# Referral Lending and Mortgage Market Power: The Role of Realtors

Dayin Zhang  Panle Jia Barwick  Lu Han  Jon Kroah*

February 2026

**Abstract**

Despite high competition in mortgage lending, significant price dispersion persists. We argue that realtor-loan-officer referral networks steer homebuyers toward a limited set of loan officers, restricting borrower choice even in competitive markets. Using a novel dataset linking 81,306 realtors to 102,860 loan officers in 41 states, we document substantial concentration within these networks: 85% of realtors direct over 40% of their clients to fewer than four loan officers. This concentration persists and even *increases* in markets with more lenders, suggesting that referrals constrain choice regardless of market structure. Borrowers working with high-concentration realtors pay significantly higher mortgage rates, even after controlling for borrower, loan, property, and market characteristics. IV estimates confirm that using a referred loan officer imposes a causal premium of 18.6 basis points, equivalent to $2,609 in upfront costs for the average loan size of $304k. The referral premium is especially pronounced for Black borrowers and financially constrained households, and nearly three times as high for Hispanic borrowers. We identify two mechanisms: referrals reduce borrowers' search intensity for lenders, and referred loan officers exercise pricing power even within the same lending institution. Efficiency arguments (faster processing) and mediation of denial risks cannot fully justify the referral premium. Our findings suggest that the referral networks reinforce mortgage market power, imposing significant financial burdens and raising equity concerns for borrowers.

**Keywords:** realtors, mortgage brokers, network, steering

**JEL** Classification Codes: L91, R13, R21

*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 the Singapore Management Univeristy, WFA, ASSA-AREUEA Conference for their helpful comments. All remaining errors are our own.

# 1 Introduction

In the U.S. mortgage market, extensive research has documented 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 concentrations. 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 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.

The crucial question is whether such concentration translates into higher borrowing costs. We

[1] We use "loan officers" to refer to both in-house loan officers of mortgage lenders, and independent mortgage brokers.

find that it does – substantially. Borrowers working with realtors with top-quintile concentration (by $CR4$) pay 28.9 basis points more than those with bottom-quintile realtors, with rate differentials decreasing 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, etc.), and ZIP code 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-driven 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. 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/load 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.

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 assess magnitude, consider that observably identical loans in our sample exhibit 86 basis points of rate dispersion at the 90th–10th percentiles, after controlling for market-specific trends, borrower-, mortgage- and house- characteristics, as well as ZIP code fixed effects. Our estimates suggest that referral premium alone accounts for 22% of this dispersion. In other words, more than one-fifth 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- and middle-income borrowers pay 19–27 basis points in referral premia compared to 11 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 interpret the baseline empirical patterns through a search-and-bargaining framework in which mortgage origination unfolds in two stages. In the first stage, borrowers and lenders search to identify feasible counterparties. In the second stage, conditional on a match, mortgage rates are determined through bilateral bargaining, with the division of surplus governed by each side's outside options. Realtor–loan-officer referrals operate primarily by compressing the borrower's choice set in the search stage, reducing the expected gains from additional search and leading borrowers to optimally contact fewer lenders. This weakening of borrower outside options shifts bargaining power toward lenders in the contracting stage, resulting in higher negotiated mortgage rates.

To test this mechanism, we use credit-bureau inquiries to measure mortgage search and 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 reflects an optimal response to weakened outside options, which shifts bargaining power toward lenders. The joint pattern - less search and higher prices - is difficult to reconcile with alternative explanations based on borrower sorting or lender quality, which would predict equal or greater comparison shopping among referred

borrowers.

Importantly, referral-induced market power operates not only across lenders but also within lending institutions. Even conditional on borrowing from the same lender, referred borrowers pay 7.5 basis points more than comparable non-referred borrowers. This within-lender premium indicates that referrals confer pricing power to individual loan officers by insulating referred borrowers from competition with other lenders and, in some cases, 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). Together, these findings show that referral networks soften price competition both across and within firms, creating localized monopoly power over borrowers.

A natural question is whether realtor–loan-officer referrals reduce processing frictions, potentially making referrals efficiency-enhancing in the contracting stage. 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: documentation is streamlined and loan officers have incentives to prioritize referred clients. 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 $870 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. To isolate this channel, we 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 approval certainty as the primary explanation.

Taken together, the evidence consistently points to localized market power created by referral networks as the central driver of elevated borrowing costs. Ex ante, referrals restrict borrower choice, reduce search intensity, and weaken outside options, shifting bargaining power toward lenders and raising negotiated interest rates. Ex post, conditional on being matched, referrals reduce coordination and processing frictions, generating faster closings. Yet neither modest time savings nor reduced denial risk can justify the magnitude or persistence of the referral premium. What appears to borrowers as a helpful shortcut thus carries an unintended consequence: higher borrowing costs

driven by localized lender market power.

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, 2020a; 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 away from market power toward alternative explanations, such as borrower sophistication (Bhutta et al., 2025), adverse selection arising from 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 merges on local mortgage competition.

We show that this view misses a key mechanism linking concentration and mortgage pricing. Traditional market-level lender 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. In this sense, the source of lender market power lies not in market structure per se, but in the home purchase 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 therefore 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 is also related to the literature on referral networks. While 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), 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 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 realtors and mortgage brokers. On the realtor side, the literature has shown that buyers' agents often steer buyers toward properties listed by agents 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; Barwick et al., 2017). On the mortgage side, prior studies show that 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 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 a limited branch network and lower brand recognition, increasing competition in the mortgage market. Agarwal et al. (2021) further shows that sole 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 examine the 100 vertical integration of realtors and mortgage brokers and shows that home borrowers experience an increase of 6 basis points in borrowing costs after the merger. 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 the distributional implications), uncover the underlying mechanisms, and evaluate non-financial borrower outcomes, including 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. Section 5 documents heterogeneous effects across homebuyers, and Section 6 explores the underlying mechanisms. Section **??** discusses potential efficiency rationales and justifications for the use of referral networks. Section 7 concludes with implications for policy and future research.

# 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 home buyers survey. According to the 2018–2022 quarterly National Survey of Mortgage Originations (NSMO), jointly managed 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, including both in-house officers employed by banks and mortgage companies and independent mortgage brokers, are intermediaries between homebuyers and lenders. Their primary responsibility is to help clients secure mortgage financing tailored to their financial situation and home purchase goals. Mortgage brokers, in particular, do not lend directly; instead, they work with multiple lenders to identify and arrange competitive loan options on behalf of the borrower. These professionals assist with every stage of the mortgage process: explaining loan products, collecting required documentation, submitting applications, coordinating with underwriters, and managing the timeline through to closing.

By navigating complex lending criteria and, in the case of independent mortgage brokers, offering access to a broad range of loan products from multiple lenders, these professionals can simplify the borrowing process. However, their recommendations may be influenced by ongoing relationships with specific lenders or realtors. For example, in-house loan officers can only offer mortgage products from the bank or institution that employs them.

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

create efficiency gains by fostering long-term, cooperative relationships that improve coordination, reduce transaction frictions, and enhance the overall client experience. On the other hand, such relationships can also lead to steering, where realtors guide homebuyers toward a narrow set of preferred loan officers, potentially limiting borrower choice and weakening price competition among lenders.

Although the Real Estate Settlement Procedures Act (RESPA) Section 8 prohibits kickbacks and unearned fees to realtors in exchange for referrals in mortgage lending, several common industry practices exist in 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, which makes 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. As a result, the referral system may inadvertently facilitate steering, raising concerns about transparency, borrower autonomy, and regulatory oversight.

## 2.2 Data

**CoreLogic Multiple Listings, Ownership Transfer, and Mortgage Data** We link the three CoreLogic data products – data on Mortgage, Owner Transfer, and Multiple Listing Service (MLS) transactions – to create a panel dataset of mortgage-financed home purchases and their 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—including interest rates and mortgage fees—in the HMDA data, 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 starting in 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.[2] 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 for at least 50% of listings. After applying these filters, the final sample comprises 703 counties across 41 states.

[2]Surprisingly, many MLS systems do not include buyer agent IDs, which would require us to use agent names, an approach prone to inconsistencies and ambiguities due to name variations.

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 would 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 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 IA3.

**Referral Network Proxy** Since direct observations of realtor–loan-officer referrals (e.g., explicit recommendations or formal agreements) are unavailable, we infer these networks indirectly by measuring the concentration of loan officers within each realtor's lender network. The intuition is straightforward: if buyers independently select 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 his/her 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 market share within that network – the fraction of the realtor's homebuyer clients whose loans were underwritten by that officer. The $CR4$ is simply 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 $CR4 < 0.4$ as low concentration and

those with $CR4 \geq 0.7$ as high concentration. Similarly, we follow the 2010 and 2023 DOJ Horizontal Merger Guidelines and use $HHI < 1500$ and $HHI \geq 2500$ as cutoffs for low and high concentration. High values of $CR4$ or $HHI$ provide suggestive evidence of referral networks. Below, we incorporate both measures in our regression analyses to test for the impact of referral networks.

It's important to emphasize that our measures of concentration are based on the frequency of interactions between realtors and individual loan officers, instead of interactions between realtors and lending institutions (or lenders). First, referrals are inherently interpersonal, occurring at the realtor and 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) officers within any given institution. Third, as discussed above, in a typical county-year, there are 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 details 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 the period from January 2018 to December 2021.

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

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 construc-

tion).[3] The resulting full regression sample includes 1.43 million transactions, intermediated by 81,306 distinct realtors.[4]

**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 tradeline records containing transaction-level details such as origination dates. The data also include credit scores, which we use as additional controls, 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 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 14,688 observations.

# 3 Summary Statistics and Descriptive Evidence

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

## 3.1 Attributes of Homebuyers and Mortgage Loans

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

Our sample comprises 1,432,602 home purchases, with an average loan amount of $285K. On the borrowing cost margin, we focus on the annual percentage rate (APR), which incorporates both the contract interest rate and up-front finance charges (fees and discount points). The average APR spread over the benchmark prime mortgage rate is 55 basis points. Joint applications account for

[3]The final sample also drops the observations in "thin" cells by the full set of fixed effects in the regressions.

[4]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.

43% of mortgages, and 97% of the mortgage amount 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 homebuyers in the 1% random sample display very similar mortgage characteristics, with the main difference being 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, 1.45 lenders on average receive their applications and pull mortgage-related credit inquiries. The number of credit inquiries serves as a useful proxy for a borrower's mortgage search intensity.

Demographically, homebuyers in our sample earn on average 1.45 times their county's median income, where the average county median is \$73.8K. 13% of borrowers are Hispanic. By race, 73% are White and 27% are minorities (9% Black, 5% Asian, and 13% other minority groups). Regarding property attributes, 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 45 days. Homebuyers in the 1% random subsample exhibit very similar demographic and housing characteristics, supporting the randomness and representativeness of this matched sample.

## 3.2 Descriptive Evidence on Realtor-Level Lender Concentration

Figure 2 plots the 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 ($CR4 \geq 0.4$). Panel B uses the Herfindahl–Hirschman Index and reports a similar pattern: 27% of realtors register $HHI \geq 2500$ (high concentration) and 52% exceed $HHI \geq 1500$ (medium concentration).

It is worth highlighting that high loan officer concentration among realtors does not arise mechanically because of a small transaction count or a limited number of local officers. The typical realtor completes 20 home purchases over our study period. More over, there are 232 loan officers (including both lender in-house officers and mortgage brokers) affiliated with 73 active lending institutions in a county–year.[5] Next, we present further evidence that the high loan officer concentration

[5] 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 with originating new purchase loans, refinancings, 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.

is likely driven by referrals.

## 3.3 Comparison between Realtor-Level and Market-level Lender Concentration

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

Figure 3 histograms of loan officer concentration ($CR4$) – the combined market share of the top four loan officers by transaction volume – measured at the realtor level (the red line) and at the county–year level (the blue line). Realtor-level concentration is substantially higher, averaging 59 percent compared to 36 percent at the market level. The stark contrast between these distributions suggests that realtor–loan-officer pairings are far from random and reflect systematic matching patterns.

To provide further evidence that these elevated realtor-level concentration measures reflect deliberate referral choices, Figure 3A plots realtor-level and market-level CR4 against market size (number of mortgage originations, binned into 100 groups). Consistent with the existing literature on industries with free entry (e.g., Bresnahan and Reiss, 1991; Dunne et al., 1988), market-level CR4 (blue triangles) falls as the number of originations rises: larger markets draw in new lenders, eroding individual shares and intensifying competition. At first glance, this appears to validate policies that focus on expanding market participation to erode lender market power.

However, when we turn to the lender concentration at the realtor level, a different pattern emerges. Within-realtor $CR4$ (red circles) remains *above 0.4* across all market sizes and actually *increases* with market size, providing strong evidence that referrals concentrate borrowers on a narrow set of officers within a realtor's network despite broader lender availability. Panel B repeats this exercise using the count of active lenders rather than originations. Again, market $CR4$ declines as lender counts grow, but realtor-level CR4 stays persistently high and rises in markets with more lenders. These patterns confirm that high lender concentration among realtors is not a mechanical consequence of limited options but is shaped by institutionalized relationships between realtors and a subset of loan officers, who maintain dominant positions among a realtor's clients even in highly competitive environments.

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 officers, simply increasing the number of lenders 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 intensity and presents OLS estimates of its effect on mortgage pricing. Section 4.3 then develops the IV strategy and reports causal estimates of the referral effect.

## 4.1 Effect of Loan Officer Concentration on Mortgage Rates

A natural benchmark in the mortgage-competition literature is *market-level lender concentration*, typically measured by a county-year $CR4$. If mortgage pricing were primarily shaped by the overall competitiveness of local lending markets, then this traditional concentration measure should have substantial explanatory power for borrowing costs. We begin by testing this prediction in a placebo exercise. 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 alone 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, and thus provides a more comprehensive measure of borrowing cost 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 directly observe the lock date for individual mortgages, which typically occurs 30–60 days before the 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 $q$th quintile of the distribution, with the bottom quintile omitted as the reference group. The set of controls $X_{irl}$ includes market-specific trends in mortgages costs (county, year-month fixed effects)[6], 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 (integer bins for house age, the number of bedrooms and bathrooms, as well as eleven square-footage bins).

Table 2 reports estimates of (1). We find no evidence that borrowers in markets with more concentrated lenders face higher borrowing costs. Specifically, 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. This likely reflects the empirical pattern that markets with more lenders tend to have higher average loan-officer concentration within realtors (Figure 3, Panel B). 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 finds limited effects of market-level lender concentration on mortgage pricing. Taken together, the placebo evidence supports the interpretation that our main results are driven by the concentration of realtors' loan–officer networks rather than by market-level lender structure.

Figure 4 contrasts the relationship between mortgage rate spreads and concentration measured at the realtor level (Panel A) versus the market level (Panel B), using 1% random sample estimates from Tables 3 and 2. Panel A plots 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. Spreads increase monotonically with realtor-level concentration. By contrast, Panel B plots spreads across quintiles of market-level lender $CR4$ and shows no systematic pattern. Together, the two panels visually display our regression evidence that the key margin for pricing is the concentration of realtors' loan–officer networks, not the overall concentration of lenders in the local market.

**Does Realtor-level Loan Officer Concentration Predict Borrowing Costs?** We next test whether the concentration of a realtor's loan-officer referral network helps explain mortgage pricing. Similar 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

[6] In this placebo analysis, $m$ denotes a county–year market. Because $CR4_m^{\text{lender}}$ is defined at the market level and exhibits little variation within a county–year–month cell, we include county and year–month fixed effects separately (rather than county×year–month fixed effects) so that the market-level regressors are identified.

borrower, loan, and market characteristics. Specifically, we sort realtors into quintiles based on the $CR4$ of their loan–officer network and estimate:

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

$Y_{irl}$ 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 $Quintile_q(CR4_r)$ are indicators equal to one if realtor $r$ falls into the $q$th quintile of the distribution, with the bottom quintile omitted as the reference group. The set of controls $X_{irl}$ includes market-specific trends in mortgages costs (county*year-month fixed effects), as well as all other borrower demographics, loan and house characteristics .

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 (mean spread = 55 bps), with rate gaps rising monotonically across quintiles. Columns (2)–(5) then add controls sequentially to show how much of this gap is explained by observables. Adding county*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 the DTI and LTV bins significantly reduces the top-quintile premium to 14.6 bps. Finally, Column (5) adds house characteristics as well as ZIP Code fixed effects, which futher reduce the the top-quintile premium to 12.5 bps.

Even though we include an extensive set of controls in the full-sample analysis, we are missing 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 that is matched to credit-bureau records, where FICO scores are observed. Panel B reports the corresponding estimates using this credit-enhanced sample.

First, we 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, which reassures us that the subsample is effectively random. As expected, standard errors are 2–3 times larger due to the smaller sample size. Nonetheless, the positive association between borrower mortgage costs and realtor–loan-officer concentration remains clear 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—about 16% lower than the corresponding

estimate without FICO controls. This attenuation indicates that part of the cost differential reflects selection on credit quality, and that controlling for borrowers' creditworthiness is important for isolating the effect of realtor–loan-officer concentration.

At the sample's average loan size ($306K), an 11.7 bp premium translates into $358 in extra annual interest payment. Equivalently, for a borrower to obtain the same rate as those working with bottom-quintile realtors, they would need to pay about $1,641 in upfront fees (assuming a point-to-rate reduction factor of 4.6 as estimated in Bhutta et al. (2020b)). 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.

## 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 referrals from realtors to a limited set of loan officers.

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 operates in a highly concentrated loan officer network, meaning that a large share of her past transactions are intermediated by a small set of loan officers. This pattern is consistent with the realtor routinely steering homebuyers toward a stable group of preferred financing partners rather than searching broadly on behalf of each homebuyer. 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, so mortgages originated through these pairs are classified as *Referral* loans:

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

Figure 2 suggests that 85% of realtors have loan-officer concentration above the 0.4 threshold. 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.

Table 1 shows that the same level of referral network prevalence on the borrower side: 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 22% reject the referral and initiate their own search for loan officers. Taken together, 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 now 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-year–month fixed effects, borrower demographics, loan characteristics, and house hedonics.

We begin by estimating these two specifications using OLS, with results reported in Table 4. Column (1) estimates Equation (3) in the full sample and finds that the coefficient on the Referral indicator is 0.099, implying that borrowers who borrow through a realtor's concentrated 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 (*Refer, But Rejected*). The estimated premium for this group is economically small (1.8 bps) while the premium for the "compliant" referral group increases to 11.6 bps. The sharp gap between these coefficients suggests that the 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.[7]

Columns (3) and (4) repeat the analysis in the 1% random subsample matched to credit-bureau records, which allows 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 "Refer, But Rejected" 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.

Taken together, the evidence indicates that steering through realtor-loan-officer relationships is associated with systematically higher borrowing costs: borrowers who borrow within concentrated referral networks and follow the referred loan officers pay a meaningful mortgage premium.

[7] One reason the "Refer, But Rejected" estimate in column (2) may be slightly positive is that public HMDA does not report FICO scores; once we can control for FICO in the credit-matched subsample, this coefficient becomes effectively zero (column (4)).

## 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 other forces that influence which loan officer a borrower ends up using. In particular, choosing a high-cost loan officer may be related to factors that are correlated with the realtor 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. To find a valid IV, 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 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. Here we use the Herfindahl–Hirschman Index ($HHI$) as our concentration measure instead of $CR4$ to avoid a mechanical correlation between $Referral_{rl}$ and IV, since both would otherwise be direct functions of $CR4_r$. This choice is not crucial. As shown in Table IA4 in the Intenet Appendix, when we construct the instrument using alternative concentration measures ($CR4$), the

estimated referral effect 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 take two steps to achieve this. First, we estimate a flexible BLP demand model that controls for the main endogenous 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 the refinance-mortgage sample, where realtor involvement is minimal. This strategy allows $\widehat{HHI}_r$ to reflect the loan officer concentration that would arise in the absence of referral from realtors.

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 study 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 is to recover borrower preferences over lenders in the absence of referrals, we estimate the following BLP demand model (Berry et al., 1995) using refinanced loans state by state and incorporate a rich set of controls to allow flexible preferences among borrowers. Specifically, we assume that the utility of borrower $i$ who secures a loan from lender $j$ in market $m$ is the following:

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

where $X_{jm}$ is a vector of lender-market characteristics. Specically, 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 a state).[8] 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, $\xi_{jm}$ represents an unobserved vertical component (such as quality of customer service) specific to lender $j$ in mar-

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

ket $m$, and $\varepsilon_{ijm}$ 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 $\xi_{jm}$ is inverted such that the model-predicted lender shares are equal to the observed lender shares in each market.

Note that the lender-choice model above implicitly assumes that borrowers choose a lending institution rather than a specific loan officer. This is for two reasons. First, we do not observe demographic or socio-economic characteristics of individual loan officers, which makes 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 therefore first choose a lender according to this model and are then probabilistically allocated across that lender's loan officers in the same market based on their observed lending volumes.

Finally, after estimating the parameters in Equation (6), we apply the model to the purchase-mortgage sample used in the main analysis and predict, for each borrower, the choice probabilities over all lenders available in her market. Using these predicted probabilities, we simulate 200 draws of lender choice for each homebuyer. To obtain the loan-officer choices needed to construct $\widehat{HHI}_r$, we then assume that, conditional on choosing a given lender, the borrower is randomly assigned to that lender's loan officers in the same market in proportion to their observed lending volumes. Aggregating these simulated loan officer choices at the realtor level yields a counterfactual loan officer distribution for each realtor, from which we compute $\widehat{HHI}_r$, as the average Herfindahl–Hirschman Index across the 200 simulations.

$$\widehat{HHI}_r = \frac{1}{S} \sum_{sim=1}^{200} \hat{HHI}_r^{sim}, \tag{6}$$

We describe the data construction, BLP estimation, and counterfactual simulations in greater detail in Section A of the Internet Appendix.

Figure 5 compares the counterfactual and observed distributions of realtor-level loan officer concentration. 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. By contrast, the observed distribution of $HHI_r$ (blue solid line) is markedly shifted to the right, indicating substantially higher concentration in actual data. The stark gap between the two distributions suggests that referral networks play an important role in shaping homebuyers' financing choices. In particular, in the counterfactual world the probability that a realtor has a highly concentrated loan officer network is close to zero, whereas in the observed data 52.36% of realtors have $HHI_r$ above the medium-

concentration threshold of 1500. For completeness, Panel B plots the distributions of simulated and observed $CR4$, which display a very similar pattern.

**IV Results** With the constructed IV, we then run reduced form regressions with the specification:

$$Y_{irl} = \kappa(HHI_r - \widehat{HHI}_r) + X'_{irl}\delta^{RF} + \varepsilon_{irl} \tag{7}$$

We then estimate the second stage of IV regressions, replacing the endogenous referral indicators with its first-stage fitted value:

$$Y_{irl} = \alpha^{2SLS}\widehat{Referral}_{rl} + X'_{irl}\delta^{2SLS} + \varepsilon_{irl} \tag{8}$$

This approach allows us to interpret $\alpha^{2SLS}$ as the causal effect of referral 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) relies on 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, 18.6 bps higher mortgage rates; this premium remains sizable at 18.6 bps after controlling for FICO. At the sample's average loan size ($285K), 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 if the borrower does not follow the referral.

The estimates imply a sizable referral effect. First, relative to the average APR spread of 51 bps, the referral premium represents a 35% increase in borrowing costs. A complementary way to gauge magnitude is to compare it with the overall mortgage price dispersion. In our sample, we esitmate the difference between the 90th and 10th percentile mortgage rate is 86 bps.[9] Our 18.6 bp referral premium accounts for roughly 22% of this dispersion, suggesting that more than one-fifth of the

[9]This mortgage rate dispersion exceeds the 55 bps estimated by Bhutta et al. (2025). Three differences in measurement and sample construction likely contribute to the gap. First, we measure pricing using the APR spread, which incorporates both the contract interest rate and upfront origination charges, whereas Bhutta et al. (2025) focus on the locked contract rate. As a result, our measure captures additional cross-borrower variation in total borrowing costs. Second, Bhutta et al. (2025) observe the lock date and can absorb high-frequency pricing movements using lock-date fixed effects; we do not observe lock dates and instead control for the coarser proxy of origination-month fixed effects. Third, Bhutta et al. (2025) include both purchase and refinance mortgages, while our analysis is restricted to purchase loans, which typically exhibit greater dispersion in borrower-paid pricing.

cross-borrower price variation may be attributable to loan-officer market power operating through realtor referral networks.

# 5 Heterogeneous Effects of Referrals

We explore heterogeneity in these effects across different market and borrower characteristics.

## 5.1 Heterogeneous Effects of Risky Borrowers

Table 8 examines how the referral effect varies across homebuyers with different risk profiles. Column (1) focuses on borrowers with low credit scores. For homebuyers with FICO scores below 660, the estimated referral premium is 11.9 bps higher than the benchmark effect of 17.1 bps for the rest of the sample. Although this incremental spread is not statistically significant – likely due to the limited size of the 1% credit-matched sample – it is economically meaningful and suggests that lower–credit-quality borrowers may be particularly vulnerable to higher borrowing costs in referral-driven relationships.

Columns (2)–(3) examine borrower risk along the debt-to-income (DTI) margin. Column (3) shows that for borrowers with DTI above 45% – about one fifth of homebuyers in our sample – referrals raise mortgage costs by an additional 10.9 bps relative to the benchmark referral effect of 16 bps for the rest of the sample. This incremental spread corresponds to roughly $1,529 in upfront points at the average loan size. Columns (4)–(5) focus on high–loan-to-value (LTV) borrowers. Column (5) indicates that borrowers with LTV above 95% – more than one third of homebuyers – pay an additional 21 bps on top of the 10.1 bps benchmark referral effect. This incremental cost is equivalent to about $2,946 in upfront fees, underscoring that highly leveraged borrowers are particularly exposed to the financial consequences of referral-driven steering.

These heterogeneous effects point to an important distributional implication of referral-driven steering: the financial burden is not borne uniformly across borrowers, but is instead tilted toward those who are already most vulnerable. Low-FICO, high-DTI, and high-LTV borrowers are households with thinner buffers and fewer outside options. They have to systematically pay larger referral premia than better-positioned homebuyers. In other words, the same referral practice that modestly raises costs for average borrowers can meaningfully amplify financing strains for those at the margin of credit access. This pattern suggests that realtor–loan-officer networks may not only reduce overall allocative efficiency, but also embed an inequitable price of advice, with the steepest toll falling on the riskiest and least resilient homebuyers.

## 5.2 Heterogeneous Effects of Homebuyer Demographics

Table 9 examines how the referral effect varies across homebuyers of different socio-economic backgrounds. Columns (1)–(2) group homebuyers by ethnicity and reveal very stark heterogeneity. For Hispanic homebuyers, the referral effect is 49.1 bps, translating into roughly $6,888 in additional upfront costs. This is substantially larger than for non-Hispanic borrowers and is plausibly related to language barriers, differential access to information, and a greater tendency to rely on realtors' recommendations as trusted intermediaries—especially when intermediaries share cultural or local ties, as documented in work on trust and advice-taking in financial and labor-market decisions (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 (or $2,413) for white borrowers, 24.8 bps (or $3,479) for Black borrowers, and 21.5 bps (or $2,736) for Asian borrowers

These patterns indicate that referral-driven costs are systematically higher for ethnic and racial minority borrowers than for otherwise comparable white borrowers. In other words, when homebuyers rely on realtor–loan-officer networks, minority households are more likely to be steered into costlier mortgage arrangements and pay larger rate premia. This suggests that steering through referral networks may not only raise average borrowing costs, but also amplify existing racial and ethnic disparities in mortgage pricing and, by extension, in the overall cost of homeownership. These results call for greater regulatory attention to the microstructure of referral networks and the institutional incentives that sustain them.

Finally, columns (5)–(6) further examine heterogeneity across income ratio terciles:[10] the referral premium is smallest for high-income borrowers (about 11–12 bps) and markedly larger for low- and medium-income borrowers (about 19–27 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 are associated with higher mortgage costs for borrowers. In this section, we interpret these findings through a unified economic mechanism, arguing that referral networks shape mortgage outcomes through a search-based channel that confers localized market power on lenders.

[10]The sample split is based on borrowers' income relative to the county median income. The high-income tercile has an average income of $166k (income ratio 2.4), the middle-income tercile has an average income of $86k (income ratio 1.2), and the low-income tercile has an average income of $53k (income ratio 0.7).

To fix ideas, consider a two-stage environment. In the first stage, borrowers search across lenders at a cost to form a feasible choice set. In the second stage, conditional on matching with a lender, the mortgage rate is determined through Nash bargaining, with each side's bargaining power governed by its outside options.

Realtor referrals operate primarily by compressing the borrower's choice set in the search stage, steering borrowers toward a limited number of loan officers. This restriction weakens borrowers' outside options and shifts bargaining power toward lenders, resulting in higher negotiated mortgage rates. At the same time, referrals may reduce coordination and screening frictions within a match - such as information transmission or document processing - leading to faster completion once a lender is chosen. This framework naturally generates a joint set of empirical patterns that are difficult to reconcile with alternative explanations: higher mortgage rates, reduced search intensity, and possibly faster closings. We have established robust causal evidence on pricing effects in the main analysis; we now turn to evidence on search behavior and processing speed to complete the interpretation.

**Search Intensity** A central implication of the search-and-bargaining framework is that referrals reduce borrowers' incentives to search by compressing their effective choice sets. When borrowers face fewer credible alternatives, the expected return to continued search declines, and they optimally contact fewer lenders. Reduced search, in turn, weakens borrowers' outside options and their bargaining position in subsequent rate negotiations.

We test this mechanism directly by examining the causal effect of referrals on borrowers' mortgage search activity. Measuring mortgage search is challenging because standard mortgage datasets record only the final choice of lender and loan officer. We address this data 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 by the credit bureau as mortgage-related inquiries, which also enter the credit score calculation. We use the number of such inquiries as a proxy for the intensity of (serious) mortgage search.

Table 6 reports IV estimates of the effect of referrals on the number of mortgage-related credit inquiries. Column (1) examines the number of lenders approached within a 90-day window prior to mortgage closing. Borrowers who use a referral loan officer make, on average, 0.168 fewer inquiries, which represents an 12% decline relative to the baseline mean of 1.45 inquiries. 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 inquiries made within a 60-day window prior to mortgage closing. The referral effects are even stronger in relative terms: the number of credit inquiries falls by 13%, and the share of homebuyers with at least two inquiries declines by 38%.

These results provide direct evidence that referrals meaningfully dampen borrowers' search intensity. Homebuyers who rely on realtor-referred loan officers contact fewer lenders and are substantially less likely to obtain competing quotes. This behavior is consistent with an equilibrium response to restricted outside options rather than borrower inattention or sorting. Importantly, the joint pattern of reduced search and higher borrowing costs documented earlier is difficult to reconcile with alternative explanations based on lender quality or borrower sophistication, which would predict equal or greater comparison shopping among referred borrowers.

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 calculation likely understates the full effect for two reasons. First, credit inquiries capture only late-stage, committed shopping and miss earlier information-gathering that may be more responsive to referrals. Second, the marginal return to search is plausibly larger for borrowers facing unfavorable initial offers from referred loan officers.

**Pricing Power Conditional on Lender Choice** In a search-based framework, referrals constrains borrowers' choice sets, with implications that extend beyond lender selection to pricing conditional on a match being formed. By limiting borrowers' outside options, referrals in the search stage weaken competitive discipline in the subsequent bargaining stage when a lender is chosen. As a result, loan officers working with referred borrowers retain greater pricing power relative to those working with non-referred borrowers within the same lending institution. We test this core implication by examining whether referred borrower are charged higher rates than non-referred borrowers who obtain mortgages from the same lender.

To do so, we re-estimate the referral regressions including lender fixed effects, thereby comparing referred and non-referred borrowers who obtain mortgages from the same lending institution. Table 7 reports the results with 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 even more flexibly by interacting lender fixed effects with year–month–county fixed effects, 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 mortgage cost 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 mortgage cost premium is 7.5 bps, down by 60% from the premium of 18.6 bps without controlling lender fixed effects.[11]

Taken together, these results imply that approximately three-fifths of the referral effect operates through a between-lender channel, while the remaining two-fifths reflects differential pricing within lenders. Even holding the lender constant, referred borrowers pay higher rates, consistent with that referrals insulate loan officers from competitive pressure by limiting borrowers' credible outside options. Further, weakened outside options continue to matter after matching has occurred, ruling out explanations based solely on borrower sorting or lender quality. Referral networks thus soften competition both across and within lending institutions, reinforcing the localized market power generated in the search stage.

**Days to Close Conditional on Lender Choice** A natural question is whether referral networks generate offsetting efficiency gains once a borrower has matched with a lender. In particular, referrals may reduce coordination and screening frictions in the contracting stage, potentially leading to faster processing and greater certainty. In this case, referrals may reallocate frictions across stages: they raise search frictions in the search stage while reducing coordination frictions conditional on matching. We assess the magnitude of these potential efficiency benefits by examining time to close and mortgage denial risk.

Mortgage transactions are not only financially costly but also time-consuming, and some borrowers may value speed and certainty. Table 10 reports IV estimates of the effect of referrals on time to close, measured as the number of days between contract acceptance and mortgage closing. Column (1) presents full sample results, conditional on a successful closing: borrowers who use referred loan officers complete their purchase about 1.75 days sooner. Column (2) uses the 1% matched random sample and adds credit score controls; the estimated effect on days to close is not statistically significant, and the economic magnitude is small. Relative to a baseline mean of 40 days, the point estimate of 1.5 days implies a little under a 4% faster closing time.

Columns (3)–(6) examine whether a transaction exceeds specific closing-time thresholds. Columns (3) and (4) test whether closing takes more than 30 days, while Columns (5) and (6) use a 60-day threshold. In our sample, 72% of transactions take more than 30 days to close, and 9% exceed 60

[11] In this subsample we cannot include lender×year–month×county fixed effects because the interacted cells become too sparse, but 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.

days. Referrals are associated with a 5–6 percentage point lower probability of exceeding 30 days and about a 1 percentage point lower probability of exceeding 60 days.

These results suggest that referral networks may deliver modest efficiency gains by streamlining parts of the mortgage process. Such benefits likely stem from the established working relationships between realtors and referred loan officers, which can facilitate smoother communication, faster document exchange, and more responsive processing. However, while referrals help reduce moderate delays, they do not appear to meaningfully prevent longer or more severe closing delays. Thus, although referral networks offer some convenience, the time savings are relatively limited and unlikely to offset the higher borrowing costs documented earlier.

**Mortgage Denial Risk** Another potential benefit is that referral networks may offer borrowers greater certainty, helping them avoid mortgage denial and ensuring that transactions close successfully. Because we do not observe denied applications, we test this by focusing on borrowers who face minimal denial risk. If even these low-risk borrowers pay more when using referred loan officers, then the premium is unlikely to be driven by approval risk mitigation alone.

Table 11 reports referral effect for high–credit-quality borrowers subsamples. Because we must condition on credit scores, this analysis uses 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 we raise the FICO threshold, the sample size shrinks substantially, but the estimated referral premium remains economically large and statistically significant. Even for borrowers with FICO scores above 780, who face minimal risk of denial, the referral premium is as large as 9.3 bps (roughly $1,305 in equivalent upfront costs).

Column (4) further tightens the high-FICO sample (above 700) by excluding borrowers with high debt-to-income ratios (DTI $>$ 36%) and high loan-to-value ratios (LTV $>$ 90%). Even in this very low-risk group, the referral premium remains statistically significant at 16 bps, underscoring that the referral effect is not driven solely by marginal or constrained borrowers.

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 referral premium appears to stem from reduced competition—an effect that persists even among borrowers with the lowest likelihood of denial.

**Summary** Overall, the evidence points to a search-based mechanism through which referral networks generate localized market power. By compressing borrowers' choice sets in the search stage, referrals reduce search intensity and weaken outside options, which in turn shifts bargaining power toward lenders and loan officers, raising negotiated mortgage rates. While referrals modestly reduce

processing frictions after a match is formed, these efficiency gains are economically small relative to the estimated referral premium. What appears to borrowers as a convenient shortcut thus operates as a distortionary market structure, softening competition and elevating prices in otherwise competitive mortgage markets.

# 7 Conclusion

This paper has shown that the U.S. mortgage market is not only shaped 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. More importantly, we show that this concentration is not an innocuous byproduct of efficiency or borrower convenience. Instead, it systematically raises borrowing costs, even in markets that appear competitive on the lender side, and even for households who are otherwise strong credit risks.

Our key contribution is to quantify the economic bite of these referral networks and to disentangle the underlying mechanisms. Causally, being steered to a realtor-linked loan officer increases mortgage costs by about 12 basis points, which accounts for a nontrivial share of the overall price dispersion documented in the literature. We further show that this premium operates through two distinct channels. First, referrals materially depress search: steered borrowers obtain fewer quotes and are significantly less likely to contact multiple lenders, leaving money on the table simply because they do not shop. Second, even conditional on lender choice, referred loan officers charge higher rates than their colleagues at the same institution, implying that referral links confer a degree of within-lender market power. Together, these channels imply that "who you are sent to" matters as much as "which lender you choose" for the price you pay.

At the same time, referrals are not without benefits. We find that referred borrowers close about a day faster on average, suggesting that realtor–loan-officer relationships can facilitate smoother processing, reduce uncertainty, and provide valuable hand-holding, especially for less experienced borrowers. But in our data the procedural gain is small relative to the recurring financial cost. Even among high-FICO, low-DTI, low-LTV borrowers who face virtually no underwriting risk, referred loans remain meaningfully more expensive. This pattern is hard to reconcile with a purely efficiency-enhancing view of referrals and instead points to a pricing advantage sustained by intermediated, relationship-based access.

The distributional implications of these networks are also consequential. We find that the referral premium is systematically larger for Hispanic borrowers and, more broadly, for racial and ethnic

minorities, while remaining substantial for white borrowers as well. Riskier borrowers – those with high DTI or very high LTV – also bear disproportionately higher premia when steered, precisely the groups least able to absorb extra costs. By contrast, there is little systematic difference by gender. Taken together, these facts suggest that referral-driven market power does not create inequalities from nothing, but it does tend to lean against the same borrowers who are already disadvantaged in terms of information, language, or bargaining position.

Conceptually, our findings speak to a broader point about how competition translates—or fails to translate—into consumer outcomes in retail financial markets. On paper, the mortgage sector features many active lenders, sophisticated secondary markets, and relatively low regulatory barriers to entry. Yet from the borrower's perspective, choice is mediated by real estate agents whose business model is only loosely aligned with minimizing financing costs. Referral networks then act as local bottlenecks: they concentrate flow on a small subset of loan officers, allow those officers to partially insulate themselves from comparison, and weaken borrowers' incentives to search. This helps explain why substantial price dispersion persists even among observably similar borrowers in the same market and loan segment, and why improvements in lender-side competition alone have had limited impact on the effective prices households pay.

From a policy standpoint, the results suggest that focusing solely on lender concentration or headline rate competition is incomplete. Regulation such as RESPA already prohibits explicit kickbacks, but our evidence indicates that informal referral structures – shaped by relationships, repeated interaction, and opaque expectations – can still support economically meaningful markups without violating formal rules. A more effective approach may require shifting attention toward the intermediated structure of choice: enhancing transparency around referral relationships, making it easier for borrowers to observe and compare alternative offers, and scrutinizing compensation and incentive schemes that tie realtors and loan officers together. At minimum, regulators and policymakers should recognize that "competition on the shelf" can be undermined by "steering at the counter."

Finally, our results open several avenues for future work. One is to study how emerging technologies – online platforms, digital brokerages, and automated pre-approval tools – interact with traditional referral networks: do they discipline steering or simply repackage it? Another is to explore the dynamics of these networks over the housing cycle, including whether periods of stress (such as refinancing booms or credit crunches) strengthen or weaken referral-based market power. More generally, our analysis underscores that understanding household debt markets requires looking beyond contracts and institutions to the informal ties that connect them. Realtor–loan-officer networks are one such set of ties; similar structures likely exist in small-business lending, auto finance, and student loans. Mapping and quantifying these relationships is essential for a realistic

view of how market power is exercised, how price dispersion arises, and which households bear the costs.

# 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, 2020a, Paying Too Much? Price Dispersion in the U.S. Mortgage Market, *Finance and Economics Discussion Series* 2020.

Bhutta, Neil, Andreas Fuster, and Aurel Hizmo, 2020b, Paying too much? price dispersion in the us mortgage market FEDS Working Paper No. 2020-62.

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

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, 2016, Understanding in-house transactions in the real estate brokerage industry, *RAND Journal of Economics* 47, 1057–1086.

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, Forthcoming, Racial disparities in housing returns, *American Economic Review* .

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 .

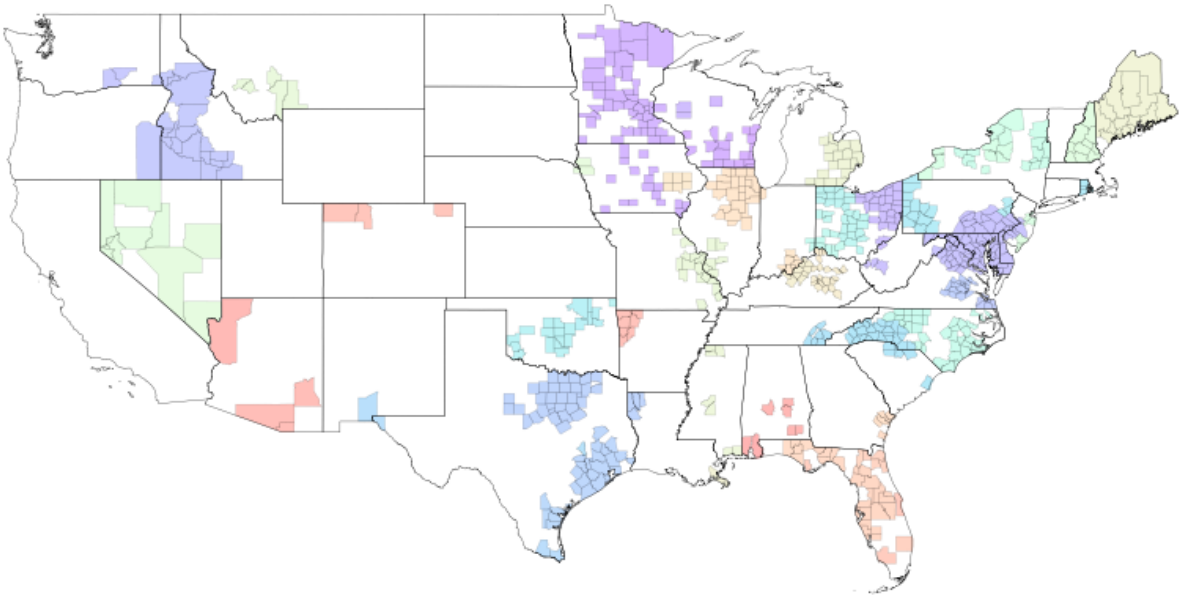
Spader, Jonathan S., and Roberto G. Quercia, 2011, Mortgage Brokers and the Refinancing Transaction: Evidence from CRA Borrowers, *The Journal of Real Estate Finance and Economics* 42, 181–210.

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.

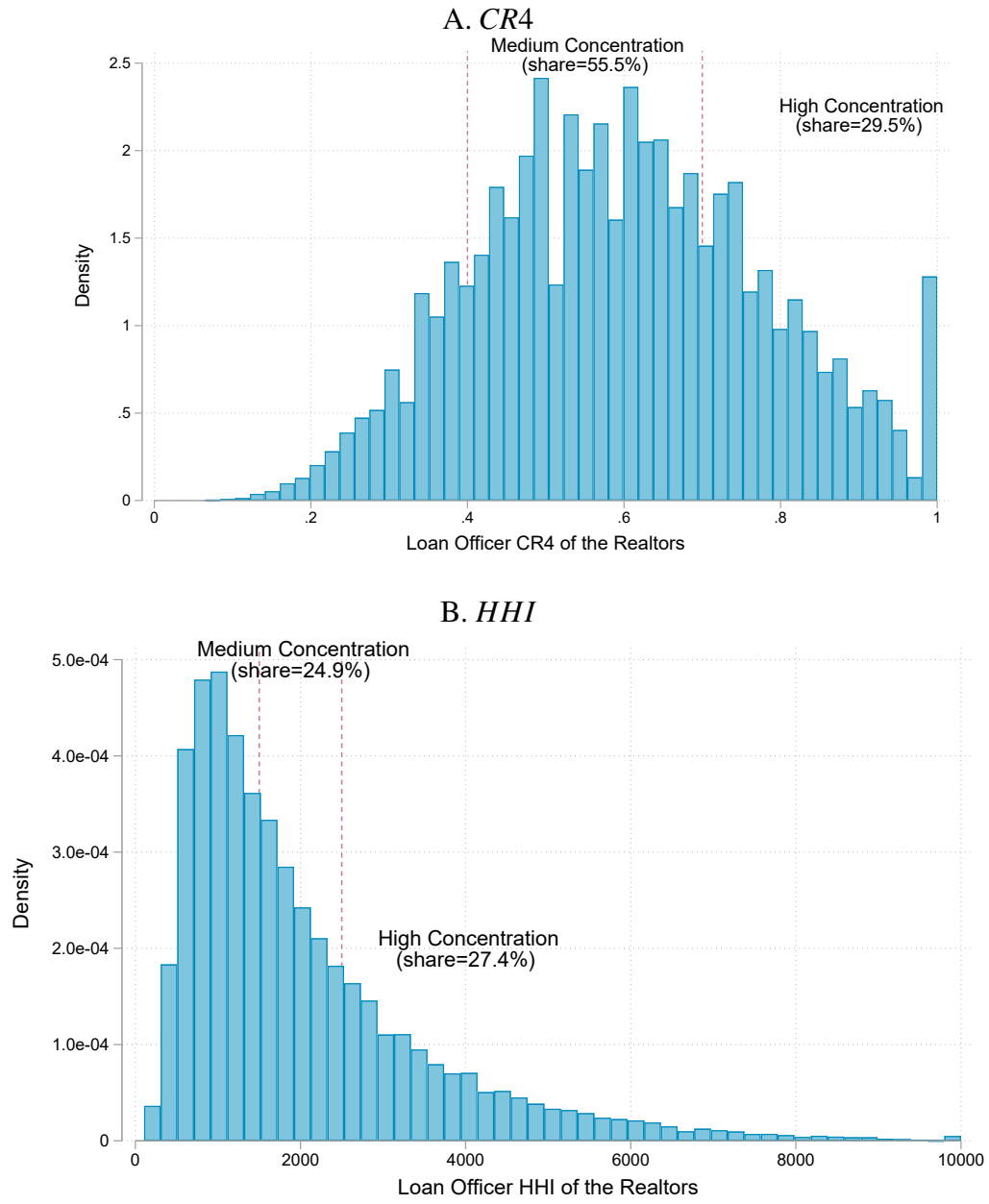
# Figures

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.

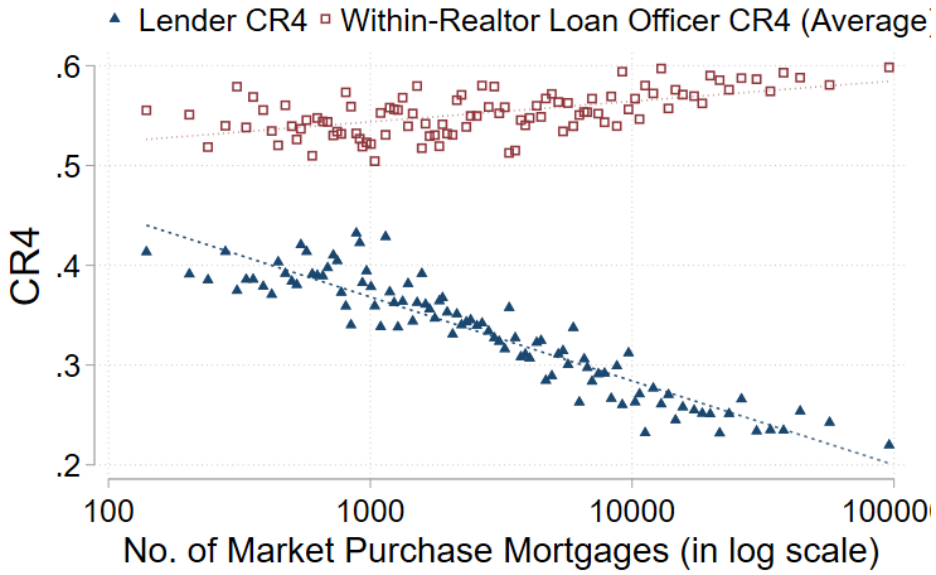
Figure 2: Distribution of Realtors' Loan Officer Concentration



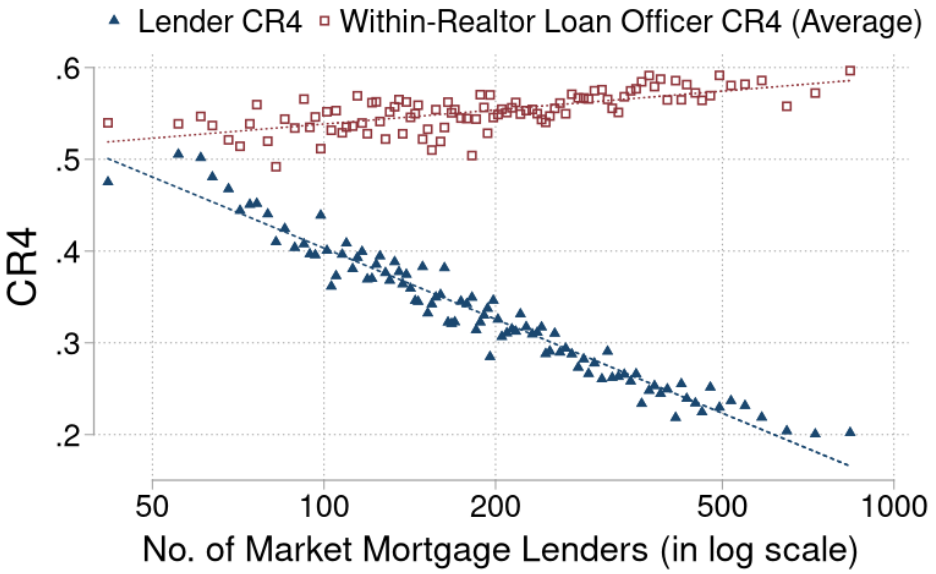

*Note*: This figure plots the distributions of realtors' loan officer concentration measures ($CR4$ in panel A, $HHI$ in panel B). For each realtor, we define his/her lender network as the set of all loan officers who financed loans for at least one of the realtor's buyer clients. We then calculate each loan officer's market share – the fraction of that realtor's buyer clients whose mortgages the officer underwrites. The $CR4$ is simply the sum of the top four loan officers' market shares for a given realtor, while the $HHI$ is the sum of squared market shares across the realtor's entire loan officer network.

Figure 3: Loan Officer Concentration and Lender Concentratration by Market Size

A. By the Size of Purchase Mortgages



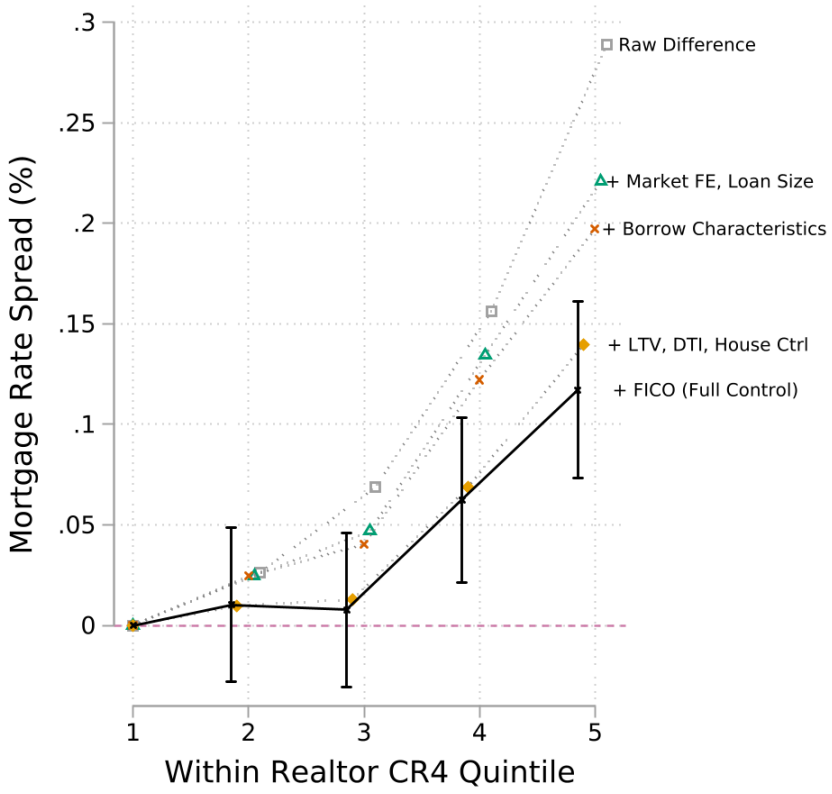

B. By the Number of Mortgage Lenders



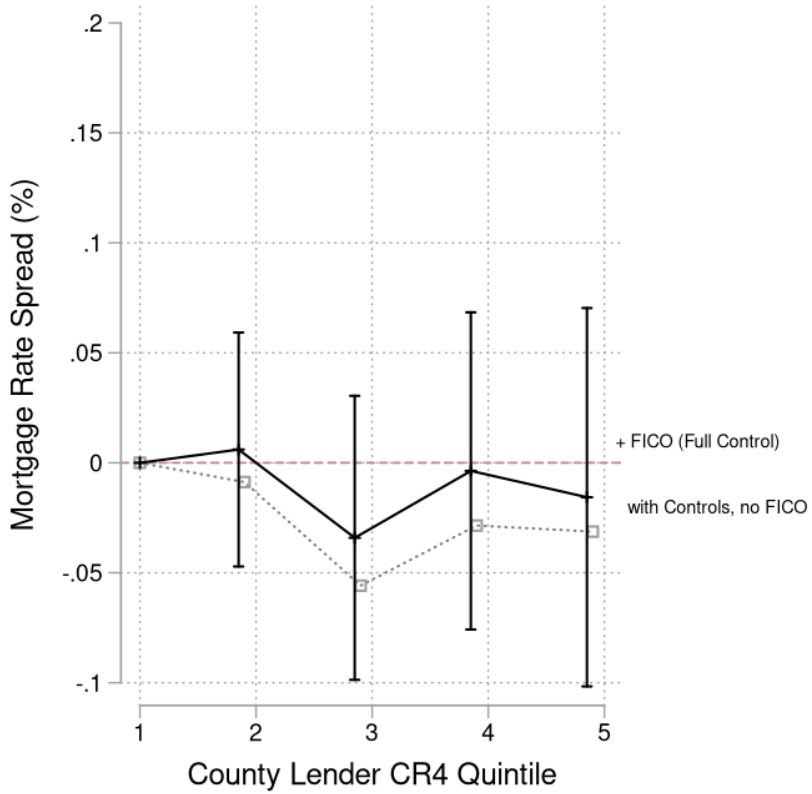

*Note*: These figures depict the relationship between average within-realtor loan officer concentration (measured by *CR*4) and lender concentration (*CR*4) with market size. Panel A sorts markets (county-year) based on the number of purchase mortgages in our sample, while Panel B sorts markets (county-year) according to the number of lenders from HMDA. Both figures present binned scatter plots (100 bins) of maket-level concentration measures alongside their respective linear fit lines. The red scatter points and lines represent the average within-realtor loan officer *CR*4. When calculating within-realtor concentration, we restrict the analysis to realtors with a minimum of 10 transactions.

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

A. By Within-Realtor Loan Officer *CR*4



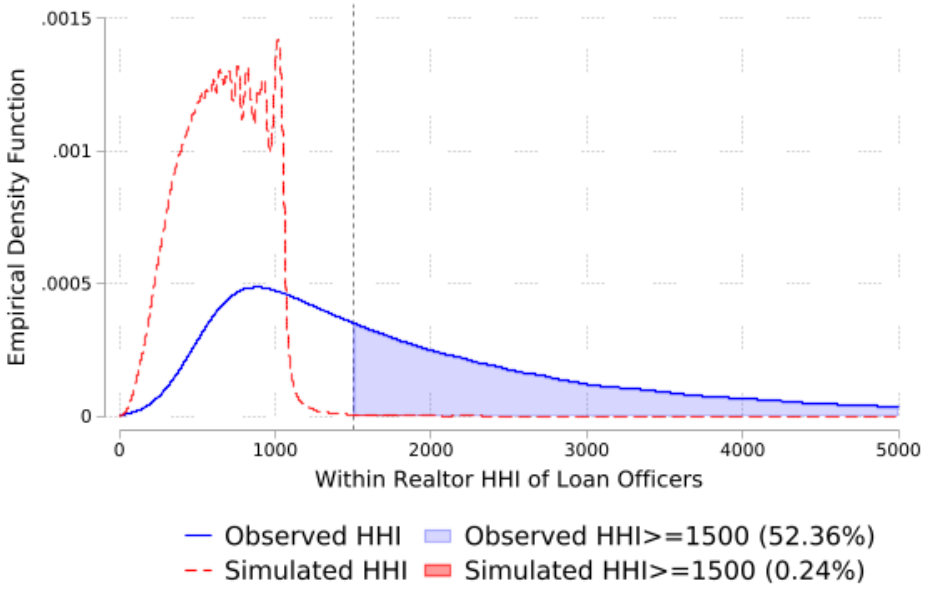

B. By Market Lender *CR*4



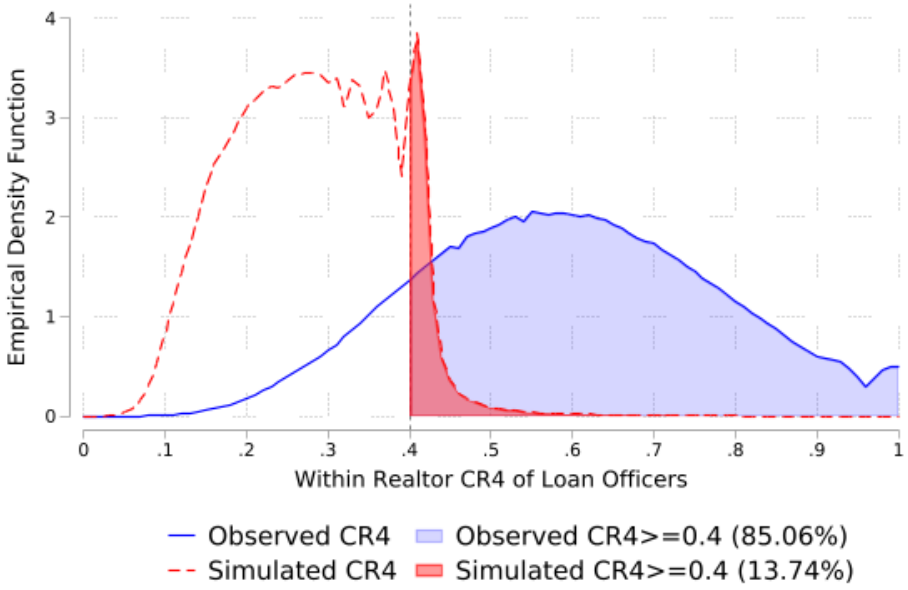

*Note*: This figure presents the relationship between mortgage APR spreads and lender concentration across different concentration quintiles. The y-axis represents the mortgage APR spread, defined as the deviation from the benchmark rate offered on prime mortgage loans of a comparable type. In Panel A, concentration is measured by the within-realtor *CR*4 of loan officers. In Panel B, concentration is measured by the market-wide lender *CR*4.

Figure 5: Distribution of Network Concentration of Realtors (Simulation vs Observation)

A. Measured by *HHI*




B. Measured by *CR*4




*Note*: This figure plots the distributions of loan-officer concentration of realtors from both the observed and simulated sample.

# Tables

Table 1: Homebuyer Attributes

| | Full Sample (1) | | 1% Random Sample (Matched to Credit Records) (2) | |
|---|---|---|---|---|
| | mean | s.d. | mean | s.d. |
| 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) |
| Close 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) |
| Refer, But Rejected | 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 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 lenders. "Refer, But Rejected" 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 lenders. The FICO score and credit inquiry counts are only available in the random sample.

Table 2: Placebo Test: Mortgage Rate Spreads By Market 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 realtor concentration quintiles. The concentration is measured by within realtor *CR*4 of the loan officer mortgage shares. 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 median income of the same year. 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 *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 concentration quintiles. The concentration is measured by within realtor *CR*4 of the loan officer mortgage shares. 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 median income of the same year. "Joint Application" takes value 1 if there is more than one applicant on the mortgage loan. Standard errors are clustered at the county level, and reported in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table 4: Effect of Referral 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) |
| Refer, But Rejected | | 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 realtor-loan officer referral network effect on mortgage costs. The concentration is measured by within realtor $CR4$ of the loan officer mortgage shares. The mortgage costs are measured by both 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 lenders. "Refer, But Rejected" 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 lenders. The income ratio is defined as the reported income over the county median income of the same year. Standard errors are clustered at the county level, and reported in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

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

| | Referral (First Stage) | | Mortgage Rate Spread (off Prime Loans, in %) | |
|---|---|---|---|---|
| | | 1% Random Sample | | 1% Random Sample |
| | (1) | (2) | 2SLS (3) | 2SLS (4) |
| $(HHI\text{-}\widehat{HHI})/10000$ | 1.228*** | 1.152*** | | |
| | (0.016) | (0.035) | | |
| Referral | | | 0.165*** | 0.186*** |
| ($CR4 \geq 0.4$, 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 regression estimates of realtor-loan officer referral network effect on mortgage costs. The concentration is measure by within realtor $CR4$ of the loan officer mortgage shares. The mortgage costs are measured by APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. $Referral$ equals 1 if the homebuyer's realtor $CR4$ is above 0.4, and she works with a top 4 loan officer of the realtor. It is instrumented with borrower preference adjusted loan officer concentration measure $CR4\text{-}\widehat{CR4}$. The income ratio is defined as the reported income over the county median income of the same year. Standard errors are clustered at the county level, and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: 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 regression estimates of realtor-loan officer referral network effect on credit Inquiries. The concentration is measure by within realtor $CR4$ of the loan officer mortgage shares. $Referral$ equals 1 if the homebuyer's realtor $CR4$ is above 0.4, and she works with a top 4 loan officer of the realtor. It is instrumented with borrower preference adjusted loan officer concentration measure $CR4$-$\widehat{CR4}$. The income ratio is defined as the reported income over the county median income of the same year. Standard errors are clustered at the county level, and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Effect of Referral 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 regression estimates of realtor-loan officer referral network effect on mortgage costs. The concentration is measure by within realtor $CR4$ of the loan officer mortgage shares. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. $Referral$ equals 1 if the homebuyer's realtor $CR4$ is above 0.4, and she works with a top 4 loan officer of the realtor. It is instrumented with borrower preference adjusted loan officer concentration measure $CR4$-$\widehat{CR4}$. The income ratio is defined as the reported income over the county median income of the same year. Standard errors are clustered at the county level, and reported in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table 8: Effect of Referral 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 regression estimates of realtor-loan officer referral network effect on mortgage costs across different borrower groups. The concentration is measure by within realtor $CR4$ of the loan officer mortgage shares. The dependent variable is the mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. $Referral$ equals 1 if the homebuyer's realtor $CR4$ is above 0.4, and she works with a top 4 loan officer of the realtor. It and its interactions are instrumented with borrower preference adjusted loan officer concentration measure $CR4$-$\widehat{CR4}$ and interaction terms. The income ratio is defined as the reported income over the county median income of the same year. Standard errors are clustered at the county level, and reported in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table 9: Heterogeneous Effect of Referral 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 regression estimates of realtor-loan officer referral network effect on mortgage costs across different borrower groups. The concentration is measure by within realtor $CR4$ of the loan officer mortgage shares. The dependent variable is the mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. $Referral$ equals 1 if the homebuyer's realtor $CR4$ is above 0.4, and she works with a top 4 loan officer of the realtor. It and its interactions are instrumented with borrower preference adjusted loan officer concentration measure $CR4$-$\widehat{CR4}$ and interaction terms. The income ratio is defined as the reported income over the county median income of the same year. Standard errors are clustered at the county level, and reported in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

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

| | 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 regression estimates of realtor-loan officer referral network effect on home purchase close time. The concentration is measure by within realtor $CR4$ of the loan officer mortgage shares. The dependent variables are the days between purchase contract accepted and purchase close, and the share of days above 30 (45) days. $Referral$ equals 1 if the homebuyer's realtor $CR4$ is above 0.4, and she works with a top 4 loan officer of the realtor. It is instrumented with borrower preference adjusted loan officer concentration measure $\widehat{CR4}$. The income ratio is defined as the reported income over the county median income of the same year. Robust standard errors in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table 11: Effect of Referral for Borrowers with Miminal Mortgage Denial Concerns (IV: $HHI$-$\widehat{HHI}$)

| | 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$>$700, DTI$\leq$36%, LTV $\leq$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 regression estimates of realtor-loan officer referral network effect on mortgage costs for high-quality borrowers. The concentration is measure by within realtor $CR4$ of the loan officer mortgage shares. The dependent variable is the mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. $Referral$ equals 1 if the homebuyer's realtor $CR4$ is above 0.4, and she works with a top 4 loan officer of the realtor. It and its interactions are instrumented with borrower preference adjusted loan officer concentration measure $CR4$-$\widehat{CR4}$ and interaction terms. The income ratio is defined as the reported income over the county median income of the same year. Standard errors are clustered at the county level, and reported in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

# Internet Appendix

# A BLP Estimation

## A.1 Model

We assume consumer $i$'s utility from choosing lender $j$ in market $t$ is given by

$$u_{ijt} = X'_{jt}\beta_1 + X'_{ijt}\beta_2 + \xi_{jt} + \varepsilon_{ijt}$$

where $X_{jt}$ is a vector of lender-market characteristics; $X_{ijt}$ is a vector of consumer-lender-market characteristics; $\xi_{jt}$ is an unobserved vertical lender characteristic. A market, $t$, is defined as a county-year pair. We assume $\varepsilon_{ijt}$ is an i.i.d. Type I Extreme Value shock, which yields the consumer choice probabilities:

$$s_{ijt}(\beta_1, \beta_2) = \frac{\exp(X'_{jt}\beta_1 + X'_{ijt}\beta_2 + \xi_{jt})}{\sum_{l=1}^{J_t} \exp(X'_{lt}\beta_1 + X'_{ilt}\beta_2 + \xi_{lt})} \tag{9}$$

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

$$s_{ijt}(\delta_{\cdot t}, \beta_2) = \frac{\exp(\delta_{jt} + X'_{ijt}\beta_2)}{\sum_{l=1}^{J_t} \exp(\delta_{lt} + X'_{ilt}\beta_2)} \tag{10}$$

We observe individuals $i = 1, \ldots, N_t$ in each market. Given $\delta_{\cdot t}$ and $\beta_2$, we approximate the model-predicted market shares of lender $j$ in market $t$ as

$$s_{jt}(\delta_{\cdot t}, \beta_2) \approx \frac{1}{N_t} \sum_{i=1}^{N_t} s_{ijt}(\delta_{\cdot t}, \beta_2) \tag{11}$$

## A.2 Estimation Procedure

We estimate $(\beta_1, \beta_2)$ in two steps. We first estimate $\beta_2$ via maximum likelihood with a nested fixed-point. At each guess of $\beta_2$, we invert out the $\delta$ terms using the contraction mapping developed by Berry et al. (1995):

$$\delta_{jt}^{(h+1)} = \delta_{jt}^{(h)} + \log\left(\frac{\tilde{s}_{jt}}{s_{jt}(\delta_{\cdot t}^{(h)}, \beta_2)}\right) \tag{12}$$

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

Given our MLE estimate $\hat{\beta}_2$, we estimate $\beta_1$ via OLS. Our estimating equation is:

$$\hat{\delta}_{jt} = X'_{jt}\beta_1 + \alpha_t + \xi_{jt} \tag{13}$$

where $\alpha_t$ is a market fixed effect; we include this because our model has no outside good, and thus $\delta_{jt}$ is only identified up to a constant in each market. $\beta_1$ is identified under the assumption that $E(\xi_{jt}|X_{jt}, \alpha_t) = 0$.

## A.3 Empirical Specification

We include the following in $X_{jt}$:

- 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 $t$)
- Dummy for whether lender $j$ is the #1-ranked lender in market $t$ (ranked by transaction volume)
- Dummy for whether lender $j$ is the #2-ranked lender in market $t$ (ranked by transaction volume)
- Dummy for whether lender $j$ is the #3-ranked lender in market $t$ (ranked by transaction volume)

We include the following in $X_{ijt}$:

- 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

## A.4 Summary Statistics

We estimate the model on a sample of mortgage refinances. 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 IA1 presents summary statistics for the estimation sample, before imputing missing values.

Table IA1: Summary Statistics for Lender Choice Model Estimation Sample

| Var | Obs | Mean | SD | Min | P5 | P50 | P95 | Max |
|---|---|---|---|---|---|---|---|---|
| #1 Lender in Market | 89,146 | 0.059 | 0.236 | 0.000 | 0.000 | 0.000 | 1.000 | 1.000 |
| #2 Lender in Market | 89,146 | 0.046 | 0.210 | 0.000 | 0.000 | 0.000 | 0.000 | 1.000 |
| #3 Lender in Market | 89,146 | 0.033 | 0.179 | 0.000 | 0.000 | 0.000 | 0.000 | 1.000 |
| Fintech | 89,146 | 0.181 | 0.385 | 0.000 | 0.000 | 0.000 | 1.000 | 1.000 |
| Bank | 88,516 | 0.395 | 0.489 | 0.000 | 0.000 | 0.000 | 1.000 | 1.000 |
| Out-of-State | 89,146 | 0.192 | 0.394 | 0.000 | 0.000 | 0.000 | 1.000 | 1.000 |
| FICO | 89,146 | 763.744 | 59.017 | 357.000 | 649.000 | 780.000 | 838.000 | 850.000 |
| Combined Loan-to-Value Ratio | 80,703 | 68.739 | 16.675 | 0.810 | 37.640 | 71.000 | 94.650 | 191.030 |
| Debt-to-Income Ratio | 77,029 | 29.660 | 12.259 | 10.000 | 10.000 | 30.000 | 48.000 | 60.000 |
| Distance \| Distance>0 | 86,372 | 25.458 | 52.820 | 0.001 | 0.805 | 7.489 | 134.776 | 668.887 |

## A.5 Estimates

Table IA2 presents the estimates of $(\beta_1, \beta_2)$.

# B 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 simulation, 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.

## B.1 Simulating lender choice

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

Table IA2: 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 |

Standard errors in parentheses. $^{*}p<0.1$, $^{**}p<0.05$, $^{***}p<0.01$

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. (Note: If a lender $j$ has no refinance transactions in market $m$, then we will not have a corresponding $\hat{\xi}_{jm}$. To these lenders, we assign the unconditional mean of the $\hat{\xi}_{jm}$.)

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.

## B.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.

## B.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.

# C Additional Tables

Table IA3: 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) |
| Market Loan Officer $CR4$ | 0.34 | (0.30) |
| Market Loan Officer $HHI$ | 914.39 | (1453.31) |

Note: This table presents the summary statistics of markets at the county*year level used in this study.

Table IA4: Effect of Referral 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 |
| | (1) | (2) | 2SLS (3) | 2SLS (4) |
| $CR4\text{-}\widehat{CR4}$ | 1.167*** | 1.123*** | | |
| | (0.007) | (0.026) | | |
| Referral | | | 0.193*** | 0.203*** |
| ($CR4 \geq 0.7$, 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 regression estimates of realtor-loan officer referral network effect on mortgage costs. The concentration is measure by within realtor $CR4$ of the loan officer mortgage shares. The mortgage costs are measured by both mortgage interest rate spreads and APR spreads, both off the benchmark rate offered on prime mortgage loans of a comparable type. $Referral$ equals 1 if the homebuyer's realtor $CR4$ is above 0.4, and she works with a top 4 loan officer of the realtor. It is instrumented with borrower preference adjusted loan officer concentration measure $CR4\text{-}\widehat{CR4}$. The income ratio is defined as the reported income over the county median income of the same year. Standard errors are clustered at the county level, and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.