Credit Without Proximity: Informational Frictions and Unequal Gains from Technology *

Erica Xuewei Jiang † Yeonjoon Lee ‡ Quinn Maingi §

Abstract

We study how the organization of information production—and its response to economic and technological forces—affects informational efficiency, credit allocation, and borrower risk. Using U.S. administrative data linking mortgage applications to loan officers and subsequent loan performance, we show that underwriting facilitated by officers located close to the borrower increases approval rates without worsening ex-post performance or processing speed, but is not always deployed where it is most valuable, because lenders allocate loan-officer labor elastically with respect to local wages. These gains are especially large for observably riskier borrowers. We develop and estimate a model that combines a core information-production problem over latent borrower risk, an endogenous choice over local versus remote underwriting, and equilibrium in mortgage and labor markets. We find substantial baseline credit rationing—up to 15 percent in high-risk segments—with local officers eliminating roughly half of it while also reducing excessively risky approvals. A technology shock that raises the processing productivity of remote officers induces lenders to substitute away from local screening, lowering informational efficiency, increasing excessively risky approvals and expected defaults, and tightening rationing for marginal borrowers despite only modest reductions in interest rates.

Keywords: informational frictions, mortgages, banking, technology
JEL classification : G21, G51, J24, L15, L23, R12, R23

*We thank Jason Allen, Vitaly Bord, Robert Clark, Chuqing Jin, Naz Koont, Noah Stoffman, Ryan Westphal the seminar participants at the Bank of Canada, Syracuse University, the University of Lugano, the University of Zurich, USC Marshall, the Richmond Fed RQS Brownbag, and Stress Testing Research Seminar, and the conference participants at Texas Alumni Conference, UBC Winter Conference, FMA Napa Valley Conference, St. Louis Fed System Economic Heterogeneity Conference, GA Tech/Atlanta Fed Household Finance Conference, UDel/Philly Fed Fintech Conference, IIOC, and Philly Fed Consumer Credit & Macro for helpful comments. We are grateful to the Conference of State Bank Supervisors (CSBS) for granting us access to data from NMLS Consumer Access SM. The views expressed are those of the authors and not necessarily those of the CSBS, the Structured Finance Association or any of its member firms, the Federal Reserve Bank of Richmond, or any other entity within the Federal Reserve System.

†UCLA Anderson School of Management, erica.jiang@anderson.ucla.edu.
‡Federal Reserve Bank of Richmond, yeonjoon.lee1@gmail.com.
§University of Southern California, Marshall School of Business, qmaingi@marshall.usc.edu.

Financial intermediation creates value by producing and using information about borrower risk. Because borrower quality is imperfectly observed, informational frictions can distort credit allocation, and intermediaries mitigate these frictions through costly screening (Diamond, 1984; Leland and Pyle, 1977; Ramakrishnan and Thakor, 1984). But the production of information is itself an economic activity: it depends on labor, technology, and organizational design, and therefore need not be concentrated where information is most valuable. As a result, understanding how information production is organized—and how it responds to economic forces—is central to understanding credit allocation.

Traditionally, information production in credit markets has relied on local presence, but technological change increasingly allows lenders to process applications remotely. If remote technologies lower the cost of underwriting without preserving the informational value of proximity, lenders may substitute away from higher-precision local information production toward lower-cost remote processing, even in settings where local information is especially valuable. In that case, technological progress can raise processing productivity while reducing informational efficiency, with important consequences for credit access and risk.

We study how the organization of information production responds to economic and technological forces, and how those responses shape informational efficiency, credit allocation, and borrower risk. We combine new reduced-form evidence with a structural model of the U.S. mortgage market, using administrative data that link the near-universe of mortgage applications to the loan officers who review them and to the subsequent performance of approved loans. We show that underwriting facilitated by officers located close to the borrower increases approval rates without worsening ex-post performance or processing speed, but is not always deployed where it is most valuable, because lenders allocate loan-officer labor elastically with respect to local wages. We then develop and estimate a model that combines a core information-production problem over latent borrower risk, an endogenous choice over local versus remote underwriting, and equilibrium in mortgage and labor markets. The estimated model implies substantial baseline credit rationing, especially in high-risk segments, and shows that local information production alleviates a large share of that rationing while reducing excessively risky approvals. In counterfactual exercises, we show that technologies that make remote processing cheaper can weaken informational efficiency, increasing false approvals, expected defaults, and rationing for marginal borrowers.

We begin by documenting two reduced-form facts about information production and the allocation of origination activity in mortgage lending, together with evidence on the underlying mechanism. These results are informative in their own right and also guide the structural analysis that follows.

Fact 1 establishes the value of local information production. Applications handled by local loan officers are rejected less often in both home-purchase and refinance markets. Crucially, these additional approvals do not come at the expense of worse ex-post performance: default is lower for home-purchase loans and statistically indistinguishable from zero for refinance loans in the most saturated specifications. By contrast, conditional differences in interest rates are economically small. Local handling also delivers materially faster decisions in refinance markets, where workflows are more standardized; the corresponding home-purchase effects become small under saturated controls. These patterns remain robust to saturated borrower and loan controls, including AUS recommendations, as well as county–month, lender–month, and loan-officer fixed effects, indicating that they do not simply reflect observable borrower composition, local market conditions, lender-specific policies, or differences in officer quality.

Fact 2 documents how lenders allocate loan origination activity. Using MSA-level finance-sector wages from the Bureau of Labor Statistics, we show that applications from higher-wage markets are less likely to be handled by local loan officers, even after controlling for borrower and loan characteristics and absorbing lender–year and state–year fixed effects. At the lender–market level, this translates into systematic wedges between each market's share of applications and its share of loan-officer capacity, with the wedges largest in higher-wage markets. Lenders therefore do not allocate local loan-officer capacity solely according to where the gains from proximity are largest; labor-market forces instead push origination activity away from some of the markets where local information would be most valuable.

Facts 1 and 2 show that lenders face an optimization problem where they balance the underwriting benefits of local officers against spatial wage differentials. Evaluating the welfare consequences of the resulting geographic mismatch requires unpacking the precise mechanism behind these benefits. We use a series of empirical tests—including one that exploits the labor-allocation patterns in Fact 2—to show that superior information production is the primary driver of the local advantage, with no detectable role for advantageous borrower sorting (i.e., local loan officers do not simply attract a systematically safer applicant pool).

First, we provide direct evidence isolating the role of information. We show that approval decisions made under local handling depend less mechanically on standardized borrower and loan observables; the same set of hard-information variables has systematically lower explanatory power for rejection decisions made by local officers. This is consistent with local officers bringing in additional soft information not captured in standardized application fields. Both the institutional setting and our conceptual framework then distinguish where this additional information should manifest in the data. In the U.S. mortgage market, rates are quoted before borrower information is fully verified, and upward repricing is strictly limited once underwriting is underway. We show theoretically that if superior underwriting information can still be incorporated into final prices, better information should make prices more tightly linked to ex-post default among observably similar borrowers; if ex-post repricing is curtailed, informational differences must appear primarily through the approval margin. The data align with this second case: local handling delivers a noticeably flatter residualized rejection–default relationship than non-local handling, while the relationship between residualized loan prices and default is nearly identical across the two groups. Consistent with an approval-margin channel, the local advantage is concentrated among borrowers who appear risky on observables—those with low FICO, high DTI, or high LTV. In these segments, codified hard information is less useful for distinguishing safer from riskier applicants, allowing local officers to safely expand approvals without worsening realized performance.

The remaining concern is that local officers may simply face a more favorable applicant pool—for example, because better-quality borrowers self-select into in-person rather than remote applications. If this kind of advantageous sorting were driving our results, the local advantage should be largest in markets where local officers are scarce relative to demand and best positioned to cream-skim the applicant pool. Building on Fact 2, which shows that local finance wages strongly predict local-officer scarcity, we test whether the local rejection advantage is larger in higher-wage markets. Within lenders, we find no such pattern: the wage interaction is statistically and economically indistinguishable from zero, indicating that within-lender assignments are not systematically tilted toward better borrowers. If anything, the across-lender evidence points in the opposite direction, with lenders that employ more local officers facing a somewhat less favorable borrower pool.

Taken together, the reduced-form evidence shows that proximity improves both informational precision and processing speed, yet lenders allocate loan-officer labor elastically to local wages. Because informational gains and labor costs vary across markets, this responsiveness need not place underwriting capacity where local information is most valuable. These findings raise two central questions. When lenders rely more heavily on lower-precision remote underwriting, how much distortion appears as credit rationing, with safer marginal borrowers being excluded, and how much appears as false approvals, with riskier borrowers being funded? And how do technologies that increase the processing productivity of remote officers—without improving their informational precision—reshape information production and, in turn, credit allocation, access, and risk?

Addressing these questions requires a framework in which pricing, borrower selection, labor allocation, and information production interact endogenously. We develop a structural model that links mortgage-market competition to the labor market for loan officers. At its core, the model is an information-production problem over latent borrower risk: borrowers differ in an unobserved type that affects expected repayment and thus lender profits, but lenders observe only a prior distribution when setting posted mortgage rates. Borrowers choose among lenders based on posted rates and lender–market-specific amenities, and because price sensitivity varies with latent risk, this choice generates both demand elasticities and selection on unobserved risk.

The model's central innovation is that underwriting-stage information affects profits through the approval margin. Consistent with the institutional environment in U.S. mortgage origination, prices are set before borrower-specific information is collected, so additional information generated during underwriting is used primarily to determine which applications are funded rather than to reprice loans ex post. We endogenize this information production by allowing lenders to choose between local and remote loan officers, trading off the higher signal precision of local underwriting against the lower cost and greater scalability of remote processing. Loan officers choose among lenders and markets based on posted wages and idiosyncratic location preferences, so wage-setting determines each lender's mix of local and remote underwriting capacity.

These ingredients are embedded in equilibrium. Once applications arrive, lenders assign them to officer types subject to capacity constraints, officers receive noisy signals of borrower risk, and lenders approve only those applications whose expected default losses fall below the posted net interest margin. In equilibrium, borrowers sort across lenders, lenders compete in mortgage and labor markets, and loan officers supply labor across lenders and locations. The framework therefore jointly determines posted rates, applicant composition, information-production choices, approval decisions, labor allocations, and realized borrower risk. It allows us to quantify the informational value of local underwriting, the degree of credit rationing that arises when lenders rely more heavily on lower-precision remote officers, and the equilibrium consequences of shocks—such as improved processing productivity of remote labor—for credit access, default risk, and welfare.

We calibrate the model to quantify baseline credit misallocation due to informational frictions, which requires recovering the information environment and borrower risk distribution faced by lenders. The key primitives are the distribution of latent borrower types, the precision of signals, and the correlation between default risk and price sensitivity. While these parameters are jointly determined in equilibrium, their identification relies intuitively on different effects across moments. All else equal, a riskier average pool raises default and lowers origination; a more dispersed pool can lower default through better approval decisions; the screening precision of local and non-local officers jointly pin down the local rejection and default advantages. We identify these primitives by jointly matching market-level origination rates, expected default patterns, the empirical relationship between residualized interest rates and residualized default, and the local rejection and default advantages documented in our reduced-form analysis.

Conditional on this information environment, we then identify the technological differences between local and remote underwriting. The key object is processing productivity, which governs how much underwriting can be performed with a given amount of labor. We identify processing productivity by matching the observed shares of local versus remote underwriting across lender types. Remaining market-level shocks are chosen to align model-implied default levels with realized performance across mortgage segments. Together, these moments pin down the informational and operational differences between local and remote underwriting that drive misallocation in the model.

The structural model provides a clear characterization of how information frictions shape equilibrium credit allocation. In the baseline, informational gains primarily shift approval cutoffs downward, reducing inefficient rejections, while only modestly tightening cutoffs on the risky tail. Because interest rates are posted before high-quality borrower signals are observed, lenders manage risk almost exclusively through the approval margin. Consequently, informational advantages translate disproportionately into reductions in false rejections: the model implies substantial rationing in equilibrium—up to 15 percent of applicants in the most risk-intensive refinance segments—and local officers eliminate roughly half of this rationing. Local officers also reduce the comparatively small volume of false approvals, consistent with their ex-post default advantage in the data. Taken together, the results mirror the reduced-form patterns: proximity expands credit supply for marginal borrowers while improving realized performance, with the largest gains in markets where unobserved heterogeneity is most severe.

We next use the model to study how technological improvements to remote underwriting—modeled as increases in the processing efficiency of non-local officers without any enhancement in their informational precision—reshape equilibrium credit provision. Because lenders' labor choices respond elastically to relative labor costs, even modest efficiency gains prompt substitution away from information-rich local screening toward lower-cost remote labor. This reallocation reduces informational efficiency and affects both margins of origination. Rationing rises materially across markets, with the largest increases in subprime and near-prime purchase segments where local screening is most valuable ex ante. At the same time, false approvals expand, raising expected default. Posted interest rates fall slightly, consistent with a standard cost-reducing shock, but the gains are small—typically no more than about $25 per month. In short, improvements in processing productivity do not translate into meaningful benefits for borrowers once the induced loss of information is taken into account.

Finally, the model reveals how informational losses amplify aggregate credit risk. When remote processing becomes more efficient and lenders substitute away from local screening, the marginal loans that are newly approved are closer to the default boundary. Their default probabilities therefore rise sharply in mild downturns, generating realized defaults that exceed the baseline. In sufficiently severe downturns, however, this effect becomes nonlinear: once these marginal loans cross the steep region of the default risk curve, their probabilities approach saturation, and the loans that would have been rejected in the counterfactual become those whose risk increases most. As a result, the impact of the technology shock on realized defaults is non-monotonic in aggregate conditions. Benchmarking the model to GFC-era default realizations, informational frictions increase aggregate risk at shock magnitudes comparable to, and in some cases larger than, those experienced during the crisis. Overall, the model highlights an important economic trade-off: expanding remote underwriting capacity lowers production costs, but it also erodes the information discipline that constrains credit supply, increasing rationing and aggregate risk in economically meaningful ways that are only weakly offset by modest reductions in mortgage rates.

Literature Review. Our paper contributes to a structural literature on informational frictions in credit markets. A central idea in this literature is that asymmetric information can distort credit-market allocation and give rise to credit rationing (Stiglitz and Weiss, 1981; Jaffee and Modigliani, 1969). One response to this problem is that lenders rely on soft, local, and private information when standardized borrower characteristics are incomplete guides to repayment risk (Petersen and Rajan, 2002; Berger et al., 2005; Agarwal and Hauswald, 2010; Liberti and Mian, 2008). In the U.S. mortgage market specifically, Bosshardt et al. (2025) provide evidence that lenders price default risk beyond what is captured by the automated underwriting systems, consistent with lender-level information production. Structural work then studies how borrower heterogeneity, asymmetric information, pricing, and credit-market outcomes interact (Crawford et al., 2018; Adams et al., 2009; Einav et al., 2012; DeFusco et al., 2022). Our paper is also closely related to Agarwal et al. (2024), which highlights the economic importance of the approval margin in credit markets. We build on this literature but shift the focus from credit-market outcomes under fixed screening technologies to lenders' endogenous choice of how to produce underwriting information, embedding that choice in joint mortgage-market and labor-market equilibrium.

A related emerging literature studies how technology changes the information available to lenders, showing that improvements or disruptions in information production can meaningfully reshape credit allocation (Fuster et al., 2022; Berg et al., 2020; Blattner and Nelson, 2021; Blattner et al., 2021, 2022; Cherry, 2025). Our contribution is to show that technological advances in lending can reshape information production even when they do not directly improve informational accuracy. By lowering the cost and increasing the scalability of remote underwriting, such technologies induce lenders to substitute away from local information production in equilibrium. This endogenous reallocation distorts credit allocation through both tighter rationing for some borrowers and greater risk in the system through false approvals, even when posted borrowing costs fall slightly.

Finally, our paper contributes to the literature that applies structural IO and spatial economics tools to consumer finance and financial product markets. One strand uses IO tools to study how competition shapes consumer welfare in mortgages (Allen et al., 2014, 2025, 2019; Buchak et al., 2018; Benetton, 2021; Jiang, 2023), deposits (Egan et al., 2017; Xiao, 2020), payments (Wang, 2025; Whited et al., 2022), and credit cards (Nelson, 2025). We combine these tools with tools from the spatial economics literature on banking (Ji et al., 2023; Maingi, 2026; D'Amico and Alekseev, 2024; Morelli et al., 2025), production networks (Bernard et al., 2019; Oberfield et al., 2024; Giroud et al., 2024; Arkolakis et al., 2025), and multinational production and outsourcing (Antràs et al., 2006; Costinot et al., 2012; Ramondo and Rodríguez-Clare, 2013; Boehm et al., 2019). In particular, Aguirregabiria et al. (2025) study the spatial flow of deposits — a key input to lending — and show how branch networks, scope economies, and local competition shape whether funding reaches high-loan-demand markets. We focus instead on labor, a complementary input, and on its geographic distribution specifically for the purpose of information production. Relatedly, Huang et al. (2025) provide reduced-form evidence that local loan officers increase mortgage originations and improve refinancing outcomes, and attribute the limited spatial reallocation of loan officers to labor-market frictions. Our contribution to this literature is twofold. First, we show that the welfare effects of productivity and technology shocks in credit markets depend critically on how those shocks interact with equilibrium information production. Second, we identify a novel source of spatial misallocation in outsourcing and production-network settings: when inputs such as loan-officer labor differ along two dimensions — processing efficiency and informational precision — outsourcing decisions generate consumer-facing externalities that firms forming those networks do not fully internalize. More broadly, our results suggest that the welfare consequences of production networks and outsourcing depend not only on whether firms reallocate toward cheaper inputs, but also on whether those inputs are equally informative.

1 Data

Our analysis combines three primary data sources: (i) the Nationwide Mortgage Licensing System and Registry (NMLS) loan officer database, (ii) confidential Home Mortgage Disclosure Act (HMDA) loan-level application records, and (iii) Black Knight McDash loan performance data. Together, these datasets allow us to observe where loan officers work, which loan officers process which applications, origination outcomes, and how originated loans subsequently perform.

Loan officer data (NMLS). We begin with administrative records from the Nationwide Mortgage Licensing System and Registry (NMLS), which—under the Secure and Fair Enforcement for Mortgage Licensing Act of 2008 (SAFE Act)—requires every residential mortgage loan officer to maintain a unique license or registration. We obtain the universe of registered and federally licensed mortgage loan officers from 2015 onward.

From these records, we construct a longitudinal dataset containing each loan officer's unique identifier, employer, and work location. Reported locations take one of three forms: *Branch*, *Main*, or *Work*. "Branch" refers to a specific branch office for state-regulated institutions; "Main" refers to the corporate address; and "Work" reflects the individual's actual work address for federally regulated institutions rather than the corporate headquarters. When multiple locations are reported for state-regulated lenders, we prioritize branch locations over main-office addresses. These data allow us to geocode each officer's work location and measure geographic proximity to each applicant.

Loan application and underwriting data (Confidential HMDA). We merge loan officer locations to confidential HMDA, an administrative dataset maintained by the Federal Reserve System that contains near-universe coverage of U.S. mortgage applications. Confidential HMDA includes lender identity, borrower and loan characteristics, applicant location, application and action dates, application outcomes, Automated Underwriting System (AUS) recommendations, and—critically for our purposes—a unique identifier for the loan officer who processed each application beginning in 2018. These data allow us to match individual applications to the specific loan officers who handled them.

Loan performance data (McDash). To measure ex-post outcomes, we merge originated loans in confidential HMDA to monthly servicing records from the Black Knight McDash dataset. Following Rosen (2011), we match loans on origination details, loan terms, and borrower characteristics. The merged dataset covers approximately 36% of approved loans in confidential HMDA and 68% of loans in McDash. For each matched loan, we construct a two-year delinquency indicator equal to one if the loan becomes 60+ days delinquent within 24 months of origination. To ensure a full performance window, we restrict the main estimation sample to applications submitted during 2018–2019.

Merged Sample and Measures After merging NMLS, confidential HMDA, and McDash, we obtain a loan-level dataset in which we observe the loan officer who processed each application, the officer's geographic location, rich borrower- and loan-level observables at application, lender identity and local market characteristics, approval or denial outcomes, interest rates on originated loans, and subsequent loan performance for matched originations. We classify loan officers as either *local* or *remote*. In the baseline specification, a loan officer is *local* if the officer's work location is in the same county as the applicant, and *remote* otherwise. We also consider a continuous measure of proximity given by the geographic distance between the borrower and the loan officer's office. Appendix B provides additional details on sample construction.

Table 1 presents the summary statistics. These linked data provide the key objects needed to study the production of borrower information at the loan-officer level, including approval decisions, processing times, pricing, ex-post loan performance, and the allocation of underwriting labor across space.

2 Institutional Setting

Mortgage origination in the United States separates *pricing* from *underwriting*. Lenders post rate sheets and loan officers quote interest rates before any verified information is collected. Once a borrower submits a formal application, the lender must issue a Loan Estimate within three business days, after which TRID rules sharply restrict upward repricing.[1] Because lenders generally cannot freely adjust pricing upward after reviewing documents, approval decisions—rather than ex-post price adjustments—are the primary mechanism for responding to borrower risk.

Loan officers play a central role in information production. They assemble income, asset, employment, and collateral documentation; underwriters rely on these materials and generally do not collect additional information themselves. Incomplete or inconsistent files frequently lead to denials, making the quality of information produced by the loan officer a key determinant of approval outcomes.

[1] Specifically, changing terms generally requires restarting the origination and disclosure process, which can delay closing and is costly for lenders. For example, Chase Bank offers a $5,000 guarantee that it will meet an agreed upon closing date (J.P. Morgan Chase & Co., 2025). See Appendix A for additional institutional and regulatory details.

This role of loan officers makes the distinction between *local* and *remote* officers central for our analysis. Local officers operate within the borrower's market and can coordinate directly with local employers, real estate agents, appraisers, and title companies. Remote officers—often located in centralized hubs—rely primarily on phone or digital communication and face greater frictions in resolving documentation issues. Because prices cannot be freely adjusted after underwriting begins, these differences in information-production efficiency are more likely to affect approval outcomes and processing time than ex-post loan pricing.

In Appendix C, we formally model these institutional features in a simple model of mortgage origination to derive testable empirical predictions. Rather than assuming a priori that TRID rules fully eliminate ex-post repricing, we first consider a general case where banks can still incorporate ex-post information into their pricing decisions. In this case, the model makes a sharp prediction: superior information leads prices to be more tightly linked to ex-post default among observably similar borrowers. Conversely, when ex-post repricing is curtailed, informational differences manifest primarily through the approval margin, jointly affecting realized default rates and approval rates. This distinction organizes our interpretation of the empirical results and illustrates why joint patterns in approval, default, and pricing are critical to understanding the quality of information production.

3 Reduced-Form Evidence

We begin by documenting two reduced-form facts about information production and the allocation of origination activity in mortgage lending, together with evidence on the underlying mechanism.

3.1 Fact 1: Local Origination Improves Loan Outcomes

We begin by showing that local loan officers are associated with superior lending outcomes along two dimensions: the *quality* of underwriting decisions— reflected in approvals, pricing, and ex-post default—and the *quantity productivity* of underwriting, reflected in the speed with which applications move through the pipeline. We formalize this comparison using the following specification:

$$Y_{i\ell ct} = \beta\,\mathrm{Local}_{i\ell ct} + \gamma' X_{i\ell ct} + \delta_{c\tau} + \lambda_{\ell\tau} + \mu_o + \varepsilon_{i\ell ct}, \tag{1}$$

where $Y_{i\ell ct}$ is alternatively an indicator for rejection, an indicator for default within two years of origination, the interest rate on originated loans, or the number of days between application and action on originated loans; $\mathrm{Local}_{i\ell ct}$ equals one when the reviewing officer is local to the borrower; $X_{i\ell ct}$ includes detailed borrower and loan observables, including AUS recommendations; $\delta_{c\tau}$ and $\lambda_{\ell\tau}$ are county–month and lender–month fixed effects; and μ_o are loan-officer fixed effects. This specification isolates variation in whether an application is handled locally or remotely, conditional on rich borrower and loan controls, county–month fixed effects, lender–month fixed effects, and loan-officer fixed effects. County–month fixed effects absorb local market conditions, while lender–month fixed effects absorb lender-specific pricing and underwriting policies at a point in time. Most importantly, our ability to include loan-officer fixed effects as a robustness test ensures that the estimates are identified from within-officer variation in proximity rather than from permanent differences in officer quality.

Tables 2 and 3 present the results. Panels A–C of Table 2 report rejection, interest-rate, and ex-post default outcomes, respectively, while Table 3 reports processing-time outcomes.

Approval decisions. Panel A of Table 2 shows that local handling is associated with significantly lower rejection rates in both home-purchase and refinance markets. In column (1), the local coefficient is −1.881 percentage points for home-purchase loans and −7.898 percentage points for refinance loans. Once lender–month fixed effects are added in column (2), the magnitudes decline to −0.643 and −2.855 percentage points, indicating that part of the raw local/non-local difference reflects which lenders rely more heavily on local versus remote handling. In column (3), after adding loan-officer fixed effects, the coefficients remain statistically significant at −0.202 percentage points for home-purchase loans and −0.655 percentage points for refinance loans. Thus, even comparing the same officer handling observably similar applications, local proximity continues to expand approval. Appendix Table A1 shows the same qualitative pattern using continuous distance.

Loan pricing. Panel B of Table 2 shows that the pricing effects are much smaller than the approval effects, but not uniformly zero. In the less saturated specifications, local handling is associated with slightly higher interest rates: for example, in column (2), the local coefficient is 0.010 percentage points for home-purchase loans and 0.037 percentage points for refinance loans. A natural interpretation is that lenders with better local information production may be willing to post or sustain somewhat higher rates, knowing that they can better screen the riskier applicant pool those rates attract. However, once lender–month and especially loan-officer fixed effects are included, the remaining within-officer pricing effect is very small: -0.002 percentage points for home-purchase loans and 0.009 percentage points for refinance loans. Appendix Table A1 delivers the same conclusion using the continuous distance measure. Overall, the evidence suggests that any pricing response is modest, while the main advantage of local information production appears in the approval margin.

Ex-post performance. Panel C of Table 2 shows that the additional approvals under local handling do not worsen realized loan performance. For home-purchase loans, local handling is consistently associated with lower ex-post default, implying that local information production not only expands approval but also improves the selection of approved borrowers. For refinance loans, the raw cross-sectional association is positive, but it falls sharply once lender–month and loan-officer fixed effects are included and becomes economically and statistically close to zero in the most saturated specification. This pattern suggests that the raw refinance default difference is largely driven by sorting across lenders and officers rather than by proximity itself. Appendix Table A1 shows the same qualitative pattern using continuous distance: greater distance predicts higher default for home-purchase loans, while the refinance relationship vanishes in the saturated specifications.

Processing speed. The first three margins concern the *quality* of underwriting decisions. Table 3 turns to a complementary *productivity* margin: how quickly applications move through the underwriting pipeline. For home-purchase loans, local handling appears faster in the less saturated specifications, but the effect becomes small and statistically insignificant once lender–month and loan-officer fixed effects are included. For refinance loans, by contrast, the local advantage remains economically meaningful and robust: in the most saturated specification, local handling shortens processing time by more than one day. The distance-based specifications in Table A2 yield the same qualitative conclusion, with greater borrower–officer distance associated with slower processing. Taken together with Panels A–C of Table 2, these results show that local handling delivers faster processing alongside—not at the expense of—higher-quality underwriting decisions.

3.2 Fact 2: Lender Labor Allocation Choices

Fact 1 establishes the *benefits* of local handling. These benefits are not free: local underwriting capacity must be paid local wages, which vary substantially across markets. We next examine how lenders trade off these costs and benefits when allocating loan officers across space.

We begin by examining how individual applications are allocated across local and remote loan officers. Columns (1)–(2) of Table 4 show that applications from higher-wage markets are less likely to be handled by local loan officers. We measure local labor-market conditions using the MSA-level finance-sector hourly wage from the Bureau of Labor Statistics, assigning each borrower the wage of the MSA in which the borrower's county is located. This measure is less susceptible than loan-officer wages to reverse-causality concerns arising from lenders' own local hiring decisions. The relationship remains negative after controlling for borrower and loan characteristics and absorbing lender–year and state–year fixed effects.

These assignment patterns aggregate into systematic market-level mismatch. To measure this, we construct a lender-level spatial imbalance index:

$$\text{Misalignment}_{jkt} = \frac{m_{jkt}}{M_{jt}} - \frac{l_{jkt}}{L_{jt}}, \tag{2}$$

where $\frac{m_{jkt}}{M_{jt}}$ is the share of lender j's applications from MSA k in year t, and $\frac{l_{jkt}}{L_{jt}}$ is the share of its loan officers located there. Positive values indicate that a market contributes more to a lender's application flow than to its local officer capacity. We then estimate:

$$\text{Misalignment}_{jkt} = \beta\,\text{Wage}_{kt} + \gamma_t + \gamma_{jt} + \mathbf{X}_{kt}\delta + \varepsilon_{jkt}, \tag{3}$$

where Wage_{kt} is measured using local finance-sector wages, γ_t are year fixed effects, γ_{jt} are lender–year fixed effects, and $\mathbf{X}_{kt}$ denotes MSA-level controls, including employment rates, income per capita, and earnings per job.

Columns (3)–(5) of Table 4 show that higher local wages are strongly associated with greater misalignment. Lenders therefore place relatively less local underwriting capacity in higher-wage markets and rely more heavily on remote information production in those areas.

Figure 1 shows that these patterns generate systematic geographic mismatch: many high-demand counties are served by relatively little local underwriting capacity, while some lower-demand counties host more officers than demand would predict. Figure 2 shows that these wedges are especially pronounced in the highest-demand areas.

Thus, even though local information production improves approval decisions and sometimes speeds processing, lenders do not allocate local underwriting capacity solely according to where it is most valuable. Instead, labor allocation responds strongly to local wage differentials, creating spatial mismatches between where mortgage demand arises and where local information-production capacity is located.

3.3 Mechanism: Evidence for Informational Advantage

The evidence thus far outlines a clear optimization problem: lenders reap substantial private underwriting benefits from local loan officers (Fact 1), but must weigh these gains against spatial wage differentials that often raise the cost of local processing (Fact 2). The welfare implications of the resulting geographic mismatch depend crucially on the underlying mechanism driving those private benefits. To formalize competing channels, Appendix C develops a screening model in which a lender sets an ex-ante rate, observes a noisy signal of borrower risk after the application is submitted, and decides whether to approve. A central result of this framework is that, holding the applicant pool fixed, the outcomes documented in Fact 1 constitute a sufficient condition for an informational advantage. The primary alternative explanation is therefore that local officers face a more favorable applicant pool (*advantageous sorting*), implying that the observed advantage reflects borrower selection rather than superior information. To distinguish between these channels, we present a series of qualitative tests, complemented by a parametric test for sorting that builds on Fact 2. Taken together, our results below point to information production as the primary driver of the local advantage, with no detectable role for advantageous sorting within lenders.

Use of codified hard information. A long literature in finance has argued that "soft information," best collected in nearby, face-to-face interactions, can yield superior informational efficiency (Petersen and Rajan, 1994; Berger and Udell, 1995). While soft information is not directly observable, we can test whether local officers collect it by comparing how much of the rejection decision is explained by codified hard information for local versus non-local handling. Table 5 shows that across increasingly saturated specifications, the R^2 is consistently lower for local handling, even though both sets of regressions condition on the same borrower and loan characteristics (including AUS recommendations) and the richer specifications absorb county–month, lender–month, and loan-officer fixed effects. Local approval decisions therefore rely less on codified information, consistent with local officers bringing in soft information not captured in standardized application fields.

Information appears in approvals, not pricing. The framework in Appendix C delivers a sharper prediction about *where* an informational advantage shows up. The key distinction is whether the lender can adjust the rate after observing the underwriting signal. If it can, the lender uses the signal both to screen out bad risks and to price the remaining ones more finely; better signals then tighten the relationship between realized rates and realized default. If it cannot—as under TRID, which constrains upward repricing after the loan estimate is issued—the ex-ante rate is pinned down by the break-even condition over the ex-ante risk distribution, so the underwriting signal can affect only the approval decision. A more precise signal therefore expands approval at the same realized default rate, while leaving the rate–default relationship unchanged. Figure 3 shows the pattern predicted by the constrained-pricing case: the (residualized) rejection–default relationship is markedly flatter for local officers than for non-local officers, while the relationship between (residualized) interest rates and default is nearly identical across the two groups. This bifurcation is difficult to reconcile with explanations in which superior information is incorporated into pricing, or in which officers steer better borrowers to apply locally—both of which would predict a tighter rate–default relationship for local officers.

Heterogeneity and risk. If local information helps distinguish among applicants who look similar on codified hard information, its value should be greatest where residual uncertainty is most severe. Because unobserved heterogeneity in credit risk tends to be positively correlated with the average credit risk of a segment (e.g., Iyer et al., 2016), the local advantage should be largest in high-risk segments. Table 6 confirms this: interacting *Local* with indicators for Subprime, High DTI, and High LTV yields significantly larger reductions in rejection for these groups, with no deterioration—and if anything, an improvement—in realized performance. Heterogeneity in pricing is limited by comparison. Figure 4 provides complementary nonparametric evidence within narrow FICO, DTI, and LTV bins. Together, these results indicate that local information production is most valuable precisely where codified hard information is least sufficient for distinguishing safer from riskier applicants.

Direct test for sorting. The patterns above are consistent with superior information production but do not directly speak to whether local officers also face a more favorable borrower pool. Nonparametric tests for sorting are infeasible here, since selection operates through rejection and censors the default risk of screened-out applicants. We therefore develop a parametric test that builds on Fact 2. The logic is a limit argument: if the local advantage were driven by advantageous sorting, its magnitude should diminish as local officers process a larger share of the market. Because Fact 2 shows that local finance wages strongly predict local-officer scarcity, a sorting explanation predicts that the local advantage should be larger in high-wage markets, where local officers are scarce and best positioned to cream-skim. The design also distinguishes sorting *across* lenders from sorting *within* lenders—a distinction relevant for the equilibrium analysis in Section 4.

Table 7 reports the results. Without lender–month fixed effects, the local rejection advantage is somewhat smaller in higher-wage markets, suggesting that lenders employing more local officers may face a less favorable borrower pool on average—a pattern more consistent with adverse than advantageous sorting across lenders. Once lender–month fixed effects are introduced, these interactions attenuate toward zero and are statistically insignificant. The within-lender estimates emphasized throughout this section therefore appear unlikely to be driven by advantageous sorting, while some across-lender selection remains possible and is revisited in the structural analysis.

Interpretation. Taken together, these results match the patterns implied by an informational advantage when ex-post risk cannot be priced. Because lenders set ex-ante prices and use underwriting-stage information only on the approval margin, better information shows up primarily as more approvals at similar or better realized performance, rather than as more refined pricing. This is precisely what we find: local handling materially expands approval, has modest effects on pricing, does not worsen default, relies less on codified hard information, is most valuable in observably risky segments, and is not appreciably weakened where sorting incentives would be strongest. The reduced-form evidence is therefore naturally interpreted as reflecting an informational advantage of local information production.

3.4 Discussion and Implications for Modeling

The reduced-form evidence documents both the benefits and the costs of local information production, and characterizes the mechanism behind the benefits. Fact 1 establishes the benefits: applications handled locally are approved more often without worse ex-post performance, and refinance applications are processed faster, with both results identified from within-officer variation in proximity. Fact 2 establishes the costs: local information production requires paying local wages, and lenders respond strongly to wage differentials, allocating local capacity away from high-wage markets and creating spatial mismatches between where mortgage demand arises and where local information-production capacity is located. Section 3.3 identifies the mechanism: a battery of qualitative tests, together with a parametric test for sorting, indicates that the local advantage reflects superior information production rather than a more favorable applicant pool.

These patterns have direct implications for the industry equilibrium characterized in our structural model. Local loan officers produce superior information, which simultaneously expands credit access and reduces realized risk—a private benefit to the lender that is also a social benefit. At the same time, the absence of advantageous sorting toward local officers within lenders suggests that borrowers do not internalize the informational benefits of proximity when choosing where to apply.[2] Combined with lenders' high sensitivity to local outside-option wages, this raises the possibility that lenders over-rely on remote underwriting in equilibrium, generating credit-rationing externalities that market competition does not fully discipline. The structural model developed in Section 4, disciplined by the moments documented above, allows us to quantify this wedge and to evaluate how shocks that improve remote processing productivity alter credit access, default risk, and welfare.

[2] The across-lender evidence in Section 3.3 suggests some adverse sorting across lenders, but within-lender estimates provide a clean measure of informational advantages that is uncontaminated by these selection effects. We therefore abstract from across-lender sorting in the quantitative model. Extending the model to incorporate such sorting is straightforward; we make this choice out of conservatism, since incorporating across-lender sorting would further widen the wedge between the private and social benefits of local loan officers and strengthen our equilibrium-misallocation results.

4 Structural Model

To quantify the impact of information frictions and technological changes on credit allocation, we develop a structural model that links product market competition over mortgages with the labor market for loan officers. The core of the model is an information-production problem over latent borrower risk. Borrowers differ in an unobserved type that affects expected repayment and, consequently, lender profits. Consistent with TRID rules, lenders must set posted prices based strictly on ex-ante information; therefore, any additional information generated during underwriting is only used to manage risk on the approval margin. We endogenize this information production through the lender's organizational choice between local and remote loan officers, trading off the higher signal precision of local underwriting against the lower cost of remote processing. Finally, we embed this organizational choice within a broader market equilibrium. Borrowers sort across lenders based on posted rates and lender characteristics, lenders compete simultaneously in the mortgage and labor markets, and loan officers optimally supply their labor. As a result, prices, applicant composition, information-production choices, approval decisions, and realized borrower default rates are all jointly determined in equilibrium.

4.1 Primitives and Timing

A market is defined by a unique combination of credit quality, loan purpose, county, and year-quarter (e.g., subprime $\times$ home purchase $\times$ Los Angeles County $\times$ 2018Q1). Markets are therefore segmented by ex-ante observable hard information. The economy consists of a collection of independent markets indexed by $i \in \mathcal{N}$. Each market i is associated with a geographic region $g(i) \in \mathcal{K}$, where $\mathcal{K}$ denotes the set of regions in which loan officers may live. In each market, there is a mass $\mathcal{M}_i$ of potential borrowers and a set of lenders $\mathcal{J}_i$. Each lender j operates in the set of markets $\mathcal{I}_j \subseteq \mathcal{N}$.

A potential borrower η in market i is characterized by price sensitivity θ_η and latent default risk x_η, drawn from the joint distribution $F_i^{x,\theta}$, with idiosyncratic preferences over lenders drawn from F_i^{ϵ}. A mass $\bar{L}$ of potential loan officers chooses a job and where to live by selecting a tuple (i, j, k)—a market i in which to process applications, a lender j to work for, and a region $k \in \mathcal{K}$ in which to live—or a region-specific outside option paying w_{ook}.

An officer who chooses tuple (i, j, k) has screening noise standard deviation σ_{ik}, processing capacity e_{ijk} applications, and idiosyncratic preferences drawn from F^A_{ijk}. Lenders know all primitive distributions but not individual realizations; in particular, x_η depends on hard-to-verify characteristics that are recoverable only through costly screening.

The timing in market i unfolds as follows:

1. **Pricing and wage posting:** Lenders simultaneously post rates and wages under Bertrand–Nash competition in both markets.
2. **Borrower application:** Each borrower observes posted rates and chooses one lender.
3. **Officer choice:** Each officer observes the full wage vector and chooses her tuple.
4. **Screening:** Lenders assign incoming applications across hired officers, who then generate signals of borrower risk.
5. **Origination:** Lenders observe the signals and decide which applications to approve. Because TRID prevents repricing, approval is the only post-application adjustment margin.

With these primitives and the timing in place, we now turn to each building block of the equilibrium. The lender's problem (Section 4.2) develops the information-production problem in which lenders choose how to allocate screening labor across regions that differ in signal precision. Borrower demand (Section 4.3) and loan officer labor supply (Section 4.4) close the model by characterizing the application share $s_{ij}(r_{ij})$ and the supply function $l_{ijk}(\mathbf{w})$ that the lender takes as given.

4.2 Lender's Problem

We first state the lender's optimization problem; we then derive the screening technology that determines the origination probability p_{ijk} and expected default rate d_{ijk}; finally, we characterize the optimal pricing and assignment decisions, with the latter capturing the local-versus-remote tradeoff at the heart of the paper.

Lenders maximize expected profits under rational expectations. Each lender j chooses interest rates r_{ij}, wages w_{ijk}, and loan officer assignment shares $\mathcal{S}_{ijk}$ to maximize profits across all markets in which it operates, subject to loan officer capacity constraints.[3]

Define

$$p_{ijk} \coloneqq \mathbb{P}_{ij}\{orig; r_{ij}, k\}, \qquad d_{ijk} \coloneqq \mathbb{E}_{ij}[d_i(x_\eta) \mid orig; r_{ij}, k]$$

as the region-k origination probability and expected default rate. The lender solves

$$\begin{aligned} \max_{\{r_{ij}, w_{ijk}, \mathcal{S}_{ijk}\}_{i \in \mathcal{I}_j, k \in \mathcal{K}}} \sum_{i \in \mathcal{I}_j} \sum_{k \in \mathcal{K}} & \Big\{ \mathcal{M}_i s_{ij}(r_{ij})\, \mathcal{S}_{ijk}\, p_{ijk}\, (r_{ij} - f_{ij} - d_{ijk}) - w_{ijk} l_{ijk}(w_{ijk}) \Big\} \\ \text{s.t.} \;\; & \mathcal{M}_i s_{ij}(r_{ij}) \mathcal{S}_{ijk} \leq l_{ijk}(w_{ijk}) e_{ijk} \quad \forall i \in \mathcal{I}_j,\, k \in \mathcal{K}. \end{aligned} \tag{4}$$

The objective sums, across regions k, the volume of originated loans $\mathcal{M}_i s_{ij} \mathcal{S}_{ijk} p_{ijk}$ multiplied by the expected profit margin per loan $(r_{ij} - f_{ij} - d_{ijk})$, net of the wage bill $w_{ijk} l_{ijk}$. The constraint requires applications assigned to loan officers living in region k not to exceed the screening capacity available there.[4] Both p_{ijk} and d_{ijk} depend on the rate r_{ij}, through adverse selection in applicant composition (Section 4.3), and on the region k, through signal precision σ^2_{ik}; we derive them from the screening technology below.

Information Production and Approval An officer in region k screening an application in market i observes a noisy signal of the borrower's latent risk,

$$\hat{x}_\eta = x_\eta + \tilde{x}_{\eta k}, \qquad \mathbb{E}[\tilde{x}_{\eta k}] = 0, \qquad \mathrm{Var}(\tilde{x}_{\eta k}) = \sigma^2_{ik},$$

where the noise variance σ^2_{ik} exhibits spatial decay—motivated by Section 3.1 and minimized at $k = g(i)$. Default losses depend on latent risk through a monotone mapping $d_i(x_\eta)$ with $d_i' > 0$. Given the posted rate r_{ij} and per-loan funding cost f_{ij}, the lender approves if and only if

$$\mathbb{E}\big[d_i(x_\eta) \mid \hat{x}_\eta, k\big] \leq r_{ij} - f_{ij},$$

[3]We show in Appendix D.2 that treating the origination decision as a balance-sheet profit-maximization problem is equivalent to selling the loan in an informationally efficient securitization market. Appendix D.2 details the institutional mechanisms—such as put-back covenants—that support this benchmark and provides a framework for modeling cases in which this internalization is incomplete. When the market is not informationally efficient, the externalities from under-employment of geographically proximate screening labor become larger, leading to greater inefficiencies in equilibrium.

[4]For expositional simplicity, we develop the model with a recovery rate of zero on defaulted loans. In the calibration and quantitative exercises, we account for the relatively high recovery rate on first mortgages in the United States.

which defines a cutoff set $\hat{\Delta}_{ij}(r_{ij}, k)$ in signal space. The origination probability and expected default rate introduced above can therefore be written as

$$p_{ijk} = \mathbb{P}\Big(\hat{x}_\eta \in \hat{\Delta}_{ij}(r_{ij}, k)\Big), \qquad d_{ijk} = \mathbb{E}\Big[d_i(x_\eta) \mid \hat{x}_\eta \in \hat{\Delta}_{ij}(r_{ij}, k), k\Big].$$

Two channels jointly determine p_{ijk} and d_{ijk}. The posted rate r_{ij} shifts the risk composition of the applicant pool through adverse selection, governing *who* the lender screens. The signal precision σ^2_{ik} governs the lender's ability to discern risky from safe applicants, determining *which* of them are approved. Together they make screening central to how informational frictions translate into credit rationing.

4.3 Borrower Demand

Borrowers in market i choose which lender to apply to based on posted rates and lender characteristics. Their decisions deliver the application share $s_{ij}(r_{ij})$ that enters the lender's problem (4), and—because price sensitivity and default risk are correlated—also shape the risk composition of each lender's applicant pool.

Borrower η receives random indirect utility

$$u_{ij\eta} = -\theta_\eta r_{ij} + \xi_{ij} + \epsilon_{ij\eta} \tag{5}$$

from applying to lender j, where θ_η is the price sensitivity introduced in Section 4.1, ξ_{ij} is a lender-market demand shifter capturing local presence, staffing, and other residual demand components, and $\epsilon_{ij\eta}$ is an idiosyncratic shock. Each borrower applies to the utility-maximizing lender. Conditional on θ, the application share is

$$s_{ij}(\theta, r_{ij}) = \mathbb{P}_\eta\left(\bigcap_{j' \in \mathcal{J}_i} \left\{-\theta_\eta(r_{ij} - r_{ij'}) + (\xi_{ij} - \xi_{ij'}) \geq \epsilon_{ij'\eta} - \epsilon_{ij\eta}\right\}\right), \tag{6}$$

and integrating over the marginal distribution F_i^θ yields the aggregate share

$$s_{ij}(r_{ij}) \coloneqq \int s_{ij}(\theta, r_{ij})\, F_i^\theta(d\theta). \tag{7}$$

The posted rate also shapes the composition of applicants. The distributions of price-sensitivity and default-risk types among lender j's applicants are

$$\begin{aligned} F_{ij}^{\theta}(t; r_{ij}) &= s_{ij}(r_{ij})^{-1} \int_{-\infty}^{t} s_{ij}(\theta, r_{ij}) F_i^{\theta}(d\theta), \\ F_{ij}^{x}(t; r_{ij}) &= s_{ij}(r_{ij})^{-1} \int_{-\infty}^{t}\int_{-\infty}^{\infty} s_{ij}(\theta, r_{ij}) F_i^{x,\theta}(dx, d\theta). \end{aligned} \tag{8}$$

The wedge between F_{ij}^{θ} and the population distribution F_i^{θ} is the standard selection-on-price effect: less price-sensitive borrowers are disproportionately likely to apply to higher-rate lenders. When x and θ are correlated, this selection propagates to default risk through F_{ij}^{x}. Independence ($x \perp \theta$) reduces (8) to the population distribution $F_i^{x}(t)$; negative correlation produces adverse selection, in which raising r_{ij} worsens the applicant pool and thereby raises the expected default rate d_{ijk}.

4.4 Loan Officer Labor Supply

Officers' tuple choices (i, j, k) over markets, lenders, and regions, introduced in Section 4.1, deliver the supply function $l_{ijk}(\mathbf{w})$ that enters the lender's problem (4).

Officer λ's indirect utility from choice (i, j, k) is

$$u_{ijk\lambda} = a_{ijk\lambda}\, w_{ijk}, \tag{9}$$

where $a_{ijk\lambda}$ is an idiosyncratic preference shifter and w_{ijk} is the wage offered for the tuple. The outside option in region k pays w_{ook}.[5] Following Giroud et al. (2024), $\{a_{ijk\lambda}\}$ follows a joint Fréchet distribution with standard Fréchet marginals and a triple-nested Gumbel copula with region-specific outside options. The mortgage sector is atomistic relative to the broader labor market, so banks face upward-sloping labor supply at the firm level but take outside wages as given.[6] Aggregating over λ delivers the supply function $l_{ijk}(\mathbf{w})$.

[5] As in Giroud et al. (2024), this constant outside-option wage is also consistent with heterogeneous wages paid by outside options, where w_{ook} is the weighted average amenity-adjusted wage. Thus, while we could easily allow for heterogeneous outside options depending also on (destination) market i and firm j, accounting for this does not affect the industry equilibrium.

[6] According to Bureau of Labor Statistics (2023), there are roughly 320,000 loan officers in the United States, compared to 136 million private-sector employees. The framework can accommodate cross-regional migration by loan officers with elasticity ϵ, but the atomistic assumption effectively shuts down this margin, consistent with evidence from Huang et al. (2025) that loan officers rarely move. Calibrating this elasticity to Giroud et al. (2024) leaves the

Labor Cost. Appendix D shows that the cost of processing $\tilde{l}_{ijk}$ loans with officers serving market i from region k reduces to

$$A_{ijk}^{-1}\,\tilde{l}_{ijk}^{1+1/\varrho}, \qquad A_{ijk} \propto \frac{e_{ijk}}{\Psi_k},$$

where Ψ_k is a region-specific wage index summarizing outside-option wages in region k. The composite parameter A_{ijk} summarizes regional labor cost: high outside-option wages drive Ψ_k up and A_{ijk} down, making screening labor in that region expensive on a per-dollar basis.

4.5 Solution to the Lender's Problem

With the demand system $s_{ij}(r_{ij})$ and labor supply system $l_{ijk}(\mathbf{w})$ in hand, we now characterize the lender's optimal pricing and assignment decisions. We use one result from Section 4.4: the cost of processing $\tilde{l}_{ijk}$ loans with officers in region k reduces to $A_{ijk}^{-1}\tilde{l}_{ijk}^{1+1/\varrho}$, where $A_{ijk} \propto e_{ijk}/\Psi_k$ embeds processing capacity and the regional wage index Ψ_k. The lender's problem is separable across markets, so we analyze each market i independently.

Optimal Pricing. The first-order condition with respect to r_{ij} is a fixed point:

$$\sum_{k\in\mathcal{K}} A_{ijk}^{\varrho} p_{ijk}^{\varrho+1}(r_{ij}-f_{ij}-d_{ijk})^{\varrho}\Big\{ r_{ij}(\varepsilon_{p_{ijk},r}+\varepsilon_{s_{ij},r}+1) - f_{ij}(\varepsilon_{p_{ijk},r}+\varepsilon_{s_{ij},r}) \\ - d_{ijk}(\varepsilon_{p_{ijk},r}+\varepsilon_{s_{ij},r}+\varepsilon_{d_{ijk},r}) - \varepsilon_{s_{ij},r}\bar{c}_{ij}(r_{ij}-f_{ij}-d_{ijk})\Big\} = 0, \quad (10)$$

where $\varepsilon_{x,y} = \partial \ln x/\partial \ln y$ and the capacity cost shifter is

$$\bar{c}_{ij} := \left(\frac{\mathcal{M}_i s_{ij}}{\sum_{k\in\mathcal{K}}\left(A_{ijk}p_{ijk}(r_{ij}-f_{ij}-d_{ijk})\right)^{\varrho}}\right)^{1/\varrho}.$$

The economic content is sharpest when screening precision is shut down. If σ_{ik}^2 is constant across k, then $p_{ijk} = \bar{p}_{ij}$ and $d_{ijk} = \bar{d}_{ij}$, and (10) collapses to a closed-form decomposition of

implied bank-market labor supply elasticities nearly unchanged. See Appendix D for details.

the optimal rate:

$$r_{ij} = \underbrace{f_{ij}\frac{\varepsilon_{p_{ij},r} + \varepsilon_{s_{ij},r}}{\varepsilon_{p_{ij},r} + \varepsilon_{s_{ij},r} + 1}}_{\text{Funding cost}} + \underbrace{d_{ij}\frac{\varepsilon_{p_{ij},r} + \varepsilon_{s_{ij},r} + \varepsilon_{d_{ij},r}}{\varepsilon_{p_{ij},r} + \varepsilon_{s_{ij},r} + 1}}_{\text{Default risk}} + \underbrace{p_{ij}^{-1}\left(\frac{\mathcal{M}_i s_{ij}}{\sum_{k\in\mathcal{K}} A_{ijk}^{\varrho}}\right)^{1/\varrho}\frac{\varepsilon_{s_{ij},r}}{\varepsilon_{p_{ij},r} + \varepsilon_{s_{ij},r} + 1}}_{\text{Capacity cost}}. \tag{11}$$

Funding costs pass through weighted by the combined elasticity of origination and market share. Default losses are priced in with an extra adverse-selection wedge $\varepsilon_{d_{ij},r}$, reflecting that higher rates worsen the applicant pool. The capacity component prices the marginal cost of screening labor and is paid for both originated and rejected applications, since screening is sunk at the application stage.

The general case (10) differs from this benchmark in one substantive way: elasticities and net margins are aggregated across heterogeneous officer types, with weights that depend on signal quality. This aggregation is how lenders internalize the differential value of precise information across regions—the *private value of information*.

Optimal Assignment. The first-order condition with respect to $\mathcal{S}_{ijk}$ delivers the assignment rule:

$$\mathcal{S}_{ijk} = \frac{A_{ijk}^{\varrho}\, p_{ijk}^{\varrho}\, (r_{ij} - f_{ij} - d_{ijk})^{\varrho}}{\sum_{m\in\mathcal{K}} A_{ijm}^{\varrho}\, p_{ijm}^{\varrho}\, (r_{ij} - f_{ij} - d_{ijm})^{\varrho}}. \tag{12}$$

Equation (12) is the private value of information made explicit. Each application is routed in proportion to the expected gross profit per application, $p_{ijk}(r_{ij} - f_{ij} - d_{ijk})$, scaled by screening capacity per dollar, A_{ijk}. The first factor rewards precision: nearby officers with low σ_{ik}^2 generate more accurate signals, raising p_{ijk} and lowering d_{ijk}, hence raising expected profit per application. The second factor penalizes cost: in regions where outside options pay high wages, Ψ_k is high and A_{ijk} is low, so each dollar buys less screening capacity. The lender substitutes away from precise but expensive local labor toward cheaper but noisier remote labor whenever the wage gap dominates the precision gap.

5 Estimation

We now take the model to the data. There are two purposes: to discipline the model's novel components—the precision and processing efficiency of local versus remote loan officers—using the empirical moments documented in Section 3.1, and to recover the borrower-side primitives that govern adverse selection and default. Some parameters are calibrated externally to standard values from the literature; the rest are estimated internally by matching a set of empirical moments that we argue identify them. We begin with the parametric assumptions used to take the model to data, then describe each estimation step, and conclude by validating the estimated model against untargeted heterogeneity in the data.

5.1 Parameterization

To take the model to data, we make the following parametric assumptions. Borrower types (θ_η, x_η) are jointly normally distributed,

$$\begin{pmatrix} \theta_\eta \\ x_\eta \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_\theta \\ \mu_i^x \end{pmatrix}, \begin{pmatrix} \sigma_\theta^2 & \gamma\sigma_\theta\sigma_i^x \\ \gamma\sigma_\theta\sigma_i^x & (\sigma_i^x)^2 \end{pmatrix} \right).$$

The demand shocks $\epsilon_{ij\eta}$ are i.i.d. Type I extreme value, and screening noise is normally distributed, $\tilde{x}_{\eta k} \sim \mathcal{N}(0, \sigma_{ik}^2)$. We assume signal precision and processing efficiency depend on whether a loan officer is local or remote: $\sigma_{ik} = \sigma_{\text{local}}$ and $e_{ijk} = e_{ij,\text{local}}$ if $g(i) = k$, and $\sigma_{ik} = \sigma_{\text{remote}}$ and $e_{ijk} = e_{ij,\text{remote}}$ otherwise.[7] Finally, d_i follows a probit process with correlated shocks: borrower η defaults if $\epsilon_\eta + \epsilon_{it} \le x_\eta$, where ϵ_η and ϵ_{it} are i.i.d. standard normal idiosyncratic and aggregate shocks, respectively.[8]

[7]It is straightforward to show that this implies bank j will offer the same wage to all remote workers processing applications in market i, regardless of where they live. Further, due to max-stability, remote loan officer supply acts as-if it follows a single "remote" market. We abstract from further heterogeneity in remote informational and processing efficiencies to map directly to our empirical setting and avoid overparameterizing the model.

[8]The inclusion of aggregate shocks is motivated by correlated default rate spikes across credit grades and loan purposes, as shown in Appendix Figure A1.

5.2 Estimation Strategy

The estimation proceeds in three steps. We first calibrate parameters with widely accepted values from prior work; we then estimate the remaining parameters internally by matching a set of empirical moments; finally, we describe the identification of the novel loan officer parameters that are central to the paper.

Externally Calibrated Parameters. We calibrate the logit price coefficients to match Buchak et al. (2018), giving $\mu_\theta = 165$ and $\sigma_\theta \approx 41$. The within-region-sector labor supply elasticity $\varrho = 8$ matches Giroud et al. (2024). Funding costs are set to 2.75% to match the spreads reported in Janus Henderson Investors (2019). The market size $\mathcal{M}_i$ equals the total observed application volume in market i. To account for the fact that there are significant recoveries on defaulted loans, we calibrate losses-given-default to 17.7% using Fannie Mae securitization data.[9]

Internally Estimated Parameters. The remaining parameters govern objects for which off-the-shelf estimates are unavailable: the borrower type distributions, the local-vs-remote screening technology, and the bank-specific demand and labor-cost shifters. We estimate these jointly by matching moments from the empirical analysis. We consider six product type buckets, $\{\text{subprime}, \text{near-prime}, \text{prime}\} \times \{\text{purchase}, \text{refi}\}$, and three bank types—"Brick and Mortar," "Fintech," and "Fringe"—mapping to the categories in Table A3.[10] The labor supply efficiency parameters $\{A_{i,j,k}\}$ and demand shifters $\{\xi_{ij}\}$ are estimated to exactly match observed labor shares and market shares using standard techniques from the IO and spatial economics literatures. We then jointly estimate the loan officer signal precisions $\{\sigma_{\text{local}}, \sigma_{\text{remote}}\}$, the correlation between default risk and price sensitivity γ, the borrower type distributions $\{\mu_i^x, \sigma_i^x\}$, and the realizations of the aggregate shocks $\epsilon_{i,t}$ to match the expected default rates, realized default rates, rejection rates, the local origination and default advantages, and the strength of the ex-post correlation between default and interest rates reported in Tables 2 and A3 as well as Figures 3 and A1.

[9] Specifically, we take the realized average loss rate on all loans originated in 2000 and later over 10 years, for which there were 10 years of history available in 2018, after two missed payments in the first two years.

[10] To map model primitives to available data, we make two additional measurement assumptions. First, for simplicity, we assume demand shifters $\{\xi_{ij}\}_{\forall i,j}$ vary only at the bank-type level. Second, we assume that banks' expected default rates correspond to post-crisis historical average default rates (2012–2017).

Identification of Critical Parameters. A natural concern is whether the novel loan officer parameters $\{A_{i,j,k}\}$ and $\{\sigma_k\}$ are jointly pinned down by the data. The signal precisions are identified by the local rejection rate and default advantages: holding fixed the default advantage of local screening, the rejection rate advantage falls as baseline noise rises, because higher noise pushes the lender toward a pooling equilibrium in which all loans are originated and improvements to precision reduce pooled originations rather than credit rationing. Figure 5 provides a numerical example illustrating this non-monotonicity, showing how the ratio of the rejection advantage to the default advantage changes with precision. Given the level of imprecision, the magnitude of the default advantage then pins down the overall informational advantage of local loan officers. The $\{A_{i,j,k}\}$ parameters capture the wage-adjusted physical efficiency of officers for each bank-market-type cell: their relative magnitudes are pinned down by the local share each bank hires, and their absolute level by the lender's first-order condition (10). Intuitively, a bank with higher physical efficiency for officer type k hires more officers of that type, internalizing the impact on its expected profits.

5.3 Estimation Results

We report the estimated values of the borrower type distributions and the aggregate shocks in Table 8. For ease of interpretation, we transform the mean of the latent x type into a "default rate" that captures what the expected per-year default rate would be in the relevant mortgage market if 100 percent of applications were approved.[11] The implied unconditional default rates and latent heterogeneity are substantially higher in the refi markets; the model requires this to jointly rationalize the combination of low realized default rates and low approval rates observed in those markets. Because the model is exactly identified, the estimated parameters reproduce the targeted moments throughout the paper. Two results are particularly noteworthy: the joint estimation yields $\gamma = -0.4$, consistent with adverse selection, and signal precisions of $\sigma_{\text{local}} = 1.4$ and $\sigma_{\text{remote}} = 2.45$, consistent with a substantial local informational advantage.

[11] That is, the reported default rate is equal to $\int d_i(x_\eta) F_i^x(dx_\eta) \equiv \int_{-\infty}^{\infty} \Phi\left(2^{-1/2} x_\eta\right) f_i^x(x_\eta)\, dx_\eta$ for a standard normal CDF Φ.

5.4 Model Validation

For simplicity, the parameterization above assumes that the information technology is the same in every market. This restriction leaves untargeted heterogeneity in effect sizes available as a tool to validate the model. To do so, we compute model-implied, market-specific rejection rate and default rate advantages and compare them to the data. The model generates a local rejection rate advantage about 3 percentage points higher in the refi market than the purchase market, and a local default rate advantage about 0.4 percentage points higher in the purchase market than the refi market—roughly consistent with the magnitudes in Table 2. The local rejection rate advantage (about 3 percentage points) and default rate advantage (about 1 percentage point) are also much larger for subprime loans than for prime loans, consistent with Figure 4, although the relevant coefficients are estimated noisily. These patterns are driven entirely by differences in borrower type distributions and aggregate shocks, which the model backs out from the observed default and origination data without targeting the heterogeneity itself.

6 Model Results and Counterfactuals

In this section, we use the model to quantify baseline informational efficiency and to study how shocks to physical efficiency alter equilibrium outcomes. We begin by showing that superior information is used primarily to offset credit rationing in equilibrium. We then consider a counterfactual increase in the physical efficiency of distant loan officers and show that this reduces informational efficiency and increases the aggregate risk borne by banks.

6.1 Credit Rationing and Pooled Originations

Lenders with more precise signals can use their information to (i) screen in creditworthy borrowers (reducing "credit rationing"), and (ii) screen out high–risk applicants (reducing "pooled origination"). The model in Section 4 provides quantitative magnitudes for these forces in equilibrium.

Consider a lender j in market i that hires a perfectly informed loan officer with $\sigma_k = 0$. This lender originates a loan if and only if $E[d(x_\eta)] \leq r_{ij} - f_{ij}$, where the borrower's risk type x_η is perfectly observed. Intuitively, an application is approved if and only if its full–information expected loss rate is below the lender's net interest margin. Credit rationing occurs when a low–loss loan is rejected due to noise in the signal—a "false" rejection. Pooled origination is the symmetric "false" approval driven by an unusually favorable signal realization for a high–loss borrower.

Table 9 reports the model–implied frequency of credit rationing and pooled origination by market. Consistent with the large rejection–rate advantages, relative to default–rate advantages, documented in Table 2, the model implies substantial credit rationing, with up to 15 percent of applicants rationed in the subprime refinance market. Panel B shows that local officers eliminate roughly *half* of this rationing, with the largest effects in lower–credit–quality and refinance markets.

Panel B shows that local officers achieve this expansion in credit while also *reducing* pooled origination, consistent with their default advantage. Because default risk in the model is highly convex, lenders behave conservatively: pooled origination is quantitatively small relative to rationing. Thus, informational advantages operate mainly through reductions in credit rationing.

6.2 Technology Shock

Having established that informational efficiency plays an important role in shaping credit access—especially for riskier borrowers—we now study the effects of a technological shock. We consider an increase in the labor efficiency e_{ijk} of distant loan officers that does not affect their informational precision. We calibrate this shock as the difference in physical efficiency between the "Fintech" and "Brick and Mortar" banks and consider two counterfactuals.[12]

In counterfactual 1, only the Brick and Mortar bank receives the efficiency improvement. In counterfactual 2, all banks receive the same efficiency improvement. After applying each shock, we recompute the equilibrium, allowing banks to re–optimize their labor inputs and interest rates. Consistent with equation (12), shocked banks substitute distant labor for local labor.

[12]Given the structure in Section 4, the difference in physical efficiency is identified from the ratio of the A_{ijk} parameters for the two banks. While the shock varies across markets, it corresponds to roughly a twenty percent increase in physical productivity.

Credit Provision. Table 10 presents the results. Panel A reports outcomes when only the Brick and Mortar bank is shocked; Panel B reports outcomes when all banks are shocked. For shock 1, rejection rates sometimes decline, but this is driven by a larger increase in pooled origination offsetting a sizable rise in rationing. Across specifications, credit rationing increases meaningfully, including increases of up to two percentage points for subprime purchase borrowers.

Expected Defaults and Rates. The increase in pooled origination naturally raises expected defaults. Table 10, Panel C reports the results. Default rates rise across all markets under both shocks, with particularly strong effects in the subprime and near–prime purchase segments—the same segments experiencing the largest increases in pooled origination.

Do borrowers nonetheless benefit? As with standard productivity shocks, interest rates fall in equilibrium. Panel C also reports changes in monthly payments.[13] Effects are modest: at most $25 per month. Screening is already physically efficient ex ante, so moderate improvements in the productivity of distant officers generate limited cost savings.

Interestingly, the largest increase in defaults arises when only the Brick and Mortar bank is shocked, while the aggregate shock produces larger cost reductions. Two forces drive this. First, substitution away from local labor is largest for the Brick and Mortar bank, since other banks employ little local labor initially. Second, the decline in informational efficiency restricts the Brick and Mortar bank's ability to lower rates: higher default and rejection rates increase its marginal cost. By contrast, shocking the fintech and fringe banks acts more like a standard TFP shock, with lower costs and relatively little substitution.

Aggregate Shocks. We next examine how the technology shock interacts with aggregate conditions. The model allows us to compute realized default rates across aggregate states. Figure 6 plots the difference in realized default rates between the baseline and shocked economies as a function of the aggregate state for the purchase markets.[14]

The effect is non–monotonic. For small aggregate shocks, pooled–origination loans expe-

[13] Payments are computed for a median–priced home of $330,000 with 20 percent down and a 30–year fixed–rate mortgage.

[14] For refinancing markets, informational advantages are smaller and can be offset by smaller aggregate shocks. This is consistent with the model's prediction that the refinance market should have a relatively small default advantage, despite its large rejection rate advantage.

rience rapid increases in default risk, raising realized defaults. For sufficiently large aggregate shocks, however, default rates for these marginal loans exceed the point of inflection of the probit curve. At this point, loans that were rationed, which tend to have significantly higher baseline expected default rates than always-approved loans, start to see their default probabilities increase rapidly, leading to the attenuation or even reversal of the information advantage. In this sense, credit rationing seems to serve as a moderate protective force against very large aggregate shocks.

Given the short time series we have to calculate default rates, we cannot scale the aggregate shocks to relative frequencies. However, we can benchmark these aggregate shocks against realized default rates during the GFC. We include vertical lines in Figure 6 to show the size aggregate shock that matches the 2007 realized default rates in each market. In the subprime and near–prime markets, the informational advantage persists even under shocks larger than the GFC. In the prime market, informational effects are smaller and eliminated by a shock roughly the size of the GFC. Quantitatively, large aggregate shocks like the GFC can roughly double the effect on realized defaults.

Economic Intuition. The dominant effect of the technology shock is the increase in credit rationing. The intuition is straightforward. While banks internalize the cost of higher defaults and borrowers are rate–sensitive, credit rationing is largely an externality from the bank's perspective. Rationed loans lie close to the break–even point on net interest income, so rejecting them sacrifices relatively little revenue. On the other hand, because labor supply is highly elastic, even moderate relative cost shocks—such as improvements in distant–officer productivity—induce sizable reallocations of labor away from local officers. These reallocations reduce informational efficiency and increase default risk without providing commensurate reductions in mortgage rates.

7 Conclusion

In this paper, we study how changes in technology, which have enabled loan officers to originate loans at a distance, have impacted physical and informational efficiency in the mortgage market. We show that, while there is an increase in physical efficiency, this comes at a significant informational cost, as banks take advantage of the opportunity to hire workers in lower wage areas. This generates an externality on would-be, otherwise credit-worthy borrowers, who end up rationed in equilibrium. Local lending officers can strongly mitigate these effects, but increases in efficiency of distant officers can worsen these effects.

References

Adams, William, Liran Einav, and Jonathan Levin, 2009, Liquidity constraints and imperfect information in subprime lending, *American Economic Review* 99, 49–84.

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

Agarwal, Sumit, and Robert Hauswald, 2010, Distance and private information in lending, *The Review of Financial Studies* 23, 2757–2788.

Aguirregabiria, Victor, Robert Clark, and Hui Wang, 2025, The geographic flow of bank funding and access to credit: Branch networks, synergies, and local competition, *American Economic Review* 115, 1818–1856.

Allen, Jason, Robert Clark, and Jean-François Houde, 2014, The effect of mergers in search markets: Evidence from the canadian mortgage industry, *American Economic Review* 104, 3365–3396.

Allen, Jason, Robert Clark, and Jean-François Houde, 2019, Search frictions and market power in negotiated-price markets, *Journal of Political Economy* 127, 1550–1598.

Allen, Jason, Robert Clark, Jean-François Houde, Shaoteng Li, and Anna Trubnikova, 2025, The role of intermediaries in selection markets: Evidence from mortgage lending, *The Review of Financial Studies* 38, 3284–3328.

Antràs, Pol, Luis Garicano, and Esteban Rossi-Hansberg, 2006, Offshoring in a knowledge economy*, *The Quarterly Journal of Economics* 121, 31–77.

Arkolakis, Costas, Federico Huneeus, and Yuhei Miyauchi, 2025, Production network formation, trade, and welfare, Technical report, CESifo Working Paper.

Benetton, Matteo, 2021, Leverage regulation and market structure: A structural model of the uk mortgage market, *The Journal of Finance* 76, 2997–3053.

Berg, Tobias, Valentin Burg, Ana Gombović, and Manju Puri, 2020, On the rise of fintechs: Credit scoring using digital footprints, *The Review of Financial Studies* 33, 2845–2897.

Berger, Allen N, Nathan H Miller, Mitchell A Petersen, Raghuram G Rajan, and Jeremy C Stein, 2005, Does function follow organizational form? evidence from the lending practices of large and small banks, *Journal of Financial economics* 76, 237–269.

Berger, Allen N, and Gregory F Udell, 1995, Relationship lending and lines of credit in small firm finance, *Journal of business* 351–381.

Bernard, Andrew B, Andreas Moxnes, and Yukiko U Saito, 2019, Production networks, geography, and firm performance, *Journal of Political Economy* 127, 639–688.

Bhutta, Neil, and Aurel Hizmo, 2021, Do minorities pay more for mortgages?, *The Review of Financial Studies* 34, 763–789.

Blattner, Laura, Jacob Hartwig, and Scott Nelson, 2022, Information design in consumer credit markets, *American Economic Review* .

Blattner, Laura, and Scott Nelson, 2021, How costly is noise? data and disparities in consumer credit, *arXiv preprint arXiv:2105.07554* .

Blattner, Laura, Scott Nelson, and Jann Spiess, 2021, Unpacking the black box: Regulating algorithmic decisions, *arXiv preprint arXiv:2110.03443* .

Boehm, Christoph E, Aaron Flaaen, and Nitya Pandalai-Nayar, 2019, Input linkages and the transmission of shocks: Firm-level evidence from the 2011 tōhoku earthquake, *Review of Economics and Statistics* 101, 60–75.

Bosshardt, Joshua, Ali Kakhbod, and Amir Kermani, 2025, Do intermediaries improve gse lending? evidence from proprietary gse data, *Journal of Financial Economics* 170, 104082.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru, 2018, Beyond the balance sheet model of banking: Implications for bank regulation and monetary policy, Technical report, National Bureau of Economic Research.

Bureau of Labor Statistics, 2023, Occupational employment and wages, may 2023: 13-2072 loan officers, https://www.bls.gov/oes/2023/may/oes132072.htm, Employment estimate: 321,090; last modified April 3 2024.

Cherry, Susan, 2025, Regulating credit: The impact of price regulations and lender technologies on financial inclusion, *Available at SSRN 5229936* .

Costinot, Arnaud, Jonathan Vogel, and Su Wang, 2012, An elementary theory of global supply chains, *The Review of Economic Studies* 80, 109–144.

Crawford, Gregory S, Nicola Pavanini, and Fabiano Schivardi, 2018, Asymmetric information and imperfect competition in lending markets, *American Economic Review* 108, 1659–1701.

DeFusco, Anthony A, Huan Tang, and Constantine Yannelis, 2022, Measuring the welfare cost of asymmetric information in consumer credit markets, *Journal of Financial Economics* 146, 821–840.

Diamond, Douglas W, 1984, Financial intermediation and delegated monitoring, *The review of economic studies* 51, 393–414.

D'Amico, Leonardo, and Maxim Alekseev, 2024, Capital market integration and growth across the united states, Technical report, Working Paper.

Egan, Mark, Ali Hortaçsu, and Gregor Matvos, 2017, Deposit competition and financial fragility: Evidence from the us banking sector, *American Economic Review* 107, 169–216.

Einav, Liran, Mark Jenkins, and Jonathan Levin, 2012, Contract pricing in consumer credit markets, *Econometrica* 80, 1387–1432.

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther, 2022, Predictably unequal? the effects of machine learning on credit markets, *The Journal of Finance* 77, 5–47.

Giroud, Xavier, Simone Lenzu, Quinn Maingi, and Holger Mueller, 2024, Propagation and amplification of local productivity spillovers, *Econometrica* 92, 1589–1619.

Huang, Ruidi, Erik J Mayer, Sheridan Titman, and David X Xu, 2025, Human capital and local credit supply: Evidence from the mortgage industry, *SMU Cox School of Business Research Paper* .

Iyer, Rajkamal, Asim Ijaz Khwaja, Erzo FP Luttmer, and Kelly Shue, 2016, Screening peers softly: Inferring the quality of small borrowers, *Management Science* 62, 1554–1577.

Jaffee, Dwight M, and Franco Modigliani, 1969, A theory and test of credit rationing, *The American Economic Review* 59, 850–872.

Janus Henderson Investors, 2019, Mortgage-backed securities: Understanding the option-adjusted spread, Technical report, Janus Henderson Investors, White paper.

Ji, Yan, Songyuan Teng, and Robert M Townsend, 2023, Dynamic bank expansion: spatial growth, financial access, and inequality, *Journal of Political Economy* 131, 2209–2275.

Jiang, Erica Xuewei, 2023, Financing competitors: Shadow banks' funding and mortgage market competition, *The Review of Financial Studies* 36, 3861–3905.

J.P. Morgan Chase & Co., 2025, Chase closing guarantee – guaranteed on-time closing or get $5,000, https://www.chase.com/personal/mortgage/closing-guarantee, Accessed: 2025-11-23.

Leland, Hayne E, and David H Pyle, 1977, Informational asymmetries, financial structure, and financial intermediation, *The journal of Finance* 32, 371–387.

Liberti, Jose M, and Atif R Mian, 2008, Estimating the effect of hierarchies on information use, *The Review of Financial Studies* 22, 4057–4090.

Maingi, Quinn, 2026, Regional banks, aggregate effects, *Journal of Financial Economics* 176, 104226.

Morelli, Juan M, Matias Moretti, and Venky Venkateswaran, 2025, Geographic funding risk and market power in deposit markets .

Nelson, Scott T, 2025, Private information and price regulation in the us credit card market, *Econometrica* 93, 1371–1410.

Oberfield, Ezra, Esteban Rossi-Hansberg, Pierre-Daniel Sarte, and Nicholas Trachter, 2024, Plants in space, *Journal of Political Economy* 132, 867–909.

Petersen, Mitchell A, and Raghuram G Rajan, 1994, The benefits of lending relationships: Evidence from small business data, *The journal of finance* 49, 3–37.

Petersen, Mitchell A, and Raghuram G Rajan, 2002, Does distance still matter? the information revolution in small business lending, *The journal of Finance* 57, 2533–2570.

Ramakrishnan, Ram TS, and Anjan V Thakor, 1984, Information reliability and a theory of financial intermediation, *The Review of Economic Studies* 51, 415–432.

Ramondo, Natalia, and Andrés Rodríguez-Clare, 2013, Trade, multinational production, and the gains from openness, *Journal of Political Economy* 121, 273–322.

Rosen, Richard J, 2011, Competition in mortgage markets: The effect of lender type on loan characteristics, *Economic Perspectives* 35.

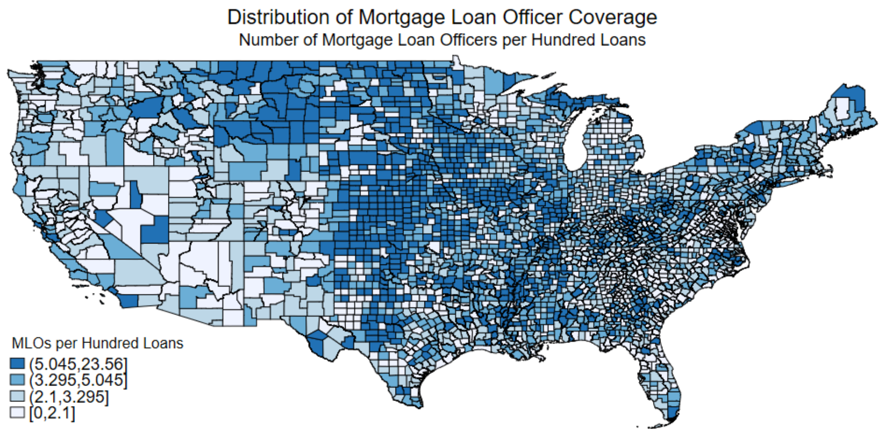
Stiglitz, Joseph E, and Andrew Weiss, 1981, Credit rationing in markets with imperfect information, *The American economic review* 71, 393–410.

Wang, Lulu, 2025, Regulating competing payment networks, *Kilts Center at Chicago Booth Marketing Data Center Paper* .

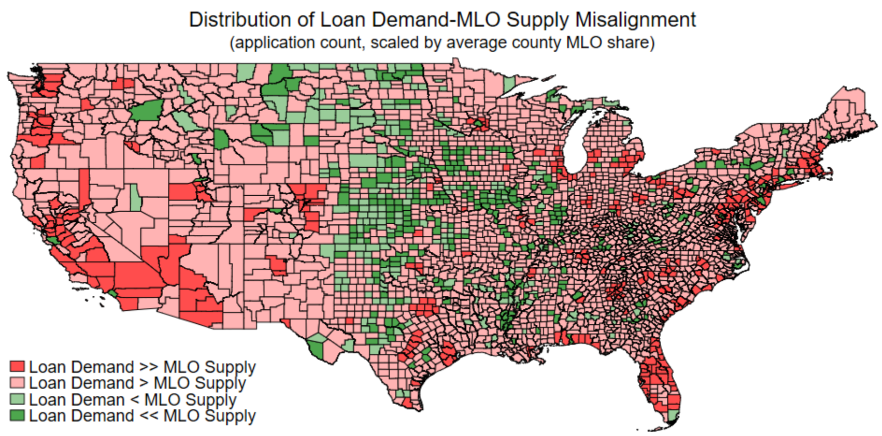
Whited, Toni M, Yufeng Wu, and Kairong Xiao, 2022, Central bank digital currency and banks, *Available at SSRN 4112644* .

Xiao, Kairong, 2020, Monetary transmission through shadow banks, *The Review of Financial Studies* 33, 2379–2420.

Figure 1. Geographic Misalignment of Mortgage Demand and Loan Officer Supply



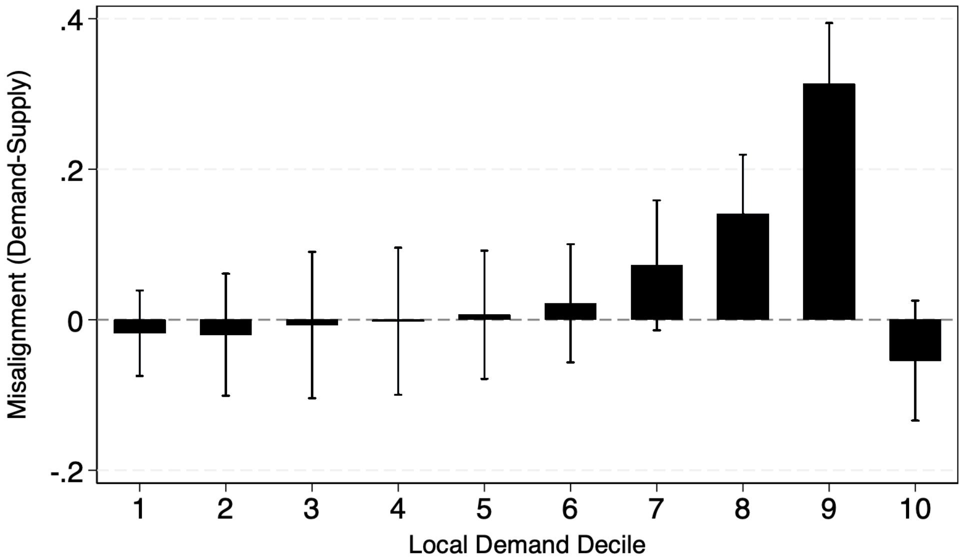
(a) Geographic Distribution of Mortgage Loan Officer Coverage



(b) Geographic Misalignment

Note: Panel A plots the county-level number of loan officers per hundred mortgage applications during 2018-2019, divided into 4 equal-sized buckets and highlighted with different color intensity. Panel B plots the county-level misalignment of mortgage demand and loan officer supply during 2018-2019. For every county, we calculate the county's share of mortgage applications out of national mortgage applications; similarly, we calculate the county's share of registered mortgage loan officers out of national total number of registered mortgage loan officers. Based on these shares, county-level misalignment index is computed as the difference between the county's share of loan demand and its share of mortgage loan officers. The index is then divided into four buckets, highlighted with two different colors in the figure. Red color marks counties with a positive value of the misalignment index, indicating that its loan demand exceeds its loan officer supply. Green color marks counties with a negative value of the misalignment index, indicating that its loan demand is below its loan officer supply.

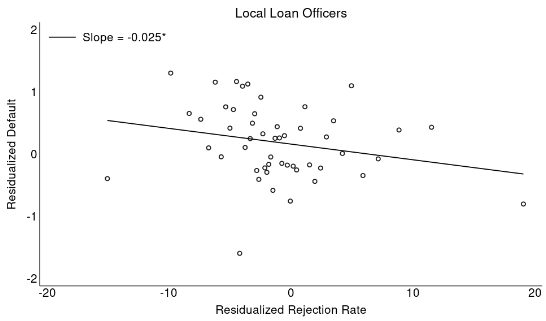
Figure 2. Loan Demand-Loan Officer Supply Misalignment Across Counties



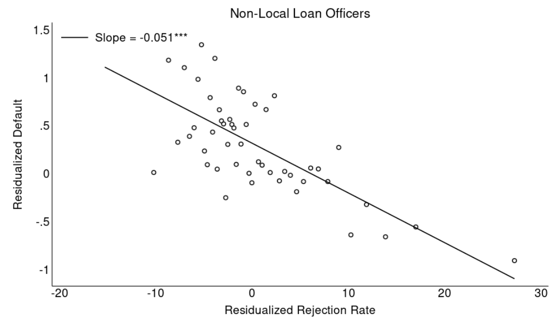
(a) Excess Loan Demand Relative to Loan Officer Supply

Note: This figure plots the average county-level misalignment between mortgage demand and loan officer supply across county mortgage demand deciles for 2018–2019. For each county, we compute (i) the county's share of national mortgage applications and (ii) the county's share of registered mortgage loan officers; the misalignment index is defined as the difference between these two shares. Counties are then grouped into ten bins based on their share of national mortgage applications (from lowest to highest demand). The bars display the mean misalignment index within each demand decile, and the capped spikes depict 95% confidence intervals from a regression of the misalignment index on the ten demand-decile indicators (with no constant term). A value above zero indicates that mortgage demand exceeds local loan officer supply ("labor-short" areas), while a value below zero indicates an oversupply of underwriting labor relative to local demand.

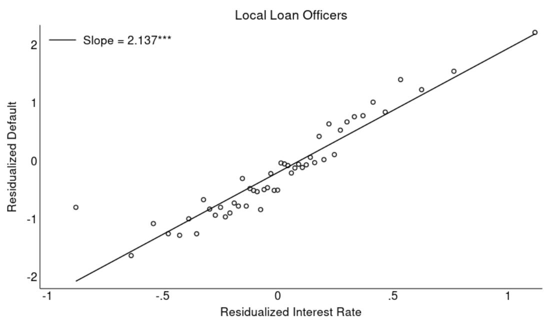
Figure 3. Advantage of Local Information Production



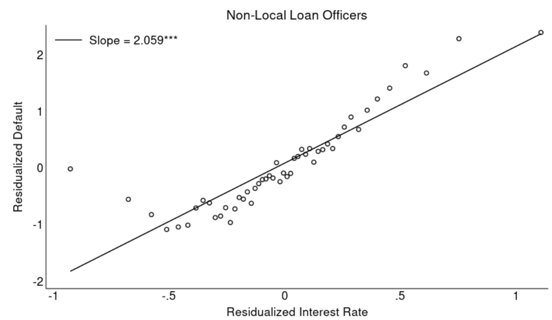
(a) Rejection: Local Loan Officer



(b) Rejection: Non-Local Loan Officer



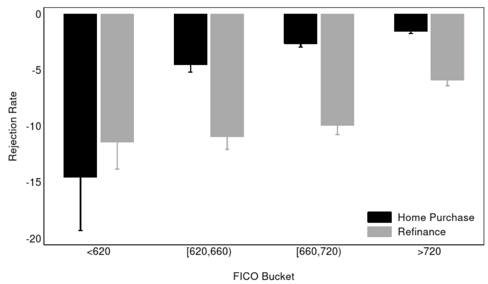
(c) Rate: Local Loan Officer



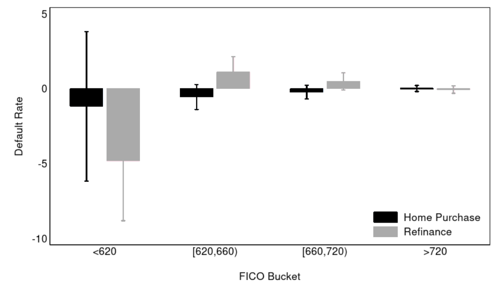
(d) Rate: Non-Local Loan Officer

Note: This figure compares local and non-local loan officers along the approval and pricing margins. Panels A and B plot residualized default against residualized rejection rates for loans handled by local and non-local loan officers, respectively. Panels C and D plot residualized default against residualized interest rates for originated home-purchase loans handled by local and non-local loan officers, respectively. Residualized rejection rates, interest rates, and default outcomes are obtained by regressing the corresponding loan-level variables on borrower-level risk controls, including FICO, LTV, DTI, AUS status, and their flexible polynomials, as well as county-by-application-month fixed effects. Residuals are then averaged at the lender–county–local/non-local level, so that each underlying observation represents the mean residualized outcome for a given lender in a given county, separately for local and non-local officers. For visualization, these lender–county averages are sorted into 50 equally sized bins within each subsample, and each point plots the mean default residual against the mean rejection-rate or interest-rate residual within a bin. The fitted line shows the linear relationship estimated separately in each panel.

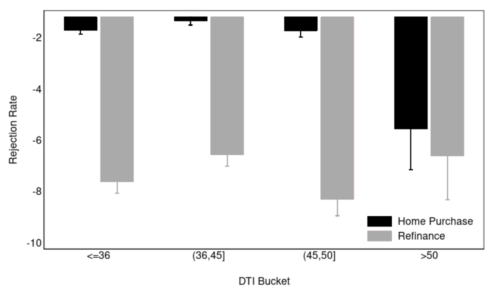
Figure 4. Rejection and Default by Hard Credit Information: Local vs Non-Local



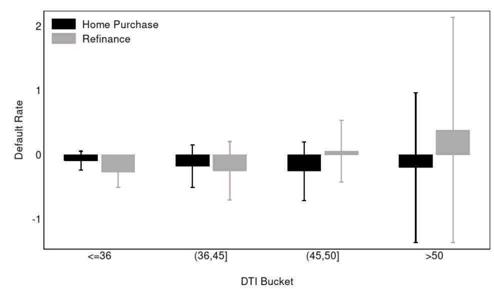
(a) Rejection-FICO



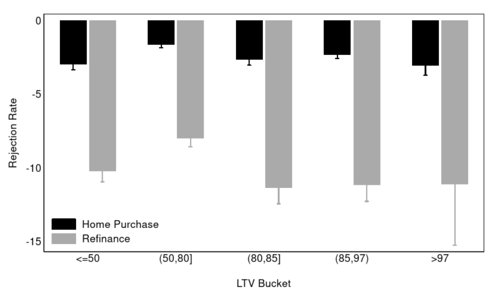
(b) Default-FICO



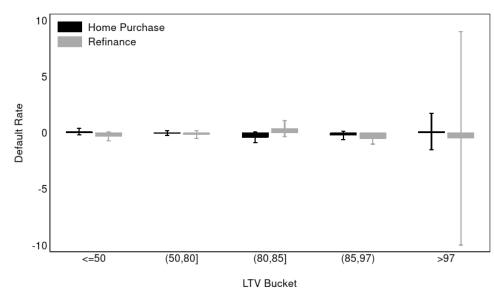
(c) Rejection-DTI



(d) Default-DTI



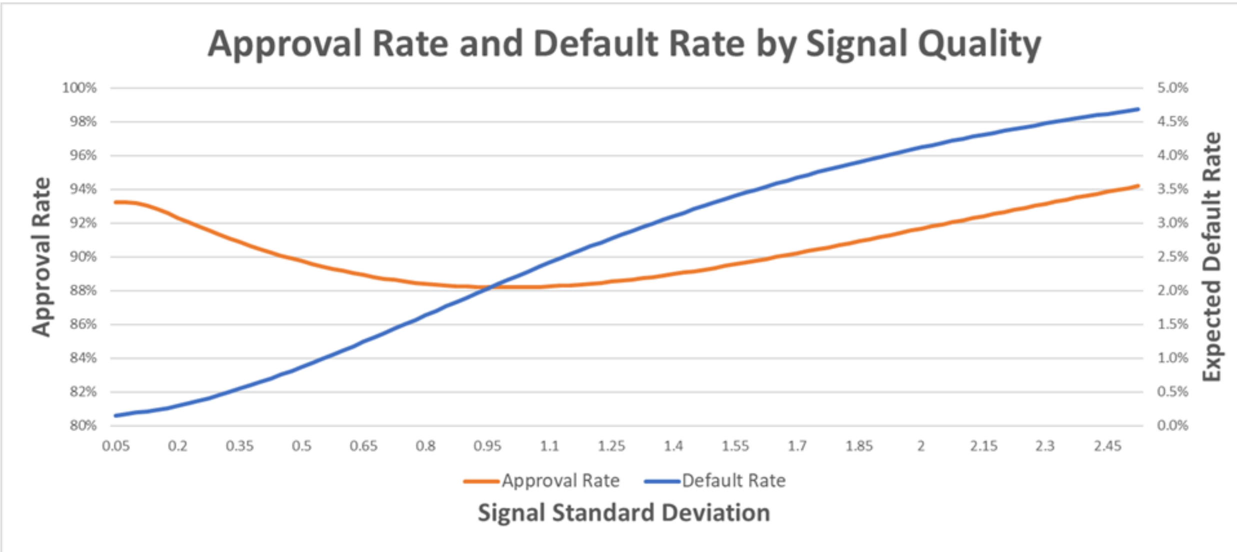
(e) Rejection-LTV



(f) Default-LTV

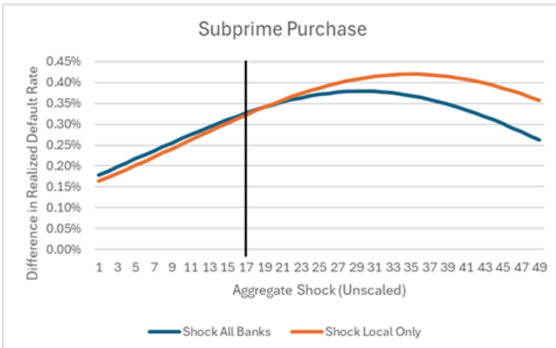
Note: This figure compares rejection rates (Panels A, C, and E) and two-year default rates (Panels B, D, and F) for home-purchase and refinance mortgages across discrete buckets of borrower hard information: FICO score (Panels A–B), debt-to-income (DTI) ratio (Panels C–D), and loan-to-value (LTV) ratio (Panels E–F). For each credit variable, the sample is divided into regulatory or industry-relevant buckets: four FICO groups (< 620, $[620, 660)$, $[660, 720)$, ≥ 720), four DTI groups ($\leq 36\%$, $(36, 45]\%$, $(45, 50]\%$, $> 50\%$), and five LTV groups ($\leq 50\%$, $(50, 80]\%$, $(80, 85]\%$, $(85, 97]\%$, $> 97\%$). Within each bucket, we estimate separate regressions of rejection (or default) on indicators for loans handled by local versus non-local loan officers. The plotted coefficients represent the average rejection (or default) rate for loans handled by local officers in each bucket, with 95% confidence intervals; analogous estimates for refinance loans are shown side-by-side for comparison. Standard errors are clustered at the borrower-county level.

Figure 5. Identification Intuition

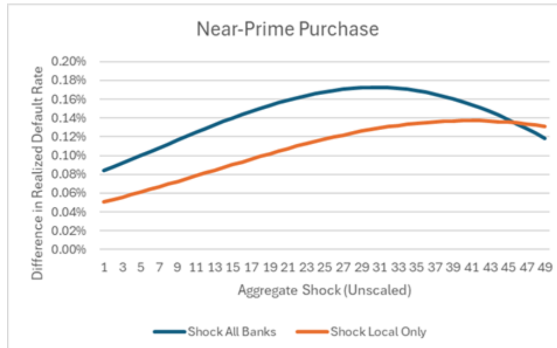


Note: This figure shows how the approval rate and default rate predicted by the model changes with the level of precision with which the loans are processed. For this figure, we assume an interest rate of 5%, a funding cost of 2.75%, a population average default rate of 5.5%, and an LGD of 17%

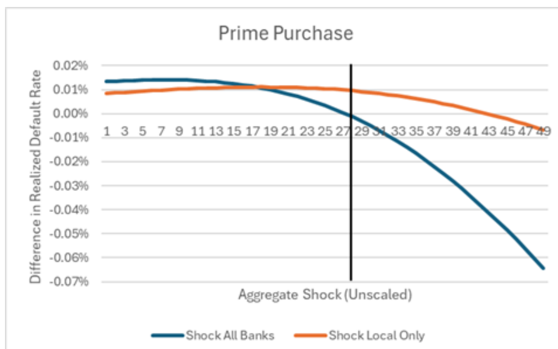
Figure 6. Technology Shock and Aggregate Risk



(a) Subprime Purchase Market



(b) Near-Prime Purchase Market



(c) Prime Purchase Market

Note: This figure shows how realized default rates change, given a technological shock, for different realizations of the aggregate state. Credit markets are differentiated by borrower credit quality. The black line indicates the value of the aggregate shock that generates a default rate equivalent to the Global Financial Crisis in 2007. See Section 6.2 for more details.

Table 1: Summary Statistics

This table reports summary statistics. Panel A reports statistics for the full application sample. Panel B reports statistics for originated loans. Panel C reports statistics for the HMDA–McDash matched sample used to study ex-post loan performance. Continuous variables are winsorized at the 1st and 99th percentiles. Local share, rejection rate, default rate, and interest rate are reported in percentage points. Local share is based on an indicator equal to one if the reviewing loan officer is located in the same county as the borrower.

	Home Purchase			Refinance		
	Mean	Median	SD	Mean	Median	SD
Panel A: Full Application Sample						
Local Share (%)	40.79	-	-	31.14	-	-
Log Distance	2.64	3.18	2.47	3.78	3.94	2.88
Rejection Rate (%)	7.52	0.00	26.36	24.72	0.00	43.14
FICO	730	739	59	719	726	65
DTI	38.42	39.23	10.58	38.77	39.18	12.52
LTV	87.14	93.55	13.67	71.63	75.00	17.48
Observations		5,346,682			3,036,391	
Panel B: Originated Loans						
Local Share (%)	41.48	-	49.27	34.63	-	47.58
Log Distance	2.59	3.15	2.45	3.48	3.61	2.86
Interest Rate (%)	4.45	4.50	0.59	4.33	4.25	0.66
FICO	732	742	57	732	740	56
DTI	37.89	38.91	9.94	36.66	37.83	10.30
LTV	87.07	93.00	13.64	70.84	74.89	17.06
Observations		4,811,565			2,201,219	
Panel C: HMDA–McDash Matched Sample						
Local Share (%)	40.77	0.00	49.14	32.59	0.00	46.87
Log Distance	2.65	3.17	2.47	3.65	3.82	2.87
Default Rate (%)	8.70	0.00	28.18	6.22	0.00	24.15
Interest Rate (%)	4.43	4.50	0.58	4.33	4.25	0.63
FICO	730	739	57	730	736	54
DTI	38.41	39.52	9.69	37.10	38.35	9.87
LTV	87.69	95.00	13.04	71.82	75.00	15.76
Observations		3,054,156			1,215,441	

Table 2

Local Information Production Across Approval, Pricing, and Default Margins

This table presents the results of the main local-indicator specifications across the three key outcome margins. In Panel A, the dependent variable is an indicator equal to 100 if a loan application is rejected. In Panel B, the dependent variable is the interest rate on the originated loan, in percentage points. In Panel C, the dependent variable is an indicator equal to 100 if a loan becomes 60 days delinquent within two years of origination. The sample in Panel A includes all HMDA applications in 2018–2019. Panel B includes all originated HMDA loans in 2018–2019. Panel C includes approved HMDA loans in 2018–2019 that are matched to McDash performance records. *Local* equals one if the reviewing loan officer's work location is in the same county as the applicant. All specifications include a saturated set of hard-information controls: loan type interacted with polynomials in FICO, LTV, and DTI (including squared terms), as well as AUS status indicators. Standard errors are clustered at the borrower–county level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	Home Purchase			Refinance		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Rejection rate						
Local	-1.881***	-0.643***	-0.202***	-7.898***	-2.855***	-0.655***
	(0.09)	(0.05)	(0.03)	(0.22)	(0.15)	(0.10)
Observations	3,721,849	3,716,598	3,704,260	1,915,240	1,907,786	1,886,996
R^2	0.229	0.279	0.316	0.380	0.456	0.506
Panel B. Interest rate						
Local	0.033***	0.010***	-0.002**	0.028***	0.037***	0.009***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
Observations	3,370,947	3,365,679	3,354,197	1,450,704	1,442,695	1,421,876
R^2	0.564	0.661	0.707	0.546	0.655	0.735
Panel C. Default within two years						
Local	-0.332***	-0.463***	-0.344***	0.492***	0.037	0.000
	(0.07)	(0.06)	(0.05)	(0.10)	(0.08)	(0.08)
Observations	2,138,596	2,135,269	2,123,285	765,822	760,305	742,317
R^2	0.107	0.124	0.165	0.079	0.110	0.195
Controls	Yes	Yes	Yes	Yes	Yes	Yes
County × month FE	Yes	Yes	Yes	Yes	Yes	Yes
Lender × month FE	No	Yes	Yes	No	Yes	Yes
Loan officer FE	No	No	Yes	No	No	Yes

Table 3: Local Advantage in Loan Processing Time

This table reports loan-level regressions examining how borrower–officer proximity affects mortgage processing time for originated loans. The dependent variable is processing time, defined as the number of days between the loan application date and the origination date. The explanatory variable is an indicator for whether the reviewing loan officer is local. Columns (1)–(3) report results for home-purchase loans; columns (4)–(6) report results for refinance loans. All specifications include a saturated set of hard-information controls: loan type interacted with polynomials in FICO, LTV, and DTI ratios (including squared terms), as well as AUS status indicators. The three specifications correspond to increasingly saturated fixed effects as indicated in the table. The sample includes all originated home-purchase and refinance loans in confidential HMDA for 2018–2019. Standard errors are clustered at the borrower–county level. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Home Purchase			Refinance		
	(1)	(2)	(3)	(4)	(5)	(6)
Local	-2.383***	-0.176	-0.028	0.500**	-1.652***	-1.174***
	(0.37)	(0.15)	(0.07)	(0.20)	(0.13)	(0.08)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
County-Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Lender-Month FE		Yes	Yes		Yes	Yes
Loan Officer FE			Yes			Yes
N	3,380,205	3,374,927	3,363,438	1,452,425	1,444,421	1,423,612
R^2	0.090	0.299	0.371	0.129	0.280	0.395

Table 4: Local Wage Conditions and the Allocation of Underwriting Capacity

This table examines how local wage conditions shape the allocation of underwriting capacity across markets. In columns (1)–(2), the dependent variable equals one if a mortgage application is handled by a loan officer whose work location is in the same county as the borrower. The explanatory variable is the standardized finance-sector hourly wage in the borrower's market. Column (1) includes year fixed effects. Column (2) includes lender × year and state × year fixed effects. All specifications in columns (1)–(2) include borrower and loan controls. Standard errors in columns (1)–(2) are clustered at the borrower-county level. In columns (3)–(5), the unit of observation is lender–MSA–year and the dependent variable is the *Misalignment Index*, defined as $(m_{jkt}/M_{jt}) - (l_{jkt}/L_{jt})$, where m_{jkt}/M_{jt} is the share of lender j's mortgage applications from MSA k and l_{jkt}/L_{jt} is the share of its registered loan officers located there. A higher value indicates that a market contributes more to a lender's mortgage demand than to its local underwriting capacity. MSA-level controls include employment rates, income per capita, and earnings per job. Standard errors in columns (3)–(5) are clustered at the lender level. In all columns, a higher wage indicates a more expensive local labor market. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Local Assignment		Misalignment Index		
	(1)	(2)	(3)	(4)	(5)
Finance-Sector Hourly Wage (std.)	-0.0388***	-0.0197**	1.589***	1.814***	2.326***
	(0.010)	(0.009)	(0.21)	(0.25)	(0.27)
Borrower / Loan Controls	Yes	Yes			
Year FE	Yes		Yes		
State × Year FE		Yes			
Lender × Year FE		Yes		Yes	Yes
MSA Controls					Yes
N	4,970,686	4,970,611	187,671	186,631	185,709
R^2	0.0263	0.1573	0.002	0.049	0.050

Table 5: Soft Information Used in Loan Approval Decisions

This table presents the R^2 analysis results from regressions of loan rejection on hard-information variables, estimated separately for applications processed by local and non-local loan officers. Panel A reports R^2, while Panel B reports adjusted R^2. The dependent variable in all specifications is an indicator for whether a home purchase loan application is rejected. The underlying sample includes all home purchase loan applications in the confidential HMDA for 2018–2019. All specifications include a saturated set of hard-information: loan type interacted with a polynomial in FICO, loan-to-value (LTV), and debt-to-income (DTI) ratios (including squared terms), as well as indicators for automated underwriting system (AUS) status. We estimate six increasingly saturated specifications. Specification 1 includes only these hard-information. Specification 2 adds application-month fixed effects. Specification 3 replaces month fixed effects with borrower-county-by-month fixed effects. Specification 4 adds lender-by-month fixed effects in addition to borrower-county-by-month fixed effects. Specification 5 replaces these with lender-by-county-by-month fixed effects. Specification 6 further adds loan officer fixed effects. The first two columns report the R^2 (or adjusted R^2) from the full sample. The final three columns report results from 100 bootstrap replications, each using a random 10% subsample of applications. For each specification, we report the mean R^2 (or adjusted R^2) for local and non-local officers, and the mean difference between the two. Bootstrap standard errors are reported in parentheses, and t-statistics for the mean difference are reported in brackets. Standard errors are clustered at the borrower-county level. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: R^2

	Full Sample		Bootstrap Sample		
	Local Loan Officer	Non-Local Loan Officer	Local Loan Officer	Non-Local Loan Officer	Difference
Specification 1	0.191	0.217	0.191	0.217	-0.026***
			(0.004)	(0.003)	[-51.36]
Specification 2	0.191	0.217	0.191	0.217	-0.026***
			(0.004)	(0.003)	[-51.49]
Specification 3	0.214	0.246	0.269	0.316	-0.046***
			(0.004)	(0.003)	[-93.23]
Specification 4	0.265	0.303	0.376	0.409	-0.034***
			(0.004)	(0.003)	[-65.54]
Specification 5	0.340	0.438	0.478	0.561	-0.082***
			(0.005)	(0.004)	[-136.08]
Specification 6	0.374	0.477	0.589	0.695	-0.106***
			(0.004)	(0.004)	-182.83]
Bootstrap Samples			100	100	

Panel B: Adjusted R^2

	Full Sample		Bootstrap Sample		
	Local Loan Officer	Non-Local Loan Officer	Local Loan Officer	Non-Local Loan Officer	Difference
Specification 1	0.191	0.217	0.191	0.217	-0.026***
			(0.004)	(0.003)	[-51.50]
Specification 2	0.191	0.217	0.191	0.217	-0.026***
			(0.004)	(0.003)	[-51.70]
Specification 3	0.198	0.226	0.191	0.225	-0.034***
			(0.005)	(0.003)	[-61.66]
Specification 4	0.231	0.271	0.216	0.262	-0.046***
			(0.006)	(0.003)	[-72.13]
Specification 5	0.244	0.284	0.234	0.280	-0.046***
			(0.007)	(0.006)	[-49.90]
Specification 6	0.249	0.299	0.145	0.226	-0.081***
			(0.010)	(0.009)	-59.99]
Bootstrap Samples			100	100	

Table 6: Cross-Sectional Heterogeneity in Local Officer Effects by Borrower Risk

This table examines how the impact of local loan officers varies with borrower risk across three outcomes: rejection (columns 1–3), default (columns 4–6), and posted interest rates (columns 7–9), all in percentage points. The sample consists of conventional and jumbo home-purchase mortgage applications in confidential HMDA for 2018–2019 (excluding FHA, VA, and RHS). Columns 1–3 use all applications; columns 4–6 use originated loans matched to McDash performance records; and columns 7–9 use all originated loans. Borrower risk is captured by (i) Subprime (FICO < 670), (ii) High DTI (DTI > 43), and (iii) High LTV (LTV > 80). Each specification includes indicators for Local, a high-risk indicator, and their interaction; loan-type interactions with polynomials in FICO, LTV, and DTI (including squared terms); AUS status indicators; county-month fixed effects; and lender-month fixed effects. Standard errors are clustered at the borrower–county level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	Rejection (pp)			Default (pp)			Interest Rate (pp)		
	Subprime (1)	High DTI (2)	High LTV (3)	Subprime (4)	High DTI (5)	High LTV (6)	Subprime (7)	High DTI (8)	High LTV (9)
Local	-0.369***	-0.336***	-0.337***	-0.242***	-0.186***	-0.165**	0.015***	0.014***	0.017***
	(0.06)	(0.07)	(0.09)	(0.06)	(0.06)	(0.07)	(0.00)	(0.00)	(0.00)
Local×Subprime	-2.927***			-0.428			-0.010*		
	(0.24)			(0.30)			(0.01)		
Local×High DTI		-0.761***			-0.273**			0.003*	
		(0.13)			(0.12)			(0.00)	
Local×High LTV			-0.386***			-0.188*			-0.003
			(0.10)			(0.11)			(0.00)
County-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lender-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2,611,212	2,611,212	2,611,212	1,419,546	1,419,546	1,419,546	2,385,885	2,385,885	2,385,885
R^2	0.252	0.250	0.250	0.093	0.092	0.092	0.670	0.664	0.665

Table 7: Testing for Borrower Sorting

This table tests borrower sorting. The dependent variable is an indicator for whether the mortgage application is rejected (scaled so coefficients are reported in percentage points). *Local* equals one if the loan officer handling the application works in the same county as the borrower, and zero otherwise. *Finance-Sector Hourly Wage (std.)* is the standardized finance-sector hourly wage in the borrower's market. Columns (1)–(2) restrict the sample to home-purchase applications, and columns (3)–(4) restrict the sample to refinance applications. Panel A uses the full sample of markets, while Panel B restricts attention to metropolitan (MSA) markets. All specifications include loan-type dummies fully interacted with borrower and loan controls (FICO, $FICO^2$, LTV, LTV^2, DTI, DTI^2, and AUS status). Columns (1) and (3) include borrower-county × application-month fixed effects; columns (2) and (4) additionally include lender × application-month fixed effects. Standard errors, reported in parentheses, are clustered at the borrower-county level. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Home Purchase		Refinance	
	(1)	(2)	(4)	(5)
Panel A: Full Sample				
Local	-4.2377***	-1.2761***	-18.3222***	-4.3123***
	(0.526)	(0.378)	(1.248)	(0.861)
Local×Finance-Sector Hourly Wage (std.)	0.3215***	0.0859	1.3809***	0.1911
	(0.070)	(0.052)	(0.164)	(0.118)
Borrower / Loan Controls	Yes	Yes	Yes	Yes
Borrower-County × Month FE	Yes	Yes	Yes	Yes
Lender × Month FE		Yes		Yes
N	3,657,215	3,652,044	1,888,481	1,881,160
R^2	0.2274	0.2779	0.3804	0.4561
Panel B: MSA-Area				
Local	-4.0562***	-1.0904***	-17.6405***	-3.5586***
	(0.597)	(0.410)	(1.385)	(0.919)
Local×Finance-Sector Hourly Wage (std.)	0.2981***	0.0597	1.2913***	0.0909
	(0.077)	(0.056)	(0.180)	(0.124)
Borrower / Loan Controls	Yes	Yes	Yes	Yes
Borrower-County × Month FE	Yes	Yes	Yes	Yes
Lender × Month FE		Yes		Yes
N	3,253,393	3,247,847	1,714,927	1,707,332
R^2	0.2177	0.2693	0.3727	0.4489

Table 8: Internally Calibrated Default Parameters

This table reports calibrated default rates, latent heterogeneity, and aggregate shock values for each market for the model in Section 4. Markets are segmented based on borrower risk type (Subprime (FICO $<$ 660), Near Prime ($660 \leq$ FICO $<$ 720), and Prime (FICO $\geq$ 720)) and loan purpose (Purchase vs Refi). See Section 5.2 for more details.

		Baseline risk measures		
Market	Risk tier	Default rate	σ_i^x	$\epsilon_{i,t}$
Refi	Subprime	31.6%	4.41	-0.99
	Near-Prime	12.9%	5.16	-1.98
	Prime	8.0%	9.19	-2.95
Purchase	Subprime	7.9%	1.21	0.12
	Near-Prime	3.9%	1.92	-0.68
	Prime	2.1%	4.75	-2.30

Table 9: Baseline Informational Efficiency: Credit Rationing and Pooled Origination

This table reports the amount of credit rationing and pooled origination in the baseline equilibrium, as implied by the model in Section 4. Markets are segmented based on borrower risk type (Subprime (FICO < 660), Near Prime ($660 \leq$ FICO < 720), and Prime (FICO ≥ 720)) and loan purpose (Purchase vs Refi). Panel A reports aggregate rationing and pooling rates; Panel B breaks these out by loan officer type. See Section 6.1 for more details.

		Share of applications (%)	
Market	Risk tier	Credit rationing	Pooled origination
Panel A: Overall rates			
Refi	Subprime	15.3%	2.4%
	Alt-A	8.4%	2.0%
	Prime	3.6%	0.5%
Purchase	Subprime	10.3%	10.6%
	Alt-A	5.6%	4.3%
	Prime	2.3%	0.8%

Baseline Informational Efficiency (continued)

		Credit rationing (%)		Pooled origination (%)	
Market	Risk tier	Local officers	Distant officers	Local officers	Distant officers
Panel B: By loan officer type					
Refi	Subprime	8.7%	17.8%	2.0%	2.5%
	Alt-A	5.0%	9.9%	1.6%	2.2%
	Prime	1.9%	4.4%	0.4%	0.6%
Purchase	Subprime	10.1%	10.4%	8.6%	12.2%
	Alt-A	5.1%	6.0%	3.2%	5.1%
	Prime	1.5%	2.8%	0.6%	0.9%

Table 10: Technology Shock and Lending Outcomes

This table reports the results of a counterfactual shock to distant loan officers' quantity productivities. Markets are segmented based on borrower risk type (Subprime (FICO $<$ 660), Near Prime (660 $\leq$ FICO $<$ 720), and Prime (FICO $\geq$ 720)) and loan purpose (Purchase vs Refi). The table reports the results of only shocking Brick and Mortar banks, as well as shocking all banks. See Section 6.2 for more details.

		Change (percentage points)		
Market	Risk tier	Rejection rate	Rationing	Pooled origination
Panel A: Shock 1 – Brick and Mortar				
Refi	Subprime	0.32%	0.34%	0.02%
	Alt-A	0.29%	0.30%	0.01%
	Prime	0.08%	0.10%	0.02%
Purchase	Subprime	-0.16%	0.44%	0.60%
	Alt-A	-0.11%	0.14%	0.25%
	Prime	0.11%	0.13%	0.02%
Panel B: Shock 2 – All Banks				
Refi	Subprime	1.13%	1.13%	-0.01%
	Alt-A	0.91%	0.87%	-0.04%
	Prime	0.21%	0.26%	0.06%
Purchase	Subprime	1.29%	2.02%	0.73%
	Alt-A	0.28%	0.73%	0.45%
	Prime	0.33%	0.37%	0.05%

Technology Shock and Lending Outcomes (continued)

		Change in expected default		Change in monthly payment	
Market	Risk tier	Shock 1	Shock 2	Shock 1	Shock 2
Panel C: Expected Defaults and Monthly Payments					
Refi	Subprime	0.06%	0.03%	-\$3.07	-\$16.55
	Alt-A	0.02%	0.01%	-\$2.70	-\$14.09
	Prime	0.01%	0.01%	-\$1.87	-\$9.98
Purchase	Subprime	0.25%	0.28%	-\$4.74	-\$23.23
	Alt-A	0.07%	0.13%	-\$3.10	-\$16.43
	Prime	0.01%	0.02%	-\$2.58	-\$13.45

Appendix for Online Publication

A Detailed Institutional Background

The U.S. residential mortgage market operates under a standardized and highly regulated origination process governed by the Truth-in-Lending Act (TILA), the Real Estate Settlement Procedures Act (RESPA), and their integration through the TILA–RESPA Integrated Disclosure (TRID) rules administered by the Consumer Financial Protection Bureau (CFPB).[15] Although internal procedures vary across lenders, the fundamental sequence is uniform:

$$\text{rate setting} \rightarrow \text{formal application} \rightarrow \text{information collection} \rightarrow \text{underwriting} \rightarrow \text{origination.}$$

A defining institutional feature is the separation between (i) the pricing stage, which occurs before verified information is available, and (ii) the underwriting stage, which evaluates the hard information assembled throughout the application pipeline.

A.1 Rate Setting Prior to Information Acquisition

Origination begins with *rate setting*. Lenders publish rate sheets daily, and loan officers (LOs) quote interest rates based on borrower-reported characteristics, program eligibility, and prevailing market conditions.[16] At this stage, lenders know only self-reported borrower information; no verified income, asset, employment, or collateral documents have been collected.

Under TRID, lenders must issue a Loan Estimate (LE) within three business days of receiving a formal application.[17] After the LE is issued, TRID tightly restricts circumstances under which lenders may increase interest rates or fees. Upward repricing is allowed only in narrowly defined "changed circumstances" such as borrower-initiated changes or corrections to borrower-provided information.[18] Thus, lenders generally *cannot* reprice loans upward after underwriting reveals adverse information. This contrasts with canonical screening models in industrial organization, where lenders can always adjust prices after privately observing borrower risk.

[15] See CFPB, "TILA–RESPA Integrated Disclosure Rule (TRID): Guide to Forms" (2022).
[16] Mortgage Bankers Association (MBA), "Mortgage Origination Survey," various years.
[17] 12 C.F.R.§1026.19(e)(1)(iii).
[18] 12 C.F.R.§1026.19(e)(3)(iv).

A.2 Formal Application and Information Collection

If the borrower proceeds, they submit a *formal application*, triggering the issuance of the LE and initiating the *information-collection* stage. Loan officers gather income documentation, bank statements, credit reports, appraisals, verifications of employment and assets, and third-party reports following detailed agency and investor requirements.[19]

Industry manuals consistently emphasize that loan officers are responsible for file completeness, accuracy, and timeliness. LOs coordinate with borrowers, employers, appraisers, title companies, and verification vendors; missing or inconsistent documents frequently delay or derail underwriting.

A.3 Underwriting and the Role of Loan Officers in Approval

During *underwriting*, human underwriters and automated underwriting systems (AUS)—Fannie Mae Desktop Underwriter®, Freddie Mac Loan Product Advisor®, and proprietary lender systems—evaluate the collected information. Underwriters assess DTI, LTV, credit history, collateral value, and program eligibility.

Although underwriters hold formal approval authority, their decisions depend entirely on the information produced earlier in the pipeline. Underwriters do not independently collect additional documents; if LOs provide incomplete or inconsistent files, denials or conditional approvals frequently follow.[20] Thus, loan officers play a *first-order* role in shaping approval outcomes by determining the quality and completeness of the underwriting file.

A.4 Local Versus Remote Loan Officers

The rise of centralized and online lending platforms has sharpened the distinction between *local* and *remote* LOs. Local LOs are based in the same geographic market as borrowers and interact regularly with local employers, real estate agents, appraisers, and title companies. This proximity reduces frictions in document collection, verification, and communication.

Remote LOs—often located in call centers or out-of-state hubs—communicate primarily

[19] See Fannie Mae Selling Guide B1–1–01 (2024); Freddie Mac Seller/Servicer Guide (2024).
[20] Freddie Mac Seller/Servicer Guide, Section 5101.2 (2024).

by phone or online channels and may face greater difficulty securing timely verifications or coordinating with local service providers.[21] Because lenders cannot freely raise rates after underwriting and because documentation quality affects approval outcomes, these information frictions translate directly into differences in rejection probabilities and processing times.

A.5 Origination and Final Rate Setting

Once underwriting conditions are satisfied, the lender issues the Closing Disclosure (CD) and proceeds to *origination*. TRID restricts rate changes at this stage. Downward renegotiation is common when market rates decline or borrowers present competing offers, but upward repricing is generally prohibited absent a qualifying changed circumstance.[22]

Hence, the initial pricing decision carries independent economic significance, and underwriting preserves a meaningful approval margin that cannot be offset through ex-post price adjustments.

B Sample Construction and Data Cleaning

This appendix describes how we construct the loan-level dataset used in the analysis by combining confidential HMDA application records, NMLS loan officer registrations, and McDash performance data. Our cleaning procedures follow standard practice in the mortgage literature, including Bhutta and Hizmo (2021), and impose additional restrictions to ensure comparability across borrowers and consistency in loan structure.

We begin with the confidential HMDA application files. To obtain a homogeneous sample of loans subject to consistent underwriting standards, we restrict attention to completed applications for first-lien, 30-year, fixed-rate mortgages secured by owner-occupied properties. Applications sourced through mortgage brokers or purchased from other lenders are excluded because these channels do not reflect the originating lender's internal screening technology or loan officer assignment. These filters remove products with distinct risk profiles, heterogeneous documentation requirements, or limited information on the underwriting process.

[21] MBA, "Technology & Origination Report," 2020.

[22] 12 C.F.R.§1026.19(e)(3).

We retain both home-purchase and refinance applications that meet these criteria. For originated loans in 2018–2019, the confidential version of HMDA provides the identity of the loan officer who processed the application. Using this identifier, we merge each application to the NMLS registry to obtain the officer's physical work location.

Next, we merge the HMDA originations to monthly loan-level performance data from the Black Knight McDash servicing database. Following the methodology of Rosen (2011), we treat a loan in HMDA and a loan in McDash as the same origination only when several characteristics match almost exactly. First, the reported origination dates in the two datasets must lie within five calendar days of each other, and the HMDA action date must be within five days of the McDash origination date to ensure consistent application–closing timing. Second, the origination amounts must differ by less than $10. Third, the property ZIP code, lien type, loan purpose (purchase or refinance), loan type, and occupancy type must agree exactly across datasets. These conditions minimize the possibility of false matches while retaining a large and representative subset of the market.

Using this procedure, we successfully merge approximately 36 percent of originated loans in confidential HMDA to 68 percent of loans in McDash. As in prior work, the imperfect overlap reflects the fact that not all HMDA-reporting lenders service loans in McDash and not all McDash servicers appear as HMDA reporters. For matched loans, we construct a two-year performance history and define a delinquency indicator equal to one if a loan becomes sixty or more days delinquent within twenty-four months of origination. To ensure complete performance histories for all loans, we restrict the analysis to applications submitted in 2018–2019.

The final dataset links each loan application to (i) borrower and loan characteristics at the time of application, (ii) the lender and loan officer responsible for the file, (iii) the physical location of the loan officer obtained from NMLS, (iv) underwriting decisions and posted interest rates, and (v) loan performance up to two years after origination. This merged dataset forms the basis for our analysis of geographic misallocation, screening efficiency, labor-input choices, and credit-access outcomes.

C Conceptual Framework

C.1 Primitives

This section presents a simple conceptual framework that rationalizes the reduced form tests in Section 2. There is a unit mass of ex-ante identical consumers. A bank sets an ex-ante rate $\bar{r}$ under a break-even constraint. A consumer is differentiated by their default rate $\theta \sim F$. Applications follow a cutoff rule: all consumers with default rate $\theta \geq \underline{\mathrm{d}}\left(\bar{r}\right)$ apply to the bank. Assume $\underline{\mathrm{d}}$'$(r)\geq 0$. The bank pays a fixed processing cost of c per application.

If the consumer applies, the bank observes a signal $\hat{\theta} = \theta + \epsilon$ with $\epsilon \sim G$, and can reject or approve. If the bank approves, it may adjust the rate to r^*, after which the loan is originated with probability $\varphi\left(r^*, r\right)$. The bank earns $r^* - \theta$ if the loan is originated and zero otherwise.

Regularity Conditions:

- F and G admit densities f and g, respectively;
- $\theta \perp \epsilon$;
- The signal satisfies the MLRP. Since the noise is additive, this is equivalent to:

$$\frac{g\left(x\right)}{g\left(x+\Delta\right)} \text{ is decreasing in } x \text{ for all } \Delta > 0.$$

If g is twice differentiable, this is equivalent to log-concavity: $\frac{\partial^2 lng}{\partial \epsilon^2} \leq 0$.

Signal Precision

Assume G belongs to a family of noise distributions ordered by parameter $\sigma \in \mathbb{R}$ in the Blackwell sense. Formally: there exists a random variable $\eta \perp \epsilon$ with density $h_{\sigma',\sigma}$ such that $g\left(\epsilon;\sigma'\right) = \int g\left(\epsilon - \eta;\sigma\right) h_{\sigma',\sigma}\left(\eta\right) d\eta$.[23] Let $m\left(\hat{\theta},\sigma\right) := \mathbb{E}\left[\theta|\hat{\theta},\sigma\right]$ denote the posterior mean.

[23]For example, if $G \sim \mathcal{N}\left(0,\sigma^2\right)$,then $\begin{bmatrix}\epsilon\\ \eta\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma^2 & 0\\ 0 & \left(\sigma'\right)^2-\sigma^2\end{bmatrix}\right)$

C.2 Results With Ex-post Pricing

Suppose the bank faces no constraint in ex-post repricing, other than the fact that consumers may walk away if it ex-post raises prices. For simplicity, we can model this walking behavior through an exponential decay function:[24]

$$\varphi\left(r^{*}, \bar{r}\right) = exp\left\{-\lambda^{-1} r^{*}\right\}$$

for hazard rate $\lambda \geq 0$.

Lemma 1. *(Optimal Ex-Post Pricing).*The bank never rejects borrowers and sets $r^{*} = m\left(\hat{\theta}, \sigma\right) + \lambda$.

Proof. Given signal $\hat{\theta}$,the bank's expected profits from setting rate r^{*} are:

$$\varphi\left(r^{*}; \bar{r}\right)\left(r^{*} - m\left(\hat{\theta}, \sigma\right)\right)$$

Taking the FOC wrt r^{*} gives the interior solution:

$$r^{*} = m\left(\hat{\theta}, \sigma\right) + \lambda$$

Note that $\left(m\left(\hat{\theta}, \sigma\right) + \lambda - m\left(\hat{\theta}, \sigma\right)\right)\varphi\left(r^{*}; \bar{r}\right) = \lambda\varphi\left(r^{*}; \bar{r}\right) > 0\ \forall r^{*} \in \mathbb{R}$, meaning rejection is never optimal. $\square$

For the next result, we consider a setting without adverse selection to cleanly isolate the properties of ex-post risk pricing.

Proposition 1. *(A Test for Ex-Post Repricing).* Suppose $\underline{\text{d}}$'=0.Then the correlation between realized default rates and realized interest rates is decreasing in σ.

Proof. By the law of total covariance, given that $r^{*}$is fully determined by $\hat{\theta}$, $Cov\left(\theta, r^{*}\right) = Cov\left(m, r^{*}\right)$. Thus:

[24] It is straightforward to replace this parametric assumption with a set of general regularity conditions on consumer "walking" behavior. A sufficient set of conditions is weak log-concavity and a weakly convex hazard function. The necessary condition is that the hazard rate doesn't change "too fast" in the tails of the walking function.

$$corr\left(\theta, r^*\right) = \frac{cov\left(m, r^*\right)}{\sqrt{Var\left(\theta\right)Var\left(r^*\right)}}$$

Applying Lemma 1 gives:

$$\frac{cov\left(m, r^*\right)}{\sqrt{Var\left(\theta\right)Var\left(r^*\right)}} = \sqrt{\frac{Var\left(m\right)}{Var\left(\theta\right)}}$$

By the Blackwell ordering on G, and the law of iterated expectations, $Var\left(m\right)$ is decreasing in σ, completing the proof. □

C.3 Results Without Ex-Post Repricing

Suppose regulatory constraints (e.g. TRID) make origination with an upward re-priced loan infeasible. Suppose also that consumers never walk from an approved loan.[25] Then:

$$\varphi\left(r^*, \bar{r}\right) = \begin{cases} 0 & r^* > \bar{r} \\ 1 & r^* \leq \bar{r} \end{cases}$$

For this section, let $\bar{d}$ denote the realized default rate on originated loans and let $\bar{p}$ denote the approval probability.

Lemma 2. *(Optimal Approval Rule).* The bank approves a loan if and only if $\hat{\theta} \leq \hat{\theta}^*\left(\bar{r}, \sigma\right)$

Proof. Approval is optimal if and only if expected profit is nonnegative. Upward repricing is infeasible and downward repricing yields no additional surplus. This implies the decision rule

$$\text{Approve} \iff m\left(\hat{\theta}, \sigma\right) \leq \bar{r}.$$

By the MLRP, $m\left(\hat{\theta}, \sigma\right)$ is weakly increasing in $\hat{\theta}$. Hence, the approval set, $\mathcal{D} := \left\{\hat{\theta} : \text{approve}\right\}$ is an interval. Defining $\hat{\theta}^*\left(\bar{r}, \sigma\right) = \sup \mathcal{D}$ completes the proof. □

Lemma 3. *(Signal Precision and Default Rates).* Suppose $\sigma_2 < \sigma_1$. Then, for any approval cutoffs $\hat{\theta}_1^*$and $\hat{\theta}_2^*$such that $\bar{p}_2 \leq \bar{p}_1$, $\bar{d}_2 < \bar{d}_1$.

[25]In our sample, fewer than five percent of approved consumers walk away.

Proof. For expositional simplicity, assume $p_1 = p_2$. Let $A_i(\theta) = G\left(\hat{\theta}_i^* - \theta; \sigma_i\right)$ denote the approval probability of a consumer with true default rate θ. Note that A_i is strictly decreasing in θ.

Since $\bar{p}_1 = \bar{p}_2$, $\int A_1(\theta) F(d\theta) = \int A_2(\theta) F(d\theta)$. Further, since $\sigma_2 < \sigma_1$, there exists η such that $A_1(\theta) = \int G\left(\hat{\theta}_1^* - \theta - \eta; \sigma_2\right) H(d\eta)$. Thus, by the MLRP on G, $A_1(\theta) - A_2(\theta)$ satisfies a single crossing property from below: there exists a unique θ^* such that $A_1(\theta) - A_2(\theta) \geq 0$ for all $\theta \geq \theta^*$ and $A_1(\theta) - A_2(\theta) < 0$ for all $\theta < \theta^*$.

Note $\bar{d}_i = \frac{\int \theta A_i(\theta) F(d\theta)}{\bar{p}_i} = \frac{\int_{\underline{\mathrm{d}}}^{\theta^*} \theta A_i(\theta) F(d\theta) + \int_{\theta^*}^{\infty} \theta A_i(\theta) F(d\theta)}{\bar{p}_i}$. Thus:

$$\begin{aligned}
\bar{p}_1\left(\bar{d}_2 - \bar{d}_1\right) = \int_{\underline{\mathrm{d}}}^{\theta^*} \theta\left(A_2(\theta) - A_1(\theta)\right) F(d\theta) + \int_{\theta^*}^{\infty} \theta\left(A_2(\theta) - A_1(\theta)\right) F(d\theta) \\
< \theta^* \int A_2(\theta) - A_1(\theta) F(d\theta) = 0
\end{aligned}$$

Finally, if $\bar{p}_2 < \bar{p}_1$, the required cutoff $\hat{\theta}_2^*$is strictly lower than the cutoff which equates approval probabilities. By the MLRP, $\bar{d}_2$ is strictly increasing in the cutoff, implying $\bar{d}_2$must be strictly lower. □

Proposition 2. *(Ex-Ante Pricing).* Realized prices depend on the ex-ante risk and signal distributions and are independent of the realized signal during origination.

Proof. Anticipating its optimal decision rule from Lemma 2, the bank sets $\bar{r} = \bar{d} + \frac{c}{\bar{p}}$ to satisfy its break even constraint. This price depends only on expectations over F and G, and not on realized signal $\hat{\theta}$. □

Proposition 3. *A sufficient condition for informational superiority.*

Suppose two distinct underwriting technologies (either within or across banks) face the same potential applicant distribution. If underwriting technology 1 exhibits (i) (weakly) lower rejection rates, (ii) (weakly) lower realized default rates, and (iii) (weakly) higher interest rates than underwriting technology 2, with at least one inequality strict, then technology 1 must use a more precise signal than technology 2.

Proof. By way of contradiction, suppose $\sigma_1 \geq \sigma_2$. Since $r_1 \geq r_2$ and $\underline{\mathrm{d}}$'$\geq$0,we have

$\underline{\mathrm{d}}\,(r_1) \geq \underline{\mathrm{d}}\,(r_2)$. Thus, the applicant pool for technology 1 is (weakly) riskier in a FOSD sense.

Case 1: If $\sigma_1 = \sigma_2$, since $\bar{p}_1 \geq \bar{p}_2$,it must be the case that technology 1 is originated with a (weakly) less strict cutoff. This contradicts $\bar{d}_1 \leq \bar{d}_2$, as long as at least one of the inequalities is strict.

Case 2: If $\sigma_1 > \sigma_2$, let $\bar{d}_i\,(\underline{\mathrm{d}})$denote the realized default rate of technology i if it faced application cutoff $\underline{\mathrm{d}}$. Since $\bar{p}_1 \geq \bar{p}_2$, by Lemma 3, it must be the case that $\bar{d}_1\left(\underline{\mathrm{d}}\,(r_2)\right) > \bar{d}_2$. Yet, by the MLRP, lowering the lower cutoff of the default type integral can only lower the realized default rate, meaning $\bar{d}_2 < \bar{d}_1$. $\rightarrow\leftarrow$. $\square$

D Structural Model Appendix

D.1 Labor Supply Function

In this section, we fully describe and derive the labor supply function from the text.

Environment. There is a mass of $\bar{L}$ potential loan officers, who can choose to be either a loan officer, or an outside option. There are N_k potential outside options in region k, which pay a wage of w_{ook}. We assume that the CDF of amenity draws, A_{ijk} is given by:

$$\mathbb{P}\left(\left(\bigcap_{\forall i,j,k}\{a_{ijk} \leq t_{ijk}\}\right)\right) = exp\left\{\left(-\sum_{k}\left\{\left(\sum_{i,j} t_{ijk}^{-\frac{\epsilon}{(1-\rho)(1-\phi)}}\right)^{1-\rho} + N_{k,oo}t_{k,oo}^{-\frac{\epsilon}{1-\phi}}\right\}^{1-\phi}\right)\right\}$$

with composite parameter $\varrho := \frac{\epsilon}{(1-\rho)(1-\phi)}$. This CDF implies that loan officers idiosyncratic preference draws are correlated within living location, consistent with Giroud et al. (2024). This implies that, by posting a wage of w_{ijk}, bank j hires:

$$l_{ijk}(w_{ijk}) = \bar{L}w_{ijk}^{\varrho}\Psi_k^{-\varrho},$$

for local wage index Ψ_k following:

$$\Psi_k := \left\{ \left(\sum_{i,j} w_{ijk}^{\frac{\epsilon}{(1-\rho)(1-\phi)}} \right)^{\rho} \left[\left(\sum_{i,j} w_{ijk}^{\frac{\epsilon}{(1-\rho)(1-\phi)}} \right)^{1-\rho} + N_{k,oo} \left(w_{ook} \right)^{\frac{\epsilon}{1-\phi}} \right]^{\phi} \left(\sum_{m} \left[\left(\sum_{i,j} w_{ijm}^{\frac{\epsilon}{(1-\rho)(1-\phi)}} \right)^{1-\rho} + N_{m,oo} \left(w_{m,oo} \right)^{\frac{\epsilon}{1-\phi}} \right]^{1-\phi} \right) \right\}^{\frac{1}{\varrho}}$$

It can be shown that this implies the labor cost function:

$$C_{ijk}(\tilde{l}_{ijk}) = \underbrace{\bar{L}^{\varrho^{-1}} \Psi_k e_{ijk}^{-1}}_{A_{ijk}^{-1}} \tilde{l}_{ijk}^{1+\varrho^{-1}}$$

.

Finally, note that, in the limit, as $\bar{L}$ and $N_{k,oo}$ jointly all approach infinity (i.e., loan officers are atomistic in the labor market), the local wage index becomes exogenous and heterogeneity in the index is driven entirely by heterogeneity in w_{ook}.

D.2 Securitization

This subsection details how securitization decisions can be incorporated into the quantitative model. The level of distortion induced by securitization depends on the extent of information shared with the secondary market. While multiple channels encourage some degree of truth-telling, if these incentives are imperfect, the frictions exacerbate the informational externalities documented in the analysis.

Consider a loan with an interest rate r and a full-information per-period default hazard rate d, at which point a fraction lgd of the balance is lost and the loan stops making payments. If the loan is funded at cost f, the NPV of the loan is approximately proportional to its expected per-period excess interest income of $r - f - d \times lgd$.[26] When

[26] Formally, as a continuous-time perpetuity, the NPV is exactly $\frac{r-f-d\times lgd}{\rho+d}$ for discount rate ρ. In the empirically relevant range, where loan terms are long and default is relatively low, the distortion from the exact continuous-time

the loan is sold on a secondary market, the bank receives a lump sum payment proportional to $\mathbb{E}[r - f - d \times lgd \mid \mathcal{I}]$, where $\mathcal{I}$ represents the capital markets information set. Under full informational efficiency—where information generation effort is credibly communicated—the bank's problem remains consistent with the baseline specification in the main text.

Informational Inefficiencies. When informational inefficiencies persist, the bank's profit function can be specified as $r - f - \alpha(d \times lgd) - (1-\alpha)(\overline{d \times lgd})$, where $\overline{d \times lgd}$ is the average loss rate on loans in the market. The parameter α captures the degree to which banks internalize their impact on expected default rates when making approval and rate decisions. If $\alpha < 1$, a new externality is introduced, which increases the distortions imposed by the under-employment of local labor and worsens the resulting inefficiencies. While it could be valuable to quantify the impact of this additional inefficiency, a major identification challenge arises: even perfectly exogenous variation in securitization behavior cannot identify α. The intuition is straightforward. Securitizing a loan affects the lender's problem through two distinct channels. First, if $\alpha < 1$, the lower internalization of default risk leads to looser origination standards. On the other hand, *even if* $\alpha = 1$, securitization also affects origination standards through funding costs. Since both channels enter the bank's problem in the same way, and jointly affect two unobservables, additional variation would be required to identify α. Thus, we work with the conservative assumption that $\alpha = 1$.

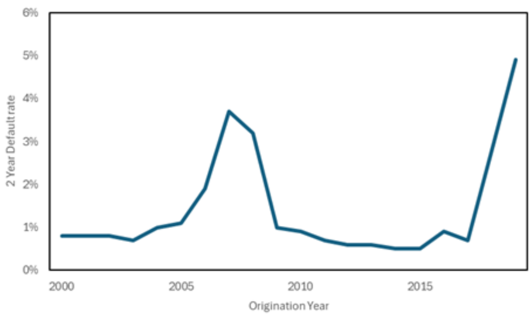
Default Internalization Mechanisms. Why might it be reasonable to assume that banks internalize default risk? The literature has shown that there are significant differences in banks' screening effort, even in securitization, consistent with skin-in-the-game (Bosshardt et al., 2025). Consistent with this, Securitization markets incorporate several mechanisms to improve the effective informativeness of signals, or alternatively, implement ex-post transfers that dynamically ensure the same outcome. Three prominent mechanisms are put-backs, reputational concerns, and servicing rights. Under put-back covenants, defaults from an unseasoned loan that did not follow proper underwriting procedures can be returned to the bank's balance sheet, requiring the bank to

perpetuity is minimal, meaning the proportionality to the numerator is nearly exact.

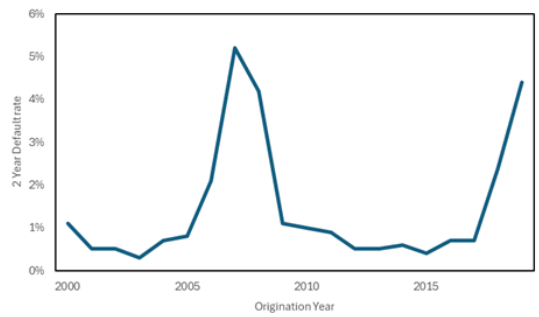
internalize the full ex-post default loss. Consequently, banks can signal informational efficiency through their stated underwriting processes. Motivated by this feature, default in the reduced-form analysis is defined as a typical put-back covenant violation (i.e., 60+ days past due within the first 24 months of the loan), with the LGD calibrated accordingly. Additionally, reputational concerns may constrain banks that diverge significantly from good-faith underwriting standards. Finally, most securitizing banks maintain servicing rights for their originated loans; these rights are valuable precisely when the loan is unobservably safer, since the servicer bears the liquidity and operational costs of delinquency. Notably, the empirical results of this paper control for automated underwriting system recommendations; the residual variation in rejection and default observed may therefore be driven by the resulting internalization of excess defaults induced by such informational efficiency measures.

E Additional Figures and Tables

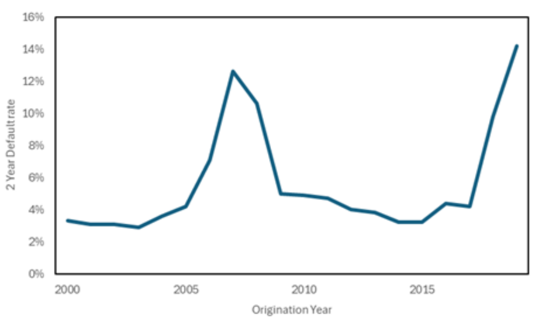
Figure A1. Time Series of Default by Credit Grade and Loan Purpose



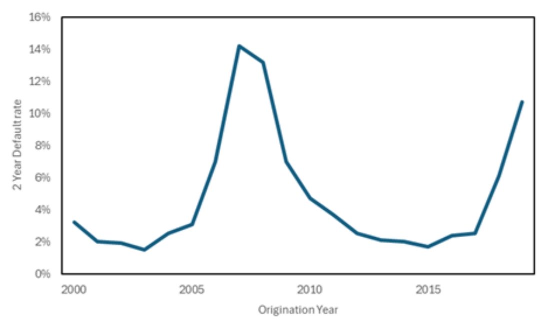
(a) Prime Purchase Market



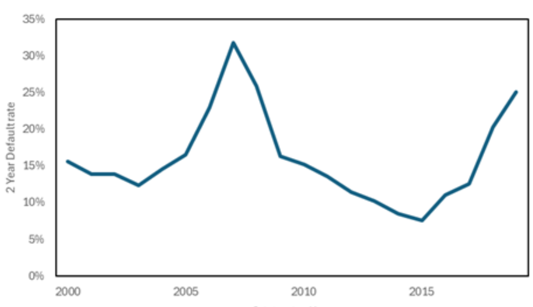
(b) Prime Refi Market



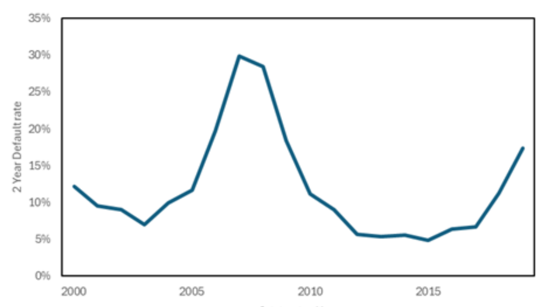
(c) Near-Prime Purchase Market



(d) Near-Prime Refi Market



(e) Subprime Purchase Market



(f) Subprime Refi Market

Note: This figure plots the time series of 2 year mortgage default rates in the McDash data, by credit grade, loan purpose, and origination year

Table A1 Distance and Loan Outcomes

This table reports the log-distance specifications corresponding to the main-text results. In Panel A, the dependent variable is an indicator equal to 100 if a loan application is rejected. In Panel B, the dependent variable is the interest rate on the originated loan, in percentage points. In Panel C, the dependent variable is an indicator equal to 100 if a loan becomes 60 days delinquent within two years of origination. The sample in Panel A includes all HMDA applications in 2018–2019. Panel B includes all originated HMDA loans in 2018–2019. Panel C includes approved HMDA loans in 2018–2019 that are matched to McDash performance records. *Log Distance* is the log distance between the borrower's location and the reviewing loan officer's office. All specifications include a saturated set of hard-information controls: loan type interacted with polynomials in FICO, LTV, and DTI (including squared terms), as well as AUS status indicators. Standard errors are clustered at the borrower–county level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	Home Purchase			Refinance		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Rejection rate						
Log Distance	0.592***	0.228***	0.069***	1.822***	0.797***	0.236***
	(0.02)	(0.01)	(0.01)	(0.05)	(0.03)	(0.02)
Observations	3,721,799	3,716,547	3,704,210	1,915,218	1,907,764	1,886,977
R^2	0.231	0.279	0.316	0.387	0.457	0.506
Panel B. Interest rate						
Log Distance	-0.010***	-0.003***	0.000**	-0.004***	-0.007***	-0.002***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	3,370,902	3,365,633	3,354,152	1,450,692	1,442,684	1,421,868
R^2	0.565	0.661	0.707	0.545	0.655	0.735
Panel C. Default within two years						
Log Distance	0.033***	0.096***	0.098***	-0.105***	0.008	0.010
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Observations	2,138,585	2,135,258	2,123,274	765,821	760,304	742,317
R^2	0.107	0.124	0.165	0.079	0.110	0.195
Controls	Yes	Yes	Yes	Yes	Yes	Yes
County × month FE	Yes	Yes	Yes	Yes	Yes	Yes
Lender × month FE	No	Yes	Yes	No	Yes	Yes
Loan officer FE	No	No	Yes	No	No	Yes

Table A2: Distance and Loan Processing Time

This table reports loan-level regressions examining how borrower–officer distance affects mortgage processing time for originated loans. The dependent variable is processing time, defined as the number of days between the loan application date and the origination date. The explanatory variable is the log distance between the borrower and the reviewing loan officer. Columns (1)–(3) report results for home-purchase loans; columns (4)–(6) report results for refinance loans. All specifications include a saturated set of hard-information controls: loan type interacted with polynomials in FICO, LTV, and DTI ratios (including squared terms), as well as AUS status indicators. The three specifications correspond to increasingly saturated fixed effects as indicated in the table. The sample includes all originated home-purchase and refinance loans in confidential HMDA for 2018–2019. Standard errors are clustered at the borrower–county level. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Home Purchase			Refinance		
	(1)	(2)	(3)	(4)	(5)	(6)
Log Distance	0.565***	-0.017	-0.021	-0.207***	0.349***	0.312***
	(0.07)	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
County-Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Lender-Month FE		Yes	Yes		Yes	Yes
Loan Officer FE			Yes			Yes
N	3,380,160	3,374,881	3,363,393	1,452,413	1,444,410	1,423,604
R^2	0.091	0.299	0.371	0.130	0.280	0.396

Table A3: Model Descriptive Moments

This table reports market-level summary statistics used to discipline the structural model. A "market" is defined by year × borrower county × loan purpose (home purchase vs. refinance) × FICO bin. Borrowers are grouped into three risk tiers: Subprime (FICO < 660), Near Prime ($660 \leq$ FICO < 720), and Prime (FICO ≥ 720). Within each market, each lender's market share is computed. Lenders with market share $\geq 1\%$ are designated "major" lenders. Among major lenders, those above the median of lender-specific *local-share* (the fraction of loans in that market processed by loan officers located in the same county as the borrower) are classified as *Local*, and those below median as *Distant*. All lenders with market share $< 1\%$ form the *Fringe* category. For each borrower risk tier × lender type, we compute weighted averages across markets using county–year application volume as weights. For outcome variables — rejection rate, processing time, interest rate, and default rate — the reported standard deviations are *not* the dispersion of raw values. Instead, they reflect the standard deviation of *residualized* outcomes, constructed by first regressing each variable on borrower risk controls (FICO, LTV, DTI, AUS status, and their flexible polynomials) and then computing the dispersion of the resulting residuals at the market level. Thus, the means report average raw outcomes, while the parentheses contain the variability of risk-adjusted (residual) outcomes. This construction matches the moments used for model calibration. Columns (1)–(3) report home purchase markets; columns (4)–(6) report refinance markets. Standard deviations appear in parentheses.

		Home Purchase			Refinance		
Variable	Risk tier	Local	Distant	Fringe	Local	Distant	Fringe
Total market applications							
	Prime	5091	5091	5091	3789	3789	3789
		(6311)	(6311)	(6311)	(6973)	(6973)	(6973)
	Near prime	1266	1266	1266	1402	1402	1402
		(1696)	(1696)	(1696)	(2414)	(2414)	(2414)
	Subprime	257	257	257	469	469	469
		(356)	(356)	(356)	(694)	(694)	(694)
Number of active lenders							
	Prime	9.40	13.91	160.55	8.77	11.38	133.60
		(3.88)	(3.86)	(106.10)	(3.40)	(3.74)	(101.75)
	Near prime	10.16	14.15	90.15	9.77	11.69	92.68
		(4.01)	(4.22)	(71.36)	(3.84)	(4.40)	(81.03)
	Subprime	10.66	13.76	34.86	8.75	12.94	52.17
		(5.51)	(5.92)	(40.61)	(4.62)	(5.03)	(55.64)
Market share of lender type							
	Prime	0.32	0.43	0.25	0.31	0.44	0.25
		(0.09)	(0.12)	(0.10)	(0.09)	(0.13)	(0.11)
	Near prime	0.33	0.44	0.23	0.31	0.45	0.23
		(0.10)	(0.14)	(0.12)	(0.11)	(0.14)	(0.13)
	Subprime	0.34	0.49	0.15	0.30	0.49	0.18
		(0.15)	(0.18)	(0.14)	(0.15)	(0.16)	(0.13)

Descriptive Statistics (continued)

Variable	Risk tier	Home Purchase			Refinance		
		Local	Distant	Fringe	Local	Distant	Fringe
Share of loans handled by local officers							
	Prime	0.73	0.28	0.23	0.68	0.17	0.22
		(0.22)	(0.25)	(0.18)	(0.20)	(0.20)	(0.16)
	Near prime	0.76	0.27	0.26	0.67	0.09	0.24
		(0.20)	(0.26)	(0.19)	(0.20)	(0.16)	(0.16)
	Subprime	0.81	0.24	0.32	0.67	0.05	0.26
		(0.17)	(0.27)	(0.19)	(0.21)	(0.13)	(0.15)
Rejection rate (pp)							
	Prime	3.64	5.58	5.12	11.52	16.51	12.77
		(3.54)	(3.07)	(2.23)	(5.56)	(5.65)	(3.82)
	Near prime	7.36	9.95	9.76	23.92	28.49	22.38
		(4.52)	(3.90)	(3.34)	(7.90)	(6.98)	(4.91)
	Subprime	18.89	22.68	23.01	57.77	55.48	45.24
		(8.62)	(8.30)	(6.19)	(10.28)	(8.58)	(6.88)
Processing time (days)							
	Prime	47.54	51.55	45.93	46.39	40.99	44.69
		(10.76)	(10.94)	(6.94)	(6.48)	(6.65)	(5.09)
	Near prime	45.88	48.57	45.49	45.95	39.35	44.96
		(10.40)	(10.13)	(7.00)	(6.93)	(6.38)	(5.63)
	Subprime	42.76	46.52	41.73	32.70	31.15	37.44
		(10.56)	(12.67)	(6.69)	(7.41)	(5.83)	(5.71)
Interest rate (%)							
	Prime	4.34	4.30	4.36	4.27	4.31	4.27
		(0.32)	(0.32)	(0.28)	(0.36)	(0.36)	(0.36)
	Near prime	4.62	4.59	4.64	4.61	4.68	4.62
		(0.31)	(0.30)	(0.27)	(0.35)	(0.34)	(0.34)
	Subprime	4.94	4.86	4.97	4.92	5.00	4.96
		(0.32)	(0.31)	(0.24)	(0.38)	(0.32)	(0.31)
Two-year default rate (pp)							
	Prime	3.85	3.49	4.31	3.90	3.23	4.42
		(3.70)	(2.31)	(3.03)	(4.75)	(2.96)	(3.42)
	Near prime	10.11	10.16	11.22	9.14	7.57	10.15
		(7.26)	(5.68)	(6.02)	(9.47)	(6.08)	(7.13)
	Subprime	16.49	16.08	17.23	14.67	12.04	14.36
		(13.97)	(12.60)	(10.96)	(16.65)	(11.31)	(11.64)