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

Erica Xuewei Jiang † Yeonjoon Lee ‡ Quinn Maingi §

Abstract

We study how the organization of screening activity—and its endogenous response to economic and technological forces—affects informational efficiency, credit allocation, and the distribution of borrower risk. Using US administrative data linking loan officers to mortgage applications and loan performance, we document that local officers achieve higher screening precision and faster processing; the informational benefits of proximity accrue disproportionately to borrowers with higher observable risk; and lenders allocate labor elastically with respect to local wages, resulting in systematic spatial misallocation of underwriting capacity relative to mortgage demand. We develop and estimate a structural model in which lenders set prices before observing borrower-specific signals, borrowers self-select on posted rates, and loan officers of different screening precision generate information that determines loan approvals. Lenders compete in mortgage pricing and in labor markets for local and remote officers, endogenously allocating information production across markets. We find substantial baseline credit rationing—up to 15 percent in high-risk segments—with local officers eliminating roughly half of it while also reducing excessively risky approvals. A technology shock that raises the processing productivity of remote officers induces lenders to substitute away from local screening, lowering informational efficiency, increasing excessively risky approvals and expected defaults, and tightening rationing for marginal borrowers despite only modest reductions in interest rates.

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

*We are grateful to the Conference of State Bank Supervisors (CSBS) for granting us access to data from NMLS Consumer Access SM. The views expressed are those of the authors and not necessarily those of the CSBS, the Structured Finance Association or any of its member firms, the Federal Reserve Bank of Richmond, or any other entity within the Federal Reserve System.

†UCLA Anderson School of Management, erica.jiang@anderson.ucla.edu.

‡Federal Reserve Bank of Richmond, yeonjoon.lee1@gmail.com.

§University of Southern California, Marshall School of Business, qmaingi@marshall.usc.edu.

Financial intermediation is built on the production and use of information. Informational frictions stemming from imperfect observability of borrower risk can distort credit allocation. Financial intermediaries create value by mitigating these frictions through costly screening (Diamond, 1984; Leland and Pyle, 1977; Ramakrishnan and Thakor, 1984). Yet the efficiency and distribution of screening depend on how it is organized: information production is itself an economic activity, mediated by labor, technology, and organizational design, and therefore need not align with where information is most valuable. Consequently, understanding how screening is organized—and how this organization responds to economic forces—is essential for understanding credit allocation.

Screening has historically been organized around local presence: proximity allows loan officers to generate soft information through in-person interactions, repeated contact, and local market knowledge. Technological advances relax this local-origination constraint by enabling remote processing and centralized verification, creating a national labor market for underwriting and lowering the marginal cost of loan processing. But lower processing cost does not imply comparable informational precision: whether these technologies deliver information of the same quality as proximity remains an empirical question. If they do not, productivity gains in remote processing may shift lenders' incentives, inducing substitution away from higher-precision local screening, even in markets where informational frictions are severe. In this case, technological progress can raise processing efficiency but reduce informational efficiency, tightening credit access and increasing risk.

We combine new empirical evidence with a structural framework to quantify how this tradeoff manifests in the U.S. mortgage market. Using administrative data that link loan officers to the applications they review and the loans' subsequent performance, we show that local officers achieve higher screening precision and faster processing, but lenders do not fully internalize the informational benefits in their labor allocation decision — they allocate labor elastically with respect to local wages, resulting in systematic spatial misallocation of underwriting capacity relative to mortgage demand. We then develop and estimate an equilibrium structural model in which lenders set prices before observing borrower-specific signals, borrowers self-select on posted rates, and loan officers of different screening precision generate information that determines approvals. In the model, lenders compete in mortgage pricing and in labor markets for local and remote officers, endogenously allocating information production across markets. The estimated model reveals substantial baseline credit rationing—up to 15 percent in high-risk segments—with local officers eliminating roughly *half* of this while also reducing excessively risky approvals. A technology shock that increases the processing productivity of remote officers induces lenders to substitute away from local screening, lowering informational efficiency, increasing false approvals and expected defaults, and tightening rationing for marginal borrowers despite only modest reductions in interest rates.

We begin by showing that unobserved borrower heterogeneity is economically meaningful in mortgage underwriting. After residualizing posted interest rates and ex-post default outcomes on an extensive set of hard information, higher residualized rates are strongly associated with higher residualized default rates. If hard information fully captured borrower risk, conditional pricing variation should be orthogonal to conditional default. Instead, the positive slope is consistent with classic adverse selection: among otherwise observationally similar borrowers, those that end up with higher interest rates are drawn from a pool with worse unobserved risk.

We then show that local loan officers improve informational efficiency. First, rejection decisions made by local officers depend less on hard-information variables than those made by non-local officers, as reflected in systematically lower explanatory power (R^2) of observable characteristics across increasingly saturated specifications—including borrower and loan controls, county–month and lender–month fixed effects, and loan-officer fixed effects. Second, conditional on observables, local officers reject fewer applications while the loans they approve default less ex post. This informational advantage is most pronounced at lenders with higher residualized interest rates, which presumably face more adversely selected borrower pools. Such informational advantages of proximity are concentrated among borrowers who appear risky on observables (e.g., low FICO or high DTI): local officers approve more applicants in these segments without increasing default.

Local officers also exhibit a meaningful advantage in processing speed: they process refinance applications materially faster than non-local officers. In our most saturated specifications with loan officer fixed effects, local officers move refinance applications through the underwriting pipeline more than a full day sooner, despite facing similar applicant pools and workflow constraints.

Despite these local advantages, lenders' labor choices limit where these gains are available. Counties with high loan demand 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 reliance on remote underwriting precisely where mortgage demand is most concentrated. At the lender–MSA level, a misalignment index—constructed as the difference between each market's share of applications and its share of the lender's loan officers—reveals large wedges between where demand originates and where loan officers are located in MSAs with higher loan-officer or finance-sector wages.

The reduced-form evidence shows that proximity improves both screening precision and processing speed, yet lenders allocate loan-officer labor elastically to local wages; because informational gains and labor costs vary across markets, this responsiveness produces allocations that do not necessarily coincide with where information is most valuable. These findings raise two central questions: How much credit rationing arises from lower-precision remote screening? And how do technologies that increase the processing productivity of remote officers—without improving their informational precision—reshape information production and, in turn, allocative efficiency, credit access and risk?

Addressing these questions requires a framework in which pricing, borrower selection, labor allocation, and screening interact endogenously. Posted interest rates influence not only loan volumes but the risk composition of applicants, so any model must permit adverse-selection dynamics. Screening precision depends on the composition of local versus remote officers a lender hires, implying that labor-market competition and local wage conditions affect the supply of information and, therefore, the intensity of credit rationing. Because screening capacity is finite, hiring decisions feed back into approval standards through congestion. A model that embeds these forces can trace how shifts in labor supply, borrower composition, or the productivity of remote underwriting propagate across markets and shape risk outcomes.

We develop a structural model that integrates these elements. Lenders simultaneously set posted mortgage rates and type-specific wages to recruit loan officers. Borrowers differ in demand preferences and in latent default risk conditional on observables (e.g., FICO). They choose among lenders based on posted rates and lender–market–specific amenities, and because price sensitivity varies with latent default risk, this choice generates both demand elasticities and selection on unobserved risk.

Loan officers differ in screening precision and processing capacity. 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 officers. Because screening precision is tied to the composition of officer types, wage-driven labor allocation mediates the supply of information available to the lender. Once applications arrive, lenders assign them to officer types subject to capacity limits, officers receive noisy signals of borrower risk, and lenders approve only those applications whose expected default losses fall below the posted net interest margin.

The timing reflects a key institutional feature of U.S. mortgage origination: posted rates are set before borrower-specific information is collected, so approval—rather than risk-based repricing—is the primary allocative margin. These modeling ingredients jointly determine equilibrium rates, approval thresholds, labor allocations, and lender market shares. The framework allows us to quantify the informational value of local screening, the degree of credit rationing that arises when lenders substitute toward lower-precision remote officers, and the equilibrium consequences of shocks—such as improved processing productivity of remote labor—on credit access, default risk, and welfare.

We calibrate the model to quantify baseline credit misallocation due to informational frictions, which requires constructing a full-information benchmark. The core primitives that govern this benchmark are the distribution of borrower types and the correlation between default risk and price sensitivity. The borrower-type distribution disciplines the composition of applicants within each observable risk tier, while the default–price correlation governs how adverse selection alters the pool of borrowers who remain when prices rise. We identify these objects by matching market-level origination rates, expected default patterns, and the empirical slope of residualized interest rates on residualized default.

Conditional on this information environment, we pin down the technological parameters of screening. Signal precisions for local and remote officers are identified by matching the observed rejection-rate and default-rate differences across observable risk tiers, which directly reflect the superior mapping from signals to outcomes under local screening. Officer-specific processing efficiencies are calibrated to match the observed shares of local versus remote underwriting across lender types, capturing the relative quantity productivity of remote and local labor. Remaining market-level shocks are set to align model-implied default levels with realized performance across mortgage segments. Together, these moments identify the informational and physical productivity differences that are central to understanding misallocation in approval decisions and the equilibrium consequences of substituting away from local screening.

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

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

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

Literature Review. Our paper relates to a broad literature on information frictions and credit allocation (Stiglitz and Weiss, 1981; Jaffee and Modigliani, 1969). Existing literature documents how lenders acquire and exploit 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). Our paper treats information production as an endogenous activity and studies economic forces that shape lenders' information production incentives. Literature shows that information production responds systematically to market structure, competitive pressures, and incentive design (Verrecchia, 1982; Broecker, 1990; Hauswald and Marquez, 2003; Yannelis and Zhang, 2023). We study a distinctive aspect of information production: the same input that produces informational precision—local loan-officer labor—also generates a separate form of operational efficiency, so changes that alter its cost or physical productivity create spillovers for screening quality, credit allocation, and risk. 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., 2020; Blattner and Nelson, 2021; Blattner et al., 2021, 2022). We contribute by showing that technologies that increase the physical productivity of underwriting—without improving its informational content—can induce lenders to reallocate labor away from information-rich screening. Rather than uniformly improving efficiency, such technologies weaken information production, increase rationing, and raise ex-post risk in equilibrium.

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, 2023). 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 the literature that applies structural IO and spatial economics tools to consumer finance and financial product markets. One strand of this literature uses IO tools to study 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). We combine these tools with tools from the spatial economics literature on production networks (Bernard et al., 2019; Oberfield et al., 2024; Giroud et al., 2024; Arkolakis et al., 2025), multinational production and outsourcing (Antràs et al., 2006; Costinot et al., 2012; Ramondo and Rodríguez-Clare, 2013; Boehm et al., 2019), and banking (Aguirregabiria et al., 2019; Ji et al., 2023; Maingi, 2026; D'Amico and Alekseev, 2024; Morelli et al., 2025). Our contribution to this literature is twofold. First, we show that the welfare implications of productivity and technology shocks in credit markets depend critically on how these shocks interact with equilibrium information production. Second, we provide evidence for a novel source of spatial misallocation in outsourcing and production network formulation.

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

When inputs, such as loan officer labor, are differentiated by both processing efficiency (cost to process a loan) and informational efficiency (screening precision), outsourcing decisions generate externalities on consumers that are not fully internalized by firms forming the networks. More generally, our results suggest that understanding the welfare impacts of production networks and outsourcing depends heavily on how these networks balance distinct notions of input efficiency, and that not all efficiency shocks are created equal.

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.[2] 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.[3]

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 dif-

[2] Specifically, TRID requires the restarting of the origination and disclosure process, which can delay closing and is extremely costly for banks. For example, Chase Bank offers a $5,000 guarantee that it will meet an agreed upon closing date (J.P. Morgan Chase & Co., 2025)

[3] See Appendix A for regulatory details.

ferences 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 efficiency 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: Local Loan Officers Improve Informational Efficiency

We begin by showing that local loan officers improve informational efficiency in mortgage underwriting. We proceed in three steps. First, we document that unobserved borrower heterogeneity is economically meaningful. Second, we show that local officers incorporate more soft information in processing loan applications. Third, we show that local officers' approval decisions are indeed more informative. Together, these patterns will indicate that proximity enables officers to separate safe from risky applicants in ways that improve the mapping from observable and latent borrower characteristics to realized performance.

2.1.1 Unobserved borrower risk is economically meaningful

Figure 1 plots residualized interest rates against residualized default. After conditioning on rich borrower, loan, lender, geographic, and Automated Underwriting System (AUS) recommendation–related controls, the remaining variation reflects unobserved borrower risk and lenders' pricing of that risk.[4] Under fully informative hard information, conditional pricing deviations should be orthogonal to conditional default. Instead, we observe that higher residualized rates are strongly associated with higher residualized default.

The positive slope is consistent with classic adverse selection: among observationally similar borrowers, those who end up paying higher rates are drawn from riskier pools. Higher prices disproportionately deter safer borrowers, while riskier borrowers remain—so loans priced above model-predicted levels also default above model-predicted levels. Pricing therefore does not fully compensate for latent heterogeneity in borrower quality.

[4]AUS recommendations summarize lenders' proprietary mapping from observable borrower characteristics into internal approval guidance, incorporating non-linear thresholds and interactions across FICO, DTI, LTV, income documentation, collateral, and program eligibility. Unlike simple econometric controls, AUS outputs represent the actual decision rule that lenders use at screening time, and therefore capture the high-dimensional structure of hard information as lenders themselves process it. Including AUS recommendations substantially mitigates concerns that residualized outcomes reflect econometric misspecification rather than unobserved risk, because we condition on the same hard-information summary that governs lender decision making. See Bhutta and Hizmo (2021) for a detailed discussion of how AUS reduces model misspecification by incorporating lenders' proprietary risk thresholds.

This composition effect is exactly the mechanism emphasized in adverse-selection models (Stiglitz and Weiss, 1981) and establishes a first-order informational friction—one that creates scope for additional information sources, such as proximity-based screening, to improve allocation.

2.1.2 Local officers make greater use of soft information

Table 1 compares the explanatory power of observable borrower and loan characteristics in predicting rejection decisions for applications handled by local versus non-local officers. Across six increasingly saturated specifications—including rich borrower and loan controls, county–month and lender–month fixed effects, and officer fixed effects—the R^2 is consistently and materially lower for local officers. In the baseline specification with only borrower and loan controls, the R^2 is 0.217 for non-local officers and 0.191 for local officers. With 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 holds under 100 bootstrap replications: the R^2 difference is large, negative, and statistically significant in every specification.

Because both specifications condition on lenders' own hard-information summaries—including AUS recommendations—this difference cannot be explained by omitted observables or model misspecification. Instead, it implies that local officers incorporate additional, non-codifiable information when deciding which applications to reject. In short, observable characteristics explain a smaller share of local officers' rejection decisions, consistent with greater reliance on soft information generated through proximity, interactions, and local market knowledge.

2.1.3 Local officers make more informative screening decisions

We next show that local loan officers make approval decisions that are more informative about true borrower risk. In the U.S. mortgage origination process, posted interest rates are set before borrower-specific information is collected; borrowers then choose where to apply, and loan officers decide whether to approve the application. Because lenders cannot reprice loans after observing borrower signals, informational advantages must operate through the approval margin rather than through interest rates. We therefore compare approval probabilities, ex-post loan performance, and posted interest rates for applications handled by local versus non-local officers. Specifically, we estimate the following specification:

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

where $Y_{i\ell ct}$ is either (i) an approval indicator, (ii) a default indicator for whether the underlying loan misses at least two months of payments in the two years following origination, or (iii) the interest rate. The indicator $\mathrm{Local}_{i\ell ct}$ equals one when the reviewing officer is local to the borrower. $X_{i\ell ct}$ includes detailed borrower and loan observables, including AUS recommendations. $\delta_{c\tau}$ and $\lambda_{\ell\tau}$ are county–month and lender–month fixed effects, and μ_o are loan officer fixed effects. The specification holds observables constant and isolates how proximity is associated with rejection, ex-post performance, and ex-ante pricing.

Rejection. Table 2 shows that local officers reject significantly fewer applications conditional on observables and fixed effects. For home-purchase loans, β ranges from -1.88 to -0.20 percentage points across specifications, where all estimates are statistically significant. The pattern is stronger for refinances: β ranges from -7.90 to -0.66 percentage points. These differences are not explained by borrower composition or lender heterogeneity; they imply that local officers approve strictly more applicants at the same observable risk profile.

Ex-post performance. Despite approving more applicants, local officers do not approve unobservably riskier loans. Table 3 shows that loans approved by local officers default less ex post. For home-purchase loans, the Local coefficient β ranges from -0.33 to -0.46 percentage points across specifications. For refinance loans, β becomes small and statistically insignificant once lender–month and officer fixed effects are included, consistent with refinance borrowers forming a more homogeneous pool and soft information playing a smaller role. This combination—lower rejection but lower default—is the empirical signature of more informative screening rather than looser standards.

Interest rates. Table 4 shows that conditional interest-rate differences are economically small. Because rates are posted ex ante, these coefficients cannot reflect screening choices. They instead reflect borrower sorting: local-intensive lenders attract applicants who look marginally riskier on observables. Taken together with the default results, this indicates that local officers identify borrowers whose true risk is lower than their observable profiles suggest.

Joint evidence. Finally, Figures 2–3 present these outcomes in a structured, non-parametric manner. Figure 2 shows how residualized defaults rate move with residualized rejection rates across lenders — which reflects lender's screening standard. For loans processed by non-local officers, residualized rejection and residualized default display a steep negative slope: reducing default requires tightening cutoffs and rejecting a larger share of applicants. For local officers, the slope is much flatter: they operate at lower rejection levels without commensurate increases in default. As Figure 3 shows, the difference is most pronounced in markets with high residualized interest rates, where adverse selection is strongest: non-local officers must reject sharply more applicants to maintain performance, whereas local officers preserve performance with looser cutoffs.

Taken together, the regression and nonparametric evidence demonstrate that local officers incorporate information beyond standardized observables. They approve more applicants with the same observable profiles, yet their loans default less ex post. Proximity therefore improves the mapping from borrower signals to realized credit risk, consistent with underwriting that utilizes soft information.

Several institutional features support interpreting these patterns as information rather than endogenous assignment of borrowers to local loan officers. Applications are assigned to officers before any borrower-specific soft information is generated; officers cannot pre-screen, so lenders cannot route observably identical applicants to local versus non-local officers based on unobservables that will only be learned after screening. On the borrower side, differential sorting is unlikely to rationalize the results: if local officers did not improve approval odds or identify safer borrowers, it is unclear why high- or low-risk applicants (conditional on observables) would systematically target them. The strongest alternative is taste-based preference for local interaction, but such preferences would need to be tightly correlated with latent default risk—i.e., local-preferring borrowers being systematically safer—to explain why, conditional on identical observables, local-approved loans perform better. Finally, lender–month and loan-officer fixed effects absorb differences in lender composition, workflows, and officer quality, further limiting selection-based interpretations.

2.2 Fact 2: High-Credit-Risk Borrowers Benefit More

The results in Section 2.1.3 show that, on average, local loan officers reject fewer applications while approving loans that perform better ex post. This pattern indicates 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.

Borrowers who appear risky on hard information—low FICO, high DTI, or high LTV—are the most plausible candidates. These observable measures become markedly less discriminating in high-risk regions: even within the same FICO, DTI, or LTV bucket, underlying creditworthiness varies substantially. For example, borrowers with low FICO scores can differ widely in income stability, job tenure, and informal support networks—factors not captured by standard hard-information variables.[5] When observable indicators are less informative, screening decisions necessarily rely more on soft information. Thus, if local officers indeed possess superior soft-information signals, their relative advantage should be largest precisely in these observable high-risk segments.

Table 5 formally tests whether the informational advantage of local officers is strongest for borrowers who appear risky on observables. We estimate heterogeneous effects across three hard–risk dimensions: Subprime (FICO<670), High DTI (DTI>43), and High LTV (LTV>80). Columns 1-3 shows that local officers reduce rejection rates substantially more for high-risk borrowers. Across all three splits, the interaction term is large, negative, and precisely estimated. For example, in the subprime specification, Local $\times$ Subprime $= -2.93$ p.p., several multiples of the main Local effect. High-DTI and high-LTV groups show similarly large reductions. These estimates indicate that local officers differentiate among observably risky borrowers more effectively, relaxing rejection disproportionately where hard-information signals are least informative.

Columns 4-6 show that these additional approvals do not come at the expense of higher ex-post default rates. Interaction coefficients for High-DTI and High-LTV are negative and significant: borrowers approved by local officers in these groups default less than their observables would predict. Even in the subprime split, the interaction remains negative and economically meaningful. Local officers therefore approve borrowers who look risky on observables but are safer within those buckets.

[5] 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 underlying risk.

Columns 7-9 examine posted interest rates. Because rates are set before underwriting and cannot be repriced at approval, they reflect borrower sorting rather than officer-level pricing. Consistent with this, conditional differences in posted rates between approvals screened by local and non-local officers are economically small and often statistically insignificant. This lack of meaningful variation reinforces that informational gains do not operate through rate-setting.

The nonparametric evidence in Figure 4 mirrors these patterns. Within narrow FICO, DTI, and LTV bins, local officers consistently reject fewer borrowers, with the largest gaps in high-risk regions (e.g., FICO$<$ 620, DTI$>$ 50%). Yet default outcomes remain similar to or lower than for approved loans screened by local officers in every bin, confirming that these additional approvals do not translate into worse performance. The contrast is sharpest in home-purchase markets, where borrower- and property-specific soft information is most salient.

In refinance markets, where underwriting is more standardized, the overall rejection gap between local and non-local officers is larger, but variation across risk buckets is flatter. This suggests that unobserved heterogeneity in refinance applications is spread more evenly across the credit spectrum, and local officers' informational advantage operates more uniformly.

Taken together, these results demonstrate that proximity enables local officers to extract soft information precisely where hard-information signals are weakest. They expand access for observably risky borrowers without raising default, revealing that the value of soft information—and thus geographic proximity—is highest in segments with severe unobserved heterogeneity.

2.3 Fact 3: Proximity Improves Quantity Productivity

We next show that proximity improves quantity productivity: conditional on observables, loan officers process local loan applications faster than non-local loan applications. This dimension is distinct from informational efficiency. Whereas Fact 1 documents that proximity improves the *quality* of screening, this fact shows that proximity improves the *speed* of processing. To quantify processing productivity difference, we estimate:

$$\text{Time}_{i\ell ct} = \beta^T \, \text{Local}_{i\ell ct} + \gamma^{T\prime} X_{i\ell ct} + \delta^T_{c\tau} + \lambda^T_{\ell\tau} + \mu^T_o + \varepsilon^T_{i\ell ct}, \tag{2}$$

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 t; $\text{Local}_{i\ell ct}$ 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. We estimate equation (2) separately for home purchase and refinance loans and restrict the sample to originated loans so that $\text{Time}_{i\ell ct}$ is well-defined.

Table 6 presents the results. For home purchase loans (Panel A), local officers appear faster in baseline specifications, but once lender-month and officer fixed effects are included, the advantage disappears. This suggests that the raw difference reflects lender-level workflow heterogeneity or officer-specific productivity rather than a systematic processing advantage from proximity.

For refinance loans (Panel B), the effects are larger and robust. After including lender-month fixed effects, local officers process refinance applications roughly 1–1.7 days faster than non-local officers, and the advantage persists when officer fixed effects are added. Distance-based specifications yield the same conclusion: greater physical separation between borrower and officer is associated with meaningfully slower processing.

Taken together, proximity provides a meaningful *processing-productivity* advantage in refinance markets, where workflows are standardized, borrower communication matters, and coordination frictions are especially sensitive to physical distance. By contrast, home-purchase underwriting involves property-specific frictions that attenuate processing gains from proximity. These processing-productivity gains are distinct from the informational-efficiency gains documented in Fact 1, and will be central for interpreting the effects of technological change in our model.

2.4 Fact 4: Lender Labor Choices and Spatial Imbalances in Loan-Officer Capacity

Finally, we show that lenders' allocation of local versus remote loan officers responds strongly to local labor-market conditions, generating systematic spatial mismatches between where mortgage demand arises and where screening capacity is located.

Table 7 documents a robust relationship between local wages and lender hiring by estimating the following specification:

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

Misalignment is a lender–MSA–year index defined below:[6]

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

where $\frac{m_{jkt}}{M_{jt}}$ is the share of lender j's applications from MSA k, and $\frac{l_{jkt}}{L_{jt}}$ is the share of its loan officers located there. Positive values indicate that an MSA contributes disproportionately to a lender's pipeline relative to its loan-officer capacity. The estimates suggest that higher local wages are strongly associated with larger misalignment: lenders are systematically underrepresented in high-wage MSAs, consistent with wage-driven labor allocation.

These lender-level patterns translate into pronounced geographic imbalances. Panel A of Figure 5 plots loan officers per hundred applications by county. High-demand counties tend to have comparatively few officers relative to local application volume, while some low-demand counties host more officers than demand would predict. Panel B adapts segregation-style indices used in labor and banking (e.g., Aguirregabiria et al. (2019)) to quantify the imbalance more formally. For each county, we compare its share of national loan officers to its share of national mortgage applications. The resulting county-level misalignment index reveals large and persistent wedges between demand and loan-officer capacity. Figure 6 shows that these wedges are most acute in the highest-demand deciles: counties generating a large share of national originations have too few loan officers compared to their loan demand.

2.5 Discussion

The reduced-form evidence highlights two distinct dimensions through which proximity enhances underwriting. First, proximity improves informational efficiency: local officers make approval decisions that better separate safe from risky borrowers, and these informational gains are strongest in market segments where borrowers are observably riskier. Second, proximity improves processing productivity: local officers move applications materially

[6]We do this exercise at MSA level instead of county level to match with the geographic granularity of the wage data.

faster, consistent with geographic proximity reducing coordination frictions and documentation delays.

Yet lenders do not fully internalize these informational benefits when allocating labor. Labor decisions are highly elastic to local wage conditions. Because informational gains and labor costs vary across markets, this responsiveness produces allocations that do not necessarily coincide with where information is most valuable.

Two central questions arise from these empirical patterns: How much credit rationing arises from reliance on lower-precision remote screening? And how do technologies that improve the processing efficiency of remote loan officers—without enhancing their informational precision—reshape lenders' labor allocation and, in turn, credit allocation and risk? To address these questions, we develop a model that embeds borrower selection, officer-specific screening precision, lender pricing, and endogenous labor allocation in a unified equilibrium environment.

3 Structural Model

In this section, we develop an equilibrium structural model that links the mortgage product market with the labor market for loan officers, where officers differ in their ability to resolve informational frictions. The model captures how pricing, borrower selection, screening precision, and labor allocation interact endogenously across heterogeneous markets. This framework enables us to quantify how these interactions shape credit allocation and risk, and provides the basis for counterfactual analyses of how technological changes that alter screening productivity or labor costs reorganize information production.

3.1 Model Setup

The model integrates lender pricing, borrower demand, loan officer labor supply, and heterogeneous screening technologies into a unified framework for mortgage market equilibrium. Lenders set prices and wages strategically, recognizing that these choices influence both the composition of applicants and the quality and quantity of screening capacity they attract. Borrower demand responds to pricing, where borrowers' price sensitivities can be correlated with their default risk; this gives rise to adverse selection. Officer types differ in information precision, creating variation in screening effectiveness and generating congestion and informational frictions.

3.1.1 Primitives and Timing

A market is defined by a unique combination of credit quality, product type, county, and year-quarter. Markets are thus segmented by ex-ante, observable hard information, such as FICO, LTV, and DTI. 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 i is as follows:

1. **Lenders' Ex-Ante Information Sets:** Prior to the start of the game, each lender j knows the joint distribution of price sensitivity and default risk, given by $F_i^{x,\theta}(\cdot)$. Lenders know the distribution of borrower' idiosyncratic tastes, given by $F_i^{\epsilon}(\cdot)$. Lenders know the number of loans each loan officer type can process (physical efficiency), e_{ijk}, their screening precision (informational efficiency), σ_{ik}, and the distribution of loan officers' idiosyncratic tastes, $F_{ik}^{A}(\cdot)$. At this stage, lenders do not observe detailed borrower-specific private information, such as verified income documents, appraisals, employment checks, and/or third-party verifications.
2. **Pricing and Hiring:** Given the common knowledge, 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 Bertrand-Nash equilibrium in prices and wages. Lenders anticipate how their choices affect borrower demand, their applicants' type distributions, and labor supply.
3. **Application:** Borrower η in market i observes all posted rates $\{r_{ij}\}_{j\in\mathcal{J}_i}$ by all active lenders in market i. Each borrower η has a price sensitivity coefficient θ_η and a continuous latent risk type x_η that summarizes all default-relevant characteristics, including those not observable at the pricing stage. Borrowers choose a single lender to apply to based on posted interest rates r_{ij}, observable lender characteristics and demand

shifters ξ_{ij}, and idiosyncratic preferences ϵ_η. There are $\mathcal{M}_i$ total borrowers and a share s_{ij} apply to lender j.

4. **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} and idiosyncratic location and lender preferences. This determines the mass l_{ijk} of type-k officers working for each lender j.

5. **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 $\mathrm{Var}(\tilde{x}_{\eta k}) = \sigma^2_{ik}$. Local officers have 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, as well as soft-information collection.

6. **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 d_i maps default types to default rates and 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.

7. **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, equal to net interest margins net of realized losses. Lenders pay labor costs of $\sum_{k\in\mathcal{K}} w_{ijk}l_{ijk}$. Denied applicants receive their outside option minus a rejection cost. 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 ex-post 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.

3.1.2 Borrower Demand

Each borrower η in market i chooses which lender to apply to based on posted interest rates and lender characteristics. Borrower η receives the following random indirect utility from applying to lender j:

$$u_{ij\eta} = -\theta_\eta r_{ij} + \xi_{ij} + \epsilon_{ij\eta}, \tag{4}$$

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:

$$s_{ij}(\theta, r_{ij}) = \mathbb{P}_\eta\left(\bigcap_{k\in\mathcal{J}}\left\{-\theta_\eta(r_{ij} - r_{ik}) + (\xi_{ij} - \xi_{ik}) \geq \epsilon_{ik\eta} - \epsilon_{ij\eta}\right\}\right). \tag{5}$$

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

$$s_{ij}(r_{ij}) \coloneqq \int_{-\infty}^{\infty} s_{ij}(\theta, r_{ij}) F_i^\theta(d\theta), \tag{6}$$

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

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:

$$F_{ij}^{\theta}(t;r_{ij}) = \big(s_{ij}(r_{ij})\big)^{-1}\int_{-\infty}^{t} s_{ij}(\theta, r_{ij})F_i^{\theta}(d\theta). \tag{7}$$

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 (lower 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. Formally, the selected distribution of default risk types can be written as:

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

Note that, in the special case where $x \perp\!\!\!\perp \theta$, the joint distribution is multiplicatively separable, and the integral immediately collapses to $F_i^x(t)$. If x and θ are negatively correlated, lenders face adverse selection in posted rates, where increasing their posted rate leads to a worse pool of applicants.

3.1.3 Loan Officer Labor Supply

Loan officers choose where to work and live to maximize their indirect utility. There is a mass $\bar{L}$ of potential loan officers, whose indirect utility depends on the combination of where they choose to live ("type" k) and the lender-market pair for which they process loans (market i, lender j):

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

where w_{ijk} is the wage offered by lender j to location-k officers in market i, and $a_{ij\lambda}$ is an idiosyncratic preference shock that varies across officers. Following Giroud et al. (2024), we assume the preference shocks $a_{ij\lambda}$ follow a *joint* Fréchet distribution, with standard Fréchet marginal distributions and a triple-nested Gumbel copula, featuring region-varying outside options. For simplicity of exposition, we consider the case where loan officers are atomistic in the overall labor market.[7]. Let ϱ represent the shape parameter of the inner-most nest.

[7]There are approximately 320,000 loan officers in the United states, compared to approximately 136 million private sector employees (Bureau of Labor Statistics, 2023). Formally, this is the limit as $\bar{L}$ and the indirect utility from choosing an outside option jointly approach infinity. We provide a formal characterization of the distribution, and

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_{ik}, which determines the precision of the signals they receive about borrower risk. Lower σ_{ik} indicates more precise screening. This maps to the informational advantage of local officers documented in Fact 2 (Section 2.1), 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 4 (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 2 (Section 2.2), 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_{ik} matters most for distinguishing between fundamentally different borrower types within the same observable risk category.

3.1.4 Lender's Problem

Lenders maximize expected profits under rational expectations. Each lender j chooses an interest rate r_{ij}, type-specific wages $\{w_{ijk}\}_{k=1}^{K}$, and loan officer assignment shares $\mathcal{S}_{ijk}$ to maximize profits across all markets, subject to loan officer capacity constraints.

limiting properties, in the model appendix.

The lender solves the following constrained optimization:

$$\max_{\left\{r_{ij},\left\{w_{ijk},\mathcal{S}_{ijk}\right\}_{k=1}^{K}\right\}} \sum_i \sum_{k=1}^{K} \Bigg\{ \mathcal{M}_i s_{ij}\left(r_{ij}\right) \mathcal{S}_{ijk} \mathbb{P}_{ij}\left\{orig; r_{ij}, k\right\} \Big(r_{ij} - f_{ij} - \mathbb{E}\left[d_{ij}; r_{ij}, k, orig\right]\Big)$$
$$- w_{ijk} l_{ijk} \Bigg\}$$
$$\text{s.t. } \mathcal{M}_i s_{ij}\left(r_{ij}\right) \mathcal{S}_{ijk} \leq l_{ijk} e_{ijk} \quad \forall k \tag{10}$$

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, as shown in equation (6). 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 $(r_{ij} - f_{ij} - \mathbb{E}[d_{ij}; r_{ij}, k, orig])$ captures the net interest margin per originated loan.[8] 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_{ik}^2. 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.

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

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 We show in the model appendix that the cost of processing $\tilde{l}_{ijk}$ loans with officers of type k can be expressed as $A_{ijk}^{-1}\tilde{l}_{ijk}^{1+\frac{1}{\varrho}}$, where A_{ijk} is a composite cost parameter that is proportional to $\frac{e_{ijk}}{\Psi_k}$. with local wage index Ψ_k that incorporates differences in region-specific outside options. 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}:

$$\sum_{k=1}^{K} A_{ijk}^{\varrho} p_{ijk}^{\varrho+1} \left(r_{ij} - f_{ij} - d_{ijk}\right)^{\varrho} \Bigg\{ r_{ij} \left(\varepsilon_{p_{ijk},r} + \varepsilon_{s_{ij},r} + 1\right) - f_{ij} \left(\varepsilon_{p_{ijk},r} + \varepsilon_{s_{ij},r}\right) \\ - d_{ijk} \left(\varepsilon_{p_{ijk},r} + \varepsilon_{s_{ij},r} + \varepsilon_{d_{ijk},r}\right) - \varepsilon_{s_{ij},r} \bar{c}_{ij} \left(r_{ij} - f_{ij} - d_{ijk}\right) \Bigg\} = 0 \quad (11)$$

where $\varepsilon_{x,y} = \frac{\partial \ln x}{\partial \ln y}$ denotes the elasticity of x to y, and the capacity cost shifter is defined as:

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

Further, the bank sets labor shares following:

$$\mathcal{S}_{ijk} = \frac{A_{ijk}^{\varrho} p_{ijk}^{\varrho} (r_{ij} - f_{ij} - d_{ijk})^{\varrho}}{\sum_k A_{ijk}^{\varrho} p_{ijk}^{\varrho} (r_{ij} - f_{ij} - d_{ijk})^{\varrho}} \quad (12)$$

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 hiring 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:

$$r_{ij} = \underbrace{f_{ij}\frac{\varepsilon_{p_{ij},r} + \varepsilon_{s_{ij},r}}{\varepsilon_{p_{ij},r} + \varepsilon_{s_{ij},r} + 1}}_{\text{Funding cost component}} + \underbrace{d_{ij}\frac{\varepsilon_{p_{ij},r} + \varepsilon_{s_{ij},r} + \varepsilon_{d_{ij},r}}{\varepsilon_{p_{ij},r} + \varepsilon_{s_{ij},r} + 1}}_{\text{Default risk component}} + \underbrace{p_{ij}^{-1}\left(\frac{\mathcal{M}_i s_{ij}}{\sum_{k=1}^{K} A_{ijk}^{\varrho}}\right)^{\frac{1}{\varrho}} \frac{\varepsilon_{s_{ij},r}}{\varepsilon_{p_{ij},r} + \varepsilon_{s_{ij},r} + 1}}_{\text{Capacity cost component}} \tag{13}$$

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. Notably, under imperfect competition, the degree to which each of the three components is priced in equilibrium varies, due to adverse selection ($\varepsilon_{d_{ij},r}$) and the fact that screening costs are paid for rejected loans.

The main difference between the general case (11) and the simplified case (13) is that the general case involves a mixture over elasticities and net interest margins across different 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_{ik}^2 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, i.e.:

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

The demand shocks $\epsilon_{ij\eta}$ are i.i.d. Type I extreme value. We allow for physical and informational efficiency of local and non-local officers to differ, i.e. $\sigma_{i,local} \neq \sigma_{i,distant}$ and $e_{ij,local} \neq e_{ij,distant}$. Finally, we assume d_i follows a probit process with correlated shocks, where borrower η defaults if $x_\eta \leq \epsilon_\eta + \epsilon_{it}$, where ϵ_η and ϵ_{it} are iid standard normal idiosyncratic and aggregate shocks, respectively.[9]

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). For the quantitative model, we account for the fact that there are significant recoveries on defaulted loans; we thus calibrate losses-given-default to 17.7% using Fannie Mae securitization data.[10]

Internally Calibrated Parameters. We consider 6 product type buckets:

$\{subprime, near-prime, prime\} \times \{purchase, refi\}$ and two types of loan officers $\{local, nonlocal\}$. 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 sensitivity, 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 Figures 1 and A3.

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 full pooling equilibrium, where all loans are originated, and improvements to informational precision are used to decrease pooled originations, rather than decrease credit rationing.

[9]The inclusion of aggregate shocks is motivated by correlated default rate spikes across credit grades and loan purposes, as shown in Appendix Figure A3

[10]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.

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 (11). 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 9. 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.[11] 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. Finally, our calibration involves setting $\gamma = -0.4$, indicating adverse selection, setting $\sigma_{local} = 1.4$, and setting $\sigma_{distant} = 2.45$, consistent with a substantial local informational advantage.

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 5, although the relevant coefficients are estimated somewhat noisily.

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

Note that these results are driven by the differences in borrower 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 $E[d(x_\eta)] \leq r_{ij} - f_{ij}$, 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 10 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 B shows that local officers achieve this expansion in credit while also *reducing* pooled origination, consistent with their default advantage. Because default risk in the model is highly convex, lenders behave conservatively: pooled origination is quantitatively small relative to rationing. Thus, informational advantages operate mainly through reductions in credit rationing.

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.[12]

In counterfactual 1, only the Brick and Mortar bank receives the efficiency improvement. 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 (12), shocked banks substitute distant labor for local labor.

Credit Provision. Table 11 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 11, 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.

[12]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.

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

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

Aggregate Shocks. We next examine how the technology shock interacts with aggregate conditions. The model allows us to compute realized default rates across aggregate states. Figure 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.[14]

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

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

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

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.

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.

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

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

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

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

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

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

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

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

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

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

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

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

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

Broecker, Thorsten, 1990, Credit-worthiness tests and interbank competition, *Econometrica: Journal of the Econometric Society* 429–452.

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.

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

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

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

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

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

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

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

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* .

Hauswald, Robert, and Robert Marquez, 2003, Information technology and financial services competition, *The Review of Financial Studies* 16, 921–948.

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* .

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

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* .

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

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

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

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

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.

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

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

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

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

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

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

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

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

Tirole, Jean, 2023, Competition and the industrial challenge for the digital age, *Annual Review of Economics* 15, 795–828.

Verrecchia, Robert E, 1982, Information acquisition in a noisy rational expectations economy, *Econometrica: Journal of the Econometric Society* 1415–1430.

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

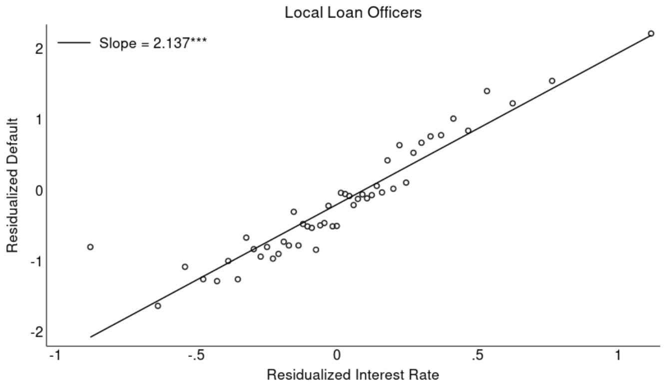
Wang, Lulu, 2025, Regulating competing payment networks, *Kilts Center at Chicago Booth Marketing Data Center Paper* .

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

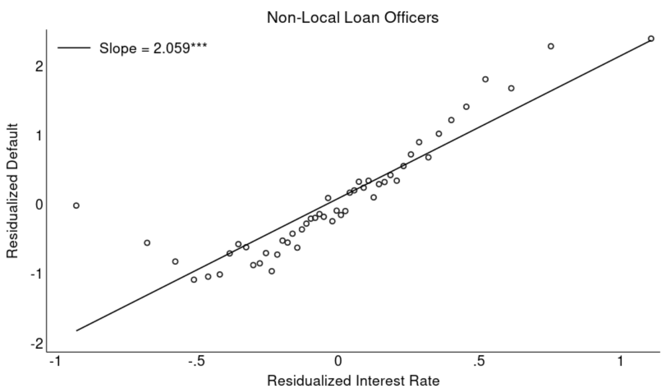
Xiao, Kairong, 2020, Monetary transmission through shadow banks, *The Review of Financial Studies* 33, 2379–2420.

Yannelis, Constantine, and Anthony Lee Zhang, 2023, Competition and selection in credit markets, *Journal of Financial Economics* 150, 103710.

Figure 1. Adverse Selection: Interest Rate and Default Relationship



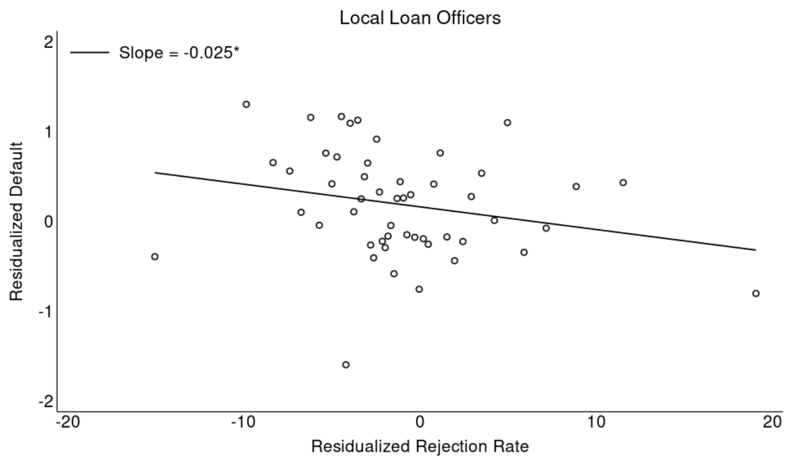
(a) Local Loan Officer



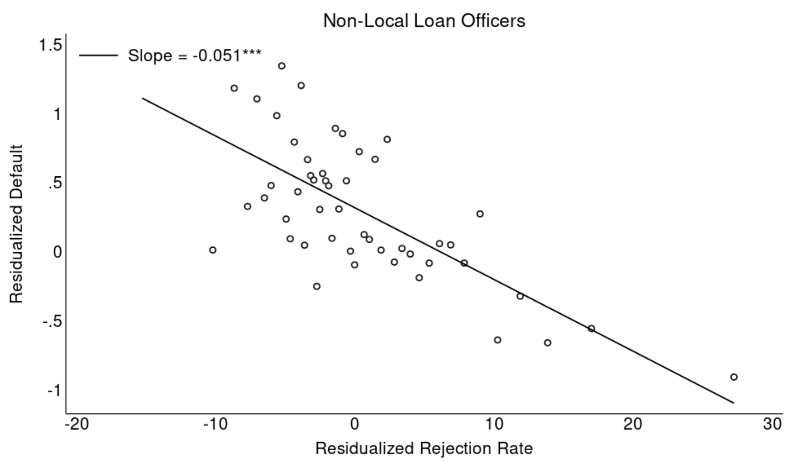
(b) Non-Local Loan Officer

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 2. 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.

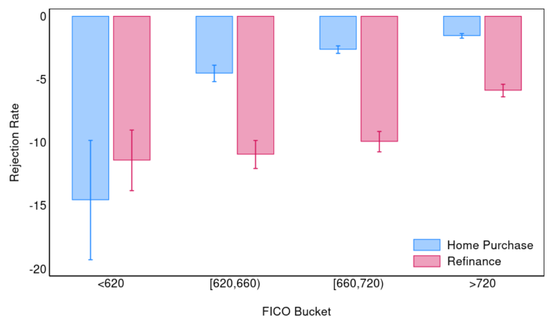
Figure 3. Loan Approval Standard and Ex-Post Default by Loan Rates

(a) High Rate Sample

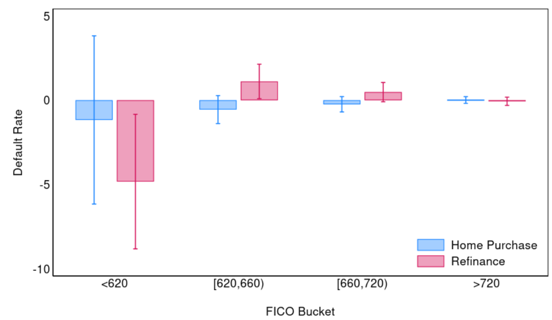
(b) Low Rate Sample

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.

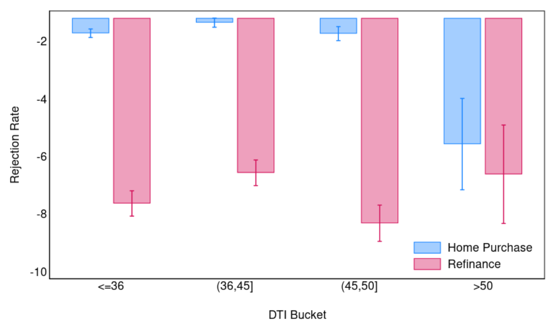
Figure 4. Rejection and Default by Hard Credit Information: Local vs Non-Local



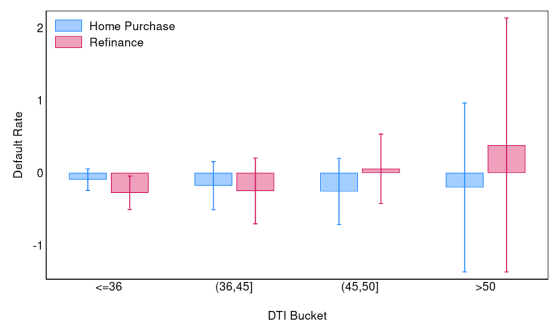
(a) Rejection-FICO



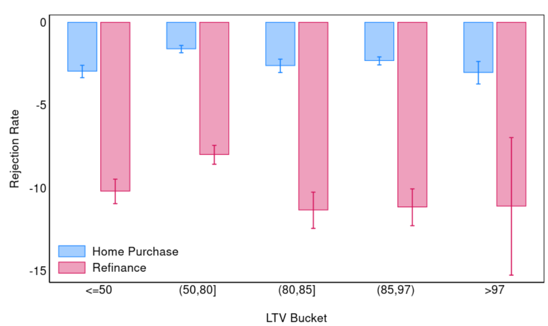
(b) Default-FICO



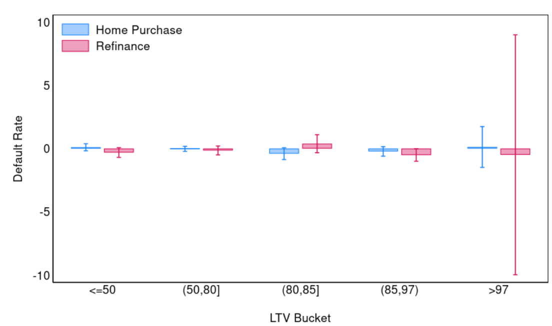
(c) Rejection-DTI



(d) Default-DTI



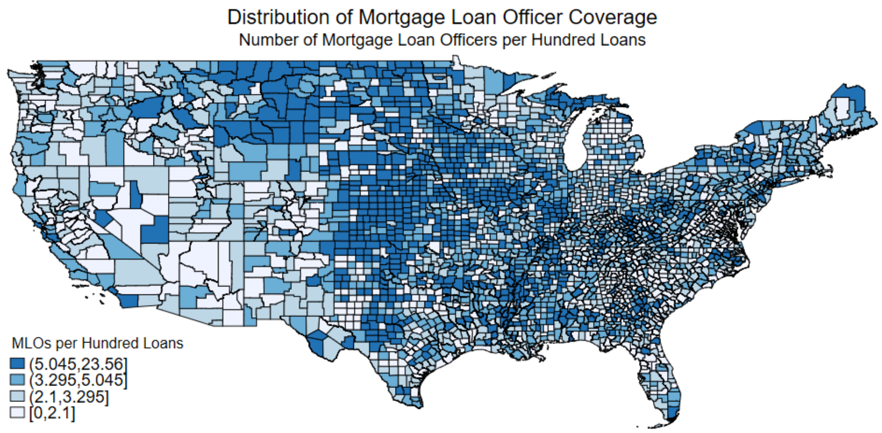
(e) Rejection-LTV



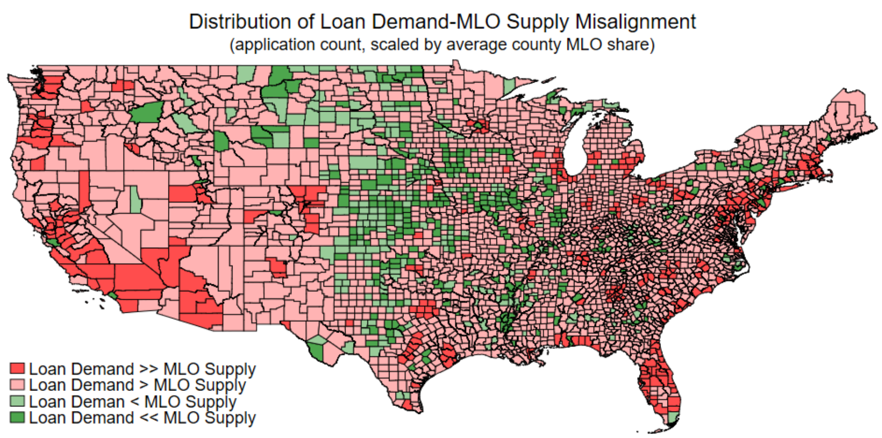
(f) Default-LTV

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

Figure 5. Geographic Misalignment of Mortgage Demand and Loan Officer Supply



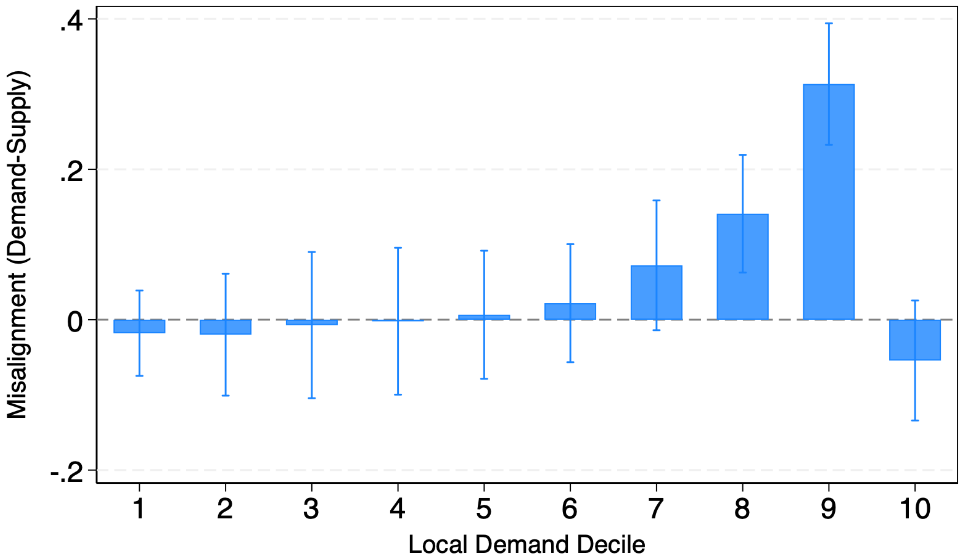
(a) Geographic Distribution of Mortgage Loan Officer Coverage



(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.

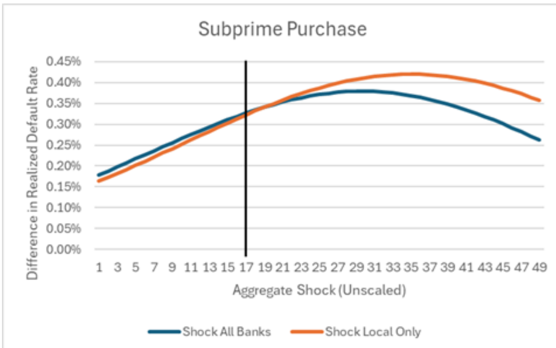
Figure 6. Loan Demand-Loan Officer Supply Misalignment Across Counties



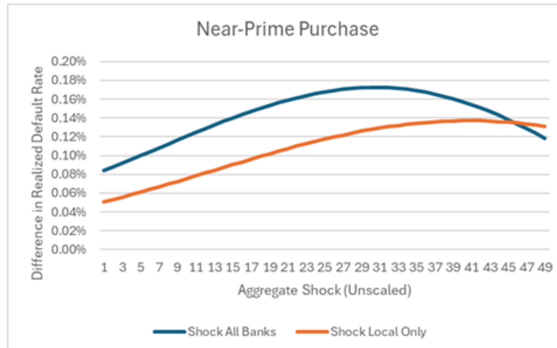
(a) Excess Loan Demand Relative to Loan Officer Supply

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.

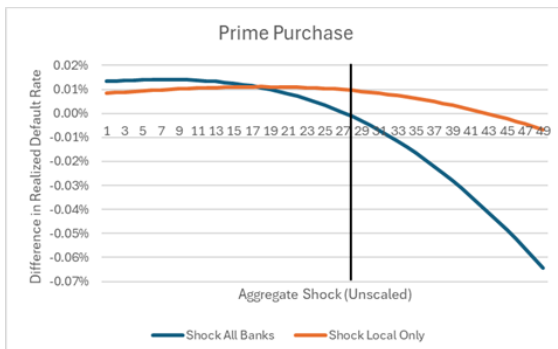
Figure 7. Technology Shock and Aggregate Risk



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

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

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

Table 6: 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 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: Interally Calibrated Default Parameters

This table reports calibrated default rates, latent heterogeneity, and aggregate shock values for each market for 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). See Section 3.2 for more details.

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

Table 10: 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 11: 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 ≥ 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).[15] Although internal procedures vary across lenders, the fundamental sequence is uniform:

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

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

A.1 Rate Setting Prior to Information Acquisition

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

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

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

A.2 Formal Application and Information Collection

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

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

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

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

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

A.4 Local Versus Remote Loan Officers

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

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

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

[20] Freddie Mac Seller/Servicer Guide, Section 5101.2 (2024).

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

A.5 Origination and Final Rate Setting

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

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

B Sample Construction and Data Cleaning

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

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

[21] MBA, "Technology & Origination Report," 2020.
[22] 12 C.F.R.§1026.19(e)(3).

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

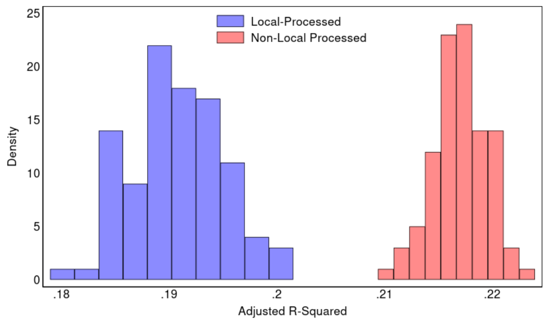
Next, we merge the HMDA originations to monthly loan-level performance data from the Black Knight McDash servicing database. Following the methodology of Rosen (2011), we treat a loan in HMDA and a loan in McDash as the same origination only when several characteristics match almost exactly. First, the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

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

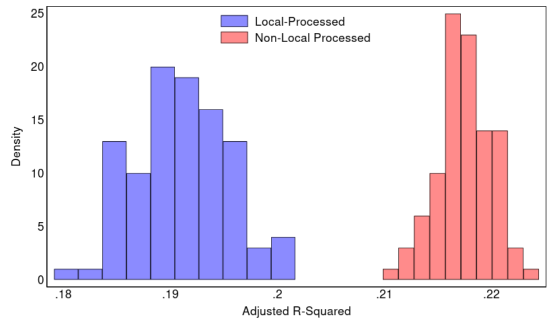
The final dataset links each loan application to (i) borrower and loan characteristics at the time of application, (ii) the lender and loan officer responsible for the file, (iii) the physical location of the loan officer obtained from NMLS, (iv) underwriting decisions and posted interest rates, and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

C Additional Figures and Tables

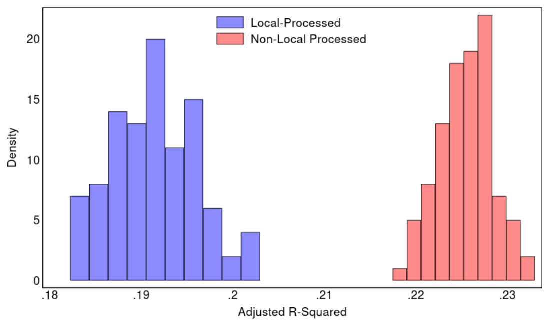
Figure A1. Explanatory Power of Hard Information in Loan Rejection



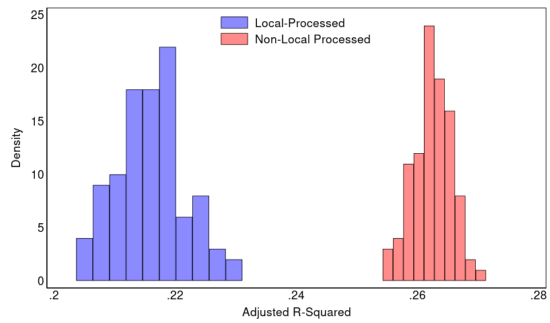
(a) Model 1



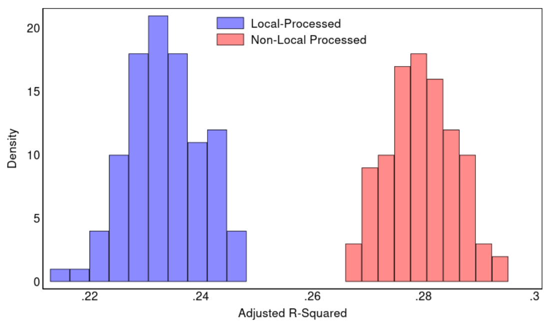
(b) Model 2



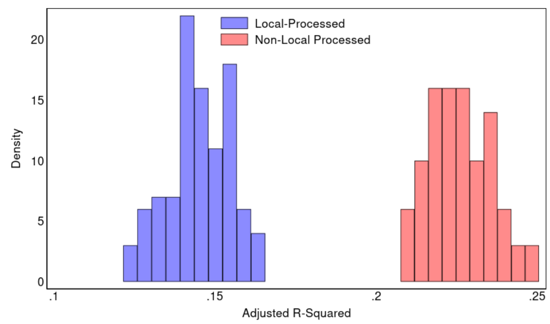
(c) Model 3



(d) Model 4



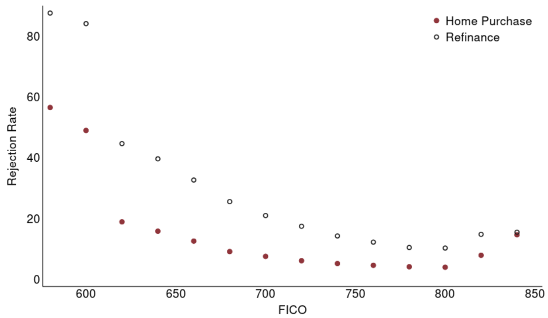
(e) Model 5



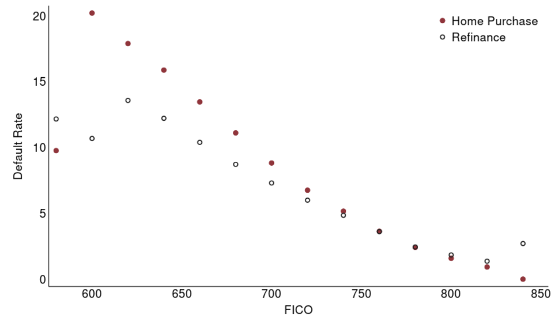
(f) Model 6

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.

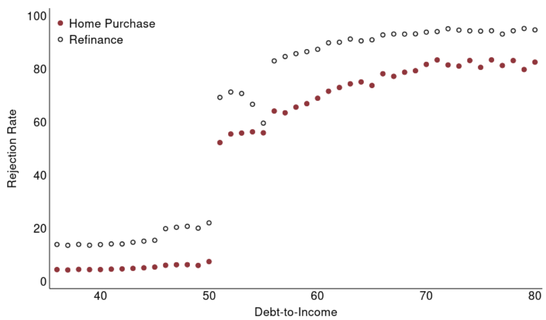
Figure A2. Rejection and Default by Hard Credit Information



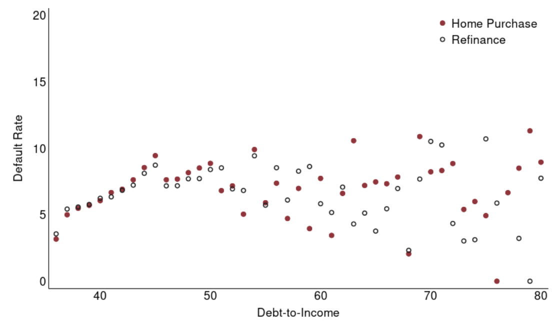
(a) Rejection-FICO



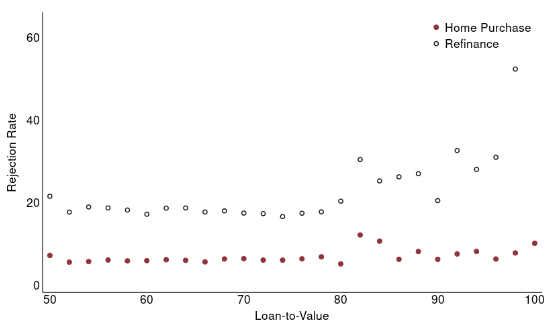
(b) Default-FICO



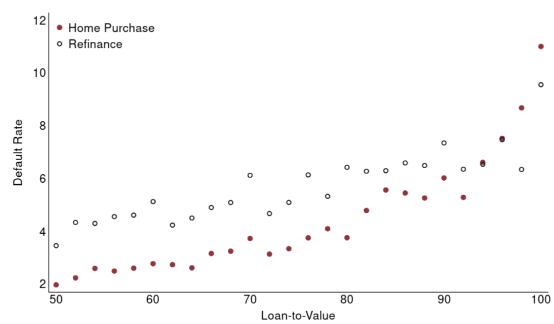
(c) Rejection-DTI



(d) Default-DTI



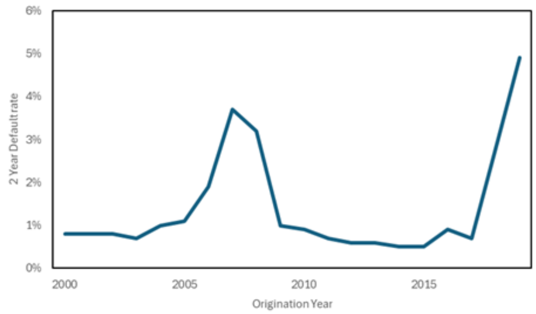
(e) Rejection-LTV



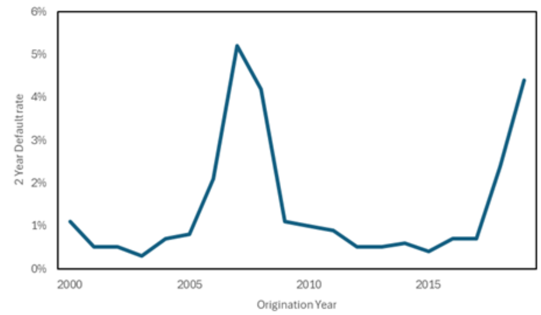
(f) Default-LTV

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.

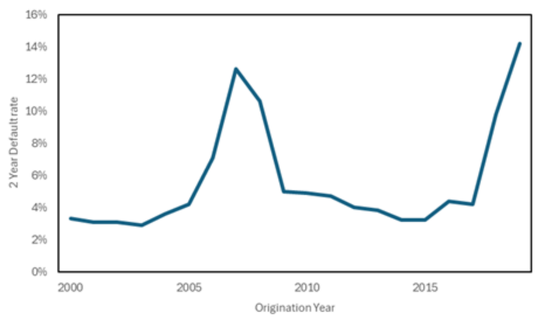
Figure A3. Time Series of Default by Credit Grade and Loan Purpose



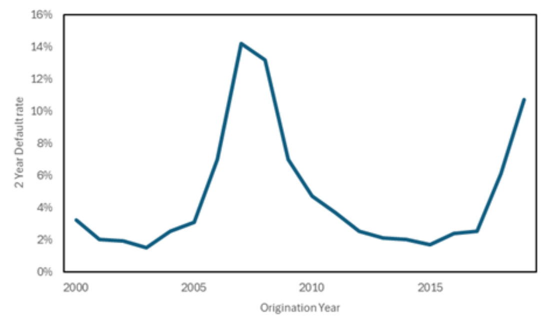
(a) Prime Purchase Market



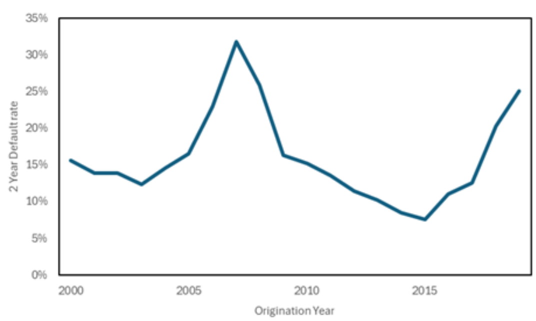
(b) Prime Refi Market



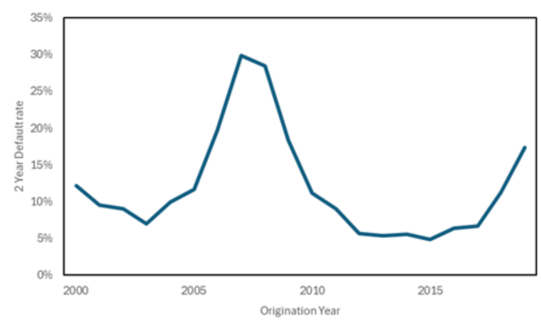
(c) Near-Prime Purchase Market



(d) Near-Prime Refi Market



(e) Subprime Purchase Market



(f) Subprime Refi Market

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

D Model Appendix

D.1 Labor Supply Function

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

Environment. There is a mass of $\bar{L}$ potential loan officers, who can choose to be either a loan officer, or an outside option. There are N_k potential outside options in region k, which pay an average (amenity) wage of $w_{k,oo}$[23] We assume that the CDF of amenity draws, A_{ijk} is given by:

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

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

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

,

for local wage index Ψ_k following:

[23] As in Giroud et al. (2024), there is a love-for-variety adjustment on average wages which is a function of market structure, but the resulting wage index is directly proportional to average wages

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

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

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

.

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