Market Power in Mortgage Pricing: the Role of Referral Lending

Dayin Zhang   Panle Jia Barwick   Lu Han   Jon Kroah*

April 2026

Abstract

Despite intense competition among mortgage lenders, borrowers continue to face elevated rate spreads and substantial price dispersion. We argue that realtor–loan officer referral networks are a key source of lender market power: by steering homebuyers toward a limited set of loan officers, these networks restrict effective borrower choice even in otherwise competitive markets. Using a novel dataset linking 81,306 realtors to 102,860 loan officers in 41 states, we document that such networks are both pervasive and highly concentrated: 85% of realtors direct over 40% of their clients to fewer than four loan officers. Strikingly, the lender concentration among realtors persists and even *increases* in markets with more lenders, suggesting that referrals constrain choices regardless of market structure. IV estimates indicate that borrowers working with referred loan officers pay 18.6 basis points higher interest rates, equivalent to $2,609 in upfront costs for the average loan of $306k. The referral premium is nearly three times as high for Hispanic borrowers as for White borrowers, and is systematically higher for Black borrowers and financially constrained households. On average, referral lending raises rate spreads by 36.5% and explains half of the (residual) standard deviation of rate spreads after controlling for lender, market, and time fixed effects. We identify two mechanisms: referrals reduce borrowers' search intensity for lenders, and referred loan officers exercise pricing power relative to other officers within the same lending institution. Efficiency arguments (faster processing) and mediation of denial risks don't fully justify the referral premium. Our findings reveal referral networks as a hidden source of market power, imposing substantial financial costs and raising equity concerns for borrowers.

*Zhang: Wisconsin School of Business, UW-Madison, dayin.zhang@wisc.edu. Barwick: Department of Economics, UW-Madison, NBER, and CEPR. pbarwick@wisc.edu. Han: Wisconsin School of Business, UW-Madison, lu.han@wisc.edu. Kroah: Department of Economics, UW-Madison, jkroah@wisc.edu. We would like to thank the discussants, Luis Lopez and David Zhang, as well as seminar participants at WFA, ASSA-AREUEA Conference, the Singapore Management University, UW-Madison and University of Hawaiì at Mānoa, for their helpful comments. All remaining errors are our own.

Keywords: realtors, mortgage brokers, referral network, steering

JEL Classification Codes: D40, G21, L14, L85, R31

1 Introduction

In the U.S. mortgage market, extensive research has documented significant rate spread and substantial price dispersion in mortgage rates (Bhutta et al., 2025; Alexandrov and Koulayev, 2018), suggesting that mortgage lenders possess a certain degree of market power. However, the mortgage lending market is considered highly competitive, with many cities or local markets exhibiting low lender concentration. Moreover, existing studies have failed to find a relationship between mortgage rates and local lender concentration (Amel et al., 2018; Buchak and Jorring, 2021). This raises a fundamental question: Where does mortgage lenders' market power come from?

We argue that standard HHI or concentration measures – whether defined at the level of cities, towns, or even sub-city neighborhoods – mask an important source of market power that operates through the home purchase process itself. Mortgage originations are heavily intermediated by real estate agents, who frequently refer buyers to a limited set of loan officers. These referral practices effectively compress borrowers' choice sets, funneling them into highly concentrated local lending environments even in otherwise competitive markets.[1] In such settings, referred lenders acquire local monopoly power over borrowers.

To quantify the prevalence of this referral channel and to assess its implications for borrower outcomes, this paper provides the first national, systematic evidence on realtor–loan-officer referral networks in the mortgage market. We assemble a novel dataset linking 81,306 realtors to 102,860 loan officers across 703 counties in 41 states from January 2018 to December 2021. Using detailed transaction-level data on home purchases and mortgage loans, we map the referral network between realtors and loan officers and evaluate each realtor's reliance on particular loan officers. Specifically, we calculate the share of each realtor's transactions handled by individual loan officers and use these shares to compute the loan officer concentration ratio ($CR4$) for each realtor.

Our data reveal several striking patterns. First, 85% of realtors exhibit substantial concentration in their loan officer networks, directing over 40% of their clients to fewer than four loan officers. Second, and more importantly, this concentration persists and even *increases* in markets with greater lender competition. In other words, having more lenders in a market does not translate into broader choices for borrowers; instead, realtors channel clients to a limited set of loan officers regardless of how many alternatives exist. This pattern suggests that homebuyers' actual mortgage choices are not determined by the market structure of lenders but are heavily constrained by realtor referral practices.

Whether such concentration translates into higher borrowing costs is an open question. We find

[1] We use "lenders" for lending institutions and "loan officers" for individuals, including both in-house loan officers of mortgage lenders and independent mortgage brokers.

that it does – substantially. Borrowers working with realtors with top-quintile concentration (in terms of *CR*4) pay 28.9 basis points more than those with bottom-quintile realtors, with rate differentials increasing monotonically across concentration quintiles. Even after controlling for market-specific trends, borrower characteristics (age, income, FICO), loan characteristics (loan-to-value ratio (LTV), debt-to-income (DTI) ratio, etc.), house attributes (size, bedrooms, ZIP code, etc.), and a rich set of fixed effects, the differential remains economically and statistically significant at 11.7 basis points.

The pattern is particularly striking in light of our earlier observation: while market-level lender concentration shows no relationship with mortgage rates, realtor-level concentration exhibits a strong, monotonic relationship. This contrast suggests that broader market competition fails to discipline mortgage pricing because realtor-driven referral networks constrain borrowers from shopping across the full range of available lenders. Within-realtor concentration appears to be a more relevant measure of market power, directly contributing to higher financing costs for homebuyers.

However, this evidence alone does not establish a causal effect of the realtor-loan officer referral network on borrowing costs. The observed correlation between realtor-level concentration and homebuyers' mortgage rates could reflect borrower choices rather than realtor steering. For example, borrowers might independently gravitate toward a dominant local lender, choose loan officers located near their home, or rely on preexisting personal or professional relationships. These forces can generate high loan-officer concentration at the realtor level even in the absence of active referrals, making the concentration measure an outcome of borrower preferences rather than a constraint on borrower choice.

To address this identification challenge, we employ an instrumental variable (IV) strategy based on "excess concentration" in realtors' loan-officer usage. Our approach proceeds in two steps. First, we estimate a discrete choice demand model (Berry et al., 1995) of borrowers' lender choices using refinance mortgages, where realtors are largely absent.[2] This model recovers borrower preferences over lender characteristics and allows us to predict what lender choices would look like absent referral influence. Second, we apply these estimated preferences to our primary dataset of home purchase mortgages to simulate a counterfactual concentration level for each realtor, the concentration that would arise if their clients chose lenders based solely on borrower preferences and lender attributes. Our instrument is the difference between observed and counterfactual concentration, which isolates the extra concentration plausibly generated by referrals. Under the exclusion restriction that this excess concentration affects mortgage pricing only through referral intensity (conditional on our extensive set of controls), the IV identifies the causal effect of using a referred loan officer.

[2] In our sample, 74% of refinance borrowers use a different loan officer than they did at purchase. Of the 26% who retain the same officer, roughly 56% were matched to that officer through a realtor referral at the time of home purchase (Section 4.2), implying that only 15% of refinance borrowers carry forward a referral-influenced relationship.

The IV estimates reveal substantial borrowing costs. Homebuyers who use referred loan officers pay 18.6 basis points higher mortgage rates. At the average loan size of $306k, this translates to $569 in additional annual interest payments, or equivalently, $2,609 in upfront points needed to buy down the rate to non-referral levels. To gauge magnitude, the average rate spread is 51 basis points, and the residual standard deviation after controlling for lender, market, and quarter fixed effects is 37 basis points. Our estimates imply that referral lending raises rate spreads by 36.5% and accounts for 50% of this residual variation. Put differently, *half* of the cross-sectional variation in mortgage prices stems from loan-officer market power mediated by realtor referral networks.

The burden of realtor referral costs falls unevenly across borrowers. Referral premia are systematically higher for riskier borrowers on standard underwriting dimensions. Borrowers with DTI above 45% or LTV above 95% pay additional referral markups of 10.9 to 21 basis points beyond the baseline referral effect, suggesting that financially fragile households, those least able to absorb extra costs, are most exposed to the costs of steering. Even starker disparities emerge along racial and ethnic lines. Hispanic borrowers face referral premia of 49.1 basis points ($6,888 in upfront costs), nearly triple the baseline effect. Black borrowers pay premia of 24.8 basis points ($3,479 upfront), while White borrowers pay 17.2 basis points ($2,413 upfront). Income-based disparities follow a similar pattern: low-income borrowers pay 24 basis points in referral premia compared to 14 basis points for high-income borrowers. Taken together, these results imply that realtor–loan-officer referral networks not only raise borrowing costs on average, but also disproportionately shift the burden toward minority and financially vulnerable households, exacerbating distributional and fairness concerns.

We present two pieces of evidence that shed light on the underlying mechanisms. First, using credit-bureau inquiries to measure mortgage search, we find that borrowers who follow their realtor's referral contact 12 percent fewer lenders and are 22 percent less likely to obtain multiple quotes. This reduction in search intensity likely compresses borrowers' choice sets and weakens their outside options, shifting bargaining power toward lenders during rate negotiation. Second, referral-induced market power operates not only across lenders but also within lending institutions. Conditioning on lender fixed effects, referred borrowers pay 7.5 basis points more than comparable non-referred borrowers obtaining loans from the same lending institution. This within-lender premium indicates that referrals confer pricing power on individual loan officers by insulating their clients from competition with other loan officers within the same firm. Decomposing the overall referral premium, approximately three-fifths reflects steering borrowers toward higher-cost lenders (a between-lender channel), while two-fifths arises from additional markups charged by referred loan officers relative to their colleagues (a within-lender channel).

A natural question is whether realtor-loan officer referrals reduce frictions in the mortgage application process, potentially making referral lending efficiency-enhancing. We test this directly. Conditional on signing a mortgage contract, borrowers who use referred loan officers close their purchases 1–2 days faster relative to a baseline of 40 days. This acceleration reflects lower coordination and screening frictions within a referral-mediated match: loan officers have incentives to prioritize referred clients and documentation is streamlined. However, while this represents a 4 percent reduction in processing time, the benefit is modest relative to its cost. The one-to-two-day time savings is accompanied by $2,609 in higher upfront borrowing costs, or roughly $1,300 per day saved – a steep premium for a marginal convenience gain.

We further test whether the referral premium compensates for reduced approval uncertainty. If referred loan officers help marginal borrowers secure financing they might otherwise be denied, higher interest rates could reflect compensation for lower denial risk. Because our data do not record failed home purchases associated with denied applications, we take an indirect approach and focus on high-quality borrowers with minimal approval uncertainty — those with FICO scores above 780, DTI below 36 percent, and LTV below 90 percent. Even among these borrowers, referred loans remain 9.3 basis points more expensive. The persistence of the referral premium among borrowers facing virtually no risk of denial rules out mediation of approval risks as the primary explanation.

Our paper contributes to three strands of literature. First, it adds to the growing body of research on price dispersion and market power in the mortgage market. Numerous studies have documented substantial dispersion in mortgage rates (Bhutta et al., 2025; Agarwal et al., 2024; Alexandrov and Koulayev, 2018), yet research has generally failed to uncover a cross-sectional link between mortgage pricing and local lender concentration (Amel et al., 2018; Buchak and Jorring, 2021). This has shifted attention toward alternative explanations, such as borrower sophistication (Bhutta et al., 2025), adverse selection arising from the costly mortgage process (Agarwal et al., 2024), and deceptive advertising (Gurun et al., 2016). From a policy perspective, the absence of evidence connecting mortgage rates to local lender concentration has led the Federal Reserve to treat mortgage markets as national in scope and to discount the potential impact of bank mergers on local mortgage competition.

We show that this view misses a key mechanism linking concentration and mortgage pricing. Traditional market-level concentration measures fail to capture how borrowers actually interact with lenders. In practice, realtor referral networks shape the set of options available to buyers, limiting their effective choice to a handful of loan officers even in markets with many competing lenders. The source of lender market power lies not in market structure per se but in the mortgage application process itself: realtors frequently steer buyers toward preferred lenders, creating localized pockets of market power within their referral networks. Measuring concentration at the realtor level provides a more relevant indicator of market power, one that directly translates into higher financing costs for homebuyers. Our findings complement those of Allen et al. (2014) and Allen et al. (2019), who study the more concentrated Canadian mortgage market and attribute lender market power to search frictions and brand loyalty.

Second, our paper also contributes to the literature on referral networks. Referral mechanisms have been studied in labor markets (Burks et al., 2015; Pallais and Sands, 2016; Chen-Zion and Rauch, 2020; Barwick et al., 2023), health care (Ho and Pakes, 2014; O'Malley et al., 2021; Sarsons, 2024), and education (Card and Giuliano, 2016; Cestau et al., 2017), but evidence outside these domains remains scarce. The limited attention is due in part to the informal nature of referral ties and the difficulty of observing them systematically. To our knowledge, we provide the first systematic evidence on referral networks that link realtors to mortgage lenders, an important but largely overlooked dimension of household financial decision-making.

Finally, our paper advances the literature on intermediation and brokerage in both housing and mortgage market. On the housing side, prior work has shown that buyers' realtors often steer buyers toward properties listed by realtors from the same brokerage firm (in-house listings) or properties paying high commissions, allowing brokerages to capture higher commissions and close deals more quickly (Han and Hong, 2016, 2011; Han and William, 2014; Barwick et al., 2017). On the mortgage side, the existing literature has focused primarily on independent mortgage brokers, who played a central role in mortgage origination before the Global Financial Crisis but have since declined to a relatively small share of front-line origination activity. Previous studies show that mortgage brokers raise the cost of financing (Ambrose and Conklin, 2014; LaCour-Little, 2009; Ernst et al., 2008) and steer borrowers toward higher-rate or riskier mortgage products (Spader and Quercia, 2011; Berndt et al., 2010). Woodward and Hall (2010) and Woodward and Hall (2012) provide direct evidence that mortgage brokers capture a substantial share of the "yield-spread premium" as profit. In contrast, Robles-Garcia (2020) argues that mortgage brokers facilitate the entry of small, new lenders with limited branch networks and lower brand recognition, increasing competition in the mortgage market. Agarwal et al. (2021) further shows that sole mortgage brokers, under financial regulatory oversight, adopt a more stringent screening process that improves loan performance.

Most of the existing literature, however, studies realtors and brokers separately. The exception is Jorgensen (2024), who studies over 100 mergers between real estate agencies and mortgage lenders and finds that borrowers using the jointly owned lender pay roughly 9 bps more in interest, illustrating the increased market power that vertical integration confers. Our paper differs by providing the first comprehensive nationwide mapping of the referral network between realtors and mortgage brokers. We quantify the effect of these networks on mortgage rates, including their distributional implications, shed light on the underlying mechanisms, and evaluate non-financial borrower outcomes, such as the duration of the mortgage application process and denial risks.

The remainder of the paper is structured as follows. Section 2 describes the data sources and provides institutional background on realtor–loan-officer referral networks. Section 3 presents stylized empirical facts on the structure and prevalence of these networks. Section 4 examines their implications for homebuyers' financing costs and Section 5 documents heterogeneous effects across homebuyers. Section 6 explores the underlying mechanisms and potential efficiency rationales for referral networks. Section 7 concludes.

2 Institutional Background and Data

2.1 Institutional Background

Realtors, or real estate agents, play a central role in U.S. residential property transactions. They offer expertise on local housing markets and provide a wide range of services tied to the purchase and sale of homes, typically on a commission basis. For sellers, realtors handle tasks such as marketing the property, recommending listing prices, hosting open houses, and negotiating with potential buyers. For buyers, realtors identify homes that meet client preferences, schedule property showings, and negotiate purchase terms with sellers. Importantly, beyond assisting with the property search itself, realtors often guide clients through critical components of the transaction process, including recommending home inspectors, real estate attorneys, and, notably, mortgage loan officers. According to a 2016 survey conducted by Freddie Mac (FreddieMac, 2016), 84% of real estate professionals maintain a preferred network of lenders to whom they typically refer clients. Among these agents, 73% reported working with just one to three lenders, while another 24% had networks comprising four to six lenders. These referrals appear to be highly effective: 76% of all agents surveyed said their clients "always" or "often" follow their lender recommendations. This share rises to 87% among high-volume realtors who complete more than 20 home transactions annually, underscoring the influential role that realtors play in shaping borrowers' financing decisions.

The importance of realtors in homebuyers' choice of mortgage lenders is also reflected by the home buyers' survey. According to the 2018–2022 quarterly National Survey of Mortgage Originations (NSMO), jointly administered by FHFA and CFPB, more than half of homebuyers consider realtors' recommendations important when shopping for mortgages. Specifically, for the question "How important were each of the following in choosing the mortgage lender/broker you used for the mortgage you took out," 5,578 (54%) out of 10,343 survey subjects checked "recommendation from a real estate agent/home builder," ahead of all other factors, including an established banking relationship, having a local office, or recommendations from friends, relatives, or co-workers.

Mortgage loan officers intermediate between homebuyers and lenders. They take applications, explain loan products, collect documentation, coordinate with underwriters, and manage the transaction through closing; the lender (an institution) formally issues and funds the loan. In-house loan officers, employed by banks or non-bank mortgage companies, offer only their employer's products, while independent loan officers work with multiple lenders on the borrower's behalf. The role of independent loan officers contracted sharply after the Global Financial Crisis, from an estimated 68% of originations at their 2004 peak to about 9% by 2014.[3] In our sample period, more than 90% of loan officers are in-house originators.

Within a lender's pricing guidelines, loan officers exercise substantial discretion over the rate quoted, points charged, and speed of processing. This discretion can simplify the borrowing process, but it also creates scope for recommendations to be shaped by ongoing relationships with specific realtors or lenders.

The collaboration between realtors and loan officers is often reciprocal. Realtors frequently refer their clients to loan officers to streamline mortgage financing, while loan officers may direct pre-approved borrowers to trusted realtors to complete the home purchase. This referral system can generate efficiency gains by fostering long-term, cooperative relationships that improve coordination, reduce transaction frictions, and enhance the overall client experience. On the other hand, these same relationships can facilitate steering, where realtors guide homebuyers toward a narrow set of preferred loan officers, potentially limiting borrower choice and weakening price competition among lenders.

Although Section 8 of the Real Estate Settlement Procedures Act (RESPA) prohibits kickbacks and unearned fees to realtors in exchange for referrals in mortgage lending, several common industry practices occupy legal and regulatory gray areas. Loan officers, for instance, are permitted to give gifts to realtors, provided the gifts are not explicitly tied to referral activity. Yet such arrangements are difficult to monitor, making law enforcement challenging. In addition, loan officers often sponsor realtor events or provide financial support for joint advertising campaigns. These co-marketing efforts can shape borrower flows and deepen referral-based relationships, even when no direct compensation is exchanged.

[3] See Reuters, "Mortgage brokers cashed in on US housing bounty," https://www.reuters.com/article/world/mortgage-brokers-cashed-in-on-us-housing-bounty-idUSKIM462166/. The decline reflected the collapse of the wholesale channel, the withdrawal of major banks from broker relationships, and regulatory reforms including the SAFE Act (2008), Dodd-Frank's ban on yield-spread premiums (2010), and the LO Compensation Rule (2011). See National Mortgage Professional, https://nationalmortgageprofessional.com/news/54078/perceptual-and-regulatory-changes-spark-resurgence-brokers-mortgage-industry.

2.2 Data

CoreLogic Multiple Listings, Ownership Transfer, and Mortgage Data We link the three CoreLogic data products – data on Mortgage, Ownership Transfer, and Multiple Listing Service (MLS) transactions – to create a panel dataset of mortgage-financed home purchases and the associated buyer agents. The sample covers listings dated between July 1, 2017, and December 31, 2021. Our choice of time window is guided by the availability of detailed loan information in the HMDA data, especially interest rates and mortgage fees, which became more comprehensive beginning in January 2018 (as described below). Because mortgage origination typically occurs near the end of the home purchase process, we include housing listings since mid-2017 to ensure adequate coverage of transactions that culminate in mortgage activity in early 2018.

Geographically, we restrict our analysis to counties that meet three key criteria. First, over 80% of housing listings in the county must originate from the same Multiple Listing Service (MLS), allowing us to avoid complications associated with cross-listings on multiple platforms. Second, the dominant MLS must provide unique realtor identifiers.[4] Third, because our analysis requires information on the number of days between purchase contract acceptance and closing, we limit the sample to counties where the dominant MLS reports non-missing contract dates[5] for at least 50% of listings. After applying these filters, the final sample comprises 703 counties across 41 states. Figure 1 maps these counties, with different colors indicating different MLSs. The sample covers most large population centers and accounts for roughly 40% of the US population.

We then apply several filters to the housing listings to ensure data consistency. First, we exclude rental listings and split property listings (e.g., duplexes where each unit is listed separately). Second, to avoid duplication and capture only distinct sale attempts, we group sequential listings of the same property occurring within a 90-day window into a single "listing event". Third, we retain only listings that resulted in a successful sale financed by a mortgage, which accounts for approximately 76% of home purchases. These restrictions are necessary because we do not observe buyer agents or loan officers otherwise. Finally, we exclude any transactions missing either the buyer agent's identifier or the loan officer's Nationwide Multistate Licensing System (NMLS) ID, ensuring the reliability of the resulting dataset.

After implementing these filters, our dataset comprises 5.1 million mortgage-financed home purchases with both realtor and loan-officer identifiers. To guard against spurious concentration

[4] Many MLS systems do not include buyer agent IDs, which would require us to merge datasets using agent names, an approach prone to inconsistencies and ambiguities due to name variations.

[5] The contract date (sometimes called the contract-acceptance date or purchase-agreement date) is the date on which the buyer and seller sign the purchase agreement. It marks the start of the window – typically 60 to 90 days – in which the buyer arranges mortgage financing and completes closing.

arising from low transaction counts, we drop realtors associated with fewer than ten purchase loans during the sample period. The resulting core sample contains 3.6 million observations. All of our realtor–loan-officer network measures are constructed based on this sample. On average, each county-year has 220 active realtors, 73 lending institutions, and 232 loan officers, as shown in the market-level statistics in Table A1.

Referral Network Proxy Since direct observations of realtor–loan-officer referrals (e.g., explicit recommendations or formal agreements) are unavailable, we infer these networks by measuring the concentration of loan officers within each realtor's lender network. The intuition is straightforward: if buyers independently selected their lenders, loan officer usage should be dispersed across lenders. In contrast, a realtor who routinely channels clients to the same small group of loan officers will exhibit a highly concentrated usage pattern – our proxy for systematic referral behavior.

To that end, we employ two standard measures of concentration: the four-firm concentration ratio ($CR4$) and the Herfindahl-Hirschman Index (HHI). For each realtor, we define the loan officer network as the set of all loan officers who financed mortgages for at least one of the realtor's buyer clients. We then calculate each loan officer's share within that network – the fraction of the realtor's homebuyer clients whose loans were underwritten by that officer. The $CR4$ is the sum of the shares of the top four loan officers for a given realtor, while the HHI is the sum of squared shares across all loan officers in the realtor's network.

Following Dranove et al. (2017), we classify realtors with $CR4 \geq 0.4$ as medium concentration and those with $CR4 \geq 0.7$ as high concentration. Similarly, following the 2010 and 2023 DOJ Horizontal Merger Guidelines, we use $HHI \geq 1500$ and $HHI \geq 2500$ as cutoffs for medium and high concentration. High values of $CR4$ or HHI provide suggestive evidence of referral networks. We incorporate both measures in our regression analyses below.

It's important to emphasize that our measures of concentration are based on interactions between realtors and individual loan officers, rather than between realtors and lending institutions. First, referrals are inherently interpersonal, occurring at the realtor-loan-officer level. Second, large lending institutions may employ hundreds of officers, yet realtors typically refer business to only a handful – often one to three – within any given institution. Third, as discussed above, a typical county-year has on average 232 active loan officers and 220 active realtors. Absent a deliberate referral relationship, it is difficult to imagine that as few as four officers could account for 40 percent or more of a realtor's purchase loans ($CR4 \geq 0.4$).

HMDA Data We next obtain detailed mortgage characteristics from the Home Mortgage Disclosure Act (HMDA) data, which provide loan-level information for the majority of mortgages issued in the U.S. A major expansion of HMDA reporting took effect in 2018, significantly enhancing the level of detail available. Specifically, the post-2018 HMDA data include key financial terms such as mortgage interest rates and origination fees, as well as borrower- and loan-level characteristics, including loan-to-value (LTV) and debt-to-income (DTI) ratios. Our analysis uses HMDA panel data covering January 2018 to December 2021.

Following standard practice in the literature (e.g., Kermani and Wong, 2026), we merge originated mortgages from HMDA with CoreLogic mortgage data using overlapping fields — lender name, loan amount, LTV, and property census tract — and require the mortgages to be originated in the same year. To ensure match quality, we retain only one-to-one matches, where exactly one mortgage in each dataset shares these same attributes.

We successfully match 2.5 million of the 3.6 million mortgage-financed home purchases in our core sample to records from HMDA. To construct our final regression sample, we impose several additional restrictions. First, we retain only first-lien, 30-year fixed-rate mortgages, which are the most common mortgage product in the U.S. Second, we exclude loans with an original balance below $100,000. Third, we restrict the sample to mortgages with non-missing values for key borrower and loan characteristics, including loan-to-value (LTV) ratio, debt-to-income (DTI) ratio, borrower age group, income, and joint application status. Finally, we drop realtors who represent fewer than ten purchases in the CoreLogic–HMDA matched sample, to ensure sufficient observations for simulating each realtor's counterfactual loan-officer concentration measure (used in the IV construction).The resulting full regression sample includes 1.43 million transactions, intermediated by 81,306 distinct realtors.[6]

Credit Bureau Data We augment our full regression sample by matching it to a 1% random sample from an anonymous national credit bureau, one of the three major U.S. credit reporting agencies. For each individual in this random sample, we observe annual consumer credit snapshots (as of year-end) that report balances and payment information across major types of formal debt – mortgages, student loans, credit cards, and other installment credit. These snapshots are linked to the underlying account-level records – known as tradelines – which report details of each individual credit account, such as origination date, credit limit or original balance, and monthly payment status. The data also include credit scores, which we use as additional controls for borrower creditworthiness, and records of credit inquiries, which we use to proxy for mortgage search intensity.

We match the full mortgage sample to this 1% credit-bureau sample using three key variables: mortgage origination date, loan amount (rounded to the nearest $100), and borrower ZIP code. The

[6]While our regression analyses are based on the CoreLogic-HMDA matched sample, we use the MLS core sample of 3.6 million home purchases to construct the realtor–loan-officer network, as it offers broader coverage.

match quality is exceptionally high. Given the low probability that multiple mortgages in the same ZIP code on the same day share exactly the same loan amount, the scope for ambiguous matches is minimal. The resulting 1% matched random sample contains 13,400 observations.

3 Summary Statistics and Descriptive Evidence

In this section, we first describe loan attributes and homebuyer demographics, then turn to the structure of realtors' referral networks.

3.1 Attributes of Mortgage Loans and Homebuyers

Table 1 presents summary statistics for the home purchase data used in this study. Column (1) reports statistics for the full regression sample, and Column (2) summarizes the 1% random sample matched to credit records.

Our full sample comprises 1,432,602 home purchases, with an average loan amount of $285K. Following Bhutta and Hizmo (2021), we measure borrowing costs using the annual percentage rate (APR) spread. APR is a composite cost measure that incorporates both the contract interest rate and upfront finance charges, such as fees and discount points, and is thus intended to capture the total cost of borrowing more comprehensively. HMDA reports the spread between a mortgage's APR and the benchmark prime mortgage rate for a comparable loan originated at the same time.

The average APR spread in the full sample is 55 basis points. Joint applications account for 43% of mortgages, and 97% of loans fall within the conforming limit. Borrowers exhibit an average loan-to-value ratio of 89 percent and a debt-to-income ratio of 35 percent.

The 1% random sample displays very similar mortgage characteristics, with an average loan size of $306K and an average APR spread of 51 basis points. The main difference from the full sample is a higher share of joint applications. This arises mechanically: for jointly held mortgages, each co-borrower appears separately in the credit files but shares the same mortgage tradeline, effectively doubling the probability that such loans are matched to the credit-bureau data. In this subsample, we observe additional information such as FICO scores and credit inquiry records. The average borrower has a credit score of 743, and, in the 90 days prior to closing, applies to lenders and triggers 1.45 mortgage-related credit inquiries on average. The number of credit inquiries serves as a useful proxy for borrowers' mortgage search intensity. We report results for both the full sample and the 1% sample, but rely on regression coefficients for the 1% sample when quantifying the impact on borrowing costs, as controlling for FICO scores is essential for mortgage pricing (though results are similar using the full sample).

Demographically, homebuyers in our sample earn on average 1.45 times their county's median income, where the average county median is $73.8K. About 13% of borrowers are Hispanic. By race, 73% are White and 27% are minorities (9% Black, 5% Asian, and 13% other minority groups). The average home has a purchase price of $330K, a living area of 2,102 ft^2, 3.4 bedrooms, 2.5 bathrooms, and an age of 40 years. The average time from signing the purchase agreement to closing is 40 days. Homebuyers in the 1% random subsample exhibit very similar demographic and housing characteristics, supporting the representativeness of this matched sample.

3.2 Descriptive Evidence on Realtor-Level Lender Concentration

Figure 2 plots histograms of realtor-level concentration measures across all realtors in our sample. Panel A shows that 30% of agents exceed the high-concentration benchmark ($CR4 \geq 0.7$), and 85% exceed the medium threshold ($CR4 \geq 0.4$). This prevalence closely matches external survey evidence from Freddie Mac (FreddieMac, 2016), which reports that 84% of real estate professionals maintain a preferred network of lenders to whom they typically refer clients. Panel B reports a similar pattern using the Herfindahl–Hirschman Index: 27% of realtors register $HHI \geq 2500$ (high concentration) and 52% exceed $HHI \geq 1500$ (medium concentration).

High loan officer concentration among realtors does not arise mechanically from small transaction counts or a limited number of local officers. The typical realtor completes 20 home purchases over our study period. Moreover, there are 232 loan officers affiliated with 73 active lending institutions in a typical county-year.[7] We next present further evidence that high loan officer concentration is likely driven by referrals.

3.3 Realtor-level vs. County-level Lender Concentration

A key insight from our analysis is that realtor-driven referrals systematically channel homebuyers toward certain loan officers and institutions, causing the distribution of loan officer market shares among homebuyers working with referring realtors to diverge sharply from the aggregate market.

Figure 3 compares loan officer concentration ($CR4$) at the county and realtor levels. Panel A plots the distributions of the two measures: the blue curve shows the combined market share of the top four loan officers in each county, while the red curve shows the combined share of the top four loan officers within each realtor's transactions. Concentration is substantially higher at the

[7] According to HMDA data for 2018–2021 (before merging with our core sample), each county–year features, on average, 197 active lending institutions – that is, all entities involved in originating new purchase loans, refinances, and home-equity lines. In contrast, our core sample is restricted to purchase loans and captures only a subset of those market participants. In either case, the sheer number of active lenders indicates a highly competitive lending environment.

realtor level, averaging 58 percent, compared with 19 percent at the county level. This pronounced gap suggests that realtor–loan officer relationships are unlikely to be random and instead reflect systematic, concentrated referral patterns.

To provide further evidence that the elevated realtor-level concentration reflects deliberate referral choices rather than mechanical market structure, Panel B plots both realtor-level and county-level $CR4$ against the number of loan officers in the county, with counties grouped into 100 bins. Consistent with the literature on industries with free entry (e.g., Bresnahan and Reiss, 1991; Dunne et al., 1988), county-level $CR4$ (blue triangles) declines as the number of loan officers rises: entry erodes individual market shares and intensifies competition. At first glance, this pattern appears to support policies that seek to reduce market power by expanding participation.

The pattern at the realtor level, however, is strikingly different. Realtor-level $CR4$ (red squares) remains above 0.4 across markets of all sizes and, if anything, increases with the number of loan officers in the market. Thus, even as the set of available loan officers expands, borrowers connected to a given realtor continue to be concentrated among a small subset of officers within that realtor's network. One possible interpretation is that greater competition among loan officers strengthens the incentive to cultivate realtor referral relationships, making those networks more rather than less important in determining borrower allocation.

These results carry important policy implications. Conventional antitrust and consumer-protection strategies have focused on encouraging firm entry and expanding market participation, assuming that more entrants will naturally broaden choice and lower costs. However, if realtor referrals steer borrowers toward a small subset of loan officers, simply increasing the number of lenders or loan officers may not deliver better pricing or enhanced borrower choice. In the next section, we examine how these referral-driven concentration patterns translate into differences in interest rates and closing costs.

4 Effect of Referral Network on Borrowing Costs

We next examine how realtors' loan-officer concentration relates to homebuyers' borrowing costs. Section 4.1 documents the correlation between realtors' network concentration and mortgage rate spreads. Section 4.2 introduces our preferred proxy for referral lending and presents OLS estimates of its effect on mortgage pricing. Section 4.3 develops the IV strategy and reports causal estimates of the referral premium.

4.1 Effect of Loan Officer Concentration on Mortgage Rates

A traditional benchmark in the mortgage-competition literature is *market-level lender concentration*, typically measured by the county-year $CR4$. If mortgage pricing were primarily shaped by the overall competitiveness of local lending markets, then this concentration measure should have explanatory power for borrowing costs. We begin by testing this prediction in a placebo exercise and present null results. We then show that our new measure—*realtor-level loan-officer network concentration*—has clear and economically meaningful explanatory power for mortgage rates.

Does Traditional Market-Level Concentration Explain Mortgage Rate Spreads? To assess whether local market structure explains mortgage pricing, we estimate differences in mortgage costs across counties with more- versus less-concentrated lender markets. Specifically, we sort county-years into quintiles based on market-level lender $CR4$ and estimate:

$$Y_{ilm} = \sum_{q=2}^{5} \tilde{\alpha}_q \, Quintile_q(CR4_m^{\text{lender}}) + X_{ilm}'\tilde{\gamma} + \tilde{\varepsilon}_{ilm}, \tag{1}$$

where Y_{ilm} measures the cost of the purchase mortgage obtained by borrower i (originated by loan officer l) in market m. In our main specifications, we use the "Annual Percentage Rate (APR) spread" reported in HMDA. APR incorporates both the contract interest rate and upfront finance charges associated with origination, thus providing a more comprehensive measure of borrowing costs than the contract interest rate alone. The APR spread is defined as the difference between the APR and the average APR offered on "prime", first-lien, 30-year fixed-rate purchase mortgages on the rate-lock date. This benchmarking helps address the fact that we do not observe the lock date for individual mortgages, which typically occurs 30–60 days before mortgage origination.

The key regressors $Quintile_q(CR4_m^{\text{lender}})$ are indicators equal to one if the lender concentration at market m falls into the qth quintile of the distribution, with the bottom quintile omitted as the reference group. The set of controls X_{irl} includes market-specific trends in mortgage costs (county, year-month fixed effects),[8] borrower demographics (ninety-nine income-ratio percentile bins, six 10-year age bins, and, in additional analyses, eight 40-point FICO score bins), loan characteristics (loan amount in \$100K increment bins, five 20-percentage-point LTV bins, six 10-percentage-point DTI bins, a conforming loan dummy, and a joint-application indicator), and house characteristics

[8] In this placebo analysis, m denotes a county-year market. Because $CR4_m^{\text{lender}}$ is defined at the market level and exhibits no variation within a county-year-month cell, we include county and year-month fixed effects separately (rather than county$\times$year-month fixed effects).

(integer bins for house age, the number of bedrooms and bathrooms, as well as eleven square-footage bins).

Table 2 reports estimates of Equation (1). We find no evidence that borrowers in markets with more concentrated lenders face higher borrowing costs. In the full-sample analysis (Panel A), several coefficients have the opposite sign of what one might expect: higher lender concentration is associated with lower mortgage costs. Panel B uses the 1% random sample matched to credit records and adds FICO controls; it shows no statistical correlation between market-level lender concentration and mortgage costs. These results are consistent with prior work, such as Buchak and Jorring (2021) and Amel et al. (2018), which find limited effects of market-level lender concentration on mortgage pricing.

Panel A of Figure 4 visualizes the relationship between mortgage rate spreads and market-level lender concentration across quintiles of market-level lender $CR4$, using 1% random sample estimates from Table 2. The dashed line corresponds to the specification with all controls except FICO scores (column 5), while the solid black line corresponds to the fully controlled specification including FICO scores (column 6). No systematic pattern emerges between market-level lender concentration and rate spreads.

Does Realtor-level Loan Officer Concentration Predict Borrowing Costs? We next examine whether the concentration of a realtor's loan-officer referral network helps explain mortgage pricing. In contrast to the placebo test, we estimate differences in mortgage rates paid by homebuyers represented by high- versus low-concentration *realtors*, controlling for market-specific rate trends and a rich set of borrower, loan, and market characteristics. Specifically, we sort realtors into quintiles based on the $CR4$ of their loan–officer network and estimate:

$$Y_{irl} = \sum_{q=2}^{5} \alpha_q Quintile_q(CR4_r^{\text{lender}}) + X'_{irl}\gamma + \varepsilon_{irl}, \tag{2}$$

where Y_{irl} measures the cost of the purchase mortgage obtained by borrower i, originated by loan officer l, and represented by realtor r. We again use the HMDA APR spread as the primary outcome. The key regressors $Quintile_q(CR4_r^{\text{lender}})$ are indicators equal to one if realtor r's lender concentration falls into the qth quintile of the distribution, with the bottom quintile omitted as the reference group. The set of controls X_{irl} is the same as in Table 2, except that we control for county $\times$ year-month FE, which is more demanding than county FE and year-month FE in Table 2.

Table 3 reports estimates of Equation (2). Panel A uses the full mortgage sample. Column (1) shows raw differences: borrowers represented by top-quintile realtors pay 25.9 bps more than those represented by bottom-quintile realtors (full sample mean spread = 55 bps), with rate premia rising monotonically across quintiles. Columns (2)–(5) add controls sequentially to examine how much of these premia are explained by observables. Adding county $\times$ year-month and loan amount bin fixed effects in Column (2) reduces the top-bottom quintile gap to 21 bps. Column (3) further adds borrower and loan characteristics, shrinking the gap to 18.7 bps. In Column (4), including DTI and LTV bins reduces the top-quintile premium to 14.6 bps. Finally, Column (5) adds house characteristics and ZIP code fixed effects, further reducing the top-quintile premium to 12.5 bps.

This extensive set of controls in the full-sample analysis misses one important borrower characteristic that lenders use in pricing mortgages – FICO scores. Because public HMDA data do not report FICO, we cannot control for credit scores in the baseline regressions. To assess how this omission affects the estimated cost differentials between borrowers represented by high- versus low-concentration realtors, we turn to a 1% random subsample matched to credit-bureau records, where FICO scores are observed. Panel B reports the corresponding estimates using this credit-enhanced sample.

We first replicate the specifications without FICO controls in columns (1)–(5). The point estimates and patterns of mortgage cost differentials are extremely similar to those in the full sample, confirming that the subsample is representative. As expected, standard errors are 2-3 times larger due to the smaller sample size. Nonetheless, the positive association between mortgage costs and realtor-loan officer concentration remains pronounced and robust. In column (6), we add FICO bin fixed effects and find that borrowers represented by top-quintile realtors pay 11.7 bps more than those represented by bottom-quintile realtors, slightly lower than but comparable to the corresponding estimates without FICO controls.

At the average loan size ($306K) for the 1% sample, an 11.7 bp premium translates into $358 in extra annual interest payments. Equivalently, for a borrower to obtain the same rate as those working with bottom-quintile realtors, they would need to pay about $1,647 in upfront fees (assuming a point-to-rate reduction factor of 4.6 as estimated in Bhutta et al. (2025)). These magnitudes underscore the tangible financial cost to homebuyers of working with highly concentrated referral networks and highlight the potential for steering to materially affect borrowing outcomes.

Panel B of Figure 4 plots these estimates visually, showing mortgage rate spreads across quintiles of realtors' loan-officer $CR4$. The black solid line reports estimates from the full specification with all controls, including FICO scores. Spreads increase monotonically with realtor-level concentration, in stark contrast to the null pattern for market-level concentration shown in Panel A.

4.2 Effect on Borrowing Costs: OLS

The preceding analysis shows that higher realtor-level concentration is associated with higher borrowing costs for homebuyers. In this subsection, we use an OLS framework to more directly quantify the effect of referral lending.

We proxy for referral relationships between realtors and loan officers by imposing two conditions on realtor-loan officer pairs. First, we require that the realtor operate a highly concentrated loan officer network, meaning that a large share of his/her past transactions are intermediated by a small set of loan officers. Second, within this concentrated network, we focus on loan officers who rank among the realtor's top four most frequently used counterparts. These top-ranked loan officers are the ones most plausibly receiving systematic referrals; hence, mortgages originated through these realtor-loan officer pairs are classified as *Referral* loans:

$$Referral_{rl} = \begin{cases} 1 & \text{if } CR4_r \geq 0.4 \text{ and } l \text{ is realtor } r\text{'s top-4 loan officer} \\ 0 & \text{otherwise.} \end{cases}$$

Two pieces of evidence support our *Referral* definition. First, 85% of realtors have loan-officer concentration above the 0.4 threshold (Figure 2). This is consistent with external survey evidence from Freddie Mac (FreddieMac, 2016), where 84% of real estate professionals maintain a preferred network of lenders. Second, a similar level of prevalence is observed on the borrower side (Table 1): 88% of homebuyers contract with a realtor who has a referral network under our definition. Among these borrowers, 56% follow the realtor's referral (choosing one of the top four loan officers), while the remaining 32% are matched with a loan officer outside the realtor's top four, whether the referral was declined or the borrower shopped independently. These patterns imply that 56% of all homebuyers in our sample are effectively influenced by realtor–loan officer referral networks. This share is both economically large and closely aligned with NSMO survey evidence that 54% of borrowers choose their lenders based on recommendations from a real estate agent (or home builder).

With these definitions, we estimate the effect of referrals on borrowing costs:

$$Y_{irl} = \alpha^{OLS} Referral_{rl} + X'_{irl}\delta^{OLS} + \varepsilon_{irl} \tag{3}$$

where Y_{irl} is the mortgage APR spread, and X_{irl} includes the same controls as in Equation (2): county $\times$ year–month fixed effects, borrower demographics, loan characteristics, and house attributes.

Table 4 reports OLS results. Column (1) estimates Equation (3) in the full sample and finds a coefficient of 0.099 on the Referral indicator, implying that borrowers who obtain loans through their realtors' loan-officer network and use a top-four loan officer pay about 9.9 bps higher mortgage rate spreads relative to the omitted group. Column (2) further separates borrowers who are in a concentrated realtor network but do not use one of the realtor's top-four loan officers (*Non-Compliant Referral*). The estimated premium for this group is economically small (1.2 bps) while the premium for the "compliant" referral group increases to 10.7 bps. The sharp gap between these coefficients suggests that higher borrowing costs are not simply a feature of being matched to a realtor who operates a concentrated network; rather, they are concentrated among borrowers who borrow through the realtor's most-connected loan officers.[9]

Columns (3) and (4) repeat the analysis using the 1% random subsample matched to credit-bureau records, allowing us to include FICO bin fixed effects. The referral premium remains statistically strong and similar in magnitude: 8.7–8.8 bps. In contrast to Column (2), the "Non-Compliant Referral" coefficient becomes economically and statistically negligible once FICO is controlled for, reinforcing the interpretation that the cost difference is tightly linked to following the referral channel rather than to unobserved differences in homebuyer characteristics across realtors.

4.3 Addressing Endogeneity Concerns: IV Estimates

The OLS estimates by themselves cannot be interpreted as causal effects of realtor-loan officer referral networks. They may be biased by factors that influence which loan officer a borrower ends up using, but have nothing to do with the realtor's referral. For example, if a single lender is dominant in a local market, many borrowers may use its loan officers regardless of what their realtor suggests. Similarly, borrowers may choose loan officers who are geographically close to their home or with whom they already have personal or professional ties. These forces can generate high concentration in observed realtor-loan officer pairings even in the absence of active steering, so OLS estimates alone do not cleanly isolate the effect of referrals.

To identify the causal effect of referrals, we implement an instrumental-variable (IV) strategy. We aim to isolate the component of realtor-level loan officer concentration that is driven by forces outside the borrower's own choice of lending representative. An ideal IV in this context shifts the concentration of loan officers in a realtor's network but has no direct effect on borrowers' mortgage costs other than through its impact on referral intensity.

The IV is constructed as the difference between (i) the observed concentration of loan officers in a realtor's network (e.g., $CR4$ or HHI) and (ii) a hypothetical concentration that would arise

[9] One reason the "Non-Compliant Referral" estimate in Column (2) may be slightly positive is that public HMDA does not report FICO scores; once we control for FICO in the credit-matched subsample, this coefficient becomes effectively zero (Column (4)).

if borrowers selected loan officers independently of any realtor influence. Intuitively, this "excess concentration" measures how much more concentrated a realtor's loan officer usage is than what we would expect in a world with no referrals. We then use this excess concentration as the source of quasi-experimental variation in referral intensity.

Our preferred IV takes the form $HHI_r - \widehat{HHI}_r$, and the corresponding first-stage specification is

$$Referral_{rl} = \gamma(HHI_r - \widehat{HHI}_r) + X'_{irl}\delta^{FS} + \varepsilon_{irl} \tag{4}$$

where HHI_r is the observed Herfindahl–Hirschman Index of realtor r's loan-officer network, and $\widehat{HHI}_r$ is the counterfactual Herfindahl–Hirschman Index that realtor r would obtain if borrowers selected loan officers independently of any realtor referrals, holding other conditions fixed (*ceteris paribus*). The difference $HHI_r - \widehat{HHI}_r$ therefore captures the "excess concentration" in a realtor's network attributable to referral behavior, which we use as the source of exogenous variation in referral intensity. We use the Herfindahl–Hirschman Index (HHI) as our concentration measure instead of $CR4$ to avoid a mechanical correlation between $Referral_{rl}$ and IV, though this choice is not crucial. As shown in Appendix Table A2, when we construct the instrument using alternative concentration measures ($CR4$), the estimated referral effects remain robust and, if anything, appear stronger.

Construction of $\widehat{HHI}_r$ The validity of our IV hinges on whether $\widehat{HHI}_r$ accurately captures the counterfactual concentration of loan officers. We combine an estimation and a simulation procedure to achieve this. First, we estimate a flexible BLP demand model that captures key determinants of borrowers' lender choice, including lending market structure, spatial proximity, and a rich set of borrower demographics and risk characteristics. Second, recognizing that loan officer choices in the purchase-mortgage sample are already influenced by realtor referrals, we estimate the model parameters using a separate *refinance*-mortgage sample, where realtor involvement is likely modest and referral pressure at the time of lender choice is largely absent. Empirically, about 26% of refinance borrowers in our sample retain the same loan officer they used at purchase. Given the 56% referral-compliance rate at purchase (Table 1), only 15% of refinance borrowers carry forward a referral-influenced relationship. While not perfectly free of referral influence, the refinance sample provides a reasonable basis for estimating $\widehat{HHI}_r$, the counterfactual loan officer concentration that would arise in the absence of realtor referrals.

We describe the data construction, BLP estimation, and counterfactual simulations in greater detail in Appendix Section B. Operationally, the construction involves three steps: (i) building a refinance–mortgage sample, (ii) estimating the BLP demand model, and (iii) running counterfactual simulations. In the first step, we extract all refinance mortgages originated in our markets over the sample period. We restrict to the same counties, years, property types, and mortgage characteristic filters as in the purchase sample, and retain only borrower–lender pairs with valid identifiers, producing a refinance dataset suitable for estimating the BLP demand model.

The second step recovers borrower preferences over lenders in the absence of referrals. We estimate the following BLP demand model (Berry et al., 1995) using refinance loans and incorporate a rich set of controls to allow flexible preferences across borrowers. Specifically, we assume that the utility of borrower i who secures a loan from lender j in market m is:

$$u_{ijm} = X'_{jm}\beta_1 + X'_{ijm}\beta_2 + \xi_{jm} + \varepsilon_{ijm}, \tag{5}$$

where X_{jm} is a vector of lender-market characteristics. We include dummies for whether lender j is a bank, a fintech firm, or an out-of-state lender (defined as having no branches within the state).[10] We also control for whether lender j is the first, second, or third-largest lender in market m (county-year).

The second set of regressors, X_{ijm}, is a vector of borrower-lender-market-specific characteristics. We include the distance between borrower i's property and lender j's nearest branch and interactions between $\text{age}_i \geq 65$ and the "bank" and "fintech" dummies. We also control for interactions between FICO score and the "bank", "fintech", and "top 1/2/3 lender" dummies as well as interactions between loan-to-value ratio and the "fintech" and "top 1/2/3 lender" dummies. Finally, ξ_{jm} represents an unobserved vertical component (such as quality of customer service) specific to lender j in market m, and ε_{ijm} is a Type I Extreme Value shock. Aggregating over choices by all borrowers delivers lender j's market share s_{jm}. We estimate preference parameters via maximum likelihood with a nested fixed point, where ξ_{jm} is inverted such that the model-predicted lender shares equal the observed lender shares in each market.

The lender-choice model implicitly assumes that borrowers choose a lending institution rather than a specific loan officer. This assumption is motivated by two considerations. First, we do not observe demographic or socioeconomic characteristics of individual loan officers, making it difficult to model choice among many observably similar officers. Second, absent referrals, borrowers are plausibly indifferent across loan officers within the same institution as long as contract terms are comparable. In the counterfactual simulations, borrowers first choose a lender according to this model and are then probabilistically assigned to one of that lender's loan officers in the same market, in proportion to each officer's observed lending volumes.

[10] We use the list of fintech firms from Fuster et al. (2019), available at https://pages.stern.nyu.edu/~pschnabl/data/data_fintech.htm.

After estimating the parameters in Equation (5), we apply them to the purchase-mortgage sample used in the main analysis and predict, for each borrower, the choice probabilities over all lenders available in the market. Using these predicted probabilities, we simulate 200 draws of lender and loan-officer choices for each homebuyer. Aggregating these simulated loan officer choices at the realtor level yields a counterfactual loan officer distribution for each realtor, from which we compute $\widehat{HHI}_r$, the average Herfindahl–Hirschman Index across 200 simulations: $\widehat{HHI}_r = \frac{1}{S}\sum_{sim=1}^{200}\widehat{HHI}_r^{sim}$.

Figure 5 compares the counterfactual and observed distributions of realtor-level loan officer concentration, and the contrast is striking. Panel A plots the distribution of $\widehat{HHI}_r$ (red dashed line), showing that, in the absence of referrals, most realtors would be matched to a highly diversified set of loan officers. The counterfactual distribution is mostly unconcentrated, with the probability that a realtor exceeds the medium-concentration threshold ($HHI \geq 1500$) close to zero. By contrast, the observed distribution of HHI_r (blue solid line) is dramatically shifted to the right: 52% of realtors exceed the medium-concentration threshold, and the distribution extends well into the high-concentration range. The gulf between these two distributions is far larger than what borrower preferences, geography, or lender market structure could plausibly generate, providing compelling visual evidence that referral networks are the dominant force shaping the concentration of loan officers within realtors' client bases. Panel B plots the distributions of simulated and observed $CR4$, which display a very similar pattern.

IV Results We next estimate the second-stage IV regression by replacing the endogenous referral indicator with its fitted value from the first stage in Equation (4):

$$Y_{irl} = \alpha^{2SLS}\widehat{Referral}_{rl} + X_{irl}'\delta^{2SLS} + \varepsilon_{irl} \tag{6}$$

The coefficient α^{2SLS} captures the causal effect of referral lending on mortgage outcomes.

Table 5 reports the IV estimates. Columns (1)–(2) present the first-stage regressions corresponding to Equation (4). Column (1) uses the full sample without FICO-bin controls, while Column (2) uses the 1% random sample and additionally controls for borrower credit quality. In both cases, the first stage is extremely strong: the "excess concentration" of a realtor's loan officer network has a large and precisely estimated effect on the probability that the borrower uses a loan officer within that network, indicating substantial predictive power for referral intensity.

Columns (3)–(4) report the IV estimates of the referral effect on borrowing costs. Without FICO controls, borrowers who use referred loan officers pay, on average, 16.5 bps higher mortgage rates; this premium remains sizable at 18.6 bps after controlling for FICO. At the sample's average loan size ($306K), this implies roughly $569 in additional interest payments per year. Equivalently, a referred borrower would need to pay about $2,609 in upfront points to buy down the rate to the level of a non-referred borrower.

The magnitude of the referral effect is economically large. Relative to the average APR spread of 51 bps, the referral premium represents a 36.5% increase in borrowing costs. Moreover, the residual standard deviation of rate spreads after controlling for lender, market, and quarter fixed effects is 37 basis points. Our estimates imply that referral lending accounts for 50% of this residual variation. Put differently, *half* of the cross-sectional variation in mortgage prices stems from loan-officer market power mediated by realtor referral networks.

5 Heterogeneous Effects of Referrals

The preceding analysis establishes that referral networks raise borrowing costs on average. A natural follow-up question is whether these costs fall evenly across borrowers. Here, we explore heterogeneity in the referral premium along two dimensions: borrower risk profiles (Section 5.1) and socio-demographic characteristics (Section 5.2).

5.1 Heterogeneous Effects by Borrower Risks

Table 6 examines how the referral premium varies across homebuyers with different risk profiles. Column (1) focuses on borrowers with FICO scores below 660. The estimated referral premium for this group is 11.9 bps higher than the 17.1 bps benchmark for the rest of the sample. Although the incremental spread is not statistically significant — likely due to the limited size of the 1% credit-matched sample — it is economically meaningful, suggesting that lower-credit-quality borrowers may be particularly vulnerable to referral-driven costs.

Columns (2)–(3) examine heterogeneity along the debt-to-income (DTI) margin. For borrowers with DTI above 45% – about one fifth of homebuyers in our sample – the referral premium is 10.9 bps larger than the rest of the sample, an increment equivalent to roughly $1,529 in upfront points at the average loan size. Columns (4)–(5) turn to high–loan-to-value (LTV) borrowers. Those with LTV above 95% – more than one third of homebuyers – pay an additional 21 bps beyond the 10.1 bps benchmark, equivalent to about $2,946 in upfront fees.

These results reveal a clear distributional pattern: the financial burden of referral-driven steering falls disproportionately on the most vulnerable borrowers. High-DTI, high-LTV, and low-FICO households — those with thinner financial buffers and fewer outside options — systematically pay larger referral premia than better-positioned homebuyers. The same referral practice that modestly raises costs for average borrowers can meaningfully amplify financing strains for those at the margin of credit access.

5.2 Heterogeneous Effects by Homebuyer Demographics

Table 7 examines how the referral effect varies across homebuyers of different socio-economic backgrounds. Columns (1)–(2) group homebuyers by ethnicity and reveal stark heterogeneity. For Hispanic homebuyers, the referral premium is 49.1 bps, translating into roughly $6,888 in additional upfront costs – substantially larger than for non-Hispanic borrowers. This gap may reflect language barriers, differential access to information, and a greater tendency to rely on realtors as trusted intermediaries, especially when intermediaries share cultural or local ties (e.g., Edin et al., 2003; Matvos et al., 2023).

Columns (3)–(4) examine heterogeneity by race. The estimated referral premia are 17.2 bps ($2,413) for White borrowers, 24.8 bps ($3,479) for Black borrowers, and 21.5 bps ($2,736) for Asian borrowers. Referral-driven costs are thus systematically higher for racial and ethnic minority borrowers, suggesting that steering through referral networks may amplify existing disparities in mortgage pricing and, by extension, in the overall cost of homeownership.

Finally, columns (5)–(6) examine heterogeneity between high- and low-income borrowers, defined as those with income ratio above versus below the county median. The referral premium is smaller for high-income borrowers (about 14 bps) and larger for low-income borrowers (about 19 to 24 bps), indicating that referral-driven costs disproportionately burden lower-income households.

6 Mechanisms and Interpretations

So far, we have established that realtor–loan-officer referral networks raise borrowing costs for homebuyers. In this section, we present several pieces of evidence that shed light on the sources of this premium. We examine whether referrals affect borrowers' mortgage search behavior across lending institutions (Section 6.1) and whether referred loan officers exercise pricing power within lending institutions (Section 6.2). We also assess whether referral lending generates offsetting benefits through faster closing times (Section 6.3) and reduced denial risk (Section 6.4).

6.1 Reduced Search Across Lenders

One potential explanation for the referral premium is that referred borrowers search less intensively across lenders, which limits their options. We test this directly by examining borrowers' mortgage search activity. Measuring search is challenging because standard mortgage datasets record only the final choice of lender and loan officer. We address this limitation by exploiting credit inquiry information in the 1% random sample matched to credit-bureau records. When a borrower seriously shops for a mortgage, each lender she approaches must obtain authorization to pull her credit report, and these inquiries are recorded as mortgage-related credit inquiries. We use the number of such inquiries as a proxy for the intensity of mortgage search.

Table 8 reports IV estimates of the effect of referrals on mortgage-related credit inquiries. Column (1) examines the number of lenders approached within a 90-day window prior to closing. Borrowers who use a referral loan officer make, on average, 0.168 fewer inquiries, a 12% decline relative to the baseline mean of 1.45. Column (2) considers the extensive margin — whether the borrower shops with at least two lenders. Referral use reduces the probability of contacting multiple lenders by 8 percentage points, a 22% decline from the baseline rate of 35 percent. Columns (3)–(4) repeat the analysis using a 60-day window. The referral effects are even stronger in relative terms: inquiries fall by 0.13, and the share of homebuyers with at least two inquiries declines by 38%. These results provide direct evidence that referrals compress borrowers' effective choice sets. Referred homebuyers contact fewer lenders and are substantially less likely to obtain competing quotes, weakening their outside options and shifting bargaining power toward lenders during rate negotiation.[11]

To gauge the quantitative importance of reduced search, we provide a back-of-the-envelope calculation. Bhutta et al. (2025) estimate that the expected gain from one additional mortgage search is approximately 18 basis points. Applying this estimate to the observed reduction of 0.168 inquiries implies an interest-rate premium of roughly 3 basis points. This likely understates the full effect for two reasons. First, credit inquiries capture only late-stage, committed shopping and miss earlier information-gathering that may also be curtailed by referrals. Second, the marginal return to search is plausibly larger for borrowers facing unfavorable initial offers from referred loan officers.

The evidence above establishes that referrals limit search across lenders. We next ask whether the resulting loss of outside options also enables referred loan officers to exercise pricing power within their own lending institutions.

6.2 Pricing Power within Lenders

To examine loan officers' pricing power within their lending institutions, we re-estimate the referral regressions with lender fixed effects, comparing referred and non-referred borrowers who obtain

[11] We also examine whether the demographic groups experiencing the largest referral premia display the sharpest reductions in search intensity (Table A3 and Appendix Section D.1). They do. Hispanic-referred borrowers make 0.254 fewer inquiries within 60 days of closing – more than twice the reduction for non-Hispanic-referred borrowers – and are 12.1 percentage points less likely to contact at least two lenders. Similar patterns hold for low-income borrowers.

mortgages from the same lending institution.

Table 9 reports the results using the same IV strategy. Relative to Column (1) of Table 5, adding lender fixed effects in Column (1) reduces the estimated referral premium from 16.5 bps to 4.2 bps. Column (2) controls for lenders more flexibly by interacting lender fixed effects with county $\times$ year-month fixed effects (an extremely demanding specification), thus comparing referred and non-referred borrowers who obtain loans from the same lender in the same local market and month. The estimated within-lender premium remains very similar at 3.7 bps. Column (3) restricts attention to the 1% matched sample and further controls for borrower credit scores. The estimated within-lender premium is 7.5 bps, about 40% of the full premium of 18.6 bps without lender fixed effects.[12]

These estimates imply that approximately three-fifths of the referral effect operates through a between-lender channel — referrals steer borrowers toward higher-cost lenders — while the remaining two-fifths reflects differential pricing within lenders. Even holding the lender constant, referred borrowers pay higher rates, indicating that referrals insulate individual loan officers from competitive pressure within their own institutions. Referral networks thus soften competition at each stage of the borrowing process: across lenders by compressing borrowers' choice sets, and within lenders by shielding referred officers from internal competition.

6.3 Closing Speed

Mortgage transactions are not only financially costly but also time-consuming, and borrowers may value speed alongside low rates. We next examine whether referral networks generate offsetting efficiency gains by reducing coordination frictions in the contracting stage, leading to faster processing.

Table 10 reports IV estimates of the effect of referrals on time to close, measured as the number of days between contract acceptance (i.e., the contract date) and mortgage closing. Column (1) presents full sample results: borrowers who use referred loan officers complete their purchase about 1.75 days sooner. Column (2) uses the 1% matched sample and adds credit score controls. The point estimate of 1.5 days implies a 4% reduction relative to the baseline mean of 40 days, though the effect is statistically insignificant. Columns (3)–(6) examine whether referrals reduce the probability of exceeding specific closing-time thresholds. In our sample, 72% of transactions take more than 30 days to close, and 9% exceed 60 days. Referrals are associated with a 5–6 percentage point lower probability of exceeding 30 days and a 1 percentage point lower probability of exceeding 60 days.

Overall, the time savings from referrals are modest. The established working relationships be-

[12] In this subsample, we cannot include lender $\times$ year-month $\times$ county fixed effects because the interacted cells become too sparse. The stability of the estimate relative to Columns (1)–(2) suggests that the within-lender referral premium is not driven by differences in borrower credit quality.

tween realtors and referred loan officers may facilitate smoother communication and faster document processing, but these convenience gains are small relative to the higher borrowing costs documented earlier.

6.4 Mortgage Denial Risk

Another potential benefit of referral networks is greater certainty of mortgage approval. If referred loan officers help marginal borrowers secure financing they might otherwise be denied, higher rates could reflect compensation for reduced denial risk. Because we do not observe denied applications, we test this indirectly by focusing on borrowers who face minimal denial risk. If even these borrowers pay more when using referred loan officers, approval risk mitigation is unlikely to be the primary explanation.

Table 11 reports the referral effect for high-credit-quality borrower subsamples, using the 1% random sample matched to credit records. Columns (1)–(3) report IV estimates for borrowers with FICO scores above 700, 740, and 780, respectively. As the FICO threshold rises, the sample shrinks substantially, but the estimated referral premium remains economically large and statistically significant. Even for borrowers with FICO scores above 780 (and minimal denial risk), the referral premium is sizeable at 9.3 bps (roughly $1,305 in equivalent upfront costs). Column (4) further tightens the sample by restricting to borrowers with FICO above 700, DTI below 36%, and LTV below 90%. Even in this very low-risk group, the referral premium remains statistically significant at 16 bps.

The persistence of the referral premium among the most creditworthy borrowers rules out approval certainty as the primary driver of higher borrowing costs. Rather than reflecting risk-based pricing, the premium appears instead to stem from reduced competition — an effect that persists even among borrowers with the lowest likelihood of denial.

7 Conclusion

This paper shows that the U.S. mortgage market is shaped not only by competition among lenders, but also by a quieter layer of intermediation: realtor–loan-officer referral networks. We document that these networks are pervasive, highly concentrated, and strongly predictive of which loan officers homebuyers ultimately use. Referral lending raises borrowing costs, even in markets that appear competitive on the lender side. It raises rate spreads by 36.5% and explains half of the (residual) standard deviation of rate spreads after controlling for lender, market, and time fixed effects. The referral premium is significantly larger for Hispanic borrowers, racial and ethnic minorities, and riskier borrowers with high DTI or high LTV — precisely the groups least able to absorb extra costs.

In terms of the underlying mechanism, referrals depress borrowers' search intensity across lenders and shield referred officers from internal competition within a lending institution. While referrals modestly accelerate closing times, these convenience gains are small relative to the additional financial cost, and the premium persists even among the most creditworthy borrowers, ruling out denial risk mitigation as the primary explanation.

From a policy standpoint, our results suggest that focusing solely on lender concentration is inadequate. A more effective approach may require enhancing transparency around referral relationships, making it easier for borrowers to obtain and compare alternative offers, and scrutinizing the incentive schemes that tie realtors and loan officers together. Looking ahead, important open questions include how emerging technologies — online platforms, digital brokerages, and automated pre-approval tools — reshape traditional referral networks, and whether these networks strengthen or weaken over the housing cycle.

References

Agarwal, Sumit, Swee Hoon Ang, Yongheng Deng, and Yonglin Wang, 2021, Mortgage Brokers and the Effectiveness of Regulatory Oversights, *Management Science* 67, 5278–5300.

Agarwal, Sumit, John Grigsby, Ali Hortaçsu, Gregor Matvos, Amit Seru, and Vincent Yao, 2024, Searching for Approval, *Econometrica* 92, 1195–1231.

Alexandrov, Alexei, and Sergei Koulayev, 2018, No shopping in the u.s. mortgage market: Direct and strategic effects of providing information Consumer Financial Protection Bureau Office of Research Working Paper No. 2017-01.

Allen, Jason, Robert Clark, and Jean-François Houde, 2014, The Effect of Mergers in Search Markets: Evidence from the Canadian Mortgage Industry, *American Economic Review* 104, 3365–3396.

Allen, Jason, Robert Clark, and Jean-François Houde, 2019, Search Frictions and Market Power in Negotiated-Price Markets, *Journal of Political Economy* 127, 1550–1598.

Ambrose, Brent W., and James N. Conklin, 2014, Mortgage Brokers, Origination Fees, Price Transparency and Competition: Mortgage Brokers, Origination Fees, Price Transparency and Competition, *Real Estate Economics* 42, 363–421.

Amel, Dean, Elliot Anenberg, and Rebecca A Jorgensen, 2018, On the Geographic Scope of Retail Mortgage Markets, *FEDS Notes Washington: Board of Governors of the Federal Reserve System, June 15, 2018* .

Barwick, Panle, Parag Pathak, and Maisy Wong, 2017, Conflicts of interest and steering in residential brokerage, *American Economic Journal: Applied Economics* 9, 191–222.

Barwick, Panle Jia, Yanyan Liu, Eleonora Patacchini, and Qi Wu, 2023, Information, Mobile Communication, and Referral Effects, *American Economic Review* 113, 1170–1207.

Berndt, Antje, Burton Hollifield, and Patrik Sandås, 2010, The Role of Mortgage Brokers in the Subprime Crisis, Technical Report w16175, National Bureau of Economic Research, Cambridge, MA.

Berry, Steve, James Levinsohn, and Ariel Pakes, 1995, Automobile prices in market equilibrium, *Econometrica* 63(4), 841–890.

Bhutta, Neil, Andreas Fuster, and Aurel Hizmo, 2025, Paying too much? borrower sophistication and overpayment in the u.s. mortgage market, *Journal of Finance* .

Bhutta, Neil, and Aurel Hizmo, 2021, Do Minorities Pay More for Mortgages?, *The Review of Financial Studies* 34, 763–789.

Bresnahan, Timothy F, and Peter C Reiss, 1991, Entry and competition in concentrated markets, *Journal of political economy* 99, 977–1009.

Buchak, Greg, and Adam Jorring, 2021, Competition with Multi-Dimensional Pricing: Evidence from U.S. Mortgages, *SSRN* .

Burks, Stephen V, Bo Cowgill, Mitchell Hoffman, and Michael Housman, 2015, The value of hiring through employee referrals, *The Quarterly Journal of Economics* 130, 805–839.

Card, David, and Laura Giuliano, 2016, Universal screening increases the representation of low-income and minority students in gifted education, *Proceedings of the National Academy of Sciences* 113, 13678–13683.

Cestau, Dario, Dennis Epple, and Holger Sieg, 2017, Admitting Students to Selective Education Programs: Merit, Profiling, and Affirmative Action, *Journal of Political Economy* 125.

Chen-Zion, Ayal, and James E. Rauch, 2020, History dependence, cohort attachment, and job referrals in networks of close relationships, *Journal of Economic Behavior & Organization* 170, 75–95.

Dranove, David, David Besanko, Mark Shanley, and Scott Schaefer, 2017, *Economics of Strategy* (Wiley).

Dunne, Timothy, Mark J Roberts, and Larry Samuelson, 1988, Patterns of firm entry and exit in us manufacturing industries, *The RAND journal of Economics* 495–515.

Edin, Per-Anders, Peter Fredriksson, and Olof Åslund, 2003, Ethnic enclaves and the economic success of immigrants—evidence from a natural experiment, *The Quarterly Journal of Economics* 118, 329–357.

Ernst, Keith, Debbie Bocian, and Wei Li, 2008, Steered wrong: Brokers, borrowers, and subprime loans, *Center for Responsible Lending* 8.

FreddieMac, 2016, Real estate professionals influence clients' mortgage lender choice: Freddie mac survey Media Room.

Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery, 2019, The role of technology in mortgage lending, *Review of Financial Studies* 32, 1854–1899.

Gurun, Umit G., Gregor Matvos, and Amit Seru, 2016, Advertising Expensive Mortgages, *The Journal of Finance* 71, 2371–2416.

Han, Lu, and Seung-Hyun Hong, 2011, Testing cost inefficiency under free entry in the real estate brokerage industry, *Journal of Business & Economic Statistics* 29, 564–578.

Han, Lu, and Seung-Hyun Hong, 2016, Understanding in-house transactions in the real estate brokerage industry, *RAND Journal of Economics* 47, 1057–1086.

Han, Lu, and Strange William, 2014, The microstructure of housing markets: Search, bargaining, and brokerage, *Handbook of Regional And Urban Economics (Duranton, Henderson, and Strange (eds.))* 5.

Ho, Kate, and Ariel Pakes, 2014, Hospital Choices, Hospital Prices, and Financial Incentives to Physicians, *American Economic Review* 104, 3841–3884.

Jorgensen, Rebecca A, 2024, The Economic Consequences of Mergers Between Real Estate Agencies and Mortgage Lenders .

Kermani, Amir, and Francis Wong, 2026, Racial disparities in housing returns, *American Economic Review* 116, 287–331.

LaCour-Little, Michael, 2009, The Pricing of Mortgages by Brokers: An Agency Problem?, *Journal of Real Estate Research* 31, 235–264.

Matvos, Gregor, Amit Seru, and Lulu Wang, 2023, Minority specialized lenders, *Northwestern University Kellogg working paper* .

O'Malley, A. James, Thomas Bubolz, and Jonathan Skinner, 2021, The Diffusion of Health Care Fraud: A Network Analysis, Technical Report w28560, National Bureau of Economic Research, Cambridge, MA.

Pallais, Amanda, and Emily Glassberg Sands, 2016, Why the Referential Treatment? Evidence from Field Experiments on Referrals, *Journal of Political Economy* 124.

Robles-Garcia, Claudia, 2020, Competition and Incentives in Mortgage Markets: The Role of Brokers 68.

Sarsons, Heather, 2024, Interpreting Signals in the Labor Market: Evidence from Medical Referrals .

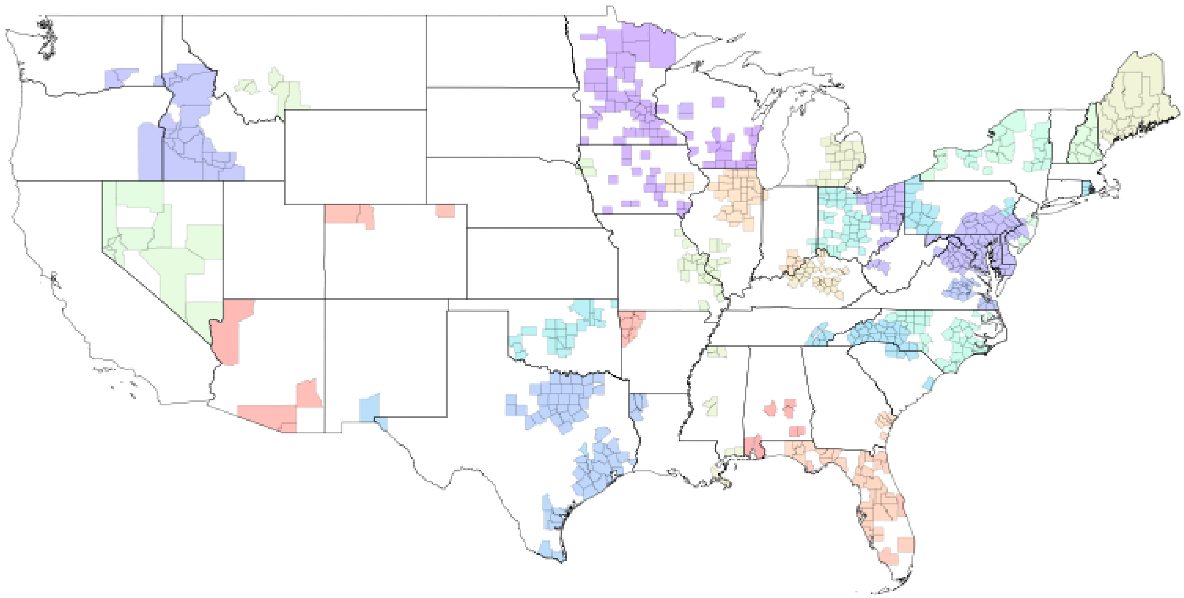
Spader, Jonathan S., and Roberto G. Quercia, 2011, Mortgage Brokers and the Refinancing Transaction: Evidence from CRA Borrowers, *The Journal of Real Estate Finance and Economics* 42, 181–210.

Woodward, Susan E, and Robert E Hall, 2010, Consumer Confusion in the Mortgage Market: Evidence of Less than a Perfectly Transparent and Competitive Market, *American Economic Review: Papers & Proceedings* 100, 511–515.

Woodward, Susan E, and Robert E Hall, 2012, Diagnosing Consumer Confusion and Sub-Optimal Shopping Effort: Theory and Mortgage-Market Evidence, *American Economic Review* 102, 3249–3276.

Figures

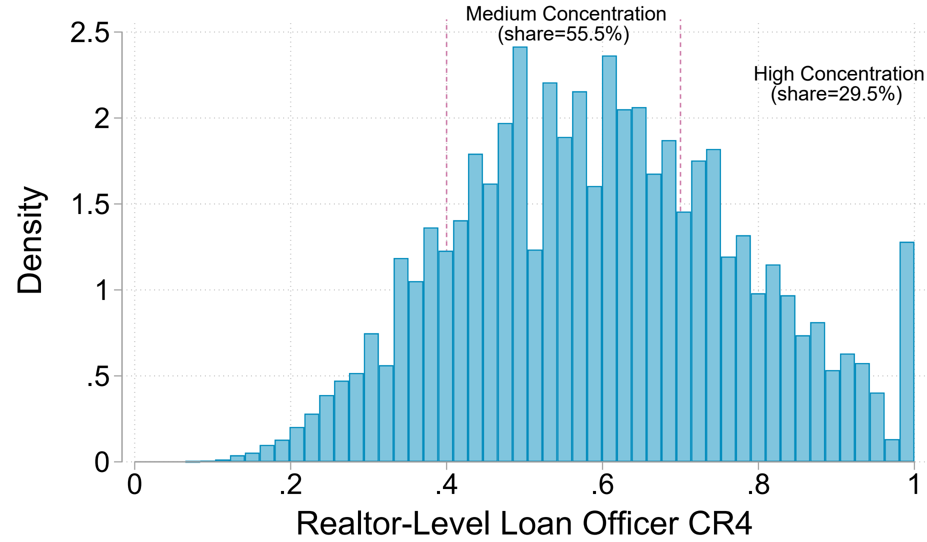
Figure 1: Geographic Coverage of the Data Sample



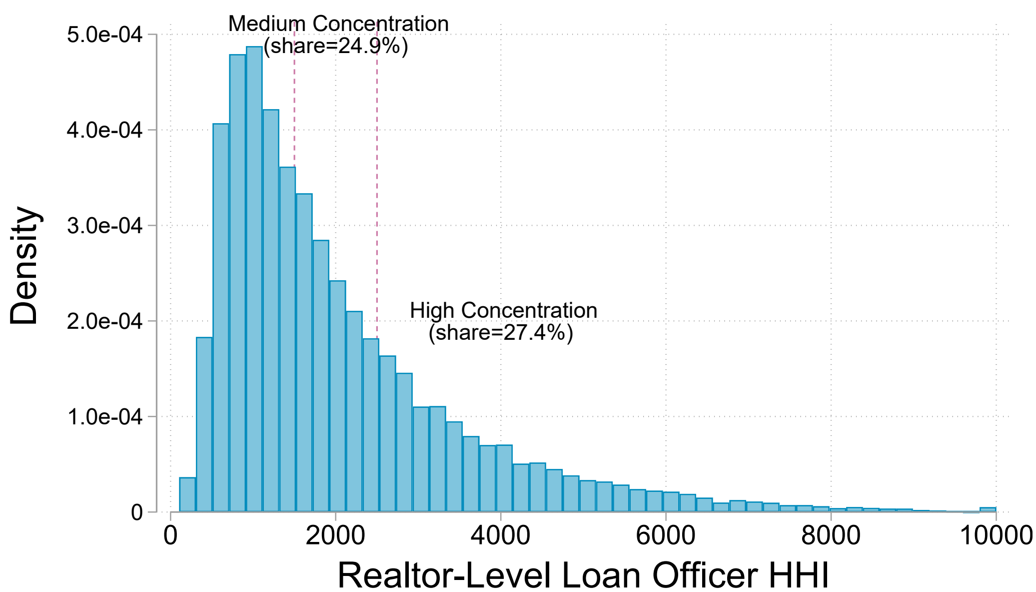
Note: This figure illustrates the counties covered in our data sample, with different colors representing the dominant MLS in each county. In total, our sample includes 728 counties across 41 states. These areas collectively represent 40% of the U.S. population.

Figure 2: Distribution of Realtors' Loan Officer Concentration

A. *CR*4



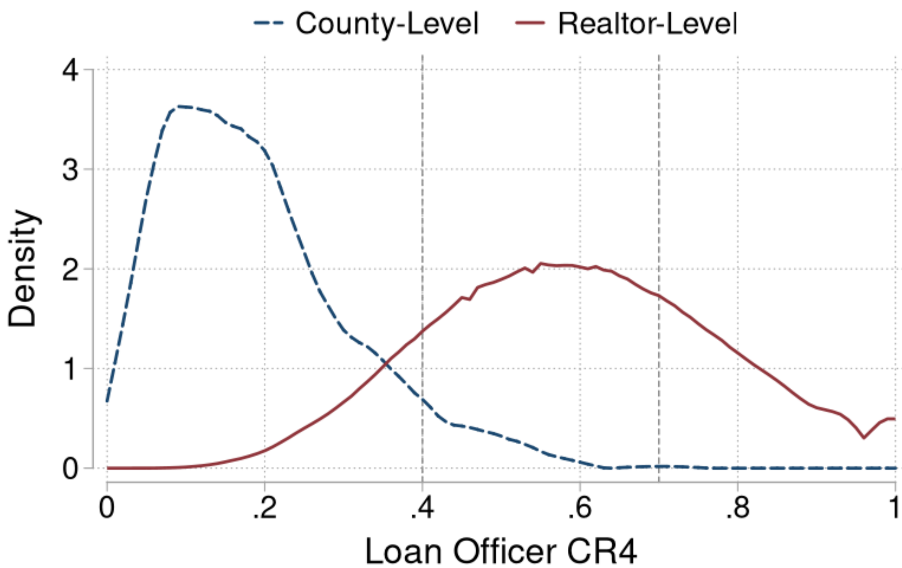
B. *HHI*



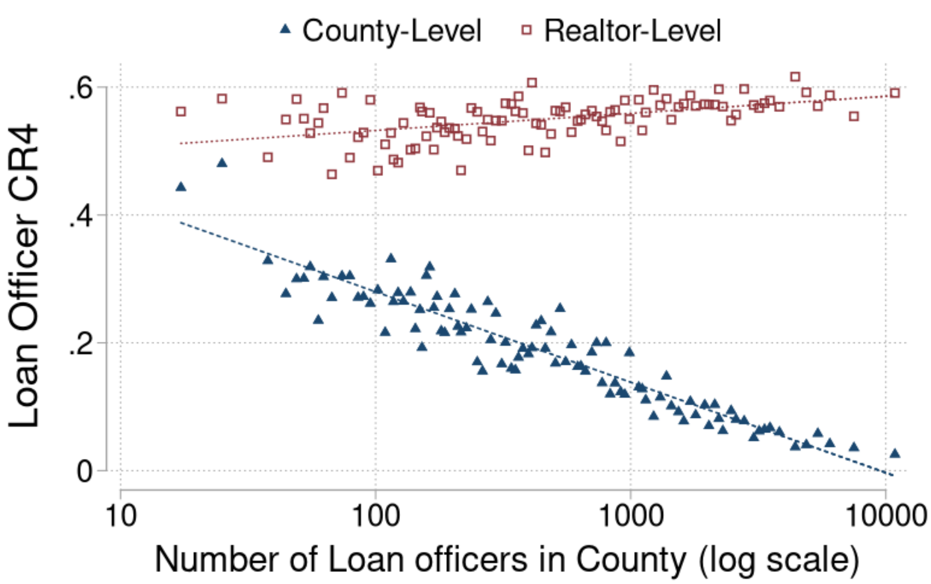
Note: This figure plots the distributions of realtors' loan officer concentration measures (*CR*4 in panel A, *HHI* in panel B). For each realtor, we define his/her loan officer network as the set of all loan officers who financed loans for at least one of the realtor's buyer clients. We then calculate each loan officer's market share – the fraction of that realtor's buyer clients whose mortgages the officer underwrites. The *CR*4 is simply the sum of the top four loan officers' shares for a given realtor, while the *HHI* is the sum of squared shares across the realtor's entire loan officer network.

Figure 3: County-Level and Realtor-Level Loan Officer Concentration

A. Distributions of County-Level and Realtor-Level Loan Officer Concentration

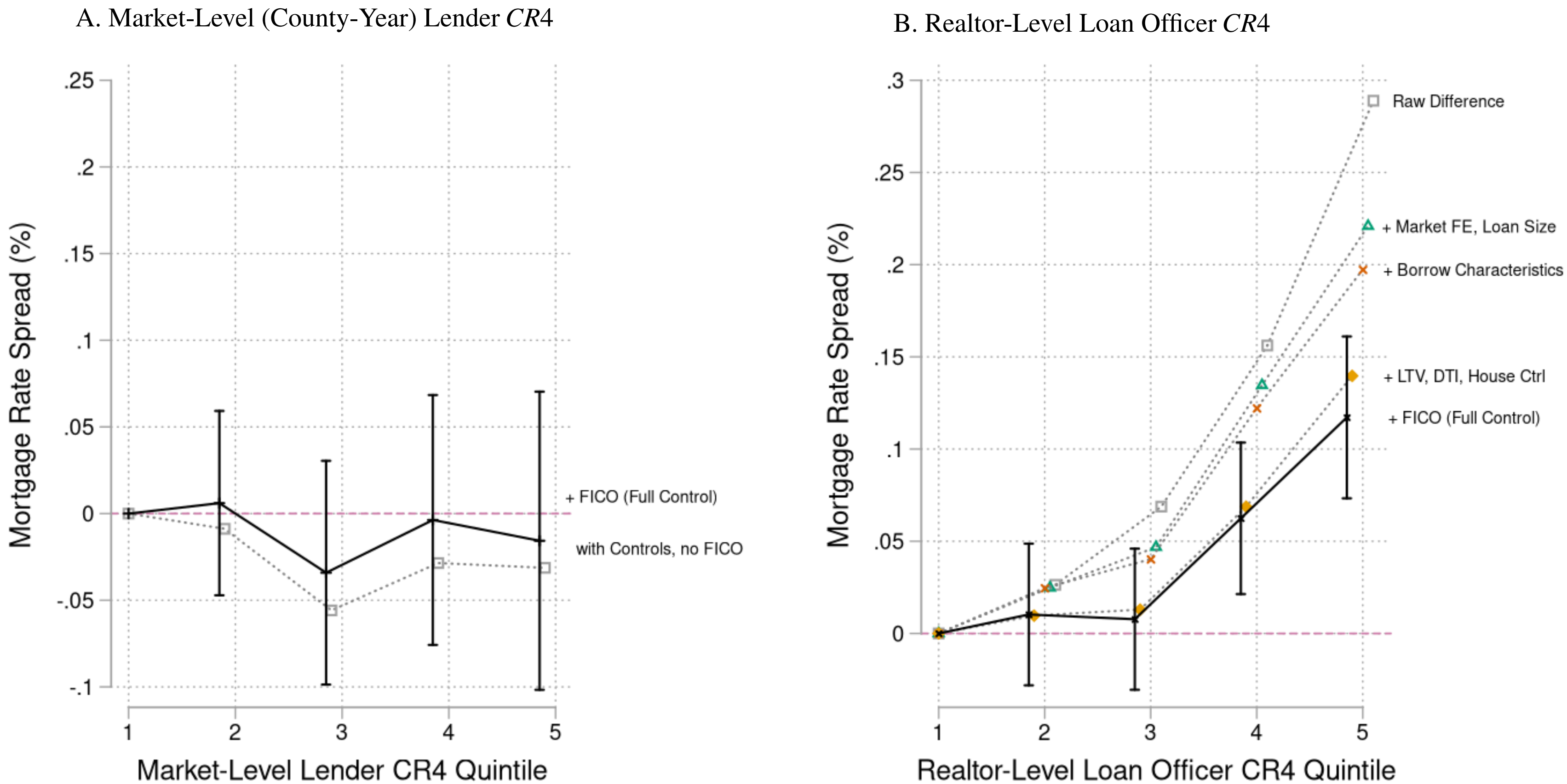


B. Loan Officer Concentration against Market Size



Note: This figure compares loan officer concentration (*CR*4) measured at the county level (blue) and the realtor level (red). Panel A plots the distributions of the two measures (dashed blue for county-level CR4 and solid red for realtor-level *CR*4). Panel B plots both measures against county size (blue-triangle for county-level *CR*4 and red-diamonds for realtor-level *CR*4), with counties sorted by the number of loan officers and grouped into 100 bins; dashed lines show linear fits. Within-realtor concentration is computed for realtors and counties with at least 10 transactions.

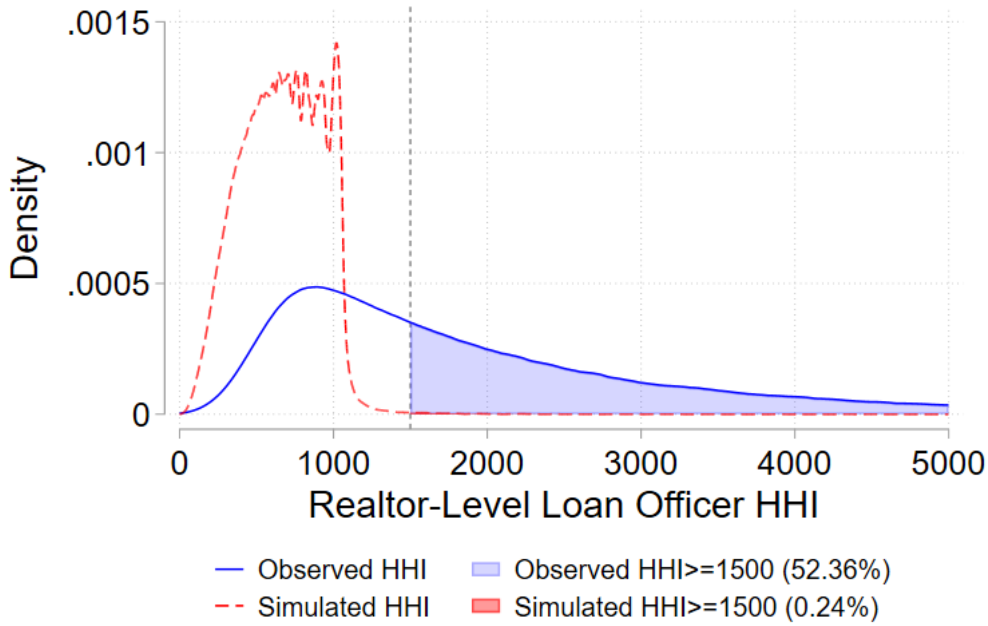
Figure 4: Mortgage Rate Spread By Lender/Loan Officer Concentration



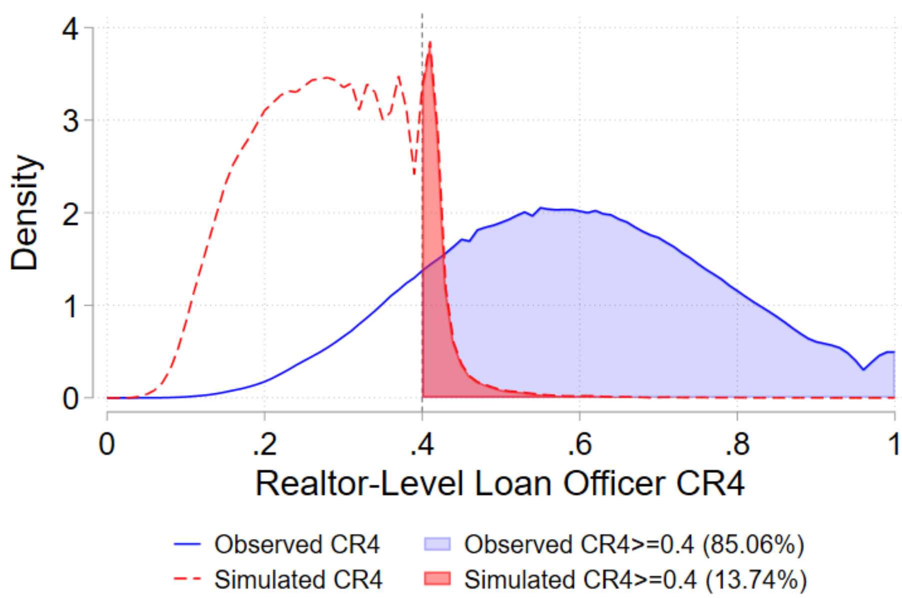
Note: This figure presents the relationship between mortgage APR spreads and lender concentration across different concentration quintiles. The 5th quintile corresponds to the highest concentration. The y-axis represents the mortgage APR spread, defined as the deviation from the benchmark rate offered on prime mortgage loans of a comparable type. In Panel A, concentration is measured by the market-wide lender *CR*4. The dashed line corresponds to the specification with all controls except FICO scores, while the solid black line corresponds to the fully controlled specification including FICO scores. In Panel B, concentration is measured by the within-realtor *CR*4 of loan officers.Gray squares plot the raw differences. Green triangles add county × year-month and loan amount bin fixed effects. Red crosses further control for borrower and loan characteristics. Orange diamonds add an even richer set of controls, including DTI and LTV bin fixed effects, house characteristics, and ZIP code fixed effects. Black crosses, connected by the solid black line, add FICO scores and represent the fully controlled specification.

Figure 5: Simulated vs Observed Distribution of the Realtor-Level Loan Officer Concentration

A. *HHI*



B. *CR*4



Note: This figure plots the distributions of loan-officer concentration at the realtor-level for both the observed and simulated sample. The simulated concentration is constructed via three steps: (i) building a refinance–mortgage sample (where referral lending is minimal), (ii) estimating a flexible BLP model of borrowers' lender choice, and (iii) running counterfactual simulations where we apply borrower preference parameters recovered in the refinance sample to the purchase-mortgage sample and predict borrowers' choice probabilities over all lenders in the market (absent referral lending). The spikes at HHI= 1,000 and $CR4 = 0.4$ in the counterfactual distribution reflect realtors at the minimum transaction threshold of 10. When the model predicts near-complete dispersion across loan officer choices, 10 borrowers each selecting a different loan officer yields $HHI = 1{,}000$ and $CR4 = 0.4$ mechanically.

Tables

Table 1: Homebuyer Attributes

	Full Sample (1)		1% Random Sample (Matched to Credit Records) (2)	
	mean	s.d.	mean	s.d.
Number of Observations (Homebuyers)	1,432,602		13,400	
APR Spread (%)	0.55	(0.68)	0.51	(0.66)
Loan Amount ($1K)	285.44	(155.04)	305.83	(163.56)
Joint Application	0.43	(0.50)	0.60	(0.49)
Conforming	0.97	(0.18)	0.96	(0.21)
LTV (%)	89.13	(11.50)	88.40	(11.74)
DTI (%)	35.15	(11.34)	34.65	(11.45)
FICO Score			742.74	(61.78)
Credit Inquires (90 Days Pre-Close)			1.45	(1.00)
Credit Inquires (60 Days Pre-Close)			1.03	(0.96)
County Median Income	73.83	(18.75)	75.53	(18.86)
Income Ratio	1.45	(1.04)	1.56	(1.12)
Hispanic	0.13	(0.34)	0.13	(0.34)
Race				
-White	0.73	(0.44)	0.74	(0.44)
-Black	0.09	(0.29)	0.08	(0.27)
-Asian	0.05	(0.22)	0.06	(0.23)
-Other Minority	0.13	(0.33)	0.13	(0.33)
Days to Close	40.13	(18.98)	39.73	(17.59)
Closing Price ($1K)	329.77	(201.61)	356.32	(213.29)
Square Feet	2102.12	(4957.47)	2163.18	(988.69)
Bedrooms	3.40	(0.84)	3.46	(0.84)
Bathrooms	2.52	(1.08)	2.60	(1.09)
House Age	39.73	(29.15)	39.19	(28.06)
Referral	0.56	(0.50)	0.57	(0.50)
Non-Compliant Referral	0.32	(0.47)	0.32	(0.47)

Note: This table presents the summary statistics of the home purchase data used in this study. Column (1) reports statistics for the full sample, while Column (2) summarizes the 1% random sample matched to credit records. “Income ratio” is defined as the reported income over the county median income of the same year. “Referral” equals one if the borrower’s realtor has a high-concentration network ($CR4 \geq 0.4$) and the loan is financed by one of that realtor’s top four loan officers. “Non-Compliant Referral” equals one if the borrower’s realtor has a high-concentration network ($CR4 \geq 0.4$) and the loan is not financed by one of that realtor’s top four loan officers. The FICO score and credit inquiry counts are only available in the 1% random sample.

Table 2: **Placebo Test** – Mortgage Rate Spreads By **Market-Level** Lender *CR*4 Quintiles

	Mortgage Rate Spread (off Prime Loans, in %)					
	No Control	+Market FE	+Borrower Ctrl	+DTI, LTV	+House Ctrl	+FICO
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Full Sample						
Market Lender *CR*4 Quintile 5	-0.072	-0.025*	-0.026**	-0.028**	-0.029**	
	(0.045)	(0.013)	(0.012)	(0.013)	(0.013)	
Market Lender *CR*4 Quintile 4	-0.079*	-0.031***	-0.033***	-0.029**	-0.030***	
	(0.045)	(0.011)	(0.010)	(0.012)	(0.011)	
Market Lender *CR*4 Quintile 3	-0.113**	-0.024**	-0.026**	-0.022**	-0.025**	
	(0.047)	(0.011)	(0.010)	(0.011)	(0.011)	
Market Lender *CR*4 Quintile 2	-0.104***	-0.009	-0.011	-0.009	-0.010	
	(0.039)	(0.007)	(0.007)	(0.007)	(0.007)	
Observations	1,437,243	1,437,243	1,437,243	1,437,243	1,437,243	
R-squared	0.004	0.164	0.187	0.293	0.320	
Dep. Var. Mean	0.55	0.55	0.55	0.55	0.55	
Panel B: 1% Random Sample Matched with Credit Records						
Market Lender *CR*4 Quintile 5	-0.090*	-0.004	0.002	-0.021	-0.031	-0.016
	(0.047)	(0.041)	(0.040)	(0.040)	(0.048)	(0.044)
Market Lender *CR*4 Quintile 4	-0.093**	-0.014	-0.011	-0.004	-0.029	-0.004
	(0.045)	(0.034)	(0.034)	(0.035)	(0.040)	(0.037)
Market Lender *CR*4 Quintile 3	-0.135***	-0.036	-0.032	-0.024	-0.056	-0.034
	(0.047)	(0.031)	(0.031)	(0.033)	(0.036)	(0.033)
Market Lender *CR*4 Quintile 2	-0.117***	-0.013	-0.012	-0.005	-0.009	0.006
	(0.040)	(0.025)	(0.025)	(0.025)	(0.028)	(0.027)
Observations	17,444	17,444	17,444	17,444	17,444	17,444
R-squared	0.005	0.197	0.228	0.338	0.470	0.545
Dep. Var. Mean	0.51	0.51	0.51	0.51	0.51	0.51
Year-Month FE		✓	✓	✓	✓	✓
County FE		✓	✓	✓	✓	✓
Loan Amount Bin FE		✓	✓	✓	✓	✓
Borrower Age Bin FE			✓	✓	✓	✓
Income Ratio Pencentile FE			✓	✓	✓	✓
Joint Application			✓	✓	✓	✓
Conforming FE			✓	✓	✓	✓
DTI Bin FE				✓	✓	✓
LTV Bin FE				✓	✓	✓
Bedroom FE					✓	✓
Bathroom FE					✓	✓
Sqft Bin FE					✓	✓
House Age FE					✓	✓
ZIP Code FE					✓	✓
FICO Bin FE						✓

Note: This table reports the differences of mortgage APR rate spreads across market-level (county-year) lender concentration (*CR*4) quintiles. The 5th quintile corresponds to the highest concentration. The dependent variable is the mortgage APR spread off the benchmark rate offered on prime mortgage loans of a comparable type. The income ratio is defined as the reported income over the county's median income of the same year. "Joint Application" takes the value 1 if there is more than one applicant on the mortgage loan. Standard errors are clustered at the county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Mortgage Rate Spreads By **Realtor-Level** Loan Officer *CR*4 Quintiles

	Mortgage Rate Spread (off Prime Loans, in %)					
	No Control	+Market FE	+Borrower Ctrl	+DTI, LTV	+House Ctrl	+FICO
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Full Sample						
Realtor *CR*4 Quintile 5	0.259***	0.210***	0.187***	0.146***	0.125***	
	(0.024)	(0.014)	(0.013)	(0.010)	(0.007)	
Realtor *CR*4 Quintile 4	0.149***	0.125***	0.113***	0.088***	0.077***	
	(0.016)	(0.009)	(0.008)	(0.006)	(0.005)	
Realtor *CR*4 Quintile 3	0.090***	0.078***	0.071***	0.056***	0.050***	
	(0.009)	(0.006)	(0.006)	(0.004)	(0.004)	
Realtor *CR*4 Quintile 2	0.051***	0.046***	0.042***	0.036***	0.033***	
	(0.006)	(0.005)	(0.004)	(0.004)	(0.003)	
Observations	1,432,602	1,432,602	1,432,602	1,432,602	1,432,602	
R-squared	0.016	0.189	0.210	0.311	0.337	
Dep. Var. Mean	0.55	0.55	0.55	0.55	0.55	
Panel B: 1% Random Sample Matched with Credit Records						
Realtor *CR*4 Quintile 5	0.289***	0.221***	0.197***	0.152***	0.140***	0.117***
	(0.033)	(0.029)	(0.027)	(0.025)	(0.025)	(0.022)
Realtor *CR*4 Quintile 4	0.156***	0.135***	0.122***	0.088***	0.069***	0.062***
	(0.025)	(0.019)	(0.018)	(0.018)	(0.022)	(0.021)
Realtor *CR*4 Quintile 3	0.069***	0.047***	0.040**	0.027	0.013	0.008
	(0.021)	(0.017)	(0.017)	(0.018)	(0.022)	(0.020)
Realtor *CR*4 Quintile 2	0.026	0.025	0.025	0.011	0.010	0.010
	(0.017)	(0.018)	(0.018)	(0.017)	(0.020)	(0.020)
Observations	13,400	13,400	13,400	13,400	13,400	13,400
R-squared	0.024	0.400	0.424	0.510	0.640	0.688
Dep. Var. Mean	0.51	0.51	0.51	0.51	0.51	0.51
Year-Month*County FE		✓	✓	✓	✓	✓
Loan Amount Bin FE		✓	✓	✓	✓	✓
Borrower Age Bin FE			✓	✓	✓	✓
Income Ratio Pencentile FE			✓	✓	✓	✓
Joint Application			✓	✓	✓	✓
Conforming FE			✓	✓	✓	✓
DTI Bin FE				✓	✓	✓
LTV Bin FE				✓	✓	✓
Bedroom FE					✓	✓
Bathroom FE					✓	✓
Sqft Bin FE					✓	✓
House Age FE					✓	✓
ZIP Code FE					✓	✓
FICO Bin FE						✓

Note: This table reports the differences of mortgage interest rate spreads across realtor-level loan officer concentration (*CR*4) quintiles. The dependent variable is the mortgage interest rate spread off the benchmark rate offered on prime mortgage loans of a comparable type. The income ratio is defined as the reported income over the county's median income of the same year. "Joint Application" takes the value 1 if there is more than one applicant on the mortgage loan. Standard errors are clustered at the county level and are reported in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table 4: Effect of Referral Lending on Purchase Mortgage Costs (OLS)

	Mortgage Rate Spread (off Prime Loans, in %)			
			1% Random Sample	
	(1)	(2)	(3)	(4)
Referral	0.099***	0.107***	0.088***	0.087***
($CR4$ $\geq$0.4, Top 4 Loan Officer)	(0.004)	(0.006)	(0.012)	(0.015)
Non-Compliant Referral		0.012***		-0.001
($CR4$ $\geq$0.4, Not Top 4 Loan Officer)		(0.003)		(0.020)
Observations	1,432,602	1,432,602	13,400	13,400
R-squared	0.338	0.338	0.688	0.688
Dep. Var. Mean	0.55	0.55	0.51	0.51
Year-Month*County FE	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓
Income Ratio Pencentile FE	✓	✓	✓	✓
Joint Application	✓	✓	✓	✓
Conforming FE	✓	✓	✓	✓
DTI Bin FE	✓	✓	✓	✓
LTV Bin FE	✓	✓	✓	✓
Bedroom FE	✓	✓	✓	✓
Bathroom FE	✓	✓	✓	✓
Sqft Bin FE	✓	✓	✓	✓
House Age FE	✓	✓	✓	✓
ZIP Code FE	✓	✓	✓	✓
FICO Bin FE			✓	✓

Note: This table reports the OLS regression estimates of the effect of referral lending on mortgage costs. Same set of controls (FE) as in Table 3. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. "Referral" equals one if the borrower's realtor has a high-concentration network ($CR4$ $\geq$0.4) and the loan is financed by one of that realtor's top four loan officers. "Non-Compliant Referral" equals one if the borrower's realtor has a high-concentration network ($CR4$ $\geq$0.4) but the loan is not financed by one of that realtor's top four loan officers. The income ratio is defined as the reported income over the county's median income of the same year. Standard errors are clustered at the county level and are reported in parentheses. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Table 5: Causal Effect of Referral Lending on Purchase Mortgage Costs (IV: HHI-$\widehat{HHI}$)

	Referral (First Stage)		Mortgage Rate Spread (off Prime Loans, in %)	
		1% Random Sample		1% Random Sample
			2SLS	2SLS
	(1)	(2)	(3)	(4)
(HHI-$\widehat{HHI}$)/10000	1.228***	1.152***		
	(0.016)	(0.035)		
Referral			0.165***	0.186***
($CR4$ ≥0.4, Top 4 Loan Officer)			(0.011)	(0.036)
Observations	1,432,602	13,400	1,432,602	13,400
R-squared	0.190	0.542		
Dep. Var. Mean	2.95	2.68	0.55	0.51
Year-Month*County FE	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓
Income Ratio Pencentile FE	✓	✓	✓	✓
Joint Application	✓	✓	✓	✓
Conforming FE	✓	✓	✓	✓
DTI Bin FE	✓	✓	✓	✓
LTV Bin FE	✓	✓	✓	✓
Bedroom FE	✓	✓	✓	✓
Bathroom FE	✓	✓	✓	✓
Sqft Bin FE	✓	✓	✓	✓
House Age FE	✓	✓	✓	✓
ZIP Code FE	✓	✓	✓	✓
FICO Bin FE		✓		✓
FS: Cragg-Donald Wald F			218970	1566
FS: Kleibergen-Paap rk F			6018	1082
FS: Anderson-RuBin p-val			0	0

Note: This table reports the IV estimates of the effect of referral lending on mortgage costs. Same set of controls (FE) as in Table 3. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. "Referral" equals one if the borrower's realtor has a high-concentration network ($CR4$ ≥0.4) and the loan is financed by one of that realtor's top four loan officers. It is instrumented with the "excess concentration" (HHI-$\widehat{HHI}$) of a realtor's loan officer network, which is the difference between observed HHI concentration and predicted HHI in the absence of referral lending. The income ratio is defined as the reported income over the county's median income of the same year. Standard errors are clustered at the county level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Effect of Referral Lending on Higher-Risk Borrowers (IV: HHI-$\widehat{HHI}$)

	Mortgage Rate Spread (off Prime Loans, in %)				
	1% Random Sample		1% Random Sample		1% Random Sample
	2SLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)
Referral	0.171***	0.147***	0.160***	0.095***	0.101**
	(0.038)	(0.010)	(0.035)	(0.009)	(0.040)
Referral*(FICO < 660)	0.119				
	(0.130)				
Referral*(DTI > 45%)		0.077***	0.109		
		(0.009)	(0.078)		
Referral*(LTV > 95%)				0.140***	0.210**
				(0.010)	(0.085)
Observations	13,400	1,432,602	13,400	1,432,602	13,400
Dep. Var. Mean	0.51	0.55	0.51	0.55	0.51
Year-Month*County FE	✓	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓	✓
Income Ratio Pencentile FE	✓	✓	✓	✓	✓
Joint Application	✓	✓	✓	✓	✓
Conforming FE	✓	✓	✓	✓	✓
DTI Bin FE	✓	✓	✓	✓	✓
LTV Bin FE	✓	✓	✓	✓	✓
Bedroom FE	✓	✓	✓	✓	✓
Bathroom FE	✓	✓	✓	✓	✓
Sqft Bin FE	✓	✓	✓	✓	✓
House Age FE	✓	✓	✓	✓	✓
ZIP Code FE	✓	✓	✓	✓	✓
FICO Bin FE	✓		✓		✓
FS: Cragg-Donald Wald F	775	109105	778.6	109495	759.2
FS: Kleibergen-Paap rk F	547.5	3080	500.6	3061	471.2
FS: Anderson-RuBin p-val	0	0	0	0	0

Note: This table reports the IV estimates of the effect of referral lending on mortgage costs for high-risk borrowers. Same set of controls (FE) as in Table 3. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. "Referral" equals one if the borrower's realtor has a high-concentration network ($CR4 \geq 0.4$) and the loan is financed by one of that realtor's top four loan officers. It is instrumented with the "excess concentration" (HHI-$\widehat{HHI}$) of a realtor's loan officer network, which is the difference between observed HHI concentration and predicted HHI in the absence of referral lending. Interactions between "Referral" and controls are instrumented with the corresponding interactions between excess concentration and the same controls. The income ratio is defined as the reported income over the county's median income of the same year. Standard errors are clustered at the county level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Effect of Referral Lending By Demographic Characteristics (IV: HHI-$\widehat{HHI}$)

	Mortgage Rate Spread (off Prime Loans, in %)					
	(By Ethnicity)		(By Race)		(By Income)	
		1% Random Sample		1% Random Sample		1% Random Sample
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Referral*Hispanic	0.265***	0.491***				
	(0.017)	(0.081)				
Referral*Non-Hispanic	0.127***	0.121***				
	(0.009)	(0.038)				
Referral*White			0.159***	0.176***		
			(0.011)	(0.041)		
Referral*Black			0.248***	0.248*		
			(0.018)	(0.127)		
Referral*Asian			0.117***	0.219		
			(0.017)	(0.139)		
Referral*Other Minority			0.176***	0.190*		
			(0.017)	(0.105)		
Referral*Income-High					0.138***	0.144***
					(0.012)	(0.049)
Referral*Income-Low					0.190***	0.236***
					(0.012)	(0.049)
Observations	1,432,602	13,400	1,432,602	13,400	1,432,602	13,400
Dep. Var. Mean	0.55	0.51	0.55	0.51	0.55	0.51
Year-Month*County FE	✓	✓	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓	✓	✓
Income Ratio Pencentile FE	✓	✓	✓	✓	✓	✓
Joint Application	✓	✓	✓	✓	✓	✓
Conforming FE	✓	✓	✓	✓	✓	✓
DTI Bin FE	✓	✓	✓	✓	✓	✓
LTV Bin FE	✓	✓	✓	✓	✓	✓
Bedroom FE	✓	✓	✓	✓	✓	✓
Bathroom FE	✓	✓	✓	✓	✓	✓
Sqft Bin FE	✓	✓	✓	✓	✓	✓
House Age FE	✓	✓	✓	✓	✓	✓
ZIP Code FE	✓	✓	✓	✓	✓	✓
FICO Bin FE		✓		✓		✓
FS: Cragg-Donald Wald F	109188	723.3	54475	370.9	109886	788.2
FS: Kleibergen-Paap rk F	3003	367.4	1403	120.8	2967	521.9
FS: Anderson-RuBin p-val	0	0	0	0	0	0

Note: This table reports the IV estimates of the effect of referral lending on mortgage costs across demographic groups. Same set of controls (FE) as in Table 3. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. "Referral" equals one if the borrower's realtor has a high-concentration network ($CR4 \geq 0.4$) and the loan is financed by one of that realtor's top four loan officers. It is instrumented with the "excess concentration" (HHI-$\widehat{HHI}$) of a realtor's loan officer network, which is the difference between observed HHI concentration and predicted HHI in the absence of referral lending. Interactions between "Referral" and demographics are instrumented with the corresponding interactions between excess concentration and the same controls. The income ratio is defined as the reported income over the county's median income of the same year. Standard errors are clustered at the county level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Effect of Referral on Mortgage Search Activities (IV: HHI-$\widehat{HHI}$)

	Within 90 Days		Within 60 Days	
	Credit Inquiries	>1 Inquiry (in %)	Credit Inquiries	>1 Inquiry (in %)
	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)
Referral	-0.168**	-7.875**	-0.131*	-8.410***
($CR4 \geq 0.4$, Top 4 Loan Officer)	(0.068)	(3.273)	(0.070)	(2.905)
Observations	13,400	13,400	13,400	13,400
Dep. Var. Mean	1.45	35.26	1.03	22.10
Year-Month*County FE	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓
Income Ratio Pencentile FE	✓	✓	✓	✓
Joint Application	✓	✓	✓	✓
Conforming FE	✓	✓	✓	✓
DTI Bin FE	✓	✓	✓	✓
LTV Bin FE	✓	✓	✓	✓
Bedroom FE	✓	✓	✓	✓
Bathroom FE	✓	✓	✓	✓
Sqft Bin FE	✓	✓	✓	✓
House Age FE	✓	✓	✓	✓
ZIP Code FE	✓	✓	✓	✓
FICO Bin FE	✓	✓	✓	✓
FS: Cragg-Donald Wald F	1566	1566	1566	1566
FS: Kleibergen-Paap rk F	1082	1082	1082	1082
FS: Anderson-RuBin p-val	0.013	0.015	0.062	0.004

Note: This table reports the IV estimates of the effect of referral on credit inquiries. Same set of controls (FE) as in Table 3. Credit inquiries obtained through credit-bureau records. "Referral" equals one if the borrower's realtor has a high-concentration network ($CR4 \geq 0.4$) and the loan is financed by one of that realtor's top four loan officers. It is instrumented with the "excess concentration" (HHI-$\widehat{HHI}$) of a realtor's loan officer network, which is the difference between observed HHI concentration and predicted HHI in the absence of referral lending. The income ratio is defined as the reported income over the county's median income of the same year. Standard errors are clustered at the county level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Effect of Referral Lending on Purchase Mortgage Costs within the Same Lender

	Mortgage Rate Spread (off Prime Loans, in %)		
			1% Random Sample
	2SLS	2SLS	2SLS
	(1)	(2)	(3)
Referral	0.042***	0.037***	0.075**
($CR4 \geq 0.4$, Top 4 Loan Officer)	(0.008)	(0.010)	(0.035)
Observations	1,432,385	1,129,983	13,051
Dep. Var. Mean	0.55	0.54	0.51
Year-Month*County FE	✓		✓
Lender FE	✓		✓
Year-Month*County*Lender FE		✓	
Loan Amount Bin FE	✓	✓	✓
Age Bin FE	✓	✓	✓
Income Ratio Pencentile FE	✓	✓	✓
Joint Application	✓	✓	✓
Conforming FE	✓	✓	✓
DTI Bin FE	✓	✓	✓
LTV Bin FE	✓	✓	✓
Bedroom FE	✓	✓	✓
Bathroom FE	✓	✓	✓
Sqft Bin FE	✓	✓	✓
House Age FE	✓	✓	✓
ZIP Code FE	✓	✓	✓
FICO Bin FE			✓
FS: Cragg-Donald Wald F	175885	136495	1139
FS: Kleibergen-Paap rk F	5208	3550	784.1
FS: Anderson-RuBin p-val	0	0	0.038

Note: This table reports the IV estimates of the effect of referral lending on mortgage costs within the same lender. Same set of controls (FE) as in Table 3, except that Columns (1) and (3) further control for lender FE and Column (2) further adds lender $\times$ year-month $\times$ county FE. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. "Referral" equals one if the borrower's realtor has a high-concentration network ($CR4 \geq 0.4$) and the loan is financed by one of that realtor's top four loan officers. It is instrumented with the "excess concentration" (HHI-$\widehat{HHI}$) of a realtor's loan officer network, which is the difference between observed HHI concentration and predicted HHI in the absence of referral lending. The income ratio is defined as the reported income over the county's median income of the same year. Standard errors are clustered at the county level and are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 10: Effect of Referral on Days to Close (IV: $HHI\text{-}\widehat{HHI}$)

	Days to Close		>30 Days (in %)		>60 Days (in %)	
		1% Random Sample		1% Random Sample		1% Random Sample
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Referral	-1.754***	-1.522	-6.080***	-5.080	-1.117***	-0.850
($CR4 \geq 0.4$, Top 4 Loan Officer)	(0.128)	(1.140)	(0.420)	(3.480)	(0.179)	(2.192)
Observations	1,424,534	13,308	1,424,534	13,308	1,424,534	13,308
Dep. Var. Mean	40.13	39.73	72.72	72.22	9.82	9.31
Year-Month*County FE	✓	✓	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓	✓	✓
Income Ratio Pencentile FE	✓	✓	✓	✓	✓	✓
Joint Application	✓	✓	✓	✓	✓	✓
Conforming FE	✓	✓	✓	✓	✓	✓
DTI Bin FE	✓	✓	✓	✓	✓	✓
LTV Bin FE	✓	✓	✓	✓	✓	✓
Bedroom FE	✓	✓	✓	✓	✓	✓
Bathroom FE	✓	✓	✓	✓	✓	✓
Sqft Bin FE	✓	✓	✓	✓	✓	✓
House Age FE	✓	✓	✓	✓	✓	✓
ZIP Code FE	✓	✓	✓	✓	✓	✓
FICO Bin FE		✓		✓		✓
FS: Cragg-Donald Wald F	217737	1543	217737	1543	217737	1543
FS: Kleibergen-Paap rk F	6007	1017	6007	1017	6007	1017
FS: Anderson-RuBin p-val	0	0.182	0	0.147	0	0.699

Note: This table reports the IV estimates of the effect of referral on home purchase closing time. Same set of controls (FE) as in Table 3. The dependent variables are the number of days between contract acceptance and closing, and indicators for whether closing exceeds 30 (60) days. "Referral" equals one if the borrower's realtor has a high-concentration network ($CR4 \geq 0.4$) and the loan is financed by one of that realtor's top four loan officers. It is instrumented with the "excess concentration" ($HHI\text{-}\widehat{HHI}$) of a realtor's loan officer network, which is the difference between observed HHI concentration and predicted HHI in the absence of referral lending. The income ratio is defined as the reported income over the county's median income of the same year. Standard errors are clustered at the county level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 11: Effect of Referral Lending for Borrowers with Minimal Denial Risks (IV: $HHI\text{-}\widehat{HHI}$)

	Mortgage Rate Spread (off Prime Loans, in %)			
	1% Random Sample			
	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)
Referral*(FICO > 700)	0.174***			
	(0.037)			
Referral*(FICO > 740)		0.155***		
		(0.043)		
Referral*(FICO > 780)			0.093**	
			(0.046)	
Referral*(FICO>700, DTI≤36%, LTV ≤90%)				0.160***
				(0.052)
Referral*Other	0.228**	0.227***	0.230***	0.192***
	(0.092)	(0.069)	(0.047)	(0.042)
Observations	13,400	13,400	13,400	13,400
Dep. Var. Mean	0.51	0.51	0.51	0.51
High-Quality Borrower Share	0.78	0.59	0.35	0.24
Year-Month*County FE	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓
Income Ratio Pencentile FE	✓	✓	✓	✓
Joint Application	✓	✓	✓	✓
Conforming FE	✓	✓	✓	✓
DTI Bin FE	✓	✓	✓	✓
LTV Bin FE	✓	✓	✓	✓
Bedroom FE	✓	✓	✓	✓
Bathroom FE	✓	✓	✓	✓
Sqft Bin FE	✓	✓	✓	✓
House Age FE	✓	✓	✓	✓
ZIP Code FE	✓	✓	✓	✓
FICO Bin FE	✓	✓	✓	✓
FS: Cragg-Donald Wald F	783.3	784.8	769.8	787.4
FS: Kleibergen-Paap rk F	539.8	549.1	524.6	585.8
FS: Anderson-RuBin p-val	0	0	0	0

Note: This table reports the IV estimates of the effect of referral lending on mortgage costs for low-risk borrowers. Same set of controls (FE) as in Table 3. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. "Referral" equals one if the borrower's realtor has a high-concentration network ($CR4 \geq 0.4$) and the loan is financed by one of that realtor's top four loan officers. It is instrumented with the "excess concentration" ($HHI\text{-}\widehat{HHI}$) of a realtor's loan officer network, which is the difference between observed HHI concentration and predicted HHI in the absence of referral lending. Interactions between "Referral" and controls are instrumented with the corresponding interactions between excess concentration and the same controls. The income ratio is defined as the reported income over the county's median income of the same year. Standard errors are clustered at the county level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Internet Appendix

for

Market Power in Mortgage Pricing: the Role of Referral Lending

Dayin Zhang   Panle Jia Barwick   Lu Han   Jon Kroah

A Data Construction

This appendix details the construction of the main purchase-mortgage sample used throughout the paper and the refinance sample used in the BLP estimation (Appendix B).

A.1 CoreLogic House, Listing and Mortgage Data

We rely on three CoreLogic data products, each covering the period July 2017 through December 2021. The first is the **Multiple Listing Service (MLS)** data, which records housing listings reported by local MLS platforms. Each listing contains the property address, listing and closing prices, listing and contract dates, and, crucially, standardized identifiers for the listing agent and the buyer agent. This data allows us to observe the realtor who represents each homebuyer. The second is the **Ownership Transfer** data (deed records), which records all property-level transactions filed with county recorders. The third is the **Mortgage** data, which records mortgage originations extracted from the same public deed sources. We join the three datasets through a universal property identifier (composite property linkage key).

County Selection As discussed in Section 2.2, our geographic sample is restricted to counties that satisfy three conditions: (i) a dominant MLS accounts for at least 80% of listings, (ii) the dominant MLS reports contract dates for at least 50% of listings, and (iii) the dominant MLS reports buyer-agent identifiers for at least 10% of listings. Counties with fewer than 10 listings in the dominant MLS are dropped. The resulting geographic sample is 703 counties across 41 states, mapped in Figure 1.

Identifying mortgage-financed home purchase events We construct the sample of arm's-length, mortgage-financed home purchases in four steps.

Step 1: Collapse listings into listing events. A single property may appear multiple times in the MLS if the seller relists after a price change, temporary withdrawal, or expired contract. To avoid double-counting sale attempts, we sort listings by property and listing date, treat any two records posted within 7 days as a single *listing attempt*, and group consecutive listing attempts for the same property into a single *listing event* whenever the gap between attempts is no more than 90 days and no attempt in between resulted in a close. We drop (i) listings whose list or close price is below 10% of the property's historical average, a price screen that removes suspected rental listings, and (ii) properties flagged as "split" — those for which the minimum list price within a listing attempt is below 90% of the maximum, indicating that units of a multi-unit property (e.g., duplex) were sold separately.

Step 2: Match each listing event to a deed record. For each listing event that concluded in a close (i.e., a non-missing close price), we link to the CoreLogic Ownership Transfer data using the composite property linkage key and sale date. We keep only transactions that are (i) flagged as arm's-length single-property purchases and (ii) mortgage-financed.

Step 3: Match each deed to a mortgage. We link each deed record to the CoreLogic Mortgage data using the composite property linkage key, searching for candidate mortgages within a $\pm$60-day window of the deed's sale date. Although the mortgage origination date should in principle match the sale date, we allow a 60-day discrepancy to account for possible recording errors. We also exclude observations in which the mortgage origination date precedes the original listing date.

Step 4: Validate IDs. Finally, we retain only transactions with a valid buyer-agent identifier in the MLS record. We exclude identifiers that appear to be placeholders, such as `UNKNOWNMEMBER`, `NONMLS`, or names containing `TEST AGENT`. Likewise, we drop transactions for which the originating loan officer lacks a Nationwide Multistate Licensing System (NMLS) ID, the unique federally regulated identifier assigned to individual mortgage loan originators.

This process yields roughly 5.1 million mortgage-financed purchases with both realtor and loan-officer identifiers. To guard against spurious concentration arising from low transaction counts, we further drop realtors associated with fewer than ten purchase loans over the sample period. This yields our **core purchase sample** of approximately 3.6 million transactions, which we use to construct each realtor's loan-officer network and concentration measures ($CR4_r$, HHI_r).

A.2 Merging CoreLogic with HMDA

Public HMDA microdata do not include a direct link to CoreLogic transactions. Following standard practice in the literature that combines these two sources (e.g., Kermani and Wong, 2026), we construct a probabilistic crosswalk using four overlapping fields that appear in both datasets: *lender name*, *loan amount*, *loan-to-value ratio (LTV)*, and *property census tract*, with both records required to share the same origination year.

For many mortgages, this produces a unique candidate pair, but in dense tracts or for common loan sizes the same combination of attributes can describe multiple originations. We classify a candidate pair as a *one-to-one match* if and only if (i) no other HMDA record shares the same attributes with the CoreLogic record, and (ii) no other CoreLogic record shares the same attributes with the HMDA record. In other words, we discard both one-to-many and many-to-one matches, retaining only pairs that are unambiguously uniquely identified by the four matching fields.

Of the 3.6 million mortgages in the core purchase sample, approximately 2.5 million (69%) are matched one-to-one to HMDA records.

Regression Sample Restrictions From the 2.5 million CoreLogic–HMDA matched mortgages, we apply the following additional restrictions to construct the main regression sample:

- **Loan product.** Retain only first-lien, 30-year, fixed-rate conventional or government-backed mortgages, the most common product in the U.S. single-family market.
- **Loan size.** Drop loans with an original principal balance below \$100,000, which are disproportionately non-traditional originations (e.g., manufactured housing, investor loans misclassified as owner-occupied).
- **Non-missing controls.** Drop observations with missing values for any of: LTV, DTI, applicant age, applicant income, and the joint-application indicator.
- **Realtor activity in the matched sample.** Drop realtors with fewer than ten matched purchase loans in the regression sample, so that each realtor's counterfactual $\widehat{HHI}_r$ and $\widehat{CR4}_r$ (used in the IV) can be computed from a sufficient number of simulated loan-officer choices.

The resulting main regression sample comprises 1.43 million transactions intermediated by 81,306 distinct realtors and 102,860 distinct loan officers.

A.3 Credit-Bureau Matched Subsample

We obtain a 1% random sample of borrowers from one of the three major U.S. credit bureaus. For each individual, we observe annual year-end credit snapshots with tradeline-level records (mortgages, student loans, credit cards, and installment credit), credit scores, and credit inquiries. We link each credit-bureau mortgage tradeline to a CoreLogic mortgage transaction in two stages.

In the first stage, we match on three anchors: (i) the loan amount, rounded to the nearest \$100 to tolerate recording noise; (ii) the event date, using either the mortgage origination or recording date on the CoreLogic side, with day offsets of up to seven days permitted in stricter-to-looser rounds; and (iii) the property ZIP code. We further require product consistency between the two sources (e.g., CoreLogic FHA and VA flags must align with the corresponding credit-bureau mortgage type codes) and enforce one-to-one uniqueness within each round, so that stricter early-round matches are never overwritten by later, more permissive ones. In the second stage, credit-bureau records still unmatched are rescued using a property anchor: for each borrower with a first-stage match, we learn the CoreLogic property linkage keys associated with that borrower, and then search CoreLogic mortgages on those same properties whose date falls within ±60 days of the credit-bureau open date and whose amount is within \$100 of the credit-bureau mortgage amount.

For the 2018–2021 vintages used in this paper, 79–84% of primary-type credit-bureau mortgages are successfully matched to CoreLogic, with the large majority linked in the first stage. Intersecting this crosswalk with the main purchase-regression sample yields our **credit-bureau matched subsample** of 14,688 mortgage observations. This subsample is used in analyses that require FICO controls (the baseline rate-spread regressions in the 1% sample) and in analyses of mortgage search intensity (credit-inquiry outcomes in Section 6.1).

A.4 Refinance Sample for BLP Estimation

The IV strategy (Section 4.3) requires estimating a lender-choice model on transactions where realtor referrals are largely absent. We construct a parallel *refinance* sample that mirrors the purchase sample wherever possible:

- **Geographic and temporal coverage.** Same 703 counties and same 2018–2021 window.
- **Loan-product filters.** First-lien, 30-year, fixed-rate, with a loan balance of at least \$100,000.

Since refinances are not brokered by real estate agents, we do not impose any realtor-side restrictions. FICO scores are a key input to the lender-choice model but are not reported in HMDA; we therefore intersect the refinance sample with the 1% random credit-bureau panel described in Section A.3, which provides directly observed borrower FICO for each transaction. The resulting estimation sample comprises 89,146 refinance transactions and is used to estimate Equation (5); summary statistics appear in Table A4 in Appendix B.

B BLP Estimation

B.1 Model

We assume consumer i's utility from choosing lender j in market m is given by

$$u_{ijm} = X'_{jm}\beta_1 + X'_{ijm}\beta_2 + \xi_{jm} + \varepsilon_{ijm}$$

where X_{jm} is a vector of lender-market characteristics; X_{ijm} is a vector of consumer-lender-market characteristics; ξ_{jm} is an unobserved vertical lender characteristic. A market, m, is defined as a county-year pair. We assume ε_{ijm} is an i.i.d. Type I Extreme Value shock, which yields the consumer choice probabilities:

$$s_{ijm}(\beta_1,\beta_2) = \frac{\exp(X'_{jm}\beta_1 + X'_{ijm}\beta_2 + \xi_{jm})}{\sum_{l=1}^{J_m}\exp(X'_{lt}\beta_1 + X'_{ilt}\beta_2 + \xi_{lt})} \tag{7}$$

Letting $\delta_{jm} \equiv X'_{jm}\beta_1 + \xi_{jm}$, we can rewrite the choice probabilities as

$$s_{ijm}(\delta_{.t},\beta_2) = \frac{\exp(\delta_{jm} + X'_{ijm}\beta_2)}{\sum_{l=1}^{J_m}\exp(\delta_{lt} + X'_{ilt}\beta_2)} \tag{8}$$

We observe individuals $i = 1,\ldots,N_m$ in each market. Given $\delta_{.t}$ and β_2, we approximate the model-predicted market shares of lender j in market m as

$$s_{jm}(\delta_{.t},\beta_2) \approx \frac{1}{N_m}\sum_{i=1}^{N_m} s_{ijm}(\delta_{.t},\beta_2) \tag{9}$$

B.2 Estimation Procedure

We estimate (β_1,β_2) in two steps. We first estimate β_2 via maximum likelihood with a nested fixed-point. At each guess of β_2, we invert out the δ terms using the contraction mapping developed by Berry et al. (1995):

$$\delta_{jm}^{(h+1)} = \delta_{jm}^{(h)} + \log\left(\frac{\tilde{s}_{jm}}{s_{jm}(\delta_{.t}^{(h)},\beta_2)}\right) \tag{10}$$

where $\tilde{s}_{jm}$ is the observed share of lender j in market m. We iterate until $|\delta_{jm}^{(h+1)} - \delta_{jm}^{(h)}| < 10^{-12}$ for all j,m. This yields the estimated mean utilities, $\hat{\delta}(\beta_2)$, that match model-predicted and observed market shares. We then search over β_2 until the log-likelihood is maximized, using the L-BFGS-B algorithm.

Given our MLE estimate $\hat{\beta}_2$, we estimate β_1 via OLS. Our estimating equation is:

$$\hat{\delta}_{jm} = X'_{jm}\beta_1 + \alpha_m + \xi_{jm} \tag{11}$$

where α_m is a market fixed effect; we include this because our model has no outside good, and thus δ_{jm} is only identified up to a constant in each market. β_1 is identified under the assumption that $E(\xi_{jm}|X_{jm},\alpha_m) = 0$.

B.3 Empirical Specification

We include the following in X_{jm}:

- Dummy for whether lender j is a bank
- Dummy for whether lender j is a fintech
- Dummy for whether lender j is out-of-state (i.e., has no branches in the same state as market m)
- Dummy for whether lender j is the #1-ranked lender in market m (ranked by transaction volume)
- Dummy for whether lender j is the #2-ranked lender in market m (ranked by transaction volume)
- Dummy for whether lender j is the #3-ranked lender in market m (ranked by transaction volume)

We include the following in X_{ijm}:

- Distance from borrower i's property address to lender j's nearest branch
  - We set this equal to 0 for fintechs and out-of-state lenders.
- Interactions between "bank" dummy and each of: FICO, loan-to-value ratio (LTV), debt-to-income ratio (DTI)
- Interactions between "fintech" dummy and each of: FICO, LTV, DTI
- Interactions between each of the #1, #2, #3 lender dummies and each of: FICO, LTV, DTI

B.4 Summary Statistics

We estimate the model on a 1% random sample of refinance mortgages, constructed by matching CoreLogic refinance transactions to the 1% random credit-bureau panel described in Appendix A.3. The match provides observed borrower FICO scores (not reported in HMDA), which are a key input to the lender-choice model. We drop markets with fewer than 10 transactions. When we estimate the model, we replace missing values of any variables with their unconditional sample means. Table A4 presents summary statistics for the estimation sample, before imputing missing values.

B.5 Estimates

Table A5 presents the estimates of (β_1, β_2).

C Simulation of the No–Referral Counterfactual

We construct a counterfactual "no–referral" world in which borrowers choose their loan officers independently of realtor influence. In total we do 200 simulations, and each iteration has two stages. In the first stage, we simulate which lender each homebuyer would select in a world without realtor referrals; in the second stage, we allocate those simulated loans across loan officers within each lender.

C.1 Simulating lender choice

First, we use our model estimates from Section B to compute model-predicted choice probabilities on our main purchase loan sample. For each transaction i in market m handled by realtor r, we compute

$$\hat{p}_{rji} = \frac{\exp(X'_{jm}\hat{\beta}_1 + X'_{ijm}\hat{\beta}_2 + \hat{\xi}_{jm})}{\sum_{l=1}^{J_m} \exp(X'_{lm}\hat{\beta}_1 + X'_{ilm}\hat{\beta}_2 + \hat{\xi}_{lm})}, \quad j = 1, \ldots, J_m$$

where $(\hat{\beta}_1, \hat{\beta}_2, \hat{\xi})$ are the MLE model estimates.

In this proces, if lender j has no refinance transactions in market m, then $\hat{\xi}\, jm$ is unavailable; we assign the unconditional mean of $\hat{\xi}\, jm$ to such lenders. HMDA does not report applicant credit scores, so for purchase-sample borrowers without a direct credit-bureau match we impute FICO as the average credit score across all individuals in a 1% random credit-bureau panel residing in the same ZIP code in the same year.

Treating these predicted probabilities as choice weights, we draw a simulated lender for each borrower. This gives us the alternative history of who each buyer would have borrowed from if lender choice were driven only by market structure, prices, borrower characteristics, and geography—but not by realtor steering.

C.2 Allocating loan officers within lenders

The next stage is to map these simulated lender choices into simulated realtor–loan-officer pairings. For each lender in each local market, we observe how its actual loan volume is distributed across individual loan officers. We interpret these observed shares as reflecting how borrowers would be allocated across officers within the same institution in the absence of steering. Conditional on each simulated lender choice, we therefore identify all loan officers of that lender operating in the borrower's market and then randomly assign a specific loan officer, with the assignment probabilities proportional to each officer's observed share of that lender's loans in the county–year. We repeat this assignment for every borrower and every simulation draw. The result is 200 simulated datasets of realtor–loan-officer relationships that reflect borrowers' predicted lender choices and the internal composition of lenders' officer pools.

C.3 Constructing the counterfactual concentration measure

For each simulation draw, we aggregate the simulated pairings at the realtor level and compute a concentration index (such as $CR4$ or HHI) for each realtor's loan-officer network. This gives us 200 simulated concentration measures per realtor. We then average across the draws to obtain a single counterfactual concentration measure for each realtor, denoted $\widehat{HHI}_r$ (or $\widehat{CR4}_r$), which answers the question: "If borrowers chose lenders and loan officers based solely on fundamentals captured in the refinance model, how concentrated would this realtor's loan-officer relationships be?" Comparing this counterfactual $\widehat{HHI}_r$ (or $\widehat{CR4}_r$) to the observed concentration for the same realtor reveals the "excess concentration" attributable to referrals. This excess concentration is the key input to our IV strategy, allowing us to isolate the causal effect of referral intensity on borrowers' mortgage costs.

D Additional Results and Robustness Checks

D.1 Heterogeneous Search Effects of Referral by Borrower Demographics

Section 6.1 of the main text shows that referred borrowers contact fewer lenders on average. If reduced search is the channel through which referrals raise borrowing costs, then the demographic groups that bear the largest referral premia in Section 5.2 — Hispanic borrowers, racial minorities, and lower-income households — should also display the sharpest reductions in search intensity. This appendix tests that prediction directly.

We re-estimate the baseline IV specification from Table 8 after replacing the single referral indicator with separate indicators for referred borrowers in each demographic group. Each demographic-specific referral indicator is instrumented with the interaction of the excess-concentration instrument ($HHI_r - \widehat{HHI}_r$) and the corresponding demographic indicator. All other controls and fixed effects are identical to Table 8. We examine three partitions: (i) Hispanic versus non-Hispanic borrowers; (ii) White, Black, Asian, and "Other Minority" borrowers by race; and (iii) borrowers grouped into low- and high-income homebuyers relative to the county median. Table A3 reports the results, with columns (1)–(3) measuring the number of mortgage-related credit inquiries within 60 days of closing and columns (4)–(6) the probability of contacting at least two lenders in the same window.

Ethnicity Hispanic referred borrowers make 0.254 fewer inquiries than comparable non-referred Hispanic borrowers (column 1), more than twice the reduction estimated for non-Hispanic referred borrowers. On the extensive margin (column 4), Hispanic referred borrowers are 12.1 percentage points less likely to contact at least two lenders, compared with 7.7 percentage points for non-Hispanic borrowers. These gaps mirror the Hispanic–non-Hispanic gap in the referral premium documented in Section 5.2, where Hispanic borrowers face a premium of 49.1 bps, nearly triple the baseline effect.

Race On the intensive margin, the race partition (column 2) yields point estimates that are uniformly negative but individually insignificant for referred White, Black, Asian, and "Other Minority" borrowers, reflecting the small number of minority borrowers in the 1% credit-matched sample. The extensive margin (column 5) is sharper: referred Asian borrowers are 17.8 percentage points less likely to contact at least two lenders, referred Black borrowers 11.2 percentage points less likely, and referred "Other Minority" borrowers 11.8 percentage points less likely — all substantially larger than the 6.3-percentage-point reduction for referred White borrowers.

Income The income partition (columns 3 and 6) shows that low-income referred borrowers make 0.167 fewer inquiries (significant at the 10% level), compared with a smaller and statistically insignificant reduction of 0.100 for high-income referred borrowers. On the extensive margin, both groups reduce the probability of shopping with multiple lenders — by 9.6 percentage points for low-income and 7.4 percentage points for high-income borrowers — with the low-income estimate larger in magnitude. This pattern echoes the income gradient in the referral premium documented in Section 5.2, where low-income borrowers pay 24 basis points in referral premia compared to 14 basis points for high-income borrowers.

Taken together, the demographic groups that bear the highest referral premia also reduce their lender search most sharply under referral, particularly on the extensive margin where effects for Hispanic, Black, and Asian referred borrowers are all statistically significant and substantially larger than those for White borrowers. The intensive-margin estimates for racial and ethnic minorities are underpowered given the limited number of minority borrowers in the matched subsample, but the point estimates are directionally consistent. These results support the interpretation that reduced search is the common mechanism driving both the average referral premium and its uneven incidence across borrowers.

E Additional Tables

Table A1: Summary Statistics for Realtors and Loan Officers

	Markets (County*Year) (1)	
	mean	s.d.
Observation	2,579	
No of Home Purchases	556	(1296)
No of Realtors	220	(432)
No of Loan Officers	232	(428)
No of Lenders	73	(95)

Note: This table reports summary statistics at the county-year market level for the 2,579 markets in the core purchase sample.

Table A2: Effect of Referral Lending on Purchase Mortgage Costs (IV: $CR4\text{-}\widehat{CR4}$)

	Referral (First Stage)		Mortgage Rate Spread (off Prime Loans, in %)	
		1% Random Sample		1% Random Sample
			2SLS	2SLS
	(1)	(2)	(3)	(4)
$CR4\text{-}\widehat{CR4}$	1.167***	1.123***		
	(0.007)	(0.026)		
Referral			0.193***	0.203***
($CR4 \geq 0.4$, Top 4 Loan Officer)			(0.012)	(0.031)
Observations	1,432,602	13,400	1,432,602	13,400
Dep. Var. Mean	10.12	9.61	0.55	0.51
Year-Month*County FE	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓
Income Ratio Pencentile FE	✓	✓	✓	✓
Joint Application	✓	✓	✓	✓
Conforming FE	✓	✓	✓	✓
DTI Bin FE	✓	✓	✓	✓
LTV Bin FE	✓	✓	✓	✓
Bedroom FE	✓	✓	✓	✓
Bathroom FE	✓	✓	✓	✓
Sqft Bin FE	✓	✓	✓	✓
House Age FE	✓	✓	✓	✓
FICO Bin FE		✓		✓
FS: Cragg-Donald Wald F			309373	2239
FS: Kleibergen-Paap rk F			27444	1936
FS: Anderson-RuBin p-val			0	0

Note: This table reports the IV estimates of the effect of referral lending on mortgage costs. Same set of controls (FE) as in Table 3. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. "Referral" equals one if the borrower's realtor has a high-concentration network ($CR4 \geq 0.4$) and the loan is financed by one of that realtor's top four loan officers. It is instrumented with the "excess concentration" ($CR4\text{-}\widehat{CR4}$) of a realtor's loan officer network, which is the difference between observed $CR4$ concentration and predicted $CR4$ in the absence of referral lending. The income ratio is defined as the reported income over the county's median income of the same year. Standard errors are clustered at the county level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Heterogeneous Effects of Referral on Mortgage Search by Borrower Demographics (IV: HHI–$\widehat{HHI}$)

	Credit Inquires (Within 60 Days)			>1 Inquiry (in %) (Within 60 Days)		
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Referral*Hispanic	-0.254*			-12.058**		
	(0.147)			(6.077)		
Referral*Non-Hispanic	-0.108			-7.748**		
	(0.073)			(3.241)		
Referral*White		-0.103			-6.274*	
		(0.069)			(3.210)	
Referral*Black		-0.130			-11.221*	
		(0.193)			(5.962)	
Referral*Asian		-0.258			-17.767*	
		(0.231)			(10.124)	
Referral*Other Minority		-0.217			-11.812	
		(0.246)			(10.169)	
Referral*Income-High			-0.100			-7.436**
			(0.096)			(3.571)
Referral*Income-Low			-0.167*			-9.562**
			(0.090)			(3.912)
Observations	13,400	13,400	13,400	13,400	13,400	13,400
R-squared	-0.004	-0.000	-0.003	-0.007	-0.004	-0.006
Dep. Var. Mean	1.03	1.03	1.03	22.10	22.10	22.10
Year-Month*County FE	✓	✓	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓	✓	✓
Income Ratio Pencentile FE	✓	✓	✓	✓	✓	✓
Joint Application	✓	✓	✓	✓	✓	✓
Conforming FE	✓	✓	✓	✓	✓	✓
DTI Bin FE	✓	✓	✓	✓	✓	✓
LTV Bin FE	✓	✓	✓	✓	✓	✓
Bedroom FE	✓	✓	✓	✓	✓	✓
Bathroom FE	✓	✓	✓	✓	✓	✓
Sqft Bin FE	✓	✓	✓	✓	✓	✓
House Age FE	✓	✓	✓	✓	✓	✓
ZIP Code FE	✓	✓	✓	✓	✓	✓
FICO Bin FE		✓		✓		✓
FS: Cragg-Donald Wald F	723.3	370.9	788.2	723.3	370.9	788.2
FS: Kleibergen-Paap rk F	367.4	120.8	521.9	367.4	120.8	521.9
FS: Anderson-RuBin p-val	0.110	0.452	0.125	0.008	0.032	0.014

Note: This table reports the IV estimates of the effect of referral lending on borrowers' mortgage search intensity across demographic groups, using the 1% random sample matched to credit-bureau records. Same set of controls (FE) as in Table 3. The dependent variables are the number of mortgage-related credit inquiries within 60 days of closing (columns 1–3) and an indicator for contacting at least two lenders in the same window (columns 4–6). "Referral" equals one if the borrower's realtor has a high-concentration network ($CR4 \geq 0.4$) and the loan is financed by one of that realtor's top four loan officers. It is instrumented with the "excess concentration" (HHI-$\widehat{HHI}$) of a realtor's loan officer network, which is the difference between observed HHI concentration and predicted HHI in the absence of referral lending. Interactions between "Referral" and demographics are instrumented with the corresponding interactions between excess concentration and the same controls. The income ratio is defined as the reported income over the county's median income of the same year. Standard errors are clustered at the county level and are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A4: Summary Statistics for Lender Choice Model Estimation Sample

Var	Obs	Mean	SD	Min	P50	Max
#1 Lender in Market	89,146	0.059	0.236	0.000	0.000	1.000
#2 Lender in Market	89,146	0.046	0.210	0.000	0.000	1.000
#3 Lender in Market	89,146	0.033	0.179	0.000	0.000	1.000
Fintech	89,146	0.181	0.385	0.000	0.000	1.000
Bank	88,516	0.395	0.489	0.000	0.000	1.000
Out-of-State	89,146	0.192	0.394	0.000	0.000	1.000
FICO	89,146	763.744	59.017	357.000	780.000	850.000
Combined LTV	80,703	68.739	16.675	0.810	71.000	191.030
DTI Ratio	77,029	29.660	12.259	10.000	30.000	60.000
Distance | Distance>0	86,372	25.458	52.820	0.001	7.489	668.887

Note: This table reports summary statistics for the refinance-mortgage sample used to estimate the lender choice model in Equation (5). This sample is constructed by matching CoreLogic refinance transactions to the 1% random credit-bureau panel described in Appendix A.3. “Bank”, “Fintech”, and “Out-of-State” are dummies for the chosen lender’s type, where “Out-of-State” indicates that the lender has no branches in the state. “#1 / #2 / #3 Lender in Market” are dummies for whether the chosen lender is ranked first, second, or third by transaction volume in the county-year market. “Distance” is the distance from the borrower’s property to the chosen lender’s nearest branch, set to zero for fintechs and out-of-state lenders; the reported statistics are conditional on positive distance.

Table A5: Lender Choice Model Estimates

Bank	-3.077 ***
	(0.142)
Bank x Combined LTV	-2.070 ***
	(0.056)
Bank x DTI	-0.314 ***
	(0.076)
Bank x FICO	5.259 ***
	(0.161)
Distance	-3.548 ***
	(0.097)
Fintech	-1.130 ***
	(0.171)
Fintech x Combined LTV	-1.253 ***
	(0.072)
Fintech x DTI	1.106 ***
	(0.093)
Fintech x FICO	1.032 ***
	(0.188)
Out-of-State	-1.377 ***
	(0.034)
Top Lender #1	-0.757 ***
	(0.278)
Top Lender #1 x Combined LTV	0.298 ***
	(0.108)
Top Lender #1 x DTI	-0.343 **
	(0.144)
Top Lender #1 x FICO	0.916 ***
	(0.318)
Top Lender #2	-1.100 ***
	(0.310)
Top Lender #2 x Combined LTV	-0.290 **
	(0.114)
Top Lender #2 x DTI	-0.030
	(0.153)
Top Lender #2 x FICO	1.488 ***
	(0.358)
Top Lender #3	-1.064 ***
	(0.353)
Top Lender #3 x Combined LTV	-0.196
	(0.138)
Top Lender #3 x DTI	0.251
	(0.185)
Top Lender #3 x FICO	1.247 ***
	(0.405)
Market FE (Second Stage)	Yes
Observations	82,067

Note: This table reports the estimates of the lender choice model (β_1, β_2) in Equation (5), estimated state-by-state on the refinance-mortgage sample described in Table A4. Standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.