

Market Power in Mortgage Pricing: the Role of Referral Lending

Dayin Zhang Panle Jia Barwick Lu Han Jon Kroah*

September 2026

Abstract

Despite intense competition among mortgage lenders, borrowers face elevated mortgage rates and substantial price dispersion. We argue that realtor–loan-officer referral networks are a key source of lender market power: by steering homebuyers toward a limited set of loan officers, they restrict borrower choice even in otherwise competitive markets. Linking 91,988 realtors to 123,749 loan officers in 40 states, we document that such networks are pervasive and highly concentrated: 79% of realtors direct over 40% of their clients to just four loan officers, despite having on average 176 loan officers in the market. Strikingly, this concentration persists and even increases in markets with more lenders, suggesting that referrals constrain choice regardless of market structure. Controlling for borrower characteristics and lender choices, IV estimates suggest borrowers using referred loan officers pay 17.7 basis points more, equivalent to \$2,712 in upfront costs at the average loan of \$333K. The premium is more than twice as large for first-time home buyers, highlighting the role of experience and financial sophistication. A second identification strategy exploits the 2023 CFPB enforcement order against Freedom Mortgage, which terminated a set of paid referral relationships. Exploiting entirely different variation, this design confirms that referral lending raises borrowing costs, with a larger premium among the paid relationships the order disrupted. Two mechanisms operate: referrals reduce borrowers' search intensity across lenders, and they insulate individual loan officers from competition within their own institution, allowing referred officers to charge higher rate spreads than their colleagues. Neither faster processing time nor reduced denial risk justifies the premium. Our findings reveal referral networks as a hidden source of market power in the lending market, imposing substantial financial costs and raising equity concerns for borrowers.

Keywords: realtors, loan officer, referral network, steering

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

*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 thank the discussants, Luis Lopez, David Zhang, Rebecca Jorgensen and Elisa Giannone, as well as seminar participants at WFA, ASSA-AREUEA Conference, Mortgage Market Research Conference at Fed Philadelphia, UEA (North American Meeting), the Singapore Management University, University of Wisconsin-Madison, University of Hawaii at Mānoa and Freddie Mac, for their helpful comments.

1 Introduction

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

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

This paper provides the first national, systematic evidence on realtor–loan-officer referral networks in the mortgage market. We assemble a novel dataset linking 91,988 realtors to 123,749 loan officers across 703 counties in 40 states from 2018 to 2024. 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, 79% of realtors exhibit substantial concentration in their loan officer networks, directing over 40% of their clients to just 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

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

by the market structure of lenders but are heavily constrained by realtor referral practices.

Whether such concentration translates into higher borrowing costs is an open question. We find that it does, substantially. Borrowers working with realtors with top-quintile concentration (in terms of *CR4*) pay 30.8 basis points more than those with bottom-quintile realtors, with rate differentials increasing monotonically across concentration quintiles. Even after controlling for market-specific trends, borrower characteristics (age, income, FICO), loan characteristics (loan-to-value ratio, debt-to-income ratio, etc.), and house attributes (size, age, ZIP code, etc.), the differential remains economically and statistically significant at 13.2 basis points. The contrast with market-level concentration, which shows no relationship with mortgage rates, suggests that market competition fails to discipline pricing because referral networks keep borrowers from shopping across the lenders available to them.

This evidence alone does not establish a causal effect. The referral indicator may be correlated with unobserved determinants of mortgage rates on both sides of the market. On the demand side, riskier or less sophisticated borrowers may both follow referrals more often and pay higher rates for reasons unrelated to referrals, biasing estimates upward. Our rich borrower controls, including FICO, DTI, LTV, income, and detailed loan and property characteristics, substantially mitigate this concern. On the supply side, high-quality or high-volume loan officers may offer lower rates while dominating realtors' networks even absent steering, which biases OLS estimates downward. Consistent with this, controlling for superstar loan officers raises the estimated referral premium.

To address any remaining endogeneity, we instrument referral lending using the *gap* between observed loan-officer concentration and the concentration predicted in the absence of realtor referrals. This “excess concentration” isolates the portion of realtor-level loan-officer concentration that cannot be explained by borrower preferences, lender characteristics, geographic proximity, or other forces that would shape lender choice absent contemporaneous realtor involvement. Under the assumption that, conditional on our extensive set of controls, this excess concentration affects mortgage pricing only through referral intensity, the IV recovers the effect of using a referred loan officer.

We construct the instrument in two steps. We first estimate a discrete choice demand model (Berry et al., 2004, 1995) of borrowers' lender choices using refinance mortgages, where realtors are largely absent.² 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

²In our sample, 24% of refinance borrowers return to the loan officer they used at purchase. We exclude them to recover lender choices without realtor influence (Section 4.2).

arise if their clients chose lenders based solely on borrower preferences and lender attributes. Our instrument is the difference between observed and counterfactual concentration, capturing the excess concentration in purchase transactions beyond what would arise absent contemporaneous realtor referral.³

The IV estimates reveal substantial borrowing costs. Homebuyers who use referred loan officers pay 17.7 basis points higher mortgage rates. At the average loan size of \$333k, this translates to \$589 in additional annual interest payments, or equivalently, \$2,712 in upfront points needed to buy down the rate to non-referral levels. To gauge magnitude, the average rate spread is 51 basis points, and the residual standard deviation after controlling for market-month fixed effects and borrower characteristics is 37 basis points. Our estimates imply that referral lending raises rate spreads by roughly 35%, and appears to be a first-order contributor to cross-sectional variation in purchase mortgage prices.

To cross-validate this estimate of the referral premium, we turn to a second, independent identification strategy that relies on neither the refinance sample nor the demand model. In August 2023, the Consumer Financial Protection Bureau (CFPB) issued a consent order against Freedom Mortgage for violating the anti-kickback provisions of Section 8 of the Real Estate Settlement Procedures Act (RESPA), fining the lender for compensating real estate brokers and agents through marketing-services payments, free subscription services, and gifts in exchange for mortgage referrals. The order is direct evidence for the mechanism this paper aims to uncover: referral relationships generate overcharges and are ultimately paid for by these higher rates. It also abruptly terminated a set of paid relationships for reasons exogenous to the choice of borrowers or realtors.

Comparing realtors with pre-order referral ties to Freedom-employed loan officers against realtors in the same markets who did not operate a referral network, we find that referral routing breaks sharply and precisely at the order date, and that the mortgage rates paid by treated realtors' clients fall significantly over the following year and a half. Used as an instrument, the enforcement order has a strong first stage and puts the cost of referral lending at 40.5 basis points with realtor fixed effects and the full set of controls. The estimate is larger than under the first design, as one would expect from a design identified from relationships lucrative enough that a lender paid to maintain them, where the incentive to steer is strongest. That two designs resting on different identifying assumptions both find a large and precisely signed premium supports a causal interpretation of the referral effect.

³Our IV is close to a control function approach in spirit, see [Wooldridge \(2015\)](#).

A natural question is whether referrals reduce frictions in the mortgage application process, so that referral lending is partly efficiency-enhancing. We test this directly. Conditional on a signed purchase contract, borrowers who use referred loan officers close 1.4 days sooner, against a baseline of 39 days, under the excess-concentration design, and no sooner under the enforcement-order design. Even at the upper bound, a 3.7 percent reduction in processing time is a small benefit relative to its cost. The 1.4 days saved come with \$2,712 in higher upfront borrowing costs, roughly \$1,900 per day, a steep price for a marginal convenience.

We next test whether the referral premium is associated with lower approval uncertainty. If referred loan officers help marginal borrowers secure financing that they might otherwise be denied, higher interest rates could reflect compensation for lower denial risk. Because our data do not record denied mortgage applications, we take an indirect approach and focus on high-quality borrowers with minimal approval uncertainty. Among borrowers with FICO scores above 780, referred loans remain 12 basis points more expensive. Among borrowers with low debt-to-income and loan-to-value ratios and above-median income, the premium is at least 11 basis points. The persistence of the referral premium among borrowers facing virtually no risk of denial suggests that reduced approval uncertainty is unlikely to be the primary explanation.

With neither faster processing nor lower denial risk accounting for the premium, we turn to its sources and present two pieces of evidence on the underlying mechanisms. First, using credit-bureau inquiries to measure mortgage search, we find that borrowers who follow their realtor's referral contact 9 percent fewer lenders and are 19 percent less likely to obtain multiple quotes. This reduction in search intensity likely compresses borrowers' choice sets and weakens their outside options, shifting bargaining power toward lenders during rate negotiation.

Second, referral-induced market power operates not only across lenders but also within lending institutions. Conditioning on lender fixed effects, referred borrowers pay about 6 basis points more than comparable non-referred borrowers obtaining loans from the same lending institution. This within-lender premium indicates that referrals confer pricing power on individual loan officers by insulating their clients from competition with other loan officers within the same firm. Decomposing the overall referral premium, approximately two-thirds reflects steering borrowers toward higher-cost lenders (a between-lender channel), while the remaining third arises from additional markups charged by referred loan officers relative to their colleagues (a within-lender channel).

Finally, the burden of referral costs falls unevenly across borrowers. The sharpest disparity is by

borrowing experience and sophistication. First-time home buyers pay a referral premium of 27.8 basis points, roughly \$4,258 in upfront costs, against 12.7 basis points for repeat buyers. This result implies that realtor–loan-officer referral networks not only raise borrowing costs on average, but also shift the burden toward inexperienced households, exacerbating distributional and fairness concerns.

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

We show that this view misses a key mechanism: realtor referrals narrow borrowers’ effective choice set to a handful of loan officers even where many lenders compete, so market power arises in the application process rather than in market structure. [Verboven and Yontcheva \(2024\)](#) make a related point for notaries in real estate transactions, whose rents depend on who controls access to the transaction rather than on the number of nearby competitors. Our findings complement [Allen et al. \(2014, 2019\)](#), who attribute lender market power in the more concentrated Canadian market to search frictions and brand loyalty. Relative to [Agarwal et al. \(2024\)](#), who trace dispersion to the interaction of search and application approval on the borrower side, we identify a distinct supply-side channel, realtor-mediated referral networks, which operates even among borrowers who face no approval risk.

Second, our paper also contributes to the literature on referral networks. Referral mechanisms have been studied in labor markets, health care, and education, but evidence outside these domains remains scarce.⁴ The limited attention is due in part to the informal nature of referral ties and the difficulty of observing them systematically. To our knowledge, we provide the first systematic evidence on referral networks that link realtors to mortgage lenders, an important but largely overlooked

⁴See [Burks et al. \(2015\)](#); [Pallais and Sands \(2016\)](#); [Chen-Zion and Rauch \(2020\)](#); [Barwick et al. \(2023\)](#) for studies in labor market, [Ho and Pakes \(2014\)](#); [O’Malley et al. \(2021\)](#); [Sarsons \(2024\)](#) for health care, and [Card and Giuliano \(2016\)](#); [Cestau et al. \(2017\)](#) for education.

dimension of household financial decision-making.

Finally, our paper advances the literature on intermediation and brokerage in both housing and mortgage markets.⁵ Most of the existing literature, however, studies realtors and brokers separately. The exception is [Jorgensen \(2024\)](#), who studies over 100 mergers between real estate agencies and mortgage lenders and finds that borrowers using the jointly owned lender pay roughly 9 bps more in interest. Our paper differs in both scope and mechanism. Whereas [Jorgensen \(2024\)](#) studies legally codified joint ownership, we focus on informal referral relationships at the individual agent–loan-officer level, the level at which referrals are actually made. This finer resolution allows us to capture the economically important relationships that operate outside formal ownership structures and to document larger and more pervasive referral effects. We further examine distributional implications, investigate the underlying mechanisms, and evaluate non-financial borrower outcomes.

The remainder of the paper proceeds as follows. Section 2 describes the institutional setting and data. Section 3 documents the structure of realtor-loan-officer referral networks and their association with mortgage rates. Section 4 estimates the referral premium using OLS, the excess-concentration instrument, and the CFPB enforcement order. Section 5 asks what the premium pays for: it rules out faster closings and greater approval certainty, traces the premium to reduced search across lenders and pricing power within them, and shows that it falls hardest on the borrowers least able to evaluate a mortgage offer. Section 6 concludes.

2 Institutional Background and Data

2.1 Institutional Background

Realtors play a central role in U.S. residential real estate transactions. Beyond helping buyers find and negotiate for a home, they routinely guide clients through the rest of the transaction, recommending home inspectors, title companies, and, notably, mortgage loan officers. According to a survey by [FreddieMac \(2016\)](#), 84% of real estate professionals maintain a preferred network of lenders to whom they refer clients; among them, 73% work with one to three lenders and another 24% with four to six. These referrals are effective: 76% of agents report that their clients “always” or “often” follow

⁵For studies on the housing market, see [Han and Hong \(2016, 2011\)](#); [Han and Strange \(2014\)](#); [Barwick et al. \(2017\)](#). For studies on mortgage markets, see [Ambrose and Conklin \(2014\)](#); [LaCour-Little \(2009\)](#); [Ernst et al. \(2008\)](#), [Spader and Quercia \(2011\)](#); [Berndt et al. \(2010\)](#), [Woodward and Hall \(2010, 2012\)](#), [Robles-Garcia \(2020\)](#), and [Agarwal et al. \(2021\)](#).

their lender recommendations, a share that rises to 87% among realtors who complete more than 20 transactions a year.

The importance of realtors in homebuyers' choice of mortgage lenders is also reflected by the homebuyers' survey. In the 2018–2022 National Survey of Mortgage Originations (NSMO), administered by [FHFA and CFPB](#), 54% of respondents (5,578 of 10,343) cited a “recommendation from a real estate agent/home builder” as important in choosing their lender, ahead of every other factor, including an established banking relationship, a local office, or recommendations from friends, relatives, or co-workers.

Mortgage loan officers intermediate between homebuyers and lenders: they take applications, explain products, collect documentation, coordinate with underwriters, and manage the transaction through closing, while the lender formally issues and funds the loan. In-house loan officers, employed by banks or non-bank mortgage companies, offer only their employer's products; independent loan officers work with multiple lenders. The independent channel contracted sharply after the Global Financial Crisis, from an estimated 68% of originations at its 2004 peak to about 9% by 2014,⁶ and more than 90% of loan officers in our sample are in-house. Within a lender's pricing guidelines, loan officers exercise substantial discretion over the rate quoted, points charged, and speed of processing, which creates scope for pricing to be shaped by ongoing relationships with specific realtors.

Those relationships are often reciprocal: realtors refer clients to loan officers, and loan officers direct pre-approved borrowers to trusted realtors. Such ties can improve coordination and reduce transaction frictions, but they can also facilitate steering, in which realtors guide homebuyers toward a narrow set of preferred loan officers, limiting borrower choice and weakening price competition. Although Section 8 of the Real Estate Settlement Procedures Act (RESPA) prohibits kickbacks and unearned fees in exchange for mortgage referrals, common practices occupy a gray area: loan officers may give realtors gifts not explicitly tied to referrals, sponsor realtor events, and fund joint advertising, all of which are difficult to monitor and can deepen referral relationships without direct compensation. These practices are not hypothetical. In 2023 the CFPB found that Freedom Mortgage, a large nonbank lender, had provided realtors with marketing payments, free subscription services, event sponsorships, and gifts in exchange for referrals, and ordered it to stop.⁷ Section 4.3 exploits this

⁶See a 2007 Reuters article, “[Mortgage brokers cashed in on US housing bounty](#),” and 2015 [National Mortgage Professional](#).

⁷Consumer Financial Protection Bureau, *In the Matter of Freedom Mortgage Corporation*, consent order, August 17, 2023.

order as a natural experiment.

2.2 Data

CoreLogic Multiple Listings, Ownership Transfer, and Mortgage Data We link the three [CoreLogic](#) data products (data on Mortgage, Ownership Transfer, and Multiple Listing Service (MLS) transactions) to create a panel dataset of mortgage-financed home purchases and the associated buyer agents. The sample covers listings dated between July 1, 2017, and December 31, 2024. Our choice of time window is guided by the availability of detailed loan information in the HMDA data, especially interest rates and mortgage fees, which became more comprehensive beginning in January 2018 (as described below). Because mortgage origination typically occurs near the end of the home purchase process, we include housing listings since mid-2017 to ensure adequate coverage of transactions that culminate in mortgage activity in early 2018.

We restrict attention to counties meeting three criteria: over 80% of listings originate from a single MLS, avoiding complications from cross-listing; that MLS provides unique realtor identifiers;⁸ and it reports contract dates, which we need to measure days from contract to closing, for at least 50% of listings.⁹ The resulting sample comprises 703 counties in 40 states, covers most large population centers, and accounts for 42.7% of the U.S. population (Figure 1).

Within these counties, we exclude rental and multi-structure listings, group sequential listings of the same property within 90 days into a single listing event, retain only listings that result in a mortgage-financed sale (76% of purchases), since buyer realtors and loan officers are not observed otherwise, and drop transactions lacking a valid realtor identifier or the loan officer’s Nationwide Multistate Licensing System (NMLS) ID. This yields 6.7 million mortgage-financed purchases originated between 2018 and 2024. Finally, to guard against spurious concentration from small transaction counts, we drop realtors with fewer than ten purchase loans over the sample period, leaving a network sample of 5.2 million observations on which all realtor–loan-officer network measures are constructed. Appendix Figure A1 traces the observations lost at each step. The ten-loan threshold is deliberately conservative; Appendix C.1 shows that results are very similar when realtors with at least

⁸Merging datasets using agent names is prone to inconsistencies and ambiguities due to name variations. We require the dominant MLS reports buyer agent identifiers for at least 10% of listings.

⁹The contract date is the date on which the buyer and seller sign the purchase agreement. It marks the start of the window, typically 60 to 90 days, in which the buyer arranges mortgage financing and completes closing.

five loans are retained. The average market, defined at the county-year level, has 175 active realtors, 59 lending institutions, and 176 loan officers (Table [A1](#)).

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

To that end, we employ two standard measures of concentration: the four-firm concentration ratio ($CR4$) and the Herfindahl-Hirschman Index (HHI). For each realtor, we define the loan officer network as the set of all loan officers who financed mortgages for at least one of the realtor’s buyer clients. We then calculate each loan officer’s share within that network: the fraction of the realtor’s homebuyer clients whose loans were underwritten by that officer. The $CR4$ is the sum of the shares of the top four loan officers for a given realtor, while the HHI is the sum of squared shares across all loan officers in the realtor’s network. Following [Dranove et al. \(2017\)](#), we classify realtors with $CR4 > 0.4$ as medium concentration and those with $CR4 > 0.7$ as high concentration. Similarly, following the 2023 DOJ–FTC Merger Guidelines, we use $HHI > 1000$ and $HHI > 1800$ as cutoffs for medium and high concentration. High values of $CR4$ or HHI provide suggestive evidence of referral networks. We incorporate both measures in our regression analyses below.

We measure concentration at the level of individual loan officers rather than lending institutions because referrals are interpersonal: a large lender may employ hundreds of officers, yet a realtor typically refers to one to three of them. With 176 active loan officers in a typical county-year, it is difficult to see how four officers could account for 40 percent or more of a realtor’s purchase loans absent a deliberate referral relationship.

HMDA Data We obtain loan terms from the Home Mortgage Disclosure Act (HMDA, [Federal Financial Institutions Examination Council](#)) data, whose 2018 expansion added interest rates, origination fees, and borrower- and loan-level characteristics including loan-to-value (LTV) and debt-to-income (DTI) ratios. Following standard practice (e.g., [Kermani and Wong, 2026](#)), we merge HMDA

originations to CoreLogic mortgages on lender name, loan amount, LTV, property census tract, and origination year, retaining only one-to-one matches. This matches 3.1 million of the 5.2 million purchases in the network sample. For the estimation sample we further retain first-lien, 30-year fixed-rate mortgages with balances of at least \$100,000 and non-missing LTV, DTI, borrower age, income, and joint-application status, and drop realtors with fewer than ten matched purchases, so that each realtor’s counterfactual concentration can be simulated for the IV. The estimation sample includes 1,751,287 transactions intermediated by 91,988 realtors, and all baseline regressions are estimated on it.¹⁰

Credit Bureau Data We match the estimation sample to a 1% random sample from one of the three major U.S. credit bureaus ([Credit Bureau, 2016–2024](#)), which provides annual year-end credit snapshots and the underlying tradelines for each individual, including credit scores, which we use as controls for creditworthiness, and credit inquiries, which we use to proxy for mortgage search. We match on mortgage origination date, loan amount (rounded to the nearest \$100), and borrower ZIP code; because multiple mortgages in the same ZIP code on the same day rarely share the same loan amount, ambiguous matches are minimal. The matched sample contains 15,390 observations.

2.3 Attributes of Mortgage Loans and Homebuyers

Table 1 reports summary statistics for the full estimation sample (column (1)) and the 1% subsample matched to credit records (column (2)).

The estimation sample comprises 1,751,287 home purchases with an average loan amount of \$304K. Following [Bhutta and Hizmo \(2021\)](#), we measure borrowing costs using the “Annual Percentage Rate (APR) spread” reported in HMDA. APR incorporates both the contract interest rate and upfront finance charges associated with origination, thus providing a more comprehensive measure of borrowing costs than the contract interest rate alone. The APR spread is defined as the difference between a mortgage’s APR and the prime rate for a comparable loan as of the date the interest rate is set. This benchmarking helps address the fact that we do not observe the lock date for individual mortgages, which typically occurs 30–60 days before mortgage origination. Appendix B shows that the referral premium we estimate is almost entirely a contract rate effect. The average APR spread

¹⁰We use the network sample of 5,167,324 home purchases to construct the realtor–loan-officer concentration measures, because its broader coverage measures each realtor’s loan-officer shares more accurately. We use the smaller sample of 1,751,287 purchases for all regressions, for which we observe HMDA pricing and the full set of borrower and loan controls. Appendix Figure A1 traces the losses at each step.

is 51 basis points, joint applications account for 44% of loans, 97% fall within the conforming limit, and the average LTV and DTI ratios are 88% and 36%.

The 1% subsample is similar, with an average loan of \$333K and an average spread of 49 basis points; its higher share of joint applications arises mechanically because each co-borrower appears separately in the credit files.¹¹ The credit records add FICO scores (mean 743) and mortgage-related credit inquiries in the 90 days before closing (mean 1.45), which we use to proxy for search intensity. Because FICO is essential to mortgage pricing, we rely on this subsample when quantifying borrowing costs, though full sample results are similar.

Homebuyers in the subsample earn 1.66 times their county’s median income (average \$73.9K; [U.S. Census Bureau, 2010–2024](#)); 28% are first-time buyers; 28% have a high-school education or less, 31% some college, and 38% a college degree; and 16% are Hispanic, while by race 84% are White, 7% Black, 7% Asian, and 1% other.¹² The average home costs \$398K, has 2,094 square feet, 3.4 bedrooms, 2.6 bathrooms, and is 41 years old, and the average time from purchase agreement to closing is 39 days. These characteristics are close to their full-sample counterparts, supporting the representativeness of the matched sample.

3 Referral Networks: Descriptive Evidence

We now turn to the referral networks themselves. This section documents three facts that motivate the rest of the paper: realtors’ loan-officer networks are highly concentrated (Section 3.1); this concentration bears no resemblance to the market structure of the counties in which realtors operate, and if anything rises with the number of competing loan officers (Section 3.2); and it predicts the rates borrowers pay, whereas market-level concentration does not (Section 3.3).

¹¹For jointly held mortgages, each co-borrower appears separately in the credit files but shares the same mortgage tradeline, roughly doubling the probability that such loans are matched.

¹²Race and ethnicity are self-reported in HMDA, with 13% and 12% of borrowers, respectively, leaving them blank. We fill non-responses using Bayesian Improved Surname Geocoding (BISG; [Elliott et al., 2009](#)), which combines the borrower’s surname with the racial composition of the property’s ZIP code, using the surname-level race probabilities from the 2010 Census surname list ([U.S. Census Bureau, 2016](#)) and ZIP-code racial composition from the 2010 Census ([U.S. Census Bureau, 2011](#)), leaving self-reported values untouched; this reduces missing values to below 2% for both.

3.1 Realtor-Level Loan-Officer Concentration

Figure 2 plots histograms of realtor-level concentration across all realtors in our sample. Panel A shows that 20% of realtors exceed the high-concentration benchmark ($CR4 > 0.7$) and 79% exceed the medium threshold ($CR4 > 0.4$), in line with the Freddie Mac survey finding that 84% of real estate professionals maintain a preferred network of lenders (FreddieMac, 2016). Panel B shows the same pattern for the HHI : 33% of realtors exceed 1800 and 64% exceed 1000.

This concentration does not arise mechanically from small transaction counts or a thin supply of loan officers. The median realtor completes 20 home purchases over our sample period, and a typical market (county-year) has 176 loan officers at 59 active lending institutions.¹³ We next show that it also does not reflect local market structure.

3.2 Realtor-level vs. County-level Lender Concentration

If realtor-level concentration reflected local market structure rather than referrals, it should track the concentration of the local lending market. Figure 3 shows that it does not.

Panel A plots the distributions of the two measures: the blue curve shows the combined market share of the top four loan officers in each county, while the red curve shows the combined share of the top four loan officers within each realtor's transactions. Concentration is substantially higher at the realtor level, averaging 55 percent, compared with 18 percent at the county level. This pronounced gap suggests that realtor-loan-officer relationships are unlikely to be random and instead reflect systematic, concentrated referral patterns.

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

The pattern at the realtor level, however, is strikingly different. Realtor-level $CR4$ (red squares) remains above 0.4 across markets of all sizes and, if anything, increases with the number of loan

¹³Section 4.2 provides a direct benchmark: simulating each realtor's loan-officer choices from a demand model estimated where realtors are absent, only 11% of realtors would exceed the medium threshold absent referrals, against 64% in the data (Figure 5).

officers in the market. Thus, even as the set of available loan officers expands, borrowers connected to a given realtor continue to be concentrated among a small subset of officers within that realtor’s network. One possible interpretation is that greater competition among loan officers strengthens the incentive to cultivate realtor referral relationships, making those networks more rather than less important in determining borrower allocation. Whether this concentration matters for what borrowers pay is the question we turn to next.

3.3 Concentration and Mortgage Rates

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

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

$$Y_{i;c(i)y-m(i)} = \sum_{q=2}^5 \tilde{\alpha}_q \text{Quintile}_q(CR4_{cy}^{\text{lender}}) + X_i' \tilde{\gamma} + \mu_c + \mu_{y-m} + \tilde{\epsilon}_i, \quad (1)$$

where $Y_{i;c(i)y-m(i)}$ measures the APR spread of the purchase mortgage obtained by borrower i in county $c(i)$ and year-month $y-m(i)$ (see Section 2.3). The key regressors $\text{Quintile}_q(CR4_{cy}^{\text{lender}})$ are indicators equal to one if the lender concentration in county c and year y falls into the q th quintile of the distribution, with the bottom quintile omitted as the reference group. We include county fixed effects μ_c and year-month fixed effects μ_{y-m} to absorb monetary policy and other aggregate shocks.¹⁴ We also include an expansive set of controls X_i that includes borrower demographics (income-ratio percentile bins, six 10-year age bins, and, in additional analyses, eight 40-point FICO score bins), loan charac-

¹⁴Because $CR4_{cy}^{\text{lender}}$ is defined at the market (county-year) level and exhibits no variation within a county-year-month cell, we include county and year-month fixed effects separately (rather than county $\times$ year-month fixed effects).

teristics (loan amount in \$100K 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, bedrooms, bathrooms, and eleven square-footage bins).

Table 2 reports the estimates. Borrowers in more concentrated lender markets do not pay more. In the full-sample (Panel A), several coefficients are negative: higher lender concentration is associated with lower mortgage costs. Panel B uses the 1% random sample matched to credit records and adds FICO controls, and shows the same slightly negative pattern, of 9 to 12 basis points, with no monotone gradient in concentration.¹⁵ These results match existing studies (Buchak and Jorring, 2021; Amel et al., 2018), which find limited effects of market-level lender concentration on mortgage pricing. Panel A of Figure 4 plots the Panel B estimates with and without FICO controls (columns (5) and (6)); no systematic pattern emerges.

Realtor-level Concentration We next examine whether the concentration of a realtor’s loan-officer referral network helps explain mortgage pricing. Specifically, we sort realtors into quintiles based on the $CR4$ of their loan-officer network and estimate:

$$Y_{i;c(i)y-m(i)r(i)l(i)} = \sum_{q=2}^5 \alpha_q \text{Quintile}_q(CR4_r) + X_i' \gamma + \mu_{c,y-m} + \varepsilon_i, \quad (2)$$

where $Y_{i;c(i)y-m(i)r(i)l(i)}$ is the APR spread on the mortgage of borrower i , originated by loan officer l , and represented by realtor r , in county $c(i)$ and year-month $y-m(i)$, the controls X_i are as in Equation (1), and county $\times$ year-month fixed effects $\mu_{c,y-m}$ replace the separate county and year-month effects.

Table 3 reports the estimates. In the full sample (Panel A), borrowers represented by top-quintile realtors pay 25.9 bps more than those represented by bottom-quintile realtors, against a mean spread of 51 bps, with rate premia rising monotonically across quintiles. Adding controls sequentially (columns (2)–(5)) lowers the gap to 11.7 bps under the full set of market $\times$ month, borrower, loan, and house controls, so observables explain a little over half of the raw differential.

Since public HMDA data do not report FICO scores, Panel B repeats the exercise on the 1% sample matched to credit records. Without FICO controls (columns (1)–(5)) the estimates are very close

¹⁵The negative sign likely reflects the empirical pattern documented in Figure 3: markets with more lenders tend to have *higher* average loan-officer concentration within realtors. Market-level and realtor-level concentration are therefore negatively correlated, so a market-level measure partly proxies for the absence of the referral channel that drives our results.

to Panel A. The standard errors are larger due to the smaller sample size. With FICO bin fixed effects in column (6), borrowers represented by top-quintile realtors pay 13.2 bps more than those represented by bottom-quintile realtors, slightly lower than but comparable to the corresponding estimates without FICO controls. Panel B of Figure 4 plots these estimates: spreads rise monotonically with realtor-level concentration, in contrast to the null pattern for market-level concentration in Panel A.

4 Effect of Referral Network on Borrowing Costs

Section 3 established that borrowers of high-concentration realtors pay more. This section asks whether the referral itself causes the difference. Section 4.1 defines our proxy for a referred loan and presents OLS estimates. Section 4.2 develops an instrument based on the concentration of a realtor’s network in excess of what borrower preferences and lender characteristics would generate. Section 4.3 pursues a second, independent identification strategy built on a regulatory enforcement order that abruptly terminated a specific set of referral relationships.

4.1 Effect on Borrowing Costs: OLS

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

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

Two pieces of evidence support our *Referral* definition. First, 79% of realtors have loan-officer concentration above the 0.4 threshold (Figure 2). This is consistent with external survey evidence from FreddieMac (2016), where 84% of real estate professionals maintain a preferred network of

lenders. Second, a similar level of prevalence is observed on the borrower side (Table 1). Under our definition, 77% of homebuyers work with realtors who have referral networks: 48% follow the realtor’s referral (choosing one of the top four loan officers), while the remaining 29% are matched with a loan officer outside the realtor’s top four, whether the referral was declined or the borrower shopped independently. Thus, nearly half of all homebuyers in our sample are effectively influenced by realtor–loan-officer referral networks. This share is both economically large and in line with NSMO survey evidence that 54% of borrowers choose their lenders based on recommendations from a real estate agent (or home builder).

The “top four” cutoff in this definition is not essential to our findings. Appendix C.2 re-estimates the headline specifications under a stricter rule that classifies a loan as a referral only when it is originated by the realtor’s single most-used, or “top one”, loan officer. The estimated referral premium remains of very similar magnitude in both the OLS and IV specifications (Appendix Tables A8, A9).

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

$$Y_{i;c(i)y-m(i)r(i)l(i)} = \alpha^{OLS} Referral_{rl} + X_i' \delta^{OLS} + \mu_{c,y-m} + \varepsilon_i \quad (3)$$

where $Y_{i;c(i)y-m(i)r(i)l(i)}$ is the mortgage APR spread (Appendix Table A4 testing alternative cost measures), and we include the same controls as in Equation (2): county $\times$ year–month fixed effects, borrower, loan, and house characteristics.

Table 4 reports OLS results. Column (1) estimates Equation (3) in the full sample and finds a coefficient of 0.107 on the Referral indicator, implying that borrowers who obtain loans through their realtors’ loan-officer network and use a top four loan officer pay about 10.7 bps higher mortgage rate spreads relative to the omitted group. Column (2) further separates borrowers who are in a concentrated realtor network but do not use one of the realtor’s top four loan officers (*Non-Compliant Referral*). The estimated premium for this group is essentially zero while the premium for the “compliant” referral group remains at 10.6 bps. The sharp gap between these coefficients suggests that higher borrowing costs are not simply a feature of being matched to a realtor who operates a concentrated network; rather, they are concentrated among borrowers who borrow through the realtor’s most-connected loan officers.

Columns (4) and (5) repeat the analysis using the 1% random subsample matched to credit-bureau records, allowing us to include FICO bin fixed effects. The referral premium remains statistically

strong and similar in magnitude: 10.1–10.8 bps. As in Column (2), the “Non-Compliant Referral” coefficient is negligible, reinforcing the interpretation that the cost difference is tightly linked to following the referral channel rather than to unobserved differences in homebuyers across realtors.

A natural concern is that this premium reflects loan-officer quality rather than referral steering: realtors may channel clients toward unusually productive, high-volume loan officers, whose pricing, not the referral itself, could drive the rate difference. Columns (3) and (6) address this directly by adding an indicator for whether the originating loan officer is a “superstar” (a loan officer in the top 5% by purchase-loan volume) together with its interaction with the referral indicator. The evidence runs opposite to the quality story. First, the referral premium *rises* once the superstar margin is held fixed. In Column (6), which adds the superstar controls to the credit-matched sample and is therefore our most demanding specification, the premium is 14.4 bps, up from 10.8 bps without them. Superstar officers tend to offer lower rates yet mechanically inflate a realtor’s measured concentration, so omitting them biases the referral estimate downward. At the 1% sample’s average loan size of \$333K, 14.4 bps amounts to \$480 in additional interest each year, or \$2,206 in upfront points to buy the rate back down.¹⁶ Second, the premium is, if anything, smaller for superstars: the Referral $\times$ Superstar interaction is negative (-7.2 bps in Column (6)), and the superstar main effect is small (4.2 bps). The cost of referrals is thus not about being matched to a more expensive star, but is concentrated among ordinary loan officers embedded in a realtor’s captive network. More broadly, this reveals a recurring pattern: the estimated referral premium grows as we progressively refine the comparison and purge supply-side noise.

4.2 IV Estimates

The OLS estimates might not be causal. The referral indicator may be correlated with unobserved determinants of mortgage rates on both sides of the market, and the two biases run in opposite directions. On the demand side, riskier or less sophisticated borrowers may both follow referrals more often and pay higher rates for unrelated reasons, biasing OLS upward. Consistent with this, Table 3 shows that sequentially adding borrower-side risk and creditworthiness controls (FICO, DTI, LTV, income, and demographics) steadily attenuates the concentration–rate gradient, as a result of correcting an upward bias.

¹⁶We assume a point-to-rate reduction factor of 4.6 as estimated in [Bhutta et al. \(2025\)](#).

On the supply side, referrals are plausibly correlated with unobserved loan-officer quality in the opposite direction: superstar loan officers dominate many realtors’ networks even absent steering yet tend to offer lower rates, while fintech lenders originate largely outside referral networks yet post higher rates, raising the non-referral baseline. Consistent with a downward supply-side bias, the OLS premium rises from 10.6 to 13.0 bps once the superstar margin is held fixed (Table 4).

To address the endogeneity that remains after a rich set of controls, we implement an instrumental-variable (IV) strategy. We aim to isolate the component of realtor-level loan officer concentration that is driven by forces outside borrowers’ own choice of lending representative. The IV is constructed as the difference between (i) the observed concentration of loan officers in a realtor’s network, and (ii) a hypothetical concentration that would arise if borrowers selected loan officers independently of 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.

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

$$Referral_{rl} = \gamma(HHI_r - \widehat{HHI}_r) + X'_i \delta^{FS} + \mu_{c,y-m} + v_i \quad (4)$$

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 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. We use the Herfindahl–Hirschman Index (HHI) to avoid a mechanical correlation between $Referral_{rl}$ and IV, though this choice is not crucial.¹⁷ By the Frisch–Waugh–Lovell theorem, this approach is analogous to the control function approach, where the predicted $\widehat{HHI}_r$ from the demand model in effect supplies a control that helps isolate loan-officer concentration driven by realtor referrals (Wooldridge, 2015).

Construction of $\widehat{HHI}_r$ The validity of the instrument hinges on whether $\widehat{HHI}_r$ captures the counterfactual concentration of loan officers. We construct it in three steps, described in detail in Appendix D: we build a *refinance*-mortgage sample, in which realtors are absent at the time of lender choice;¹⁸ esti-

¹⁷The estimated referral effects remain robust and slightly stronger when we construct IV using CR4 (Table A2).

¹⁸We drop the 24% of refinance borrowers who returned to the loan officer they used at purchase. Appendix D.6 shows that keeping these repeat refinances changes neither the demand estimates nor the downstream IV results.

mate a flexible BLP model of lender choice on that sample, capturing lending market structure, spatial proximity, and borrower demographics and risk; and simulate the lender and loan-officer choices that purchase borrowers would make under the estimated preferences.

We estimate the following lender choice model, where the utility of borrower i who secures a loan from lender j in market (county-year) cy is:

$$u_{ij,cy} = X'_{j,cy}\beta_1 + X'_{ij,cy}\beta_2 + \xi_{j,cy} + \varepsilon_{ij,cy}, \quad (5)$$

where $X_{j,cy}$ collects lender-market characteristics (indicators for bank, fintech,¹⁹ and out-of-state lenders, and for ranking among the three largest lenders in the county-year), $X_{ij,cy}$ collects borrower-lender characteristics (the distance from the property to the lender's nearest branch,²⁰ and interactions of the borrower's LTV, DTI, and FICO with the lender indicators), $\xi_{j,cy}$ is an unobserved lender-market component such as service quality, and $\varepsilon_{ij,cy}$ is a Type I Extreme Value shock.

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

Figure 5 compares the counterfactual and observed distributions of realtor-level loan officer concentration. Absent referrals, only 11% of realtors would exceed the medium-concentration threshold ($HHI > 1000$; Panel A, red dashed line), whereas in the data 64% exceed 1,000 (blue solid line), and the observed distribution extends well into the high-concentration range. Panel B shows the same pattern for $CR4$. Borrower preferences, geography, and lender market structure cannot generate a gap of this size; referral networks are the dominant force behind the concentration of loan officers within realtors' client bases.

¹⁹We use the list of fintech firms from Fuster et al. (2019), available at https://pages.stern.nyu.edu/~pschnabl/data/data_fintech.htm.

²⁰We use Nationwide Multistate Licensing System B2B Access data to link loan officers to their branches.

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

$$Y_{i;c(i)y-m(i)r(i)l(i)} = \alpha^{2SLS} \widehat{Referral}_{rl} + X_i' \delta^{2SLS} + \mu_{c,y-m} + e_i \quad (6)$$

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

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

Columns (3)–(4) report the IV estimates of the referral effect on borrowing costs. Without FICO controls, borrowers who use referred loan officers pay, on average, 15.1 bps higher mortgage rates; this premium remains sizable at 17.7 bps after controlling for FICO. At the 1% sample’s average loan size (\$333K), this implies roughly \$589 in additional interest payments per year. Equivalently, a referred borrower would need to pay about \$2,712 in upfront points to buy down the rate to the level of a non-referred borrower. Appendix B decomposes the APR spread into three components, the contract rate spread, origination fees, and discount points, and shows that the premium is almost entirely a contract rate effect.

The magnitude of the referral effect is economically large. Relative to the average APR spread of 51 bps, the referral premium represents a roughly 35% increase in borrowing costs. Moreover, relative to the residual standard deviation of 37 bps that remains after market-month fixed effects and borrower controls, it is an effect of about half a standard deviation. Put differently, referral networks appear to be a first-order contributor to cross-sectional variation in purchase mortgage prices.

4.3 A Second Identification Strategy: CFPB Enforcement Order

The excess-concentration instrument rests on the refinance-based counterfactual. This section pursues a second strategy that relies on neither the refinance sample nor the demand model: a regulatory order that abruptly ended payment for a specific set of realtor–loan-officer referral relationships.

The Enforcement Order In August 2023, the CFPB issued a consent order against Freedom Mortgage Corporation for violating Section 8 of RESPA, the anti-kickback provision discussed in Section 2.1.²¹ Freedom agreed to pay a \$1.75 million civil penalty and to stop providing things of value in exchange for referrals. The order serves two purposes for us. It documents that referral relationships exist and are paid for, and it provides a natural experiment: a set of established relationships stopped being paid for on a date set by a regulator rather than by borrowers or realtors.

Research Design We build a difference-in-differences design around the order. The treated units are realtors. Following Section 4.1, a realtor has a referral network if she has at least ten pre-order transactions and her pre-order top four loan officers account for more than 40 percent of them; those four officers are her referral loan officers. A realtor is *Freedom-tied*, and hence treated, if at least one of her referral loan officers was employed at Freedom in both 2021 and 2022, so that the relationship was active in the two years leading up to the order.²² The treatment group consists of the homebuyers these realtors represent.

The control group consists of homebuyers represented by realtors in the same markets who did not operate a referral network before the order, that is, realtors with at least ten pre-order transactions whose top four officers account for less than 40 percent. These homebuyers were not steered, so their mortgage rates trace out pricing conditions in the same local markets over the same months. We deliberately exclude realtors with referral networks at lenders other than Freedom. The enforcement action was the CFPB’s first public RESPA Section 8 action since 2017 and was widely read as a warning to the industry, so realtors with referral ties to other lenders may respond to it as well.²³

We restrict the sample to the 20 counties where Freedom’s referral business operated at scale and keep realtors, treated or control, with at least a quarter of their pre-order transactions in those counties. Every element of the design — whether a realtor has a referral network, who her referral

²¹Consumer Financial Protection Bureau, *In the Matter of Freedom Mortgage Corporation*, Administrative Proceeding File No. 2023-CFPB-0008, consent order issued August 17, 2023. The CFPB found that Freedom Mortgage provided real estate brokers and agents with (1) monthly payments under marketing services agreements, (2) free access to industry subscription services, and (3) event sponsorships, entertainment, and gifts, all in connection with mortgage referrals. Recipients of these payments made more than a thousand mortgage referrals to Freedom beginning in 2017.

²²A loan officer is Freedom-employed in a year if she originates at least one purchase loan at Freedom Mortgage (NMLS 2767) or at its wholly owned subsidiary RoundPoint Mortgage Servicing (NMLS 18188).

²³Trade coverage called it “a wakeup call” for lenders (HousingWire, October 10, 2023), state realtor associations advised members to enter marketing agreements only “with extreme caution” (Illinois REALTORS, August 2023), law firms issued client alerts anticipating further actions (Orrick, August 22, 2023), and in December 2024 the CFPB sued Rocket Homes over referral incentives paid to real estate agents.

loan officers are, whether she is Freedom-tied — is fixed using pre-order data and does not respond to changes after the order. The regression sample contains 512 treated realtors with 5,102 purchase loans and 2,311 control realtors with 49,094 loans. Treated and control realtors differ on observables: control realtors close more and larger loans, for borrowers with higher incomes and lower leverage (Appendix Table A13). To address this imbalance, we weight the control group by entropy balancing (Hainmueller, 2012), which chooses weights so that the weighted control means of loan attributes match the treated means. The main text reports weighted results (event studies and regression tables); unweighted results, which are very similar, are in Appendix F.

Event-Study Evidence Figure 6 traces the referral outcome (Panel A) and the purchase mortgage rate spread (Panel B) year by year for homebuyers represented by Freedom-tied realtors relative to control homebuyers. The referral outcome is an indicator equal to one if the homebuyer’s loan is financed by one of her realtor’s referral loan officers, the realtor’s top four officers defined using her pre-order loans; that set is fixed before the order and never redefined. Both panels are estimated with realtor and county $\times$ year-month fixed effects and the full set of borrower, loan, and house controls. The year 2023 is split at the order, and January–July 2023, the seven months immediately preceding it, is the reference period.

Panel A shows no pre-trend: before the order, the share of treated realtors’ loans financed by their referral loan officers was stable, with the coefficients for 2018 through 2022 lying between -1 and $+7$ percentage points, so routing showed no decay running into the order. From August 2023, the share of treated realtors’ loans financed by a referral officer falls by 18.4 percentage points, and by 22.1 points over 2024, against a pre-order level of 65 percent. Because control realtors have no referral relationships, this decline is identified from treated realtors’ own behavior: roughly a third of the business they had been routing to their referral officers went elsewhere, a large response to the order.

Panel B repeats the exercise for the mortgage rate spread on referral loans and shows the same timing. Before the order, the coefficients are statistically indistinguishable from zero. After the order, the spread paid by Freedom-tied realtors’ borrowers falls by 16.6 basis points relative to control borrowers over the remaining months of 2023, and by 11.6 basis points over 2024.

IV Estimates The enforcement order provides an alternative instrument for referral lending: the interaction of Freedom-tied status with an indicator for originations from August 2023 onward. Panel A of Figure 6 is then the first stage, Panel B the reduced form, and Table 6 combines them. The exclusion restriction is that, conditional on the fixed effects and controls, the order changed borrowing costs for Freedom-tied realtors’ clients only through the referral relationships it disrupted. Three features of the setting support it. The order targeted the lender’s referral arrangements, not its pricing, underwriting, or capacity to lend; it arrived as a legal event on a date no borrower or realtor controlled; and the control group consists of realtors in the same counties whose clients were not being steered, so county-month movements in mortgage pricing are differenced out.

Columns (1) and (2) of Table 6 report the first stage with realtor and county $\times$ year-month fixed effects alone and with the full set of borrower, loan, and house controls. After the order, a Freedom-tied realtor’s loan is about 25 percentage points less likely to be financed by one of her referral loan officers, with Kleibergen–Paap F statistics of 41 and 42. Columns (3) and (4) report the corresponding 2SLS estimates: referral lending raises the mortgage rate spread by 32.3 basis points without controls and 40.5 basis points with them.

These estimates exceed the 17.7 basis points of Section 4.2. The gap is consistent with the two designs recovering local average treatment effects for different populations. The CFPB design is identified from the referral relationships of realtors embedded in arrangements lucrative enough that a lender paid to maintain them, the setting in which the incentive to steer is strongest, so the premium on these relationships should exceed the average across all referral lending. What matters for our purposes is that two designs resting on largely non-overlapping assumptions, one on cross-sectional differences in realtors’ loan-officer concentration benchmarked against a demand model, the other on the timing of a legal event, deliver premia of the same sign and the same order of magnitude.

Appendix G reports a third source of variation, operating through the realtor rather than the lender side of the relationship: the exit of a locally dominant realtor, which severs the referral pipeline she maintained. We compare homebuyers within ten miles of a sharp exit by a realtor who had written at least a quarter of local transactions with their neighbors in the adjacent 10–20 mile ring. Exposure to the exit lowers the probability of a referral loan by 18 percentage points, and using this exposure as the instrument for referral lending yields a premium of 30.5 basis points. Because the design requires a single realtor to dominate a local market, it is confined to thin markets and we treat it as corroborating evidence.

Across the three designs, the referral premium ranges from 17.7 to 40.5 basis points, with the designs identified from the most valuable relationships producing the larger estimates.

5 Sources and Incidence of the Referral Premium

Sections 4.2 and 4.3 establish that referred borrowers pay more. This section asks what they are paying for, and who pays most. We first consider two benign explanations: that referral relationships deliver faster closings (Section 5.1), and that referred loan officers help marginal applicants secure financing they would otherwise be denied, so that the premium compensates for reduced approval risk (Section 5.2). Finding little support for either, we turn to the two channels through which referrals soften competition: reduced search across lenders (Section 5.3) and pricing power within them (Section 5.4). Section 5.5 then shows that the premium is largest for the borrowers least able to evaluate a mortgage offer.

5.1 Closing Speed

Mortgage transactions are not only financially costly but also time-consuming, and borrowers may value speed alongside low rates. Table 7 reports IV estimates of the effect of referrals on time to close, measured as the number of days between contract acceptance and mortgage closing, under both instruments. Columns (1) and (2) use the excess-concentration instrument. On the full sample, borrowers who use referred loan officers close about 1.4 days sooner, a 3.7 percent reduction relative to the baseline mean of 39 days; on the 1% sample matched to credit records, with credit score controls, the point estimate is similar, 1.6 days sooner, but statistically insignificant. Column (3) uses the enforcement-order instrument with the full-control specification of Table 6 and finds no speed advantage: the point estimate is positive and imprecise.

Overall, referral has little impact on days to close. The established working relationships between realtors and referred loan officers may facilitate smoother communication and faster document processing, but the gain is at most a day or two under one design and absent under the other, and in either case small relative to the higher borrowing costs documented earlier.

5.2 Approval Certainty

If referred loan officers help marginal borrowers secure financing they might otherwise be denied, higher rates could compensate for reduced denial risk. We do not observe denied applications, so we test this indirectly by asking whether borrowers who face essentially no denial risk also pay the premium. If they do, approval risk cannot be its primary explanation.

Table 8 reports the referral premium for low-risk borrowers under three definitions. Each column is a single regression on the full sample of its design in which the referral indicator enters separately for low-risk and for all other borrowers, so the two coefficients are the premium paid by each group. Columns (1) and (2) use the 1% sample matched to credit records, with FICO bin fixed effects and the excess-concentration instrument, and define low risk as a FICO score above 700 and above 780. The premium is 18.2 basis points for borrowers above 700 and 12.4 basis points for borrowers above 780, both precisely estimated and neither statistically different from the premium paid by borrowers below the threshold. Columns (3) and (4) use a definition available without a credit match: a DTI at or below 36 percent, an LTV at or below 80 percent, and income above the county median. Under the excess-concentration instrument (Column (3)), these borrowers pay a premium of 11.0 basis points, about 5 basis points less than other borrowers but still significant above zero. Column (4) repeats the exercise with the enforcement-order instrument and confirms that the premium for low-risk borrowers does not vanish under the second design.²⁴

The persistence of the referral premium among the most creditworthy borrowers rules out approval certainty as its primary driver, just as Section 5.1 ruled out faster closings. If the premium is a price for neither speed nor approval certainty, it most likely reflects a weakening of competition. The next two subsections show where that weakening occurs.

5.3 Reduced Search Across Lenders

One potential explanation for the referral premium is that referrals reshape borrowers' effective choice sets. By directing borrowers toward a small group of loan officers, referrals raise the effective cost of reaching lenders outside the network, leaving borrowers with fewer readily accessible alternatives even in otherwise competitive markets. This mechanism predicts less realized mortgage search and

²⁴The point estimate for low-risk borrowers suggests a large premium, though the first stage for the interacted instrument is weak (Kleibergen–Paap F of 3.1) in this smaller sample with a demanding specification. Without weighting the sample, the premium for low-risk borrowers is 79.6 basis points with a Kleibergen–Paap F of 9.7 (Appendix Table A14).

fewer competing quotes. We test this directly by examining borrowers' mortgage search activity. Measuring search is challenging because standard mortgage datasets record only the final choice of lender and loan officer. We address this limitation by exploiting credit inquiry information in the 1% random sample matched to credit-bureau records. When a borrower seriously shops for a mortgage, each lender she approaches must obtain authorization to pull her credit report, and these inquiries are recorded as mortgage-related credit inquiries. We use the number of such inquiries as a proxy for the intensity of mortgage search.

Table 9 reports IV estimates of the effect of referrals on mortgage-related credit inquiries. Column (1) examines the number of lenders approached within a 90-day window prior to closing. Borrowers who use a referral loan officer make, on average, 0.129 fewer inquiries, a 9% decline relative to the baseline mean of 1.45. Column (2) considers the extensive margin: whether the borrower shops with at least two lenders. Referral use reduces the probability of contacting multiple lenders by 6.6 percentage points, a 19% decline from the baseline rate of 35 percent. Columns (3)–(4) repeat the analysis using a 60-day window. The referral effects are even stronger in relative terms: inquiries fall by 0.15, and the share of homebuyers with at least two inquiries declines by 31%. These results provide direct evidence that referrals compress borrowers' effective choice sets. Referred homebuyers contact fewer lenders and are substantially less likely to obtain competing quotes, weakening their outside options and shifting bargaining power toward lenders during rate negotiation.

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 0.129 inquiries implies an interest-rate premium of roughly 2 basis points. This likely understates the full effect for two reasons. First, credit inquiries capture only late-stage, committed shopping and miss earlier information-gathering that may also be curtailed by referrals. Second, the marginal return to search is plausibly larger for borrowers facing unfavorable initial offers from referred loan officers.

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

5.4 Pricing Power within Lenders

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

Table 10 reports the results using the same IV strategy. Relative to Column (3) of Table 5, adding lender fixed effects in Column (1) reduces the estimated referral premium from 15.1 bps to 2.7 bps. Column (2) controls for lenders more flexibly by interacting lender fixed effects with county $\times$ year-month fixed effects (an extremely demanding specification), thus comparing referred and non-referred borrowers who obtain loans from the same lender in the same local market and month. The estimated within-lender premium remains very similar at 2.1 bps. Column (3) restricts attention to the 1% matched sample and further controls for borrower credit scores. The estimated within-lender premium is 6.0 bps, about one-third of the full premium of 17.7 bps without lender fixed effects.²⁵

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

5.5 Incidence of the Referral Premium

The premium documented so far is an average. If referrals raise costs by substituting the realtor's recommendation for the borrower's own search, as Sections 5.3 and 5.4 indicate, the premium should be largest for borrowers least equipped to evaluate a mortgage offer on their own. We test that prediction across borrower experience, education, and ethnicity.

Table 11 examines how the referral effect varies across homebuyers. All columns are estimated on the 1% random sample matched to credit records, which covers more demographic information. In each column the referral indicator is fully interacted with one mutually exclusive and exhaustive

²⁵In this subsample, we cannot include lender $\times$ year-month $\times$ county fixed effects because the interacted cells become too sparse. The stability of the estimate relative to Columns (1)–(2) suggests that the within-lender referral premium is not driven by differences in borrower credit quality.

set of groups, so the outcome of interest is the referral premium for that group relative to non-referred borrowers in the same group.

Column (1) splits homebuyers by whether they are buying for the first time. We classify a borrower as a first-time home buyer if the credit file shows no first-lien mortgage opened before the purchase and no homeowner indicator in any earlier demographic snapshot; 28% of the sample qualifies, close to the national share. The contrast is sharp. First-time home buyers pay a referral premium of 27.8 bps, translating into roughly \$4,258 in additional upfront costs, against 12.7 bps for repeat buyers, and the difference is significant at the 1% level. The reading is intuitive: a borrower who has never held a mortgage has no prior lender relationship and no benchmark for what a rate should be, so the realtor's recommendation carries more weight and is priced accordingly.

Column (2) turns to education, which is a noisy proxy for financial sophistication. The premium falls monotonically with schooling, from 19.1 bps for borrowers with a high school education or less, to 15.2 bps for those with some college, to 14.2 bps for college graduates. The ordering matches the first-time home buyer result, in that borrowers with less capacity to evaluate a mortgage offer pay more, but the differences are not statistically distinguishable from one another.

Columns (3) and (4) group borrowers by ethnicity and by race. Hispanic homebuyers pay 21.8 bps against 15.4 bps for non-Hispanic homebuyers, a gap consistent with language barriers, differential access to information, and greater reliance on realtors as trusted intermediaries, especially when intermediaries share cultural or local ties (e.g., [Edin et al., 2003](#); [Matvos et al., 2023](#)). The race split in column (4) is estimated far less precisely: instrumenting a separate interaction for each racial group leaves little identifying variation in a sample of this size, the first-stage statistic falls dramatically, and none of the minority coefficients is distinguishable from the White estimate. We therefore do not draw conclusions from the race column here. Appendix Table [A3](#) reports the ethnicity and race splits on the full estimation sample, where all coefficients are precisely estimated. It illustrates much higher premia paid by Hispanic and Black borrowers.

Referral networks thus raise borrowing costs on average and shift the burden toward the households with the least experience and the fewest alternatives, which is what a mechanism operating through displaced search implies.

6 Conclusion

This paper shows that the U.S. mortgage market is shaped not only by competition among lenders, but also by a quieter layer of intermediation: realtor–loan-officer referral networks. We document that these networks are pervasive, highly concentrated, and strongly predictive of which loan officers homebuyers ultimately use. Referral lending raises borrowing costs, even in markets that appear competitive on the lender side. Two independent identification strategies, an excess-concentration instrument benchmarked against lender choices estimated from refinance loans, and the 2023 CFPB enforcement order against Freedom Mortgage, which severed a set of paid referral relationships, point to the same conclusion despite resting on largely non-overlapping assumptions. A third, smaller-scale design based on the exit of locally dominant realtors points the same way (Appendix G). Referral lending raises rate spreads by roughly 35%, an effect equal to about half the residual standard deviation of spreads after market-month fixed effects and borrower controls, making referral networks a first-order contributor to cross-sectional variation in purchase mortgage prices. The referral premium is significantly larger for first-time home buyers, a group that relies most on the realtor’s recommendation and is least able to evaluate it.

In terms of the underlying mechanism, referrals depress borrowers’ search intensity across lenders and shield referred officers from internal competition within a lending institution. Referrals shorten closings by at most a day or two under one design and not at all under the other, so any convenience gain is small relative to the additional financial cost, and the premium persists even among the most creditworthy borrowers, ruling out denial risk mitigation as the primary explanation.

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

References

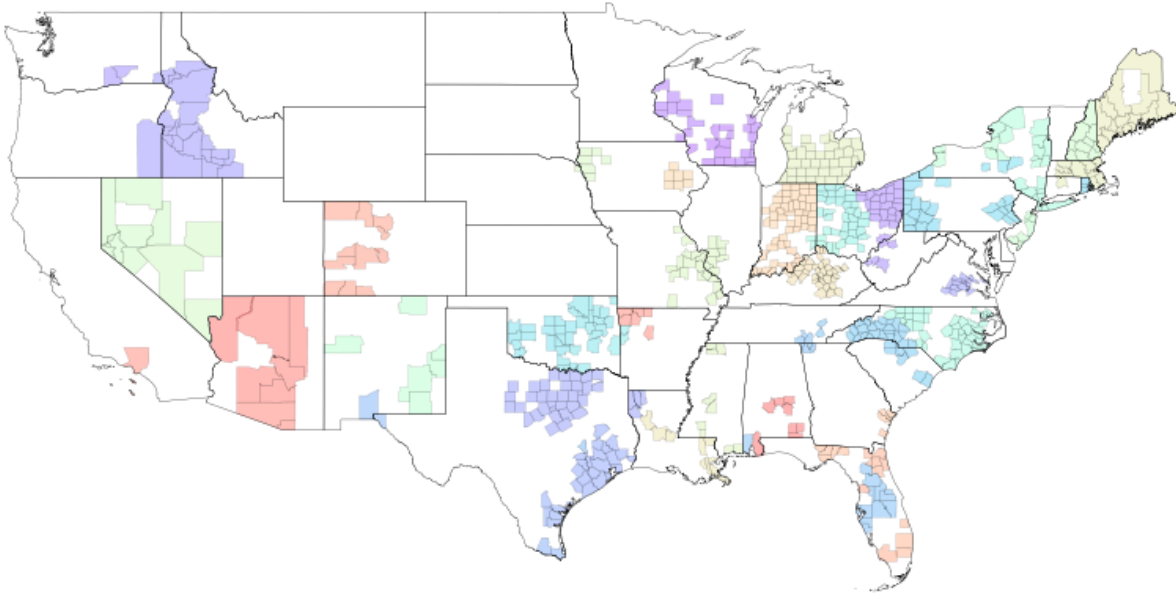
- Agarwal, Sumit, Swee Hoon Ang, Yongheng Deng, and Yonglin Wang, 2021, Mortgage Brokers and the Effectiveness of Regulatory Oversight, *Management Science* 67, 5278–5300.
- Agarwal, Sumit, John Grigsby, Ali Hortaçsu, Gregor Matvos, Amit Seru, and Vincent Yao, 2024, Searching for Approval, *Econometrica* 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.
- Berry, Steven, James Levinsohn, and Ariel Pakes, 2004, Differentiated products demand systems from a combination of micro and macro data: The new car market, *Journal of Political Economy* 112, 68–105.
- Bhutta, Neil, Andreas Fuster, and Aurel Hizmo, 2025, Paying too much? borrower sophistication and overpayment in the u.s. mortgage market, *Journal of Finance*.
- Bhutta, Neil, and Aurel Hizmo, 2021, Do Minorities Pay More for Mortgages?, *The Review of Financial Studies* 34, 763–789.
- Bresnahan, Timothy F, and Peter C Reiss, 1991, Entry and competition in concentrated markets, *Journal of political economy* 99, 977–1009.
- Buchak, Greg, and Adam Jorring, 2021, Competition with Multi-Dimensional Pricing: Evidence from U.S. Mortgages, *SSRN*.
- Burks, Stephen V, Bo Cowgill, Mitchell Hoffman, and Michael Housman, 2015, The value of hiring through employee referrals, *The Quarterly Journal of Economics* 130, 805–839.
- Card, David, and Laura Giuliano, 2016, Universal screening increases the representation of low-income and minority students in gifted education, *Proceedings of the National Academy of Sciences* 113, 13678–13683.
- Cestau, Dario, Dennis Epple, and Holger Sieg, 2017, Admitting Students to Selective Education Programs: Merit, Profiling, and Affirmative Action, *Journal of Political Economy* 125.
- Chen-Zion, Ayal, and James E. Rauch, 2020, History dependence, cohort attachment, and job referrals in networks of close relationships, *Journal of Economic Behavior & Organization* 170, 75–95.

- CoreLogic, 2025, Multiple Listing Service, Ownership Transfer, and Mortgage Data, 2017–2024, CoreLogic, Inc. Proprietary data obtained under license (accessed [September 2025]).
- Credit Bureau, 2016–2024, Consumer Credit Panel: 1% Random Sample of Annual Credit Snapshots and Tradelines, One of the three major U.S. credit bureaus, anonymized under the data use agreement. Proprietary data obtained under license (accessed [April 2025]).
- 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.
- Elliott, Marc N., Peter A. Morrison, Allen Fremont, Daniel F. McCaffrey, Philip Pantoja, and Nicole Lurie, 2009, Using the Census Bureau’s surname list to improve estimates of race/ethnicity and associated disparities, *Health Services and Outcomes Research Methodology* 9, 69–83.
- Ernst, Keith, Debbie Bocian, and Wei Li, 2008, Steered wrong: Brokers, borrowers, and subprime loans, *Center for Responsible Lending* 8.
- Federal Financial Institutions Examination Council, 2018–2024, Home Mortgage Disclosure Act Loan/Application Records, 2018–2024, Consumer Financial Protection Bureau, <https://ffiec.cfpb.gov/data-publication/> (accessed [September 2025]).
- FHFA and CFPB, 2024, National Survey of Mortgage Originations Public Use File, 2018–2022 Waves, <https://www.fhfa.gov/data/national-survey-of-mortgage-originations-nsmo-public-use-file> (accessed [September 2025]).
- 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.
- Hainmueller, Jens, 2012, Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies, *Political Analysis* 20, 25–46.
- Han, Lu, and Seung-Hyun Hong, 2011, Testing cost inefficiency under free entry in the real estate brokerage industry, *Journal of Business & Economic Statistics* 29, 564–578.
- Han, Lu, and Seung-Hyun Hong, 2016, Understanding in-house transactions in the real estate brokerage industry, *RAND Journal of Economics* 47, 1057–1086.
- Han, Lu, and William Strange, 2014, The microstructure of housing markets: Search, bargaining, and brokerage, *Handbook of Regional And Urban Economics* (Duranton, Henderson, and Strange (eds.)) 5.
- Ho, Kate, and Ariel Pakes, 2014, Hospital Choices, Hospital Prices, and Financial Incentives to Physicians, *American Economic Review* 104, 3841–3884.
- Jorgensen, Rebecca A, 2024, The Economic Consequences of Mergers Between Real Estate Agencies and Mortgage Lenders .
- Kermani, Amir, and Francis Wong, 2026, Racial disparities in housing returns, *American Economic Review* 116, 287–331.
- LaCour-Little, Michael, 2009, The Pricing of Mortgages by Brokers: An Agency Problem?, *Journal of Real Estate Research* 31, 235–264.

- Matvos, Gregor, Amit Seru, and Lulu Wang, 2023, Minority specialized lenders, *Northwestern University Kellogg working paper* .
- Nationwide Multistate Licensing System, 2012–2024, NMLS Business-to-Business Data: Individual Registration, Sponsorship, Company, and Branch Files, Conference of State Bank Supervisors. Proprietary data obtained under license (accessed [December 2024]).
- 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, *Working Paper* .
- 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.
- U.S. Census Bureau, 2010–2024, Small Area Income and Poverty Estimates: Median Household Income by County, <https://www.census.gov/programs-surveys/saipe.html> (accessed [September 2025]).
- U.S. Census Bureau, 2011, 2010 Census Summary File 1: Race and Hispanic Origin by ZIP Code Tabulation Area, <https://data.census.gov> (accessed [September 2025]).
- U.S. Census Bureau, 2016, Frequently Occurring Surnames from the 2010 Census, https://www.census.gov/topics/population/genealogy/data/2010_surnames.html (accessed [September 2025]).
- Verboven, Frank, and Biliana Yontcheva, 2024, Private monopoly and restricted entry: Evidence from the notary profession, *Journal of Political Economy* 132, 3658–3707.
- 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.
- Wooldridge, Jeffrey M., 2015, Control function methods in applied econometrics, *Journal of Human Resources* 50, 420–445.

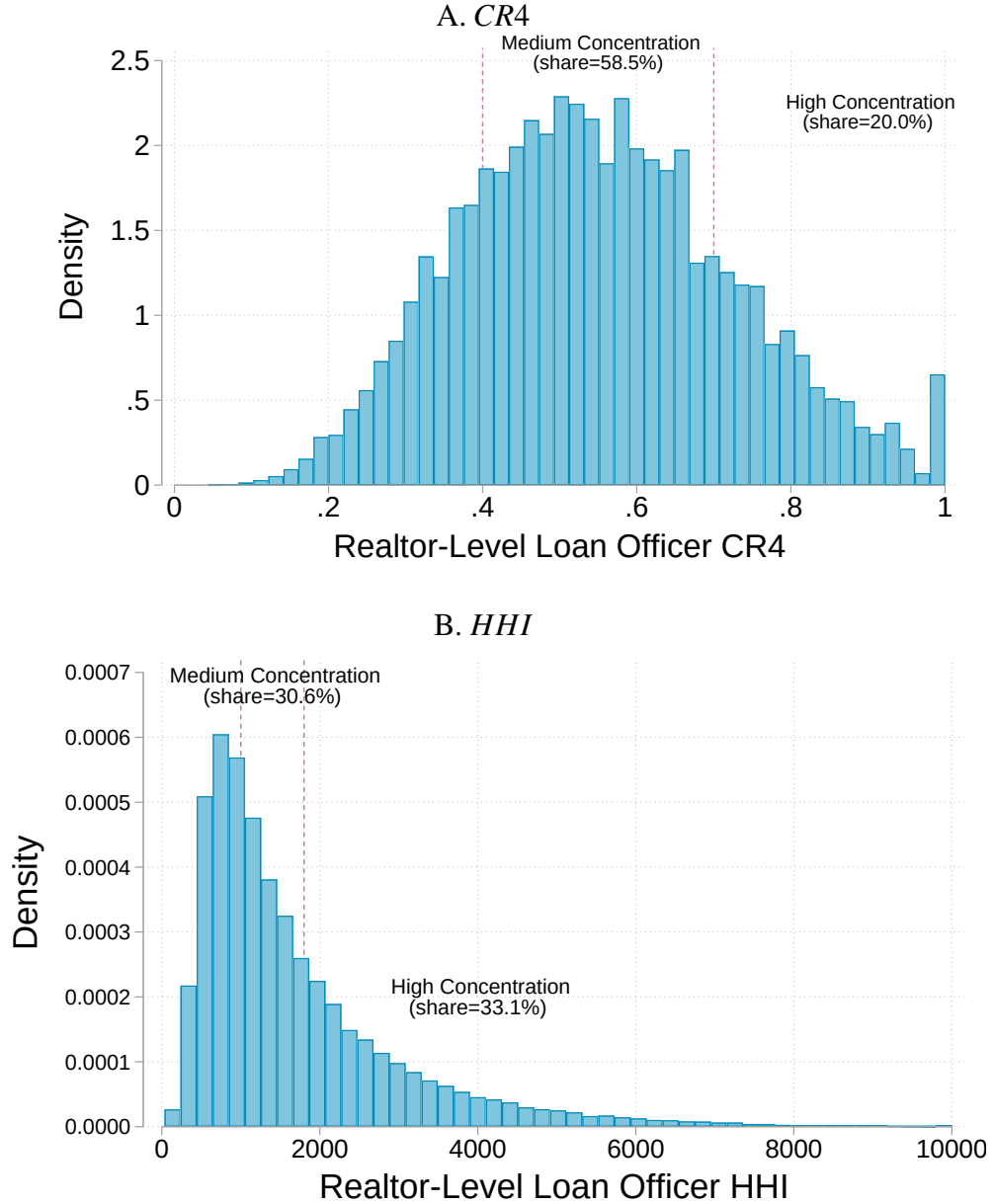
Figures

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. We restrict the sample to counties in which a single dominant MLS accounts for at least 80% of housing listings, that dominant MLS provides buyer-agent identifiers, and reports contract (purchase-agreement) dates for at least 50% of its listings. In total, our sample includes 703 counties across 40 states, which collectively represent 42.7% of the U.S. population.

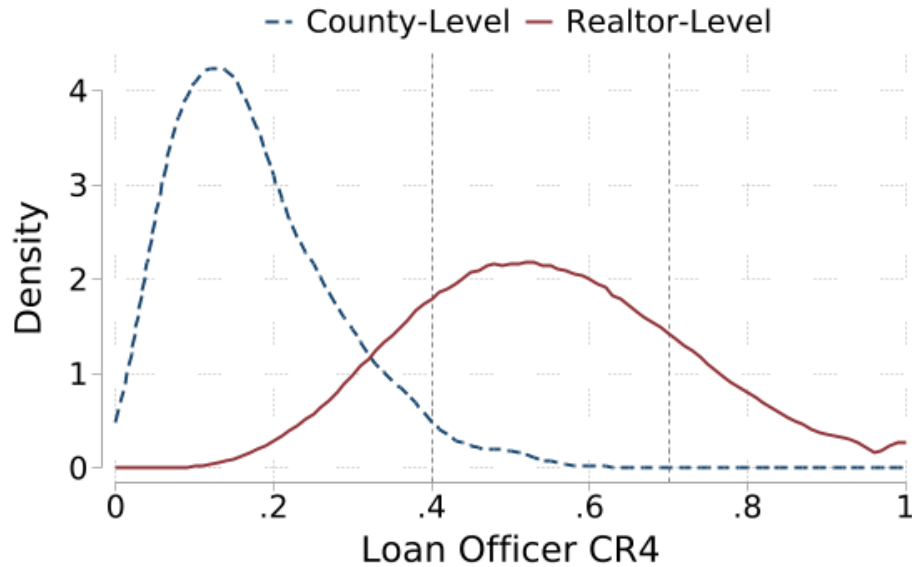
Figure 2: Distribution of Realtors' Loan Officer Concentration



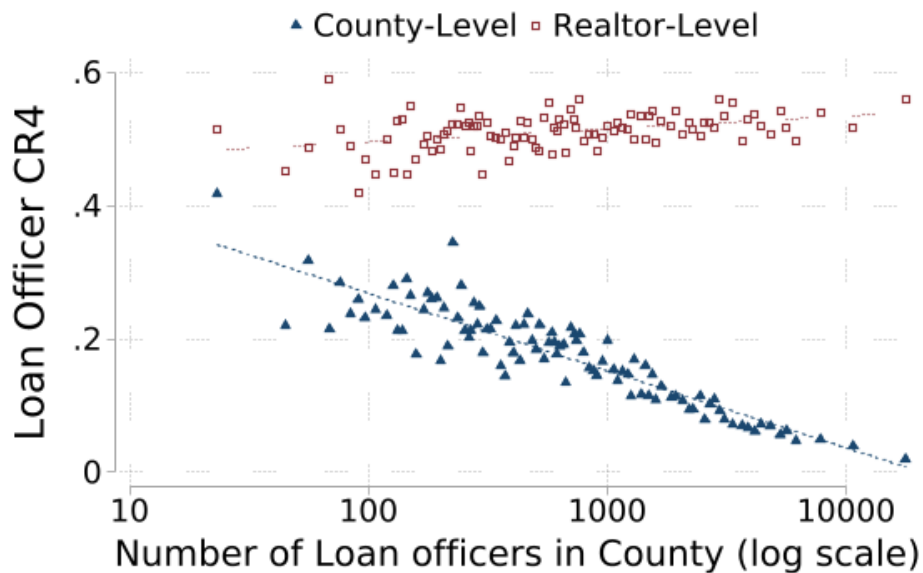
Note: This figure plots the distributions of realtors' loan officer concentration measures ($CR4$ in Panel A, HHI in Panel B). For each realtor, we define her loan officer network as the set of all loan officers who financed loans for at least one of the realtor's buyer clients. We then calculate each loan officer's 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' shares for a given realtor, while the HHI is the sum of squared shares across the realtor's entire loan officer network.

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

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

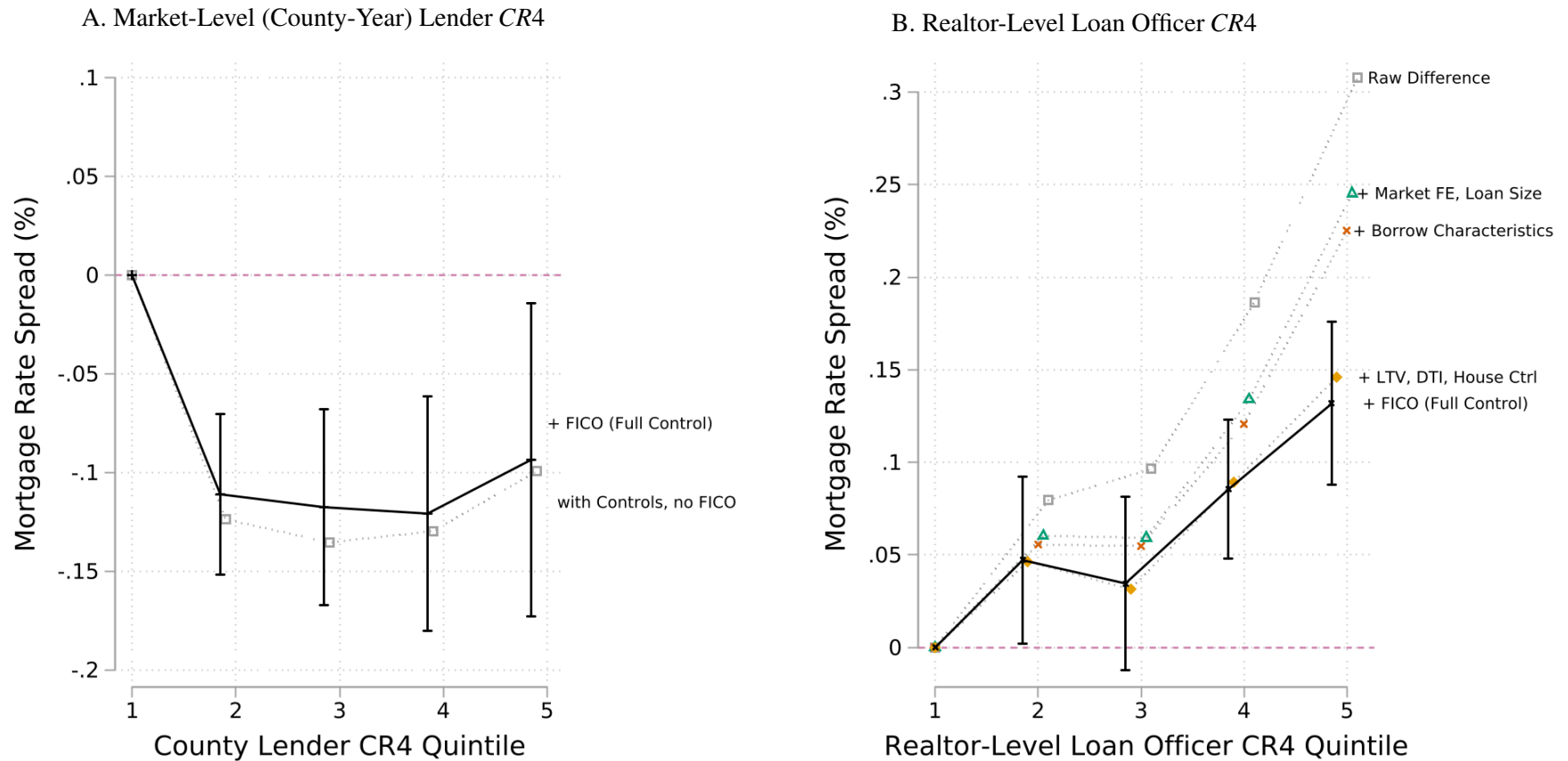


B. Loan Officer Concentration against Market Size



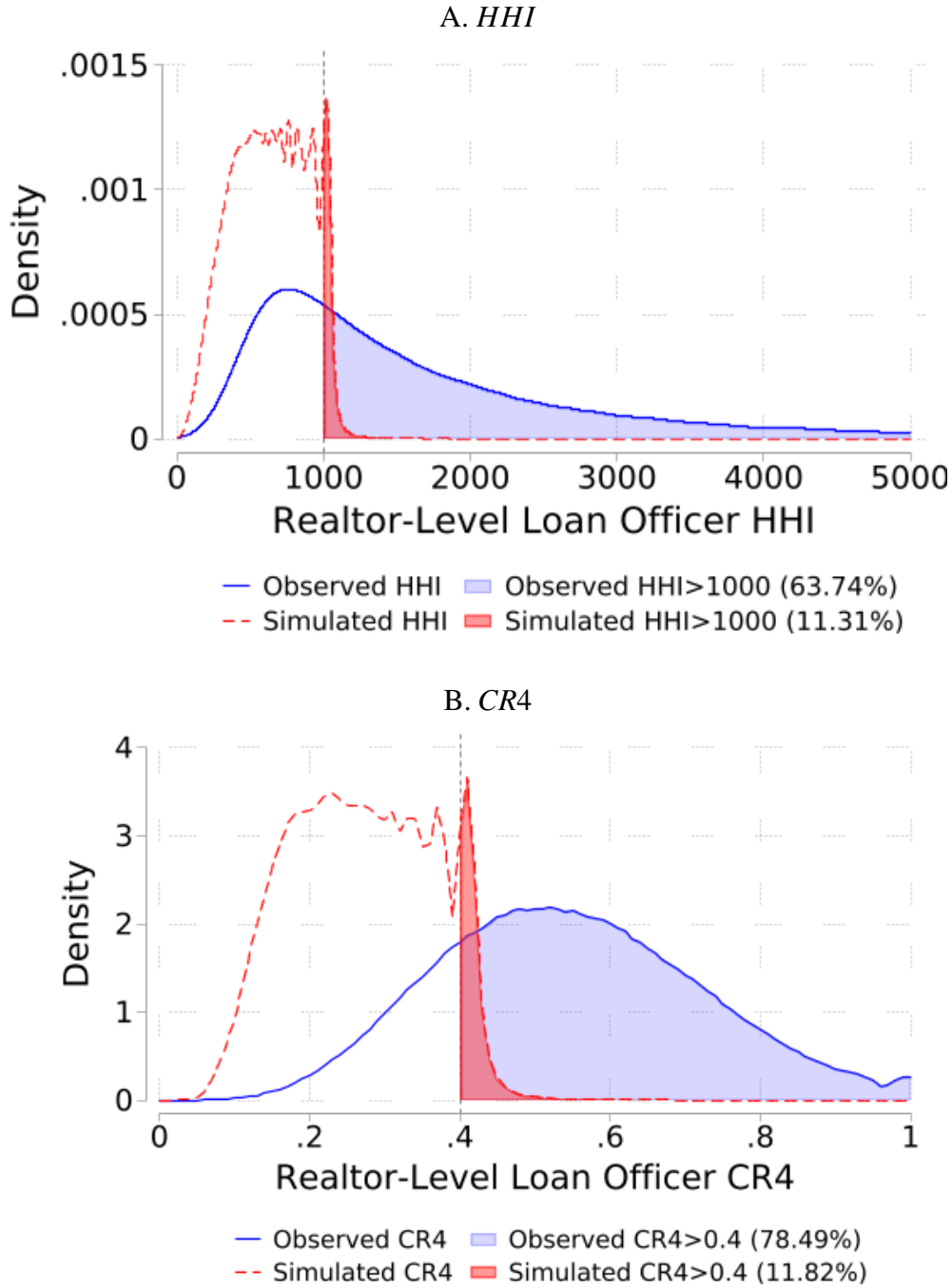
Note: This figure compares loan officer concentration (CR4) measured at the county level and the realtor level. Panel A plots the distributions of the two measures (dashed blue for county-level CR4 and solid red for realtor-level CR4). Panel B plots both measures against county size (blue triangles for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

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



Note: This figure presents the relationship between mortgage costs and lender concentration across different concentration quintiles. The 5th quintile corresponds to the highest concentration. The y-axis represents the mortgage APR spread, defined as the deviation from the benchmark rate offered on prime mortgage loans of a comparable type. In Panel A, concentration is measured by the market-wide lender CR4. The dashed line corresponds to the specification with all controls except FICO scores, while the solid black line corresponds to the fully controlled specification including FICO scores. In Panel B, concentration is measured by the within-realtor CR4 of loan officers. Gray squares plot the raw differences. Green triangles add county $\times$ year-month and loan amount bin fixed effects. Red crosses further control for borrower and loan characteristics. Orange diamonds add an even richer set of controls, including DTI and LTV bin fixed effects, house characteristics, and ZIP code fixed effects. Black crosses, connected by the solid black line, add FICO scores and represent the fully controlled specification.

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



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

Figure 6: Event Studies around the CFPB Enforcement Order



Note: This figure plots the coefficients on Freedom-tied realtor, estimated at the loan level, for the probability that a loan is financed by one of the realtor's pre-order referral loan officers (Panel A) and the purchase mortgage APR spread (Panel B) off the benchmark rate offered on prime mortgage loans of a comparable type, in percentage points. Treated realtors are referral realtors with a pre-order referral tie to a loan officer employed at Freedom in both 2021 and 2022; control realtors are realtors in the same counties without a referral network. Each panel is one regression with realtor, county $\times$ year-month fixed effects and the full set of borrower, loan and house controls, all absorbed. Treated realtors receive a weight of one and each control realtor receives an entropy-balancing weight, applied to all of her loans (Appendix F.2). The dashed line marks the CFPB order of August 2023, splitting 2023 into January–July and August–December. January–July 2023 is the reference period. Error bars are 95% confidence intervals. Standard errors are clustered at the county level.

Tables

Table 1: Homebuyer Attributes

	Full Sample		1% Random Sample (Matched to Credit Records)	
	(1)		(2)	
	mean	s.d.	mean	s.d.
Number of Observations (Homebuyers)	1,751,287		15,390	
APR Spread (%)	0.51	(0.69)	0.49	(0.69)
Loan Amount (\$1K)	304.18	(185.45)	332.79	(202.54)
Joint Application	0.44	(0.50)	0.61	(0.49)
Conforming	0.97	(0.18)	0.95	(0.21)
LTV (%)	87.83	(12.81)	87.13	(12.94)
DTI (%)	35.58	(11.26)	35.03	(11.49)
FICO Score			742.64	(61.69)
Credit Inquires (90 Days Pre-Close)			1.45	(0.98)
Credit Inquires (60 Days Pre-Close)			1.05	(0.93)
County Median Income	72.44	(16.21)	73.86	(16.10)
Income Ratio	1.54	(1.11)	1.66	(1.17)
First-Time Buyer			0.28	(0.45)
Education				
-HS or Less			0.28	(0.45)
-Some College			0.31	(0.46)
-College or Above			0.38	(0.49)
Hispanic	0.15	(0.36)	0.16	(0.37)
Race				
-White	0.84	(0.37)	0.84	(0.37)
-Black	0.08	(0.27)	0.07	(0.25)
-Asian	0.06	(0.23)	0.07	(0.25)
-Other Minority	0.01	(0.10)	0.01	(0.10)
Days to Close	39.03	(19.09)	38.55	(17.99)
Closing Price (\$1K)	360.61	(253.16)	397.54	(276.44)
Square Feet	2034.11	(21598.21)	2094.04	(1183.61)
Bedrooms	3.33	(0.83)	3.39	(0.83)
Bathrooms	2.49	(0.94)	2.58	(0.95)
House Age	42.60	(30.64)	40.76	(29.34)
Referral	0.48	(0.50)	0.49	(0.50)
Non-Compliant Referral	0.29	(0.45)	0.28	(0.45)

Note: This table presents the summary statistics of the home purchase data used in this study. Column (1) reports statistics for the full sample, while column (2) summarizes the 1% random sample matched to credit records. “Income ratio” is defined as the reported income over the county median income of the same year. A borrower is a “first-time home buyer” if the credit file shows no first-lien mortgage opened before the purchase and no record of homeownership predating it. “Ethnicity” and “Race” are self-reported in HMDA. We impute only missing values using Bayesian Improved Surname Geocoding (BISG), which combines surname and ZIP-code racial composition while preserving all self-reported responses. “Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network ($CR4 > 0.4$) and the loan is financed by one of that realtor’s top four loan officers. “Non-Compliant Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network ($CR4 > 0.4$) and the loan is not financed by one of that realtor’s top four loan officers. The FICO score, credit inquiry counts, first-time buyer status and education are only available in the 1% random sample.

Table 2: **Placebo Test**—Mortgage Rate Spreads By **Market-Level** Lender *CR4* Quintiles

	Mortgage Rate Spread (off Prime Loans, in %)					
	No Control (1)	+Market FE (2)	+Borrower Ctrl (3)	+DTI, LTV (4)	+House Ctrl (5)	+FICO (6)
Panel A: Full Sample						
Market Lender <i>CR4</i> Quintile 5	-0.116*** (0.041)	-0.068*** (0.021)	-0.072*** (0.023)	-0.063*** (0.019)	-0.069*** (0.024)	
Market Lender <i>CR4</i> Quintile 4	-0.129*** (0.044)	-0.071*** (0.020)	-0.076*** (0.022)	-0.065*** (0.018)	-0.072*** (0.023)	
Market Lender <i>CR4</i> Quintile 3	-0.174*** (0.036)	-0.069*** (0.017)	-0.075*** (0.019)	-0.061*** (0.015)	-0.068*** (0.019)	
Market Lender <i>CR4</i> Quintile 2	-0.160*** (0.043)	-0.046*** (0.011)	-0.048*** (0.010)	-0.040*** (0.010)	-0.045*** (0.011)	
Observations	1,751,287	1,751,287	1,751,287	1,751,287	1,751,287	
R-squared	0.008	0.144	0.165	0.269	0.292	
Dep. Var. Mean	0.51	0.51	0.51	0.51	0.51	
Panel B: 1% Random Sample Matched with Credit Records						
Market Lender <i>CR4</i> Quintile 5	-0.159*** (0.054)	-0.086** (0.042)	-0.093** (0.044)	-0.099*** (0.036)	-0.099** (0.042)	-0.093** (0.040)
Market Lender <i>CR4</i> Quintile 4	-0.167*** (0.056)	-0.116*** (0.034)	-0.121*** (0.035)	-0.120*** (0.026)	-0.130*** (0.034)	-0.121*** (0.030)
Market Lender <i>CR4</i> Quintile 3	-0.214*** (0.041)	-0.128*** (0.023)	-0.129*** (0.025)	-0.117*** (0.020)	-0.135*** (0.026)	-0.118*** (0.025)
Market Lender <i>CR4</i> Quintile 2	-0.209*** (0.050)	-0.107*** (0.019)	-0.106*** (0.019)	-0.104*** (0.018)	-0.124*** (0.021)	-0.111*** (0.021)
Observations	15,390	15,390	15,390	15,390	15,390	15,390
R-squared	0.016	0.185	0.216	0.324	0.449	0.513
Dep. Var. Mean	0.49	0.49	0.49	0.49	0.49	0.49
Year-Month FE		✓	✓	✓	✓	✓
County FE		✓	✓	✓	✓	✓
Loan Amount Bin FE		✓	✓	✓	✓	✓
Borrower Age Bin FE			✓	✓	✓	✓
Income Ratio Percentile FE			✓	✓	✓	✓
Joint Application			✓	✓	✓	✓
Conforming FE			✓	✓	✓	✓
DTI Bin FE				✓	✓	✓
LTV Bin FE				✓	✓	✓
Bedroom FE					✓	✓
Bathroom FE					✓	✓
Sqft Bin FE					✓	✓
House Age FE					✓	✓
ZIP Code FE					✓	✓
FICO Bin FE						✓

Note: This table reports the differences of mortgage APR rate spreads across market-level (county-year) lender concentration (*CR4*) quintiles. The 5th quintile corresponds to the highest concentration. The dependent variable is the mortgage APR spread off the benchmark rate offered on prime mortgage loans of a comparable type. Standard errors are clustered at the county level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Mortgage Rate Spreads By **Realtor-Level** Loan Officer *CR4* 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 <i>CR4</i> Quintile 5	0.259*** (0.022)	0.198*** (0.015)	0.181*** (0.013)	0.139*** (0.011)	0.117*** (0.008)	
Realtor <i>CR4</i> Quintile 4	0.156*** (0.014)	0.113*** (0.009)	0.104*** (0.008)	0.083*** (0.006)	0.073*** (0.005)	
Realtor <i>CR4</i> Quintile 3	0.111*** (0.010)	0.077*** (0.006)	0.071*** (0.006)	0.058*** (0.005)	0.053*** (0.004)	
Realtor <i>CR4</i> Quintile 2	0.067*** (0.008)	0.047*** (0.005)	0.044*** (0.005)	0.038*** (0.004)	0.036*** (0.003)	
Observations	1,751,287	1,751,287	1,751,287	1,751,287	1,751,287	
R-squared	0.015	0.177	0.195	0.295	0.317	
Dep. Var. Mean	0.51	0.51	0.51	0.51	0.51	
Panel B: 1% Random Sample Matched with Credit Records						
Realtor <i>CR4</i> Quintile 5	0.308*** (0.029)	0.245*** (0.025)	0.225*** (0.023)	0.171*** (0.021)	0.146*** (0.022)	0.132*** (0.023)
Realtor <i>CR4</i> Quintile 4	0.186*** (0.026)	0.134*** (0.024)	0.121*** (0.022)	0.097*** (0.020)	0.089*** (0.021)	0.085*** (0.019)
Realtor <i>CR4</i> Quintile 3	0.096*** (0.020)	0.059*** (0.017)	0.055*** (0.018)	0.041*** (0.019)	0.032 (0.024)	0.035 (0.024)
Realtor <i>CR4</i> Quintile 2	0.080*** (0.017)	0.060*** (0.020)	0.056*** (0.020)	0.052*** (0.019)	0.046*** (0.023)	0.047*** (0.023)
Observations	15,390	15,390	15,390	15,390	15,390	15,390
R-squared	0.023	0.380	0.405	0.486	0.600	0.646
Dep. Var. Mean	0.49	0.49	0.49	0.49	0.49	0.49
Year-Month*County FE		✓	✓	✓	✓	✓
Loan Amount Bin FE		✓	✓	✓	✓	✓
Borrower Age Bin FE			✓	✓	✓	✓
Income Ratio Percentile 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 spreads across realtor-level loan officer concentration (*CR4*) quintiles. The dependent variable is the mortgage APR spread off the benchmark rate offered on prime mortgage loans of a comparable type. Standard errors are clustered at the county level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

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

	Mortgage Rate Spread (off Prime Loans, in %)					
	Full Sample			1% Random Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Referral	0.107***	0.106***	0.130***	0.101***	0.108***	0.144***
(CR4 >0.4, Top 4 LO)	(0.004)	(0.006)	(0.005)	(0.015)	(0.019)	(0.023)
Non-Compliant Referral		-0.000	0.002		0.010	0.013
(CR4 >0.4, Not Top 4 LO)		(0.003)	(0.003)		(0.018)	(0.018)
Referral × Superstar LO			-0.047***			-0.072***
			(0.003)			(0.025)
Superstar LO			0.030***			0.042***
			(0.006)			(0.015)
Observations	1,751,287	1,751,287	1,751,287	15,390	15,390	15,390
R-squared	0.320	0.320	0.320	0.647	0.647	0.647
Dep. Var. Mean	0.51	0.51	0.51	0.49	0.49	0.49
Year-Month*County FE	✓	✓	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓	✓	✓
Income Ratio Percentile FE	✓	✓	✓	✓	✓	✓
Joint Application	✓	✓	✓	✓	✓	✓
Conforming FE	✓	✓	✓	✓	✓	✓
DTI Bin FE	✓	✓	✓	✓	✓	✓
LTV Bin FE	✓	✓	✓	✓	✓	✓
Bedroom FE	✓	✓	✓	✓	✓	✓
Bathroom FE	✓	✓	✓	✓	✓	✓
Sqft Bin FE	✓	✓	✓	✓	✓	✓
House Age FE	✓	✓	✓	✓	✓	✓
ZIP Code FE	✓	✓	✓	✓	✓	✓
FICO Bin FE				✓	✓	✓

Note: This table reports the OLS regression estimates of the effect of referral lending on mortgage costs. Columns (1)–(3) use the full sample and columns (4)–(6) the 1% random sample matched to credit records (which adds FICO-bin fixed effects); within each block, the third column adds the superstar loan-officer interaction. Same set of controls (FE) as in Table 3. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type (see Appendix B for alternative cost measures). “Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network ($CR4 > 0.4$) and the loan is financed by one of that realtor’s top four loan officers. “Non-Compliant Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network ($CR4 > 0.4$) but the loan is not financed by one of that realtor’s top four loan officers. “Superstar LO” equals one if the loan officer is in the top 5% of loan officers by purchase-loan volume (at least 129 purchase loans over the sample period); “Referral × Superstar LO” is the differential referral premium when the originating top-four loan officer is a superstar. Standard errors are clustered at the county level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

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

	Referral (First Stage)		Mortgage Rate Spread (off Prime Loans, in %)	
	1% Random Sample		1% Random Sample	
	(1)	(2)	2SLS (3)	2SLS (4)
$(\widehat{HHI}-\widehat{HHI})/10000$	1.451*** (0.036)	1.414*** (0.052)		
Referral ($CR4 > 0.4$, Top 4 Loan Officer)			0.151*** (0.010)	0.177*** (0.036)
Observations	1,751,287	15,390	1,751,287	15,390
R-squared	0.205	0.521		
Dep. Var. Mean	0.51	0.49	0.51	0.49
Year-Month*County FE	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓
Income Ratio Percentile 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			298031	2226
FS: Kleibergen-Paap rk F			1641	750

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

Table 6: Causal Effect of Referral Lending on Mortgage Costs (IV: CFPB Enforcement Order)

	Referral (First Stage)		Mortgage Rate Spread (off Prime Loans, in %)	
	(1)	(2)	2SLS (3)	2SLS (4)
Freedom-tied Realtor*Post (IV)	-0.257*** (0.040)	-0.251*** (0.038)		
Referral (CR4 > 0.4, Top 4 Loan Officer, pre-order)			0.323* (0.168)	0.405** (0.167)
Observations	54,196	54,196	54,196	54,196
R-squared	0.617	0.662		
Dep. Var. Mean	0.06	0.06	0.32	0.32
Year-Month*County FE	✓	✓	✓	✓
Realtor FE	✓	✓	✓	✓
Loan Amount Bin FE		✓		✓
Age Bin FE		✓		✓
Income Ratio Percentile FE		✓		✓
Joint Application		✓		✓
Conforming FE		✓		✓
DTI Bin FE		✓		✓
LTV Bin FE		✓		✓
Bedroom FE		✓		✓
Bathroom FE		✓		✓
Sqft Bin FE		✓		✓
House Age FE		✓		✓
ZIP Code FE		✓		✓
FS: Cragg-Donald Wald F			824	779
FS: Kleibergen-Paap rk F			41	42

Note: This table reports the IV estimates of the effect of referral lending on mortgage costs, identified from the CFPB's August 2023 RESPA enforcement order against Freedom Mortgage. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. The sample is realtors with at least a quarter of their pre-order business in the 20 counties where Freedom Mortgage's referral business operated at scale. Treated realtors are referral realtors with a pre-order referral tie to a loan officer employed at Freedom in both 2021 and 2022, and control realtors are realtors in the same counties without a referral network (see Appendix F). The post period begins with originations in August 2023. "Referral" equals one if the loan is financed by one of the realtor's top four pre-order loan officers, and is instrumented with "Freedom-tied Realtor*Post", the interaction of Freedom-tied status with an indicator for originations from August 2023 onward. Columns (1)–(2) report the first stage and columns (3)–(4) the 2SLS estimates, under two conditioning sets: columns (1) and (3) include realtor and county $\times$ year-month fixed effects only; columns (2) and (4) add the full set of borrower, loan and property controls. Treated realtors receive a weight of one and each control realtor receives an entropy-balancing weight, applied to all of her loans (Appendix F.2). Standard errors are clustered at the county level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: Effect of Referral on Days to Close

	Days to Close		
	IV: $\widehat{HHI}-HHI$		IV: CFPB Enforcement Order
	1% Random Sample		
	2SLS (1)	2SLS (2)	2SLS (3)
Referral	-1.429*** (0.145)	-1.608 (1.028)	2.880 (3.454)
Observations	1,744,026	15,294	53,964
Dep. Var. Mean	39.03	38.54	47.55
Year-Month*County FE	✓	✓	✓
Realtor FE			✓
Loan Amount Bin FE	✓	✓	✓
Age Bin FE	✓	✓	✓
Income Ratio Percentile 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	296856	2218	770
FS: Kleibergen-Paap rk F	1640	734	43

Note: This table reports the IV estimates of the effect of referral on home purchase closing time, measured as the number of days between contract acceptance and closing. Columns (1) and (2) use the same set of controls (FE) as in Table 3; column (2) is estimated on the 1% random sample matched to credit records and adds FICO bin fixed effects. “Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network ($CR4 > 0.4$) and the loan is financed by one of that realtor’s top four loan officers. In columns (1) and (2) it is instrumented with the “excess concentration” ($\widehat{HHI}-HHI$) of a realtor’s loan officer network, which is the difference between observed HHI concentration and predicted HHI in the absence of referral lending. Column (3) instruments referral with the 2023 CFPB enforcement order against Freedom Mortgage, following the design of Section 4.3 and the specification and matching weights of Table 6. Standard errors are clustered at the county level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 8: Causal Effect of Referral Lending for Borrowers with Minimal Denial Risks

Low-Risk Borrowers:	Mortgage Rate Spread (off Prime Loans, in %)			
	IV: $HHI-\widehat{HHI}$			IV: CFPB Enforcement Order
	FICO > 700	FICO > 780	DTI ≤ 36%, LTV ≤ 80%, Income > Median	DTI ≤ 36%, LTV ≤ 80%, Income > Median
	1% Random Sample			
	2SLS (1)	2SLS (2)	2SLS (3)	2SLS (4)
Referral × Low-Risk Borrowers	0.182*** (0.034)	0.123*** (0.041)	0.110*** (0.013)	1.337** (0.535)
Referral × Other Borrowers	0.163** (0.074)	0.203*** (0.048)	0.156*** (0.010)	0.318* (0.178)
Observations	15,390	15,390	1,751,287	54,196
Dep. Var. Mean	0.488	0.488	0.515	0.324
Share of Low-Risk Borrowers	0.779	0.340	0.139	0.206
Year-Month*County FE	✓	✓	✓	✓
Realtor FE				✓
Loan Amount Bin FE	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓
Income Ratio Percentile 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	1113	1104	149289	120
FS: Kleibergen-Paap rk F	394	313	848	3.1

Note: This table reports the IV estimates of the effect of referral lending on mortgage costs across risk groups. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. “Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network ($CR4 > 0.4$) and the loan is financed by one of that realtor’s top four loan officers. It is interacted with an indicator for the low-risk borrowers defined in the column header and with its complement (“Other Borrowers”). Each coefficient is the referral premium for that group relative to non-referred borrowers in the same group. “Share of Low-Risk Borrowers” is the fraction of loans in the low-risk group. Columns (1) and (2) are estimated on the 1% random sample matched to credit records, with low-risk borrowers defined by FICO scores above 700 and above 780, respectively. Columns (3) and (4) define low-risk borrowers as those with a DTI at or below 36%, an LTV at or below 80%, and income above the county median, with no FICO restriction; column (3) uses the full estimation sample and column (4) the sample of Table 6. In columns (1)–(3) the two referral terms are instrumented with the “excess concentration” ($HHI-\widehat{HHI}$) of a realtor’s loan officer network and its interaction with the low-risk indicator. Column (4) instruments them with the 2023 CFPB enforcement order against Freedom Mortgage and its interaction with the low-risk indicator, following the specification and matching weights of Table 6. Standard errors are clustered at the county level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 9: Effect of Referral on Mortgage Search Activities (IV: $\widehat{HHI-HHI}$)

	Within 90 Days		Within 60 Days	
	Credit Inquiries	>1 Inquiry (in %)	Credit Inquiries	>1 Inquiry (in %)
	2SLS (1)	2SLS (2)	2SLS (3)	2SLS (4)
Referral (CR4 >0.4, Top 4 Loan Officer)	-0.129** (0.063)	-6.602** (2.844)	-0.149*** (0.045)	-7.091*** (2.104)
Observations	15,390	15,390	15,390	15,390
Dep. Var. Mean	1.45	35.17	1.05	22.61
Year-Month*County FE	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓
Income Ratio Percentile 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	2226	2226	2226	2226
FS: Kleibergen-Paap rk F	750	750	750	750

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

Table 10: Effect of Referral Lending on Mortgage Costs within the Same Lender (IV: $\widehat{HHI}-HHI$)

	Mortgage Rate Spread (off Prime Loans, in %)		
	2SLS (1)	2SLS (2)	1% Random Sample
			2SLS (3)
Referral ($CR4 > 0.4$, Top 4 Loan Officer)	0.027*** (0.006)	0.021*** (0.007)	0.060* (0.036)
Observations	1,751,066	1,355,500	14,945
Dep. Var. Mean	0.51	0.52	0.49
Year-Month*County FE	✓		✓
Lender FE	✓		✓
Year-Month*County*Lender FE		✓	
Loan Amount Bin FE	✓	✓	✓
Age Bin FE	✓	✓	✓
Income Ratio Percentile 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	252292	191683	1762
FS: Kleibergen-Paap rk F	1735	1173	462

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

Table 11: Causal Effect of Referral Lending By Borrower Characteristics (IV: $\widehat{HHI-HHI}$)

	Mortgage Rate Spread (off Prime Loans, in %)			
	1% Random Sample			
	(First-Time)	(Education)	(Ethnicity)	(Race)
	2SLS (1)	2SLS (2)	2SLS (3)	2SLS (4)
Referral*First-Time Buyer	0.278*** (0.040)			
Referral*Repeat Buyer	0.127*** (0.037)			
Referral*HS or Less		0.191*** (0.062)		
Referral*Some College		0.152** (0.063)		
Referral*College or Above		0.142*** (0.051)		
Referral*Hispanic			0.218* (0.120)	
Referral*Non-Hispanic			0.154*** (0.038)	
Referral*White				0.184*** (0.036)
Referral*Black				0.093 (0.119)
Referral*Asian				0.207* (0.124)
Observations	15,390	15,390	15,390	15,390
Dep. Var. Mean	0.49	0.49	0.49	0.49
Full Controls	✓	✓	✓	✓
FS: Cragg-Donald Wald F	1121	533	690	288
FS: Kleibergen-Paap rk F	375	24	99	1.1

Note: This table reports the IV estimates of the effect of referral lending on mortgage costs across demographic groups. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. “Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network ($CR4 > 0.4$) and the loan is financed by one of that realtor’s top four loan officers. The endogenous variables are instrumented with the “excess concentration” ($\widehat{HHI-HHI}$) of a realtor’s loan officer network and its interaction with the demographic group indicators. All columns are estimated on the 1% random sample matched to credit records. Appendix Table A3 reports the ethnicity and race splits on the full estimation sample. In each column, the referral indicator is fully interacted with one mutually exclusive and exhaustive set of borrower groups, so each coefficient is the referral premium for that group relative to non-referred borrowers in the same group. On each split, the unclassified borrowers are retained as their own category, which is included in every regression but omitted from the table for readability. For unreported race and ethnicity in HMDA, we reclassify them using Bayesian Improved Surname Geocoding (BISG), which combines the borrower’s surname with the racial composition of the property’s ZIP code; every self-reported value is left untouched. Standard errors are clustered at the county level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix

for

Market Power in Mortgage Pricing: the Role of Referral Lending

Dayin Zhang Panle Jia Barwick Lu Han Jon Kroah

A Data Construction

This appendix details the construction of the main purchase-mortgage sample used throughout the paper and the refinance sample used in the BLP estimation (Appendix D). Figure A1 summarizes the full sample-construction sequence and the observations lost at each step.

A.1 CoreLogic House, Listing and Mortgage Data

We rely on three CoreLogic data products, each covering the period July 2017 through December 2024. The first is the **Multiple Listing Service (MLS)** data, which records housing listings reported by local MLS platforms. Each listing contains the property address, listing and closing prices, listing and contract dates, and, crucially, standardized identifiers for the listing agent and the buyer agent. These data allow us to observe the realtor who represents each homebuyer. The second is the **Ownership Transfer** data (deed records), which records all property-level transactions filed with county recorders. The third is the **Mortgage** data, which records mortgage originations extracted from the same public deed sources. We join the three datasets through a universal property identifier (composite property linkage key).

County Selection As discussed in Section 2.2, our geographic sample is restricted to counties that satisfy three conditions: (i) a dominant MLS accounts for at least 80% of listings, (ii) the dominant MLS reports buyer-agent identifiers for at least 10% of listings, and (iii) the dominant MLS reports contract dates for at least 50% of listings. Counties with fewer than 10 listings in the dominant MLS are dropped. The resulting geographic sample comprises 703 counties across 40 states, mapped in Figure 1.

Identifying mortgage-financed home purchase events We construct the sample of arm’s-length, mortgage-financed home purchases in four steps. Figure A1 traces the observations surviving each step from Step 2 onward, together with the sample restrictions described in the remainder of this appendix.

Step 1: Collapse listings into listing events. A single property may appear multiple times in the MLS if the seller relists after a price change, temporary withdrawal, or expired contract. To avoid double-counting sale attempts, we sort listings by property and listing date, treat any two records

posted within 7 days as a single *listing attempt*, and group consecutive listing attempts for the same property into a single *listing event* whenever the gap between attempts is no more than 90 days and no attempt in between resulted in a close. Before collapsing, we drop (i) listings whose list or close price is below 10% of the property’s historical average, a price screen that removes suspected rental listings, and (ii) multi-structure properties—those whose list prices vary widely within a single listing attempt,¹ indicating that the units of a multi-unit property (e.g., a main house with a separate cottage) were sold separately.

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

Step 3: Match each deed to a mortgage. We link each house transaction record to the CoreLogic Mortgage data using the property ID, searching for candidate mortgages within a ± 60 -day window of the deed’s sale date. Although the mortgage origination date should in principle match the sale date, we allow a 60-day discrepancy to account for possible recording errors. We also exclude observations in which the mortgage origination date precedes the original listing date. These screens leave 7,930,568 transactions.

Step 4: Validate identifiers. Finally, we retain only transactions carrying valid individual realtor and loan-officer identifiers. On the realtor side, we exclude (i) placeholder identifiers such as UNKNOWNMEMBER, NONMLS, or TEST AGENT; and (ii) IDs that record more than 100 transactions in a single calendar year, which are almost always team- or brokerage-level identifiers rather than individual agents. On the loan-officer side, we drop transactions for which the originating loan officer lacks a Nationwide Multistate Licensing System (NMLS) ID, the unique federally regulated identifier assigned to individual mortgage loan originators. Together with the restriction to mortgages originated between January 2018 and December 2024, this step yields 6,743,632 mortgage-financed home purchases observed with both the realtor and the loan officer who intermediated them.

Step 5: Drop low-volume realtors. To guard against spurious concentration arising from low transaction counts, we drop realtors associated with fewer than ten transactions over the sample period. The result is our *network sample* of 5,167,324 transactions, on which we construct each realtor’s

¹Specifically, the minimum list price within a listing attempt is below 90% of the maximum.

loan-officer network and concentration measures ($CR4_r$, HHI_r).

A.2 Merging CoreLogic with HMDA

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

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

Of the 5,167,324 mortgages in the network sample, 3,127,141 (61%) are matched one-to-one to HMDA records.

Estimation Sample Restrictions From the 3,127,141 CoreLogic–HMDA matched mortgages, we apply the following additional restrictions to construct the estimation sample. We impose them in the order listed; the count in parentheses reports the number of purchases surviving each one:

- **Loan product.** Retain only first-lien, 30-year, fixed-rate conventional or government-backed mortgages, the most common product in the U.S. single-family market. We additionally require the loan to be secured by owner-occupied, single-unit, and site-built properties.
- **Loan size and income.** Drop loans with an original principal balance below \$100,000, which are disproportionately non-traditional originations (e.g., manufactured housing, investor loans misclassified as owner-occupied), and loans with invalid reported income.
- **Non-missing outcome.** Drop observations with a missing APR spread. (2,480,755)
- **Realtor activity in the matched sample.** Drop realtors with fewer than ten matched purchase loans, so that each realtor’s counterfactual $\widehat{HHI}_r$ and $\widehat{CR4}_r$ (used in the IV) can be computed

from a sufficient number of simulated loan-officer choices. This retains 91,988 of the 179,567 realtors present in the matched data. (1,972,578)

- **Non-missing controls.** Drop observations with missing values for any of: LTV, DTI, applicant age, applicant income, and the joint-application indicator, together with 10,055 observations falling in singleton fixed-effect cells. (1,751,287)

The resulting estimation sample comprises 1,751,287 transactions intermediated by 91,988 distinct realtors and 123,749 distinct loan officers. The baseline regressions reported in the paper are estimated on this sample; specifications requiring FICO scores or credit-inquiry counts use the credit-bureau matched subsample described next.

A.3 Credit-Bureau Matched Subsample

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

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

For the 2018–2024 vintages, 73–84% of primary-type credit-bureau mortgages are successfully matched to CoreLogic, with the large majority linked in the first stage. Intersecting this crosswalk with the main purchase-regression sample yields our credit-bureau matched subsample of 15,390

mortgage observations. This subsample is used in analyses that require FICO controls (the baseline rate-spread regressions in the 1% sample) and in analyses of mortgage search intensity (credit-inquiry outcomes in Section 5.3).

A.4 Refinance Sample for BLP Estimation

The IV strategy (Section 4.2) requires estimating a lender-choice model on transactions where realtor referrals are largely absent. We construct a parallel *refinance* sample that mirrors the purchase sample wherever possible:

- **Geographic and temporal coverage.** Same 703 counties and same 2018–2024 window.
- **Loan-product filters.** First-lien, 30-year, fixed-rate, with a loan balance of at least \$100,000.

Because refinances are not intermediated by real estate agents, we impose no realtor-side restrictions. FICO scores are a key input to the lender-choice model but are not reported in HMDA; we therefore intersect the refinance sample with the 1% random credit-bureau panel described in Section A.3, which provides directly observed borrower FICO for each transaction. The resulting refinance sample comprises 105,808 transactions. Our baseline drops the 11,298 refinances that return to the loan officer who originated the borrower’s purchase mortgage (Appendix D.6), leaving 94,510; summary statistics for this sample appear in Table A10 in Appendix D. Dropping county-year markets with fewer than ten refinances leaves the 84,205 transactions on which Equation (5) is estimated (Table A11, Column (1)); keeping the relationship refinances gives 95,394.

B Decomposing the Referral Premium into Price Components

The headline results measure the price of a mortgage by its APR spread. The APR is a to-maturity concept: it converts upfront charges into an annualized rate by amortizing them over the full contractual term, and so implicitly assumes the borrower holds the loan to maturity. Because the average purchase mortgage prepays well before its thirtieth year, the APR overstates the annualized cost of a dollar of upfront charges for the typical borrower. This appendix therefore reports the referral premium separately for each component of the price a borrower actually faces, leaving them in their own units rather than aggregating them through an amortization assumption.

We use three outcomes:

- **Contract rate spread**, the contract interest rate net of the average prime offer rate (APOR). It removes the finance charges from the APR spread. In practice, the APOR is quoted on the interest rate lock day, which is not observed. So we derived the lock-day APOR by subtracting the APR spread from the APR, which is derived from its reported terms and charges and note rate.
- **Discount points**, HMDA's reported discount points, as a percentage of the loan amount. A borrower who pays points buys down the note rate, so points and the rate spread are substitutes along a menu the lender offers.
- **Origination fees net of points**, HMDA's reported origination charges, which include discount points. We subtract the discount points to measure the fees paid directly to the lenders.

We report the decomposition twice, once per estimator, so that the two samples used for the headline results sit side by side. Table A4 reports OLS and Table A5 reports 2SLS, instrumenting referral with excess concentration ($HHI - \widehat{HHI}$). In both, each price component spans a pair of columns: the odd-numbered column is the full estimation sample and the even-numbered column the 1% credit-bureau matched subsample with FICO bin fixed effects, the same pairing as columns (3) and (6) of Table 4 and the two corresponding columns of Table 5. Within each sample the three components are estimated on one common set of loans, those with all three observed.

Three findings emerge, and each holds in both samples and under both estimators. First, the referral premium is essentially all rate, and therefore not an artifact of amortization. The contract rate spread rises by 12.4 basis points in the full sample and 13.3 in the credit-matched subsample under OLS, and by 14.2 and 16.4 basis points under the IV. The referral effect on the unadjusted APR spread on the same rows is 12.9 and 14.1 under OLS and 15.0 and 17.0 under the IV, so stripping the annualized finance charge out of the outcome moves the coefficient by less than one basis point everywhere: the contract rate spread is 94% to 97% of the total in all four cells. Whatever the correct assumption about how long a borrower holds the loan, it is close to immaterial for the estimates in this paper, because the referral premium was almost entirely a rate premium to begin with.

Second, referred borrowers also pay higher origination fees: 0.093% of the loan amount in the full sample and 0.105% in the credit-matched subsample under OLS, and 0.109% and 0.110% under the

IV, against sample means of roughly 0.47%.

Third, and least expected, referred borrowers purchase *fewer* discount points: -0.045% and -0.040% of the loan amount under OLS, and -0.022% and -0.071% under the IV, against means near 0.31%. This is the opposite of what a rate–points menu would produce. If referred borrowers simply chose a lower-upfront-cost point on the same menu, they would pay a higher rate and fewer points, but they would not simultaneously pay higher origination fees; if they were being steered into a costlier product, the rate and the fees would both rise, which is what we observe.

These pieces fit together. Netting the two upfront components, referred borrowers pay between 0.038% and 0.087% more of the loan amount in total upfront charges depending on the sample and estimator. Converting that to a rate equivalent at the paper’s factor of 0.218 percentage points per point of upfront cost gives between 0.8 and 1.9 basis points, and adding it to the contract rate spread relatively accurately reproduces the APR spread effect.

C Additional Results and Robustness Checks

C.1 Sensitivity to the Minimum Realtor Transaction Count

To limit spurious concentration arising from realtors with very few transactions, our main analysis restricts attention to realtors with at least ten purchase loans over the study period (Section 2.2). A natural concern is whether the headline estimates are sensitive to this cutoff. We therefore re-estimate the two central specifications – the referral OLS in Table 4 and the instrumented referral effect in Table 5 – on a substantially broader sample that retains every realtor with at least *five* purchase transactions. Lowering the threshold from ten to five roughly doubles the number of realtors in the estimation sample, from 91,988 to 183,629, and expands the regression sample from 1.75 to 2.28 million purchase mortgages (and the credit-matched subsample from 15,390 to 21,507 observations).

Tables A6 and A7 report the results, which are nearly identical to the baseline. The OLS referral premium is 10.6 basis points in the full sample and 9.6 basis points in the credit-matched sample, compared with 10.7 and 10.1 basis points in the main analysis. The instrumented (2SLS) referral premium is 15.9 basis points in the full sample and 18.1 basis points in the credit-matched sample, compared with 15.1 and 17.7 basis points in Table 5, and the first stage remains very strong. The stability of the estimates – despite roughly doubling the set of realtors – indicates that the documented

referral premium is not an artifact of restricting attention to higher-volume realtors.

C.2 Sensitivity to the Referral Definition: The Realtor’s Top 1 Loan Officer

Our baseline classifies a loan as a referral when the originating loan officer is among the realtor’s *top four* most-used officers within a high-concentration network ($CR4 > 0.4$). One might worry that the top-four cutoff is arbitrary, or that it dilutes the referral signal by pooling the realtor’s single most-trusted officer with three less-frequently used ones. We therefore re-estimate the headline OLS and IV specifications under a stricter definition that counts a loan as a referral only if it is originated by the realtor’s *single most-used* (rank-one, “top 1”) loan officer; loans in a high-concentration network that are originated by any other officer – including the realtor’s second- through fourth-ranked officers – are reclassified as non-compliant.

Tables A8 and A9 report the results. The compliant referral premium is essentially unchanged from the baseline: once non-compliant loans are separated, the OLS premium for borrowers routed to the realtor’s top-1 officer is 10.5 basis points in the full sample and 10.0 basis points in the credit-matched sample, very close to the 10.6 and 10.8 basis points under the top-four definition (Table 4). The instrumented (2SLS) premium is 15.1 basis points in the full sample and 18.4 basis points in the credit-matched sample (compared with 15.1 and 17.7 in Table 5), and the first stage remains very strong. Two further patterns reinforce the referral interpretation. First, the simple referral indicator *without* the non-compliant control is *smaller* under the top-1 definition (7.8 versus 10.7 basis points in the full sample), precisely because the implicit comparison group now absorbs the realtor’s second-through fourth-ranked officers, who themselves carry an elevated premium. Second, that premium is visible directly: the non-compliant coefficient rises to 3.9 basis points in the full sample, compared with a near-zero estimate under the top-four definition, consistent with referral intensity that declines smoothly with an officer’s rank in the realtor’s network rather than dropping discretely after the single top officer. Columns (3) and (6) of Table A8 additionally add the superstar loan-officer interaction, mirroring the main-text Table 4: holding the superstar margin fixed *raises* the top-1 referral coefficient (to 13.8 basis points in the full sample), and the Referral $\times$ Superstar interaction is negative (–5.9 basis points), so the top-1 premium is likewise not driven by superstar officers. Overall, the cost of referral is concentrated among the realtor’s most-used loan officers and is robust to tightening the referral definition all the way to the single top-1 officer.

D BLP Estimation

D.1 Model

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

$$u_{ij,cy} = X'_{j,cy}\beta_1 + X'_{ij,cy}\beta_2 + \xi_{j,cy} + \varepsilon_{ij,cy}$$

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

$$s_{ij,cy}(\beta_1, \beta_2) = \frac{\exp(X'_{j,cy}\beta_1 + X'_{ij,cy}\beta_2 + \xi_{j,cy})}{\sum_{l=1}^{J_{cy}} \exp(X'_{l,cy}\beta_1 + X'_{il,cy}\beta_2 + \xi_{l,cy})} \quad (7)$$

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

$$s_{ij,cy}(\delta_{j,cy}, \beta_2) = \frac{\exp(\delta_{j,cy} + X'_{ij,cy}\beta_2)}{\sum_{l=1}^{J_{cy}} \exp(\delta_{l,cy} + X'_{il,cy}\beta_2)} \quad (8)$$

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

$$s_{j,cy}(\delta_{j,cy}, \beta_2) \approx \frac{1}{N_{cy}} \sum_{i=1}^{N_{cy}} s_{ij,cy}(\delta_{j,cy}, \beta_2) \quad (9)$$

D.2 Estimation Procedure

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

$$\delta_{j,cy}^{(h+1)} = \delta_{j,cy}^{(h)} + \log \left(\frac{\tilde{s}_{j,cy}}{s_{j,cy}(\delta_{j,cy}^{(h)}, \beta_2)} \right) \quad (10)$$

where $\tilde{s}_{j,cy}$ is the observed share of lender j in market cy . We iterate until $|\delta_{j,cy}^{(h+1)} - \delta_{j,cy}^{(h)}| < 10^{-12}$ for all j, cy . This yields the estimated mean utilities, $\hat{\delta}(\beta_2)$, that match model-predicted and observed

market shares. We then search over β_2 until the log-likelihood is maximized, using the L-BFGS-B algorithm.

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

$$\hat{\delta}_{j,cy} = X'_{j,cy}\beta_1 + \alpha_{cy} + \xi_{j,cy} \quad (11)$$

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

D.3 Empirical Specification

To avoid estimating mean utilities for many small lenders, we treat the bottom five percent of lenders in each market as a single composite lender. When we predict choice probabilities in the purchase loan sample, we apply the estimated mean utility for the composite lender to all its constituent lenders.

We include the following lender-market-specific characteristics in $X_{j,cy}$:

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

We include the following borrower-lender-market-specific characteristics in $X_{ij,cy}$:

- 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 combined loan-to-value ratio and the “bank”, “fintech”, and each of the #1, #2, #3 lender dummies

- Interactions between debt-to-income ratio and the “bank”, “fintech”, and each of the #1, #2, #3 lender dummies;
- Interactions between FICO score and the “bank”, “fintech”, and each of the #1, #2, #3 lender dummies.

D.4 Summary Statistics

We estimate the model on a 1% random sample of refinance mortgages, constructed by matching CoreLogic refinance transactions to the 1% random credit-bureau panel described in Appendix A.3. The match provides observed borrower FICO scores (not reported in HMDA), which are a key input to the lender-choice model. Table A10 presents summary statistics for the 94,510 refinances that remain after dropping relationship refinances (Appendix D.6), before imputing missing values. We then drop county-year markets with fewer than 10 transactions, which leaves the 84,205 transactions used in estimation, and replace missing values of any variables with their unconditional sample means.

D.5 Estimates

Column (1) of Table A11 presents the estimates of (β_1, β_2) .

D.6 Robustness to Relationship Refinances

Because we estimate the lender-choice model on refinance mortgages, the estimated demand system could in principle be contaminated by refinances that do not reflect a fresh search over lenders. In particular, a borrower may refinance with the same loan officer who originated their purchase mortgage, so that the observed choice encodes a pre-existing relationship rather than a *de novo* demand decision. Such “relationship” refinances are a minority of the sample: among refinances that we can link to an earlier non-refinance loan on the same property and for which the earlier loan officer is identifiable, 24% return to the same individual loan officer. If these transactions systematically reinforce incumbent lenders, they could inflate the estimated taste for large, established lenders and, through the counterfactual, the simulated instrument. Our baseline therefore excludes them: Column (1) of Table A11 is estimated on the 84,205 refinances that do not return to the purchase loan officer.

To assess whether this exclusion matters, Column (2) of Table A11 re-estimates the model keeping every refinance (95,394 loans) and reports the estimates alongside the baseline. The two sets of

coefficients are very close where it matters: the economically central parameters – the bank effects, the distance disutility and the FICO-sorting gradients – move only slightly. Several of the top-lender interaction terms, which are imprecisely estimated in both columns, do move in sign or significance. Propagating this set of estimates through the counterfactual simulation of Appendix E yields a simulated instrument that is virtually indistinguishable from the baseline. Re-running the headline IV regressions of Table 5 with this reconstructed instrument leaves the estimates essentially intact: as Table A12 reports, the instrumented referral premium is 17.8 basis points with relationship refinances kept, against 17.7 in the baseline, with first stages just as strong. Because relationship refinances would, if anything, bias us toward finding excess concentration, their immateriality indicates that residual contamination of the refinance sample by pre-existing borrower–loan-officer relationships does not drive our demand estimates or our downstream results.

E Simulation of the No–Referral Counterfactual

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

E.1 Simulating lender choice

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

$$\hat{p}_{rji} = \frac{\exp(X'_{j,cy}\hat{\beta}_1 + X'_{ij,cy}\hat{\beta}_2 + \hat{\xi}_{j,cy})}{\sum_{l=1}^{J_{cy}} \exp(X'_{l,cy}\hat{\beta}_1 + X'_{il,cy}\hat{\beta}_2 + \hat{\xi}_{l,cy})}, \quad j = 1, \dots, J_{cy}$$

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

In this process, if lender j has no refinance transactions in market cy , then $\hat{\xi}_{j,cy}$ is unavailable; we assign the unconditional mean of $\hat{\xi}_{j,cy}$ to such lenders. HMDA does not report applicant credit scores, so for purchase-sample borrowers without a direct credit-bureau match we impute FICO as the

average credit score across all individuals in a 1% random credit-bureau panel residing in the same ZIP code in the same year.

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

As a robustness check, we also construct the counterfactual choice probabilities using the full estimated market–lender mean utility $\hat{\delta}_{j,cy} = X'_{j,cy}\hat{\beta}_1 + \hat{\alpha}_{cy} + \hat{\xi}_{j,cy}$ from Appendix D directly – i.e., the estimated fixed effect that exactly rationalizes each lender’s observed market share – in place of reconstructing it from its components $X'_{j,cy}\hat{\beta}_1 + \hat{\xi}_{j,cy}$, adjusting only the referral-related terms for the counterfactual. This “fixed-effect” specification and the baseline specification yield counterfactual concentration measures, and downstream instrumented estimates, that are nearly identical: the headline second-stage referral premium is unchanged, and across every mechanism and heterogeneity specification the instrumented coefficients move by less than two percent and never change sign or statistical significance. We conclude that our findings do not depend on how the unobserved component of lender mean utility is treated in the counterfactual.

E.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.

E.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.

F The CFPB Enforcement Order: Design Details and Supporting Evidence

F.1 Sample Construction

This appendix records the construction of the sample used in Section 4.3 and the counts at each step.

The Freedom group and its loan officers. We identify each loan officer’s employer from the NMLS registry: the officer and company identifiers recorded on the deed are matched to the officer’s NMLS registration and sponsorship spells, and a company is assigned to a loan only if the mortgage date falls within an active spell at that company. As stated in the consent order, Freedom’s traditional retail unit, the officers who maintained the realtor relationships at issue, was moved into Freedom’s wholly owned subsidiary RoundPoint Mortgage Servicing in August 2021. In our data the same officers’ originations carry RoundPoint’s identifier over that window, so we treat the two entities (NMLS 2767 and 18188) as one Freedom group. The roster is the set of loan officers who originated at least one purchase loan before August 2023 while employed at the group: 561 officers. A roster officer is employed at Freedom in both 2021 and 2022 if she originated at least one loan while at the group in each of those calendar years, so that she was active at the group through the two years before the order.

Treated realtors. All network measures are computed on originations before August 2023. A realtor is a referral realtor if her pre-order top-4 loan-officer share exceeds 0.4 and she has at least ten pre-order transactions; her pre-order top-4 officers are her referral loan officers, and a loan is a referral loan if its loan officer is one of them, in any period. A referral realtor is *treated* if at least one of her referral loan officers is employed at Freedom in both 2021 and 2022.

Market restrictions. A county enters if at least ten pre-order loans were written by a treated realtor and financed by a roster officer; 20 counties qualify, in New York, New Jersey, Arizona, Nevada, California, Massachusetts and Ohio. A realtor, treated or control, enters the sample if at least 25 percent of her pre-order loans lie in those counties, and her entire book is then kept, so that the panel is balanced within realtor.

Control realtors. Controls are realtors with at least ten pre-order transactions whose pre-order top-4 share is at or below 0.4, i.e. realtors without a referral network under the paper’s definition, who pass the same market restrictions. With the full control set, the regression sample contains 512 treated realtors with 5,102 purchase loans and 2,311 control realtors with 49,094 loans.

Specification. Every regression absorbs realtor fixed effects, county $\times$ year-month fixed effects, ZIP code fixed effects, and the borrower, loan and property controls of Table 3. Every regression in Section 4.3, the event studies, the first stages, and the 2SLS regressions, is estimated with analytic weights: a treated realtor’s loans have weight one, and a control realtor’s loans carry her entropy-balancing weight, constructed in Appendix F.2 from the realtor-level pre-order characteristics of Table A13 and the county shares, and scaled so that the control weights sum to the number of treated realtors. Figure A2 and Table A14 report the unweighted versions.

F.2 Matching on Pre-Order Characteristics

Treated and control realtors differ by construction in whether they operate a referral network, and Table A13 shows that their pre-order business differs as well. Treated realtors closed somewhat fewer purchase loans per year (3.0 to 4.2 against 4.1 to 5.4 for controls) and smaller ones (\$409,000 against \$465,000), and their clients borrowed at higher loan-to-value and debt-to-income ratios (86 against 83 percent, and 38.5 against 35.5 percent), with lower incomes relative to the county median (1.29

against 1.80 times) and almost entirely in conforming loans (99 against 90 percent). The design does not require the two groups to be alike: the identifying assumption is that, absent the order, their outcomes would have moved in parallel, and the regressions of Section 4.3 absorb these characteristics loan by loan. One could still worry that realtors who serve more leveraged, lower-income borrowers faced different movements in mortgage pricing than the controls as rates rose over 2022–2023, which would surface as a pre-trend. We therefore make the two groups comparable on these characteristics and use the matched controls throughout Section 4.3.

We do so by matching, implemented as entropy balancing (Hainmueller, 2012). Each control realtor receives a weight, chosen to be as close as possible to uniform weights, such that the weighted means of every characteristic in Table A13, and of the shares of each of the 20 treated counties, equal the corresponding treated means; treated realtors receive a weight of one, and the control weights are scaled to sum to the number of treated realtors.² Balancing is done at the realtor level and every loan of a realtor carries her weight, so no control realtor is dropped and the panel is unchanged. Column (5) shows that balance is exact. The weights are concentrated, because controls whose clients resemble the treated realtors' clients are relatively few: the largest weight is 8.4 and the weighted effective number of control realtors is 343 of the 2,463 that enter the matching (2,311 remain once the fixed-effects regressions drop singleton realtors), so the matched estimates are less precise than the unweighted ones.

Figure A2 repeats Figure 6 without the matching weights, and the picture is the same. In Panel A, treated realtors' routing to their referral loan officers shows the same pre-order stability and a sharp drop at the order of 22.0 percentage points, and of 25.6 points over 2024. In Panel B, the rate spread paid by treated realtors' borrowers falls after the order by 13.1 bps, and by 14.6 bps over 2024. The one difference is the pre-order coefficient for 2020, which is negative and significant without the weights: matching the two groups removes this discrepancy and leaves the post-order drop intact.

Table A14 re-estimates the IV specifications of Tables 6, 7 and 8 without the matching weights. The first stage is somewhat stronger, and the referral premium is 35.5 basis points, slightly below and statistically indistinguishable from the 40.5 basis points in Table 6. Days to close remain small and insignificant, and the premium for low-risk borrowers from the pooled interacted specification of Table 8 is 79.6 basis points, again larger than for other borrowers and more precisely estimated than

²Counties with at least one treated and five control realtors enter as separate shares; the remaining counties are pooled into one share.

with the weights, on a first stage with a Kleibergen–Paap F of 9.7. Reweighting the control realtors therefore changes none of the conclusions: the results of Section 4.3 do not hinge on the pre-order differences between treated and control realtors.

G An Alternative Identification Strategy: Local Realtor Exits

This appendix reports a third source of variation in referral lending, one that operates through the realtor side of the relationship rather than the lender side. When a high-volume realtor exits the market, the pipeline she maintained toward particular loan officers is severed, and homebuyers purchasing in her former territory shortly afterwards face a local intermediation environment with fewer referral channels. Comparing these homebuyers with neighbors just outside the exposure zone yields an instrument for referral lending that is plausibly unrelated to the characteristics of any individual borrower. Because the design requires a single realtor to have dominated a local market, it is confined to thin markets and is far less precise than the two designs in the main text; we present it as corroborating evidence.

G.1 Design

A homebuyer is *treated* if a qualifying realtor exit occurred within the previous two years and the exiting realtor(s) accounted for at least 25% of purchase transactions within a 10-mile radius of the property, measured in the year preceding the exit. This exposure bar is deliberately high: it selects markets in which the exiting realtor was locally dominant, so that her exit materially changes the referral environment. *Control* homebuyers are those whose local exit exposure never reaches the 25% threshold at any lag and whose nearest same-year treated homebuyer lies between 10 and 20 miles away, the ring immediately adjacent to the exposure zone; we further require that the control’s county-year cell contain at least one treated homebuyer. Figure A3 illustrates the design around a single treated homebuyer.

G.2 Sample Construction

The design draws on the CoreLogic MLS transaction history. Because exposure is measured as the exiting realtor’s share of local transactions, that share is meaningful only where the MLS captures

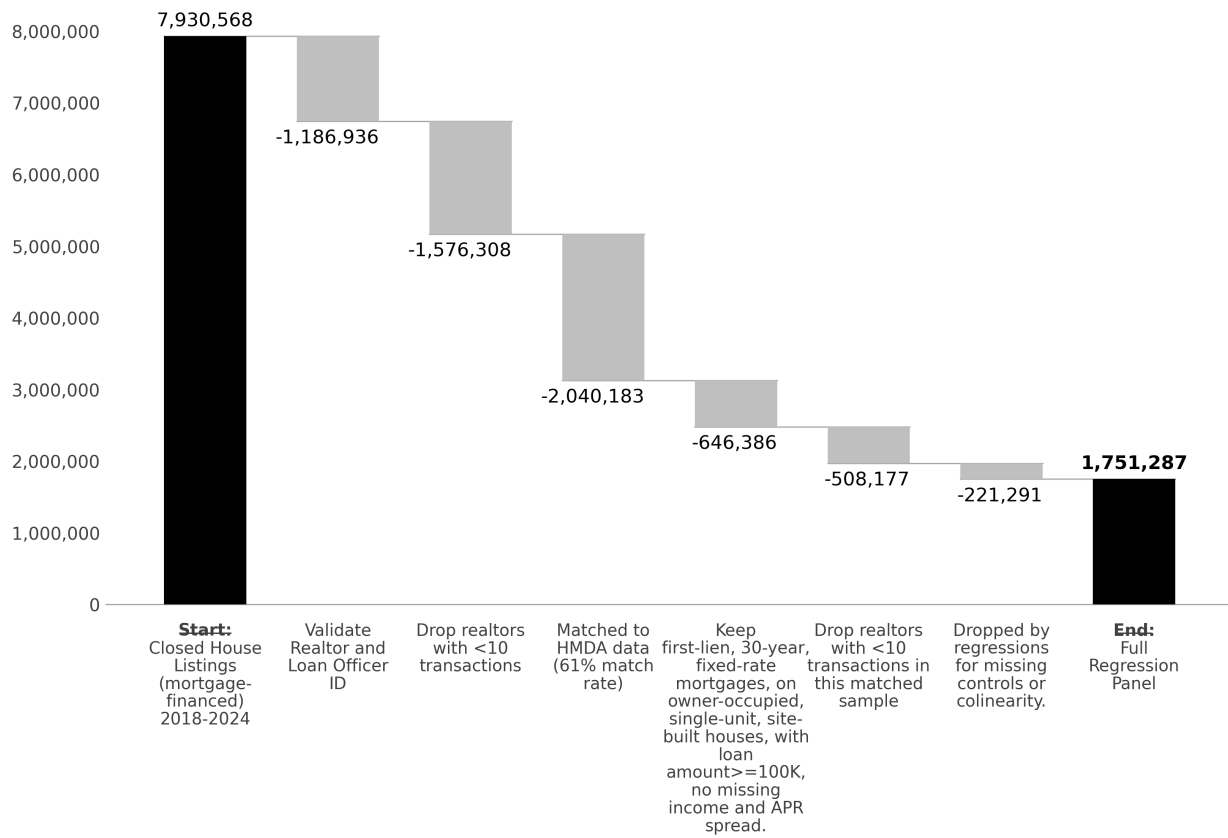
most of what actually transacts: a neighborhood that appears dominated by one realtor in a thinly covered MLS may not be dominated at all. We therefore restrict the sample to ZIP Codes in which MLS closed transactions exceed 50% of deed records in the years used to measure exposure. For every realtor we date her last closed transaction and call the exit *confirmed* if at least twelve months of inactivity separate it from the end of the panel (December 2024). A confirmed exit *qualifies* if the realtor was operating at scale and stopped abruptly: she averaged at least four transactions a year from her first active year through the exit year, and her volume in the year before the exit was at least 80% of her average annual volume over the preceding active years. The second condition separates realtors who went from full activity to none from those who were already winding down, whose clients would have been redirected gradually rather than at a single date. The panel yields 12,150 qualifying exits. At the exposure settings above, the estimation sample contains 5,767 homebuyers, 386 treated and 5,381 controls, in 13 counties (64 of them fall in singleton fixed-effect cells and drop out of the regressions, which therefore use 5,703).

Table A15 compares treated and control homebuyers on the borrower and property characteristics that the regressions absorb as fixed effects, reporting group means, the raw difference, and the difference within county-by-year-month cells, the comparison the fixed-effects specification actually uses. Within cells the two groups are statistically indistinguishable on every covariate: loan amount, DTI, LTV, income, joint application, conforming status, house age, and house size.

G.3 Results

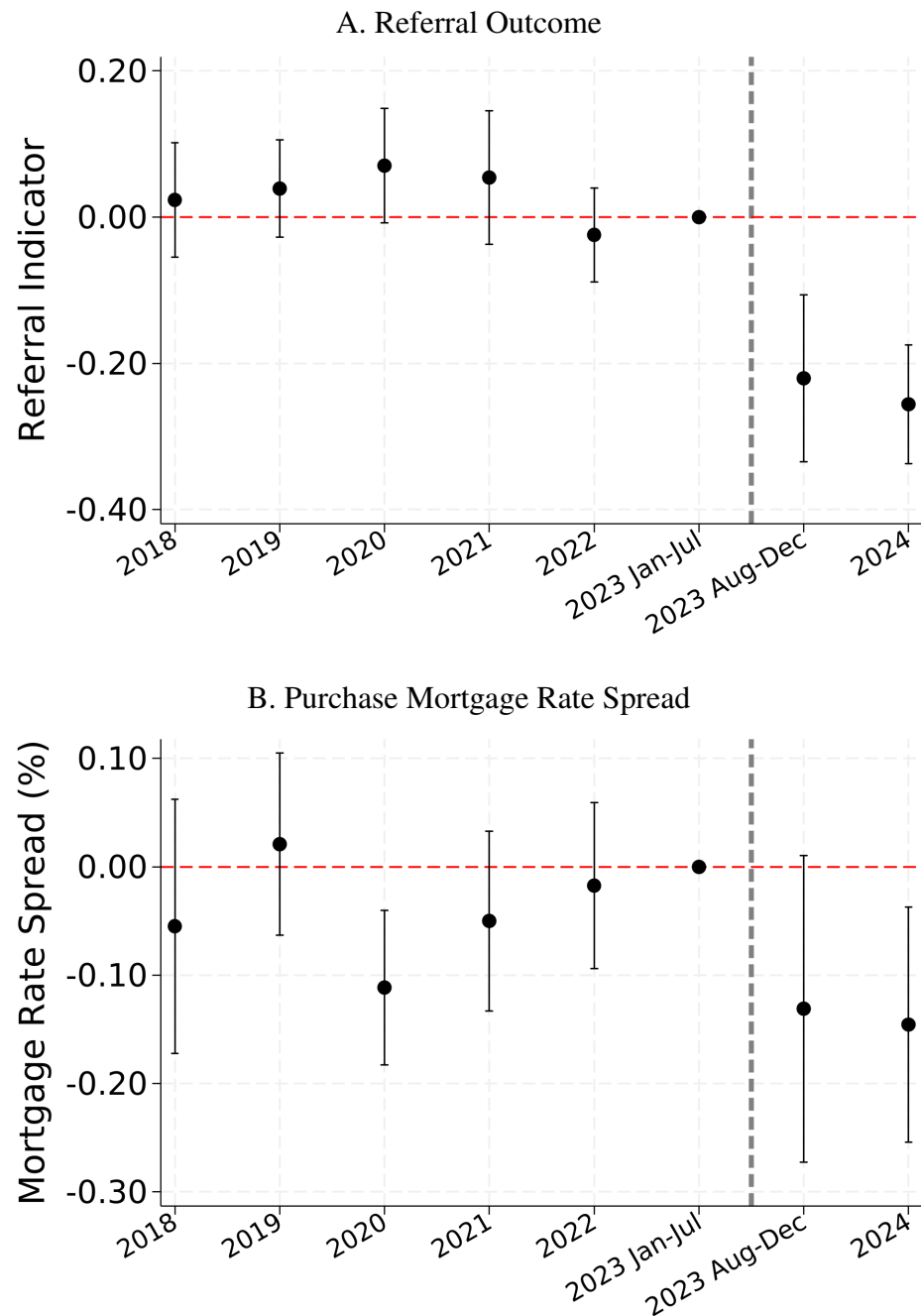
Table A16 reports the results. Column (1) is the first stage: local exposure to a dominant realtor's sharp exit one to two years before origination lowers the probability of a referral loan by 18.1 percentage points. The cluster-robust Kleibergen–Paap F is 30.6. Column (2) instruments referral with exit exposure: referral lending raises rate spreads by 30.5 basis points. The estimate is less precise, but its sign and magnitude line up with the excess-concentration and enforcement-order designs of Sections 4.2 and 4.3. As with the enforcement order, the exit design is identified from relationships that mattered, those maintained by a locally dominant realtor, so a somewhat larger local average treatment effect is what one would expect.

Figure A1: Sample Construction



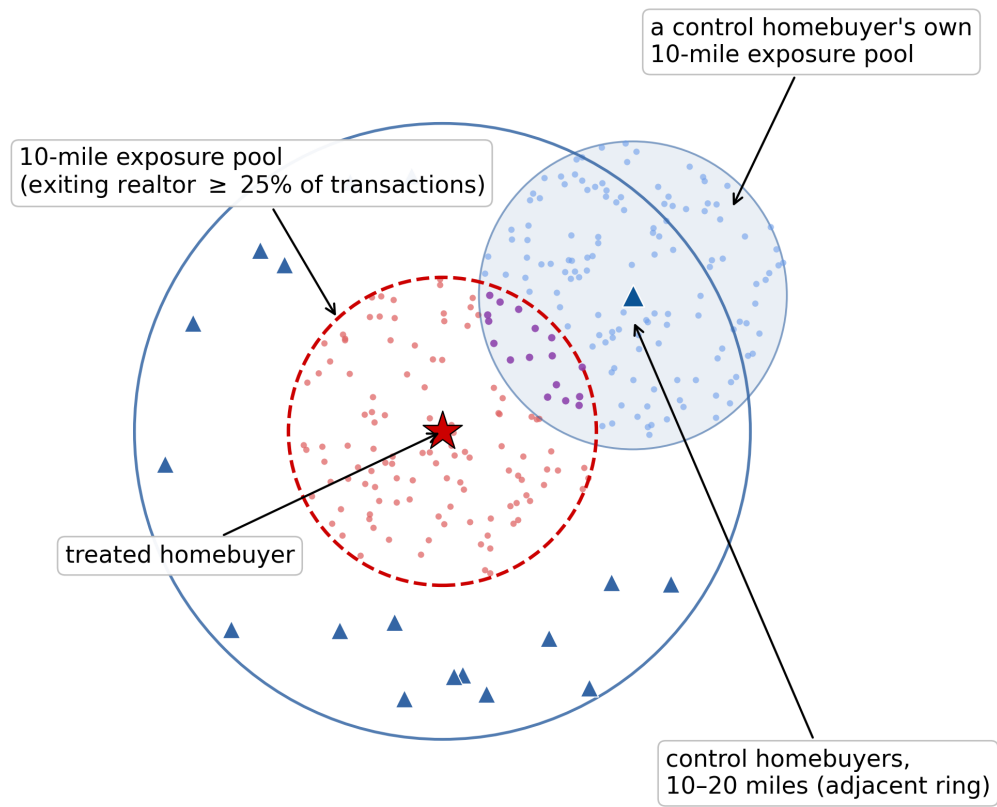
Note: This figure traces the construction of the analysis samples and the number of home purchases lost at each step. Additional detail on each step is described in Section A. Specifications that require FICO scores or credit-inquiry counts use the 1% credit-bureau match (15,390 purchases).

Figure A2: Event Studies around the CFPB Enforcement Order, Unweighted Controls



Note: This figure repeats Figure 6 without the matching weights of Table A13, so that every loan of treated and control realtors carries the same weight. Outcomes, sample, fixed effects, controls, reference period, error bars and clustering are as in Figure 6.

Figure A3: The Realtor-Exit Design: Exposure Pool and Adjacent Control Ring



Note: Schematic of the realtor-exit design of Appendix G around a single treated homebuyer (red star). Red dots mark the treated homebuyer's exposure-pool transactions, within 10 miles in the years before origination (dashed red circle); the exiting realtor(s) must account for at least 25% of this pool. Triangles mark control homebuyers, who lie 10–20 miles from the nearest same-year treated homebuyer (solid blue circle), the ring immediately adjacent to the exposure zone. One control homebuyer's own 10-mile exposure pool is highlighted and shaded blue; its exposure loans outside the overlap with the treated pool are shown as blue dots.

H Additional Tables

Table A1: Summary Statistics for Realtors and Loan Officers

	Markets (County*Year)	
	(1)	
	mean	s.d.
Observation	4,225	
No of Home Purchases	415	(1244)
No of Realtors	175	(405)
No of Loan Officers	176	(385)
No of Lenders	59	(86)

Note: This table reports summary statistics on local market structure for the 4,225 county-year markets that make up the core home-purchase sample. A “market” is defined as a county-year cell, and the underlying sample consists of the mortgage-financed home purchases retained in our main regression sample. For each market, we count the number of home purchases, the number of distinct realtors (buyer agents), the number of distinct loan officers, and the number of distinct lenders (the lending institutions that employ those officers) active in that market.

Table A2: **Robustness**—Causal Effect of Referral Lending on Mortgage Costs (IV: $CR4 - \widehat{CR4}$)

	Referral (First Stage)		Mortgage Rate Spread (off Prime Loans, in %)	
	1% Random Sample		1% Random Sample	
	(1)	(2)	2SLS (3)	2SLS (4)
$(CR4 - \widehat{CR4})$	1.256*** (0.015)	1.211*** (0.025)		
Referral ($CR4 > 0.4$, Top 4 Loan Officer)			0.171*** (0.011)	0.190*** (0.027)
Observations	1,751,287	15,390	1,751,287	15,390
R-squared	0.250	0.543		
Dep. Var. Mean	0.51	0.49	0.51	0.49
Year-Month*County FE	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓
Income Ratio Percentile 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			420321	2944
FS: Kleibergen-Paap rk F			7293	2306

Note: This table replicates Table 5 with the instrument constructed from $CR4$ instead of HHI . Same set of controls (FE) as in Table 3. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. “Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network ($CR4 > 0.4$) and the loan is financed by one of that realtor’s top four loan officers. It is instrumented with the “excess concentration” ($CR4 - \widehat{CR4}$) of a realtor’s loan officer network, which is the difference between observed $CR4$ concentration and predicted $CR4$ in the absence of referral lending. Standard errors are clustered at the county level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A3: Causal Effect of Referral Lending By Borrower Characteristics (IV: $\widehat{HHI}-\widehat{HHI}$; Full Estimation Sample)

	Mortgage Rate Spread (off Prime Loans, in %)	
	Full Estimation Sample	
	(Ethnicity) 2SLS (1)	(Race) 2SLS (2)
Referral*Hispanic	0.249*** (0.020)	
Referral*Non-Hispanic	0.114*** (0.008)	
Referral*White		0.153*** (0.011)
Referral*Black		0.202*** (0.019)
Referral*Asian		0.080*** (0.024)
Observations	1,751,287	1,751,287
Dep. Var. Mean	0.51	0.51
Full Controls	✓	✓
FS: Cragg-Donald Wald F	98396	59229
FS: Kleibergen-Paap rk F	406	293

Note: This table reports the full-sample counterpart to Table 11, which is estimated on the 1% random sample matched to credit records, for the two splits observable outside the credit-matched data: ethnicity in column (1) and race in column (2). The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. “Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network ($CR4 > 0.4$) and the loan is financed by one of that realtor’s top four loan officers. The endogenous variables are instrumented with the “excess concentration” ($\widehat{HHI}-\widehat{HHI}$) of a realtor’s loan officer network and its interaction with the demographic group indicators. In each column, the referral indicator is fully interacted with one mutually exclusive and exhaustive set of borrower groups, so each coefficient is the referral premium for that group relative to non-referred borrowers in the same group. The unclassified borrowers and the residual other-minority group are retained as their own categories, which are included in every regression but omitted from the table for readability. For unreported race and ethnicity in HMDA, we reclassify them using Bayesian Improved Surname Geocoding (BISG), as in Table 11. First-time home buyer status and education are observed only in the 1% random sample and have no full-sample counterpart. Same set of controls (FE) as in Table 3, without FICO bin fixed effects. Standard errors are clustered at the county level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A4: The Referral Premium by Price Component (OLS)

	Contract Rate Spread (p.p.)		Origination Fees (% of loan)		Discount Points (% of loan)	
		1% Random Sample		1% Random Sample		1% Random Sample
	(1)	(2)	(3)	(4)	(5)	(6)
Referral	0.124*** (0.005)	0.133*** (0.020)	0.093*** (0.004)	0.105*** (0.022)	-0.045*** (0.004)	-0.040** (0.018)
Referral × Superstar LO	-0.045*** (0.003)	-0.059** (0.024)	-0.027*** (0.004)	-0.042* (0.024)	0.026*** (0.004)	-0.032 (0.026)
Non-Compliant Referral	0.003 (0.003)	0.015 (0.016)	0.003 (0.002)	0.013 (0.014)	-0.009*** (0.002)	-0.013 (0.019)
Superstar LO	0.031*** (0.005)	0.038*** (0.013)	0.012*** (0.003)	0.024* (0.014)	-0.037*** (0.004)	0.028 (0.022)
Observations	1,735,525	15,161	1,735,525	15,161	1,735,525	15,161
R-squared	0.321	0.649	0.138	0.492	0.146	0.485
Dep. Var. Mean	0.426	0.406	0.491	0.453	0.318	0.296
Year-Month*County FE	✓	✓	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓	✓	✓
Income Ratio Percentile 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 decomposes the referral premium into the three components of the price of a mortgage, estimated by OLS. Each component spans a pair of columns. The outcome in columns (1)–(2) it is the reported mortgage contract rate over the average prime offer rate (APOR), in percentage points; in columns (3)–(4) it is HMDA origination charges net of discount points, as a percentage of the loan amount; and in columns (5)–(6) it is HMDA discount points, also as a percentage of the loan amount. Within each pair the first column is the full estimation sample and the second the 1% credit-bureau matched subsample with FICO bin fixed effects, matching specifications in columns (3) and (6) of Table 4. “Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network ($CR4 > 0.4$) and the loan is financed by one of that realtor’s top four loan officers; “Non-Compliant Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network ($CR4 > 0.4$) and the loan is not financed by one of that realtor’s top four loan officers. “Superstar LO” equals one if the loan officer is in the top 5% of loan officers by purchase-loan volume. Within each sample all three columns are estimated on one common set of loans with all three components observed. The rate column and the two upfront columns are in different units and are not additive; the text converts the upfront components to a rate equivalent using the paper’s points-to-rate factor of 0.218. Standard errors are clustered at the county level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A5: The Referral Premium by Price Component (IV: $\widehat{HHI}-\widehat{HHI}$)

	Contract Rate Spread (p.p.)		Origination Fees (% of loan)		Discount Points (% of loan)	
		1% Random Sample		1% Random Sample		1% Random Sample
	(1)	(2)	(3)	(4)	(5)	(6)
Referral	0.142*** (0.010)	0.164*** (0.036)	0.109*** (0.007)	0.110*** (0.030)	-0.022*** (0.008)	-0.071** (0.031)
Observations	1,735,525	15,161	1,735,525	15,161	1,735,525	15,161
R-squared	0.006	0.004	0.004	0.005	0.001	0.001
Dep. Var. Mean	0.426	0.406	0.491	0.453	0.318	0.296
First-stage KP rk F	1668	740.9	1668	740.9	1668	740.9
Year-Month*County FE	✓	✓	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓	✓	✓
Income Ratio Percentile 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 repeats the decomposition of Table A4 by two-stage least squares, instrumenting “Referral” with the “excess concentration” ($\widehat{HHI}-\widehat{HHI}$) of a realtor’s loan officer network as in Table 5. Outcomes, samples, controls and the referral definition are as in Table A4, except that the non-compliant referral and superstar terms are not included here: columns (1)–(2) report the reported mortgage contract rate over the average prime offer rate (APOR), in percentage points; columns (3)–(4) HMDA origination charges net of discount points, as a percentage of the loan amount; and columns (5)–(6) HMDA discount points, also as a percentage of the loan amount. Within each pair the first column is the full estimation sample and the second the 1% credit-bureau matched subsample with FICO bin fixed effects, matching specifications in columns (3) and (6) of Table 5. “Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network ($CR4 > 0.4$) and the loan is financed by one of that realtor’s top four loan officers; “Non-Compliant Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network ($CR4 > 0.4$) and the loan is not financed by one of that realtor’s top four loan officers. “Superstar LO” equals one if the loan officer is in the top 5% of loan officers by purchase-loan volume. Within each sample all three columns are estimated on one common set of loans with all three components observed. The rate column and the two upfront columns are in different units and are not additive; the text converts the upfront components to a rate equivalent using the paper’s points-to-rate factor of 0.218. Standard errors are clustered at the county level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A6: **Robustness**—Effect of Referral Lending on Mortgage Costs (OLS; Realtors with ≥ 5 Transactions)

	Mortgage Rate Spread (off Prime Loans, in %)					
	Full Sample			1% Random Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Referral	0.106***	0.104***	0.123***	0.096***	0.104***	0.127***
(CR4 >0.4, Top 4 LO)	(0.004)	(0.005)	(0.005)	(0.012)	(0.017)	(0.018)
Non-Compliant Referral		-0.003	-0.000		0.013	0.016
(CR4 >0.4, Not Top 4 LO)		(0.003)	(0.003)		(0.014)	(0.015)
Referral $\times$ Superstar LO			-0.039***			-0.048**
			(0.003)			(0.020)
Superstar LO			0.029***			0.031**
			(0.005)			(0.014)
Observations	2,276,132	2,276,132	2,276,132	21,507	21,507	21,507
R-squared	0.314	0.314	0.315	0.619	0.619	0.619
Dep. Var. Mean	0.51	0.51	0.51	0.48	0.48	0.48
Year-Month*County FE	✓	✓	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓	✓	✓
Income Ratio Percentile 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 replicates Table 4 on the broader sample of realtors with at least five purchase transactions over the study period; the main analysis restricts to realtors with at least ten. Columns (1)–(3) use the full sample and columns (4)–(6) the 1% random sample matched to credit records (which adds FICO bin fixed effects); within each block, the third column adds the superstar loan-officer interaction. Same set of controls (FE) as in Table 3. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. “Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network (CR4 >0.4) and the loan is financed by one of that realtor’s top four loan officers; “Non-Compliant Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network (CR4 >0.4) and the loan is not financed by one of that realtor’s top four loan officers. “Superstar LO” equals one if the loan officer is in the top 5% of loan officers by purchase-loan volume; “Referral $\times$ Superstar LO” is the differential referral premium when the referral loan officer is a superstar. Standard errors are clustered at the county level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: **Robustness**—Causal Effect of Referral Lending on Mortgage Costs (IV: $HHI - \widehat{HHI}$; Realtors with ≥ 5 Transactions)

	Referral (First Stage)		Mortgage Rate Spread (off Prime Loans, in %)	
	1% Random Sample		1% Random Sample	
	(1)	(2)	2SLS (3)	2SLS (4)
$(HHI - \widehat{HHI})/10000$	1.265*** (0.028)	1.230*** (0.033)		
Referral ($CR4 > 0.4$, Top 4 Loan Officer)			0.159*** (0.010)	0.181*** (0.037)
Observations	2,276,132	21,507	2,276,132	21,507
R-squared	0.176	0.471		
Dep. Var. Mean	0.51	0.48	0.51	0.48
Year-Month*County FE	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓
Income Ratio Percentile 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			312774	2538
FS: Kleibergen-Paap rk F			2039	1369

Note: This table replicates Table 5 on the broader sample of realtors with at least five purchase transactions over the study period; the main analysis restricts to realtors with at least ten. Same set of controls (FE) as in Table 3. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. “Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network ($CR4 > 0.4$) and the loan is financed by one of that realtor’s top four loan officers. It is instrumented with the “excess concentration” ($HHI - \widehat{HHI}$) of a realtor’s loan officer network, which is the difference between observed HHI concentration and predicted HHI in the absence of referral lending. Standard errors are clustered at the county level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A8: **Robustness**—Effect of Referral Lending on Mortgage Costs (OLS; Referral = Realtor’s Top 1 Loan Officer)

	Mortgage Rate Spread (off Prime Loans, in %)					
	Full Sample			1% Random Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Referral (CR4 >0.4, Top 1 LO)	0.078*** (0.003)	0.105*** (0.005)	0.138*** (0.005)	0.062*** (0.014)	0.100*** (0.020)	0.134*** (0.029)
Non-Compliant Referral (CR4 >0.4, Not Top 1 LO)		0.039*** (0.004)	0.040*** (0.004)		0.052*** (0.018)	0.052*** (0.018)
Referral × Superstar LO			-0.059*** (0.004)			-0.064* (0.033)
Superstar LO			0.034*** (0.006)			0.035*** (0.012)
Observations	1,751,287	1,751,287	1,751,287	15,390	15,390	15,390
R-squared	0.317	0.317	0.318	0.645	0.645	0.645
Dep. Var. Mean	0.51	0.51	0.51	0.49	0.49	0.49
Year-Month*County FE	✓	✓	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓	✓	✓
Income Ratio Percentile 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 replicates Table 4 using a stricter referral definition: a loan counts as a referral only if it is originated by the realtor’s *single most-used* (rank-one, “top 1”) loan officer, rather than any of the realtor’s top four. Columns (1)–(3) use the full sample and columns (4)–(6) the 1% random sample matched to credit records (which adds FICO bin fixed effects); within each block, the third column adds the superstar loan-officer interaction. “Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network (CR4 >0.4) and the loan is financed by that realtor’s top-1 loan officer; “Non-Compliant Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network (CR4 >0.4) and the loan is not financed by that realtor’s top-1 loan officer. “Superstar LO” equals one if the loan officer is in the top 5% of loan officers by purchase-loan volume; “Referral × Superstar LO” is the differential referral premium when the referral loan officer is a superstar. Same set of controls (FE) as in Table 3. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. Standard errors are clustered at the county level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A9: **Robustness**—Causal Effect of Referral Lending on Mortgage Costs (IV: $HHI-\widehat{HHI}$; Referral = Realtor's Top 1 Loan Officer)

	Referral (First Stage)		Mortgage Rate Spread (off Prime Loans, in %)	
	1% Random Sample		1% Random Sample	
	(1)	(2)	2SLS (3)	2SLS (4)
$(HHI-\widehat{HHI})/10000$	1.447*** (0.020)	1.364*** (0.034)		
Referral ($CR4 > 0.4$, Top 1 Loan Officer)			0.151*** (0.010)	0.184*** (0.039)
Observations	1,751,287	15,390	1,751,287	15,390
R-squared	0.210	0.509		
Dep. Var. Mean	0.51	0.49	0.51	0.49
Year-Month*County FE	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓
Income Ratio Percentile 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			353385	2372
FS: Kleibergen-Paap rk F			5258	1582

Note: This table replicates Table 5 using a stricter referral definition: a loan counts as a referral only if it is originated by the realtor's *single most-used* (rank-one) loan officer, rather than any of the realtor's top four. "Referral" equals one if the borrower's realtor has a medium-to-high-concentration network ($CR4 > 0.4$) and the loan is financed by that realtor's top-1 loan officer. It is instrumented with the "excess concentration" ($HHI-\widehat{HHI}$) of a realtor's loan officer network, which is the difference between observed HHI concentration and predicted HHI in the absence of referral lending. Same set of controls (FE) as in Table 3. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. Standard errors are clustered at the county level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A10: Summary Statistics for Lender Choice Model Estimation Sample

Var	Obs	Mean	SD	Min	P50	Max
#1 Lender in Market	94,510	0.065	0.247	0.000	0.000	1.000
#2 Lender in Market	94,510	0.043	0.204	0.000	0.000	1.000
#3 Lender in Market	94,510	0.036	0.186	0.000	0.000	1.000
Fintech	94,510	0.202	0.401	0.000	0.000	1.000
Bank	93,956	0.364	0.481	0.000	0.000	1.000
Out-of-State	94,510	0.187	0.390	0.000	0.000	1.000
FICO	94,510	760.068	60.962	446.000	776.000	850.000
Combined LTV	87,654	65.579	17.358	0.810	68.077	191.030
DTI Ratio	84,354	30.898	12.269	10.000	30.000	60.000
Distance Distance>0	91,566	27.773	56.442	0.006	7.836	668.887

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

Table A11: Lender Choice Model Estimates

	(1) Baseline (drops same-LO refis)	(2) Keeping same-LO refis
Bank	-3.637*** (0.139)	-3.385*** (0.134)
Bank x Combined LTV	-1.620*** (0.080)	-1.813*** (0.068)
Bank x DTI	-0.831*** (0.078)	-0.792*** (0.075)
Bank x FICO	5.741*** (0.175)	5.617*** (0.166)
Distance	-3.230*** (0.210)	-3.330*** (0.203)
Fintech	-1.489*** (0.164)	-1.068*** (0.175)
Fintech x Combined LTV	-0.537*** (0.068)	-0.900*** (0.062)
Fintech x DTI	0.558*** (0.119)	0.619*** (0.120)
Fintech x FICO	1.149*** (0.188)	0.857*** (0.191)
Out-of-State	-1.330*** (0.080)	-1.350*** (0.078)
Top Lender #1	0.260 (0.289)	0.014 (0.291)
Top Lender #1 x Combined LTV	-0.242** (0.110)	0.050 (0.110)
Top Lender #1 x DTI	0.064 (0.152)	-0.075 (0.142)
Top Lender #1 x FICO	0.075 (0.330)	0.222 (0.316)
Top Lender #2	-0.220 (0.357)	-0.563* (0.334)
Top Lender #2 x Combined LTV	-0.371*** (0.136)	-0.004 (0.126)
Top Lender #2 x DTI	-0.140 (0.197)	-0.187 (0.163)
Top Lender #2 x FICO	0.625 (0.393)	0.791** (0.357)
Top Lender #3	-0.338 (0.409)	-0.560 (0.369)
Top Lender #3 x Combined LTV	-0.469*** (0.149)	-0.225 (0.149)
Top Lender #3 x DTI	0.241 (0.184)	0.223 (0.178)
Top Lender #3 x FICO	0.735 (0.451)	0.818** (0.410)
Market FE (Second Stage)	Yes	Yes
Observations	84,205	95,394

Note: This table reports the estimates of the lender choice model (β_1, β_2) in Equation (5), estimated by maximum likelihood on the refinance-mortgage sample described in Table A10, with county-year market fixed effects. Column (1) is the baseline, estimated after dropping every refinance that returns to the same loan officer who originated the borrower's original purchase mortgage (see Appendix D); these "relationship" refinances are a potential source of contamination because the lender choice reflects a pre-existing match rather than a fresh search. Column (2) re-estimates the model keeping them. The two sets of estimates are nearly identical. Bootstrapped standard errors are reported in parentheses and are estimated by resampling markets with replacement 200 times. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A12: **Robustness**—Causal Effect of Referral Lending on Mortgage Costs (IV: $\widehat{HHI}-HHI$; Keeping Relationship Refinances in the Lender-Choice Model)

	Referral (First Stage)		Mortgage Rate Spread (off Prime Loans, in %)	
	1% Random Sample		1% Random Sample	
	(1)	(2)	2SLS (3)	2SLS (4)
$(\widehat{HHI}-HHI)/10000$	1.451*** (0.036)	1.414*** (0.052)		
Referral ($CR4 > 0.4$, Top 4 Loan Officer)			0.151*** (0.010)	0.178*** (0.037)
Observations	1,751,287	15,390	1,751,287	15,390
R-squared	0.205	0.521		
Dep. Var. Mean	0.51	0.49	0.51	0.49
Year-Month*County FE	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓
Income Ratio Percentile 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			297974	2225
FS: Kleibergen-Paap rk F			1643	752

Note: This table replicates the headline 2SLS of Table 5 on the same purchase-mortgage sample, but instruments referral lending with the “excess concentration” ($\widehat{HHI}-HHI$) constructed from a lender-choice model estimated on all refinances, including those that return to the borrower’s original purchase loan officer, which the baseline instrument of Table 5 excludes (see Appendix D.6). Only the instrument changes; the regression sample and controls are identical to Table 5, so the estimates are directly comparable. “Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network ($CR4 > 0.4$) and the loan is financed by one of that realtor’s top four loan officers. Same set of controls (FE) as in Table 3. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. Standard errors are clustered at the county level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A13: Pre-Order Characteristics of Treated and Control Realtors, Before and After Matching

	Treated mean	Control mean		Standardized difference	
		unweighted	matched	unweighted	matched
	(1)	(2)	(3)	(4)	(5)
Purchase loans, 2018	3.0	4.1	3.0	-0.27	-0.00
Purchase loans, 2019	3.7	4.7	3.7	-0.25	0.00
Purchase loans, 2020	4.0	5.2	4.0	-0.28	0.00
Purchase loans, 2021	4.2	5.4	4.2	-0.29	0.00
Purchase loans, 2022	2.3	3.2	2.3	-0.31	0.00
Loan amount (\$1,000)	408.6	465.1	408.6	-0.35	0.00
LTV (%)	86.1	83.1	86.1	0.51	-0.00
DTI (%)	38.5	35.5	38.5	0.80	0.00
Income ratio	1.3	1.8	1.3	-0.90	0.00
Joint application	0.537	0.510	0.537	0.17	-0.00
Conforming	0.986	0.903	0.986	0.71	-0.00

Note: This table reports realtor-level pre-order characteristics of the treated and control realtors of Section 4.3: the number of purchase loans closed in each year from 2018 to 2022, and the means over the realtor's pre-order loans of the loan amount, loan-to-value ratio, debt-to-income ratio, income ratio (reported income over the county's median income of the same year), joint-application indicator and conforming-loan indicator. Columns (1) and (2) report unweighted means for treated and control realtors. Column (3) reports the control mean after matching by entropy balancing, which reweights control realtors so that the weighted means of every characteristic in the table and of the shares of the 20 treated counties equal the treated means (Appendix F.2). Columns (4) and (5) report the standardized difference, the difference in means divided by the pooled standard deviation across realtors, before and after matching.

Table A14: Causal Effect of Referral Lending on Mortgage Costs (IV: CFPB Enforcement Order), Unweighted Controls

	Referral (First Stage)	Mortgage Rate Spread (off Prime Loans, in %)		Days to Close
			DTI ≤ 36%, LTV ≤ 80%, Income > Median	
	(1)	2SLS (2)	2SLS (3)	2SLS (4)
Freedom-tied Realtor*Post (IV)	-0.278*** (0.037)			
Referral (CR4 > 0.4, Top 4 Loan Officer, pre-order)		0.355** (0.154)		4.840 (3.432)
Referral × Low-Risk Borrowers			0.796*** (0.147)	
Referral × Other Borrowers			0.271 (0.174)	
Observations	54,196	54,196	54,196	53,964
Dep. Var. Mean	0.06	0.32	0.32	47.5
Share of Low-Risk Borrowers			0.206	
Year-Month*County FE	✓	✓	✓	✓
Realtor FE	✓	✓	✓	✓
Loan Amount Bin FE	✓	✓	✓	✓
Age Bin FE	✓	✓	✓	✓
Income Ratio Percentile FE	✓	✓	✓	✓
Joint Application	✓	✓	✓	✓
Conforming FE	✓	✓	✓	✓
DTI Bin FE	✓	✓	✓	✓
LTV Bin FE	✓	✓	✓	✓
Bedroom FE	✓	✓	✓	✓
Bathroom FE	✓	✓	✓	✓
Sqft Bin FE	✓	✓	✓	✓
House Age FE	✓	✓	✓	✓
ZIP Code FE	✓	✓	✓	✓
FS: Cragg-Donald Wald F		1407	455	1388
FS: Kleibergen-Paap rk F		58	9.7	56

Note: This table repeats the fully controlled specifications of the CFPB enforcement-order design without the matching weights of Table A13, so that every loan of treated and control realtors carries the same weight. Column (1) is the first stage and column (2) the 2SLS estimate for the mortgage APR spread, as in columns (2) and (4) of Table 6; column (3) repeats the pooled specification of column (4) of Table 8, with “Referral” interacted with an indicator for low-risk borrowers ($DTI \leq 36\%$, $LTV \leq 80\%$ and income above the county median) and with its complement, so that the two coefficients are the premium paid by low-risk borrowers and by all other borrowers; column (4) is the 2SLS estimate for days from contract to closing, as in column (3) of Table 7. Sample, instrument, endogenous variable, fixed effects, controls and clustering are as in those tables. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A15: Realtor-Exit Design: Covariate Balance between Treated and Control Homebuyers

	Treated Mean	Control Mean	Raw Difference	Difference within County×Month
	(1)	(2)	(3)	(4)
Loan Amount (\$k)	271.8	460.4	-188.6** ($t = -2.27$)	+3.2 ($t = 0.03$)
DTI (%)	32.9	36.3	-3.37 ($t = -1.30$)	-2.99 ($t = -1.56$)
LTV (%)	86.7	86.0	+0.68 ($t = 0.68$)	+1.50 ($t = 0.71$)
Income Ratio	1.81	1.89	-0.08 ($t = -1.14$)	+0.28 ($t = 0.59$)
Joint Application	0.44	0.51	-0.065** ($t = -2.23$)	-0.055 ($t = -1.00$)
Conforming	0.98	0.95	+0.023 ($t = 1.31$)	-0.079 ($t = -0.81$)
House Age (years)	52.8	34.6	+18.2*** ($t = 3.79$)	+9.03 ($t = 1.48$)
Bedrooms	3.47	3.51	-0.04 ($t = -0.38$)	-0.08 ($t = -0.34$)
Bathrooms	2.47	2.83	-0.36*** ($t = -2.87$)	-0.02 ($t = -0.06$)

Note: This table compares treated and control homebuyers in the realtor-exit design of Appendix G. Treated homebuyers experienced a qualifying realtor exit within the previous two years, with the exiting realtor(s) accounting for at least 25% of purchase transactions within 10 miles of the property in the year before the exit; controls lie 10–20 miles from the nearest same-year treated homebuyer in the same county. Column (3) reports the coefficient on the treatment indicator from a regression of each covariate on treatment alone; column (4) adds county $\times$ year-month fixed effects, matching the within-cell comparison used in estimation. The income ratio is defined as the reported income over the county's median income of the same year. Standard errors are clustered at the county level; t -statistics are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A16: Effect of Referral Lending (IV: Local Realtor-Exit Exposure)

	Referral	Mortgage Rate Spread (%)
	Full Sample First Stage (1)	Full Sample 2SLS (2)
Local Exit Exposure	-0.181*** (0.033)	
Referral		0.305** (0.150)
Observations	5,703	5,703
R-squared	0.117	
Dep. Var. Mean	0.61	0.36
Year-Month*County FE	✓	✓
Loan Amount Bin FE	✓	✓
Age Bin FE	✓	✓
Income Ratio Percentile FE	✓	✓
Joint Application	✓	✓
Conforming FE	✓	✓
DTI Bin FE	✓	✓
LTV Bin FE	✓	✓
Bedroom FE	✓	✓
Bathroom FE	✓	✓
Sqft Bin FE	✓	✓
House Age FE	✓	✓
ZIP Code FE	✓	✓
FS: Cragg-Donald Wald F		1.8
FS: Kleibergen-Paap rk F		31

Note: This table reports the realtor-exit design of Appendix G. Local Exit Exposure equals one if at least 25% of purchase transactions within 10 miles of the property, one or two years before origination, were written by referral-active realtors who subsequently exited sharply. Controls are homebuyers 10–20 miles from the nearest same-year treated homebuyer in the same county. “Referral” equals one if the borrower’s realtor has a medium-to-high-concentration network ($CR4 > 0.4$) and the loan is financed by one of that realtor’s top four loan officers. Column (1) is the first stage and column (2) instruments “Referral” with local exit exposure. The mortgage costs are measured by mortgage APR spreads off the benchmark rate offered on prime mortgage loans of a comparable type. Same set of controls (FE) as in Table 3. Standard errors are clustered at the county level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.