Credit Without Proximity: Informational Frictions and Unequal Gains from Technology

Abstract

We study how technological advances that change the relative efficiency of local and remote screening affect information production, credit rationing, and allocative efficiency. Using administrative data linking loan officers to applications and loan outcomes, we document that informational frictions are first-order; local and remote loan officers differ sharply in screening precision and processing speed; and lenders' labor-allocation decisions respond strongly to local wage differentials, generating systematic spatial misalignment between mortgage demand and local underwriting capacity. Motivated by these patterns, we develop and estimate a structural model in which lenders compete in mortgage pricing and in labor markets for heterogeneous loan officers, borrowers with different unobserved default risks sort on prices, and screening precision varies with officer type. The model implies substantial baseline credit rationing—up to 15 percent in high-risk segments—with local officers eliminating roughly half while also reducing false approvals. A technology shock that increases the physical efficiency of remote work induces lenders to substitute away from local screening, reducing informational efficiency, raising pooled origination and expected defaults, and tightening rationing for marginal borrowers despite only modest reductions in rates.

Keywords: informational frictions, adverse selection, mortgages, banking

Financial intermediation hinges on the production and use of information about borrower risk. When lenders cannot perfectly observe borrower quality, they rely on costly screening technologies, giving rise to adverse selection, credit rationing, and disparities in credit access. Critically, the production of this information is itself an economic activity—mediated through labor, technology, and organizational design—whose supply need not be aligned with where information is most valuable.

Technological developments that expand remote underwriting capabilities have the potential to fundamentally change how lenders produce information. By relaxing geographic constraints in underwriting labor markets, these developments create a national market for screening services that could improve lenders' cost efficiency. However, this geographic decoupling of screening from borrowing locations risks degrading information quality, as remote officers may lack access to soft information available to local underwriters. This paper examines the equilibrium consequences of this ongoing transformation.

Using newly linked administrative data that connect the universe of U.S. mortgage loan officers to the applications they handle and the loans' subsequent performance, we show that informational frictions are first-order, local and remote officers differ sharply in screening precision, and lenders' labor allocation responds strongly to wage differentials. We build a structural model to quantify the consequences. The calibrated model suggests substantial baseline credit rationing in the US mortgage market, with local officers eliminating about *half* of the baseline credit rationing while also reducing false approvals. A technology shock that raises the physical efficiency of remote work leads lenders to substitute away from local labor, reducing informational efficiency, increasing pooled origination and expected defaults, and tightening rationing for marginal borrowers, even though rates fall only modestly. These results show that technological progress in underwriting can improve cost efficiency while simultaneously degrading information quality and altering the distribution of credit access.

We begin by documenting a systematic spatial misalignment of underwriting labor. High-demand counties, especially those with riskier borrowers, are disproportionately served by loan officers located elsewhere, while some low-demand counties have an oversupply of officers. This mismatch is not random: lenders' hiring decisions respond strongly to local wage conditions, placing relatively fewer loan officers in high-wage MSAs and thereby increasing the reliance on remote underwriting exactly where mortgage demand is most concentrated.

We then show that this allocation has first-order consequences for information production and credit access. Local loan officers rely less on hard information, reject fewer applications, and yet the loans they approve default less, indicating that proximity improves the precision of screening on unobservables and mitigates adverse selection. These informational gains are strongest for borrowers who appear risky on observables (low FICO, high DTI, high LTV), for whom local officers approve more loans without raising default, and they are complemented in refinance markets by a clear quantity advantage, as local officers process applications meaningfully faster. Together, the facts reveal that local underwriting capacity delivers both quality and quantity efficiency gains, but is least available in the markets where informational frictions are most severe and where borrowers rely most on soft information for credit access.

The reduced-form evidence shows that informational frictions are first-order, that proximity improves both screening precision and processing speed, and that lenders allocate labor in response to local wage conditions, which may not always align with the local value of information. These patterns highlight two efficiency channels through which proximity matters—higher-quality screening and faster processing—and show that both operate systematically across markets and borrower types. They also raise a set of central questions that require an industry equilibrium framework: How much credit rationing arises from lenders' reliance on lower-precision remote screening? How do labor-allocation decisions interact with pricing and borrower sorting to shape default risk and market composition? And how will emerging technologies that enhance the physical efficiency of remote underwriting—without improving its information content—alter the balance between cost savings, credit access, and aggregate risk? To answer these questions, we develop a structural model that embeds borrower selection, officer-specific screening precision, lender pricing, and endogenous labor allocation in a unified environment.

Addressing these questions requires a framework in which pricing, borrower selection, labor allocation, and screening technology interact endogenously across heterogeneous markets. Because posted interest rates shape both the volume and the risk composition of applicants, a model must allow for classic adverse-selection forces. Since screening precision depends on the mix of local and remote officers a lender hires, labor-market competition and local wage conditions must influence the supply of information—and thus the intensity of credit rationing—in each market. And because screening capacity is limited, lenders' hiring decisions must feed back into approval standards through congestion and capacity constraints. A model with these ingredients allows local labor supply shocks, technological changes in remote underwriting, or shifts in borrower composition to propagate across markets, screening margins, and risk outcomes. It also enables counterfactual experiments that quantify how much information is lost when lenders substitute away from local officers, how these changes shape credit access for different borrower segments.

In this model, lenders compete simultaneously in mortgage pricing and in labor markets for heterogeneous loan-officer types. Borrowers choose where to apply based on posted interest rates, generating both demand elasticities and selection on unobserved risk. Loan officers differ in their screening precision and processing capacity, and they choose among lenders and markets based on posted wages and idiosyncratic location preferences, which endogenously determines each lender's mix of local versus remote loan officers. After applications arrive, lenders allocate them across officer types subject to capacity constraints, and officers observe noisy signals of borrower risk; lenders approve applications when expected default losses fall below the posted net interest margin, consistent with institutional constraints that prevent repricing after soft information is collected. These elements jointly determine equilibrium prices, approval thresholds, labor allocations, and market shares across lenders. The structure delivers a rich set of cross-market interactions and selection patterns, and provides a unified environment for quantifying the informational value of local screening, the extent of credit rationing generated by remote underwriting, and the consequences of technological or labor-market shocks for credit access, default risk, and welfare.

The key parameters governing informational frictions and screening technology are calibrated using moments that directly reflect the reduced-form facts. The precision of local and remote screening signals is pinned down by matching the observed rejection-rate and default-rate advantages of local officers across product segments and borrower-risk tiers. Conditional on these informational parameters, the physical efficiency of each officer type—captured by bank-market-officer–specific labor-efficiency terms—is calibrated to match the observed local versus remote labor shares of Brick-and-Mortar, Fintech, and Fringe lenders. Borrower-type distributions and the correlation between default risk and price sensitivity are jointly disciplined by matching market-level origination rates, expected default rates, and the strength of the empirical relationship between residualized interest rates and default. Finally, market-specific aggregate shocks are chosen to align model-implied default levels with observed performance in each of the six mortgage segments. Together, these moments allow the model to recover the informational and physical productivity differences across officer types that the reduced-form evidence points to as central for understanding screening, credit rationing, and market composition.

The structural model delivers a clear characterization of how information frictions shape equilibrium credit allocation. We find that, in the baseline, lenders use superior information primarily to offset credit rationing rather than to avoid pooled originations with excessive default risk. Because interest rates are posted before high-quality information is collected, lenders rely almost entirely on the approval margin to manage risk. As a result, informational advantages translate disproportionately into reductions in "false" rejections: the model implies substantial rationing in equilibrium—up to 15 percent of applicants in the riskiest refinance segments—and local officers eliminate roughly *half* of this rationing. At the same time, local officers also reduce the much smaller amount of pooled origination, consistent with their default advantage in the data. These findings echo the reduced-form evidence: proximity expands credit supply for marginal borrowers while simultaneously improving ex-post performance, especially in markets where unobserved heterogeneity is most severe.

We next use the model to study how technological improvements to remote underwriting—captured as increases in the physical efficiency of distant officers without improving their information precision—reshape equilibrium credit provision. Because lenders' labor choices respond strongly to relative labor costs, even modest efficiency gains induce them to substitute away from local, information-rich screening. This reallocation reduces informational efficiency and affects both margins of origination. Rationing increases meaningfully across markets, particularly in subprime and near-prime purchase segments, where local screening plays the largest role ex ante. At the same time, the rise in pooled origination increases expected default rates. While posted interest rates fall slightly, consistent with a standard productivity shock, the savings are limited—typically no more than about $25 per month—because screening is already relatively physically efficient in the baseline. In short, improvements in physical efficiency do not translate into commensurate gains for borrowers once the induced loss of information is taken into account.

Finally, the model highlights how these informational losses interact with aggregate credit risk. At first, the marginal excess pooled originations are significantly closer to the default boundary, meaning that their default probabilities rise sharply in mild downturns, raising realized defaults relative to the baseline. However, in sufficiently severe downturns, default rates for these marginal loans saturate—exceeding the inflection point of the probit curve—and the loans that are rationed in the counterfactual become the ones whose risk increases most. Thus, the effect of the technology shock on realized default rates is non-monotonic in aggregate conditions. Benchmarking the model to GFC-era default rates, we find that informational frictions remain quantitatively meaningful at aggregate shock magnitudes comparable to or even larger than those experienced during the crisis. Overall, the model reveals an important economic trade-off: expanding remote underwriting capabilities lowers production costs but erodes the information that disciplines credit supply, increasing rationing and aggregate risk in ways that are economically significant yet only weakly offset by lower mortgage rates.

Literature Review. Out paper relates to a broad literature on information frictions and credit allocation. Foundational theories show how imperfect information generates credit rationing and inefficient borrower selection (Stiglitz and Weiss, 1981; Jaffee and Modigliani, 1969), while empirical work documents how lenders acquire and exploit information—especially soft information—to screen opaque borrowers (Petersen and Rajan, 1994, 2002; Berger and Udell, 1995; Berger et al., 2005; Agarwal and Hauswald, 2010; Liberti and Mian, 2008). A related set of studies highlights the role of organizational structure, monitoring, and loan-officer heterogeneity in shaping information production inside financial institutions (Hertzberg et al., 2010; Berg et al., 2020b; Frame et al., 2025). Our contribution is to show that information production is itself an equilibrium outcome: screening precision is produced with heterogeneous labor inputs, and lenders' allocation of local versus remote officers responds to local wage conditions rather than to where information is most valuable, generating systematic geographic misallocation in underwriting capacity. Within this broader agenda, an emerging literature studies how technology interacts with informational frictions, showing that improvements or disruptions in information production can meaningfully reshape credit allocation (Fuster et al., 2022; Berg et al., 2020a; Blattner and Nelson, 2021; Blattner et al., 2021, 2022). We empirically document and structurally quantify how the production of information itself—via heterogeneous local versus remote underwriting technologies—affects credit rationing, pooled origination, and the spatial distribution of credit.

Our paper also contributes to the literature examining how technological change reshapes production technologies and competitive structure in financial intermediation (Eizenberg, 2014; Kogan et al., 2017; Stulz, 2019; Vives, 2019; Tirole, 2020). A large body of work emphasizes that the traditional banking model is fundamentally local: branch networks and geographically embedded loan officers play a central role in producing information, expanding credit access, and supporting local economic activity.[1] Recent work shows that digital disruption—including the rise of fintech lenders and remote-processing technologies—is transforming this localized production structure and altering competition, cost efficiency, and the geography of credit supply (Buchak et al., 2018b; Fuster et al., 2019; Chen et al., 2019; Goldstein et al., 2019; Berg et al., 2022; He et al., 2021). Related papers document how these technologies reshape branch-based competition and financial inclusion (Jiang et al., 2022; Haendler, 2022; Koont, 2023; Narayanan et al., 2025). Our contribution is to show that technological change in underwriting reshapes not only lenders' cost structures but also the production of information itself. As remote-processing technologies improve, lenders substitute away from local, information-rich screening and toward cheaper, standardized remote labor—altering the precision with which borrower risk is assessed. We quantify how shifts in screening capacity translate into changes in credit rationing, influencing allocative efficiency and resulting in unequal gains from technology.

Finally, our paper contributes to a growing literature that applies structural IO tools to consumer finance and financial product markets. Prior work studies how competition shapes consumer welfare in mortgages (Allen et al., 2014, 2025; Agarwal et al., 2024; Buchak et al., 2018a; Benetton, 2021; Jiang, 2023), deposits (Egan et al., 2017; Xiao, 2020), payments (Wang, 2025; Whited et al., 2022), and credit cards (Nelson, 2025). A related and expanding literature in spatial IO analyzes how geographically segmented financial markets shape competitive conduct and allocative outcomes (Aguirregabiria et al., 2019; Ji et al., 2023; Maingi, 2025; Cox et al., 2021; D'Amico and Alekseev, 2024; Morelli et al., 2025). We contribute to this agenda by showing that informational precision varies sharply across space and is a first-order determinant of the marginal value of labor in underwriting—generating meaningful implications for efficiency, credit access, and aggregate risk.

[1] On the role of loan officers in information acquisition, see Hertzberg et al. (2010).

1 Institutional Background and Data

1.1 Mortgage Origination in the U.S.

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. Because lenders generally cannot raise rates after reviewing documents, approval decisions—rather than ex-post price adjustments—are the primary mechanism for responding to borrower risk.[2]

Loan officers play a central role in producing the information used in underwriting. They assemble income, asset, employment, and collateral documentation; underwriters rely on these materials and 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.

A distinguishing feature for our analysis is the contrast between *local* and *remote* loan officers. 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 solely on phone or digital communication and face greater frictions in resolving documentation issues. Because prices cannot be freely adjusted after underwriting, these differences in information-collection efficiency translate directly into differences in screening quality and processing time. Additional institutional detail is provided in Appendix A.

Conceptual Implications. The institutional features of mortgage origination—rate setting before information acquisition, limited scope for ex-post repricing, and the central role of loan officers in information production—create a sharp separation between pricing and approval that is absent from canonical industrial organization screening models. Because lenders cannot adjust prices to reflect borrower-specific risks revealed during underwriting, the approval decision becomes a distinct allocative margin shaped by the quality and effi-

[2]See Appendix A for regulatory details.

ciency of the loan officer's information-gathering activities.

1.2 Data

Our analysis combines three primary data sources: (i) the Nationwide Mortgage Licensing System and Registry (NMLS) loan officer database, (ii) the 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 complete 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 employment location for federally regulated institutions, explicitly reported as the loan officer's actual work address 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 to measure their 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 (e.g., FICO, loan-to-value, debt-to-income), applicant location, application and action dates, application outcomes (approved or denied), Automated Underwriting System (AUS) recommendations, and—critically for our analysis—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 underwrote 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 the linkage procedure in 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 two-year performance window, we restrict the main estimation sample to applications submitted during 2018–2019.

Combined dataset. After merging NMLS, confidential HMDA, and McDash, we obtain a loan-level panel in which we observe: the loan officer who processed each application, the geographic location of that loan officer, rich borrower- and loan-level observables at application, lender identity and local market characteristics, action taken (approval or denial), final interest rate for originated loans, and subsequent loan performance for matched originations. This combined dataset enables us to measure geographic misalignment between underwriting labor and mortgage demand, quantify screening differences across local and remote loan officers, and evaluate how these differences translate into credit access and ex-post default outcomes. Appendix B provides additional details on sample construction, data cleaning, and the standard filters used in the literature.

2 Reduced Form Facts

This section documents new facts on how lenders allocate local versus remote loan officers and how these choices shape screening quality, information production, and access to credit.

2.1 Fact 1: Spatial Misallocation of Underwriting Labor

We begin by comparing the spatial distribution of mortgage applications with the locations of loan officers. Panel A of Figure 1 plots the number of registered mortgage loan officers per hundred mortgage applications across U.S. counties. The map reveals substantial geographic dispersion in this loan-officer coverage ratio: many high-demand counties have a low number of loan officers relative to the volume of applications they receive, while some low-demand counties exhibit relative oversupply, hosting more loan officers than their local demand would predict.

To quantify this geographic misalignment more formally, we adapt classic measures of residential segregation from sociology and labor economics, recently applied to banking by Aguirregabiria et al. (2019) to capture local deposit–loan imbalances. For each county, we compute its share of all registered mortgage loan officers nationwide and its share of national mortgage applications. The misalignment index is defined as the difference between these two shares, capturing whether a county has more or fewer loan officers than would be expected given its contribution to national credit demand. Panel B of Figure 1 visualizes this index and shows pronounced imbalances: many counties exhibit positive misalignment values, indicating that local demand exceeds the available loan officer supply.

Figure 2 links this misalignment to local loan demand. Counties in the highest demand deciles exhibit the largest positive wedges between their share of national applications and their share of loan officers, reinforcing the pattern that local demand and local underwriting capacity are systematically misaligned.

Overall, the cross-sectional patterns reveal a clear mismatch between where mortgage demand is concentrated and where loan officers are located. As a result, applications originating in counties with high demand relative to local underwriting personnel are more likely to be processed by loan officers located outside the applicant's county.

2.2 Fact 2: Local Loan Officers Improve Informational Efficiency

We next show that local loan officers improve informational efficiency in mortgage underwriting. We establish this in three steps. First, we document that unobserved borrower heterogeneity is a first-order feature of the market by showing a strong positive relationship between residualized interest rates and residualized default, consistent with classic adverse selection. Second, we demonstrate that local officers make greater use of soft information: hard-information variables explain substantially less of their rejection decisions than for non-local officers. Third, we show that local officers' approval decisions are more predictive of ex-post performance: they reject fewer applications yet the loans they approve default less, and this pattern holds even though the borrowers they approve often face slightly higher posted rates.

2.2.1 Adverse selection: unobserved borrower risk is economically meaningful

Figure 3 provides evidence of adverse selection by plotting the relationship between residualized interest rates and residualized default rates separately for loans processed by local and non-local officers. Both interest rates and default outcomes are residualized using a comprehensive set of borrower, loan, lender, and geographic controls, so the remaining variation captures unobserved borrower risk and lenders' pricing of that risk. In both panels, we observe a clear positive relationship: borrowers paying higher-than-predicted rates are more likely to default ex post.

This pattern is consistent with the classic adverse-selection framework of Stiglitz and Weiss (1981): among otherwise observationally similar loans, those that end up with higher interest rates are drawn from a pool with worse unobserved risk. The fact that higher residualized rates are *not* accompanied by lower residualized default indicates that pricing adjustments are incomplete and that unobserved heterogeneity plays a substantial role in shaping loan outcomes. In economic terms, moving to a higher posted rate does not simply compensate for risk and restore default to a flat baseline; instead, it tilts the composition of borrowers toward those with worse unobservables, raising both price and default. This establishes the core informational problem that local officers may be particularly well-equipped to resolve.

2.2.2 Adverse selection: unobserved borrower risk is economically meaningful

Figure 3 provides evidence of adverse selection by plotting the relationship between residualized interest rates and residualized default rates separately for loans processed by local and non-local officers. Both interest rates and default outcomes are residualized using a

comprehensive set of borrower, loan, lender, and geographic controls, so the remaining variation captures unobserved borrower risk and lenders' pricing of that risk. In both panels, we observe a clear positive relationship: borrowers paying higher-than-predicted rates are more likely to default ex post.

This pattern is a hallmark of adverse selection. In a frictionless market with fully informative hard information, lenders would set higher interest rates precisely for borrowers who are riskier *on observables*. Once we control for all such observables, any remaining variation in rates should be driven by noise; importantly, higher residualized rates should *not* systematically predict higher residualized default.

Under adverse selection, however, raising the interest rate affects not only the price of a loan but also the composition of borrowers who apply and are approved. Borrowers with better unobserved characteristics tend to be more sensitive to price—they are more willing to walk away when rates are high—while borrowers with worse unobserved characteristics remain in the pool. As a result, loans with higher residualized rates disproportionately attract (or get approved for) borrowers with worse unobservable risk.

Economically, the positive relationship between residualized interest rates and residualized default reflects classic adverse selection. Because both variables are residualized using the same rich set of controls, the remaining variation captures unobserved borrower risk and lenders' pricing of that risk. If lenders fully inferred and priced all such unobservables, then—conditional on observables—higher residualized rates would not be associated with higher residualized default; pricing would already incorporate all relevant risk differences.

Instead, borrowers who pay higher-than-predicted rates are also more likely to default relative to their predicted default probability. This indicates that pricing does not fully adjust for unobserved risk, and that increases in posted rates tilt the applicant pool toward riskier borrowers. Such composition effects—higher price drawing in worse unobservables—are a hallmark of adverse-selection models such as Stiglitz and Weiss (1981). This establishes the core informational problem that local officers may be particularly well-equipped to resolve.

2.2.3 Local officers make greater use of soft information

Table 1 compares the explanatory power of hard-information variables in predicting loan rejection across applications handled by local versus non-local officers. Across six increasingly saturated specifications—including rich borrower and loan controls, month fixed effects, county and lender fixed effects, and eventually loan officer fixed effects—the R^2 for local officers is consistently and substantially lower than for non-local officers.

For example, in the baseline specification with borrower and loan controls only, the R^2 is 0.217 for non-local officers and only 0.191 for local officers. When we add county-month and lender-month fixed effects (Specification 5), the gap widens further: 0.438 for non-local versus 0.340 for local. The pattern is robust across 100 bootstrap replications: the average R^2 difference is large, negative, and statistically significant in every specification.

These results indicate that observable borrower and loan characteristics explain a smaller share of rejection decisions made by local officers. This implies that local underwriters rely more heavily on additional, non-codifiable information—"soft information"—that is not captured by standard hard-information variables, such as FICO, LTV, DTI, income, or loan type.

2.2.4 Local officers make more informative screening decisions

We next show that local loan officers' approval decisions are more predictive of ex-post performance, consistent with more accurate screening on unobservables. In specific, local officers reject fewer applications, yet the loans they approve perform better ex post. In the US residential mortgage market, lenders post interest rates ex ante, borrowers choose where to apply, and loan officers then decide which applications to approve. Loan officers do not renegotiate rates at the approval stage, so any informational advantages of local officers must operate through the approval margin rather than through pricing.

Rejection. To study approval decisions, we estimate regressions of the form

$$\text{Reject}_{i\ell ct} = \beta^R \,\text{Local}_{i\ell ct} + \gamma^{R\prime} X_{i\ell ct} + \delta^R_{c\tau} + \lambda^R_{\ell\tau} + \mu^R_o + \varepsilon^R_{i\ell ct}, \tag{1}$$

where $\text{Reject}_{i\ell ct}$ is an indicator equal to 1 if application i at lender ℓ in county c and month τ is rejected and 0 otherwise; $\text{Local}_{i\ell ct}$ indicates that the reviewing officer o is local to the applicant; $X_{i\ell ct}$ is a saturated set of borrower and loan characteristics; $\delta^R_{c\tau}$ are county-month fixed effects; $\lambda^R_{\ell\tau}$ are lender-month fixed effects; and μ^R_o are loan officer fixed effects in the most saturated specifications.

Table 2 reports estimates of equation (1) for home purchase and refinance loans. For home purchase loans, the coefficient β^R on the Local indicator ranges from -1.88 to -0.20 percentage points across specifications (1)–(3), all statistically significant, indicating that local officers reject fewer applications conditional on observables and fixed effects. For refinance loans, the pattern is similar, with β^R ranging from -7.90 to -0.66 percentage points, again highly significant.

Ex-Post Performance. We then relate local approval decisions to ex-post performance by estimating

$$\text{Default}_{i\ell ct} = \beta^D \, \text{Local}_{i\ell ct} + \gamma^{D\prime} X_{i\ell ct} + \delta^D_{c\tau} + \lambda^D_{\ell\tau} + \mu^D_o + \varepsilon^D_{i\ell ct}, \tag{2}$$

where $\text{Default}_{i\ell ct}$ is an indicator equal to 1 if the loan misses two consecutive payments within two years of origination and 0 otherwise, and the remaining covariates and fixed effects mirror those in equation (1).

Table 3 presents estimates of equation (2). For home purchase loans, the coefficient β^D on Local is negative and statistically significant in all specifications, ranging from -0.33 to -0.46 percentage points. Thus, even though local officers approve more applications, the loans they approve are *less* likely to default. For refinance loans, β^D becomes small and statistically insignificant once we include lender-month and loan officer fixed effects, consistent with refinance borrowers forming a more homogeneous pool in which soft information plays a more limited role.

Taken together, equations (1) and (2) and the estimates in Tables 2 and 3 imply that local officers approve *fewer* high-risk borrowers and *more* low-risk borrowers, conditional on observables. Their approval decisions are therefore more informative about true borrower quality than those of non-local officers.

Interest Rates. Moreover, although local officers have lower rejection rates, the loans they approve carry slightly higher *posted* interest rates. To examine this formally, we estimate:

$$\text{Rate}_{i\ell ct} = \beta \, \text{Local}_{i\ell ct} + \gamma' X_{i\ell ct} + \delta_{c\tau} + \lambda_{\ell\tau} + \mu_o + \varepsilon_{i\ell ct}, \tag{3}$$

where $\text{Rate}_{i\ell ct}$ is the posted interest rate for loan i reviewed by officer o at lender ℓ in county c and month τ; $\text{Local}_{i\ell ct}$ indicates that the reviewing officer is local to the applicant; $X_{i\ell ct}$ includes borrower and loan observables; $\delta_{c\tau}$ are county-month fixed effects; $\lambda_{\ell\tau}$ are lender-month fixed effects; and μ_o are loan officer fixed effects in the most saturated specification.

Table 4 reports estimates of equation (3) across specifications (1)–(6), which sequentially add these fixed effects. For home purchase loans, the coefficient on the Local indicator ranges from 0.033 percentage points in specification (1) to a small negative value, −0.002, once officer fixed effects are included in specification (3). For refinance loans, the Local coefficient remains positive and statistically significant across all specifications.

Because lenders post interest rates *before* loan officers review applications and observe soft information, these coefficients cannot reflect officer-level pricing. Instead, they reflect differences in the observable risk characteristics of borrowers who choose to apply to lenders relying more on local officers. Local officers tend to approve borrowers who appear riskier on observables (and thus face higher posted rates), yet—as shown in Table 3—these borrowers perform *better* ex post. This combination of slightly higher posted rates and lower realized default is precisely what we would expect if local officers identify borrowers whose true risk is lower than what hard-information variables imply.

Joint Behavior. Finally, Figures 4–5 and 5 illustrate how these informational advantages manifest in the joint behavior of rejection, pricing, and default.

Figure 4 plots the relationship between residualized rejection rates and residualized default rates, separately for applications handled by local versus non-local officers. Each point represents a lender–county average, so the figure can be interpreted as comparing how different approval standards translate into ex post performance across local and non-local production technologies. For non-local officers, we see a steep negative gradient: in counties and lenders where residualized rejection is higher, residualized default is substantially lower. This pattern is consistent with a screening technology that relies heavily on hard cutoffs: to reduce default, non-local officers must tighten observable criteria and reject a larger share of applications. By contrast, the local-officer gradient is much flatter. Local officers operate at lower rejection levels, yet the associated increase in residualized default is modest. Economically, this pattern is exactly what we would expect if local officers possess better information about borrower quality: they can approve applicants who look marginal on observables without generating large increases in default.

Figure 5 further decomposes this relationship by separating markets into those with above-median versus below-median residualized interest rates. In the high-rate subsample, where adverse selection documented in Figure 3 is most pronounced, the contrast between local and non-local officers is stark. For non-local officers, higher residualized rejection is strongly associated with lower residualized default, indicating that they must be very strict on the approval margin to keep losses in check. For local officers, the slope is considerably flatter: even in high-rate markets, they can maintain relatively low rejection rates without a commensurate increase in default. In low-rate markets, where adverse selection is weaker, the difference in slopes between local and non-local officers is smaller. Together, these patterns suggest that the informational advantage of local officers is most valuable precisely where unobserved risk is most severe.

Taken together, the regression evidence in Tables 2–4 and the patterns in Figures 3–5 paint a consistent picture. Local officers operate at looser observable approval standards, especially in segments with worse observable risk and higher rates, but they do so in a way that preserves or improves ex post performance. This is precisely the signature of underwriting that incorporates soft information: proximity allows officers to better distinguish safe from risky borrowers within the same hard-information cell, mitigating adverse selection and improving the mapping from approval decisions to true borrower quality.

2.3 Fact 3: Local Loan Officers Are More Quantity-Productive

We next show that local loan officers are more *quantity-productive*: conditional on observables, they process local mortgage applications faster than non-local officers. This fact captures a distinct dimension of performance relative to informational efficiency. Whereas Fact 2 focuses on how proximity improves the *quality* of screening, Fact 3 focuses on how proximity affects the *speed* with which loan officers move applications through the underwriting pipeline. In our model, we treat these as separate technological dimensions. This distinction will be crucial when we study technological change that improves the quantity productivity of remote workers without improving their informational efficiency.

To quantify quantity productivity, we estimate regressions of the form

$$\text{Time}_{i\ell ct} = \beta^T \, \text{Local}_{i\ell ct} + \gamma^{T\prime} X_{i\ell ct} + \delta^T_{c\tau} + \lambda^T_{\ell\tau} + \mu^T_o + \varepsilon^T_{i\ell ct}, \tag{4}$$

where $\text{Time}_{i\ell ct}$ is the number of days between the application date and the action date for loan i at lender ℓ in county c and month τ; $\text{Local}_{i\ell ct}$ indicates that the reviewing officer o is local to the borrower; $X_{i\ell ct}$ is the full set of borrower and loan characteristics; $\delta^T_{c\tau}$ are county-month fixed effects; $\lambda^T_{\ell\tau}$ are lender-month fixed effects; and μ^T_o are loan officer fixed effects in the most saturated specification. We estimate equation (4) separately for home purchase and refinance loans, and restrict the sample to originated loans so that $\text{Time}_{i\ell ct}$ is well-defined.

The estimates in Table 5 reveal an asymmetry between home purchase and refinance loans in how proximity affects processing speed. Panel A shows that for home purchase loans, local officers appear substantially faster in the baseline specification: conditional on borrower observables and county-month fixed effects, the Local coefficient is -2.38 days. However, once lender-month fixed effects are added, the coefficient becomes economically small and statistically insignificant. A similar pattern appears when replacing the Local indicator with log distance. These patterns suggest that much of the raw home-purchase advantage reflects lender-level workflow differences or officer-specific heterogeneity, rather than a systematic productivity premium from geographic proximity.

Panel B shows a markedly different pattern for refinance loans. Here, the proximity effects are not only larger but also persist even in the most saturated specifications. In the baseline, local officers process refinance applications about 0.5 days slower than non-local officers, but once lender-month fixed effects are included, the sign reverses and the Local coefficient becomes strongly negative: -1.65 days in column (2) and -1.17 days in column (3). These estimates indicate that, within the same lender and month—and even controlling for officer-specific productivity—local officers process refinance loans more than a full day faster than non-local officers. Specifications using log distance yield the same conclusion: conditional on lender-month and officer fixed effects, greater borrower–officer distance is associated with meaningfully longer processing times.

Taken together, the results show that proximity generates a robust quantity-productivity advantage primarily in the refinance market, not in the home-purchase market. This asymmetry is consistent with refinance workflows being more standardized and time-sensitive, with fewer property-specific frictions but a larger role for borrower communication and document coordination. In such environments, even modest frictions introduced by physical distance can slow down processing in ways that are not absorbed by officer-level or lender-level controls. The distinction between these quantity-productivity gains and the informational-efficiency gains documented in Fact 2 will be central for interpreting the effects of technological change in our model.

2.4 Fact 4: High-Credit-Risk Borrowers Benefit More

The results in Fact 2 show that, on average, local loan officers reject fewer applications while approving loans that perform better ex post. This pattern implies that local officers possess informational advantages that improve screening along the approval margin. A natural next question is *which borrowers benefit most from these informational gains.* The adverse-selection patterns documented in Fact 2 indicate that unobserved borrower heterogeneity is a first-order feature of mortgage underwriting across all market segments. If local officers rely more effectively on soft information, their informational edge should help them distinguish better from worse risks within observable risk categories throughout the credit spectrum.

Borrowers who appear risky on *hard* information—low FICO, high DTI, or high LTV—are natural candidates. These observable measures become less discriminating in high-risk regions, because even within the same FICO, DTI, or LTV bucket, borrower creditworthiness varies substantially. For example, borrowers with FICO scores just below 620 or DTIs just above 43 may differ widely in income stability, job tenure, informal support networks, or local market conditions—all factors not captured in standard hard-information variables.[3] Because hard information is least precise in these regions, screening decisions should rely more heavily on soft information. Thus, if local officers indeed possess better soft-information signals, their relative advantage should be largest in these observable high-risk segments. Table 6 formally test this by estimating how the effect of local officers varies across three measures of observable borrower risk: Subprime (FICO $<$ 670), High DTI (DTI $>$ 43), and High LTV (LTV $>$ 80).

[3]As shown in Figure A2, both rejection rates and unconditional default rates rise steeply as borrowers move into higher-risk observable categories, reflecting wider dispersion in true underlying risk.

Panel A reports heterogeneous rejection regressions of the form

$$\text{Reject}_{i\ell ct} = \beta\,\text{Local}_{i\ell ct} + \theta\,\text{Local}_{i\ell ct} \times \text{HighRisk}_{i\ell ct} + \Gamma' X_{i\ell ct} + \delta_{c\tau} + \lambda_{\ell\tau} + \varepsilon_{i\ell ct}. \tag{5}$$

Across all three dimensions, the interaction term θ is large, negative, and precisely estimated. For instance, in the subprime specification, the coefficient on Local $\times$ Subprime is -2.93 percentage points—several times larger than the main effect of Local. The High-DTI and High-LTV interactions are similarly negative and sizable. These results show that local officers relax rejection probabilities *far more* for borrowers who appear risky on observables, consistent with their ability to distinguish safer from riskier applicants within the most heterogeneous groups.

Panel B turns to ex-post loan performance:

$$\text{Default}_{i\ell ct} = \beta^D\,\text{Local}_{i\ell ct} + \theta^D\,\text{Local}_{i\ell ct} \times \text{HighRisk}_{i\ell ct} + \Gamma^{D\prime} X_{i\ell ct} + \delta^D_{c\tau} + \lambda^D_{\ell\tau} + \varepsilon^D_{i\ell ct}. \tag{6}$$

Local officers' more permissive approval behavior for observably risky borrowers does *not* lead to higher default. In the DTI and LTV splits, the interaction coefficients θ^D are negative and statistically significant, indicating that local officers disproportionately approve borrowers who perform *better* than expected given their hard-information profiles. Even in the subprime split—where statistical power is lower—the interaction term is negative and economically meaningful. Thus, the extra approvals granted by local officers to high-risk observable groups are primarily directed toward safer borrowers within those groups.

Panel C examines posted interest rates. Because interest rates are posted *before* loan officers observe soft information, these coefficients reflect borrower sorting rather than officer-level pricing. High-risk borrowers appear riskier on observables and thus face higher posted rates when approved by local officers. Taken together with Panel B, these results are consistent with local officers identifying borrowers whose posted rates overstate their true underlying risk.

Figure 6 provides complementary evidence by comparing rejection and default outcomes for loans handled by local and non-local officers within discrete bins of FICO, DTI, and LTV. Across all three dimensions, local officers consistently reject fewer applications within every observable-risk bucket, with the largest gaps appearing in the high-risk bins (e.g., FICO < 620, DTI $> 50\%$). These differences are most pronounced in *home-purchase* markets, where borrower- and property-specific soft information plays a central role in underwriting.

In refinance markets, the patterns differ: although the *overall* rejection gap between local and non-local officers is larger, the gap varies far less across FICO, DTI, and LTV buckets. This flatter cross-sectional pattern indicates that local officers' informational advantages in refinance markets operate more uniformly across observable risk categories, rather than being concentrated in high-risk segments. This suggests that unobserved heterogeneity in refinance applications is more evenly distributed throughout the credit spectrum, and local officers' ability to extract soft information provides value across all risk levels.

Crucially, despite these heterogeneous rejection patterns, default outcomes remain similar or lower for loans screened by local officers within every observable-risk bucket. Thus, even when local officers approve substantially more borrowers in high-risk categories, these additional approvals do not translate into higher ex-post default.

In sum, the cross-sectional evidence demonstrates that local officers' informational advantages are strongest exactly where unobserved borrower heterogeneity is most severe. Local officers approve more observably risky borrowers *without increasing* default, indicating that proximity enables lenders to identify high-quality borrowers who appear marginal on hard-information metrics. These patterns reinforce the interpretation that soft information is particularly valuable in high-risk segments and that geographic proximity plays a central role in mitigating adverse selection.

2.5 Fact 5: Lender Choices of Labor Input

Finally, we document that lenders' allocation of local versus remote loan officers is systematically related to local labor-market conditions. Table 7 shows that lenders place relatively fewer loan officers in MSAs with higher wages, generating greater geographic mismatch between where their mortgage applications originate and where their loan officers reside. To measure this mismatch, we construct a *Misalignment Index* for each lender–MSA–year observation,

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

where $\frac{m_{jkt}}{M_{jt}}$ is the share of lender j's total applications coming from MSA k, and $\frac{l_{jkt}}{L_{jt}}$ is the share of the lender's loan officers located there. A positive value indicates an *under-supply* of local loan officers relative to the lender's local demand.

We estimate the following regression specification:

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

where Wage_{kt} is either the mortgage-loan-officer hourly wage or the finance-sector hourly wage in MSA k; γ_t denotes year fixed effects; γ_{jt} denotes lender-by-year fixed effects; and $\mathbf{X}_{kt}$ includes MSA-level controls (employment rate, income per capita, earnings per job). Standard errors are clustered at the lender level.

Across all specifications, higher local labor costs are strongly associated with larger misalignment. In high-wage MSAs, lenders devote a systematically smaller fraction of their loan-officer workforce than would be proportional to those markets' share of applications. These results indicate that local labor costs play a meaningful role in shaping lenders' decisions over whether to staff a market with local officers or rely more heavily on remote production capacity.

Table 8 describes how the composition of lenders and the allocation of local versus remote loan officers vary across borrower risk tiers. "Local" lenders are defined as major lenders (market share $\geq$1%) whose within-market share of loans handled by local loan officers is above the median among major lenders in that market; those below the median are classified as "Distant" lenders. The statistics indicate that subprime markets—especially in home-purchase segments where underwriting relies more on soft information—feature a greater reliance on lenders that use local loan officers. Specifically, the number of Local lenders rises noticeably as borrower risk increases in home-purchase markets. The average number of Local lenders is about 13% higher in subprime markets relative to that in prime markets. This shift indicates that a larger share of major lenders active in subprime home-purchase segments rely more heavily on local officers. By contrast, refinance markets show little systematic movement. A similar pattern appears in local-officer shares: in home-purchase markets, Local lenders increase their local-officer usage from 0.73 (prime) to 0.81 (subprime), whereas refinance markets exhibit only minimal variation across risk tiers for both Local and Distant lenders.

2.6 Discussion

The reduced-form evidence establishes three key patterns: informational frictions are pervasive in mortgage markets, local loan officers mitigate these frictions through superior screening and processing efficiency, and lenders' labor-allocation decisions respond to local wage conditions in ways that generate systematic geographic misalignment between underwriting capacity and mortgage demand.

These findings reveal two distinct dimensions of efficiency gains from proximity. First, proximity improves *quality efficiency*: local officers extract more predictive soft information and achieve better default outcomes despite approving more marginal borrowers, with these informational advantages being particularly valuable in segments where unobserved heterogeneity is most economically significant. Second, proximity improves *quantity efficiency*: local officers process applications meaningfully faster, consistent with proximity mitigating coordination and documentation frictions. The relative importance of these efficiency gains varies across market contexts, but both dimensions contribute to the overall value of local underwriting capacity in mortgage origination.

Finally, the reduced-form evidence underscores the key frictions that a structural model must capture and motivates the need for an equilibrium framework. In the presence of informational frictions, imperfect information naturally generates *credit rationing*: when lenders cannot perfectly infer borrower risk, they tighten approval standards rather than raise rates, leaving high-friction segments underserved. Because screening precision depends on the lender's chosen mix of local and remote officers, labor-allocation decisions directly affect the intensity of credit rationing.

Inefficient screening also has downstream consequences: when lenders misidentify borrower risk, losses are borne by lenders or, in the case of GSE-insured loans, by taxpayers. Moreover, because borrowers choose where to apply and lenders compete in both pricing and screening intensity, the spatial allocation of high-precision (local) versus low-precision (remote) screening technologies can reshape the composition of borrowers served by different lenders and potentially produce spillovers across markets. These considerations motivate the structural model that follows, which embeds borrower selection, officer-specific screening precision, lender pricing, and endogenous labor decisions to quantify the efficiency, credit-access, and welfare implications of technological and labor-market changes in mortgage underwriting.

3 Structural Model

The reduced-form evidence shows that informational frictions are central, that screening precision differs sharply between local and remote loan officers, and that lenders' labor-allocation decisions respond strongly to local wage differences. As a result, lenders under-provide high-precision (local) screening in the very markets where informational frictions are most severe. These patterns motivate the structural model that follows, which integrates the mortgage product market with the labor market for loan officers and highlights that labor input matters because officer types differ in their ability to resolve informational frictions. This framework allows us to quantify how equilibrium interactions between pricing, screening, and labor allocation shape credit rationing and welfare, and sets the stage for counterfactual analyses of how changes in lending technology alter these outcomes.

3.1 Model Setup

3.1.1 Overview

The model integrates four interconnected components that jointly determine mortgage market equilibrium: lender competition, borrower demand, loan officer labor supply, and screening technology. Lenders compete strategically in both product and labor markets, making simultaneous pricing and hiring decisions while anticipating competitors' actions and equilibrium outcomes.

Borrowers choose among lenders based on posted interest rates and expected approval probabilities, and their price sensitivity influences both the level of demand and the risk composition of applicants, generating a classic adverse-selection channel.

Loan officers choose which lender and market to work in based on wages and idiosyncratic location preferences. Officer types differ in screening precision and processing capacity, so a lender's wage-setting decision affects both the quality and quantity of its underwriting capacity.

Finally, after borrowers apply, lenders assign applications to their hired officer types. Officers receive noisy signals of true borrower risk, and lenders approve applications if expected default costs lie below the posted margin. Screening precision, which varies across officer types, determines how well lenders can separate high- and low-risk borrowers.

These components interact through a Nash equilibrium where lenders simultaneously choose interest rates and wages. They take into account strategic interactions with competing lenders, borrower demand responses to posted rates, loan officer labor supply responses to posted wages, capacity constraints from limited screening resources, and information frictions in the screening process. The interaction of these forces shapes equilibrium pricing, approval standards, credit rationing, and the spatial allocation of credit across markets.

Several important economic mechanisms emerge from this framework. Adverse selection occurs because posted rates affect the risk composition of applicants, creating a classic lemons problem where higher rates attract riskier borrowers. Congestion externalities arise from limited screening capacity, creating congestion costs that affect pricing and market structure. Information value is created by heterogeneous screening technologies, influencing lenders' hiring decisions and wage differentiation. Spatial competition results from differences in screening precision between local and remote officers, creating spatial competition frictions that affect market outcomes. Finally, lenders exercise two-sided market power in both product markets through interest rates and labor markets through wages, with interactions between the two.

3.1.2 Primitives and Timing

A market is defined by a unique combination of credit quality, product type, county, and year-quarter. The economy consists of a collection of independent markets indexed by $i \in \mathcal{N}$. In each market, there is a mass $\mathcal{M}_i$ of potential borrowers, a set of lenders $\mathcal{J}_i$, and a set of loan officer types $\mathcal{K}$.

The timing within each market is as follows:

1. **Pricing and Hiring:** Each lender j simultaneously posts an interest rate r_{ij} for all loans in market i and a vector of type-specific wages $\{w_{ijk}\}_{k\in\mathcal{K}}$ to hire loan officers. These decisions are made taking as given other lenders' simultaneous decisions, forming

a Nash equilibrium in prices and wages. Lenders anticipate how their choices affect borrower demand and labor supply responses.

2. **Application:** Borrowers observe all posted rates $\{r_{ij}\}_{j \in \mathcal{J}_i}$ across all active lenders in the market. Each borrower η has a continuous latent risk type x_η summarizing all default-relevant characteristics not observable at pricing stage, observable hard-information characteristics such as FICO, LTV, and DTI embedded in demand shifter ξ_{ij}, and idiosyncratic preferences for lenders captured by random utility shocks. Borrowers choose a single lender to apply to or select the outside option based on posted interest rates r_{ij}, expected probability of being approved which depends on their risk type and lender's screening technology, observable lender characteristics and demand shifters ξ_{ij}, and idiosyncratic preferences. At this stage, lenders observe only the distribution of x_η in the market, coarse self-reported characteristics and public data such as FICO, LTV, and DTI, and the mass of applications $\mathcal{M}_i s_{ij}(r_{ij})$ they receive. Verified income documents, appraisals, employment checks, and third-party verifications are not yet available at this pricing stage.

3. **Labor Market Clearing:** Loan officers observe posted wages $\{w_{ijk}\}$ across all lenders and markets, and choose where to work based on posted wages w_{ijk}, idiosyncratic location and lender preferences, their type-specific screening precision σ^2_{ik}, and processing capacity e_{ijk}. This determines the mass l_{ijk} of type-k officers working for each lender j. Lenders compete for scarce screening capacity through their wage offers.

4. **Screening:** Lenders randomly assign incoming applications to hired loan officers using assignment shares $\mathcal{S}_{ijk}$. Assignment must respect capacity constraints such that $\mathcal{M}_i s_{ij}(r_{ij})\mathcal{S}_{ijk} \leq l_{ijk} e_{ijk}$ for all k. An officer of type k screening applicant η observes a noisy signal $\hat{x}_\eta = x_\eta + \tilde{x}_{\eta k}$ where $\mathbb{E}[\tilde{x}_{\eta k}] = 0$ and $\mathrm{Var}(\tilde{x}_{\eta k}) = \sigma^2_{ik}$. Signal precision varies by officer type, with local officers having lower σ^2_{ik} indicating more precise signals, while remote officers have higher σ^2_{ik} indicating noisier signals. This stage captures the gathering of verified documentation, appraisals, employment checks, and third-party verifications that occurs after application submission.

5. **Origination:** For each application, lender j calculates the expected default cost conditional on the observed signal and officer type. The lender approves the application if and only if $\mathbb{E}\left[d_i(x_\eta) \mid \hat{x}_\eta, k\right] \leq r_{ij} - f_{ij}$ where f_{ij} is the per-loan funding cost. This

condition defines a cutoff set $\hat{\Delta}_{ij}(r_{ij}, k)$ in the signal space. Consistent with TRID rules, lenders cannot reprice interest rates after underwriting reveals new information, making the approval decision the only margin of adjustment. Competition affects origination standards through its impact on rates, wages, and screening capacity.

6. **Payoffs and Equilibrium:** Approved loans generate utility for borrowers from obtaining mortgage financing and profits for lenders equal to $(r_{ij} - f_{ij} - d_{ijk})$ per originated loan, net of realized defaults. Lenders pay labor costs either as $\sum_{k \in \mathcal{K}} w_{ijk} l_{ijk}$ or equivalently as $A_{ijk}^{-1} \tilde{l}_{ijk}^{1+\frac{1}{\varrho}}$ from the labor supply problem. Denied applicants receive their outside option. Market equilibrium requires that all agents' decisions are mutually consistent and optimal given others' actions, forming a coherent equilibrium across all markets and decision stages.

Information Environment The model's timing captures a key institutional feature of U.S. mortgage origination: *pricing occurs before high-quality information is collected.* Unlike canonical screening models where lenders can reprice after observing private signals, here the TRID rules prevent expost repricing. This makes approval decisions, rather than risk-based pricing, the primary allocative margin. The precision of screening technology σ_{ik}^2 therefore becomes central to equilibrium outcomes, as it determines lenders' ability to separate risky from safe borrowers after rates are already set.

3.1.3 Borrower Demand

Each borrower η in market i chooses which lender to apply to (or selects the outside option of not applying) based on posted interest rates and lender characteristics. Borrower η receives the following random indirect utility from applying to lender j:

$$u_{ij\eta} = -\theta_\eta r_{ij} + \xi_{ij} + \epsilon_{ij\eta}, \tag{7}$$

where r_{ij} is the interest rate posted by lender j in market i, θ_η captures borrower η's sensitivity to interest rates, ξ_{ij} represents lender-market specific characteristics that affect borrower preferences, and $\epsilon_{ij\eta}$ is an idiosyncratic utility shock. The demand shifter ξ_{ij} can be decomposed as $\xi_{ij} = \zeta_{ij} + \tilde{\xi}_{ij}$, where ζ_{ij} captures consumer preferences for local branches (which are necessary to hire local loan officers) and $\tilde{\xi}_{ij}$ represents other residual demand factors.

Each borrower chooses the lender that provides the highest utility. The market share for lender j among borrowers with price sensitivity θ is given by:

$$s_{ij}(\theta, r_{ij}) = \mathbb{P}_\eta\left(\bigcap_{k\in\mathcal{J}}\left\{-\theta_\eta(r_{ij} - r_{ik}) + (\xi_{ij} - \xi_{ik}) \geq \epsilon_{ik\eta} - \epsilon_{ij\eta}\right\}\right). \tag{8}$$

The overall market share for lender j is obtained by integrating over the distribution of price sensitivities:

$$s_{ij}(r_{ij}) := \int_{-\infty}^{\infty} s_{ij}(\theta, r_{ij}) F_i^\theta(d\theta), \tag{9}$$

where F_i^θ is the cumulative distribution function of price sensitivity types θ in market i.

Adverse Selection Mechanism The interest rate r_{ij} affects not only the volume of applications but also their risk composition through adverse selection. The distribution of applicant types faced by lender j is given by:

$$F_{ij}^\theta(t; r_{ij}) = \left(s_{ij}(r_{ij})\right)^{-1} \int_{-\infty}^{t} s_{ij}(\theta, r_{ij}) F_i^\theta(d\theta). \tag{10}$$

This represents the conditional distribution of price sensitivity types among applicants to lender j. The wedge between $F_{ij}^\theta(t; r_{ij})$ and the population distribution $F_i^\theta(t)$ captures the standard selection-on-price effect: borrowers who are less price-sensitive (higher θ) are more likely to apply to lenders charging higher rates. Since price sensitivity may be correlated with default risk in our model, this selection effect influences the risk composition of each lender's applicant pool.

This adverse selection mechanism is empirically validated in Fact 2 (Section 2.2), where we document a strong positive relationship between residualized interest rates and residualized default rates. The model's prediction that higher rates attract riskier applicants aligns with the reduced-form evidence that borrowers paying higher-than-predicted rates are more likely to default ex post, consistent with classic adverse selection patterns.

3.1.4 Loan Officer Labor Supply

Loan officers choose where to work based on wages and their preferences for different locations and lenders. Each officer type k has a mass l_k of potential workers. An officer λ of type k chooses which lender j and market i to work for to maximize indirect utility:

$$u_\lambda = b_{ij\lambda} w_{ijk}, \tag{11}$$

where w_{ijk} is the wage offered by lender j to type-k officers in market i, and $b_{ij\lambda}$ is an idiosyncratic preference shock that varies across officers.

The preference shocks $b_{ij\lambda}$ follow a *joint* Fréchet distribution that captures the structure of labor market competition:

$$\mathbb{P}\left(\left(\bigcap_{n=1}^{N} \bigcap_{j \in E_n} \{b_{ij} \leq t_{ij}\}\right) \cap \{b_{oo} < t_{oo}\}\right) = \exp\left\{ - \left(\sum_i \left(\sum_j B_{ij} t_{ij}^{\overline{\frac{\epsilon}{(1-\rho)(1-\phi)}}} \right)^{1-\rho} - B_{oo} t_{oo}^{\frac{\epsilon}{1-\phi}} \right)^{1-\phi} \right\}.$$

This distribution structure allows for flexible substitution patterns in the labor market, as in Giroud et al. (2024). The parameters ρ and ϕ govern the degree of competition: there is typically more substitution among jobs within the same market than across markets, and more substitution among inside options (mortgage lending jobs) than between inside and outside options. The outside option pays an exogenous wage w_{oo} with scale parameter B_{oo}.

Loan officers differ along two key dimensions that affect lender productivity, both of which are empirically validated in the reduced-form analysis: screening precision and processing capacity. Specifically, officer type k has screening efficiency σ_{ijk}, which determines the precision of the signals they receive about borrower risk. Lower σ_{ijk} indicates more precise screening. This captures the informational advantage of local officers documented in Fact 2 (Section 2.2), where local officers make greater use of soft information and achieve better default outcomes despite lower rejection rates. Moreover, each officer of type k can process up to e_{ijk} applications, representing their workload capacity. This dimension aligns with the quantity productivity advantage documented in Fact 3 (Section 2.3), where local officers process refinance applications faster than non-local officers.

The labor supply system yields a well-defined supply function $l_{ijk}(\mathbf{w})$ that gives the mass of type-k officers who choose to work for lender j in market i as a function of the wage vector $\mathbf{w}$ across all lenders and markets. This structure directly relates to Fact 5 (Section 2.4), where we show that lenders' allocation of local versus remote officers responds systematically to local labor costs, with higher wages leading to greater spatial misalignment between underwriting capacity and mortgage demand.

The heterogeneous screening precision across officer types also connects to Fact 4 (Section 2.4), where we document that local officers' informational advantages are most valuable for high-credit-risk borrowers. In the model, this corresponds to settings where the signal precision σ_{ijk} matters most for distinguishing between fundamentally different borrower types within the same observable risk category.

3.1.5 Lender's Problem

Lenders are profit-maximizing and have rational expectations. Each lender j chooses an interest rate r_{ij} and type-specific wages $\{w_{ijk}\}_{k=1}^{K}$ to maximize profits across all markets, subject to loan officer capacity constraints. The lender also chooses assignment shares $\mathcal{S}_{ijk}$ that allocate incoming applications across different officer types.

The lender solves the following constrained optimization:

$$\begin{aligned}\max_{\left\{r_{ij},\left\{w_{ijk},\mathcal{S}_{ijk}\right\}_{k=1}^{K}\right\}} \sum_{i}\sum_{k=1}^{K}\Bigg\{&\mathcal{M}_i s_{ij}\left(r_{ij}\right)\mathcal{S}_{ijk}\mathbb{P}_{ij}\left\{orig; r_{ij}, k\right\}\Big(r_{ij}-f_{ij}-\mathbb{E}\left[lgd * d_{ij}; r_{ij}, k, orig\right]\Big)\\ &- w_{ijk}l_{ijk}\Bigg\}\\ \text{s.t. }\ &\mathcal{M}_i s_{ij}\left(r_{ij}\right)\mathcal{S}_{ijk}\leq l_{ijk}e_{ijk}\quad\forall k\end{aligned} \tag{12}$$

The term $\mathcal{M}_i s_{ij}(r_{ij})$ represents the total volume of applications received by lender j in market i, which equals the mass of potential borrowers $\mathcal{M}_i$ multiplied by the lender's market share $s_{ij}(r_{ij})$. The market share depends on the posted interest rate r_{ij} through borrower demand, which will be formally derived later. The assignment share $\mathcal{S}_{ijk}$ denotes the fraction of applications that lender j assigns to officer type k in market i. The term $\mathbb{P}_{ij}\{orig; r_{ij}, k\}$ represents the origination probability for applications screened by officer type k, while $(r_{ij} - f_{ij} - \mathbb{E}[lgd * d_{ij}; r_{ij}, k, orig])$ captures the net interest margin per originated loan. Finally,

$w_{ijk}l_{ijk}$ represents the labor cost of employing l_{ijk} officers of type k at wage w_{ijk}.

The capacity constraint ensures that the volume of applications assigned to each officer type $\mathcal{M}_i s_{ij}(r_{ij})\mathcal{S}_{ijk}$ does not exceed the total screening capacity $l_{ijk}e_{ijk}$ available from that officer type. Here, l_{ijk} denotes the mass of type-k loan officers working for lender j in market i, and e_{ijk} represents the processing capacity of each officer, measured in applications per officer.

To simplify notation, define the origination probability and expected default rate as:

$$p_{ijk} \coloneqq \mathbb{P}_{ij}\left\{orig; r_{ij}, k\right\}, \quad d_{ijk} \coloneqq \mathbb{E}\left[d_{ij}; r_{ij}, k, orig\right].$$

These objects depend on both the interest rate r_{ij}, which affects the risk composition of applicants through adverse selection, and the officer type k, which determines screening precision through the signal variance σ^2_{ik}. The formal derivation of how screening precision affects p_{ijk} and d_{ijk} through the lender's optimal approval rule will be provided in what follows.

Origination and Screening Screening determines both which applicants are approved and the expected default profile of originated loans. An officer of type k observes a noisy signal of applicant η's true default risk x_η:

$$\hat{x}_\eta = x_\eta + \tilde{x}_\eta, \quad \tilde{x}_\eta \sim F^{\tilde{x}}_{ijk}, \quad \mathbb{E}[\tilde{x}_\eta] = 0, \quad \mathrm{Var}(\tilde{x}_\eta) = \sigma^2_{ijk}.$$

There is a monotonic mapping from the latent risk type to the default probability, $d_i(x_\eta)$, with $d_i' > 0$.

Given the posted rate r_{ij} and funding cost f_{ij}, the lender approves an application if and only if the expected default cost does not exceed the net interest margin:

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

This condition defines a cutoff set $\hat{\Delta}_{ij}(r_{ij}, k)$ in the signal space. Applicants whose signals fall within this set are originated.

This screening technology maps officer precision and borrower characteristics into two key objects:

1. The **origination probability** for applicants screened by type k:

$$p_{ijk} = \mathbb{P}\left(\hat{x}_\eta \in \hat{\Delta}_{ij}(r_{ij}, k)\right).$$

2. The **expected default rate** among originated loans screened by type k:

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

Because borrower selection (through r_{ij}) and screening precision (through σ^2_{ijk}) jointly determine p_{ijk} and d_{ijk}, screening and origination are central to understanding how informational frictions translate into credit rationing and equilibrium outcomes. The posted rate r_{ij} shifts the distribution of applicant risk types, while the signal quality σ^2_{ijk} determines the lender's ability to discern and select among them.

Labor Cost A key feature of the model is that lenders do not internalize their full impact on local wages. From the labor supply problem, the cost of hiring l_{ijk} officers of type k can be expressed as:

$$w_{ijk} l_{ijk} = A_{ijk}^{-1} \tilde{l}_{ijk}^{1+\frac{1}{\varrho}}$$

where $\varrho = \frac{\epsilon}{(1-\rho)(1-\phi)}$ is a composite labor supply elasticity parameter, and A_{ijk} is a composite cost parameter that incorporates wage information, utility scale parameters, and labor efficiency units e_{ijk}. This formulation implies that the lender's problem is separable across markets, allowing us to analyze each market i independently.

Solution to the Bank's Problem The first-order conditions yield a fixed point problem for the optimal interest rate r_{ij}:

$$\sum_{k=1}^{K} A_{ijk}^{\varrho} p_{ijk}^{\varrho+1} \left(r_{ij} - f_{ij} - lgd * d_{ijk}\right)^{\varrho} \Bigg\{ r_{ij}\left(\varepsilon_{p_{ijk}r} + \varepsilon_{s_{ij}r} + 1\right) - f_{ij}\left(\varepsilon_{p_{ijk}r} + \varepsilon_{s_{ij}r}\right) - lgd * d_{ijk}\left(\varepsilon_{p_{ijk}r} + \varepsilon_{s_{ij}r} + \varepsilon_{d_{ijk}r}\right) - \varepsilon_{s_{ij}r} \bar{c}_{ij}\left(r_{ij} - f_{ij} - lgd * d_{ijk}\right) \Bigg\} = 0 \quad (13)$$

where $\varepsilon_{xy} = \frac{\partial \ln x}{\partial \ln y}$ denotes an elasticity, and the capacity cost shifter is defined as:[4]

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

Further, the bank sets labor shares following:

$$\mathcal{S}_{ijk} = \frac{A_{ijk}^{\varrho} p_{ijk}^{\varrho} (r_{ij} - f_{ij} - d_{ijk} * lgd)^{\varrho}}{\sum_k A_{ijk}^{\varrho} p_{ijk}^{\varrho} (r_{ij} - f_{ij} - d_{ijk} * lgd)^{\varrho}} \tag{14}$$

To build intuition on prices, consider a simplified case where screening precision is homogeneous across officer types, so that $p_{ijk} = \bar{p}_{ij}$ and $d_{ijk} = \bar{d}_{ij}$ for all k. This corresponds to a world where local and remote officers have identical screening abilities—there is no informational advantage to proximity. In this counterfactual, the lender's problem reduces to pure cost minimization: hiring decisions are driven solely by wage differences and processing capacities, not by differential screening quality.

In this simplified case, the optimal rate simplifies to:

$$r_{ij} = \underbrace{f_{ij} \frac{\varepsilon_{p_{ij}r} + \varepsilon_{s_{ij}r}}{\varepsilon_{p_{ij}r} + \varepsilon_{s_{ij}r} + 1}}_{\text{Funding cost component}} + \underbrace{lgd * d_{ij} \frac{\varepsilon_{p_{ij}r} + \varepsilon_{s_{ij}r} + \varepsilon_{d_{ij}r}}{\varepsilon_{p_{ij}r} + \varepsilon_{s_{ij}r} + 1}}_{\text{Default risk component}} + \underbrace{p_{ij}^{-1} \left(\frac{\mathcal{M}_i s_{ij}}{\sum_{k=1}^{K} A_{ijk}^{\varrho}} \right)^{\frac{1}{\varrho}} \frac{\varepsilon_{s_{ij}r}}{\varepsilon_{p_{ij}r} + \varepsilon_{s_{ij}r} + 1}}_{\text{Capacity cost component}} \tag{15}$$

This decomposition reveals three key components of mortgage pricing: (1) funding cost component, which reflects the pass-through of funding costs, weighted by the combined elasticity of origination probability and market share, (2) default risk component, which captures expected default losses, incorporating adverse selection through $\varepsilon_{d_{ij}r}$, and (3) capacity cost component: which represents the marginal cost of screening capacity, which depends on application volume and the cost of hiring loan officers.

The main difference between the general case (13) and the simplified case (15) is that the general case involves a mixture over elasticities and net interest margins across different

[4] Note that d_{ijk}, p_{ijk}, $\varepsilon_{p_{ijk}r}$, $\varepsilon_{s_{ij}r}$, $\varepsilon_{d_{ijk}r}$, and $\bar{c}_{ij}$ all depend on r_{ij} and can be calculated via Monte Carlo integration.

officer types. This mixture depends on signal quality and allows lenders to pay different marginal costs to different officer types, capturing the private value of information to the bank. The heterogeneity in screening precision σ^2_{ik} across officer types generates differential information rents that are reflected in equilibrium pricing.

3.2 Calibration

Parameterization and Identification To take the model to data, we make the following parametric assumptions. Borrower types (θ_η, x_η) are jointly normally distributed. The demand shocks $\epsilon_{ij\eta}$ are i.i.d. Type I extreme value. We consider two types of loan officers ($K = 2$), "local" and "non-local," with the former having a lower signal variance (σ^2_{ijk}). Finally, borrower η defaults if $x_\eta \leq \epsilon_\eta + \epsilon_t$, where ϵ_η and ϵ_t are iid standard normal idiosyncratic and aggregate shocks, respectively.

The parameters governing borrower preferences $(\mu_\theta, \sigma^2_\theta)$ are identified via standard Berry, Levinsohn, and Pakes (1995) moments from the demand system. The joint distribution parameters, particularly the adverse selection correlation γ, are identified by comparing the realized default rates of observably similar borrowers across different lenders and pricing regimes, leveraging the model's structure on how rates select applicants. The value of a rejected application is the utility of the forgone outside option, net of any sunk application cost. This can be parameterized as $u_{\text{reject}} = \kappa$, where $\kappa \leq 0$ is a parameter to be calibrated or estimated, potentially from data on application rates and rejection probabilities.

Externally Calibrated Parameters. We calibrate our logit price coefficients to match Buchak et al. (2018a), giving $\mu_\theta = 165$ and $\sigma_\theta \approx 41$. We calibrate the within-region-sector labor supply elasticty $\varrho = 8$ to match Giroud et al. (2024). We calibrate funding costs to 2.75% to match the spreads reported in Janus Henderson Investors (2019). We calibrate losses-given-default, lgd, to 17.7% using Fannie Mae securitization data.[5]

Internally Calibrated Parameters. We consider 6 product type buckets:

$\{subprime, near-prime, prime\} \times \{purchase, refi\}$ and two types of loan officers $\{local, nonlocal\}$.

[5] 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

The model is calibrated to match average market conditions reported in Table 8. We consider three types of banks: "Brick and Mortar", "Fintech", and "Fringe" in the representative market. We calibrate the labor supply efficiency parameters $\{A_{i,j,k}\}_{\forall i,j,k}$ and demand shifters $\{\xi_{i,j,k}\}_{\forall i,j,k}$ to exactly match observed labor shares and market shares using standard techniques from the IO and Spatial Economics literatures. We then jointly calibrate loan officer signal precisions $\{\sigma_k\}_{\forall k}$, the correlation between default type and price sensivity, γ, borrower type distributions $\mu_i^x, \sigma_i^x{}_{\forall i}$, and market-specific aggregate state shocks $\epsilon_{i,t}$ to jointly match 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, 3, and 8 as well as Figure 3

Critical Parameter Identification. Briefly, we describe how our novel loan officer parameters, $\{A_{i,j,k}\}_{\forall i,j,k}$ and $\{\sigma_k\}_{\forall k}$ are jointly pinned down by the data. First, intuitively, the magnitudes of $\{\sigma_k\}_{\forall k}$ are jointly pinned down by the local rejection rate and default advantages. Ceteris paribus, as information gets less precise, the rejection rate advantage of the same informational advantage declines. In the extreme case, the model can generate a psuedo-pooling equilibrium, where improving informational precision actually increases rejection rates; this is because, as information gets sufficiently imprecise, the benefits of rejecting bad applicants begin to dominate the benefits of approving good applicants, leading to an increase in overall rejection rates. Given the level of imprecision, improved information lowers expected default rates, pinning down the overall informational advantage of local loan officers. Given these advantages, the $\{A_{i,j,k}\}_{\forall i,j,k}$ parameters capture the wage-adjusted physical efficiency of loan officers for each bank-market-officer type. The relative physical efficiency advantage of each type is pinned down by the local share they hire, and the absolute advantage is pinned down through their first order condition in equation (13). Intuitively, a bank with a higher physical efficiency for loan officer type k will hire more officers of that type, taking into account the impact that their relative lending efficiency has on the banks' expected profits.

Parameters and Moments. We report the values of borrower distribution types and the aggregate shock in Table blank. For ease of interpretation, we transform the mean of the latent x type to a 'default rate', which captures what the expected per-year default rate would be in the relevant mortgage market, if 100 percent of applications were approved.

Default rates and latent heterogeneity are significantly higher in the refi markets, driven by the lower realized default rates and approval rates. Perhaps unsurprisingly, since the model is exactly identified, the joint set of parameters allows us to exactly match the reported moments throughout the paper, with targets as noted above.

In the above parameterization of the model, for simplicity, we assumed that the information technology is the same in every market. This leaves untargetted heterogeneity in the effect sizes as a potential tool to validate the model. To validate the model, we calculate model-implied, market-specific rejection rate and default rate advantages, and compare them to the data. We find that in the model, the local rejection rate advantage is about 3 percentage points higher in the refi market than the purchase market, roughly consistent with the magnitude in Table 2; the local default rate advantage is about 0.4 percentage points higher in the purchase market than the refi market, roughly consistent with the magnitude in Table 3; and the local rejection rate (about 3 percentage points) and default rate advantages (about 1 percentage point) are much larger for subprime loans than for prime loans, consistent with Table 6, although the relevant coefficients are estimated somewhat noisily.

Note that these results are driven by the differences in borrwer type distributions and aggregate shocks, which the model backs out from the observed default rate and origination rate data.

4 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.

4.1 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 3 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 $lgd * E[d(x_\eta)] \leq r_{ij} - f_{ij}$, 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 documented in Table 2, relative to default–rate advantages in Table 3, 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 C 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.

4.2 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.[6]

In counterfactual 1, only the Brick and Mortar bank receives the efficiency improvement.

[6] Given the structure in Section 3, 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.

In counterfactual 2, all banks recieve 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 (14), shocked banks substitute distant labor for local labor.

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.[7] 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.

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

Figure 7 plots the difference in realized default rates between the baseline and shocked economies as a function of the aggregate state for the purchase markets.[8]

The effect is non–monotonic. For small aggregate shocks, pooled–origination loans experience 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 7 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.

[8] 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.

5 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

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, 2019, The geographic flow of bank funding and access to credit: Branch networks, local synergies, and competition .

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, 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.

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, 2020a, On the rise of fintechs: Credit scoring using digital footprints, *The Review of Financial Studies* 33, 2845–2897.

Berg, Tobias, Andreas Fuster, and Manju Puri, 2022, Fintech lending, *Annual Review of Financial Economics* 14, 187–207.

Berg, Tobias, Manju Puri, and Jörg Rocholl, 2020b, Loan officer incentives, internal rating models, and default rates, *Review of Finance* 24, 529–578.

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.

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* .

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

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru, 2018b, Fintech, regulatory arbitrage, and the rise of shadow banks, *Journal of Financial Economics* 130, 453–483.

Chen, Mark A, Qinxi Wu, and Baozhong Yang, 2019, How valuable is fintech innovation?, *The Review of financial studies* 32, 2062–2106.

Cox, Natalie, Ernest Liu, and Daniel Morrison, 2021, Market power in small business lending: A two-dimensional bunching approach, Technical report.

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.

Eizenberg, Alon, 2014, Upstream innovation and product variety in the us home pc market, *Review of Economic Studies* 81, 1003–1045.

Frame, W Scott, Ruidi Huang, Erica Xuewei Jiang, Yeonjoon Lee, Will Shuo Liu, Erik J Mayer, and Adi Sunderam, 2025, The impact of minority representation at mortgage lenders, *The Journal of Finance* 80, 1209–1260.

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.

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

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

Goldstein, Itay, Wei Jiang, and G Andrew Karolyi, 2019, To fintech and beyond, *The Review of Financial Studies* 32, 1647–1661.

Haendler, Charlotte, 2022, Keeping up in the digital era: How mobile technology is reshaping the banking sector, *Available at SSRN 4287985* .

He, Zhiguo, Sheila Jiang, Douglas Xu, and Xiao Yin, 2021, Investing in lending technology: It spending in banking, *University of Chicago, Becker Friedman Institute for Economics Working Paper* .

Hertzberg, Andrew, Jose Maria Liberti, and Daniel Paravisini, 2010, Information and incentives inside the firm: Evidence from loan officer rotation, *The Journal of Finance* 65, 795–828.

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.

Jiang, Erica Xuewei, Gloria Yang Yu, and Jinyuan Zhang, 2022, Bank competition amid digital disruption: Implications for financial inclusion, *Available at SSRN 4178420* .

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, 2017, Technological innovation, resource allocation, and growth, *The quarterly journal of economics* 132, 665–712.

Koont, Naz, 2023, The digital banking revolution: Effects on competition and stability, *Available at SSRN* .

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, 2025, Regional banks, aggregate effects, *SSRN)* .

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

Narayanan, Rajesh, Dimuthu Ratnadiwakara, and Philip E Strahan, 2025, The decline of branch banking, *Available at SSRN* .

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

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.

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.

Stulz, René M, 2019, Fintech, bigtech, and the future of banks, *Journal of applied corporate finance* 31, 86–97.

Tirole, Jean, 2020, Competition and the industrial challenge for the digital age, *IFS Deaton Review on Inequalities in* .

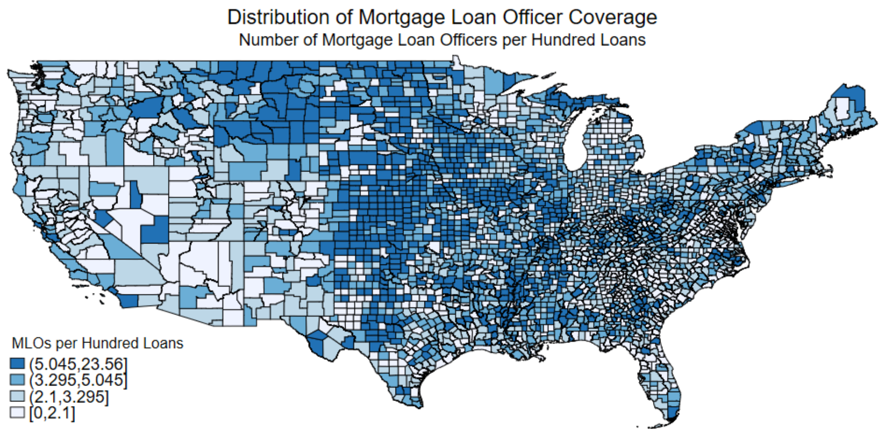
Vives, Xavier, 2019, Digital disruption in banking, *Annual Review of Financial Economics* 11, 243–272.

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

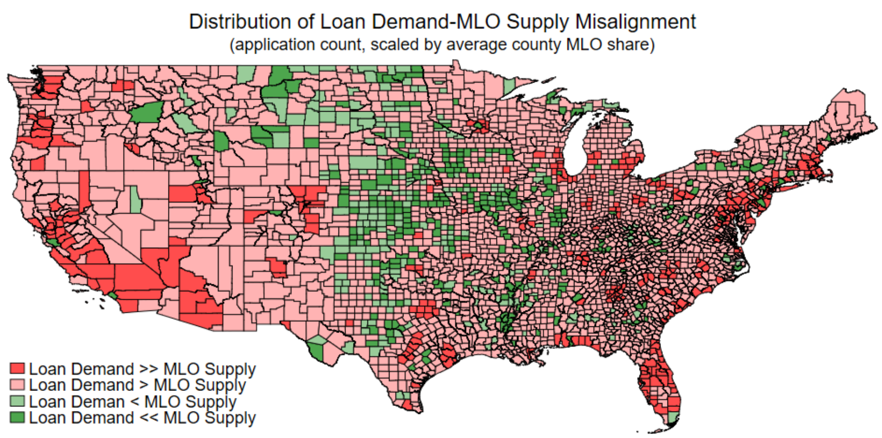
Whited, Toni M, Yufeng Wu, and Kairong Xiao, 2022, Central bank digital currency and banks, *Available at SSRN 4112644* .

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



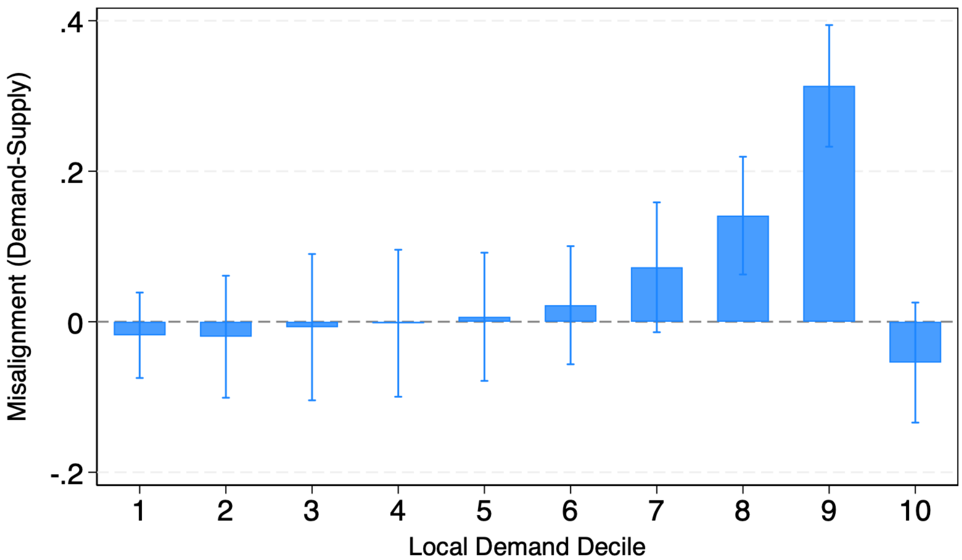
(a) Geographic Distribution of Mortgage Loan Officer Coverage



(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.

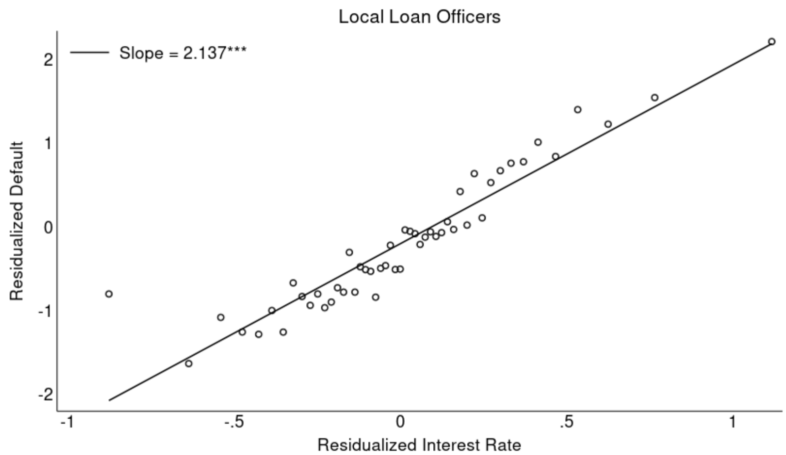
Figure 2. Loan Demand-Loan Officer Supply Misalignment Across Counties



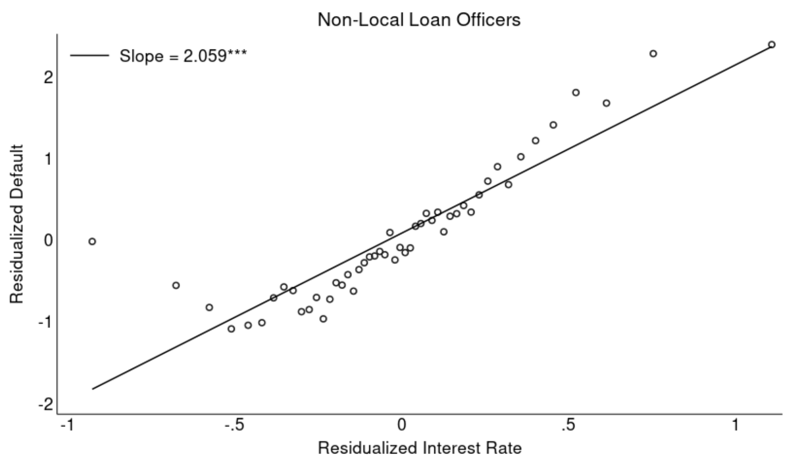
(a) Excess Loan Demand Relative to Loan Officer Supply

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. Adverse Selection: Interest Rate and Default Relationship



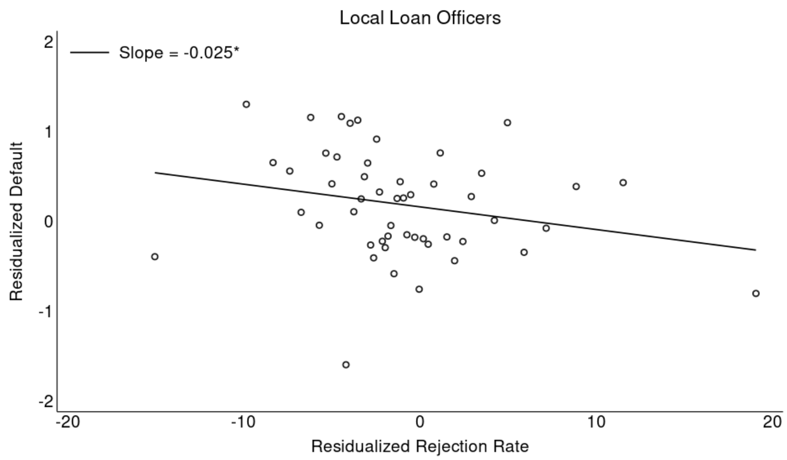
(a) Local Loan Officer



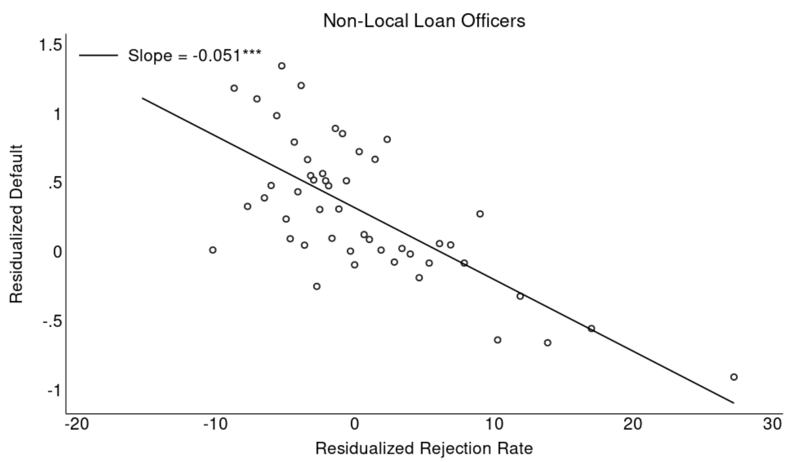
(b) Non-Local Loan Officer

Note: This figure plots the relationship between residualized interest rates and residualized default outcomes for originated home-purchase loans, separately for loans handled by local loan officers (Panel A) and non-local loan officers (Panel B). Residualized interest rates are obtained by regressing the contracted interest rate on borrower-level risk controls (FICO, LTV, DTI, AUS status, and their flexible polynomials) and county-by-application-month fixed effects. Residualized default probabilities are constructed analogously by residualizing the two-year default indicator. Residuals are first averaged at the lender–county–local/non-local level; thus each underlying observation represents the mean residualized interest rate and mean residualized default for a given lender in a given county, separately for local and non-local officers. For visualization, these lender–county averages are then sorted into 50 equally sized bins based on the residualized interest rate (within each subsample). Each point plots the mean default residual against the mean interest-rate residual within a bin. The fitted line shows the slope from a regression of residualized default on residualized interest rates estimated separately for local and non-local subsamples.

Figure 4. Loan Approval Standard and Ex-Post Default



(a) Local Loan Officer



(b) Non-Local Loan Officer

Note: This figure plots the relationship between residualized loan approval standards and residualized default outcomes for loans processed by local (Panel A) and non-local (Panel B) loan officers. Loan-level default residuals and loan-level rejection residuals (from HMDA) are first aggregated to the lender–county–local/non-local loan officer level, so each underlying observation reflects the average residualized rejection and default outcomes for a given lender in a given county, separately for local and non-local officers. For visualization, these lender–county averages are then sorted into 50 equally sized bins (within each local/non-local group), and the figure plots the mean default residual against the mean rejection residual within each bin. The fitted line shows the linear relationship between the two variables within each group.

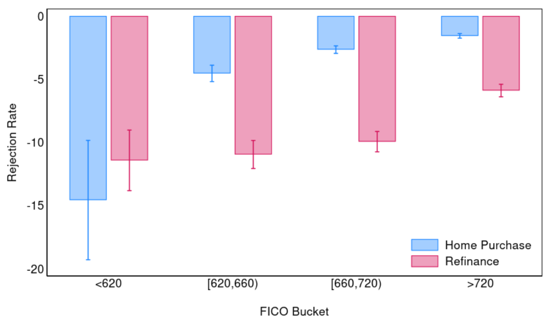
Figure 5. Loan Approval Standard and Ex-Post Default by Loan Rates

(a) High Rate Sample

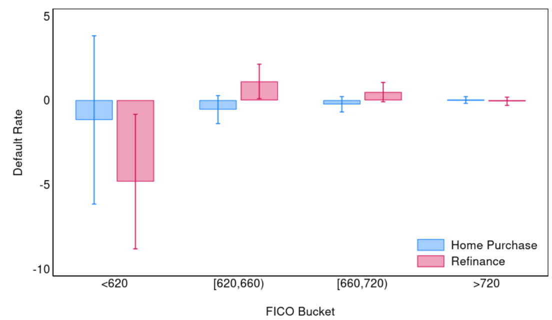
(b) Low Rate Sample

Note:This figure plots the relationship between residualized rejection rates and residualized default probabilities for home-purchase loans, separately for markets with above-median residualized interest rates (Panel A) and below-median residualized interest rates (Panel B). Residualized rejection rates are obtained by regressing the loan-level rejection indicator on borrower risk controls (FICO, LTV, DTI, AUS status, and their flexible polynomials) and county-by-application-month fixed effects. Residualized default probabilities are constructed analogously from a regression of the two-year default indicator with the same controls. The underlying data are first aggregated to the lender–county–local/non-local level, producing mean residualized rejection rates and default rates for each lender operating in each borrower county, separately for loans handled by local and non-local officers. Within each interest-rate subsample (high vs. low), these lender–county averages are sorted into 50 quantile bins of the residualized rejection rate. Each point in the figure represents the mean residualized default and mean residualized rejection rate within a bin. The fitted lines plot slopes from regressions of residualized default on residualized rejection rates estimated separately for local and non-local loan officers within each subsample.

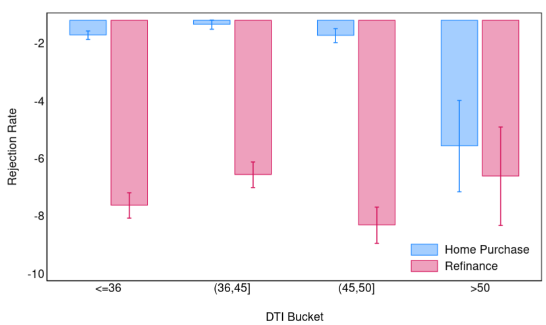
Figure 6. Rejection and Default by Hard Credit Information: Local vs Non-Local



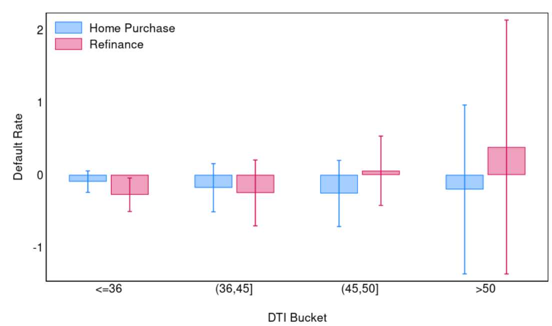
(a) Rejection-FICO



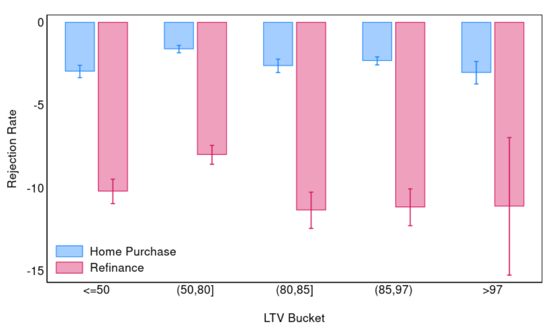
(b) Default-FICO



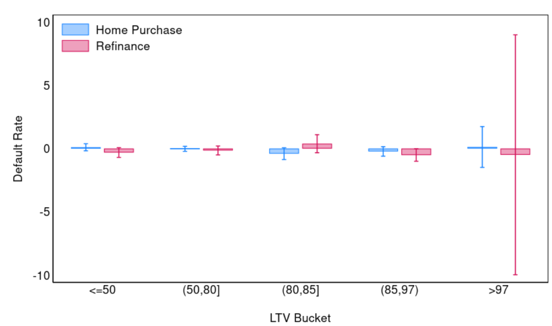
(c) Rejection-DTI



(d) Default-DTI



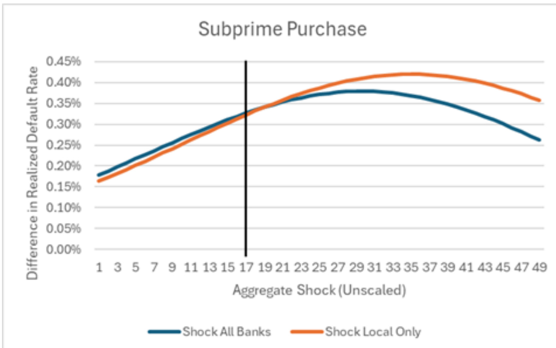
(e) Rejection-LTV



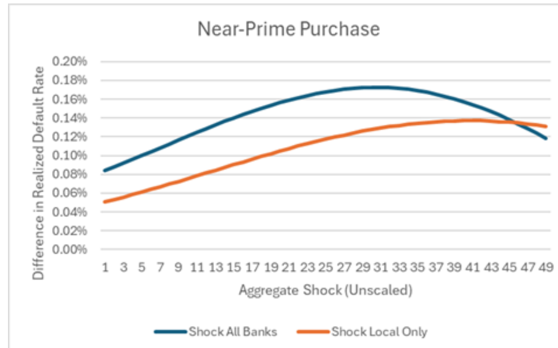
(f) Default-LTV

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.

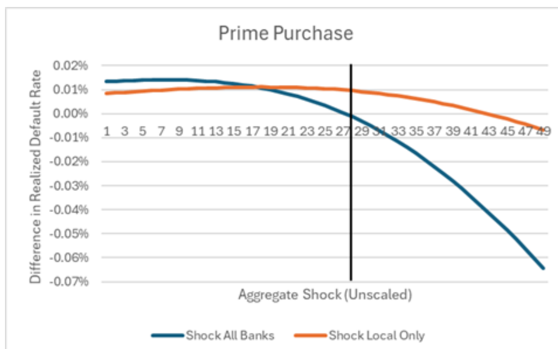
Figure 7. Technology Shock and Aggregate Risk



(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 4.2 for more details.

Table 1: 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 2: Local Advantage - Rejection

This table reports loan-level regressions examining how borrower–officer proximity affects mortgage rejection decisions. The dependent variable is an indicator equal to 100 if a loan application is rejected and 0 otherwise. The sample includes all home purchase (Panel A) and refinance (Panel B) applications in the confidential HMDA for 2018–2019. 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. Columns (1)–(3) estimate the effect of proximity using an indicator for whether the reviewing officer is local. Columns (4)–(6) replace the Local indicator with log distance between the borrower and the officer. Across both sets, the three specifications correspond to progressively richer fixed effects as indicated in the table. Standard errors are clustered at the borrower–county level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Home Purchase Loans

	Rejection Rate					
	(1)	(2)	(3)	(4)	(5)	(6)
Local	-1.881***	-0.643***	-0.202***			
	(0.09)	(0.05)	(0.03)			
Log Distance				0.592***	0.228***	0.069***
				(0.02)	(0.01)	(0.01)
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	3,721,849	3,716,598	3,704,260	3,721,799	3,716,547	3,704,210
R^2	0.229	0.279	0.316	0.231	0.279	0.316

Panel B: Refinance

	Rejection Rate: Refinance					
	(1)	(2)	(3)	(4)	(5)	(6)
Local	-7.898***	-2.855***	-0.655***			
	(0.22)	(0.15)	(0.10)			
Log Distance				1.822***	0.797***	0.236***
				(0.05)	(0.03)	(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	1,915,240	1,907,786	1,886,996	1,915,218	1,907,764	1,886,977
R^2	0.380	0.456	0.506	0.387	0.457	0.506

Table 3: Local Advantage - Default

This table reports loan-level regressions examining how borrower–officer proximity affects default outcomes. The dependent variable is an indicator equal to 100 if a loan becomes 60 days delinquent within two years of origination and 0 otherwise. The sample includes all home purchase (Panel A) and refinance (Panel B) approved loans in the confidential HMDA approved loan applications during 2018-2019 that are merged with McDash loan performance records. 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. Columns (1)–(3) estimate the effect of proximity using an indicator for whether the reviewing officer is local. Columns (4)–(6) replace this indicator with the log distance between the borrower and the officer. Across both sets, the three specifications correspond to progressively richer fixed effects as indicated in the table. Standard errors are clustered at the borrower–county level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	Default Rate: Home Purchase Loans					
	(1)	(2)	(3)	(4)	(5)	(6)
Local	-0.332***	-0.463***	-0.344***			
	(0.07)	(0.06)	(0.05)			
Log Distance				0.033***	0.096***	0.098***
				(0.01)	(0.01)	(0.01)
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	2,138,596	2,135,269	2,123,285	2,138,585	2,135,258	2,123,274
R^2	0.107	0.124	0.165	0.107	0.124	0.165

	Default Rate: Refinance					
	(1)	(2)	(3)	(4)	(5)	(6)
Local	0.492***	0.037	0.000			
	(0.10)	(0.08)	(0.08)			
Log Distance				-0.105***	0.008	0.010
				(0.02)	(0.01)	(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	765,822	760,305	742,317	765,821	760,304	742,317
R^2	0.079	0.110	0.195	0.079	0.110	0.195

Table 4: Local Advantage - Interest Rates

This table reports loan-level regressions examining how borrower–officer proximity affects mortgage interest rates for originated loans. The dependent variable is the interest rate on the originated loan (in percentage points). 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. Columns (1)–(3) estimate the effect of proximity using an indicator for whether the reviewing loan officer is local. Columns (4)–(6) replace this indicator with the log distance between the borrower and the officer. Across both sets, the three specifications correspond to increasingly saturated fixed effects as indicated in the table. The sample includes all originated home purchase (Panel A) and refinance loans (Panel B) in the confidential HMDA for 2018–2019. Standard errors are clustered at the borrower–county level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Home Purchase Loans

	Interest Rate					
	(1)	(2)	(3)	(4)	(5)	(6)
Local	0.033***	0.010***	-0.002**			
	(0.00)	(0.00)	(0.00)			
Log Distance				-0.010***	-0.003***	0.000**
				(0.00)	(0.00)	(0.00)
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	3,370,947	3,365,679	3,354,197	3,370,902	3,365,633	3,354,152
R^2	0.564	0.661	0.707	0.565	0.661	0.707

Panel B: Refinance

	Interest Rate					
	(1)	(2)	(3)	(4)	(5)	(6)
Local	0.028***	0.037***	0.009***			
	(0.01)	(0.00)	(0.00)			
Log Distance				-0.004***	-0.007***	-0.002***
				(0.00)	(0.00)	(0.00)
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	1,450,704	1,442,695	1,421,876	1,450,692	1,442,684	1,421,868
R^2	0.546	0.655	0.735	0.545	0.655	0.735

Table 5: Local Advantage - 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 loan processing time, defined as the number of days between the origination date and the loan application date. 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. Columns (1)–(3) estimate the effect of proximity using an indicator for whether the reviewing loan officer is local. Columns (4)–(6) replace this indicator with the log distance between the borrower and the officer. Across both sets, the three specifications correspond to increasingly saturated fixed effects as indicated in the table. The sample includes all originated home purchase (Panel A) and refinance loans (Panel B) in the confidential HMDA for 2018–2019. Standard errors are clustered at the borrower–county level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Home Purchase Loans

	Processing Time					
	(1)	(2)	(3)	(4)	(5)	(6)
Local	-2.383***	-0.176	-0.028			
	(0.37)	(0.15)	(0.07)			
Log Distance				0.565***	-0.017	-0.021
				(0.07)	(0.03)	(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	3,380,205	3,374,927	3,363,438	3,380,160	3,374,881	3,363,393
R^2	0.090	0.299	0.371	0.091	0.299	0.371

Panel B: Refinance Loans

	Processing Time					
	(1)	(2)	(3)	(4)	(5)	(6)
Local	0.500**	-1.652***	-1.174***			
	(0.20)	(0.13)	(0.08)			
Log Distance				-0.207***	0.349***	0.312***
				(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	1,452,425	1,444,421	1,423,612	1,452,413	1,444,410	1,423,604
R^2	0.129	0.280	0.395	0.130	0.280	0.396
R^2 Adjusted	0.104	0.240	0.323	0.104	0.240	0.323

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), interest rates (columns 4-6), and default (columns 7-9), all in percentage points. The sample focuses on all conventional and jumbo home purchase loan applications in confidential HMDA for 2018–2019 (so excluding FHA, VA, and RHS loans): columns 1-3 use all loan applications; columns 4-6 use all originated loans; and columns 7-9 use all originated loans that are matched with the loan performance records in McDash. We consider three measures of borrower risk: (i) Subprime (FICO < 670), (ii) High DTI (DTI > 43), and (iii) High LTV (LTV > 80). For each outcome, we estimate regressions including the Local indicator, a high-risk indicator, and their interaction, along with loan type interacted 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)			Interest Rate (pp)			Default (pp)		
	Subprime	High DTI	High LTV	Subprime	High DTI	High LTV	Subprime	High DTI	High LTV
Local	-0.369***	-0.336***	-0.337***	0.015***	0.014***	0.017***	-0.242***	-0.186***	-0.165**
	(0.06)	(0.07)	(0.09)	(0.00)	(0.00)	(0.00)	(0.06)	(0.06)	(0.07)
Local×Subprime	-2.927***			-0.010*			-0.428		
	(0.24)			(0.01)			(0.30)		
Local×High DTI		-0.761***			0.003*			-0.273**	
		(0.13)			(0.00)			(0.12)	
Local×High LTV			-0.386***			-0.003			-0.188*
			(0.10)			(0.00)			(0.11)
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	2,385,885	2,385,885	2,385,885	1,419,546	1,419,546	1,419,546
R^2	0.252	0.250	0.250	0.670	0.664	0.665	0.093	0.092	0.092

Table 7: Misalignment and Wage

This table reports lender–MSA–year level regressions examining the relationship between mortgage demand–MLO supply misalignment and local labor market wages. For each lender j in MSA k and year t, we compute the share of the lender's registered loan officers located in that MSA (l_{jkt}/L_{jt}) and the share of the lender's loan applications originating from that MSA (m_{jkt}/M_{jt}). The dependent variable is the *Misalignment Index*, defined as $(m_{jkt}/M_{jt}) - (l_{jkt}/L_{jt})$. A higher value indicates that lender j receives a disproportionately large share of mortgage demand from MSA k relative to the share of its loan officers located there (i.e., undersupply of MLOs in that market). MSA-level controls include employment rates, income per capita, and earnings per job. Standard errors are clustered at the lender level. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Misalignment Index (Mortgage Demand-MLO Supply)					
	(1)	(2)	(3)	(4)	(5)	(6)
Loan Officer Hourly Wage	0.343***	0.396***	0.249***			
	(0.06)	(0.07)	(0.06)			
Finance Job Hourly Wage				1.589***	1.814***	2.326***
				(0.21)	(0.25)	(0.27)
Year FE	Yes			Yes		
Lender-Year FE		Yes	Yes		Yes	Yes
MSA Controls			Yes			Yes
N	187,671	186,631	185,709	187,671	186,631	185,709
R^2	< 0.001	0.047	0.049	0.002	0.049	0.050

Table 8: Descriptive Statistics

This table reports market-level summary statistics used to discipline the structural model. A "market" is defined by year $\times$ borrower county $\times$ loan purpose (home purchase vs. refinance) $\times$ 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 $\times$ 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)

		Home Purchase			Refinance		
Variable	Risk tier	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)

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 3. 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). Panel A reports aggregate rationing and pooling rates; Panel B breaks these out by loan officer type. See Section 4.1 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 4.2 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).[9] 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.[10] 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.[11] 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.[12] 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.

[9] See CFPB, "TILA–RESPA Integrated Disclosure Rule (TRID): Guide to Forms" (2022).
[10] Mortgage Bankers Association (MBA), "Mortgage Origination Survey," various years.
[11] 12 C.F.R.§1026.19(e)(1)(iii).
[12] 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.[13]

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.[14] 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

[13]See Fannie Mae Selling Guide B1–1–01 (2024); Freddie Mac Seller/Servicer Guide (2024).

[14]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.[15] 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.[16]

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.

[15] MBA, "Technology & Origination Report," 2020.
[16] 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.

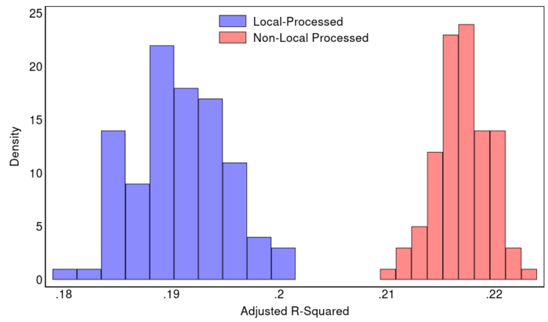
Next, we merge the HMDA originations to monthly loan-level performance data from the Black Knight McDash servicing database. Following the methodology of Rosen (2011), we treat a loan in HMDA and a loan in McDash as the same origination only when several characteristics match almost exactly. First, the 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.

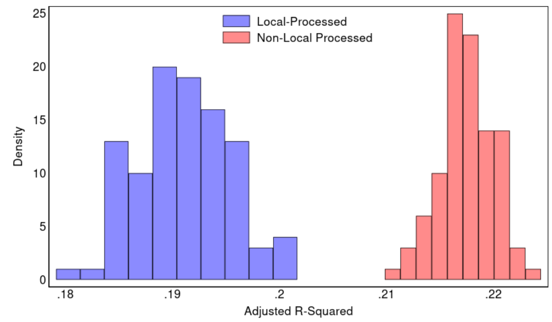
The final dataset links each loan application to (i) borrower and loan characteristics at the time of application, (ii) the lender and loan officer responsible for the file, (iii) the physical location of the loan officer obtained from NMLS, (iv) underwriting decisions and posted interest rates, and (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 Additional Figures and Tables

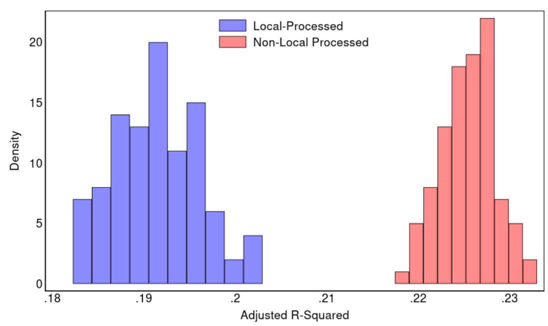
Figure A1. Explanatory Power of Hard Information in Loan Rejection



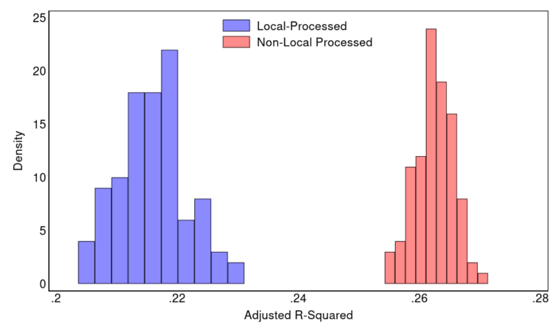
(a) Model 1



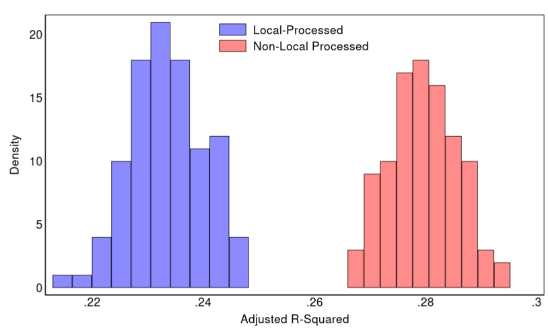
(b) Model 2



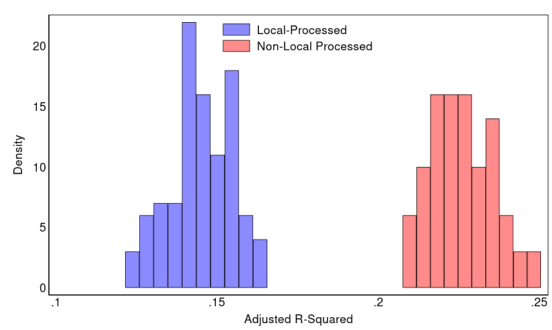
(c) Model 3



(d) Model 4



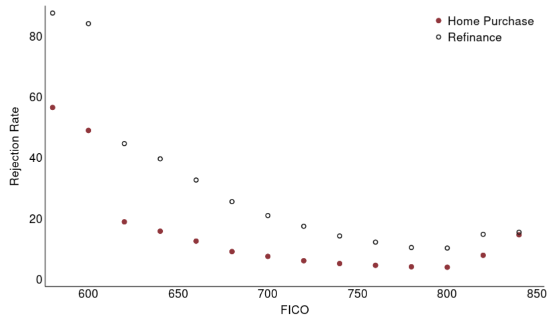
(e) Model 5



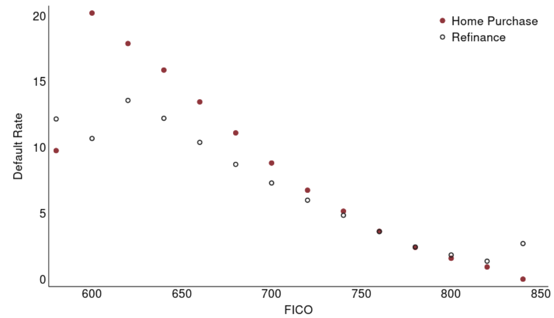
(f) Model 6

Note: This figure plots adjusted R-squared of six different regression specifications. In each regression, we regress a loan rejection indicator on observable loan and borrower characteristics, as well as different combinations of fixed effects. For each regression specification, we estimate it using 100 random xx% samples and record the adjusted R-squared of each estimation.

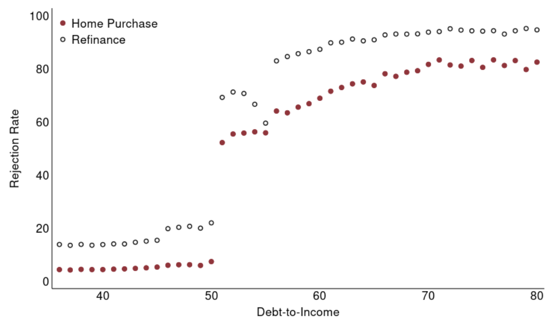
Figure A2. Rejection and Default by Hard Credit Information



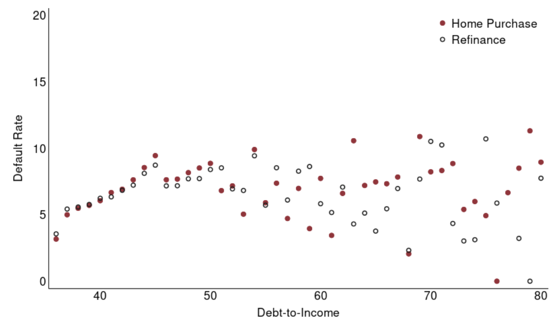
(a) Rejection-FICO



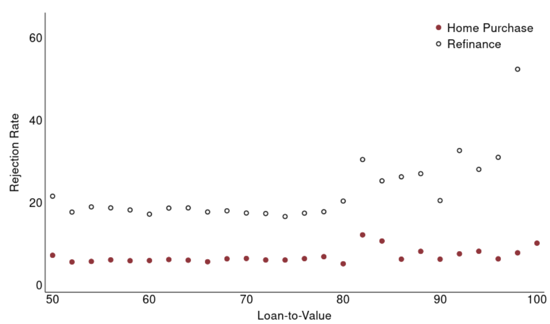
(b) Default-FICO



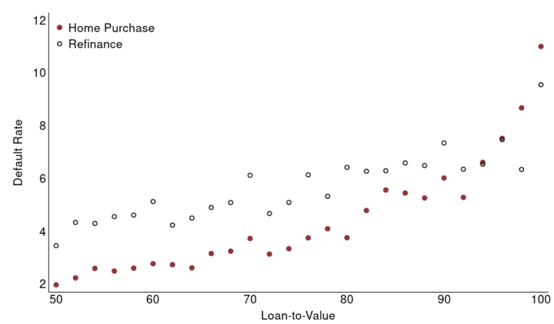
(c) Rejection-DTI



(d) Default-DTI



(e) Rejection-LTV



(f) Default-LTV

Note: This figure plots rejection rate and default rate by FICO buckets in Panels A and B, by Debt-to-Income (DTI) ratios in Panels C and D, and by Loan-to-Value (LTV) ratios in Panels E and F.