. *The Journal of Finance 71*(5), 2371–2416.

Hall, R. E. and S. Woodward (2012, December). Diagnosing consumer confusion and sub-optimal shopping effort: Theory and mortgage-market evidence. *American Economic Review 102*(6).

Honka, E. (2014). Quantifying search and switching costs in the US auto insurance industry. *The RAND Journal of Economics 45*(4), 847–884.

Ioannidou, V., N. Pavanini, and Y. Peng (2022). Collateral and asymmetric information in lending markets. *Journal of Financial Economics 144*(1), 93–121.

Jia-Barwick, P., P. Pathak, and M. Wong (2017). Conflicts of interest and steering in residential brokerage. *American Economic Journal: Applied Economics 9191-222*, 191–222.

Jiang, W., A. Nelson, and E. Vytlacil (2014). Liar's loan?: Effects of origination channel and information falsification on mortgage delinquency. *Review of Economics and Statistics 96*, 1–18.

Khorana, A., H. Servaes, and P. Tufano (2008, 04). Mutual fund fees around the world. *The Review of Financial Studies 22*(3), 1279–1310.

LaCour-Little, M. (2009). The pricing of mortgages by brokers: An agency problem? *Journal of Real Estate Research 31*.

Maritz (2012). Mortgage insights. Maritz Research Canada.

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

Myśliwski, M. and M. Rostom (2022). Value of information, search, and competition in the UK mortgage market. Bank of England working paper No. 967.

Robles-Garcia, C. (2020). Competition and incentives in mortgage markets: The role of brokers. Working paper.

Rubinstein, A. and A. Wolinsky (1987). Middlemen. *The Quarterly Journal of Economics 102*(3), 581–593.

Ruckman, K. (2003). Expense ratios of North American mutual funds. *Canadian Journal of Economics/Revue canadienne d'économique 36*(1), 192–223.

Salz, T. (2022). Intermediation and competition in search markets: An empirical case study. *Journal of Political Economy 130*, 310–345.

Schneider, H. S. (2012). Agency problems and reputation in expert services: Evidence from auto repair. *Journal of Industrial Economics 60*, 406–33.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *The Journal of Human Resources 50*(2), 420–445.

A Additional tables and figures

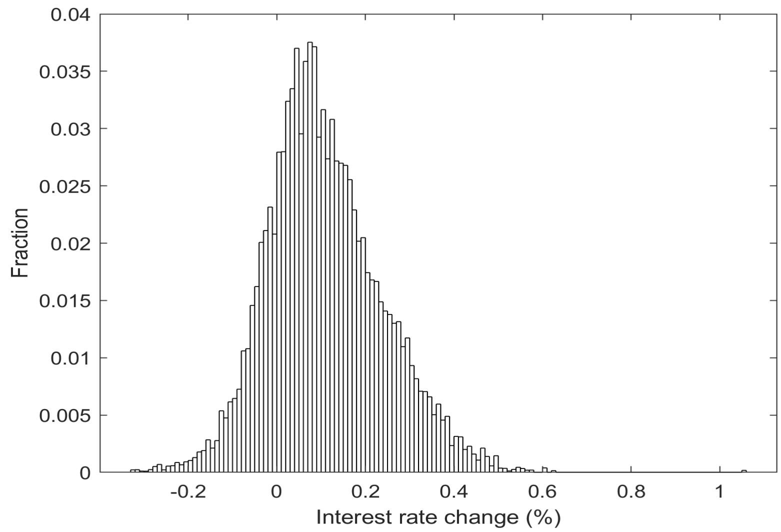
Table A1 presents the marginal effects of instrumental variables presented in Table 5 in Section 6.1. Figure 4 plots the change in mortgage rates faced by broker-clients in a counterfactual where they pay the same rate as similar branch-clients.

Table A1: Marginal effects of instrumental variables in first-stage regressions

	Broker (OLS)	Broker (Probit)	Rate: Branch	Rate: Broker
Nb. banks	-0.017***	-0.017***	-0.009***	-0.003
	(0.003)	(0.003)	(0.003)	(0.003)
Share excluding brokers	-0.153***	-0.152***	0.016	0.088*
	(0.046)	(0.047)	(0.050)	(0.050)
Broker presence	0.064***	0.065***	0.006	0.017
	(0.017)	(0.018)	(0.019)	(0.018)
Nb. brokers (log)				
Rate premium=1.2%	0.062***	0.061***	-0.025**	-0.034***
	(0.008)	(0.008)	(0.010)	(0.009)
Rate premium=1.45%	0.059***	0.057***	-0.018**	-0.021***
	(0.007)	(0.007)	(0.008)	(0.007)
Rate premium=1.75%	0.055***	0.053***	-0.009	-0.004
	(0.008)	(0.008)	(0.010)	(0.009)
Obs	48398	48398	23611	24787

Note: The variable 'rate premium' is defined as the difference between the big six banks' median posted rate and the no-haggle rate. It has a min, median, and max of 1.2%, 1.45%, and 1.75%, respectively.

Figure 4: Distribution of rate changes moving from Panel C to Panel D



Note: The horizontal axis is measured in percentage points, i.e., 0.2 is 20 basis points. The histogram plots the change in mortgage rates that broker-clients would counterfactually face if they used a bank branch instead to originate their mortgage.

B Choice of origination channel and rate determination

In section 5.1 we provide an argument for how our control function approach captures the choice of origination channel and the distribution of mortgage rates. Here we present one model that can rationalize our approach. Consider a first stage to the game where consumers choose their origination channel. That is, choose either to search on their own (obtaining quotes and finding the best interest rate) or outsourcing this process to a broker.

We first describe the procedure if a consumer searches on their own. We assume that the consumer knows their cost of contacting potential lenders, κ_i, realized from the distribution $F(\cdot)$. For simplicity, we assume that all 'own-search' consumers obtain the same number, n, of interest rate quotes. We parameterize the offer rate p_{ij} of bank j to consumer i as follows:

$$p_{ij} = c_i + \omega_j,$$

where c_i is a common lending cost and ω_j is a bank-specific cost-shock, distributed according to $G(\omega)$ with mean 0. While all mortgages are insured, the lending costs capture the cost of prepayment net of potential benefits to banks of obtaining a customer and cross-selling other products. We assume that this component is observed and known by consumers and banks. Note that we also assume that cost differentials across banks are random, so that there is no bank that consistently offers smaller/larger rates to different consumers.

The rate negotiation process we use is a simplified version Allen et al. (2019), where banks compete for consumers in an English auction. The best offer is equal to the second smallest lending cost:

$$p_i = c_i + \omega^n_{(2)},$$

where $\omega^n_{(k)}$ represents the k-th smallest value of ω among n banks. Hence, if a consumer searches on their own, her expected utility is defined as the expected utility from the best mortgage product conditional on the interest rate net of search costs:

$$\mathbb{E}_\omega[U^*_i(p_i, broker_i = 0)] - \kappa_i.$$

Alternatively, consumers can outsource the search process to a broker. In this case the consumer does not have to pay a search cost, since brokers obtain the quotes and negotiate with the banks.[31] The broker provides an additional benefit—she can contact more banks, n^b, and hence potentially obtain lower offers. As a result, the best interest rate obtained through the broker channel is:

$$p^b_i = c_i + \omega^{n_b}_{(2)},$$

which generates expected utility $\mathbb{E}_\omega[U^*_i(p^b_i, broker_i = 1)]$.

As a result, the consumer decides to use the broker if the expected utility from the broker channel is higher than searching on their own:

$$\mathbb{E}_\omega[U^*_i(p^b_i, broker_i = 1)] \geq \mathbb{E}_\omega[U^*_i(p_i, broker_i = 0)] - \kappa_i$$

This results in a cutoff rule, where consumers use a broker only if their realized search cost is high enough relative to the expected benefits from using a broker:

$$\kappa_i \geq \mathbb{E}_\omega[U^*_i(p^b_i, broker_i = 0)] - \mathbb{E}_\omega[U^*_i(p_i, broker_i = 1)]$$

Selection into the broker channel works through several mechanisms. First, consumers with high search costs are most likely to use brokers. Second, a broker's ability to negotiate on interest rates affects the consumer choice at the first stage in a heterogenous manner. While all consumers gain from lower interest rates, those who are unable to make a large down-payment, or who are sensitive to total interest costs, benefit more. More than that, the interest rate plays a crucial role in the TDS constraint, as high monthly mortgage payments constrain the amount of non-mortgage debt that a consumer can accumulate. Thus, the indirect effect of brokers negotiating lower interest rates is that consumers can more easily meet the TDS constraint ("qualify"). The effect is more pronounced for high-risk consumers with large debt, which plays a role in adverse selection into the broker channel.

[31] In other words, we normalize the search costs through the broker channel to 0, and we can interpret κ_i as the difference between search costs through direct channel and through broker channel.

C A model of product choice that does not correct for endogeneity of origination channel and rates

This section presents the estimates of the restricted model discussed in Section 6.3, which does not deal with the endogeneity of mortgage rates and choice of origination channel. The restricted model does not include the control function variables obtained from first-stage regressions in the random coefficient specifications, i.e., setting the coefficient of $\hat{e}_1$, $\hat{e}_2$, and their interaction to zero. Therefore, the restricted model shuts down the possibility of selection on unobservables in the data generating process.

Table A2 reports the maximum likelihood estimates for this restricted model. The likelihood ratio test statistic rejects the restricted model at 0.1% significance level, suggesting that selection on unobservables is important for rationalizing the product choices observed in the data.

Table A3 reports the product-specific fixed effects in the branch channel ($\mu_{j,broker=0}$) and the estimated difference between origination channels: $\mu_{j,broker=1} - \mu_{j,broker=0}$. All coefficients are relative to a baseline product with amortization of 25 years and LTV of 80%. Compared to the benchmark product dummy estimates reported in Table 8, brokers in the restricted model seem to have stronger influence on borrower preferences relative to the baseline product.

Table A4 reports the predicted product shares given the restricted model under the different counterfactual scenarios described in Section 7. Notice that in the benchmark, panel C eliminates selection on unobservables, while in the restricted model it is assumed away. Moving from panel B to panel C in Table A4, therefore, does not lead to any change in product shares.

Finally, Table A5 summarizes the contribution of broker steering for the observed product choices. Compared to the benchmark model predictions in Table 12, the restricted model seems to attribute the effect of selection on unobservables on product choices at least partially to broker steering. For instance, the contribution of broker steering in the low-LTV/low-amortization and the high-LTV/high-amortization bins are close to the sum of the contribution from both broker steering and selection on unobservables reported in Table 12. Therefore, the restricted model overestimates the contribution of broker steering away from low-LTV/low-amortization products and the contribution of broker steering towards high-LTV/high-amortization products.

Table A2: Restricted model: Product choice model estimates

	(1)	(2)	(3)
	$\ln\theta$	$\bar{d}$	$\ln\delta$
Constant	21.4666 (0.1207)	-9.0098 (0.0848)	2.9206 (0.3062)
FICO∈[680,717]	0.2108 (0.0093)	0.0882 (0.0069)	0.2815 (0.0227)
FICO∈[718,759]	0.129 (0.0082)	0.0323 (0.0061)	0.1861 (0.021)
FICO≥760	-0.1699 (0.008)	-0.1109 (0.0057)	0.0012 (0.021)
Borrower age	-0.0049 (0.0003)	0.0003 (0.0002)	-0.0003 (0.0007)
Income (log)	0.2005 (0.0086)	1.4665 (0.0071)	0.2223 (0.0242)
New property	-0.0228 (0.0064)	0.0478 (0.0049)	0.1814 (0.0161)
House price (log)	-2.1082 (0.0098)	-0.6304 (0.008)	-0.4322 (0.0276)
Bond rate	0.1871 (0.0109)	0.0129 (0.0084)	-0.1932 (0.0303)
Rate premium	-0.0867 (0.0288)	-0.0237 (0.0231)	-0.0053 (0.0785)
$y\theta$		-0.0194 (0.0008)	
σ_θ	0.6027 (0.0066)		
σ_d		0.5091 (0.0012)	
σ_δ			0.9054 (0.0088)
Log likelihood		-96,106.13	

Note: This table presents the maximum likelihood estimates of a restricted model, where we do not control for the endogeneity of mortgage rates and broker choices using the (generalized) residuals obtained from regressions reported in Table 5. Estimates on year and region dummies are not reported. Standard errors in parentheses. Given the log likelihood in Table 6 and Table A2, we can test the null hypothesis that the coefficients on the control function proxies are all 0. The LR statistic is $2 \times (96,106.13 - 95,633.69) = 944.88$. The critical value of $\chi^2(6)$ distribution associated with the 0.1% significance level is 22.46.

Table A3: Restricted model: Product preference estimates

	LTV=80%	LTV=85%	LTV=90%	LTV=95%	LTV=100%
AMT=25Y	(0,0)	(-28.04,2.28)	(-66.44,5.96)	(-110,11.01)	(-183.43,23.13)
AMT=30Y	(-4.59,1.29)	(-34.81,4.4)	(-74.95,8.9)	(-118.09,14.08)	(-186.04,26.06)
AMT=35Y	(-5.32,2.66)	(-36.54,6.12)	(-77.11,10.73)	(-120.32,16.62)	(-184.89,27.87)
AMT=40Y	(-7.86,4.98)	(-37.78,6.47)	(-78.68,11.04)	(-121.98,17.14)	(-183.8,28.01)

Note: In each cell, the first number is the estimated product preference at branch channel ($\mu_{j,broker=0}$), while the second number is the estimated difference between origination channels ($\mu_{j,broker=1} - \mu_{j,broker=0}$). In order to test the null hypothesis: $\mu_{j,broker=1} = \mu_{j,broker=0}, \forall j$, we estimate the model under the null and obtain a likelihood ratio test statistic of 664.59. The critical value of $\chi^2(19)$ distribution associated with the 0.1% significance level is 43.82.

Table A4: Restricted Model: Product shares in broker channel (%)

	LTV=80%	LTV=85%	LTV=90%	LTV=95%	LTV=100%	Total
	Panel A: Model					
AMT=25Y	0.94	2.47	6.65	15.09	3.87	29.02
AMT=30Y	0.26	0.73	2.08	5.41	2.39	10.86
AMT=35Y	0.33	1.25	3.84	12.63	5.19	23.23
AMT=40Y	0.38	1.61	4.98	18.12	11.79	36.89
Total	1.91	6.05	17.55	51.25	23.24	100.00
	Panel B: Model w/o broker steering					
AMT=25Y	1.16	3.37	8.94	19.51	3.65	36.63
AMT=30Y	0.35	0.91	2.17	5.96	1.68	11.07
AMT=35Y	0.49	1.28	3.65	10.21	3.05	18.67
AMT=40Y	0.27	1.89	5.60	17.37	8.50	33.63
Total	2.26	7.45	20.35	53.05	16.89	100.00
	Panel C: Model w/o broker steering, branch-like control function variables					
AMT=25Y	1.16	3.37	8.94	19.51	3.65	36.63
AMT=30Y	0.35	0.91	2.17	5.96	1.68	11.07
AMT=35Y	0.49	1.28	3.65	10.21	3.05	18.67
AMT=40Y	0.27	1.89	5.60	17.37	8.50	33.63
Total	2.26	7.45	20.35	53.05	16.89	100.00
	Panel D: Model w/o broker steering, branch-like control function variables, rate setting same as branch					
AMT=25Y	1.23	3.44	9.02	19.42	3.64	36.75
AMT=30Y	0.37	0.94	2.25	6.00	1.69	11.25
AMT=35Y	0.52	1.31	3.71	10.12	2.99	18.64
AMT=40Y	0.31	1.97	5.71	17.17	8.19	33.35
Total	2.42	7.67	20.69	52.72	16.50	100.00
	Panel E: Model w/o broker steering, branch-like control function variables, rate setting, and observable characteristics same as branch					
AMT=25Y	1.22	3.52	9.62	22.82	4.81	41.98
AMT=30Y	0.28	0.78	1.92	5.52	1.68	10.18
AMT=35Y	0.44	1.10	3.27	9.93	3.07	17.81
AMT=40Y	0.20	1.57	4.79	15.30	8.16	30.03
Total	2.15	6.97	19.60	53.57	17.72	100.00

Note: In panel A we present the predicted product shares in our benchmark model. Panel B presents product shares under the restriction that the product dummies across channels are equal. Panel C presents product shares in a version of the model where we set the control function variables in the random coefficients to mimic those of the branch borrowers. This leads to no change since the restricted model does not include the control function variables. In panel D we simulate broker borrowers' mortgage rates using the first-stage rate regression estimates from branch borrowers (column (3) in Table 5). For each broker borrower, we draw the residual rate from the corresponding distribution in branch channel, preserving the percentile ranking. Panel E draws observations from the branch borrower sample. The sampling scheme is weighted to mimic the share for each market-quarter.

Table A5: Restricted model: Changes in product shares from Panel A to E

	LTV≤95%	LTV=100%	**Total**
AMT=25Y	12.02 (65.13%)	0.94 (-22.46%)	12.96 (58.76%)
AMT>25Y	-6.5 (22.67%)	-6.47 (95.01%)	-12.96 (58.76%)
Total	5.52 (115.08%)	-5.52 (115.08%)	0

Note: In each cell, the first number shows the percentage point change in product share from Panel A to Panel E in Table A4. The fraction of the change that is due to broker steering (Panel A to Panel B) is shown in parentheses.