

Do Mortgage Brokers Help Borrowers Get Better Rates?*

Zirui Wang

McCombs School of Business, University of Texas at Austin

October 29, 2026

[Link to the latest version](#)

Abstract

Mortgage brokers promise to shop across lenders, yet rates on brokered loans are as dispersed as those on retail loans. Using novel U.S. mortgage data linking borrowers, brokers, and lenders, I study why similar borrowers in the same local market on the same day pay widely different rates despite using mortgage brokers. Broker firms exhibit persistent differences in the rates their borrowers pay. An IV design based on borrowers' relative proximity to broker branches estimates that using a high-cost broker raises rate spreads by 10–12 basis points. Lender choice helps explain these differences: some firms systematically choose lenders that are relatively expensive and pay brokers more, consistent with steering. Steering practices differ across firms and persist within firms: borrowers who use firms with stronger past steering practices pay higher rates. Overall, the findings show how conflicts of interest can undermine the benefits of delegated search and raise household borrowing costs.

JEL classification: G14, G21, G24, G51

Keywords: Mortgage Brokers, Price Dispersion, Steering, Conflicts of Interest

*I am deeply grateful to my dissertation committee, John M. Griffin, Samuel Kruger, and Caitlin Gorback, for their invaluable guidance and relentless support. I also thank Andres Almazan, Aydogan Altı, Susan Cherry, Jonathan Cohn, Andrew Fieldhouse, Cesare Fracassi, William Fuchs, John Hatfield, Erik Heitfield, Minjoo Kim, Eric Lee, Richard Lowery, Prateek Mahajan, Kevin Mei, Daniel Neuhann, Andrey Ordin, Aaron Pancost, Amiyatosh Purnanandam, Marius Ring, Nikolai Roussanov, Max Sacher, Lukas Schmid, Clemens Sialm, Shikhar Singla, Michael Sockin, Laura Starks, Luke Taylor, Sheridan Titman, Nancy Wallace, Toni Whited, Mindy Xiaolan, Liyan Yang, David Zhang, Xinran Zhang, Yiyuan Zhang (discussant), William Zheng (discussant), Qifei Zhu, and seminar and conference participants at the University of Texas at Austin and the Texas Finance Job Candidate Conference at Texas A&M University for their helpful comments. All errors are my own. Email: Zirui.Wang@mcombs.utexas.edu.

Mortgages are the largest liability on household balance sheets, yet borrowers with similar characteristics often pay very different interest rates. Search frictions help explain these differences: half of mortgage borrowers seriously consider only one lender, and those who know less about mortgages and shop less pay more (Stigler, 1961; Burdett and Judd, 1983; Bhutta, Fuster, and Hizmo, 2026). Mortgage brokers promise to address this problem by shopping across lenders on borrowers' behalf. Surprisingly, however, rates on brokered loans are no less dispersed than those on retail loans.¹ Similar borrowers in the same local market on the same day still pay rates that differ by tens of basis points. Why do such differences persist when a market expert searches on the borrower's behalf? Do brokers differ in the rates they deliver, and if so, why?

In this paper, I study why mortgage rates remain dispersed even when borrowers use mortgage brokers. If brokers are all the same and shop for lower rates on borrowers' behalf, similar borrowers should receive the same rate regardless of which broker they use, reducing rate dispersion. Alternatively, brokers differ in pricing and some broker firms may consistently deliver higher rates because they choose more expensive lenders, particularly when those lenders pay the broker more. Understanding brokers' role in this dispersion is challenging because existing data sources do not jointly identify the borrower, broker, and lender on the same loan together with borrower characteristics, mortgage prices, and broker compensation. I address this challenge by constructing a novel dataset that combines Optimal Blue rate locks, CoreLogic mortgage records, and Home Mortgage Disclosure Act (HMDA) records, allowing me to compare the prices similar borrowers receive across brokers and the incentives behind brokers' lender choices.

My analysis delivers four primary findings. First, similar borrowers pay widely different rates across broker firms, and these differences persist across periods and loan types. Using a high-cost broker raises rate spreads by 10–12 basis points in estimates based on borrowers' relative proximity to broker branches. Second, some firms repeatedly choose lenders that are expensive for their borrowers and pay the broker more, consistent with steering. Third, steering practices vary across firms and persist within firms: brokers that chose expensive lenders in the past continue to do so, and borrowers

¹Figure 1 summarizes brokers' advertising and compares the number of lenders that brokers advertise with the number they use, and Figure 2 compares residual rate dispersion in the two channels.

who use firms with stronger past steering practices pay higher rates. Finally, costly lender choices mainly reflect the use of lenders that are expensive in general, rather than poor borrower-lender fit, and the selection of expensive lenders from within brokers' existing menus.

I begin by documenting large and persistent differences in broker pricing. Among the largest firms, similar borrowers receive rates ranging from about 41 basis points below the retail benchmark to 32 basis points above it, after accounting for borrower characteristics, county, and the date the rate was locked. A firm's pricing before 2022 predicts its pricing from 2022 on, and firms that are expensive for purchase borrowers are also expensive for refinancing borrowers and across credit-score groups. Broker identity explains more of the remaining variation in brokered rates than lender identity does. Thus, similar average rates in the brokered and retail channels conceal substantial differences across the intermediaries that borrowers use.

A natural concern is that these differences reflect the borrowers rather than the brokers. High-cost firms may attract borrowers who are riskier in ways that lenders observe but I do not. To address this sorting concern, I instrument for using a high-cost broker with the borrower's relative distance to the nearest high-cost and low-cost broker branches. Firms are classified using their pricing in the previous year, so the current borrower's loan does not determine the classification. Distances are measured from where borrowers search for their mortgages, using previous home addresses for purchase borrowers. The identifying intuition is that the convenience of a nearby branch can influence broker choice independently of the borrower's credit risk: two borrowers may have similarly good access to brokers but differ in whether the closest branches belong to high-cost or low-cost firms. I therefore compare relative proximity while controlling for general access to branches, borrower risk, neighborhood characteristics, and market conditions at both the search origin and the mortgaged property.

Borrowers are more likely to use a high-cost firm when a high-cost branch is relatively closer, providing a strong first stage. The resulting IV estimates imply that using a high-cost firm raises rate spreads by 10–12 basis points, close to the OLS estimates and about half the average spread in the IV sample. The estimates capture the effect for borrowers whose broker choice responds to relative proximity. Their causal interpretation requires that, conditional on the controls, proximity to a

particular type of broker affect the mortgage rate through broker choice rather than through other determinants of borrowing costs. A potential violation of this exclusion restriction is that high-cost branches locate in neighborhoods whose residents would pay higher rates regardless of the broker they use. Balance tests show economically small differences in borrower and neighborhood characteristics associated with relative proximity. The chosen brokers' website claims, used as proxies for advertised service quality, are also relatively balanced, suggesting that proximity does not systematically direct borrowers toward firms advertising different services. These comparisons mitigate concerns about neighborhood sorting and differences in broker quality.

Higher rates might reflect better service. A borrower may be willing to pay more for a broker that can close a loan quickly. The linked rate-lock and mortgage records provide lock dates, closing dates, and requested lock periods, allowing me to test this alternative explanation directly. Using the same instrument, I find no statistically significant differences between borrowers at high-cost and low-cost firms in either the time from rate lock to closing or the requested lock period. Unobserved borrower characteristics may nevertheless affect estimates of counterfactual lender prices, a concern I directly address later.

Why do some broker firms deliver higher rates? One reason could be their choice of lenders. Brokers advertise that they shop across lenders to find borrowers a lower rate, making lender selection a natural place to examine what high-cost firms do differently. I therefore estimate a lender pricing model that predicts each lender's rate for the same observed borrower characteristics, accounting for broker-lender relationships and local market conditions, and compare the selected lender with alternatives in the state. The lender choices reveal estimated overpayment relative to alternatives in the market, averaging 27.7 basis points above the lowest predicted price.

A concern is that lender quotes observed by a broker may reflect pricing information not captured by the estimated lender pricing model. An honest broker who selects the lowest quote can therefore generate a positive observed gap if the selected lender does not have the lowest predicted price. I address this concern by simulating honest shopping with differences between quoted and predicted prices drawn from a distribution calibrated to out-of-sample prediction errors. Estimated overpay-

ment remains positive after this adjustment, averaging 11.9 basis points; I call this difference the *steering gap* and use it as a proxy for steering practices. Larger gaps correspond to higher rates actually paid by borrowers: within the same firm, a 10 basis point larger gap is associated with a roughly 7.3 basis point higher realized rate spread. The measure therefore captures costly lender choices that are reflected in borrowers' financing costs.

These practices differ markedly across broker firms. Some of the largest firms have average steering gaps close to zero, while others have gaps of 15–20 basis points. The differences also persist over time. Firms that used relatively expensive lenders in the previous year tend to continue doing so for new borrowers, including when comparing the same firm at different points in time. A 10 basis point increase in a firm's past gap is associated with a 4.6 basis point higher gap on its current loans. The pattern points to persistent differences in firms' lender-choice practices, beyond occasional expensive choices on individual loans.

Those differences matter for a borrower choosing a broker. Borrowers who use a firm with a higher past steering gap pay higher rates themselves, even after accounting for their characteristics and local market conditions. Within the same firm, a one-standard-deviation increase in its past gap is associated with a roughly 7 basis point higher rate spread for subsequent borrowers. Because the firm's history uses earlier customers' loans, it characterizes the broker before the focal borrower arrives. A firm's established lender-choice practices are therefore informative about the mortgage price its next customer is likely to receive, helping explain why pricing differences across brokers persist.

To understand why these lender choices are expensive, I examine two distinctions. First, a broker may repeatedly use lenders that are expensive in general, or choose lenders that are competitive for many borrowers but a poor match for a particular borrower's characteristics. Separating these sources shows that poor borrower-lender fit accounts for about a fifth of the average steering gap but almost none of its relationship with subsequent borrower rates; that relationship mainly reflects generally expensive lenders and the prices they offer through the broker's relationships. Second, an expensive choice could reflect a costly menu of lenders or the selection of an expensive lender when cheaper options exist on the same menu. Distinguishing these channels matters because expanding lender

access would not by itself correct poor choices among existing partners. Choices within the broker's menu account for roughly three-quarters of the average steering gap and four-fifths of the relationship between past gaps and subsequent rates. Thus, most of the cost comes from which lenders brokers select among their existing partners, although the composition of their menus also matters.

Why would brokers repeatedly use these lenders? The compensation results point to a conflict of interest. Lender choices associated with higher borrower costs also pay the broker more, even among loans originated by the same firm. Compensation can differ across a broker's lender agreements, so the lender that offers a worse price to the borrower may offer a better payment to the broker. This evidence connects costly lender choices to the broker's financial incentive and strengthens the steering interpretation of the proxy.

Overall, the findings have three implications. First, delegated search does not necessarily reduce price dispersion. When lenders pay the intermediary that shops on borrowers' behalf, conflicts of interest can undermine the benefits of shopping. Second, choosing a broker matters as well as choosing a lender: broker identity explains at least as much variation in brokered rates as lender identity, and brokers' past lender choices predict the rates their future borrowers pay. Third, current regulation leaves room for steering across lenders. Regulation Z restricts compensation tied to mortgage terms within a lender but allows compensation to differ across unaffiliated lenders, while its anti-steering safe harbor permits qualifying comparisons among a subset of the lenders with which a broker regularly does business. Greater transparency about lender-paid compensation and brokers' past lender choices could help borrowers and regulators identify costly practices.

Related Literature. This paper contributes to three strands of literature. First, it contributes to research on mortgage intermediaries and borrowing costs. Much of this work compares brokered and retail loans or studies competition within the broker channel. Brokered loans can be more expensive, particularly for borrowers with lower incomes and credit scores ([LaCour-Little, 2009](#)), and borrowers who compare too few brokers overpay ([Woodward and Hall, 2012](#)). Studies of broker profits and competition document market power and lower origination fees where brokers face more competition ([Berndt, Hollifield, and Sandås, 2010](#); [Ambrose and Conklin, 2014](#)). Other work examines brokered

refinances, licensing, and racial differences in fees (Spader and Quercia, 2011; Agarwal, Ang, Deng, and Wang, 2021; Ambrose, Conklin, and Lopez, 2021). Outside the United States, brokers can increase lender competition while commissions distort advice, and borrower preferences help explain product choices (Robles-Garcia, 2020; Allen, Clark, Houde, Li, and Trubnikova, 2025).² I contribute by studying differences across U.S. broker firms using data that identify the borrower, broker firm, loan officer, and lender on the same loan. I document persistent differences in pricing across firms and estimate the borrowing-cost effect of using a high-cost firm. The distinction matters because brokers that deliver lower rates and brokers that deliver higher rates can coexist even when the channel's average price is close to retail.

Second, the paper contributes to research on conflicts of interest and steering by intermediaries. Seller-paid compensation can influence the advice intermediaries give their customers (Inderst and Ottaviani, 2012). Evidence from mutual funds, convertible bonds, annuities, and life insurance links intermediary compensation to product recommendations and customer outcomes (Christoffersen, Evans, and Musto, 2013; Bergstresser, Chalmers, and Tufano, 2009; Egan, 2019; Egan, Ge, and Tang, 2022; Anagol, Cole, and Sarkar, 2017). In employer-sponsored health insurance, commissions lead brokers to steer employers toward insurers that pay more, and they also shape which insurers brokers contract with and therefore which insurers compete for an employer's business (Li, Liu, LoSasso, and Tilipman, 2026). Real estate agents also respond to commissions when directing buyers toward properties (Barwick, Pathak, and Wong, 2017; Barry, Fried, and Hatfield, 2025), and brokered markets can sustain high commissions even with low concentration and easy entry (Hatfield, Kominers, and Lowery, 2025). In lending, prior work studies steering across loan products (Guiso, Pozzi, Tsoy, Gambacorta, and Mistrulli, 2022; Agarwal, Amromin, Ben-David, and Evanoff, 2025; Agarwal, Ambrose, and Yao, 2020; Benetton, Buchak, and Robles-Garcia, 2022) and real estate agents' influence on lender choice (Lopez, McCoy, and Sah, 2022; Zhang, Barwick, Han, and Kroah, 2026; Jorgensen, 2025). I study lender choice by the intermediary hired to find the borrower a mortgage. Within conventional

²Related research examines lender screening and markups, specialization, and the role of local loan officers in credit provision (Bosshardt, Kakhbod, and Kermani, 2025; Paravisini, Rappoport, and Schnabl, 2023; Huang, Mayer, Titman, and Xu, 2026).

conforming loans, I show that some broker firms persistently use more expensive lenders and that these choices are associated with both higher borrower rates and higher broker compensation. The contribution is to connect the intermediary’s choice of lender to the prices paid by borrowers and the compensation received by the intermediary, while allowing lenders to price the same borrower differently.

Finally, the paper contributes to research on price dispersion and search in financial markets. Search costs can sustain price differences across sellers (Stigler, 1961; Burdett and Judd, 1983). In mortgage markets, similar borrowers pay different rates, and less shopping and greater lender advertising are associated with higher borrowing costs (Allen, Clark, and Houde, 2014; Alexandrov and Koulayev, 2017; Bhutta, Fuster, and Hizmo, 2026; Woodward and Hall, 2010, 2012; Gurun, Matvos, and Seru, 2016). Related evidence links search frictions to markups in auto lending and documents dispersion in commercial property capitalization rates (Argyle, Nadauld, and Palmer, 2023; Priest, 2023). Substantial price differences also arise for identical or nearly identical financial products, including municipal and corporate bonds, convertible bonds, and index funds (Green, Hollifield, and Schürhoff, 2007; Griffin, Hirschey, and Kruger, 2023, 2026; Egan, 2019; Hortaçsu and Syverson, 2004). I examine dispersion when consumers delegate search to a professional intermediary. Mortgage rates remain widely dispersed in this setting, and persistent differences in brokers’ lender choices help explain why. The results show that an intermediary’s access to multiple sellers is not sufficient to ensure that consumers receive similar prices; the intermediary’s choices and incentives also matter.

1 Data, Background, and Hypothesis

I first introduce the data, then describe the institutional setting and develop the hypotheses that guide the analysis.

1.1 Data

Studying broker pricing requires linking the borrower, broker, and lender on each loan while observing the borrower’s risk characteristics and mortgage price. No single source provides all these elements. Mortgage records identify the participants but lack detailed borrower risk measures; rate locks contain risk and pricing information but do not identify the borrower; and HMDA reports prices and

demographics but identifies the lender rather than the broker. I combine these sources at the loan level.

1.1.1 Data Sources

CoreLogic mortgage records identify borrowers and properties and report the mortgage amount, dates, and company and loan-officer identifiers. These records connect the borrower to the broker and the lender on a completed mortgage. Optimal Blue rate-lock data provide credit scores, loan-to-value and debt-to-income ratios, the date and rate of the lock, the requested lock period, and broker compensation. They measure the terms agreed upon before closing. HMDA data provide rate spreads, origination charges, income, and demographics. Linking the three sources allows me to follow the same mortgage from its rate lock to its recorded origination and compare pricing across the firms involved.

I supplement these data with NMLS licensing and branch records, Verisk and DataAxle consumer address histories, and American Community Survey neighborhood characteristics. The address histories locate purchase borrowers before they buy the mortgaged property, allowing me to measure their proximity to broker branches when searching. I also collect broker website claims to study how firms describe their services.³

1.1.2 Sample Construction

The three sources share no universal identifier for a loan or a borrower. I therefore match them on loan characteristics that they record in common, such as the property’s location, the loan amount, and the loan’s terms, and I do so in two steps. First, I link Optimal Blue locks to CoreLogic mortgages using property location, loan amount, dates, and loan characteristics. Second, I link these pairs to HMDA using geography, lender identity, loan amount, occupancy, and note rate. The match produces 15.8 million loans that closed from 2018 to 2025. Internet Appendix Section [A](#) describes the matching procedure in greater detail, and Figure [IA.2](#) visualizes the matching procedure.

As a diagnostic for accidental matches, matching the 2022 locks against the 2021 HMDA file produces matches for about 1% of locks.

³Websites come from NMLS records and supplementary searches by company name. They were collected in June and September 2026 and classified by claims about rates, lender access, approval, and service.

I classify a loan as brokered when CoreLogic identifies a non-funding broker firm and either Optimal Blue or HMDA identifies an intermediated application; the remaining matched loans are classified as retail. Distinguishing the funder from the broker is important because CoreLogic's company fields do not have a consistent lender-broker ordering. I resolve the roles using the lender of record and NMLS company identities. Each brokered loan is then assigned to the branch where its officer worked in the lock month using NMLS employment and branch records.

The main analysis focuses on conventional conforming fixed-rate mortgages, which makes the comparison less dependent on differences across mortgage products. It includes 8,150,052 loans, of which 335,997 are brokered, spanning 15,409 broker firms and 31,377 loan officers. The lender-choice analysis uses the 308,810 brokered loans for which lender prices can be predicted. The distance-based IV analysis additionally requires a located search origin and a previous-year price classification for the broker. I explain these requirements with the corresponding tests; minor filters are reported in the exhibit notes.

1.1.3 Main Variables

Mortgage prices. The main outcome is the HMDA rate spread: the loan's annual percentage rate (APR) minus the average prime offer rate (APOR) for a comparable loan when the rate is set, measured in basis points. Because APR includes points and other finance charges, the spread captures more of the borrower's cost than the note rate alone. Optimal Blue lock dates also allow me to compare borrowers who face the same market rates. Borrowers closing on the same day may have locked weeks apart; comparisons by lock date avoid this timing difference.

Broker compensation. Broker compensation is measured in basis points of the loan amount. This unit differs from the rate spread: compensation records a payment on the transaction, whereas the rate spread is an annualized measure of borrowing cost. Keeping the two units distinct is important when interpreting whether a more expensive lender choice pays the broker more.

Lender networks. A broker's lender network consists of lenders that funded its loans during the previous 12 months. I construct these networks from all brokered CoreLogic mortgages, rather than only matched loans, so a lender relationship is not lost merely because a mortgage cannot be linked to

the other sources. I measure lender concentration by the sum of squared lender shares (HHI). A high HHI means that the firm places a large share of its loans with a small number of lenders; it does not by itself indicate whether those lenders are cheap or expensive. The baseline pricing controls include income, credit score, loan amount, property value, combined loan-to-value ratio, debt-to-income ratio, origination charges, race, and sex.

1.1.4 Summary Statistics and Representativeness

Table 1 reports brokered-loan prices, borrower characteristics, loan characteristics, and lender networks. Rate spreads vary substantially, with a mean of 31.3 basis points and a standard deviation of 48.2 basis points. The 10th and 90th percentiles are -17.8 and 92.0 basis points. The average borrower has a credit score of 752, an income of \$130,000, and a loan of \$353,000. Average debt-to-income and combined loan-to-value ratios are 37.4% and 73.0%, respectively. These borrowers have substantial observable credit quality, yet face considerable variation in mortgage prices.

Retail borrowers have similar credit scores and debt-to-income ratios, although brokered loans are larger and more often cash-out refinances. These composition differences motivate the borrower and loan controls used in comparisons across channels. The matched brokered sample covers 4.8% of comparable HMDA brokered loans. It has a similar average rate spread under common sample definitions, but overrepresents 2020–2021 originations and principal residences. The data therefore provide detailed links for a subset of the broker market.⁴

1.2 Institutional Background

The wholesale mortgage market separates the intermediary that advises the borrower from the lender that funds the loan.

1.2.1 Brokers and Wholesale Lenders

A borrower can obtain a mortgage directly from a lender or through a mortgage broker. In the wholesale channel, the broker takes the application, collects documents, and submits the loan to a lender, which underwrites and funds it. Brokers work with lenders that have approved them as wholesale

⁴Table IA.1 compares brokered and retail loans. Table IA.2 compares the matched brokered sample with HMDA using common definitions.

partners. Each broker-lender agreement sets the terms of the relationship, including compensation. Brokers therefore draw on a limited lender network: for the median brokered loan in the sample, the firm used five lenders in the previous year and the loan officer used three (Table 1, Panel D).

Brokers obtain pricing through lenders' rate sheets or online systems and lock the borrower's rate with the selected lender. A lender can offer different pricing to different broker accounts, and brokers may negotiate pricing on individual loans. The price available from a lender can therefore depend on which broker submits the loan. The firm and its loan officers enter the transaction at different levels. The firm maintains the wholesale relationships, while an individual loan officer works with the borrower and submits the application. A firm may therefore have access to more lenders than an individual officer regularly uses.

Figure 1, Panel A, presents the phrases used on mortgage broker websites, with larger phrases appearing on more firms' websites. Expertise and guidance are prominent, as are access to multiple lenders and finding a lower or better rate. Brokers thus advertise both help with the mortgage process and shopping on the borrower's behalf.⁵ Panel B compares the number of lenders that broker firms advertise with the number they use. Among the 637 broker firms whose websites state a number of lenders, the median advertised number is 25 lenders, while the median number of lenders that funded a firm's brokered loans in 2025 is five. Of these firms, 94.7% advertise at least 20 lenders, 94.5% used fewer than 20 in 2025, and 97% used fewer lenders than they advertise. Most broker firms therefore place their loans with a small subset of the lenders they advertise.

Licensing and disclosure. Mortgage originators are licensed or registered through the Nationwide Multistate Licensing System (NMLS). Broker firms also obtain state licenses, and loan documents report originator and company identifiers.

1.2.2 Broker Compensation and Its Regulation

The broker is paid either by the borrower through fees at closing or by the lender, typically as a percentage of the loan amount; it cannot receive compensation from both on the same transaction. Among brokered loans whose Optimal Blue record states who pays the broker, the lender pays on

⁵Figure 1A.1 provides an example of a broker website.

63.3% and the borrower on 36.7% (Table 1, Panel A). Recorded broker compensation averages 207.6 basis points of the loan amount. Lender-paid compensation is a cost of originating the mortgage and can be reflected in the borrower's price.

Federal rules restrict how brokers can be paid. Before 2011, lenders could pay brokers more for higher-rate loans through yield spread premiums (Woodward and Hall, 2012). Regulation Z now prohibits compensation based on the interest rate or other transaction terms, or proxies for those terms, while permitting compensation based on the amount of credit extended. Compensation can nevertheless differ across agreements with unaffiliated lenders. A broker can therefore earn more by placing a loan with one lender than with another, even though a given lender cannot pay more because the borrower receives a higher-rate loan.⁶

Anti-steering rule. The anti-steering provision seeks to prevent brokers from directing borrowers to loans to increase their own compensation. Under 12 CFR 1026.36(e)(1), a broker may not direct a borrower to a particular loan because it provides greater compensation unless the loan is in the borrower's interest. The rule does not require choosing the lender that pays the broker the least. A loan from a higher-paying lender may satisfy the rule when its terms serve the borrower's interest, such as when a higher rate comes with lower upfront charges that the borrower prefers.⁷

Safe harbor. The regulation also provides a safe harbor based on a limited comparison of loan options. A broker must obtain and review options from at least three lenders with which it regularly does business, or all such lenders if it works with fewer than three, and present options meeting specified rate, fee, and loan-feature criteria for which the borrower likely qualifies.⁸ It need not review every

⁶See 12 CFR 1026.36(d) and its official commentary. For example, a lender cannot pay 1% on one loan and 2% on another because their interest rates or other transaction terms differ. Separate agreements under which unaffiliated Lender A pays 1% and Lender B pays 2% are permissible, subject to the compensation and anti-steering rules. Affiliated lenders are treated as one person under paragraph (d)(3). The CFPB also discusses brokers' ability to choose among lenders paying different compensation in its June 2013 final rule, 78 FR 35430.

⁷See official comments 36(e)(1)-2 and -3. Outside the safe harbor described below, the borrower-interest comparison considers other possible offers available through lenders with which the broker regularly does business and for which the borrower likely qualifies. Merely describing a loan as being in the borrower's interest does not establish compliance.

⁸Under 12 CFR 1026.36(e)(2)-(4), for each loan type in which the borrower expresses interest, the required options include the lowest interest rate, the lowest total dollar amount of discount points and origination points or fees, and the lowest interest rate without specified features such as negative amortization, a prepayment penalty, or a balloon payment in the first seven years. The broker must have a good-faith belief that the borrower likely qualifies. One loan can satisfy multiple criteria; the options presented need not come from three different lenders. If more than three options are

lender in its network or search the entire market. For example, a broker working with ten lenders can satisfy the comparison requirement by reviewing three, provided it meets the other safe-harbor conditions. Those three need not be the least expensive or lowest-paying lenders in its network.

Fiduciary duties. These federal compensation and anti-steering restrictions do not impose a general fiduciary duty on mortgage brokers. State laws can impose additional obligations. California, Nevada, and Washington expressly impose fiduciary duties on covered mortgage brokers or licensees, and New Mexico imposes such duties on mortgage loan originators.⁹ Compliance with the federal safe harbor does not remove applicable state duties.

These rules seek to prevent compensation-driven steering, but they leave differences in compensation across lenders and discretion over which lenders a broker compares. If the comparison is limited to relatively expensive or high-paying lenders, its lowest-rate option can still be more expensive than alternatives elsewhere in the broker’s network or in the market. The lender that pays the broker more need not offer the borrower a lower price. These features leave room for compensation to influence lender choice, even with federal anti-steering restrictions and applicable state duties.

1.3 Research Question and Hypotheses

If brokers shop across lenders for borrowers, they may reduce the search frictions associated with mortgage rate dispersion. I first compare the dispersion of brokered and retail rates after accounting for borrower characteristics and local market conditions:

$$RateSpread_i = \beta \mathbb{1}(Brokered)_i + X_i' \gamma + \delta_{c(i),d(i)} + \varepsilon_i, \quad (1)$$

where $RateSpread_i$ denotes the rate spread of loan i , in basis points, and $\mathbb{1}(Brokered)_i$ equals one for a brokered loan and zero for a retail loan. X_i denotes the baseline borrower and loan controls. The functions $c(i)$ and $d(i)$ identify the property’s county and the loan’s lock date, respectively. The fixed

presented for a loan type, the qualifying options must be highlighted.

⁹See California Civil Code [Section 2923.1](#); Nevada Revised Statutes [Section 645B.0147](#); Washington [RCW 19.146.095](#); and New Mexico Statutes Section 58-21B-20. The provisions differ in scope. For example, California requires covered brokers to place the borrower’s economic interest ahead of their own, while New Mexico specifies duties for individual mortgage loan originators. These are examples, not an exhaustive list of state-law obligations.

effect $\delta_{c,d}$ absorbs market conditions shared by borrowers in county c locking on date d , and ε_i denotes the remaining loan-level pricing variation. I estimate the model jointly on retail and brokered loans.

The coefficient β captures the average conditional difference between channels. The residuals measure how much more or less each borrower pays than the model predicts. Estimating a common pricing function and including the brokered indicator separates the comparison of dispersion from differences in average rates. In particular, a narrower brokered residual distribution would indicate that the rates of similar brokered borrowers vary less around their channel's conditional mean.

Figure 2 presents the residual distributions, with gray bars for retail loans and a red line for brokered loans. The distributions overlap closely. The residual standard deviation is 38.2 basis points for retail loans and 39.8 basis points for brokered loans. Both distributions have long right tails: some borrowers pay at least 50 basis points more than similar borrowers in the same market. Average rates are also similar across the two channels.¹⁰

The comparison motivates the paper's question: why do similar borrowers pay different rates even when brokers shop on their behalf? I examine whether systematic differences across broker firms help explain this dispersion.

Similar dispersion in the two channels does not establish the role of brokers. Even if all brokers delivered the same expected price to a given borrower, unobserved borrower characteristics and loan-specific pricing could generate dispersion. Alternatively, some brokers might deliver lower prices and others higher prices, with their differences obscured by the channel average. Distinguishing these explanations requires following broker identities across loans, rather than comparing channels alone.

Hypothesis 0. *Observably similar borrowers in the same market pay the same expected rate regardless of which broker firm they use.*

Under this null, broker firms should not have persistent differences in conditional price levels. Their estimated effects may differ in a finite sample, but a firm's effect in one group of loans should not systematically predict its effect in another. The null does not require the brokered and retail

¹⁰The estimated average difference is 0.4–1.2 basis points and statistically insignificant across the specifications in Table IA.3. Figures IA.4 and IA.3 provide comparisons by credit score and with finer geographic controls.

channels to have the same average price. The alternative concerns differences among broker firms themselves.

Hypothesis 1. *Observably similar borrowers in the same market pay different expected rates depending on which broker firm they use.*

If some brokers deliver lower rates and others higher rates, the brokered average can resemble the retail average while substantial dispersion remains. One potential explanation is lender choice. Brokers may select expensive lenders when those lenders offer higher compensation.

Hypothesis 1A. *Some broker firms systematically choose lenders that are expensive for their borrowers. These choices are associated with higher borrower rates and higher lender-paid compensation.*

I first test for persistent differences in broker pricing and estimate the effect of using a high-cost firm. I then examine lender choice and compensation as a potential explanation for those differences.

2 Are All Broker Firms the Same?

Brokered rates remain widely dispersed after accounting for borrower characteristics and local market conditions. I now ask whether this dispersion reflects systematic differences across broker firms. I estimate each firm’s price level and examine whether it persists across periods and types of loans.

2.1 Broker Firms Differ in Pricing

If broker firms deliver similar rates to similar borrowers, their estimated price levels should differ only because of sampling variation. I test this prediction by estimating a separate effect for each broker firm:

$$RateSpread_i = \sum_f \beta_f \mathbb{1}(BrokerFirm_i = f) + X_i' \gamma + \delta_{c(i),d(i)} + \varepsilon_i, \quad (2)$$

where $RateSpread_i$ denotes loan i ’s rate spread, and $\mathbb{1}(BrokerFirm_i = f)$ equals one if the loan was originated through broker firm f and zero otherwise. The sum runs over broker firms. X_i denotes borrower and loan controls, while $c(i)$ and $d(i)$ identify the county and lock date. The fixed effect $\delta_{c,d}$ absorbs common local pricing conditions, and ε_i denotes the residual. Retail loans are the omitted

group, so β_f measures how much more or less firm f 's borrowers pay than similar retail borrowers.

This specification assigns the broker a price level while allowing the underlying lender to differ across its loans. That is the relevant comparison for a borrower choosing an intermediary: the rate the firm delivers may reflect both which lenders it uses and the terms it obtains from those lenders. The estimate does not require the broker itself to fund the mortgage. At this stage, it is a conditional price comparison; the next section addresses whether borrower sorting explains the differences.

Figure 3 presents these estimates for the 170 largest broker firms. Each dot is a firm's estimated price level, with its confidence interval, and firms are ordered from lowest to highest. The zero line is the retail benchmark. Blue and red dots identify firms significantly below and above retail, respectively, after adjusting for multiple comparisons.

The estimates range from about -41 to $+32$ basis points. Fifty-nine firms deliver significantly lower rates than retail, while 30 deliver significantly higher rates. These differences are large despite comparing borrowers with similar observable characteristics in the same market. The pattern extends beyond the largest firms: among all 10,104 firms in the regression, 947 are significantly below retail and 1,190 above it after correcting for multiple comparisons. Thus, the near-zero average difference between brokered and retail loans does not describe the prices delivered by individual broker firms.

I also assess how much of the remaining rate variation is associated with firm identity. After borrower controls and county $\times$ lock-date effects, broker firm effects explain 12.1% of residual rate-spread variance among brokered loans. Lender effects explain 9.4% (Table 2). These are separate explanatory comparisons, not additive shares or causal decompositions. They show that knowing the intermediary is informative about the borrower's rate, even though a different institution ultimately funds the loan.

2.2 Broker Pricing Persists Across Periods and Loan Types

The estimated differences could reflect a temporary mix of borrowers or loans. If they instead reflect persistent differences across firms, a firm's price level in one group of loans should predict its price level in another. I split the sample by period and loan characteristics and estimate equation (2) separately within each group.

Each side of a split uses its own retail comparison group. For example, the purchase-loan effect

compares a firm’s purchase borrowers with retail purchase borrowers, and the refinance effect makes the corresponding comparison among refinances. A positive relationship between these two effects therefore cannot be attributed simply to a firm doing more of the more expensive loan type. The time split similarly asks whether relative pricing survives changes in the overall level of mortgage rates and in the composition of originations.

Figure 4 presents six binned scatter plots. In each panel, the horizontal axis measures a firm’s price level in one group and the vertical axis its price level in the other. Each dot averages firms in a bin. The panels compare periods before and from 2022 onward, purchases and refinances, conventional and government-backed loans, high and low credit scores, long and short loan terms, and high and low loan-to-value ratios.

The relationships are positive throughout. Firms that deliver higher rates in 2018–2021 also deliver higher rates in 2022–2025. The same pattern holds across loan products and borrower groups. Correlations adjusted for estimation noise range from 0.65 to 0.95 across the six splits, with a correlation of 0.65 across periods (Table IA.4). The adjustment uses effects estimated in different lock months to distinguish persistent covariance from estimation noise. Persistence is substantial, although firms’ relative prices are not perfectly fixed over time.

These results reject the null of no persistent pricing differences across broker firms. The differences extend beyond a particular period or borrower group and help explain why substantial dispersion remains within the brokered channel. They are also reflected in local market outcomes: counties with a greater share of loans from high-price firms have higher residual brokered-rate dispersion and higher brokered rates relative to retail.¹¹ Persistent sorting of borrowers could still contribute to these patterns, which motivates the identification strategy that follows.

3 Do High-Cost Broker Firms Raise Borrowers’ Rates?

A natural concern is that the persistent pricing differences reflect the borrowers rather than the brokers. High-cost firms may serve borrowers with risks that lenders observe but I do not, or borrowers

¹¹Figure IA.5 measures each firm’s price level using its loans in other counties, so the local loans do not directly determine that classification.

with different preferences for mortgage services. To address this sorting concern, I use borrowers' relative proximity to high-cost and low-cost branches as an instrument for broker choice.

3.1 Borrowers Use Nearby Broker Branches

The design relies on proximity influencing which broker a borrower uses. I first examine this relationship without distinguishing high-cost from low-cost firms. For each borrower, I form a choice set of branches active during the previous 12 months whose firms are licensed in the property's state. This includes alternatives the borrower did not choose, which are necessary for comparing the appeal of nearby and distant branches.

Distance is measured from the borrower's home when searching, which I call the search origin. For a purchase borrower, this is the previous address traced through consumer address histories; for a refinance borrower, it is the mortgaged property. Using the previous address for purchases avoids treating the newly purchased home as the place from which the borrower searched. I estimate

$$\mathbb{1}(Chosen)_{i,j} = \sum_k \beta_k \mathbb{1}(Distance_{i,j} \in Bin_k) + \alpha_i + \eta_{j,t(i)} + u_{i,j}, \quad (3)$$

where each observation pairs loan i with a potential branch j . $\mathbb{1}(Chosen)_{i,j}$ indicates whether the borrower chose that branch, expressed in percentage points, and $Distance_{i,j}$ denotes the distance from the search origin to the branch. The indicator for Bin_k identifies distance category k , with branches at least 20 miles away as the omitted category. The function $t(i)$ identifies the lock month. Loan fixed effects α_i hold the borrower and choice set fixed; branch $\times$ lock-month effects $\eta_{j,t}$ absorb a branch's overall popularity in month t ; and $u_{i,j}$ denotes the residual.

The coefficient β_k compares a nearby branch with a distant branch for the same borrower, after accounting for differences in branch appeal. A branch that is popular with borrowers throughout a market does not appear attractive merely because it is close to some of them. The specification instead asks whether its likelihood of being chosen is higher among borrowers who live nearer to it.

Figure 5, Panel A, presents the estimated choice probabilities relative to branches at least 20 miles away. Dots show the coefficients in percentage points, and bars show 95% confidence intervals. The

probability of choosing a branch falls steadily with distance. A branch within half a mile is 2.67 percentage points more likely to be used, compared with 1.30 percentage points for a branch one to two miles away and about 0.7 percentage points for a branch three to five miles away. Borrowers' tendency to use nearby branches provides the basis for the instrument.¹²

The relationship is not a claim that borrowers always choose the nearest branch. Many branches remain feasible at a distance, and most branches in a borrower's choice set are not selected. What matters for the IV is that relative proximity changes the probability of choosing one type of broker. The distance gradient provides evidence for that first step before introducing the high-cost classification.

3.2 Differential Distance as an Instrument

A borrower who lives relatively close to a high-cost branch should be more likely to use a high-cost firm. I classify firms using their estimated price levels in the previous calendar year, so the borrower's current loan does not enter the classification. I estimate firm-year price effects relative to retail and shrink them toward the monthly mean, with greater shrinkage for less precisely estimated effects. This reduces the influence of noisy price estimates from small firms. A firm is high-cost if its lagged price level is at or above the median across branches in the lock month; otherwise it is low-cost.

The instrument is the difference between the distances to the nearest low-cost and high-cost branches:

$$dd_i = \text{Distance}(\text{Home}_i, \text{nearest low-cost branch}) - \text{Distance}(\text{Home}_i, \text{nearest high-cost branch}),$$

where Home_i denotes the search origin of the borrower on loan i , and dd_i denotes differential distance. A positive value means that the high-cost branch is closer. For example, if the nearest low-cost branch is four miles away and the nearest high-cost branch is one mile away, differential distance is three miles. The measure compares the two types of branch without requiring the borrower to use either nearest branch.¹³

¹²The average choice set contains about 1,700 branches. A conditional logit specification also finds a strong distance relationship (Table [IA.5](#)).

¹³Figure [IA.7](#) illustrates the instrument for a borrower in Travis County, Texas.

I estimate the effect of using a high-cost broker by two-stage least squares:

$$\mathbb{1}(\text{HighCost})_i = \pi dd_i + X_i' \gamma_1 + W_i' \lambda_1 + \alpha_{o(i),t(i)} + \mu_{c(i),t(i)} + u_i, \quad (4)$$

$$\text{RateSpread}_i = \beta \widehat{\mathbb{1}(\text{HighCost})}_i + X_i' \gamma_2 + W_i' \lambda_2 + \alpha_{o(i),t(i)} + \mu_{c(i),t(i)} + \varepsilon_i, \quad (5)$$

where $\mathbb{1}(\text{HighCost})_i$ equals one if loan i uses a high-cost broker firm and zero otherwise, and the hat denotes the component of broker choice predicted by the first stage. RateSpread_i denotes the loan's rate spread in basis points. Borrower controls X_i include deciles of credit score, combined loan-to-value ratio, and debt-to-income ratio. Origin controls W_i include neighborhood characteristics and distance to the nearest branch of either type. The functions $o(i)$, $c(i)$, and $t(i)$ identify the origin county, mortgage county, and lock month. The fixed effects $\alpha_{o,t}$ and $\mu_{c,t}$ absorb conditions in the origin and mortgage markets, respectively; u_i and ε_i denote the first- and second-stage residuals.

Controlling for distance to the nearest branch separates relative proximity from general access to brokers. Two borrowers can have equally good access to branches but differ in whether their nearby branches belong to high-cost or low-cost firms. Neighborhood controls account for differences such as population density, income, and education. Origin and mortgage county effects allow the place where a purchase borrower searches to differ from the place where the borrower buys. The coefficient β then uses broker-choice variation associated with relative proximity after accounting for these market conditions.

3.3 Relevance and the Exclusion Restriction

The instrument must predict broker choice and affect mortgage rates through that choice. I first test relevance. Figure 5, Panel B, is a binned scatter plot of high-cost broker use against differential distance, after accounting for the controls and fixed effects. Each dot averages loans in a distance bin. High-cost broker use rises as the nearest high-cost branch becomes relatively closer. A 10-mile increase in differential distance is associated with a 5.2–5.6 percentage point increase in high-cost broker use, against a mean of 42.7%. First-stage F -statistics range from 84.8 to 101.1 (Table 3).

The exclusion restriction requires that relative proximity affect the borrower's rate only through

the type of broker used. This would fail, for example, if high-cost branches locate near borrowers who would pay higher rates at any firm. Although unobserved differences cannot be tested directly, I examine whether the instrument predicts observable borrower and neighborhood characteristics.

The timing of the classification addresses a different concern. Because firms are classified from earlier loans, an unusually expensive focal mortgage cannot make its own broker high-cost. This avoids a mechanical relationship between the treatment indicator and the outcome. It does not, however, make branch locations random. The location concern is why the design controls for both search-origin characteristics and local market conditions and examines balance on observables.

Figure 6 presents these comparisons. Each point is the coefficient from a separate regression of a standardized characteristic on the standardized instrument, with county $\times$ month fixed effects. The figure also reports associations with the chosen broker’s website claims. Those claims describe the selected firm rather than predetermined borrower characteristics, but help assess whether relative proximity leads borrowers to firms advertising different services.

The differences are economically small. A one-standard-deviation increase in differential distance changes every borrower and loan characteristic by at most 0.02 standard deviations. The largest neighborhood difference is 0.08 standard deviations for the college-educated share, and 38 of the 40 coefficients are below 0.05 standard deviations in absolute value. Fourteen are statistically significant at the 5% level, so the evidence does not imply exact balance. The IV regressions also control for borrower risk and neighborhood characteristics. These comparisons support the design, while the causal interpretation continues to require the exclusion restriction.¹⁴

3.4 High-Cost Brokers Raise Borrowers’ Rates

I next use the variation in broker choice induced by relative proximity to estimate its effect on pricing. Table 3 presents OLS estimates in Panel A and 2SLS estimates in Panel B. Across columns, I add controls for conditions in the mortgage county, the lock date, and the lender in the lock month. The lender comparisons ask whether high-cost brokers deliver higher rates even among loans funded by the same lender.

¹⁴Tables IA.8 and IA.9 and Figure IA.6 provide complementary comparisons by the sign of the instrument.

The OLS estimates range from 8.8 to 10.9 basis points. The IV estimates are similar, ranging from 10.4 to 12.2 basis points and significant at the 5% level throughout. With both origin and mortgage county $\times$ month fixed effects, the estimate is 10.9 basis points. Under the identification assumptions, these estimates indicate that using a high-cost firm raises the rate spread for borrowers whose broker choice responds to relative proximity.

The similarity of OLS and IV estimates is informative because the two comparisons use different sources of variation. OLS compares borrowers who chose different types of firms, whereas IV uses the component of that choice predicted by relative proximity. Their similarity does not establish that all sorting is absent. Under the exclusion and monotonicity assumptions, the IV estimate describes the effect for borrowers induced to use a high-cost firm by proximity, rather than an average effect for every borrower or every broker pair.

The specifications with lender $\times$ lock-month effects further compare loans funded by the same lender in the same month. The higher-rate relationship remains in those specifications. Lender identity therefore does not absorb all of the pricing difference across broker types. Because lender choice can itself respond to broker choice, these specifications are useful conditional comparisons rather than a decomposition of the total IV effect into separate causal channels.

The increase is economically meaningful: it is about half of the mean rate spread in the IV analysis. An illustrative 11 basis point increase in the note rate on a \$300,000, 30-year mortgage at 7% would increase payments by about \$267 per year.¹⁵

Higher rates might reflect better service. Purchase borrowers facing closing deadlines may be willing to pay more for a broker that can complete the loan quickly. The linked rate-lock and mortgage records provide lock dates, closing dates, and requested lock periods, allowing me to test this explanation directly. Using the same instrument, I find no statistically significant differences between borrowers at high-cost and low-cost firms in either the time from lock to closing or the requested lock period (Table [IA.7](#)).

¹⁵Discounted at 7%, the additional payments have a present value of about \$3,339 if the mortgage is held for the full 30 years. This illustration treats the estimated APR-spread difference as a note-rate change; realized costs depend on fees, prepayment, and the holding period.

The absence of a statistically significant difference in planned or realized timing weakens the explanation that high-cost firms charge more because they close loans faster. The data do not capture every aspect of service, such as assistance with documentation. I next examine lender choice as an explanation for the pricing differences, allowing for borrower information missing from my data when comparing lenders' prices.

4 Why Are Some Broker Firms High-Cost?

The previous section shows that using a high-cost broker raises borrowers' rates under the IV assumptions. Why do these firms deliver higher rates? One reason could be their choice of lenders. Brokers advertise that they shop for a lower rate, so I examine whether they select lenders that are inexpensive for their borrowers. I then ask whether costly lender choices differ persistently across firms and help explain what subsequent borrowers pay.

4.1 Measuring Lender Choice

Assessing a broker's lender choice requires comparing the selected lender with alternatives for the same borrower. The challenge is that I observe the rate only at the lender that funds the loan. Comparing lenders' average rates would not solve the problem: a lender may have high average rates because it serves riskier borrowers, yet be inexpensive for a particular borrower. I instead estimate lender-specific pricing functions and evaluate each at the same borrower's characteristics. This holds the borrower fixed while changing the lender.

The comparison set S_i consists of lenders active in borrower i 's state, together with the lenders used by the broker there in the previous year. It provides a common market benchmark while retaining the broker's existing relationships. Using a state set also allows the analysis to capture costly restrictions in the broker's network. If the comparison included only the lenders a broker already uses, a firm with an expensive network could appear to shop well simply by selecting the cheapest lender within that network.¹⁶

¹⁶The state set includes lenders that funded brokered loans for at least 10 broker firms in the state over the previous 12 months, together with the broker's own lenders in that state.

The lender pricing model is

$$RateSpread_i = X_i' \gamma_\ell + \phi_{\ell, f(i)} + \psi_{\ell, s(i), y(i)} + \tau_{\ell, t(i)} + \varepsilon_i, \quad (6)$$

where $RateSpread_i$ denotes loan i 's rate spread and X_i denotes its borrower and loan characteristics. The subscript ℓ identifies the lender, and the coefficients γ_ℓ allow each lender to price those characteristics differently. The functions $f(i)$, $s(i)$, $y(i)$, and $t(i)$ identify the broker firm, state, year, and lock month. Lender $\times$ broker effects $\phi_{\ell, f}$ capture relationship-specific pricing; lender $\times$ state $\times$ year effects $\psi_{\ell, s, y}$ capture local pricing differences; lender $\times$ lock-month effects $\tau_{\ell, t}$ capture changes in the lender's prices over time; and ε_i denotes the residual.

The 100 lenders with the most brokered loans have separate pricing models, estimated on their retail and brokered loans; the remaining lenders share a pooled model. Evaluating each model at the focal borrower's characteristics gives a predicted rate spread $\hat{r}_{i\ell}$. The chosen lender's price is also predicted, leaving out the focal loan. All lender comparisons therefore use predicted prices, and the focal loan's realized rate does not enter the prediction for its chosen lender.

Predicting a price outside the broker's existing lender network requires accounting for relationships that have not been observed. I decompose the lender $\times$ broker effect into a broker component, a lender component, and a relationship-specific component. Every lender prediction includes the broker component, while a relationship-specific component is included only for lenders on the broker's observed menu. This retains the firm's general price level when comparing lenders and allows existing relationships to offer different prices. It does not assume that an unobserved relationship has the same negotiated terms as an existing one.

The models imply substantial price differences for the same borrower. The mean predicted spread across lenders in the state set is 41.0 basis points, compared with 5.7 basis points at the lender with the lowest predicted spread. Lenders in the broker's own network are cheaper than the state average, with a mean predicted spread of 33.9 basis points. The median lender-model R^2 is 0.49, and the out-of-sample root mean squared prediction error on brokered loans is 35.1 basis points.¹⁷

¹⁷Figure [IA.10](#) reports model fit, and Figure [IA.11](#) compares lender-specific prices of borrower characteristics. Internet

I first calculate the *observed gap*, the chosen lender's predicted price relative to a benchmark price in the state set:

$$\text{ObservedGap}_i^b = \hat{r}_{i\ell^*(i)} - B_i^b, \quad (7)$$

where

$$B_i^b \in \left\{ \min_{\ell \in S_i} \hat{r}_{i\ell}, \frac{1}{|S_i|} \sum_{\ell \in S_i} \hat{r}_{i\ell} \right\}$$

is the benchmark, with $b \in \{\text{min}, \text{mean}\}$ selecting the lowest or the mean predicted price in the state lender set S_i . Here $\ell^*(i)$ denotes the lender that funded loan i , and $\hat{r}_{i\ell}$ denotes lender ℓ 's predicted spread for that loan. The gap compares the chosen lender with a common reference for the borrower, in basis points of rate spread. I report observed gaps against the minimum, for which the gap is zero if the broker chooses the lender with the lowest predicted price and positive otherwise.

A concern is that lender quotes observed by a broker may reflect pricing information not captured by the estimated model. If this information changes the ranking of lenders, an honest broker selecting the lowest quote can have a positive observed gap.

To address this concern, I simulate the lender choices of an *honest broker* that obtains a quote from every lender in S_i and selects the lowest. Each quote equals the predicted price plus an independent, mean-zero normal pricing error. I set the common standard deviation σ to 35 basis points, calibrated to the 35.1 basis point out-of-sample prediction error on brokered loans. The prediction error may include borrower risk that raises or lowers all lenders' quotes by the same amount, leaving their ranking unchanged. The variation relevant to lender choice could therefore be smaller. The simulation assigns this 35-basis-point standard deviation to lender-specific pricing errors.

The *honest-broker gap* is the observed gap an honest broker would generate on average across the possible quotes:

$$\text{HonestGap}_i^b = \sum_{\ell \in S_i} p_{i\ell} \hat{r}_{i\ell} - B_i^b, \quad (8)$$

where $p_{i\ell}$ denotes the probability that an honest broker selects lender ℓ for loan i , and B_i^b is the same

Appendix Section B records additional implementation details.

benchmark as in equation (7). The simulated broker selects a lender using its quote, but I evaluate that choice at the lender’s predicted price, just as I do for the actual broker. Averaging these predicted-price gaps gives the gap that honest shopping can generate under the assumed quote distribution.

I subtract the honest-broker gap from the observed gap to obtain the *steering gap*, a proxy for steering after allowing for these pricing errors:

$$\begin{aligned}
SteeringGap_i &= ObservedGap_i^b - HonestGap_i^b \\
&= \left(\hat{r}_{i\ell^*(i)} - B_i^b \right) - \left(\sum_{\ell \in S_i} p_{i\ell} \hat{r}_{i\ell} - B_i^b \right) \\
&= \hat{r}_{i\ell^*(i)} - \sum_{\ell \in S_i} p_{i\ell} \hat{r}_{i\ell},
\end{aligned} \tag{9}$$

where $SteeringGap_i$ denotes the part of the observed gap above the gap expected from honest shopping under the simulation. A positive value means that the chosen lender’s predicted price exceeds the expected predicted price selected by the honest broker; a negative value means it falls below that expectation. Individual honest choices can produce positive or negative values, but their expected steering gap under the model is zero. The proxy is intended to capture systematic lender-choice practices, rather than classify every individual loan as steered or not steered. I use this adjusted gap, rather than the unadjusted observed gap, in the subsequent analysis.

The benchmark B_i^b cancels from the difference, so the steering gap is the same whether the observed and honest-broker gaps are measured against the minimum or the mean. The measure compares the chosen lender directly with a probability-weighted alternative, rather than measuring how far the borrower is from a single unusually cheap prediction. Figure IA.8 illustrates the observed gap, the honest-broker gap, and their difference for one loan.

A positive average steering gap indicates lender choices that are more expensive than the honest-shopping benchmark under its assumptions. Because I do not observe competing offers or the advice given to borrowers, I interpret this gap as a proxy for steering practices. The next tests examine whether this proxy varies systematically across firms, persists over time, and predicts what borrowers pay; the compensation analysis then tests whether these costly choices also benefit the broker.

I also examine whether the results depend on the scope of the comparison set. A lender active in a state need not be accessible through every broker or willing to make every loan. The state benchmark includes the cost of differences in lender networks. Repeating the analysis using only the broker’s own lenders tests whether the findings depend entirely on comparisons with lenders outside its observed relationships.

I assess sensitivity to the pricing-error assumptions as well. Larger errors give lenders with higher predicted prices more opportunity to offer the lowest quote, increasing the honest-broker gap and reducing the steering gap. I examine lender-specific prediction errors and a common standard deviation of 111 basis points, as well as the own-network benchmark. The main relationships remain positive and significant under these alternatives (Table IA.10). Under the own-network benchmark, the association between past firm gaps and rates is about 40% smaller but remains positive. The relationship with borrowing costs therefore holds under alternative pricing-error assumptions and when the comparison is confined to lenders the broker has used.

4.2 The Steering Gap Predicts Realized Rates

If the steering gap captures costly lender choices, borrowers with larger gaps should pay higher realized rates. I test this relationship using

$$RateSpread_i = \beta SteeringGap_i + X_i' \gamma + \delta_{c(i),t(i)} + \zeta_{f(i)} + \varepsilon_i, \quad (10)$$

where $RateSpread_i$ denotes the realized rate spread on loan i , and $SteeringGap_i$ denotes its steering gap, both in basis points. X_i denotes borrower and loan controls.¹⁸ The functions $c(i)$, $t(i)$, and $f(i)$ identify the county, lock month, and broker firm. County $\times$ lock-month effects $\delta_{c,t}$ absorb local pricing conditions, broker effects ζ_f absorb persistent firm characteristics, and ε_i denotes the residual. With firm effects, β compares the rates of a firm’s borrowers whose loans have different steering gaps.

Figure 7 introduces the gaps and their relationship with realized rates. Panel A compares the observed-gap distribution with the distribution implied by honest-broker choices. Panel B shows the

¹⁸The controls are income, credit score, loan amount, property value, combined loan-to-value ratio, debt-to-income ratio, race, ethnicity, sex, occupancy, property type, number of units, loan purpose, and loan term.

steering-gap distribution in gray bars on the left axis and a binned scatter plot of realized spreads against the gap on the right axis. Each dot averages loans after accounting for borrower characteristics and local market conditions.

The observed gap exceeds the gap generated by honest shopping after accounting for the assumed pricing errors. Their means are 27.7 and 15.8 basis points, leaving an average steering gap of 11.9 basis points. Thus, honest selection of the lowest quote does not explain the observed lender choices under the calibrated benchmark. This excess supports a steering interpretation. The loan-level steering gap is positive for 71.9% of loans. Panel B also shows that realized spreads increase with the gap throughout its distribution.

Table 4 presents the regression estimates.¹⁹ The first column compares borrowers within local markets; later columns add broker firm effects, lender concentration, and lock-date effects. A one basis point higher gap is associated with a 0.79 basis point higher realized spread. Within the same broker firm, the coefficient is about 0.73 and changes little with the additional controls. A 10 basis point larger gap therefore corresponds to a roughly 7.3 basis point higher realized spread. The predicted lender-price differences are reflected in what borrowers pay.

4.3 Lender Choices Differ Across Broker Firms

The loan-level proxy can still reflect an individual pricing error or an unusual lender choice. If it captures broker practices, however, those differences should also appear when loans are aggregated within firms. I therefore calculate each broker firm's mean steering gap. A firm following the simulated honest-shopping rule has an expected mean gap of zero; persistent positive averages indicate choices that are expensive relative to that rule.

Figure 8 compares the 25 largest broker firms. It reports each firm's mean steering gap and confidence interval, ordered from lowest to highest.

Mean steering gaps range from about zero to 20 basis points. Some firms choose lenders close to the honest-shopping benchmark on average, while others have gaps of 15–20 basis points even after the adjustment for pricing errors. The honest-broker gaps vary less across these firms (Figure IA.12), so the

¹⁹Internet Appendix Tables IA.11 to IA.14 repeat Tables 4 to 7 with the observed gap, before the honest-broker adjustment, in place of the steering gap.

differences do not simply reflect different simulated allowances. These practices differ markedly across broker firms: the variation remains when lender choices are averaged across each firm’s borrowers.

4.4 Lender Choices Persist Within Firms

A firm’s average gap could reflect a temporary group of borrowers or lending conditions. If it reflects persistent lender-choice practices, the firm’s past gap should predict the gap on subsequent loans. I estimate

$$SteeringGap_i = \beta PastSteeringGap_{f(i),t(i)} + X_i' \gamma + \kappa HHI_{f(i),t(i)} + \delta_{c(i),t(i)} + \zeta_{f(i)} + \varepsilon_i, \quad (11)$$

where $SteeringGap_i$ denotes the current loan’s gap, and $PastSteeringGap_{f,t}$ denotes the mean steering gap across firm f ’s loans in months $t - 12$ through $t - 1$, measured in basis points and updated each month. The focal loan and other loans in lock month t do not enter this average. $HHI_{f,t}$ denotes the firm’s lender concentration over the same prior period, standardized by its sample standard deviation. The functions $f(i)$, $t(i)$, and $c(i)$ identify the firm, lock month, and county of loan i . The controls X_i , local market effects $\delta_{c,t}$, firm effects ζ_f , and residual ε_i are defined as in equation (10).

Without firm effects, the coefficient compares firms with different recent lender-choice histories. With firm effects, it asks whether a firm’s current choices become more expensive when its own recent choices have been more expensive. The concentration control distinguishes the prices of the lenders used from the extent to which the firm relies on a small number of lenders. A concentrated network can be inexpensive, and a broad network can still produce expensive choices.

Figure 9 presents two binned scatter plots. Panel A compares the focal loan’s gap with its firm’s past gap. Panel B compares firm averages in two successive, non-overlapping 12-month windows. Both relationships are positive and approximately linear.

The second comparison uses a firm’s mean in months $t - 12$ to $t - 1$ and its mean in months $t - 24$ to $t - 13$. The two means contain different loans, so their relationship does not mechanically arise from a loan appearing on both sides of the regression. Together, the loan-level and annual-window tests distinguish a persistent firm pattern from an isolated expensive lender choice.

Table 5 reports the corresponding estimates. In Panel A, the coefficient is 0.71 across firms and 0.46 within firms. Thus, a 10 basis point increase in a firm’s past gap is associated with a 4.6 basis point higher gap on its current loans, holding persistent firm characteristics fixed. In Panel B, coefficients across non-overlapping annual windows range from 0.60 to 0.66. Firms that used expensive lenders in the past tend to continue doing so.

4.5 Past Lender Choices Predict Borrowers’ Rates

The persistence of the steering proxy makes it possible to characterize a firm’s lender-choice practices before a new borrower arrives. I next ask whether borrowers who use firms with higher past steering gaps pay higher rates themselves. I relate the current loan’s rate spread to the same measure of past firm steering practices:

$$RateSpread_i = \beta PastSteeringGap_{f(i),t(i)} + X_i'\gamma + \kappa HHI_{f(i),t(i)} + \delta_{c(i),t(i)} + \zeta_{f(i)} + \varepsilon_i, \quad (12)$$

where $RateSpread_i$ denotes loan i ’s realized spread and $PastSteeringGap_{f,t}$ denotes the mean steering gap across firm f ’s loans in months $t-12$ through $t-1$. In this regression, the measure is standardized by its standard deviation of 18.6 basis points. The functions $f(i)$, $t(i)$, and $c(i)$ identify the broker firm, lock month, and county. $HHI_{f,t}$ denotes standardized lagged lender concentration; X_i denotes borrower and loan controls; $\delta_{c,t}$ and ζ_f denote local market and firm effects; and ε_i denotes the residual. The coefficient β measures the rate difference associated with a one-standard-deviation higher past gap.

This specification uses the broker’s history rather than the lender choice on the current loan. The historical measure can be constructed before the focal loan is locked, without observing that loan’s price. Firm effects then compare borrowers at the same firm during periods with different recent lender-choice histories. The test asks whether the persistence documented above has a corresponding pattern in subsequent borrowing costs.

Figure 10 presents this within-firm relationship as a binned scatter plot. The horizontal axis is the firm’s past steering gap and the vertical axis the realized rate spread, adjusted for controls and fixed effects. Rates rise with the past gap. Table 6 reports the estimates with and without firm effects and

lender concentration controls.

A one-standard-deviation higher past steering gap is associated with a 10.0 basis point higher rate spread across firms. Within the same firm, the estimate is 7.1 basis points, or 6.8 basis points with lock-date effects. Borrowers therefore pay more when they use a firm with a recent history of larger steering gaps, even after accounting for borrower characteristics, local market conditions, and persistent firm differences. Because the proxy uses earlier loans, this relationship extends beyond the lender choice and price on the focal loan. The firm's past lender-choice practices predict the borrowing costs of its subsequent customers.

4.6 Expensive Lenders or Poor Borrower Fit?

A lender can be expensive for a borrower for two different reasons. It may be expensive in general, charging more than other lenders for most borrowers or for most of the broker's borrowers. Or it may be inexpensive in general but price this borrower's characteristics less favorably than other lenders do. A low credit score, for example, may raise the rate more at one lender than at another. The two explanations describe different broker behavior. In the first, the broker uses lenders that are expensive in general. In the second, the broker uses lenders that can be competitive but does not place each borrower with the lender that prices that borrower best.

The pricing model in equation (6) distinguishes the two. As noted above, the lender $\times$ broker effect is the sum of a broker component, a lender component, and a relationship-specific component, $\phi_{\ell,f} = \phi_f^B + \phi_\ell^L + \phi_{\ell,f}^R$. Let X^{ref} denote a reference borrower with the average characteristics of brokered loans. Each predicted price then splits into four components:

$$\hat{r}_{i\ell} = \underbrace{X^{\text{ref}'}\hat{\gamma}_\ell + \hat{\phi}_\ell^L + \hat{\psi}_{\ell,s(i),y(i)} + \hat{\tau}_{\ell,t(i)}}_{\text{lender level } \hat{r}_{i\ell}^{\text{Lev}}} + \underbrace{\hat{\phi}_{\ell,f(i)}^R}_{\text{relationship } \hat{r}_{i\ell}^{\text{Rel}}} + \underbrace{(X_i - X^{\text{ref}})'\hat{\gamma}_\ell}_{\text{borrower fit } \hat{r}_{i\ell}^{\text{Fit}}} + \underbrace{\hat{\phi}_{f(i)}^B}_{\text{broker}}. \quad (13)$$

The lender level $\hat{r}_{i\ell}^{\text{Lev}}$ is lender ℓ 's price for the reference borrower in the loan's state, year, and lock month. The relationship component $\hat{r}_{i\ell}^{\text{Rel}}$ is the part of the lender's price specific to this broker, and it is zero for lenders outside the broker's observed menu. The borrower-fit component $\hat{r}_{i\ell}^{\text{Fit}}$ is the price difference that the borrower's own characteristics make at lender ℓ , relative to the reference borrower.

The broker component is the same for every lender. The lender level and the relationship component therefore measure how expensive a lender is in general, for any borrower or for any borrower of this broker, while the borrower-fit component measures how the lender prices this borrower's characteristics.

Substituting equation (13) into the steering gap in equation (9) splits the gap into the same parts:

$$\begin{aligned}
 \text{SteeringGap}_i = & \underbrace{\left(\hat{r}_{i\ell^*(i)}^{\text{Lev}} - \sum_{\ell \in S_i} p_{i\ell} \hat{r}_{i\ell}^{\text{Lev}} \right)}_{\text{lender level}} + \underbrace{\left(\hat{r}_{i\ell^*(i)}^{\text{Rel}} - \sum_{\ell \in S_i} p_{i\ell} \hat{r}_{i\ell}^{\text{Rel}} \right)}_{\text{relationship}} \\
 & + \underbrace{\left(\hat{r}_{i\ell^*(i)}^{\text{Fit}} - \sum_{\ell \in S_i} p_{i\ell} \hat{r}_{i\ell}^{\text{Fit}} \right)}_{\text{borrower fit}}.
 \end{aligned} \tag{14}$$

Each part compares the chosen lender's component with the component that the honest broker is expected to select. Because the selection probabilities $p_{i\ell}$ sum to one, the broker component cancels. The probabilities depend on the full predicted prices, so the decomposition is an accounting split at the benchmark's probabilities rather than a set of separate counterfactuals.

On the average loan, most of the steering gap reflects lenders that are expensive in general. Lender levels account for 62.1% of the 11.9 basis point average gap and relationship components for 16.3%. Borrower fit accounts for the remaining 21.6%, so poor fit is part of the gap but the smaller part.

Which of these components helps explain why borrowers at firms with higher past gaps pay more? To examine this, I split the past gap in equation (12) into its three parts, each a trailing 12-month firm average divided by the standard deviation of the past gap. I regress the rate spread on the three parts together, with the controls and fixed effects of Table 6, Column (4). A part's contribution is its coefficient multiplied by its share of the past gap, the coefficient from regressing the part on the past gap with the same controls and fixed effects. The contributions add up to the coefficient on the past gap.

Figure 11 reports the contributions. Of the 7.1 basis point within-firm estimate, lender levels account for 58.5%, relationship components for 38.6%, and borrower fit for 2.9%. The observed gap

against the lowest predicted spread splits similarly, with borrower fit accounting for 5.3% of its estimate. The relationship between past steering gaps and subsequent rates is concentrated in firms' use of lenders that are expensive in general; differences in how well lenders price particular borrowers account for little of that relationship. Borrower fit accounts for about a fifth of the average gap but for almost none of its relationship with subsequent rates.

4.7 Lender Choice or Menu Choice?

A broker's menu is the set of lenders it worked with in the borrower's state during the previous year. An expensive choice can arise within this menu or from the menu itself. The broker may choose a lender with a higher predicted price than others on its menu. Or it may choose well among its own lenders while the menu as a whole is expensive relative to the other lenders active in the state. This distinction matters because access to more lenders would not by itself improve choices among existing partners.

To separate the two, I compare three predicted rates for the same loan. The first is the rate at the chosen lender, $\hat{r}_{i\ell^*(i)}$. The second is the expected predicted price of the lender selected by the honest broker in equation (8), which shops the full state set, $\bar{r}_i^S = \sum_{\ell \in S_i} p_{i\ell} \hat{r}_{i\ell}$. The third is the corresponding expected predicted price when the honest broker shops only the broker's menu M_i , $\bar{r}_i^M = \sum_{\ell \in M_i} p_{i\ell}^M \hat{r}_{i\ell}$, where $p_{i\ell}^M$ denotes the probability that lender ℓ quotes the lowest price among the menu lenders, with the same pricing-error distribution as in the state-set benchmark. The chosen lender's rate can then be written as

$$\hat{r}_{i\ell^*(i)} = \underbrace{\bar{r}_i^S}_{\text{state-set benchmark}} + \underbrace{(\bar{r}_i^M - \bar{r}_i^S)}_{\text{menu choice}} + \underbrace{(\hat{r}_{i\ell^*(i)} - \bar{r}_i^M)}_{\text{lender choice}}. \quad (15)$$

The first term is the expected predicted price of the lender selected by an honest broker shopping the state set. The second measures how much higher that expected predicted price is when shopping is restricted to the broker's menu. The third compares the chosen lender's predicted price with the honest-shopping benchmark for the same menu. All three terms use predicted prices; the pricing errors affect which lender the honest broker selects.

Subtracting the state-set benchmark from both sides splits the steering gap into the two parts:

$$\text{SteeringGap}_i = \hat{r}_{i\ell^*(i)} - \bar{r}_i^S = \underbrace{(\hat{r}_{i\ell^*(i)} - \bar{r}_i^M)}_{\text{lender choice}} + \underbrace{(\bar{r}_i^M - \bar{r}_i^S)}_{\text{menu choice}}. \quad (16)$$

The observed gap splits in the same way, with the lowest predicted spreads on the menu and in the state set in place of the two honest brokers' expected predicted prices: the chosen lender's predicted spread minus the lowest on the menu, plus the lowest on the menu minus the lowest in the state set. Figure IA.9 illustrates both splits for one loan.

On the average loan, lender choice accounts for 8.8 basis points of the 11.9 basis point steering gap, or 73.5%, and menu choice for 3.2 basis points, or 26.5%. Menu choice is a larger part of the observed gap, 12.3 of 27.7 basis points. The lowest predicted spread in the larger state set is low partly because it is the minimum over more noisy predictions; the honest-broker comparison accounts for this.

Using the method of Section 4.6, I split the within-firm relationship between the past gap and borrowers' rates (Table 6, Column (4)) into the two parts. Figure 11, Panel B, reports the contributions. Of the 7.1 basis point estimate, lender choice accounts for 78.9% and menu choice for 21.1%. The observed gap splits almost identically, 80.3% and 19.7%. Per basis point, the menu part predicts higher rates than the lender-choice part, but it varies less across firms and over time, accounting for 15% of the variation in the past gap. The higher rates paid by subsequent borrowers at firms with larger past gaps are therefore mainly associated with expensive choices from the firms' existing menus; expensive menus account for about a fifth of the relationship. Table IA.10, Column (5), reports the main estimates with the lender-choice part alone as the gap.

5 Why Do Some Broker Firms Use Expensive Lenders?

Why would a broker repeatedly choose lenders that are expensive for its borrowers? Compensation provides a possible explanation. A broker's compensation agreements can differ across lenders, so a lender that charges the borrower more may also pay the broker more. I test this explanation by examining whether the costly lender choices captured by the steering gap are associated with higher lender-paid compensation.

The relevant comparison is borrower-specific. A lender that pays high compensation need not be expensive for every borrower because lenders price credit and loan characteristics differently. Relating compensation to the steering gap accounts for this distinction: the gap assesses the selected lender against alternatives for the same loan. Firm effects further separate the relationship from a tendency for some broker firms to earn more on all their loans.

Lenders pay the broker on about 63% of brokered loans (Section 1.2). For lender-paid loans, I estimate:

$$Comp_i = \beta SteeringGap_i + X_i' \gamma + \delta_{c(i),t(i)} + \zeta_{f(i)} + \varepsilon_i, \quad (17)$$

where $Comp_i$ denotes the compensation paid by the lender to the broker on loan i , in basis points of the loan amount, and $SteeringGap_i$ denotes the loan's gap in basis points of rate spread. X_i denotes borrower and loan controls. The functions $c(i)$, $t(i)$, and $f(i)$ identify the county, lock month, and broker firm. The fixed effects $\delta_{c,t}$ and ζ_f absorb local market conditions and persistent firm compensation levels, respectively, and ε_i denotes the residual. The coefficient β measures whether a firm receives more compensation on loans placed with relatively expensive lenders.

Figure 12 presents a binned scatter plot of compensation against the steering gap. Each dot averages loans after accounting for borrower characteristics, local market conditions, and broker firm effects. Table 7 reports the corresponding regressions, first without fixed effects and then with local market, firm, and lock-date effects.

Compensation increases with the steering gap. Without fixed effects, a one basis point higher gap is associated with 0.28 basis points more compensation. Within firms, the coefficient is 0.11 and remains unchanged when lock-date effects are added. A 10 basis point larger steering gap is therefore associated with about 1.1 basis points more compensation as a share of the loan amount, relative to a mean of 205.4 basis points. Although modest in magnitude, this association provides evidence of compensation-related steering: the lender choices captured by the proxy are associated with both higher broker pay and higher borrower costs. The result holds within the same firm, beyond differences in average compensation across brokers.

The compensation difference is a payment at origination, whereas the rate difference affects the

borrower's financing cost over the life of the mortgage, so the two coefficients are not a dollar-for-dollar transfer. The compensation regression identifies an association rather than the causal response to a change in a lender's payment schedule. In conjunction with the earlier results, however, it strengthens the interpretation of the steering gap as capturing a conflict of interest. Lender choices remain expensive after the simulation allows for information missing from my data, vary persistently across firms, predict higher rates for subsequent borrowers, and pay the broker more.

6 Conclusion

Mortgage brokers promise to shop across lenders for borrowers, yet rates remain as dispersed among brokered loans as among retail loans. Using novel data linking borrowers, brokers, and lenders, I document wide and persistent differences in the rates broker firms deliver. An IV design based on borrowers' relative proximity to broker branches shows that choosing a high-cost firm raises rate spreads by 10–12 basis points.

Some broker firms systematically select lenders that are expensive for borrowers and pay brokers more, consistent with compensation-related steering. These practices vary widely across firms and persist within firms: borrowers who use firms with stronger past steering practices pay higher rates. The relationship with subsequent borrowing costs is concentrated in the use of generally expensive lenders and expensive choices within firms' existing menus. Poor borrower-lender fit accounts for little of this relationship, and costly menus play a smaller role than choices among existing partners.

Delegated search can leave borrowers exposed to conflicts of interest even when intermediaries offer access to many lenders. Expanding lender access alone will not ensure lower rates if brokers continue to favor expensive partners. Greater transparency about lender-paid compensation and brokers' past lender choices could help borrowers compare firms and help regulators identify steering practices that raise household borrowing costs.

References

- Agarwal, Sumit, Brent W. Ambrose, and Vincent W. Yao, 2020, Lender Steering in Residential Mortgage Markets, *Real Estate Economics* 48, 446–475.
- Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, and Douglas D. Evanoff, 2025, Predatory Lending and Hidden Risks, *Journal of Financial and Quantitative Analysis* 60, 2526–2554.
- Agarwal, Sumit, Swee Hoon Ang, Yongheng Deng, and Yonglin Wang, 2021, Mortgage Brokers and the Effectiveness of Regulatory Oversight, *Management Science* 67, 5278–5300.
- Alexandrov, Alexei, and Sergei Koulayev, 2017, No Shopping in the U.S. Mortgage Market: Direct and Strategic Effects of Providing Information, Working Paper.
- Allen, Jason, Robert Clark, and Jean-François Houde, 2014, Price Dispersion in Mortgage Markets, *Journal of Industrial Economics* 62, 377–416.
- Allen, Jason, Robert Clark, Jean-François Houde, Shaoteng Li, and Anna Trubnikova, 2025, The Role of Intermediaries in Selection Markets: Evidence from Mortgage Lending, *Review of Financial Studies* 38, 3284–3328.
- Ambrose, Brent W., and James N. Conklin, 2014, Mortgage Brokers, Origination Fees, Price Transparency and Competition, *Real Estate Economics* 42, 363–421.
- Ambrose, Brent W., James N. Conklin, and Luis A. Lopez, 2021, Does Borrower and Broker Race Affect the Cost of Mortgage Credit?, *Review of Financial Studies* 34, 790–826.
- Anagol, Santosh, Shawn Cole, and Shayak Sarkar, 2017, Understanding the Advice of Commissions-Motivated Agents: Evidence from the Indian Life Insurance Market, *Review of Economics and Statistics* 99, 1–15.
- Argyle, Bronson, Taylor Nadauld, and Christopher Palmer, 2023, Real Effects of Search Frictions in Consumer Credit Markets, *Review of Financial Studies* 36, 2685–2720.
- Barry, Jordan M., Will Fried, and John William Hatfield, 2025, Et Tu, Agent? Commission-Based Steering in Residential Real Estate, *Iowa Law Review* 110, 1473.
- Barwick, Panle Jia, Parag A. Pathak, and Maisy Wong, 2017, Conflicts of Interest and Steering in Residential Brokerage, *American Economic Journal: Applied Economics* 9, 191–222.
- Benetton, Matteo, Greg Buchak, and Claudia Robles-Garcia, 2022, Wide or Narrow? Competition and Scope in Financial Intermediation, Working Paper.
- Bergstresser, Daniel, John M. R. Chalmers, and Peter Tufano, 2009, Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry, *Review of Financial Studies* 22, 4129–4156.
- Berndt, Antje, Burton Hollifield, and Patrik Sandås, 2010, The Role of Mortgage Brokers in the Subprime Crisis, Working Paper.
- Bhutta, Neil, Andreas Fuster, and Aurel Hizmo, 2026, Paying Too Much? Borrower Sophistication and Overpayment in the U.S. Mortgage Market, *Journal of Finance* 81, 49–90.
- Bosshardt, Joshua, Ali Kakhbod, and Amir Kermani, 2025, Do Intermediaries Improve GSE Lending? Evidence from Proprietary GSE Data, *Journal of Financial Economics* 170, 104082.

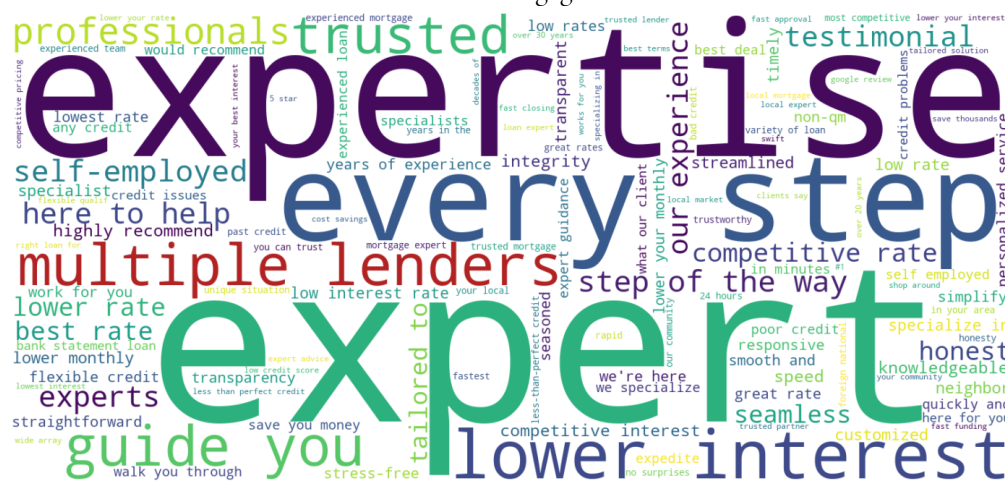
- Burdett, Kenneth, and Kenneth L. Judd, 1983, Equilibrium Price Dispersion, *Econometrica* 51, 955–969.
- Christoffersen, Susan E. K., Richard Evans, and David K. Musto, 2013, What Do Consumers’ Fund Flows Maximize? Evidence from Their Brokers’ Incentives, *Journal of Finance* 68, 201–235.
- Egan, Mark, 2019, Brokers versus Retail Investors: Conflicting Interests and Dominated Products, *Journal of Finance* 74, 1217–1260.
- Egan, Mark, Shan Ge, and Johnny Tang, 2022, Conflicting Interests and the Effect of Fiduciary Duty: Evidence from Variable Annuities, *Review of Financial Studies* 35, 5334–5386.
- Green, Richard C., Burton Hollifield, and Norman Schürhoff, 2007, Financial Intermediation and the Costs of Trading in an Opaque Market, *Review of Financial Studies* 20, 275–314.
- Griffin, John M., Nicholas Hirschey, and Samuel Kruger, 2023, Do Municipal Bond Dealers Give Their Customers “Fair and Reasonable” Pricing?, *Journal of Finance* 78, 887–934.
- Griffin, John M., Nicholas Hirschey, and Samuel Kruger, 2026, What Drives Large Corporate Bond Trading Costs?, Working Paper.
- Guiso, Luigi, Andrea Pozzi, Anton Tsoy, Leonardo Gambacorta, and Paolo Emilio Mistrulli, 2022, The Cost of Steering in Financial Markets: Evidence from the Mortgage Market, *Journal of Financial Economics* 143, 1209–1226.
- Gurun, Umit G., Gregor Matvos, and Amit Seru, 2016, Advertising Expensive Mortgages, *Journal of Finance* 71, 2371–2416.
- Hatfield, John William, Scott Duke Kominers, and Richard Lowery, 2025, Collusion in Brokered Markets, *Journal of Finance* 80, 1417–1462.
- Hortaçsu, Ali, and Chad Syverson, 2004, Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds, *Quarterly Journal of Economics* 119, 403–456.
- Huang, Ruidi, Erik J. Mayer, Sheridan Titman, and David Xu, 2026, Human Capital and Local Credit Supply: Evidence from the Mortgage Industry, Working Paper.
- Inderst, Roman, and Marco Ottaviani, 2012, Competition through Commissions and Kickbacks, *American Economic Review* 102, 780–809.
- Jorgensen, Rebecca A., 2025, The Economic Consequences of Mergers Between Real Estate Agencies and Mortgage Lenders, Working Paper.
- LaCour-Little, Michael, 2009, The Pricing of Mortgages by Brokers: An Agency Problem?, *Journal of Real Estate Research* 31, 235–264.
- Li, Anran, Tong Liu, Anthony LoSasso, and Nicholas Tilipman, 2026, Going for Broker? Intermediation in Health Insurance Markets, Working Paper.
- Lopez, Luis Arturo, Shawn J. McCoy, and Vivek Sah, 2022, Steering Consumers to Lenders in Residential Real Estate Markets, *Real Estate Economics* 50, 1596–1641.
- Paravisini, Daniel, Veronica Rappoport, and Philipp Schnabl, 2023, Specialization in Bank Lending: Evidence from Exporting Firms, *Journal of Finance* 78, 2049–2085.

- Priest, Alex, 2023, Does Credit Supply Affect Real Estate Cap Rates?, Working Paper.
- Robles-Garcia, Claudia, 2020, Competition and Incentives in Mortgage Markets: The Role of Brokers, Working Paper.
- Spader, Jonathan S., and Roberto G. Quercia, 2011, Mortgage Brokers and the Refinancing Transaction: Evidence from CRA Borrowers, *Journal of Real Estate Finance and Economics* 42, 181–210.
- Stigler, George J., 1961, The Economics of Information, *Journal of Political Economy* 69, 213–225.
- 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* 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.
- Zhang, Dayin, Panle Jia Barwick, Lu Han, and Jonathan Kroah, 2026, Market Power in Mortgage Pricing: The Role of Referral Lending, Working Paper.

Figure 1: Broker Advertising and Lender Use

This figure describes mortgage brokers' advertising. Panel A shows common phrases on broker websites; phrase size is proportional to the number of firms using it. Panel B compares, for each broker firm whose website states a number of lenders, that number with the number of distinct lenders that funded the firm's brokered CoreLogic mortgages in 2025. Bubble size is proportional to the firm's brokered loans in 2025. The gray dashed line marks equal advertised and actual counts, and the share of firms below it is reported along the line. The dark red dashed lines mark 20 lenders; the braces report the share of firms advertising at least 20 lenders (top) and the share using fewer than 20 (right). The marginal histograms show the share of firms (filled bars) and the share of their 2025 loans (hollow bars) in bins of five lenders. The axes are cut at 103 lenders; the histograms and shares use all firms. Websites are obtained from NMLS records and supplementary company-name searches and were collected in June and September 2026. Panel B uses the 637 broker firms that state a number of lenders on their website and have at least one brokered CoreLogic mortgage in 2025.

Panel A: Phrases on Mortgage Broker Websites



Panel B: Advertised and Actual Number of Lenders

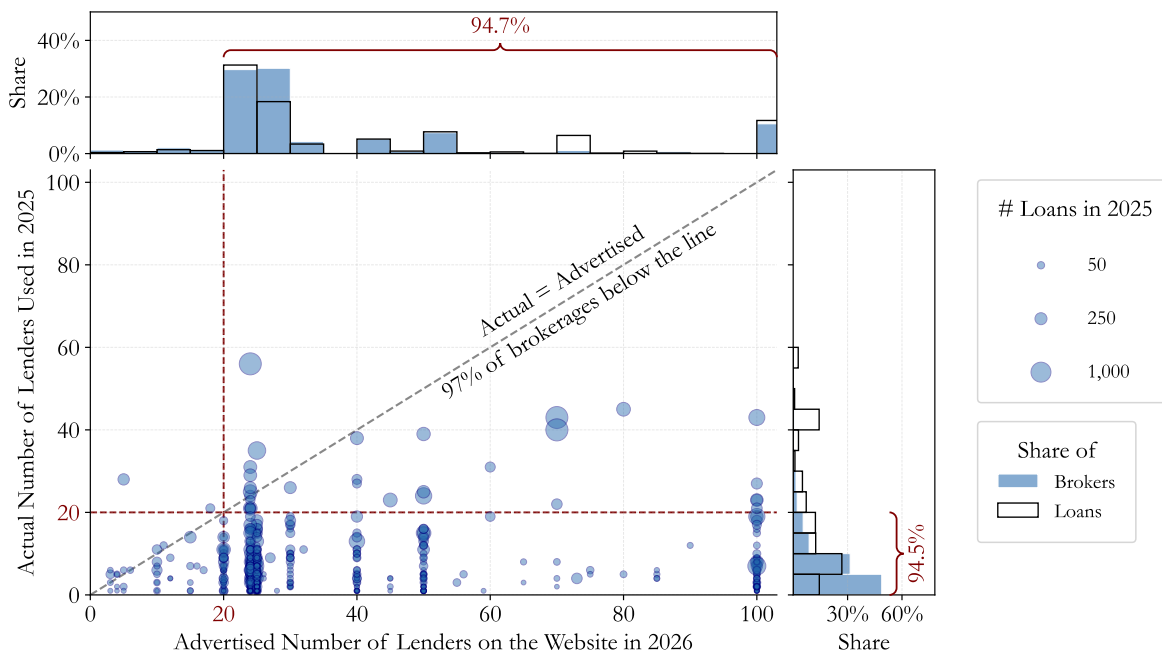


Figure 2: Residual Rate Dispersion of Retail and Brokered Loans

This figure compares rate dispersion across origination channels. It shows residual rate spreads for retail loans (gray bars) and brokered loans (red line). Residuals come from a pooled regression of the rate spread on a brokered indicator, income, credit score, loan amount, property value, combined loan-to-value ratio, debt-to-income ratio, origination charges, race, sex, and county $\times$ lock-date fixed effects. The legend reports residual standard deviations adjusted for absorbed degrees of freedom. The rate spread is APR minus APOR, in basis points. The sample is conventional conforming fixed-rate loans in the matched Optimal Blue, CoreLogic, and HMDA data, 2018–2025; the displayed range is trimmed at the pooled 0.5th and 99.5th percentiles.

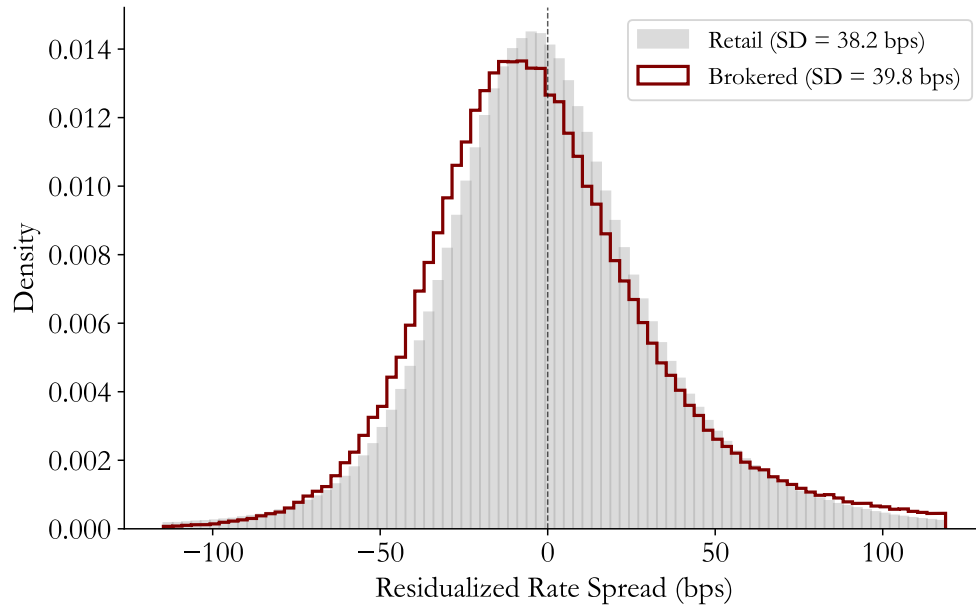


Figure 3: Broker Firm Price Levels Relative to Retail

This figure reports broker firm price levels relative to retail loans. I regress the rate spread on broker firm indicators, borrower and loan controls, and county $\times$ lock-date fixed effects, with retail loans as the omitted group. Dots show estimates for the 170 largest broker firms, ordered by their effects. Blue and red identify estimates significantly below and above zero; white identifies the remaining estimates. Confidence intervals and significance classifications use a Bonferroni correction for 170 comparisons. Controls include income, credit score, loan amount, property value, combined loan-to-value ratio, debt-to-income ratio, origination charges, race, and sex. The rate spread is APR minus APOR, in basis points. The sample is matched Optimal Blue, CoreLogic, and HMDA loans, 2018–2025, in county-days containing a brokered loan. Standard errors are clustered by broker firm (lender for retail loans), county, and lock date.

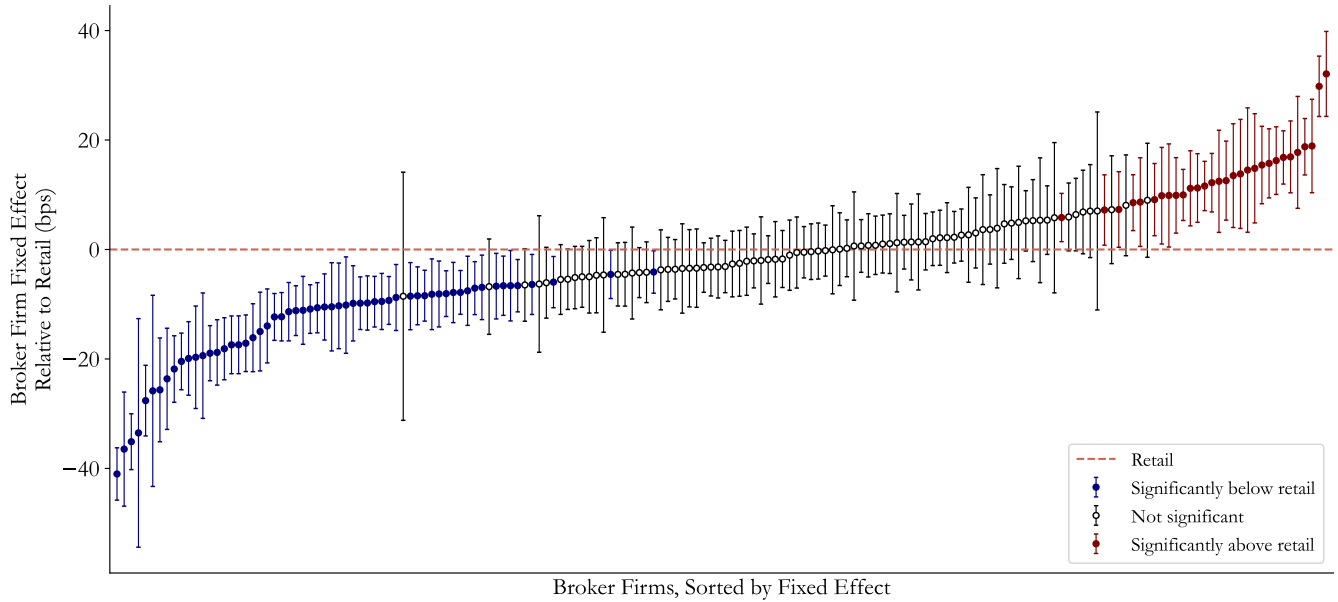


Figure 4: Broker Firm Price Levels Across Periods and Loan Types

This figure compares broker firm price levels across periods and loan types. Firm effects are estimated separately in each group from regressions of rate spread on firm indicators, borrower and loan controls, and county $\times$ lock-date fixed effects, with retail loans in that group as the benchmark. Each panel is a binned scatter plot with 20 bins, a linear fit, and a 95% confidence band. The comparisons are 2018–2021 versus 2022–2025; purchase versus refinance; conventional versus government-backed; credit scores at least 740 versus below 740; 30-year versus 10–25-year terms; and combined loan-to-value ratios at most 80% versus above 80%. Firms must have at least 10 brokered loans in each group and are weighted by the harmonic mean of their group loan counts. Controls include income, credit score, loan amount, property value, combined loan-to-value ratio, debt-to-income ratio, origination charges, race, and sex. Effects are in basis points. The sample is matched Optimal Blue, CoreLogic, and HMDA loans, 2018–2025. Standard errors are clustered by broker firm.

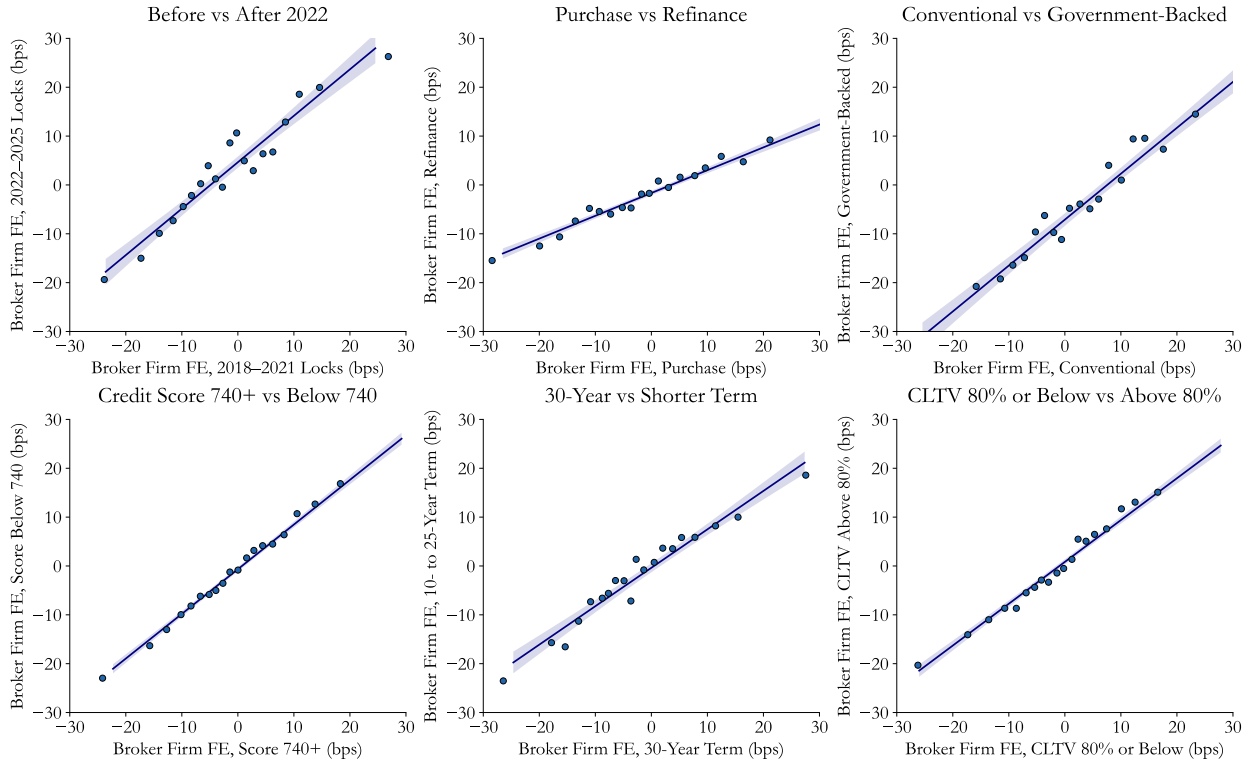


Figure 5: Distance, Broker Choice, and the First Stage

This figure examines proximity and broker choice. Panel A reports a linear probability regression of branch choice on distance-bin indicators, with loan and branch $\times$ lock-month fixed effects. Each observation pairs a loan with a potential branch. Dots are coefficients in percentage points, relative to branches at least 20 miles away, and bars are 95% confidence intervals. Alternatives are sampled and weighted by inverse selection probabilities: the chosen branch, up to 100 branches within 20 miles of the origin or property, and 30 other branches. Panel B presents the first stage of the differential distance IV. The horizontal axis is distance to the nearest low-cost branch minus distance to the nearest high-cost branch. High-cost firms have a previous-year price effect at or above the monthly median across branches. The vertical axis is high-cost broker use. Dots are 20 equal-count bins after adjustment for credit-score, loan-to-value, and debt-to-income deciles, origin characteristics, nearest-branch distance, and origin and mortgage county $\times$ lock-month effects. Bars are 95% confidence intervals. In both panels, the choice set contains branches active in the previous year whose firms are licensed in the property's state. The origin is the borrower's previous address for purchases, traced using Verisk and DataAxle, and the mortgaged property for refinances. The data link Optimal Blue, CoreLogic, HMDA, and NMLS. Panel B uses the 2019–2025 IV sample in Table 3. Standard errors are clustered by origin county in Panel A and mortgage county in Panel B.

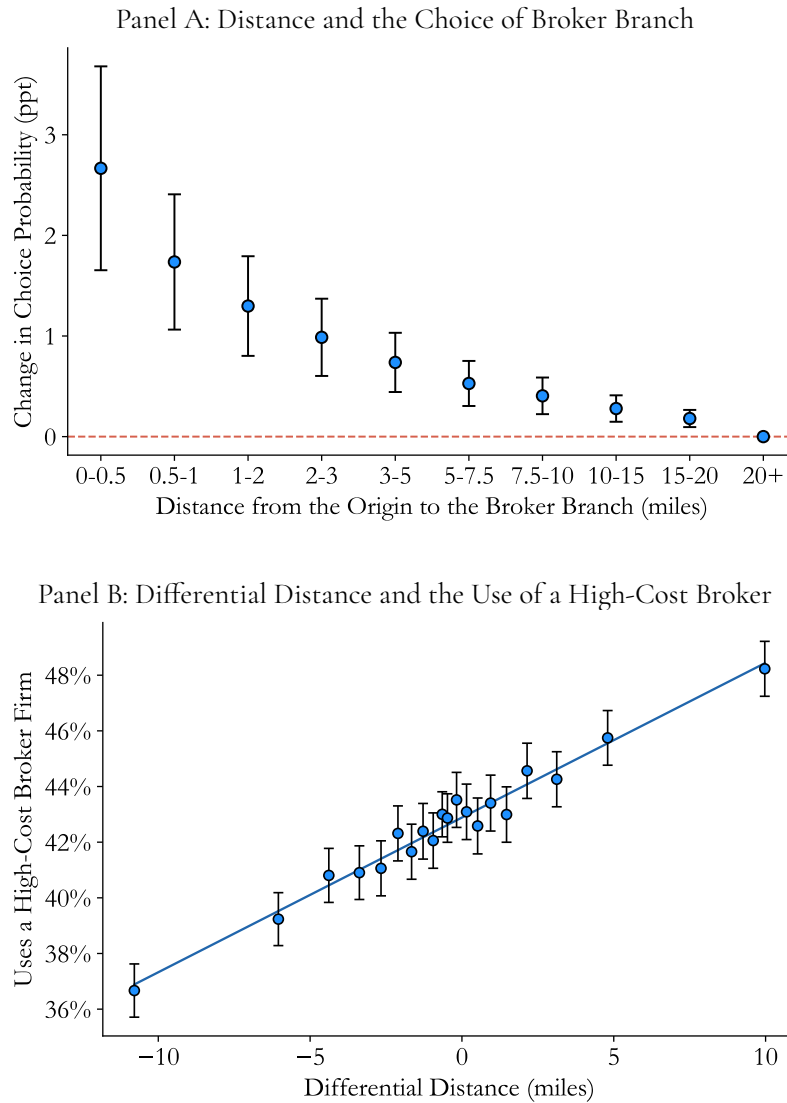


Figure 6: Balance Tests for the Differential Distance Instrument

This figure reports associations between differential distance and borrower, neighborhood, and selected-broker characteristics. Differential distance is distance to the nearest low-cost branch minus distance to the nearest high-cost branch; high-cost firms are classified using previous-year price effects. Each dot comes from a separate regression of a characteristic on differential distance, with mortgage county $\times$ lock-month fixed effects and no other controls. Both variables are divided by their standard deviations, so coefficients measure standard-deviation changes per one-standard-deviation (12.0-mile) increase in distance. Bars are 95% confidence intervals. The three groups cover borrower and loan characteristics, origin-neighborhood characteristics from the American Community Survey, and the selected broker's website claims. Claims are indicators; advertised lender count is log one plus the stated number. Websites were collected in 2026, so these measures describe the chosen broker rather than predetermined borrower characteristics. The sample is owner-occupied brokered loans in the matched Optimal Blue, CoreLogic, and HMDA data with traced origins, 2019–2025. Standard errors are clustered by mortgage county.

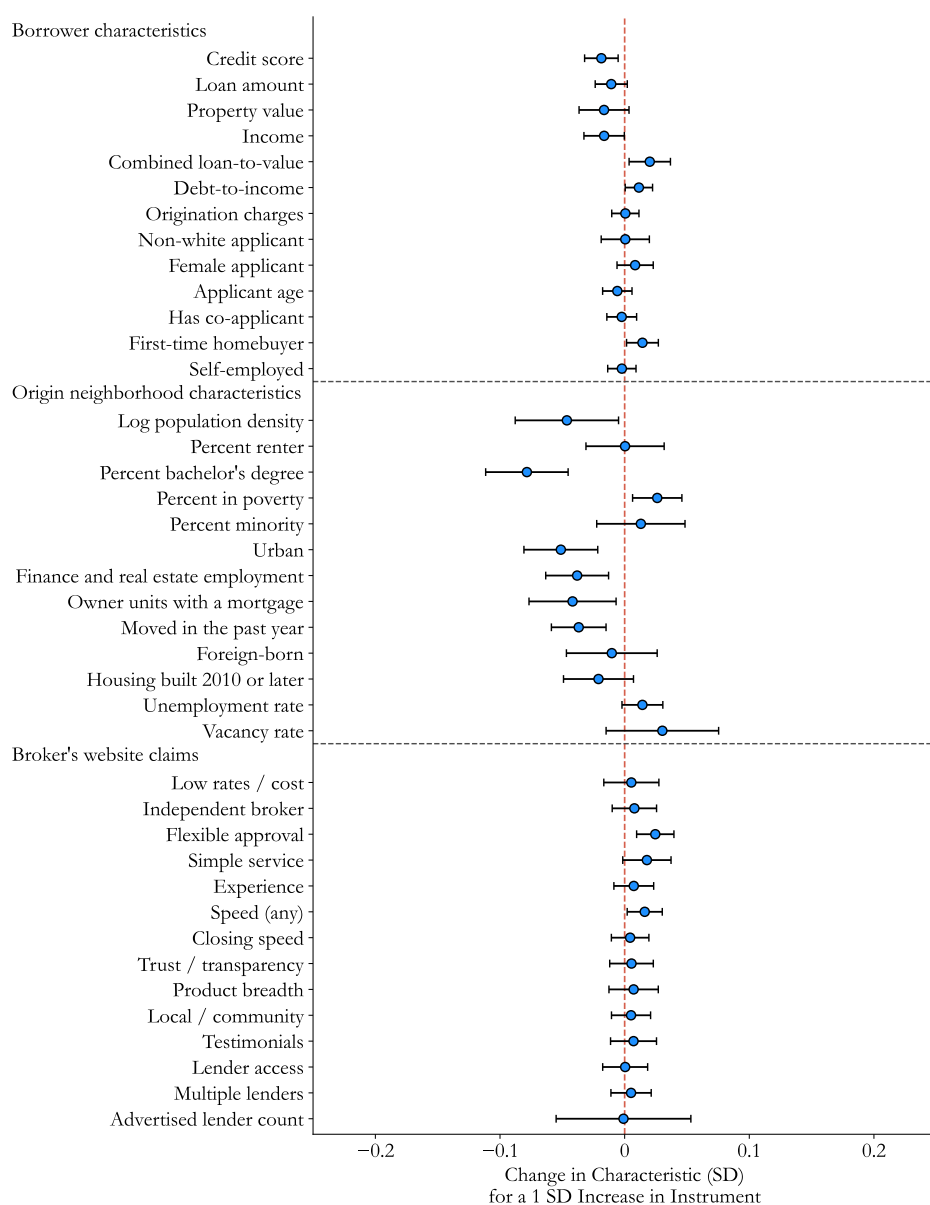
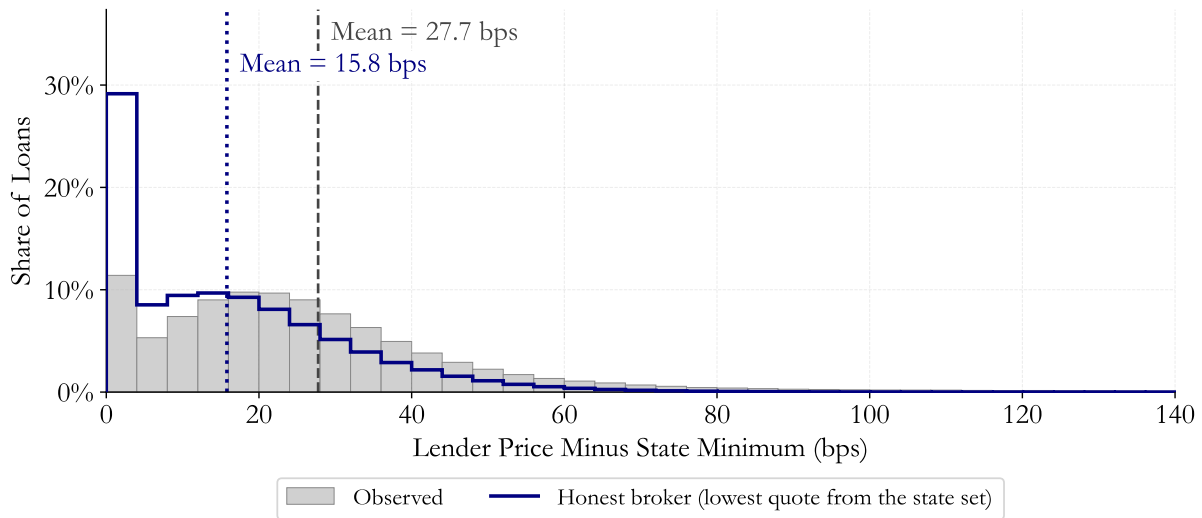


Figure 7: The Steering Gap

This figure compares lender-choice gaps and relates them to realized mortgage rates. The observed gap is the chosen lender's predicted spread minus the lowest predicted spread in the state set, which includes lenders active with at least 10 brokers in the state during the previous year and the focal broker's own lenders. Prices are predicted using lender-specific borrower-risk coefficients and broker, state-year, and lock-month effects. The honest broker selects the lowest quote, where quotes equal predicted spreads plus independent normal noise with a standard deviation of 35 basis points. Its expected observed gap is the honest-broker gap; the steering gap subtracts this expectation from the observed gap. Panel A compares observed gaps (gray bars) with gaps under honest-broker choices (dark line), weighting each possible lender by its probability of quoting lowest. Vertical lines mark the respective means. Panel B shows steering gaps in 4-basis-point bins (bars, left axis) and realized spreads in a 30-bin scatter plot (dots, right axis), adjusted for borrower and loan characteristics and county $\times$ lock-month effects; the red dashed line marks a steering gap of zero. All gaps and spreads are in basis points. The sample consists of brokered conventional conforming fixed-rate loans in the matched Optimal Blue, CoreLogic, and HMDA data, 2018–2025, with usable lender-price predictions. There are 308,810 loans; means include observations outside the displayed range.

Panel A: Observed and Honest-Broker Gaps



Panel B: Distribution of the Steering Gap and Realized Rates

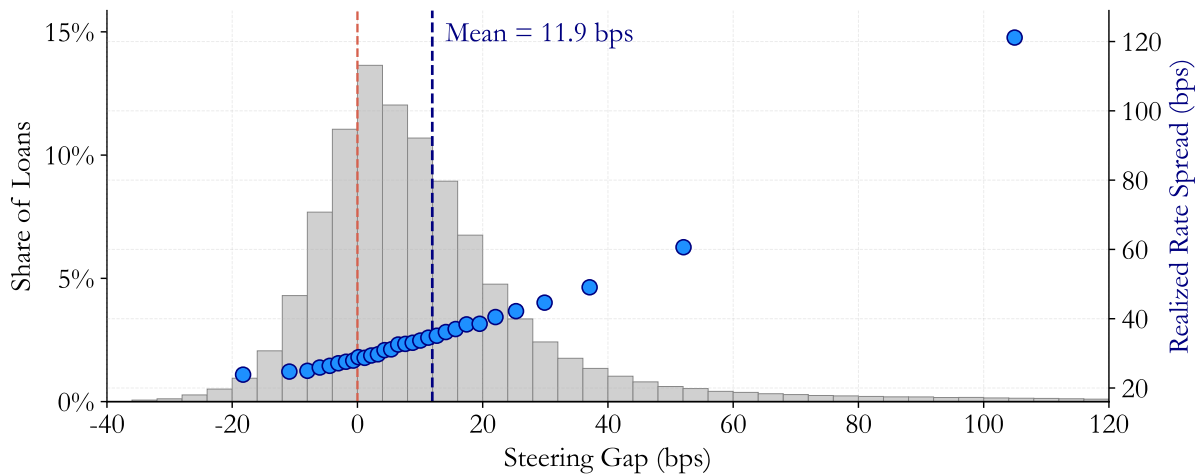


Figure 8: The Steering Gap Across the Largest Broker Firms

This figure reports the mean steering gaps of the 25 largest broker firms in the steering analysis, ordered from lowest to highest, with 95% confidence intervals. The observed gap compares the chosen lender's predicted spread with the lowest predicted spread in the state set. The honest-broker gap is the expected observed gap when the broker selects the lowest of the state lenders' noisy quotes, with independent normal noise of 35 basis points. Their difference is the steering gap. The dashed line reports the mean across all loans. Gaps are in basis points, and confidence intervals use within-firm variation. The sample consists of brokered conventional conforming fixed-rate loans in the matched Optimal Blue, CoreLogic, and HMDA data, 2018–2025, with usable lender-price predictions.

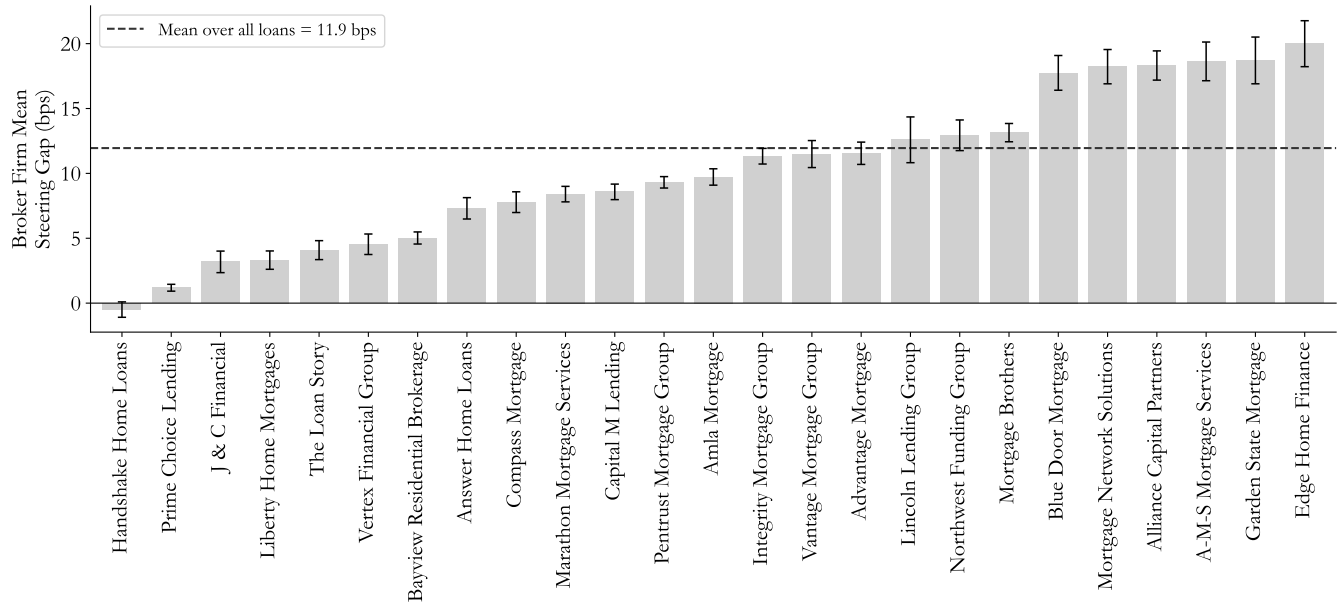


Figure 9: Past and Current Steering Gaps of Broker Firms

This figure compares current and past steering gaps. Panel A relates each loan's gap to its firm's mean gap during the previous 12 months, adjusting for borrower and loan characteristics, lender concentration, county $\times$ lock-month effects, and firm effects. Panel B compares firm means in months $t - 12$ to $t - 1$ with means in months $t - 24$ to $t - 13$, adjusting for county $\times$ month effects. Each observation in Panel B is a firm-month. Each window requires at least five loans. Dots are binned means; lines are linear fits with 95% confidence bands. The steering gap is the chosen lender's predicted spread minus the expected predicted spread selected by a broker taking the lowest quote in the state lender set. Quotes equal predicted spreads plus independent normal noise with a standard deviation of 35 basis points. Gaps are in basis points. The sample consists of brokered conventional conforming fixed-rate loans in the matched Optimal Blue, CoreLogic, and HMDA data, 2018–2025, with usable lender-price predictions. Standard errors are clustered by county.

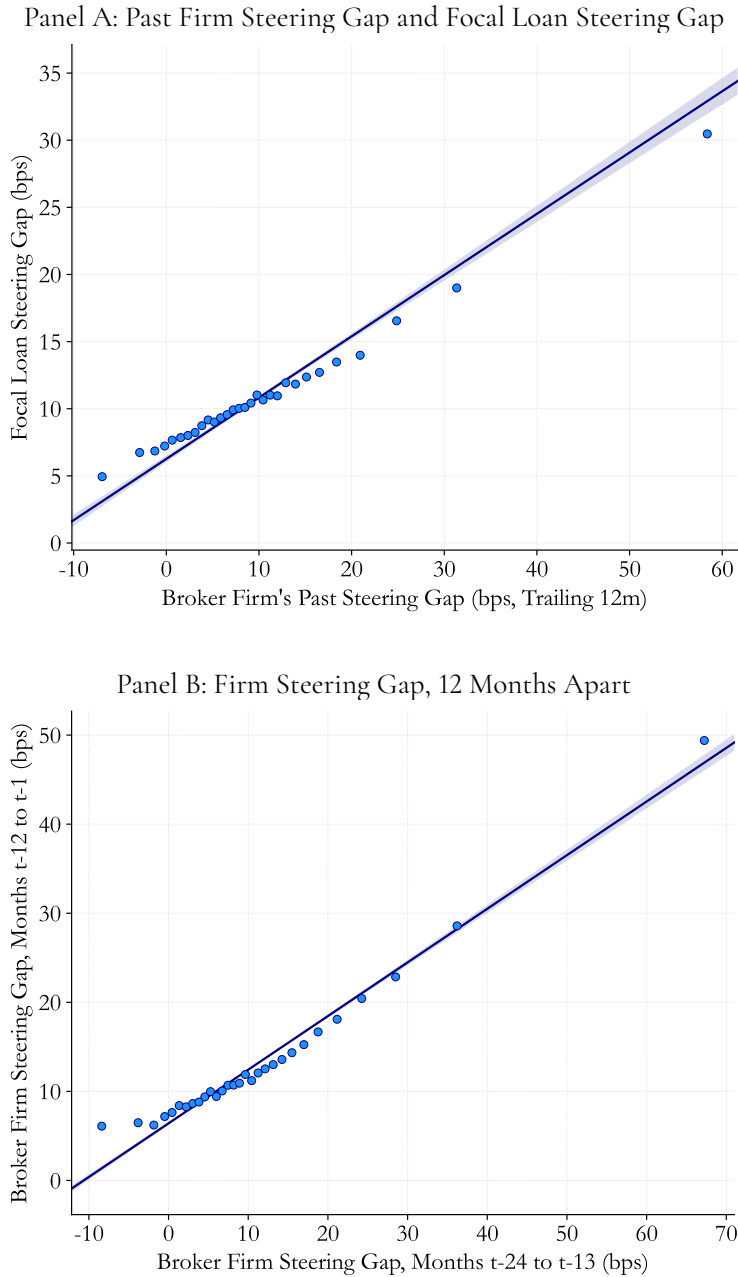


Figure 10: Past Firm Steering Gap and Mortgage Rates

This figure relates realized mortgage rates to a broker firm's past lender choices. Dots show a 30-bin scatter plot of the rate spread against the firm's mean steering gap over the previous 12 months, with at least five loans. The plot adjusts for borrower and loan characteristics, lender concentration, county $\times$ lock-month effects, and broker firm effects. The line is a linear fit with a 95% confidence band. The steering gap is the chosen lender's predicted spread minus the expected predicted spread selected by a broker taking the lowest quote in the state lender set. Quotes equal predicted spreads plus independent normal noise with a standard deviation of 35 basis points. Spreads and gaps are in basis points. The sample consists of brokered conventional conforming fixed-rate loans in the matched Optimal Blue, CoreLogic, and HMDA data, 2018–2025, with usable lender-price predictions. Standard errors are clustered by county.

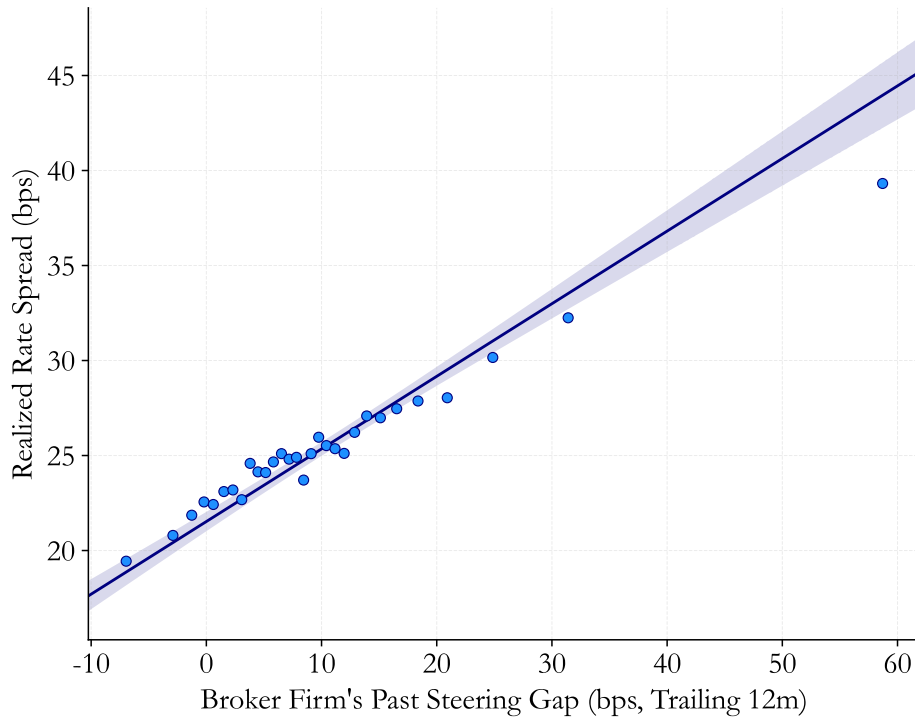


Figure 11: Past Firm Steering Gap and Mortgage Rates, by Part of the Gap

This figure decomposes the relationship between a firm's past lender choices and subsequent borrowers' rate spreads. Panel A separates general lender prices, broker-lender relationship prices, and the pricing of borrower characteristics. Panel B separates lender choice within the broker's prior-year state menu from menu choice relative to the state lender set. The observed gap compares the chosen lender's predicted spread with the lowest in the state set. The steering gap instead uses the expected predicted spread selected under honest shopping, with independent normal quote noise of 35 basis points. The menu decomposition applies these benchmarks separately to the menu and state set. Past gaps and components are firm averages over the previous 12 months, requiring at least five loans, scaled by the standard deviation of the past gap. Each contribution is the component's coefficient in a joint regression of rate spreads on all components, multiplied by its coefficient from a regression on the total past gap using the same controls. Contributions sum to the coefficient on the past gap and are in basis points per standard deviation, with 95% confidence intervals. Regressions control for borrower and loan characteristics, lender concentration, county $\times$ lock-month effects, and firm effects. The sample is brokered loans with a measured steering gap. Standard errors are clustered by county.

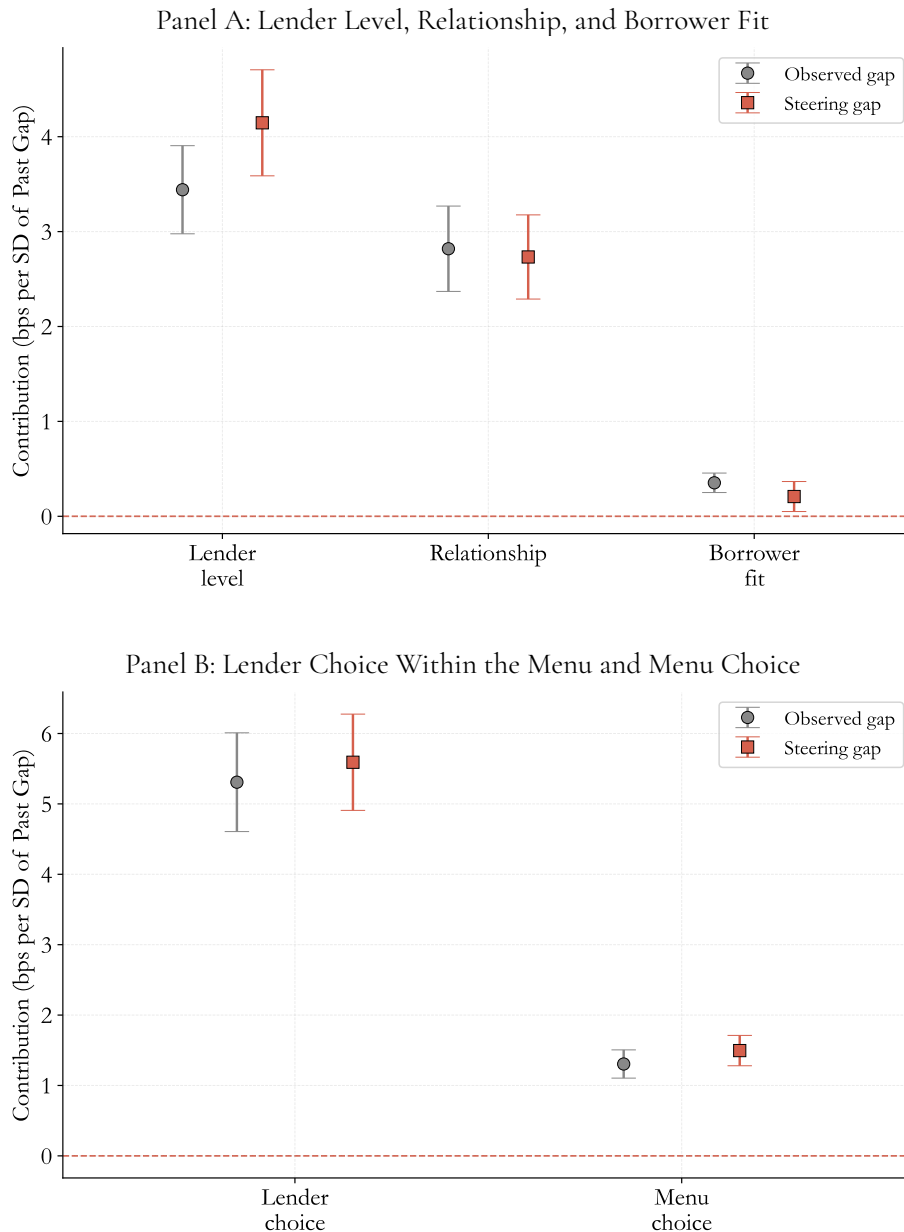


Figure 12: The Steering Gap and Broker Compensation

This figure relates lender-paid broker compensation to the steering gap. Dots show a 30-bin scatter plot after adjustment for borrower and loan characteristics, county $\times$ lock-month effects, and broker firm effects. The line is a linear fit with a 95% confidence band. Compensation is measured in basis points of the loan amount; the steering gap is in basis points of rate spread. The steering gap is the chosen lender's predicted spread minus the expected predicted spread selected by a broker taking the lowest quote in the state lender set. Quotes equal predicted spreads plus independent normal noise with a standard deviation of 35 basis points. The sample consists of brokered conventional conforming fixed-rate loans in the matched Optimal Blue, CoreLogic, and HMDA data, 2018–2025, with usable lender-price predictions. Only loans with recorded compensation are included. Standard errors are clustered by county.

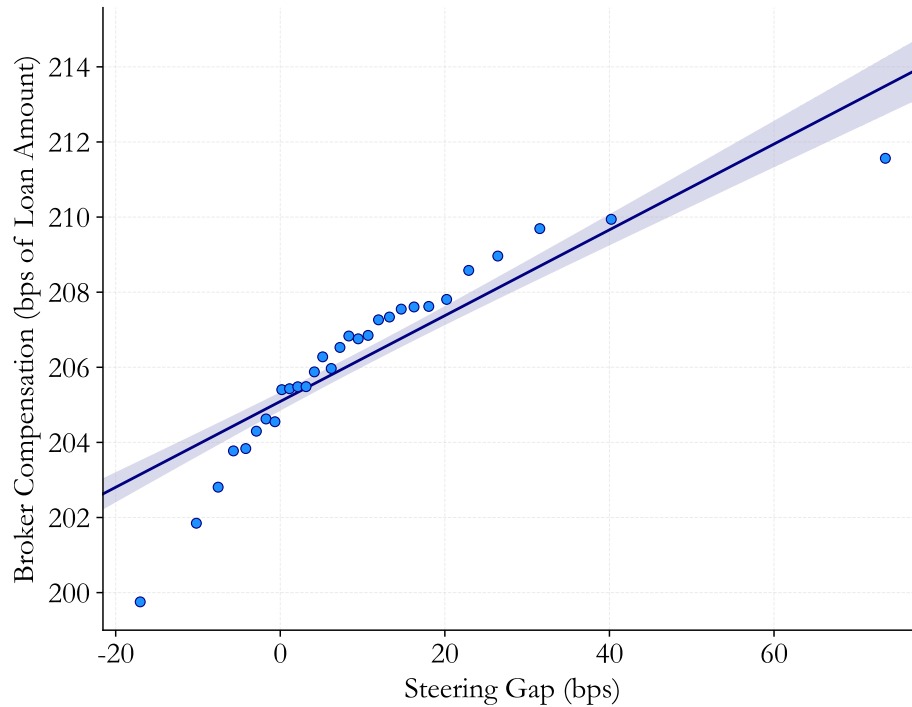


Table 1: Summary Statistics of Brokered Loans

This table reports summary statistics for brokered loans. Panel A summarizes mortgage prices and broker compensation. Rate spread is APR minus APOR. Broker compensation is recorded in Optimal Blue in basis points of the loan amount. Compensation paid by the lender equals one if the lender pays the broker and zero if the borrower does, for loans whose Optimal Blue record states who pays. Variables marked (1/0) are indicators. Panel B reports borrower characteristics. Minority identifies non-White applicants among those reporting race, with joint applications coded as non-minority. Panel C reports loan characteristics. Panel D describes broker firm and officer lender networks: lender counts and HHI use all brokered CoreLogic loans during the 12 months before the lock month. HHI is the sum of squared lender loan shares, multiplied by 10,000. The sample is conventional conforming fixed-rate loans in the matched Optimal Blue, CoreLogic, and HMDA data closing in 2018–2025, with complete borrower and loan controls. Rate spread, income, combined loan-to-value ratio, and origination charges are winsorized at the 1st and 99th percentiles by closing year. Each variable uses its available observations; Obs. reports that count.

	Mean	SD	P10	Median	P90	Obs.
<i>Panel A: Prices</i>						
Rate spread (bps)	31.3	48.2	−17.8	22.7	92.0	335,997
Note rate (%)	3.99	1.53	2.62	3.38	6.62	335,991
Origination charges (\$)	3,381	3,344	197	1,992	8,295	335,997
Broker compensation (bps of loan amount)	207.6	51.0	137.5	200.0	275.0	186,991
Compensation paid by the lender (1/0)	0.633	0.482	0.000	1.000	1.000	320,440
<i>Panel B: Borrower Characteristics</i>						
Credit score	752	44	688	761	802	335,997
Income (\$000s)	130.4	86.5	52.0	109.0	229.0	335,997
Debt-to-income ratio (%)	37.4	7.7	30.0	37.0	47.0	335,997
Combined loan-to-value ratio (%)	73.0	17.9	46.9	77.2	95.0	335,997
Minority applicant (1/0)	0.175	0.380	0.000	0.000	1.000	290,929
Female applicant (1/0)	0.236	0.424	0.000	0.000	1.000	318,253
Has co-applicant (1/0)	0.444	0.497	0.000	0.000	1.000	335,940
<i>Panel C: Loan Characteristics</i>						
Loan amount (\$000s)	352.7	164.6	160.0	330.0	569.4	335,997
Property value (\$000s)	523.1	316.8	225.0	455.0	875.0	335,997
Purchase loan (1/0)	0.485	0.500	0.000	0.000	1.000	335,997
Cash-out refinance (1/0)	0.193	0.395	0.000	0.000	1.000	335,997
30-year term (1/0)	0.857	0.350	0.000	1.000	1.000	335,997
Principal residence (1/0)	0.954	0.210	1.000	1.000	1.000	335,997
<i>Panel D: Broker Lender Networks, Previous 12 Months</i>						
Lenders used by the broker firm	6.5	6.6	2.0	5.0	13.0	324,802
Broker firm lender HHI (0–10,000)	4,895	2,480	2,099	4,357	9,232	324,802
Lenders used by the loan officer	3.6	2.4	1.0	3.0	7.0	301,859
Loan officer lender HHI (0–10,000)	5,737	2,639	2,595	5,122	10,000	301,859
Brokered loans	335,997					
Broker firms	15,409					
Broker loan officers	31,377					
Lenders	334					
Counties	1,505					

Table 2: Broker Firm Fixed Effects and the Variation in Brokered Rates

This table compares residual rate-spread variation with different intermediary fixed effects. All columns regress brokered-loan rate spreads on income, credit score, loan amount, property value, combined loan-to-value ratio, debt-to-income ratio, origination charges, race, sex, and county $\times$ lock-date effects. Column (2) adds broker firm effects, Column (3) adds lender effects, and Column (4) includes both. The share of baseline residual variance explained compares each column's degrees-of-freedom-adjusted residual variance with Column (1). Rate spread is APR minus APOR, in basis points. The sample is matched Optimal Blue, CoreLogic, and HMDA brokered loans, 2018–2025, in county-date cells with at least two brokered loans, after dropping singleton firms and lenders.

Dep. Variable:	Rate Spread (bps)			
	(1)	(2)	(3)	(4)
R^2	0.557	0.631	0.599	0.658
Adjusted R^2	0.383	0.458	0.441	0.496
Residual SD (bps)	35.07	32.89	33.38	31.72
Share of baseline residual variance explained		0.121	0.094	0.182
County $\times$ Lock-Date FE	✓	✓	✓	✓
Broker Firm FE		✓		✓
Lender FE			✓	✓
Borrower and loan controls	✓	✓	✓	✓
Broker firms		7,969		7,969
Observations	207,573	207,573	207,573	207,573
Dep. Var. Mean	27.0	27.0	27.0	27.0
Dep. Var. SD	44.7	44.7	44.7	44.7

Residual SDs are degrees-of-freedom corrected for the absorbed fixed effects.

Table 3: High-Cost Brokers and Mortgage Rates: Differential Distance IV

This table estimates the relationship between high-cost broker use and rate spreads. A firm is high-cost if its previous-year price effect, estimated relative to retail with borrower controls and local market effects and shrunk toward the monthly mean, is at or above the monthly median across branches. Panel A reports OLS estimates. Panel B instruments the high-cost indicator, marked with a hat, with distance to the nearest low-cost branch minus distance to the nearest high-cost branch, in units of 10 miles. Panel B also reports first-stage F -statistics. The dependent variable is APR minus APOR, in basis points. Column (1) includes origin county $\times$ lock-month effects; Column (2) adds mortgage county $\times$ lock-month effects; Columns (3)–(5) additionally include lock-date effects, lender $\times$ lock-month effects, or both. Borrower controls are credit-score, combined loan-to-value, and debt-to-income deciles. Origin controls include nearest-branch distance, log population density and median household income, renter, college-educated, poverty, and minority shares, an urban indicator, and tract shares for finance and real estate employment, mortgaged owner units, recent movers, foreign-born residents, recent housing, unemployment, and vacancy. The origin is the borrower's previous address for purchases, traced using Verisk and DataAxle, and the mortgaged property for refinances. The sample is owner-occupied brokered loans in the matched Optimal Blue, CoreLogic, and HMDA data, 2019–2025, with traced origins, a lagged firm classification, and absolute differential distance below its 95th percentile. Standard errors are clustered by mortgage county.

Dep. Variable:	Rate Spread (bps)				
	(1)	(2)	(3)	(4)	(5)
<i>Panel A: OLS</i>					
$\mathbb{1}(\text{High-Cost Broker Firm})$	10.86*** (0.44)	10.35*** (0.46)	10.30*** (0.45)	8.80*** (0.50)	8.83*** (0.49)
Observations	143,970	137,675	137,505	132,245	132,065
Dep. Var. Mean	23.25	22.62	22.62	22.36	22.35
Dep. Var. SD	43.15	42.29	42.27	41.91	41.89
<i>Panel B: 2SLS</i>					
$\mathbb{1}(\text{High-Cost Broker Firm})$	11.05** (4.70)	10.93** (4.67)	10.43** (4.72)	12.17** (5.10)	11.21** (5.20)
First-stage F -statistic	93.7	85.9	84.8	100.4	101.1
Observations	143,970	137,675	137,505	132,245	132,065
Dep. Var. Mean	23.25	22.62	22.62	22.36	22.35
Dep. Var. SD	43.15	42.29	42.27	41.91	41.89
Origin County $\times$ Lock-Month FE	✓	✓	✓	✓	✓
Mortgage County $\times$ Lock-Month FE		✓	✓	✓	✓
Lock-Date FE			✓		✓
Lender $\times$ Lock-Month FE				✓	✓
Borrower controls	✓	✓	✓	✓	✓
Origin controls	✓	✓	✓	✓	✓

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: The Steering Gap and Realized Rates

This table regresses realized rate spreads on the loan's steering gap, both in basis points. The steering gap is the chosen lender's predicted spread minus the expected predicted spread selected by a broker taking the lowest quote in the state lender set. Quotes equal predicted spreads plus independent normal noise with a standard deviation of 35 basis points. Column (1) includes county $\times$ lock-month effects. Column (2) adds firm effects, Column (3) adds standardized firm HHI over the previous 12 months, and Column (4) adds lock-date effects. HHI is the sum of squared lender shares in the firm's loans. Controls are income, credit score, loan amount, property value, combined loan-to-value ratio, debt-to-income ratio, race, ethnicity, sex, occupancy, property type, number of units, loan purpose, and loan term. The sample consists of brokered conventional conforming fixed-rate loans in the matched Optimal Blue, CoreLogic, and HMDA data, 2018–2025, with usable lender-price predictions. Standard errors are clustered by county.

Dep. Variable:	Rate Spread (bps)			
	(1)	(2)	(3)	(4)
Steering gap (bps)	0.786*** (0.017)	0.729*** (0.016)	0.730*** (0.016)	0.726*** (0.016)
Broker firm HHI (std.)			0.587*** (0.149)	0.479*** (0.155)
County $\times$ Lock-Month FE	✓	✓	✓	✓
Lock-Date FE				✓
Broker Firm FE		✓	✓	✓
Borrower and loan controls	✓	✓	✓	✓
Observations	308,810	304,526	297,131	296,937
Adjusted R^2	0.497	0.547	0.547	0.573
Dep. Var. Mean	30.9	30.5	30.5	30.5
Dep. Var. SD	49.7	49.1	49.0	49.0

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Past and Current Steering Gaps

This table relates current steering gaps to past firm gaps. Panel A regresses a loan's gap on its firm's mean gap during the previous 12 months. All columns include borrower and loan controls and standardized firm HHI. Column (1) includes county $\times$ lock-month effects, Column (2) adds firm effects, and Column (3) adds lock-date effects. Panel B regresses the firm's mean gap in months $t - 12$ to $t - 1$ on its mean in months $t - 24$ to $t - 13$. Observations are firm-months assigned to the firm's modal county of the year. Columns (1)–(3) include county $\times$ month, state $\times$ month, or month effects, without firm effects; singleton county-months are excluded in Column (1). Each window requires at least five loans. The steering gap is the chosen lender's predicted spread minus the expected predicted spread selected by a broker taking the lowest quote in the state lender set. Quotes equal predicted spreads plus independent normal noise with a standard deviation of 35 basis points. Gaps are in basis points. Controls are income, credit score, loan amount, property value, combined loan-to-value ratio, debt-to-income ratio, race, ethnicity, sex, occupancy, property type, number of units, loan purpose, and loan term. The sample consists of brokered conventional conforming fixed-rate loans in the matched Optimal Blue, CoreLogic, and HMDA data, 2018–2025, with usable lender-price predictions. Standard errors are clustered by county.

Panel A: Past Firm Steering Gap and Focal Loan Steering Gap			
Dep. Variable:	Focal Loan Steering Gap (bps)		
	(1)	(2)	(3)
Past firm steering gap (bps, 12 months)	0.713*** (0.010)	0.456*** (0.019)	0.455*** (0.020)
Broker firm HHI (std.)	0.038 (0.072)	0.028 (0.119)	0.036 (0.118)
County $\times$ Lock-Month FE	✓	✓	✓
Lock-Date FE			✓
Broker Firm FE		✓	✓
Borrower and loan controls	✓	✓	✓
Observations	260,440	260,013	259,822
Adjusted R^2	0.459	0.496	0.499
Dep. Var. Mean	10.89	10.88	10.88
Dep. Var. SD	19.36	19.35	19.35
Panel B: Autocorrelation of the Firm Steering Gap			
Dep. Variable:	Firm Steering Gap, Months $t - 12$ to $t - 1$ (bps)		
	(1)	(2)	(3)
Firm steering gap, months $t - 24$ to $t - 13$ (bps)	0.602*** (0.025)	0.622*** (0.019)	0.655*** (0.019)
County $\times$ Month FE	✓		
State $\times$ Month FE		✓	
Month FE			✓
Observations	64,232	74,601	74,964
Adjusted R^2	0.557	0.521	0.460
Dep. Var. Mean	13.29	13.37	13.33
Dep. Var. SD	16.41	16.46	16.44

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Past Firm Steering Gap and Mortgage Rates

This table regresses realized rate spreads on past firm steering gaps. The past gap averages the firm's loans in the previous 12 months, with at least five loans, and is divided by its standard deviation of 18.6 basis points. The steering gap is the chosen lender's predicted spread minus the expected predicted spread selected by a broker taking the lowest quote in the state lender set. Quotes equal predicted spreads plus independent normal noise with a standard deviation of 35 basis points. All columns include county $\times$ lock-month effects. Columns (3)–(5) add firm effects, and Column (5) adds lock-date effects. Columns (2), (4), and (5) control for standardized firm HHI, the sum of squared lender shares over the previous year. The dependent variable is in basis points. Controls are income, credit score, loan amount, property value, combined loan-to-value ratio, debt-to-income ratio, race, ethnicity, sex, occupancy, property type, number of units, loan purpose, and loan term. The sample consists of brokered conventional conforming fixed-rate loans in the matched Optimal Blue, CoreLogic, and HMDA data, 2018–2025, with usable lender-price predictions. Standard errors are clustered by county.

Dep. Variable:	Rate Spread (bps)				
	(1)	(2)	(3)	(4)	(5)
Past firm steering gap (std.)	9.984*** (0.460)	9.997*** (0.462)	7.056*** (0.422)	7.087*** (0.425)	6.772*** (0.397)
Broker firm HHI (std.)		−0.113 (0.167)		0.645*** (0.239)	0.520** (0.237)
County $\times$ Lock-Month FE	✓	✓	✓	✓	✓
Lock-Date FE					✓
Broker Firm FE			✓	✓	✓
Borrower and loan controls	✓	✓	✓	✓	✓
Observations	246,342	244,929	245,890	244,492	244,297
Adjusted R^2	0.435	0.436	0.486	0.486	0.518
Dep. Var. Mean	27.7	27.8	27.7	27.8	27.8
Dep. Var. SD	45.2	45.2	45.1	45.2	45.2

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: The Steering Gap and Broker Compensation

This table regresses lender-paid broker compensation on the loan's steering gap. Compensation is measured in basis points of the loan amount, and the gap in basis points of rate spread. The steering gap is the chosen lender's predicted spread minus the expected predicted spread selected by a broker taking the lowest quote in the state lender set. Quotes equal predicted spreads plus independent normal noise with a standard deviation of 35 basis points. Column (1) has no fixed effects. Column (2) includes county $\times$ lock-month and firm effects; Column (3) adds lock-date effects. Controls are income, credit score, loan amount, property value, combined loan-to-value ratio, debt-to-income ratio, race, ethnicity, sex, occupancy, property type, number of units, loan purpose, and loan term. The sample consists of brokered conventional conforming fixed-rate loans in the matched Optimal Blue, Core-Logic, and HMDA data, 2018–2025, with usable lender-price predictions. Only loans with recorded compensation are included. Standard errors are clustered by county.

Dep. Variable:	Broker Compensation (bps of Loan Amount)		
	(1)	(2)	(3)
Steering gap (bps)	0.282*** (0.012)	0.114*** (0.013)	0.114*** (0.013)
County $\times$ Lock-Month FE		✓	✓
Lock-Date FE			✓
Broker Firm FE		✓	✓
Borrower and loan controls	✓	✓	✓
Observations	197,611	178,648	178,451
Adjusted R^2	0.259	0.888	0.888
Dep. Var. Mean	207.2	205.4	205.4
Dep. Var. SD	51.1	51.0	51.0

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

For Online Publication

**Internet Appendix For:
“Do Mortgage Brokers Help Borrowers Get Better Rates?”**

A Sample Construction

The sample combines the CoreLogic, Optimal Blue, and HMDA data described in Section 1.1. Figure IA.2 summarizes the two matching steps.

Step 1: Optimal Blue and CoreLogic. Optimal Blue locks carry no borrower identity, so I match them to CoreLogic mortgages on loan characteristics that both sources record. A CoreLogic mortgage is a candidate for a lock if it has the same property ZIP code and the same loan amount and was signed between 10 days before and 240 days after the lock date. A candidate is rejected if the two records disagree on the loan purpose, the loan program, the loan term, or whether the rate is fixed or adjustable. The remaining candidates are scored on secondary features of the property and the loan, and a lock is matched only when its best candidate is clearly better than the next one. Later rounds use the county instead of the ZIP code or allow a small difference in the loan amount, for locks and mortgages that are still unmatched. When several locks match the same mortgage, as with relocks or a borrower who locked with more than one lender, I keep the best-scoring lock.

Step 2: adding HMDA. I then attach HMDA records to the matched lock-mortgage pairs. HMDA reports loan amounts in \$10,000 bins and no dates, so in the first round a candidate HMDA record must have the same activity year, census tract, and loan-amount bin as the pair, and the same note rate, occupancy, loan purpose, loan program, and term as the Optimal Blue lock. Candidates are scored on the lender and on other fields that both records report. Later rounds search the county instead of the tract, the neighboring year's HMDA file for loans closing in December or January, and, last, candidates whose note rate differs; each of these rounds also requires the same lender as the CoreLogic record. The sample keeps the most reliable matches: first-round matches whose best candidate has a high score and is the only candidate or clearly better than the next one, and county-round matches, which also require the same loan-to-value ratio. These are 14.1 million of the 15.8 million matched loans (89.4%). In a placebo that matches the 2022 locks to the 2021 HMDA file, about 1% of locks find a match, providing a diagnostic for accidental matches.

Brokered and retail loans. CoreLogic records up to two companies on each mortgage but does not record which is the lender and which is the broker, and the order of the two fields changes over time. I

assign the lender as the company whose name matches the lender of record on the deed, using NMLS registry names, and the broker as the other company, with a list of firms that fund loans to break ties. Brokered loans are then classified as in Section 1.1. Each brokered loan is assigned to the branch where its loan officer worked in the lock month, from the NMLS employment and branch records.

Sample filters. The main sample includes conventional conforming fixed-rate loans closing from 2018 to 2025, with complete borrower and loan controls, and excludes singleton county $\times$ lock-date cells. Rate spread, income, combined loan-to-value ratio, and origination charges are winsorized at the 1st and 99th percentiles within closing year. The steering analysis retains 308,810 brokered loans with usable lender-price predictions. The IV analysis uses owner-occupied brokered loans locked from 2019 to 2025, with a located search origin and a previous-year firm price level. Origins are previous addresses for purchases and mortgaged properties for refinances. Additional restrictions appear in the exhibit notes.

B Lender Pricing and the Honest-Broker Benchmark

The 100 lenders with the most brokered loans have separate pricing models, estimated on their retail and brokered loans; the remaining lenders share a pooled model. The chosen lender's predicted price leaves out the focal loan. Predictions require sufficient observations in the relevant state-year and lock-month cells. Figure IA.10 reports model fit, and Figure IA.11 compares estimated prices of borrower characteristics.

To predict prices for lenders outside a broker's observed network, I decompose the lender $\times$ broker effect into a broker component, a lender component, and a relationship-specific component. All lender predictions include the broker component. A relationship-specific component is included only for lenders on the broker's observed menu. The state set combines lenders used by at least 10 brokers in the state during the previous 12 months with the focal broker's own lenders there. It is a market benchmark, rather than a record of offers made to the borrower.

The honest-broker calculation is a robustness exercise for pricing information observed by lenders but absent from the research data. It assumes independent normal pricing errors with a common standard deviation of 35 basis points, calibrated to the 35.1 basis point out-of-sample root mean squared

prediction error on brokered loans. This assigns the full prediction-error scale to lender-specific variation, although part may reflect risk priced alike by lenders and therefore leave their ranking unchanged. I compute the probability that each lender offers the lowest quote by numerical integration over this distribution. The distribution in Figure 7, Panel A, weights each possible lender's predicted-price gap by this probability; it is not the distribution of loan-level expected gaps. The loan-level expected gap is subtracted from the observed gap to construct the steering-gap proxy in Panel B. Table IA.10 varies the pricing-error assumption and the lender set.

Figure IA.1: A Broker Firm's Website

This figure shows the home page of one broker firm's website, an example of how mortgage brokers describe their service to borrowers. The broker advertises that it compares rates across many wholesale lenders on the borrower's behalf. Figure 1, Panel A, summarizes the phrases used across all broker websites.

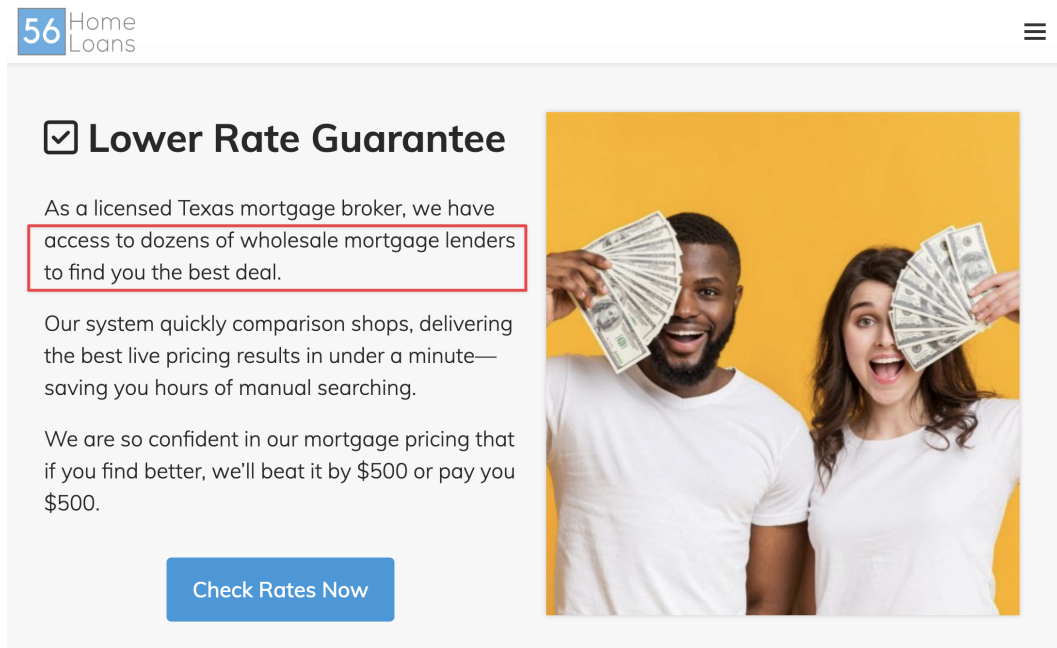
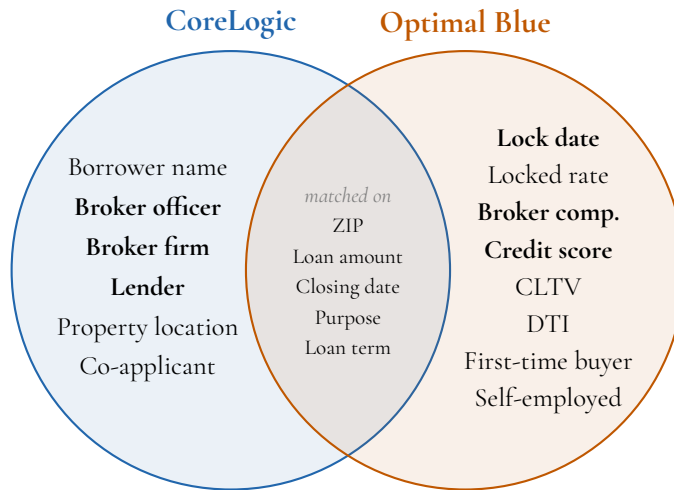


Figure IA.2: Linking Borrowers, Brokers, and Lenders

This figure shows the two steps that link the three main data sources, described in Section A. Panel A matches Optimal Blue rate locks to CoreLogic mortgages. Panel B adds HMDA records to the matched lock-mortgage pairs. In each panel, the circles list what each source records, the center lists the matching variables, and the variables in bold are the ones the analysis requires from each source.

Panel A: Matching Optimal Blue Rate Locks to CoreLogic Mortgages



Panel B: Adding HMDA Records to the Matched Lock-Mortgage Pairs

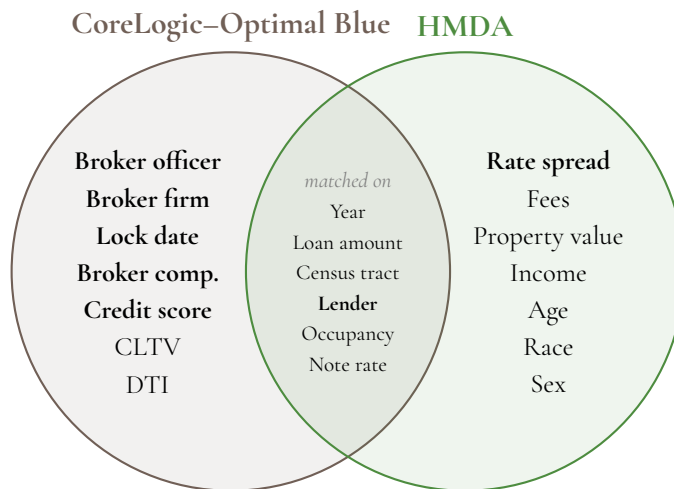


Figure IA.3: Residual Rate Dispersion with Census Tract $\times$ Lock-Date Fixed Effects

This figure compares residual rate dispersion for retail and brokered loans using census tract $\times$ lock-date fixed effects. Residuals come from a pooled regression of rate spread on a brokered indicator, income, credit score, loan amount, property value, combined loan-to-value ratio, debt-to-income ratio, origination charges, race, and sex. Gray bars represent retail loans and the red line brokered loans. Rate spread is APR minus APOR, in basis points; the legend reports residual standard deviations. Singleton tract-date cells are dropped, and the displayed range matches the county-date comparison in Figure 2. The data are matched Optimal Blue, CoreLogic, and HMDA loans, 2018–2025.

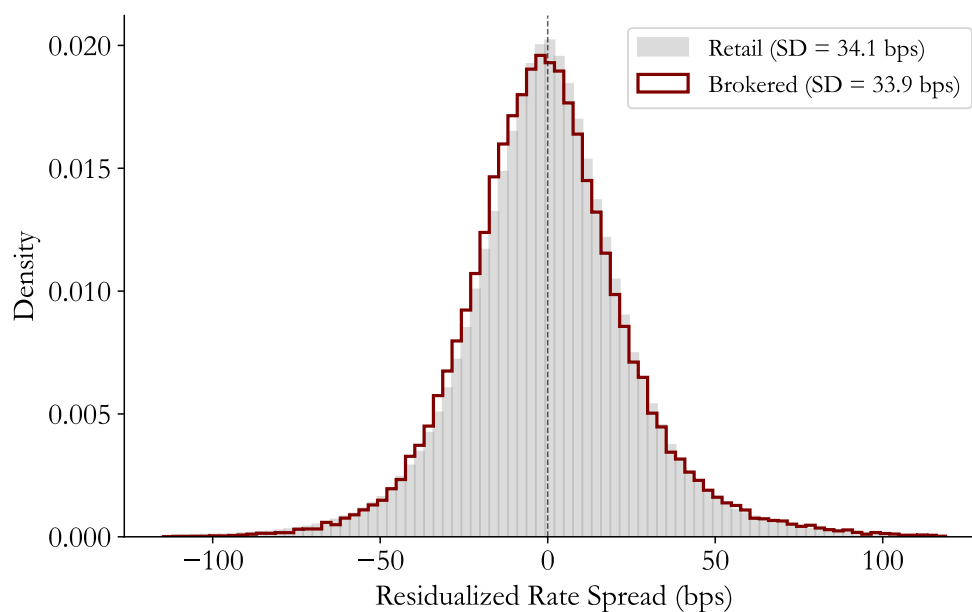


Figure IA.4: Rate Spreads of Retail and Brokered Loans by Credit Score

This figure compares the average rate spread of retail and brokered loans across borrower credit scores. I regress the rate spread on indicators for each channel and credit-score bin, with no other controls, so each point is the mean rate spread of the channel in the bin; the bars are 95% confidence intervals. Gray circles are retail loans and red squares brokered loans. The rate spread is in basis points. The sample is the matched Optimal Blue, CoreLogic, and HMDA sample of rate locks from 2018 to 2025. Standard errors are clustered by state.

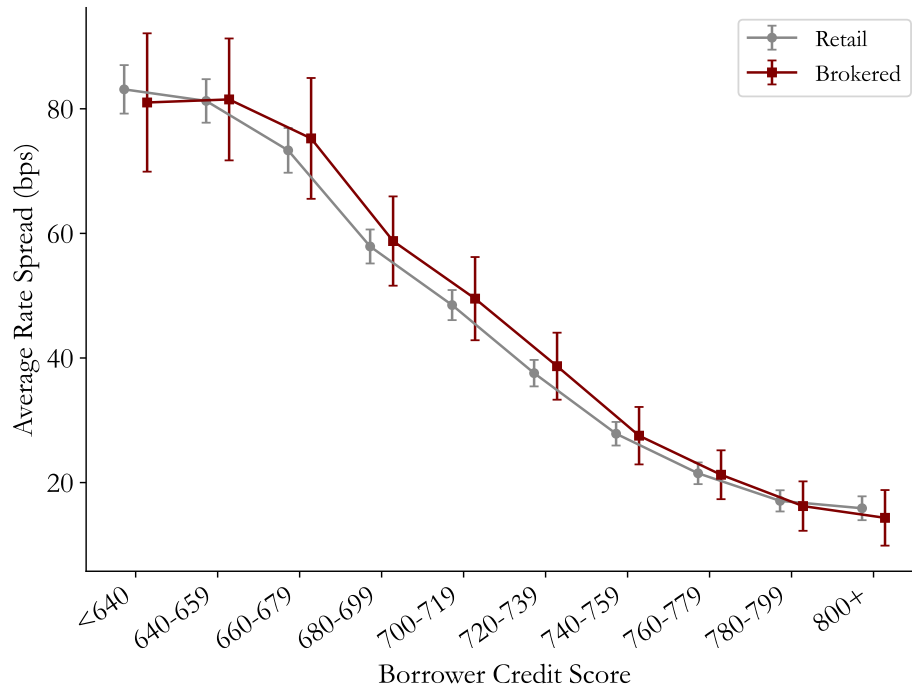


Figure IA.5: Broker Firms with High Price Levels and Brokered Rates Across Counties

This figure relates the share of a county's brokered loans made by broker firms with high price levels to the dispersion and the level of brokered rates in the county. Each dot is a county with at least 100 brokered loans, and its size is proportional to the county's brokered loans. The horizontal axis is the share of the county's brokered loans made by broker firms in the top third of price levels across brokered loans. A firm's price level is its mean brokered-minus-retail rate gap, defined below, over its brokered loans in other counties, shrunk toward the overall mean by empirical Bayes, so none of the county's own loans enters the share. Panel A plots the standard deviation of the residual rate spread of the county's brokered loans, from a regression of the rate spread on borrower and loan controls with county $\times$ lock-month fixed effects, corrected for leverage; the line is a loan-weighted quadratic fit. Panel B plots the brokered-minus-retail rate gap, the mean over the county's brokered loans of the rate spread adjusted for borrower and loan controls minus the mean adjusted rate spread of retail loans in the same county and lock month; the line is a loan-weighted linear fit, and the vertical axis is cut at -20 and 30 basis points, and the fit uses all counties. The sample is the matched Optimal Blue, CoreLogic, and HMDA sample of rate locks from 2018 to 2025, pooled over years.

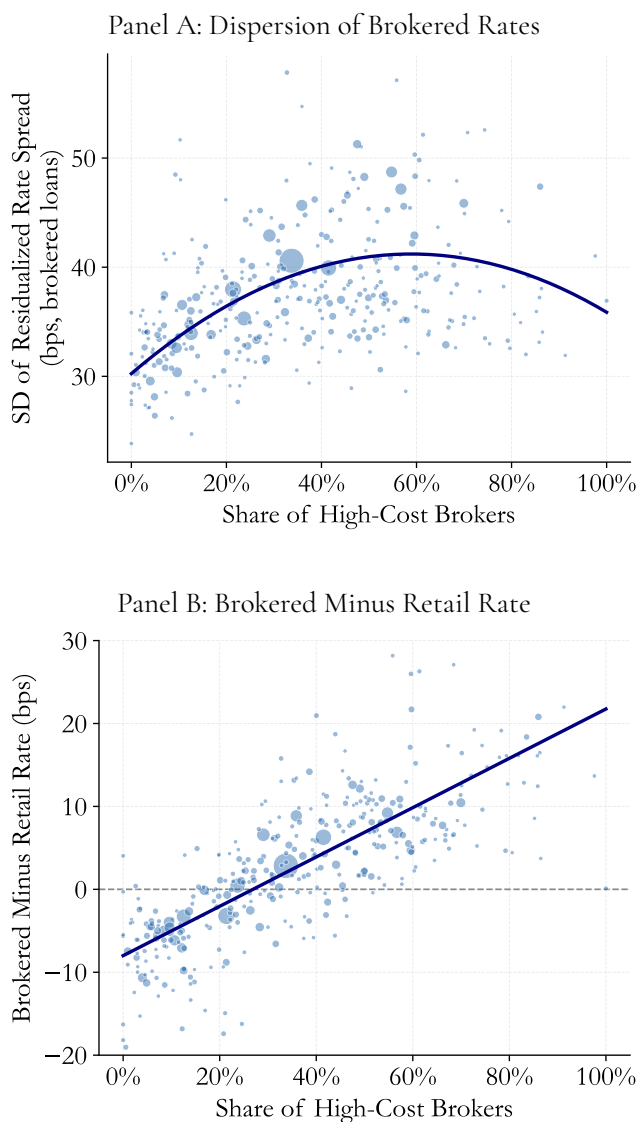


Figure IA.6: Borrower and Neighborhood Characteristics by the Sign of the Instrument

This figure compares borrower and neighborhood distributions by relative proximity to high-cost and low-cost branches. Differential distance dd is distance to the nearest low-cost branch minus distance to the nearest high-cost branch; high-cost firms are classified using previous-year price effects. Blue represents $dd < 0$ and red $dd \geq 0$. Each panel displays one characteristic without controls or fixed effects, and dashed lines mark group means. Upper panels report borrower characteristics; lower panels report origin block-group and tract characteristics from the American Community Survey. The origin is the previous address for purchases and the mortgaged property for refinances. Displayed ranges cover the pooled 2nd–98th percentiles. The sample is owner-occupied brokered loans with traced origins in the matched Optimal Blue, CoreLogic, and HMDA IV data, 2019–2025.

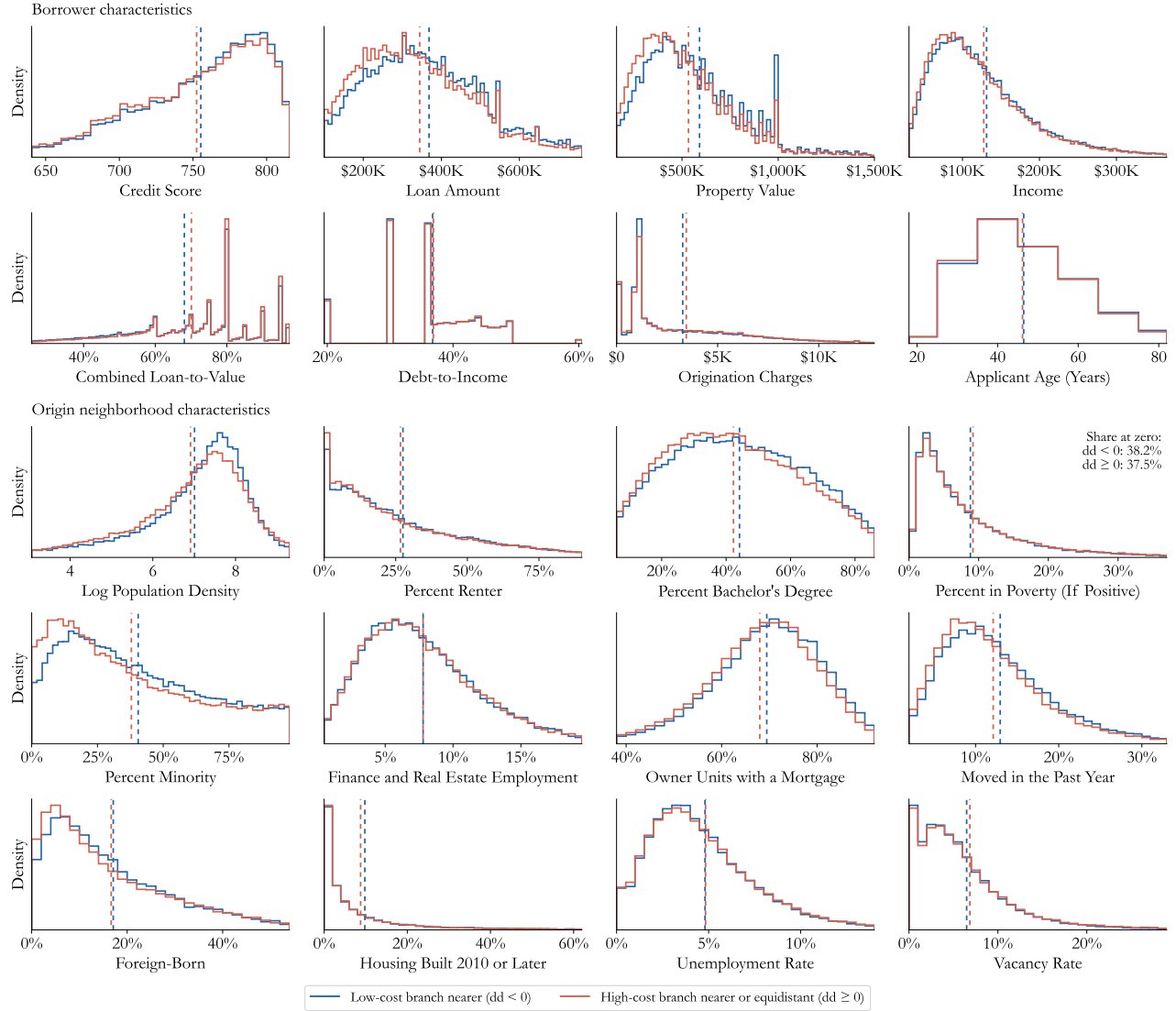


Figure IA.7: The Differential Distance Instrument: An Example

This figure illustrates differential distance for a borrower in Travis County, Texas, who locked a mortgage rate in June 2022. Triangles mark potential broker branches. Red branches belong to high-cost firms, classified by whether their previous-year price effect is at or above the monthly median across branches; blue branches belong to low-cost firms. The nearest low-cost branch is 1 mile from the origin and the nearest high-cost branch is 4 miles away, giving a differential distance of $1 - 4 = -3$ miles. The origin is the borrower's previous address, traced using Verisk and DataAxle. The map also marks the mortgaged property, and tract shading shows population density.

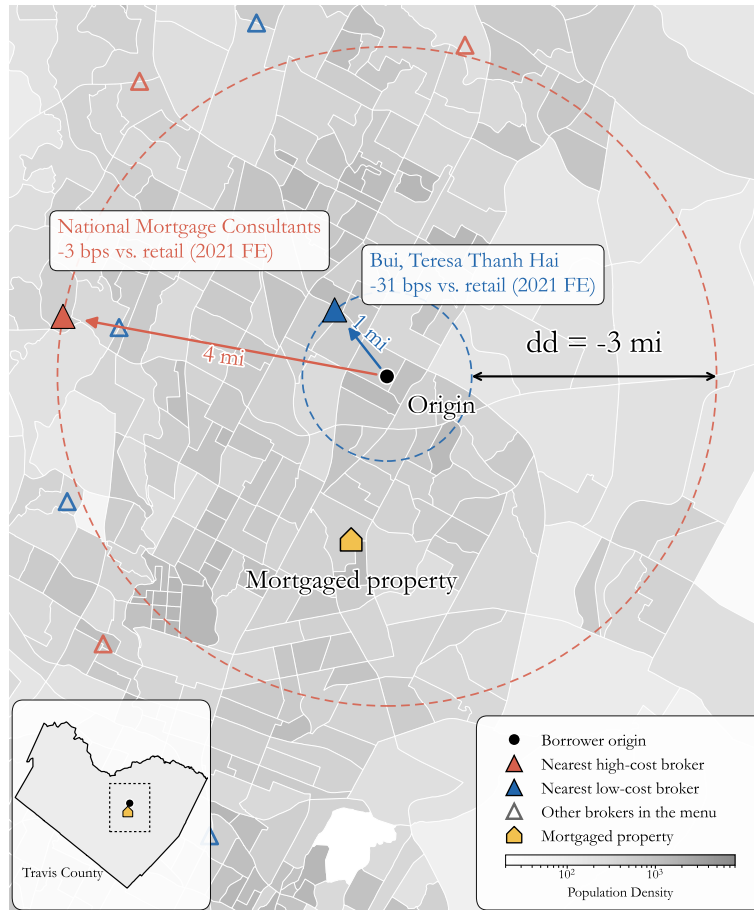
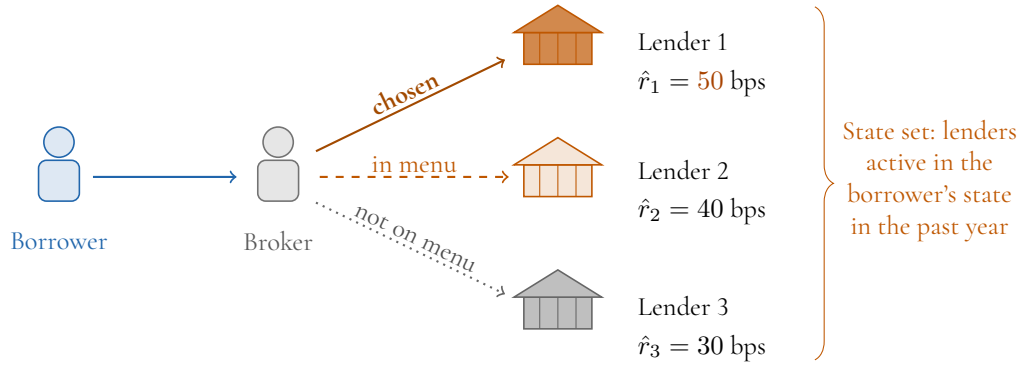


Figure IA.8: Measuring the Steering Gap

This figure illustrates the observed, honest-broker, and steering gaps for a hypothetical loan. The chosen lender is Lender 1. Lenders 1 and 2 are on the broker's menu; Lender 3 is in the state set but not on the menu and receives no relationship-specific pricing adjustment. All three prices are predicted rate spreads, including the chosen lender's. The observed gap is the chosen price minus the lowest predicted price, $50 - 30 = 20$ basis points. The honest broker obtains a quote from each lender, equal to its predicted price plus independent normal noise with a standard deviation of 35 basis points, and selects the lowest. The table shows three hypothetical draws of the noise; in Draw 2, Lender 1 quotes lowest although its predicted price is the highest. Over all draws, the three lenders quote lowest with probabilities of approximately 0.22, 0.32, and 0.46, so the expected predicted price selected by the honest broker is 37.6 basis points and the honest-broker gap is $37.6 - 30 = 7.6$. The steering gap is $20 - 7.6 = 12.4$, equivalently $50 - 37.6$. Using the mean predicted price of 40 as the reference instead gives an observed gap of 10 and an honest-broker gap of -2.4 , leaving the steering gap unchanged. All values are illustrative.



	Lender 1	Lender 2	Lender 3	Lowest quote
Predicted spread $\hat{r}_\ell$	50	40	30	Lender 3
<i>Quotes $\hat{r}_\ell + \sigma v_\ell$, three draws of the noise ($\sigma = 35$ bps)</i>				
Draw 1	$50 - 7.0 = 43.0$	$40 + 14.0 = 54.0$	$30 + 10.5 = \mathbf{40.5}$	Lender 3
Draw 2	$50 - 31.5 = \mathbf{18.5}$	$40 + 3.5 = 43.5$	$30 + 17.5 = 47.5$	Lender 1
Draw 3	$50 + 21.0 = 71.0$	$40 - 24.5 = \mathbf{15.5}$	$30 + 7.0 = 37.0$	Lender 2
Probability of the lowest quote p_ℓ	0.22	0.32	0.46	

Honest broker's expected predicted price = $0.22 \times 50 + 0.32 \times 40 + 0.46 \times 30 = 37.6$ bps

Observed gap = $50 - 30 = 20$ bps Honest-broker gap = $37.6 - 30 = 7.6$ bps

Steering gap = $20 - 7.6 = 50 - 37.6 = 12.4$ bps

Figure IA.9: Lender Choice Within the Menu and Menu Choice

This figure splits the observed and steering gaps of a hypothetical loan into lender choice within the broker's menu and menu choice. Predicted rate spreads are on the horizontal axis, in basis points. The chosen lender, Lender 1, and Lender 2 are on the broker's menu, with predicted spreads of 50 and 40 basis points; Lender 3, at 30 basis points, is in the state set but not on the menu. The loan is that of Figure IA.8. The observed gap of 20 is lender choice of $50 - 40 = 10$, the chosen price minus the lowest menu price, plus menu choice of $40 - 30 = 10$, the lowest menu price minus the lowest state-set price. The honest brokers obtain quotes equal to predicted prices plus independent normal noise with a standard deviation of 35 basis points and select the lowest. An honest broker that shops the state set selects the three lenders with probabilities of approximately 0.22, 0.32, and 0.46, an expected predicted price of 37.6; one that shops only the menu selects Lender 1 or Lender 2 with probabilities of approximately 0.42 and 0.58, an expected predicted price of 44.2. The steering gap of $50 - 37.6 = 12.4$ is lender choice of $50 - 44.2 = 5.8$ plus menu choice of $44.2 - 37.6 = 6.6$. All values are illustrative.

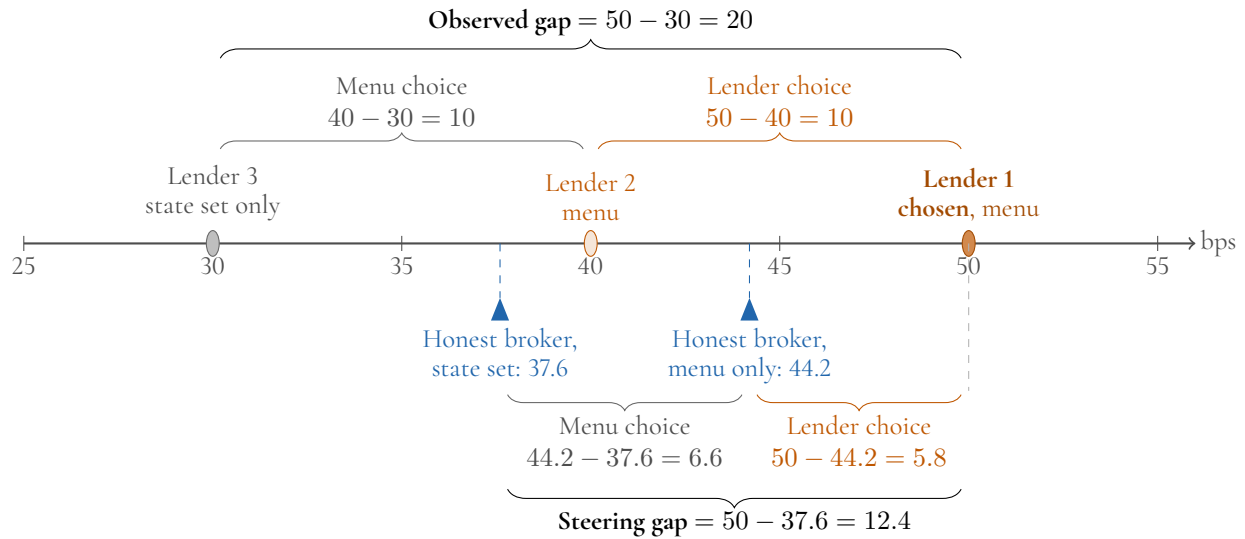


Figure IA.10: Fit of the Lender Pricing Models

This figure relates the fit of the lender pricing models of equation (6) to the number of loans each model is estimated on. Each of the 100 lenders with the most brokered loans has its own pricing model, estimated on its retail and brokered loans (dots); the remaining lenders share one pooled model (diamond). The vertical axis is the R^2 of the model over all of its loans, and the line is the median R^2 within quintiles of the number of loans across the 100 lender models. The horizontal axis is on a log scale. The sample is retail and brokered loans in the matched Optimal Blue, CoreLogic, and HMDA sample of rate locks from 2018 to 2025.

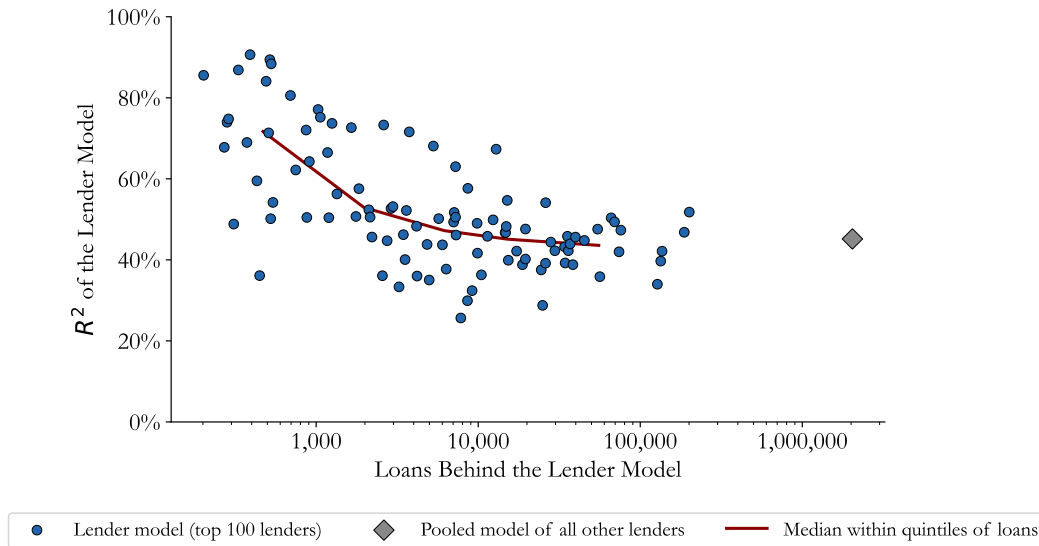


Figure IA.11: Borrower Characteristics in the Lender Pricing Models

This figure reports the coefficients on borrower and loan characteristics in the lender pricing models of equation (6), for the 30 lenders with the most brokered loans. Each panel shows one characteristic: each point is one lender's coefficient with its 95% confidence interval, and lenders are sorted by the coefficient. Coefficients on credit score, combined loan-to-value ratio, debt-to-income ratio, income, and loan amount are the change in the rate spread, in basis points, for a one-standard-deviation increase in the characteristic, with standard deviations taken over brokered loans with a measured steering gap; the loan amount coefficient holds the combined loan-to-value ratio fixed. Coefficients on investment property, second home, and two units are relative to a principal residence with one unit. The gray dashed line is the coefficient in the pooled model of the lenders outside the top 100, and the red dashed line marks zero. The sample is retail and brokered loans in the matched Optimal Blue, CoreLogic, and HMDA sample of rate locks from 2018 to 2025. Standard errors are clustered by county.

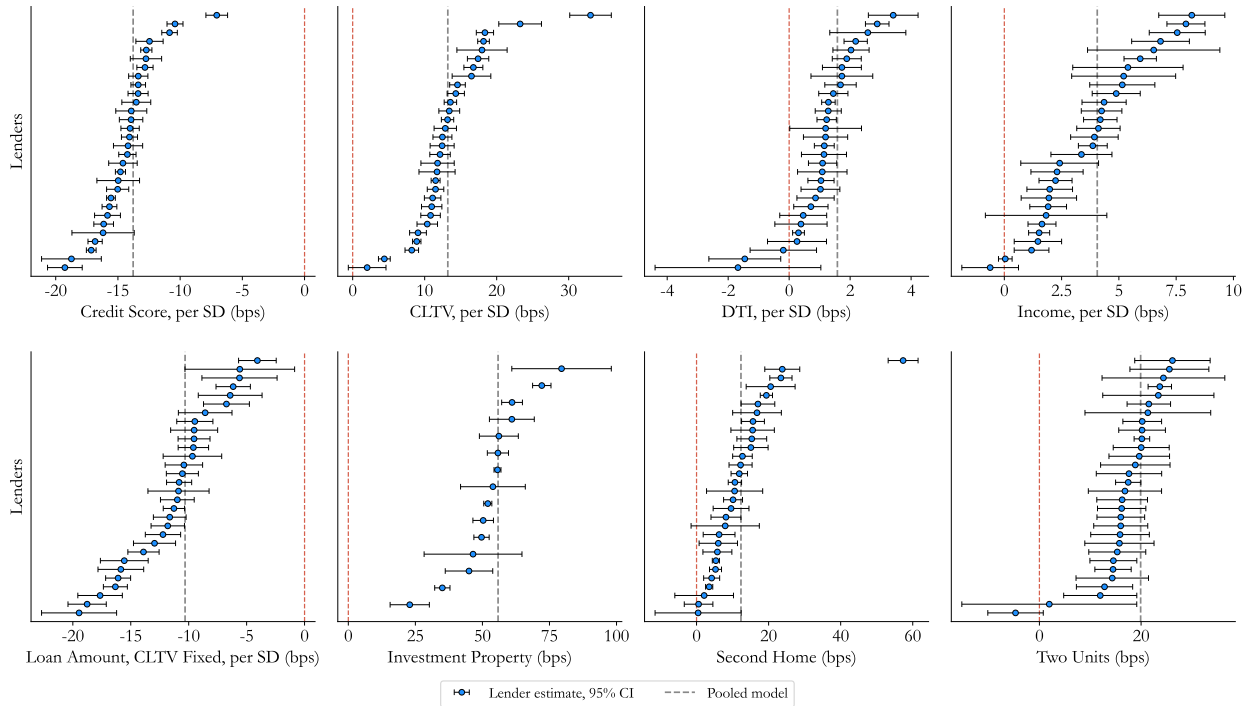


Figure IA.12: Observed and Honest-Broker Gaps of the Largest Broker Firms

This figure reports the mean observed gaps (gray) and honest-broker gaps (blue) of the 25 largest broker firms in the steering analysis, in the order of Figure 8, with 95% confidence intervals. The observed gap compares the chosen lender's predicted spread with the lowest predicted spread in the state set. The honest-broker gap is the expected observed gap when the broker selects the lowest of the state lenders' noisy quotes, with independent normal noise of 35 basis points. The steering gap in Figure 8 is their difference. The dashed and dotted lines report the observed and honest-broker means across all loans. Gaps are in basis points, and confidence intervals use within-firm variation. The sample consists of brokered conventional conforming fixed-rate loans in the matched Optimal Blue, CoreLogic, and HMDA data, 2018–2025, with usable lender-price predictions.

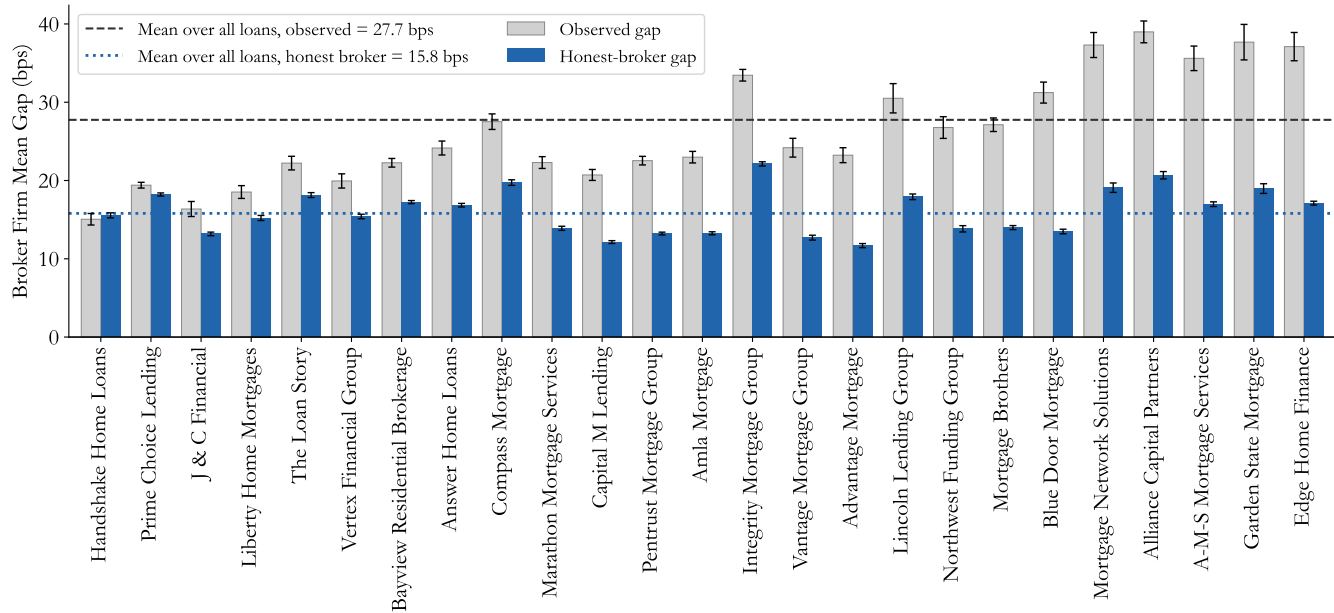


Table IA.1: Summary Statistics of Retail Loans

This table summarizes retail loans and compares their means with brokered loans. Retail loans are matched loans not classified as brokered. Panel A reports prices, Panel B borrower characteristics, and Panel C loan characteristics. Rate spread is APR minus APOR, in basis points. Minority identifies non-White applicants among those reporting race, with joint applications coded as non-minority. The final columns report the brokered mean and the brokered-minus-retail difference. The sample consists of conventional conforming fixed-rate loans in the matched Optimal Blue, CoreLogic, and HMDA data closing in 2018–2025, with complete borrower and loan controls. Rate spread, income, combined loan-to-value ratio, and origination charges are winsorized at the 1st and 99th percentiles by closing year. Statistics use the available observations for each variable.

	Retail Loans					Brokered	Brokered
	Mean	SD	P10	Median	P90	Mean	– Retail
<i>Panel A: Prices</i>							
Rate spread (bps)	31.2	47.9	−18.8	26.8	90.3	31.3	0.0
Note rate (%)	4.19	1.51	2.75	3.62	6.62	3.99	−0.20
Origination charges (\$)	2,449	2,386	684	1,579	5,449	3,381	932
<i>Panel B: Borrower Characteristics</i>							
Credit score	753	44	689	762	803	752	−1
Income (\$000s)	124.5	81.5	50.0	104.0	221.0	130.4	5.9
Debt-to-income ratio (%)	36.9	7.7	30.0	36.0	47.0	37.4	0.4
Combined loan-to-value ratio (%)	74.9	17.8	48.8	80.0	95.0	73.0	−1.9
Minority applicant (1/0)	0.130	0.337	0.000	0.000	1.000	0.175	0.045
Female applicant (1/0)	0.242	0.429	0.000	0.000	1.000	0.236	−0.007
Has co-applicant (1/0)	0.468	0.499	0.000	0.000	1.000	0.444	−0.025
<i>Panel C: Loan Characteristics</i>							
Loan amount (\$000s)	307.7	152.8	135.8	280.8	510.4	352.7	45.0
Property value (\$000s)	436.3	810.0	195.0	385.0	735.0	523.1	86.8
Purchase loan (1/0)	0.586	0.493	0.000	1.000	1.000	0.485	−0.100
Cash-out refinance (1/0)	0.143	0.350	0.000	0.000	1.000	0.193	0.050
30-year term (1/0)	0.849	0.358	0.000	1.000	1.000	0.857	0.008
Principal residence (1/0)	0.945	0.227	1.000	1.000	1.000	0.954	0.009
Retail loans	7,814,055						
Lenders	2,692						
Retail loan officers	197,250						
Counties	2,215						

Table IA.2: Matched Brokered Loans Compared with All Brokered Loans in HMDA

This table compares matched brokered loans with comparable brokered loans in HMDA. Panels A–C report prices, loan characteristics, and borrower characteristics; Panel D reports closing-year shares. The reweighted mean applies the matched sample’s year weights to HMDA means. The normalized difference divides the difference in means by the square root of the average of the two variances. Both groups use HMDA variable definitions. The HMDA benchmark consists of originated first-lien conventional conforming fixed-rate loans whose applications were not submitted directly to the reporting lender, closing in 2018–2025. Rate spread, note rate, property value, income, combined loan-to-value ratio, and origination charges are winsorized using the HMDA benchmark’s annual 1st and 99th percentiles. Debt-to-income is the upper endpoint of the reported range. Each variable uses its available observations. The matched data link Optimal Blue, CoreLogic, and HMDA.

	Matched Brokered		HMDA Brokered				
	Mean	SD	Mean	SD	Reweightd Mean	Diff. in Means	Normalized Diff.
Panel A: Prices							
Rate spread (bps)	32.4	53.7	32.2	65.2	29.3	0.2	0.004
Note rate (%)	3.99	1.53	4.52	1.84	4.07	−0.53	−0.311
Origination charges (\$)	3,446	3,526	3,228	3,404	3,088	218	0.063
Panel B: Loan Characteristics							
Loan amount (\$000s)	353.4	164.6	331.8	172.2	330.5	21.6	0.128
Property value (\$000s)	517.1	278.3	493.8	285.0	499.9	23.4	0.083
Income (\$000s)	131.1	91.0	134.4	99.0	132.5	−3.3	−0.035
Debt-to-income ratio (%)	37.4	7.7	37.3	8.0	37.0	0.1	0.008
Combined loan-to-value ratio (%)	73.0	17.9	71.9	17.8	71.0	1.0	0.059
Purchase loan (1/0)	0.485	0.500	0.501	0.500	0.448	−0.016	−0.031
Cash-out refinance (1/0)	0.193	0.395	0.185	0.389	0.198	0.008	0.021
30-year term (1/0)	0.857	0.350	0.824	0.381	0.807	0.033	0.090
Principal residence (1/0)	0.954	0.210	0.841	0.365	0.851	0.113	0.378
Panel C: Borrower Characteristics							
Minority applicant (1/0)	0.175	0.380	0.218	0.413	0.218	−0.042	−0.106
Female applicant (1/0)	0.236	0.424	0.232	0.422	0.230	0.004	0.009
Has co-applicant (1/0)	0.444	0.497	0.428	0.495	0.432	0.016	0.032
Panel D: Share of Loans by Closing Year (%)							
2018	4.3		6.8			−2.5	
2019	8.2		10.5			−2.3	
2020	26.3		22.0			4.3	
2021	33.3		24.5			8.8	
2022	12.3		10.3			2.0	
2023	4.0		7.0			−3.0	
2024	5.4		8.9			−3.6	
2025	6.3		10.0			−3.8	
Loans	335,997		7,049,923				
Matched / HMDA brokered loans (%)	4.8						

Table IA.3: Rate Levels of Brokered and Retail Loans

This table compares the rates of brokered and retail loans made to observably similar borrowers. Each column regresses the rate spread on an indicator for brokered loans and borrower and loan controls, with location-time fixed effects that become finer across columns: state $\times$ lock month in Column (1), county $\times$ lock month in Column (2), county $\times$ lock month plus lock date in Column (3), county $\times$ lock date in Column (4), and county $\times$ lock date plus census tract in Column (5). Borrower and loan controls are income, credit score, loan amount, property value, combined loan-to-value ratio, debt-to-income ratio, origination charges, race, and sex. Loans alone in a fixed-effect cell are dropped. The rate spread is in basis points. The sample is the matched Optimal Blue, CoreLogic, and HMDA sample of rate locks from 2018 to 2025. Standard errors are clustered by state.

Dep. Variable:	Rate Spread (bps)				
	(1)	(2)	(3)	(4)	(5)
1(Brokered)	1.24 (0.91)	0.85 (0.88)	0.95 (0.88)	0.81 (0.91)	0.41 (0.88)
State $\times$ Lock-Month FE	✓				
County $\times$ Lock-Month FE		✓	✓		
Lock-Date FE			✓		
County $\times$ Lock-Date FE				✓	✓
Census Tract FE					✓
Borrower and loan controls	✓	✓	✓	✓	✓
Observations	8,150,052	8,150,052	8,150,052	8,150,052	8,145,021
Adjusted R^2	0.297	0.317	0.344	0.363	0.389
Dep. Var. Mean	31.2	31.2	31.2	31.2	31.2
Dep. Var. SD	47.9	47.9	47.9	47.9	47.9

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table IA.4: Correlation of Broker Firm Price Levels Across Periods and Loan Types

This table relates a broker firm's price level on one side of a sample split to its price level on the other side. For each split, the broker firm effects of Figure 3 are estimated separately on side A and side B, each relative to retail loans on the same side in the same county on the same day. Each column is one split: lock years 2018–2021 against 2022–2025, purchase against refinance loans, rate-and-term against cash-out refinances, principal residences against second homes and investment properties, 30-year against 10- to 25-year terms, credit scores of 740 or above against below 740, combined loan-to-value ratios of 80% or below against above 80%, and conventional against government-backed loans. The raw correlation is between the two sets of estimated effects. The correlation and slope net of noise use only covariances between effects estimated on loans in different lock months, to adjust for estimation noise. The persistent standard deviation is the standard deviation of firm effects net of estimation error. Panel A uses one effect per broker firm. Panel B uses one effect per broker firm and lock year, demeaned by lock year, and compares the two sides within the same firm and year; the lock-year split has no Panel B version. Firms need at least 10 brokered loans on each side. The sample is the matched Optimal Blue, CoreLogic, and HMDA sample of rate locks from 2018 to 2025. Standard errors are from a bootstrap over broker firms.

Split	Lock year	Purpose	Refi type	Occupancy	Term	Score	CLTV	Loan type
Side A	2018–21	Purchase	Rate-term	Primary	30-year	740+	≤80%	Conv.
Side B	2022–25	Refinance	Cash-out	Other	Shorter	<740	>80%	Govt.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A: Broker Firm Effects</i>								
Correlation, raw	0.56	0.62	0.60	0.48	0.65	0.77	0.74	0.58
Correlation, net of noise	0.65	0.75	0.77	0.54	0.85	0.95	0.90	0.69
	(0.03)	(0.02)	(0.08)	(0.15)	(0.02)	(0.01)	(0.02)	(0.02)
Slope of B on A, net of noise	1.08	0.55	0.89	0.88	0.86	1.06	0.99	1.10
	(0.06)	(0.02)	(0.04)	(0.16)	(0.04)	(0.03)	(0.03)	(0.05)
Mean effect, B – A (bps)	4.73	–1.55	0.26	–2.47	0.11	–0.70	0.88	–7.28
	(0.55)	(0.33)	(0.36)	(1.59)	(0.45)	(0.21)	(0.27)	(0.61)
Persistent SD, side A (bps)	11.11	13.76	10.02	12.89	11.85	11.89	11.92	13.88
Persistent SD, side B (bps)	18.46	10.13	11.69	21.07	12.06	13.22	13.23	22.16
Broker firms	1,370	2,033	1,254	179	987	2,277	1,833	2,039
Brokered loans, side A	150,066	114,636	66,093	60,963	147,290	169,517	156,370	208,287
Brokered loans, side B	54,000	120,081	37,098	3,569	29,348	79,899	69,608	84,941
<i>Panel B: Broker Firm-Year Effects, Demeaned by Lock Year</i>								
Correlation, raw		0.63	0.64	0.35	0.67	0.74	0.70	0.54
Correlation, net of noise		0.81	0.84	0.40	0.85	0.97	0.90	0.69
		(0.03)	(0.03)	(0.17)	(0.03)	(0.02)	(0.02)	(0.04)
Slope of B on A, net of noise		0.66	0.89	1.03	0.90	1.09	0.96	1.00
		(0.03)	(0.05)	(0.20)	(0.04)	(0.03)	(0.04)	(0.07)
Mean effect, B – A (bps)		–0.17	2.02	1.38	0.28	0.17	0.34	–9.09
		(0.36)	(0.40)	(2.89)	(0.42)	(0.31)	(0.34)	(0.82)
Persistent SD, side A (bps)		12.77	9.87	9.29	11.12	11.98	12.74	15.22
Persistent SD, side B (bps)		10.32	10.34	24.11	11.75	13.42	13.51	22.08
Broker firm-years		1,665	942	94	877	2,127	1,808	1,832
Broker firms		962	643	67	608	1,121	892	839
Brokered loans, side A		50,230	40,289	17,550	66,915	100,445	88,211	96,017
Brokered loans, side B		72,222	21,406	1,560	20,563	46,524	40,999	45,004

Standard errors in parentheses.

Table IA.5: Distance and the Choice of Broker Branch

This table relates the probability that a borrower uses a broker branch to the distance between the borrower and the branch. Each observation is a loan paired with one broker branch in the loan's choice set, and the dependent variable equals one if the borrower used that branch. Distance is measured from the borrower's origin to the branch, and loan fixed effects absorb each loan's choice set. Columns (1)–(4) are linear probability models weighted by inverse sampling probabilities, with the dependent variable multiplied by 100 so that coefficients are in percentage points; Columns (1) and (2) measure distance in units of 10 miles, capped at 50 miles, and Columns (3) and (4) use the log of one plus the distance in miles. Because of the weights, the dependent-variable mean in Columns (1)–(4) is the probability that a branch in the full choice set is chosen, one over the number of branches per choice set. Columns (5) and (6) are conditional logit models with a sampling correction; Column (6) adds broker branch constants on a smaller sample of alternatives, and their dependent-variable means are the shares of chosen branches among the sampled alternatives. The choice set, the stratified sample of alternatives, and the origin are as in Figure 5. The sample is loans in the matched Optimal Blue, CoreLogic, and HMDA sample with lock dates from 2018 to 2025. Standard errors are clustered by origin county.

Dep. Variable:	$\mathbb{1}(\text{Branch Is Chosen})$					
Model:	Linear Probability (ppt)				Conditional Logit	
	(1)	(2)	(3)	(4)	(5)	(6)
Distance (10 miles)	−0.102*** (0.024)	−0.109*** (0.021)				
Log(1 + distance in miles)			−0.066*** (0.011)	−0.089*** (0.011)	−1.06*** (0.021)	−1.08*** (0.019)
Loan FE	✓	✓	✓	✓	✓	✓
Branch × Lock-Month FE		✓		✓		
Branch FE						✓
Loans (choice sets)	218,192	218,192	218,192	218,192	218,192	218,180
Branches per choice set	1,681	1,681	1,681	1,681	1,681	1,681
Observations	20,452,623	20,373,529	20,452,623	20,373,529	20,452,623	5,507,731
Dep. Var. Mean (ppt)	0.06	0.06	0.06	0.06	1.07	3.96
Dep. Var. SD (ppt)	2.44	2.44	2.44	2.44	10.3	19.5

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table IA.6: High-Cost Brokers and Mortgage Rates: First Stage

This table reports first-stage regressions of high-cost broker use on differential distance. A firm is high-cost if its shrunk previous-year price effect is at or above the monthly median across branches. Differential distance is distance to the nearest low-cost branch minus distance to the nearest high-cost branch, in units of 10 miles. The dependent variable is a zero-one indicator. Column (1) includes origin county $\times$ lock-month effects; Column (2) adds mortgage county $\times$ lock-month effects; Columns (3)–(5) add lock-date effects, lender $\times$ lock-month effects, or both. All columns include credit-score, combined loan-to-value, and debt-to-income deciles, nearest-branch distance, and origin-neighborhood controls. The origin is the borrower’s previous address for purchases, traced using Verisk and DataAxle, and the mortgaged property for refinances. The sample is owner-occupied brokered loans with traced origins and lagged firm classifications in the matched Optimal Blue, CoreLogic, and HMDA data, 2019–2025, excluding the top 5% of absolute differential distance. Standard errors are clustered by mortgage county.

Dep. Variable:	$\mathbb{1}(\text{High-Cost Broker Firm})$				
	(1)	(2)	(3)	(4)	(5)
Differential distance (10 miles)	0.0546*** (0.0056)	0.0556*** (0.0060)	0.0560*** (0.0061)	0.0520*** (0.0052)	0.0526*** (0.0052)
Origin County $\times$ Lock-Month FE	✓	✓	✓	✓	✓
Mortgage County $\times$ Lock-Month FE		✓	✓	✓	✓
Lock-Date FE			✓		✓
Lender $\times$ Lock-Month FE				✓	✓
Borrower controls	✓	✓	✓	✓	✓
Origin controls	✓	✓	✓	✓	✓
First-stage F -statistic	93.7	85.9	84.8	100.4	101.1
Observations	143,970	137,675	137,505	132,245	132,065
Dep. Var. Mean	0.427	0.426	0.426	0.424	0.424
Dep. Var. SD	0.495	0.494	0.494	0.494	0.494

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table IA.7: High-Cost Brokers and Loan Timing: Differential Distance IV

This table estimates the relationship between high-cost broker use and loan timing by 2SLS. High-cost firms have a previous-year price effect at or above the monthly median across branches. The high-cost indicator, marked with a hat, is instrumented by distance to the nearest low-cost branch minus distance to the nearest high-cost branch. Origins are previous addresses for purchases and mortgaged properties for refinances. Panel A uses days from the Optimal Blue lock date to the CoreLogic mortgage date, winsorized at the 1st and 99th percentiles (6 and 102 days). Columns add origin county $\times$ month, mortgage county $\times$ month, lock-date, and lender $\times$ month effects as indicated. Panel B uses the two county-month effects and four outcomes: lock-to-close days, requested lock-period days, an indicator for a requested period above 30 days, and slack (requested period minus lock-to-close days). All columns include credit-score, loan-to-value, and debt-to-income deciles, nearest-branch distance, and origin-neighborhood controls. The sample is owner-occupied brokered loans in the matched Optimal Blue, CoreLogic, and HMDA IV data, 2019–2025, with both timing measures. Standard errors are clustered by mortgage county.

Panel A: Lock-to-Close Days					
Dep. Variable:	Lock-to-Close Days				
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}(\text{High-Cost Broker Firm})$	−1.26 (2.51)	−1.87 (2.63)	−2.30 (2.62)	−1.41 (2.72)	−1.80 (2.74)
Origin County $\times$ Lock-Month FE	✓	✓	✓	✓	✓
Mortgage County $\times$ Lock-Month FE		✓	✓	✓	✓
Lock-Date FE			✓		✓
Lender $\times$ Lock-Month FE				✓	✓
Borrower controls	✓	✓	✓	✓	✓
Origin controls	✓	✓	✓	✓	✓
First-stage F -statistic	93.9	86.1	84.9	100.5	101.2
Observations	143,947	137,655	137,484	132,228	132,047
Dep. Var. Mean	32.27	32.36	32.37	32.32	32.33
Dep. Var. SD	17.22	17.29	17.29	17.16	17.16

Panel B: Measures of Loan Timing				
Dep. Variable:	Lock-to-Close Days	Requested Lock Period (Days)	$\mathbb{1}(\text{Lock Period} > 30 \text{ Days})$	Slack (Days)
	(1)	(2)	(3)	(4)
$\mathbb{1}(\text{High-Cost Broker Firm})$	−1.87 (2.63)	−1.58 (2.23)	−0.051 (0.082)	0.29 (1.80)
Origin County $\times$ Lock-Month FE	✓	✓	✓	✓
Mortgage County $\times$ Lock-Month FE	✓	✓	✓	✓
Borrower controls	✓	✓	✓	✓
Origin controls	✓	✓	✓	✓
First-stage F -statistic	86.1	86.1	86.1	86.1
Observations	137,655	137,655	137,655	137,655
Dep. Var. Mean	32.36	36.02	0.391	3.66
Dep. Var. SD	17.29	11.67	0.488	14.97

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table IA.8: Borrower and Neighborhood Characteristics by the Sign of the Instrument

This table compares borrower and neighborhood characteristics by relative proximity to high-cost and low-cost branches. Differential distance dd is distance to the nearest low-cost branch minus distance to the nearest high-cost branch. High-cost firms are classified using previous-year price effects. The groups are $dd < 0$ and $dd \geq 0$. The table reports unconditional means, medians, mean differences, and differences divided by the characteristic's standard deviation. Origins are previous addresses for purchases and mortgaged properties for refinances; neighborhood measures come from the origin block group and tract in the American Community Survey. Dollar values are in thousands, neighborhood shares are in percent, and variables marked (1/0) are indicators. The sample is owner-occupied brokered loans with traced origins in the matched Optimal Blue, CoreLogic, and HMDA IV data, 2019–2025. Standard errors are clustered by mortgage county.

	$dd < 0$		$dd \geq 0$		Difference		
	Mean	Median	Mean	Median	Diff.	SE	Diff./SD
Credit score	755	765	752	762	−2.83***	(0.585)	−0.064
Loan amount (\$000s)	369	355	345	325	−23.9***	(4.47)	−0.149
Property value (\$000s)	591	515	534	465	−57.3***	(12.4)	−0.167
Income (\$000s)	131	112	128	108	−3.71***	(1.31)	−0.045
Combined loan-to-value	68.1	71.2	70.1	74	1.95***	(0.397)	0.108
Debt-to-income	36.8	36	36.9	36	0.128	(0.0798)	0.016
Origination charges (\$000s)	3.27	1.84	3.45	2.1	0.177***	(0.0621)	0.053
Non-white applicant (1/0)	0.185	0	0.157	0	−0.028**	(0.012)	−0.073
Female applicant (1/0)	0.400	0	0.400	0	−0.0003	(0.0050)	−0.001
Applicant age (bin midpoint)	46.5	40	46.2	40	−0.293*	(0.155)	−0.023
Has co-applicant (1/0)	0.470	0	0.456	0	−0.014***	(0.0041)	−0.028
First-time homebuyer (1/0)	0.134	0	0.146	0	0.012***	(0.0042)	0.034
Self-employed (1/0)	0.141	0	0.144	0	0.0031	(0.0023)	0.009
Log population density	7.01	7.34	6.91	7.19	−0.096**	(0.0439)	−0.066
Percent renter	27.5	20.8	26.7	19.5	−0.865*	(0.522)	−0.036
Percent bachelor's degree	44.1	43.1	42.3	40.8	−1.87***	(0.516)	−0.089
Percent in poverty	5.44	2.5	5.74	2.62	0.292***	(0.11)	0.035
Percent minority	40.5	35.2	37.9	31	−2.59**	(1.01)	−0.094
Percent urban	91.2	100	89.8	100	−1.44***	(0.526)	−0.049
Tract: percent employed in finance and real estate	7.8	7.06	7.79	7.05	−0.0043	(0.141)	−0.001
Tract: percent of owner units with a mortgage	69.5	70.4	68	69.1	−1.46***	(0.259)	−0.119
Tract: percent moved in the past year	13	11.6	12.1	10.7	−0.841***	(0.146)	−0.116
Tract: percent foreign-born	17.1	13.5	16.7	12.5	−0.432	(0.513)	−0.032
Tract: percent of housing built 2010 or later	9.79	3.52	8.77	3.35	−1.02*	(0.536)	−0.069
Tract: unemployment rate (percent)	4.8	4.16	4.83	4.21	0.0323	(0.05)	0.010
Tract: vacancy rate (percent)	6.49	4.83	6.85	4.96	0.361**	(0.173)	0.049
Observations	88,783		86,026		174,809		

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table IA.9: Borrower and Neighborhood Characteristics by the Sign of the Instrument, Within County and Month

This table compares borrower and neighborhood characteristics by relative proximity to high-cost and low-cost branches, adjusting for county and month. Differential distance dd is distance to the nearest low-cost branch minus distance to the nearest high-cost branch; high-cost firms are classified using their previous-year price levels. Means and medians are unconditional within the $dd < 0$ and $dd \geq 0$ groups. Adjusted differences come from regressions of each characteristic on an indicator for $dd \geq 0$, with mortgage county $\times$ lock-month effects, and are also reported in standard-deviation units. Dollar values are in thousands, neighborhood shares are in percent, and variables marked (1/0) are indicators. Neighborhood characteristics are from the origin block group and tract in the American Community Survey. The sample is the matched Optimal Blue, CoreLogic, and HMDA IV sample, 2019–2025. Standard errors are clustered by mortgage county.

	$dd < 0$		$dd \geq 0$		Difference Within County-Month		
	Mean	Median	Mean	Median	Diff.	SE	Diff./SD
Credit score	755	765	752	762	−0.978***	(0.296)	−0.022
Loan amount (\$000s)	369	355	345	325	−4.51***	(1.27)	−0.028
Property value (\$000s)	591	515	534	465	−9.68***	(3.08)	−0.028
Income (\$000s)	131	112	128	108	−2.64***	(0.686)	−0.032
Combined loan-to-value	68.1	71.2	70.1	74.0	0.295*	(0.164)	0.016
Debt-to-income	36.8	36.0	36.9	36.0	0.112**	(0.049)	0.014
Origination charges (\$000s)	3.27	1.84	3.45	2.10	−0.0069	(0.022)	−0.002
Non-white applicant (1/0)	0.185	0	0.157	0	−0.0020	(0.0059)	−0.005
Female applicant (1/0)	0.400	0	0.400	0	0.0040	(0.0038)	0.008
Applicant age (bin midpoint)	46.5	40.0	46.2	40.0	−0.061	(0.092)	−0.005
Has co-applicant (1/0)	0.470	0	0.456	0	−0.0029	(0.0031)	−0.006
First-time homebuyer (1/0)	0.134	0	0.146	0	0.0065**	(0.0027)	0.019
Self-employed (1/0)	0.141	0	0.144	0	−0.0035*	(0.0020)	−0.010
Log population density	7.01	7.34	6.91	7.19	−0.020	(0.030)	−0.013
Percent renter	27.5	20.8	26.7	19.5	0.306	(0.487)	0.013
Percent bachelor's degree	44.1	43.1	42.3	40.8	−2.01***	(0.308)	−0.095
Percent in poverty	5.44	2.50	5.74	2.62	0.341***	(0.102)	0.041
Percent minority	40.5	35.2	37.9	31.0	0.820	(0.603)	0.030
Percent urban	91.2	100	89.8	100	−0.491	(0.301)	−0.017
Tract: percent employed in finance and real estate	7.80	7.06	7.79	7.05	−0.272***	(0.064)	−0.060
Tract: percent of owner units with a mortgage	69.5	70.4	68.0	69.1	−0.418**	(0.213)	−0.034
Tract: percent moved in the past year	13.0	11.6	12.1	10.7	−0.330***	(0.095)	−0.046
Tract: percent foreign-born	17.1	13.5	16.7	12.5	0.272	(0.286)	0.020
Tract: percent of housing built 2010 or later	9.79	3.52	8.77	3.35	−0.311	(0.291)	−0.021
Tract: unemployment rate (percent)	4.80	4.16	4.83	4.21	0.105***	(0.032)	0.032
Tract: vacancy rate (percent)	6.49	4.83	6.85	4.96	0.088	(0.089)	0.012
Observations	88,783		86,026		174,809		
Mortgage County $\times$ Lock-Month FE					✓		

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table IA.10: The Steering Gap Under Other Honest Brokers

This table re-estimates the main column of each table with the steering gap measured against other honest brokers. Each panel reports the coefficient of interest from one column of a main table, with the same controls and fixed effects: Panel A from Table 4, Column (2); Panels B and C from Table 5, Panel A, Column (2), and Panel B, Column (1); Panel D from Table 6, Column (3); and Panel E from Table 7, Column (2). In Panels B to D, the past firm gap is the trailing mean of the same measure. Column (1) uses the observed gap against the lowest predicted rate spread, before the honest-broker adjustment. Column (2) repeats the steering gap of the main tables, in which the honest broker shops the state set and every lender quotes with a noise standard deviation of 35 basis points. Column (3) gives each lender its own noise standard deviation, the leave-one-out root mean squared error of its pricing model, and Column (4) a common standard deviation of 111 basis points. Column (5) measures the steering gap against an honest broker that shops only the priced lenders on the broker's menu, with the same noise of 35 basis points, so it leaves out what an honest broker gains by shopping the whole state set. The value of 111 basis points is the noise level at which simulated honest brokers in Texas and Ohio, with quote noise of 30 basis points and rates priced by pricing models refit on the simulated loans, have a mean steering gap of zero. This level was calibrated for a steering gap measured against the broker's own menu rather than the state set, so here it is a sensitivity check only. The sample is brokered loans with a measured steering gap. Standard errors are clustered by county.

Gap:	Observed (Minimum)	Steering Gap (Observed minus Honest Broker)			
Honest broker shops: Quote noise SD:		State Set 35 bps	State Set Lender RMSE	State Set 111 bps	Menu 35 bps
	(1)	(2)	(3)	(4)	(5)
<i>Panel A: Rate Spread on the Gap (Table 4, Column 2)</i>					
Coefficient	0.677*** (0.016)	0.729*** (0.016)	0.740*** (0.016)	0.767*** (0.017)	0.613*** (0.015)
Dep. Var. Mean	30.53	30.53	30.53	30.53	30.53
Dep. Var. SD	49.09	49.09	49.09	49.09	49.09
Observations	304,526	304,526	304,526	304,526	304,526
<i>Panel B: Gap on the Past Firm Gap (Table 5, Panel A, Column 2)</i>					
Coefficient	0.450*** (0.019)	0.456*** (0.019)	0.450*** (0.019)	0.449*** (0.018)	0.483*** (0.019)
Dep. Var. Mean	26.71	10.88	6.90	1.39	8.25
Dep. Var. SD	21.62	19.35	19.21	16.88	19.16
Observations	260,013	260,013	260,013	260,013	260,013
<i>Panel C: Firm Gap on Its Value 12 Months Earlier (Table 5, Panel B, Column 1)</i>					
Coefficient	0.612*** (0.023)	0.602*** (0.025)	0.594*** (0.025)	0.568*** (0.029)	0.664*** (0.018)
Dep. Var. Mean	29.93	13.29	9.16	2.74	9.75
Dep. Var. SD	18.06	16.41	16.37	13.90	15.98
Observations	64,232	64,232	64,232	64,232	64,232
<i>Panel D: Rate Spread on the Past Firm Gap (Table 6, Column 3)</i>					
Coefficient	6.57*** (0.43)	7.06*** (0.42)	7.26*** (0.43)	7.89*** (0.41)	4.15*** (0.33)
Dep. Var. Mean	27.69	27.69	27.69	27.69	27.69
Dep. Var. SD	45.15	45.15	45.15	45.15	45.15
Observations	245,890	245,890	245,890	245,890	245,890
<i>Panel E: Broker Compensation on the Gap (Table 7, Column 2)</i>					
Coefficient	0.099*** (0.011)	0.114*** (0.013)	0.117*** (0.013)	0.130*** (0.014)	0.104*** (0.012)
Dep. Var. Mean	205.39	205.39	205.39	205.39	205.39
Dep. Var. SD	50.96	50.96	50.96	50.96	50.96
Observations	178,648	178,648	178,648	178,648	178,648

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table IA.11: The Observed Gap and Realized Rates

This table repeats Table 4 with the observed gap in place of the steering gap. It regresses realized rate spreads on the loan's observed gap, both in basis points. The observed gap is the chosen lender's predicted spread minus the lowest predicted spread in the state lender set, before the honest-broker adjustment. Column (1) includes county $\times$ lock-month effects. Column (2) adds firm effects, Column (3) adds standardized firm HHI over the previous 12 months, and Column (4) adds lock-date effects. HHI is the sum of squared lender shares in the firm's loans. Controls are income, credit score, loan amount, property value, combined loan-to-value ratio, debt-to-income ratio, race, ethnicity, sex, occupancy, property type, number of units, loan purpose, and loan term. The sample is brokered loans with a measured steering gap. Standard errors are clustered by county.

Dep. Variable:	Rate Spread (bps)			
	(1)	(2)	(3)	(4)
Observed gap (bps)	0.722*** (0.017)	0.677*** (0.016)	0.678*** (0.016)	0.673*** (0.015)
Broker firm HHI (std.)			0.647*** (0.149)	0.539*** (0.155)
County $\times$ Lock-Month FE	✓	✓	✓	✓
Lock-Date FE				✓
Broker Firm FE		✓	✓	✓
Borrower and loan controls	✓	✓	✓	✓
Observations	308,810	304,526	297,131	296,937
Adjusted R^2	0.491	0.544	0.543	0.569
Dep. Var. Mean	30.9	30.5	30.5	30.5
Dep. Var. SD	49.7	49.1	49.0	49.0

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table IA.12: Past and Current Observed Gaps

This table repeats Table 5 with the observed gap in place of the steering gap. Panel A regresses a loan's observed gap on its firm's mean observed gap during the previous 12 months. All columns include borrower and loan controls and standardized firm HHI. Column (1) includes county $\times$ lock-month effects, Column (2) adds firm effects, and Column (3) adds lock-date effects. Panel B regresses the firm's mean observed gap in months $t - 12$ to $t - 1$ on its mean in months $t - 24$ to $t - 13$. Observations are firm-months assigned to the firm's modal county of the year. Columns (1)–(3) include county $\times$ month, state $\times$ month, or month effects, without firm effects; singleton county-months are excluded in Column (1). Each window requires at least five loans. The observed gap is the chosen lender's predicted spread minus the lowest predicted spread in the state lender set, before the honest-broker adjustment. Gaps are in basis points. Controls are income, credit score, loan amount, property value, combined loan-to-value ratio, debt-to-income ratio, race, ethnicity, sex, occupancy, property type, number of units, loan purpose, and loan term. The sample is brokered loans with a measured steering gap. Standard errors are clustered by county.

Panel A: Past Firm Observed Gap and Focal Loan Observed Gap			
Dep. Variable:	Focal Loan Observed Gap (bps)		
	(1)	(2)	(3)
Past firm observed gap (bps, 12 months)	0.723*** (0.009)	0.450*** (0.019)	0.448*** (0.019)
Broker firm HHI (std.)	−0.091 (0.079)	−0.001 (0.141)	0.002 (0.140)
County $\times$ Lock-Month FE	✓	✓	✓
Lock-Date FE			✓
Broker Firm FE		✓	✓
Borrower and loan controls	✓	✓	✓
Observations	260,440	260,013	259,822
Adjusted R^2	0.502	0.538	0.541
Dep. Var. Mean	26.72	26.71	26.71
Dep. Var. SD	21.63	21.62	21.62
Panel B: Autocorrelation of the Firm Observed Gap			
Dep. Variable:	Firm Observed Gap, Months $t - 12$ to $t - 1$ (bps)		
	(1)	(2)	(3)
Firm observed gap, months $t - 24$ to $t - 13$ (bps)	0.612*** (0.023)	0.628*** (0.018)	0.689*** (0.018)
County $\times$ Month FE	✓		
State $\times$ Month FE		✓	
Month FE			✓
Observations	64,232	74,601	74,964
Adjusted R^2	0.596	0.567	0.485
Dep. Var. Mean	29.93	29.67	29.58
Dep. Var. SD	18.06	18.19	18.21

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table IA.13: Past Firm Observed Gap and Mortgage Rates

This table repeats Table 6 with the observed gap in place of the steering gap. It regresses realized rate spreads on past firm observed gaps. The past gap averages the firm's loans in the previous 12 months, with at least five loans, and is divided by its standard deviation. The observed gap is the chosen lender's predicted spread minus the lowest predicted spread in the state lender set, before the honest-broker adjustment. All columns include county $\times$ lock-month effects. Columns (3)–(5) add firm effects, and Column (5) adds lock-date effects. Columns (2), (4), and (5) control for standardized firm HHI, the sum of squared lender shares over the previous year. The dependent variable is in basis points. Controls are income, credit score, loan amount, property value, combined loan-to-value ratio, debt-to-income ratio, race, ethnicity, sex, occupancy, property type, number of units, loan purpose, and loan term. The sample is brokered loans with a measured steering gap. Standard errors are clustered by county.

Dep. Variable:	Rate Spread (bps)				
	(1)	(2)	(3)	(4)	(5)
Past firm observed gap (std.)	9.127*** (0.443)	9.154*** (0.446)	6.573*** (0.428)	6.613*** (0.430)	6.310*** (0.399)
Broker firm HHI (std.)		0.053 (0.166)		0.682*** (0.242)	0.555** (0.239)
County $\times$ Lock-Month FE	✓	✓	✓	✓	✓
Lock-Date FE					✓
Broker Firm FE			✓	✓	✓
Borrower and loan controls	✓	✓	✓	✓	✓
Observations	246,342	244,929	245,890	244,492	244,297
Adjusted R^2	0.434	0.434	0.486	0.485	0.518
Dep. Var. Mean	27.7	27.8	27.7	27.8	27.8
Dep. Var. SD	45.2	45.2	45.1	45.2	45.2

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table IA.14: The Observed Gap and Broker Compensation

This table repeats Table 7 with the observed gap in place of the steering gap. It regresses lender-paid broker compensation on the loan's observed gap. Compensation is measured in basis points of the loan amount, and the gap in basis points of rate spread. The observed gap is the chosen lender's predicted spread minus the lowest predicted spread in the state lender set, before the honest-broker adjustment. Column (1) has no fixed effects. Column (2) includes county $\times$ lock-month and firm effects; Column (3) adds lock-date effects. Controls are income, credit score, loan amount, property value, combined loan-to-value ratio, debt-to-income ratio, race, ethnicity, sex, occupancy, property type, number of units, loan purpose, and loan term. The sample is brokered loans with a measured steering gap. Only loans with recorded compensation are included. Standard errors are clustered by county.

Dep. Variable:	Broker Compensation (bps of Loan Amount)		
	(1)	(2)	(3)
Observed gap (bps)	0.211*** (0.014)	0.099*** (0.011)	0.099*** (0.011)
County $\times$ Lock-Month FE		✓	✓
Lock-Date FE			✓
Broker Firm FE		✓	✓
Borrower and loan controls	✓	✓	✓
Observations	197,611	178,648	178,451
Adjusted R^2	0.256	0.887	0.888
Dep. Var. Mean	207.2	205.4	205.4
Dep. Var. SD	51.1	51.0	51.0

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$