Valuation Uncertainty and Lender Frictions in Mortgage Refinancing *

Francisco Amaral
University of Zurich, Swiss Finance Institute

Gianmarco Ruzzier
Banco de España

This version: November, 2026

Click here for the latest version

Abstract

Many borrowers do not refinance even when doing so would lower their payments. We show that this is partly driven by a lender-side friction: the difficulty of valuing a house, or collateral uncertainty, restricts refinancing. Using loan performance and property records, we find that borrowers refinance less where homes are harder to value, even when refinancing would yield net savings. Mortgage application data show that greater uncertainty about local home values predicts a higher likelihood that lenders turn down a refinance application. In these markets, refinance appraisals rely on fewer and more distant comparable sales, making valuations less precise. Foreclosed homes also sell at larger discounts, pointing to weaker recovery in default. When mortgage rates fall, borrowers are less likely to refinance where homes are harder to value, leaving substantial savings unrealized.

Keywords: Housing, Securitization, Banks, Mortgages

JEL Classification codes: E32, E44, G21, G51

*Email: amaralfranc@gmail.com; gianmarco.ruzzier@bde.es. This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1-390838866. This work benefited greatly from comments by Andrea Caggese, Geraldo Cerqueiro, Eduardo Davila, Christian Eufinger, Andreas Fuster, Pedro Gete (discussant), Marco Giometti (discussant), Arpit Gupta, Dmitry Kuvshinov, Ralf Meisenzahl, David Martinez-Miera, Lakshmi Naaraayanan, Steven Ongena, Marco Pagano, Björn Richter, Glenn Schepens (discussant), Javier Suarez, Victoria Vanasco and Anthony Lee Zhang. We also thank the conference participants at the 8th Summer Conference on Financial Intermediation and Corporate Finance, BIS – CEPR – Gerzensee- SFI Conference on Financial Intermediation, SGF Conference 2026, the 14th EFI Network, at the X Madrid Barcelona Workshop on Banking and Corporate Finance, and at Seminars at the Bank of Spain, for useful comments and suggestions. The views expressed do not necessarily reflect those of the Banco de España or the Eurosystem. All remaining errors are our own.

1. Introduction

When mortgage rates fall, homeowners can lower their monthly payments by refinancing. Yet many stay in above-market contracts even when refinancing would save them money (Keys et al., 2016). Whether households refinance matters for their spending and default risk (Abel and Fuster, 2021) and for the transmission of monetary policy (Eichenbaum et al., 2022). A large literature attributes incomplete refinancing to borrower-side frictions such as inattention, limited financial literacy, and inertia (Andersen et al., 2020). But even when borrowers want to refinance, the lender can turn down their application for a new loan backed by the same house. For a purchase mortgage, the lender can anchor its valuation to the price agreed on by the buyer and seller.[1] For a refinance, the lender must make a valuation of the same home without a new sale price.[2] We study collateral value uncertainty as a lender-side friction and ask how it affects the likelihood that an existing mortgage is refinanced.

Following Jiang and Zhang (2025), we measure this uncertainty by how widely sale prices vary among homes with similar observed characteristics. Drawing on persistent differences in valuation difficulty across housing markets, we find that a one-standard-deviation increase in uncertainty is associated with a decline of -0.91 percentage points in five-year prepayment among borrowers in the Freddie Mac data who could have saved money by refinancing, even after closing costs. The same pattern holds when we use linked property transaction and mortgage records to identify refinancing separately from other forms of prepayment and compare homes within the same ZIP code. Lenders also deny refinance applications 2.2% more often for a one-standard-deviation increase in collateral uncertainty (9.3% relative to the mean), comparing applications to the same lender in the same year and controlling for borrower and loan characteristics.

Among denials, lenders list collateral first more often where homes are harder to value. Valuation uncertainty can affect both loan eligibility and expected recovery in default. In these areas, fewer and more distant comparable sales make refinance appraisals less precise. A low

[1] In 2015 during the Presidential Mortgage Group, Sales manager Matt Hodges reported that more than half of his refinance appraisals had come in low over the previous few weeks, while none of his purchase appraisals had. He pointed to a difference in what the appraiser sees. A purchase appraisal comes with the sale contract, which gives the appraiser a price to work from, while a refinance request "cannot carry any guidance about value" (Mortgage News Daily, March 20, 2015).

[2] In its report on lending and commercial real estate, Kansas City Fed warns that "challenges in valuing properties reportedly exacerbated headwinds for refinancing activity" (Beige Book, Tenth District, May 31, 2023).

appraisal can push the loan-to-value ratio above the lender's limit. Foreclosed homes also sell at larger discounts. A low foreclosure price can leave the lender with a loss, making it less willing to refinance even when the appraised value supports the requested loan.

These lender-side frictions can keep borrowers in higher-rate mortgages even when market rates fall. When market rates fall one percentage point relative to existing mortgage rates, the estimated rise in quarterly prepayment is about 55 percent larger in the lowest uncertainty decile than in the highest. We quantify the dollar value of the refinancing opportunities borrowers miss because of collateral value uncertainty. In 2011, after a sharp fall in mortgage rates, the estimated cost was about $15,300 per missed refinance in present value, net of closing costs.

In the United States, most borrowers with long-term fixed-rate mortgages can refinance without paying a prepayment penalty. When market rates fall below the rates on their existing loans, borrowers have an incentive to exercise that option and lower their monthly payments. This can generate waves of refinancing. To refinance, however, borrowers must qualify for a new loan, and the lender caps the loan amount relative to the home's appraised value.[3] To measure the valuation difficulty lenders face when applying this loan-to-value limit, we use ZTRAX sale and assessment records to estimate how widely prices vary after accounting for differences in size, age, and location. We use the same ZTRAX dataset to identify refinances by linking property transactions to later mortgage records for the same home. Freddie Mac loan histories let us estimate how much borrowers could save by refinancing and observe whether they prepay. We use Home Mortgage Disclosure Act (HMDA) records from 2007 to 2024 to observe which refinance applications lenders deny and the reasons they report.

We first examine whether Freddie Mac purchase mortgages originated between 2000 and 2024 ever prepay while we observe them. We compare loans across three-digit ZIP codes within the same commuting zone and origination quarter, accounting for borrower and loan characteristics and differences across sellers in each quarter. A one-standard-deviation increase in collateral value uncertainty is associated with a -0.35 percentage-point decline in prepayment. We then ask whether the pattern holds among borrowers who could have saved money by refinancing, even after closing costs. Starting with mortgages outstanding in January of each

[3] During the 2009 refinancing wave, Richmond lenders found refinance demand "hard to keep up with," and the process was taking longer than usual "due to more stringent appraisals and thorough underwriting." The Federal Reserve's national summary reported the same delay across Districts (Beige Book, April 15, 2009).

year, we find that among these borrowers, a one-standard-deviation increase in uncertainty is associated with -0.83 percentage points less prepayment within three years. The gap persists through year five. The estimates change little when we account for how much each borrower could save.

Selling a home can trigger prepayment, so we use ZTRAX to examine refinancing separately. We link home purchases from 2000 to 2020 to later mortgage records for the same property and follow each property until its next recorded sale or the end of the data. We count mortgages explicitly coded as refinances and other recordings consistent with refinancing based on their timing, amount, and loan type. Among similar homes bought in the same ZIP code and quarter, a one-standard-deviation increase in property-level collateral value uncertainty is associated with a decline of -0.76 percentage points in the probability of a qualifying refinance, accounting for how long we observe each property.

HMDA records let us observe lenders' decisions once borrowers have applied to refinance. We compare applications to the same lender in the same year across ZIP codes with different collateral value uncertainty, accounting for borrower and loan characteristics and conditions in each county and year. A one-standard-deviation increase in uncertainty is associated with an increase of 1 percentage point in the probability of denial. The pattern holds when we restrict the sample to conforming loans and when we use the atypicality of the local housing stock as an instrument for uncertainty. Among the applications lenders deny, they list collateral as the first reason more often where homes are harder to value.

To connect lenders' decisions on individual applications to the refinancing gap, we use HMDA application outcomes to estimate how many fewer refinances would occur over five years per 100 mortgages outstanding at the start. Under our accounting assumptions, that projected reduction accounts for about 44% of the five-year Freddie Mac prepayment gap. Most of that contribution comes from lenders turning down borrowers who had already applied to refinance. We find a similar result in ZTRAX, where we observe refinancing separately from other forms of prepayment.

In the FHFA appraisal records, purchase appraisals deviate more from contract prices where collateral values are more uncertain, and refinance appraisals draw on fewer and more distant comparable sales. Both findings point to less precise valuations. For a given loan amount, a

low appraisal raises the loan-to-value ratio and can make the lender's limit bind.[4] In CoreLogic repeat-sales records, distressed properties sell at larger discounts in higher-uncertainty counties after accounting for local house price growth. This is consistent with lenders expecting to recover less in default, which can make them less willing to approve a refinance even when it meets the appraisal-based limit.

Following the same starting group of Freddie Mac mortgages, we find that only about one-fifth of the five-year prepayment gap closes by year ten. Borrowers who refinance later can spend years making higher payments while they wait. We estimate about 34,500 additional denials in 2011 by applying the denial gaps relative to the lowest uncertainty decile to that year's applications. We value a missed refinance in each decile using Freddie Mac mortgages outstanding in January 2011. At the prevailing refinance rate, we discount the payment reductions over the remaining mortgage term, subtract a closing cost of 2 percent of the outstanding balance, and average the savings across loans with a positive net gain. Applying these values to the additional denials yields about $528 million in foregone savings for 2011.

Related literature. First, we contribute to the literature on mortgage refinancing by identifying collateral value uncertainty as a lender-side friction. Existing work on incomplete refinancing emphasizes borrower behavior, including inattention and inertia (Keys et al., 2016; Agarwal et al., 2016; Andersen et al., 2020; Belgibayeva et al., 2025; Berger et al., 2025). Other work shows that closing costs and income documentation requirements (DeFusco and Mondragon, 2020) and low home equity (Abel and Fuster, 2021; Beraja et al., 2019) limit refinancing. We show that mortgages prepay less often where collateral is harder to value, even among borrowers who could save by refinancing. Among borrowers who apply, lenders deny more applications in higher-uncertainty areas, and the relationship holds in specifications that account for loan-to-value bins.

Second, we contribute to the literature on collateral value uncertainty by studying the refinancing of existing mortgages. Jiang and Zhang (2025) show that noisier appraisals and lower expected recoveries restrict mortgage credit at purchase. When borrowers seek to replace an

[4] Richmond Fed contacts attributed appraisals that missed the agreed sale price to "limited comparable home sales" (Beige Book, Fifth District, November 29, 2017)

existing mortgage, lenders must value the same home without a new sale price. We use appraisal and foreclosure data to show how both channels can constrain refinancing, and we quantify the household savings lost when borrowers miss refinancing opportunities. A companion paper (Amaral et al., 2026) shows how the difficulty of selling mortgages backed by uncertain collateral shapes which lenders serve local markets.

Third, we contribute to the literature on monetary policy transmission through mortgage refinancing. Beraja et al. (2019) show that regional differences in home equity shape the refinancing response to lower rates, and Eichenbaum et al. (2022) show how that response depends on the distribution of refinancing incentives. We document a lender-side source of geographic variation in this transmission. Where collateral is harder to value, lenders are more likely to deny refinance applications and prepayment responds less to a given decline in market rates.

Section 2 describes the institutional setting, data, and measure of collateral value uncertainty. Section **??** presents the prepayment, refinancing, and denial results, links the denial estimates to the refinancing gap, and examines the appraisal and recovery channels. Section 4 examines how prepayment responds to lower rates and estimates foregone refinancing savings. Section 5 concludes.

2. Data and measurement

This section describes the data sources and the construction of the paper's main outcome variables as well as the key explanatory variable, collateral value uncertainty, and its instrumental variable. Data cleaning steps (e.g., outlier handling, missing-value rules, and sample exclusions by reporting quality) are delegated to the appendix.

2.1. Refinancing outcomes

We measure realized refinancing behavior using the Freddie Mac Single-Family Loan-Level Dataset, which provides loan-level origination characteristics and monthly performance outcomes for loans acquired by Freddie Mac. Because refinancing is not directly observed as a distinct label in loan performance data, we use voluntary prepayment (as recorded in the performance

files) as our primary realized refinancing outcome. Throughout, we exclude terminations associated with default resolution or liquidation and focus on prepayment events consistent with borrower-initiated payoff.[5] Our baseline loan-level outcome is an indicator for a prepayment event occurring within a given time window (month or year, depending on the specification). Summary statistics for Freddie Mac loan-level variables are reported in Table 1.

To study cross-sectional patterns and to align refinancing outcomes with county-level collateral uncertainty, we also construct an aggregate refinancing rate by county and year. Let $\mathcal{L}_{c,t}$ denote the set of active loans in county c during year t. The county-year refinancing rate is

$$\text{RefiRate}_{c,t} = \frac{1}{|\mathcal{L}_{c,t}|} \sum_{i \in \mathcal{L}_{c,t}} \mathbb{1}\{\text{Prepay}_{i,t} = 1\}. \tag{1}$$

We assign Freddie Mac loans to counties using the property geography available in the dataset (mapped to county FIPS via a standard crosswalk). The crosswalk and geographic harmonization are described in Appendix A.

ZTRAX refinancing events. To verify that the Freddie Mac prepayment results reflect refinancing rather than sale-related mortgage payoff, we construct a complementary property-level outcome from the Zillow Transaction and Assessment Database (ZTRAX). ZTRAX combines public deed and mortgage records from county recorders with parcel characteristics from county assessors. We link cleaned arm's-length single-family purchase transactions to subsequent mortgage documents on the same parcel and organize the records into purchase spells. A spell enters the risk set 90 days after purchase and ends at the next recorded sale or the end of the available records. Our outcome equals one if, before that endpoint, the parcel receives either a mortgage explicitly coded as a refinance or a qualifying mortgage recording consistent with refinancing. For the latter, we require that the recording occur at least 180 days after purchase, have a loan amount of at least $50,000 and at least 25 percent of the purchase price, and not be coded as purchase-money, home-equity, second or subordinate, modification, reverse, future-advance, or open-ended credit. The main analysis uses purchase cohorts from 2000 through 2020. Because some qualifying mortgage records do not state loan purpose explicitly, we refer

[5] The exact mapping from performance status to voluntary prepayment is reported in Appendix A.

Table 1: *Summary statistics for Freddie Mac loan-level origination data (2000-2024)*

Panel A: Full GSE prepayment sample (originations 2000–2024)						
	N	*Mean*	*SD*	*P10*	*P50*	*P90*
Prepayment indicator	8,048,748	0.681	0.466	0.000	1.000	1.000
Collateral uncertainty	8,048,748	0.295	0.081	0.201	0.284	0.405
Collateral uncertainty (SFH)	8,047,457	0.304	0.083	0.207	0.294	0.416
FICO score at origination	8,048,748	745.452	48.252	677.000	755.000	800.000
Log original UPB	8,048,748	12.159	0.634	11.327	12.191	12.941
Original loan term (months)	8,048,748	343.792	50.644	360.000	360.000	360.000
Original interest rate (%)	8,048,748	5.146	1.443	3.250	5.000	7.000
Original LTV (%)	8,048,748	79.151	15.189	59.000	80.000	95.000
Debt-to-income ratio (%)	8,048,748	42.604	87.151	21.000	36.000	48.000
First-time homebuyer indicator	8,048,748	0.311	0.463	0.000	0.000	1.000

Note: The sample consists of fixed-rate, single-family purchase loans originated between 2000 and 2024 and acquired by Freddie Mac, drawn from the Single-Family Loan-Level Dataset (SFLLD). The following loans are excluded: loans with a prepayment penalty, loans originated under the Home Possible, HFA Advantage, and Refi Possible affordable housing programs, Relief Refinance loans (including HARP), interest-only mortgages, loans with FICO scores outside the range [300, 850], and loans without a Collateral Uncertainty value for their 3-digit ZIP code and origination year. Panel A reports statistics for the full sample. Collateral Uncertainty is the GAM-based hedonic residual dispersion measure estimated at the 5-digit ZIP-code-by-year level from ZTRAX records and aggregated to the 3-digit ZIP-code-by-year level as a transaction-weighted mean. Collateral Uncertainty (SFH) is the single-family-only variant used in IV specifications. The prepayment indicator equals one if the loan was paid off in full before its contractual maturity during its observed life in the Freddie Mac monthly performance files. Original loan term is the contractual term in months. All other variables are recorded at origination in the Freddie Mac origination files.

to the broad outcome as a mortgage recording consistent with refinancing. The sale censoring is nevertheless useful: a property sale cannot mechanically generate the measured refinancing event. Appendix B.1 provides the transaction cleaning and event-construction details.

2.2. Loan application and securitization

To measure mortgage application activity, underwriting outcomes, and loan sales/securitization, we use the Home Mortgage Disclosure Act (HMDA) data. HMDA covers the near-universe of U.S. mortgage applications and reports action taken (originated, approved-but-not-originated, denied, etc.) as well as whether an originated loan is sold within the calendar year of origination and to which purchaser type.[6]

We use HMDA to analyse both mortgage applications and secondary-market outcomes. We

[6]HMDA's "sold within year" flag captures sales reported by year-end and is therefore an informative (and potentially conservative) measure of secondary-market activity.

restrict the sample to observations with applicant income above $1,000 and loan amount above $10,000 (both are left-censored in HMDA; additionally, very small loans are unlikely to reflect standard mortgage financing). We exclude observations outside the 50 states and Washington, D.C. We focus on conventional (non-government-guaranteed) loans and single-family properties, which account for the large majority of mortgage applications and originations in the relevant markets. All additional HMDA cleaning rules are in Appendix A.

Table 2: *Summary statistics for HMDA loan-level data (2007-2024)*

HMDA — Loan-Level Sample (2007–2024)					
	Mean (1)	St. Dev. (2)	p10 (3)	p50 (4)	p90 (5)
Loan characteristics					
Loan amount (000s)	258.088	179.406	85.000	213.000	485.000
Applicant income (000s)	142.964	7.4e+04	41.000	96.000	230.000
Log loan amount	5.340	0.660	4.443	5.361	6.184
Log applicant income	4.572	0.705	3.714	4.564	5.438
Loan-to-income ratio	2.665	2.819	0.896	2.340	4.474
Refinancing					
Refinancing loan	0.590	0.492	0.000	1.000	1.000
Refinancing-cash-out loan	0.193	0.394	0.000	0.000	1.000
Among refinancing applications:					
Application denied	0.244	0.430	0.000	0.000	1.000
Denied: collateral (primary reason)	0.043	0.203	0.000	0.000	0.000
Denied: collateral (any reason)	0.049	0.216	0.000	0.000	0.000
Holding rates (originated loans)					
Retained on balance sheet	0.309	0.462	0.000	0.000	1.000
Observations	102,374,877				

Note: The sample consists of conventional, single-family mortgage applications for home purchase, home improvement, or refinancing from the Home Mortgage Disclosure Act (HMDA) database for 2007–2024, restricted to applications that were either originated or denied. The following applications are excluded: loan amounts below $50,000 or above the 99th percentile of the annual loan-amount distribution, those with applicant income below $1,000, and those whose census tract cannot be matched to a ZIP code with a Collateral Uncertainty measure from ZTRAX. Statistics for denial and collateral-denial rates are computed over refinancing applications only. Refinancing loans include cash-out refinancing. The cash-out refinancing indicator is available only for 2018 onward and is computed over all applications in those years. Denied: collateral (primary reason) equals one if the application was denied and collateral was listed as the primary denial reason, and Denied: collateral (any reason) equals one if collateral was listed among any of the reported denial reasons. Retained on balance sheet is defined only for originated loans and equals one if the loan was not sold within the calendar year or was sold to an affiliate institution.

We define a refinancing denial indicator at the application level if the application is denied, and zero if it is originated (or, in robustness exercises, if it is approved but not originated). We

also report county-level denial rates and lender-level heterogeneity in denial patterns. Summary statistics for HMDA outcomes and covariates are provided in Table 2. For originated loans, HMDA reports whether the loan is sold within the same calendar year and (when sold) the purchaser type. We define a binary indicator for whether an originated loan is sold within the year.

2.3. Collateral value uncertainty

We approximate collateral value uncertainty by measuring how difficult it is to value the property underlying a mortgage contract. Following Jiang and Zhang (2025), we proxy this uncertainty with idiosyncratic house price dispersion in transaction prices: when observably similar homes sell at more dispersed prices, collateral values are less predictable. Appendix **??** shows that our measure is strongly correlated with appraisal-based uncertainty, including the appraisal–sale price gap and the number and distance of comparable properties used in the appraisal process. Appendix B.3 further shows that areas with higher collateral uncertainty also exhibit longer time on market and lower sale-to-list price ratios. We interpret these patterns as evidence that our measure captures valuation difficulty in thinner and less liquid housing markets, where comparable transactions are less informative, rather than merely systematic differences in average house prices or housing quality. We construct this measure from ZTRAX transaction and assessment records in two steps: (i) estimate transaction-level pricing residuals from a hedonic model, and (ii) convert realized dispersion into an ex ante proxy by predicting expected squared residuals as a flexible function of observable property characteristics and time. The ZTRAX cleaning procedure is described in Appendix B.1.

Step 1: For each county separately, we estimate a hedonic pricing model on cleaned ZTRAX transactions:

$$\ln(p_{it}) \;=\; f_c(\ell_i, s_i, a_i, r_i, t) + u_{it}, \tag{2}$$

where p_{it} is the transaction price of property i at time t, ℓ_i is location, s_i is square footage, a_i is year built, and r_i contains bedrooms and bathrooms. We implement $f_c(\cdot)$ using a generalized

additive model (GAM) with year-month effects, a smooth location-by-time surface, and smooth time-varying functions of square footage, year built, bedrooms, and bathrooms. The residual u_{it} captures the deviation of the realized transaction price from the value predicted by observable property characteristics and local time variation; u_{it}^2 therefore measures transaction-level idiosyncratic dispersion. The exact specification is provided in Appendix B.1.

Step 2: The dispersion u_{it}^2 is observed ex post. To approximate ex ante uncertainty (i.e., what a lender could infer from observables and contemporaneous nearby transactions), we follow Jiang and Zhang (2025) and predict the expected squared residual using the same flexible set of property characteristics and time:

$$u_{it}^2 = g_c(\ell_i, s_i, a_i, r_i, t) + \epsilon_{it}, \tag{3}$$

$$\hat{\sigma}_{it} = \sqrt{\max\{\hat{g}_c(\ell_i, s_i, a_i, r_i, t), 0\}}, \tag{4}$$

where $g_c(\cdot)$ is estimated with a second county-specific GAM. We winsorize its fitted variance before taking the square root and then aggregate property-level predicted uncertainty to the ZIP-code-by-year level:

$$\sigma_{z,t} = \frac{1}{N_{z,t}} \sum_{i \in (z,t)} \hat{\sigma}_{it}, \tag{5}$$

where $N_{z,t}$ is the number of transactions in ZIP code z during year t. We interpret $\sigma_{z,t}$ as average predicted idiosyncratic price dispersion in ZIP code z and year t, i.e., collateral value uncertainty. The complete ZTRAX panel used in the analysis spans 2000–2024; we exclude the partial 2025 vintage.

Table 3 provides summary statistics for our baseline measure, *collateral uncertainty*, and the alternative one, *collateral uncertainty (SFH)*, estimated on single-family homes only. The latter is used in our instrumented variable analysis. The table shows that collateral uncertainty averages 0.354 across ZIP-code-years and has relatively high dispersion. In Appendix B.2 we show that our measure varies sharply across and within states, and is highly persistent across time and within ZIP codes.

Table 3: *Summary statistics for collateral uncertainty (2000-2024)*

Summary Statistics of House Collateral Uncertainty (ZIP Code Level)								
	Observations	Mean	St. Dev	p5	p10	p50	p90	p95
Collateral uncertainty	334,768	0.3540	0.1237	0.1646	0.1964	0.3467	0.5224	0.5679
Collateral uncertainty (SFH)	334,163	0.3627	0.1232	0.1709	0.2040	0.3579	0.5291	0.5736

Note: The sample consists of ZIP-code-by-year estimates of collateral uncertainty from Zillow Transaction and Assessment Dataset (ZTRAX) transactions for 2000–2024. Collateral Uncertainty is the average, across transactions in the ZIP code and year, of each property's predicted idiosyncratic price dispersion from a county-specific generalized additive model (GAM) hedonic price regression, predicted from property characteristics, location, and time. Collateral Uncertainty (SFH) is the single-family-only variant. Each measure is trimmed separately at the 0.8th and 99.5th percentiles of its 2000–2024 distribution across ZIP-code-years.

2.3.1. Instrumental-variables strategy and identification

The main identification concern is that local collateral uncertainty may co-move with borrower composition, credit demand, or other unobserved local factors that also affect refinancing outcomes. To isolate plausibly exogenous variation in collateral uncertainty, we instrument for σ_{zt} using residualized measures of local housing-stock atypicality, following Jiang and Zhang (2025). The core idea is that valuation becomes more difficult when the local housing stock is more heterogeneous: nonstandard properties have fewer close comparables, making transaction prices and appraisals noisier even for observationally similar borrowers.

We construct the instrument in three steps. First, for each transacted property, we compute the squared distance of key housing attributes – building age, square footage, bedrooms and geographic location – from the county-year mean. Second, we residualize these distances on the levels of the same housing characteristics and transaction-date fixed effects, so that the instrument captures relative nonstandardization rather than levels of housing quality. Third, we aggregate the residualized distances to the ZIP-year level. The resulting instrument therefore captures the extent to which the local housing stock is atypical along slow-moving dimensions of the built environment.[7] Consistent with this interpretation, collateral uncertainty is highly persistent across locations over time: ZIP codes with high uncertainty early in the sample tend to remain high-uncertainty later on, while low-uncertainty ZIP codes remain comparatively stable.[8] This pattern is consistent with collateral uncertainty reflecting durable local housing-market

[7] Full construction details and the 2SLS equations are provided in Appendix C.
[8] These results are show in Appendix B.2.

characteristics rather than purely transitory shocks.

The identifying assumption is that, conditional on borrower controls, lender fixed effects, and local-time fixed effects, the housing-stock-driven component of collateral uncertainty affects refinancing outcomes through valuation difficulty rather than through borrower credit risk or other omitted determinants of refinancing. While this assumption is not directly testable, Jiang and Zhang (2025) provide supporting evidence for the same instrument in the mortgage-lending context: they show that it is not systematically related to observable ex-ante borrower characteristics and does not predict ex-post mortgage delinquency. Given this validation evidence, together with the first-stage strength reported in our main tables, we interpret the 2SLS estimates as supportive of a causal interpretation, while recognizing that analogous balance and placebo tests in our own setting would further strengthen this claim.

3. Collateral uncertainty, prepayment risk and refinancing frictions

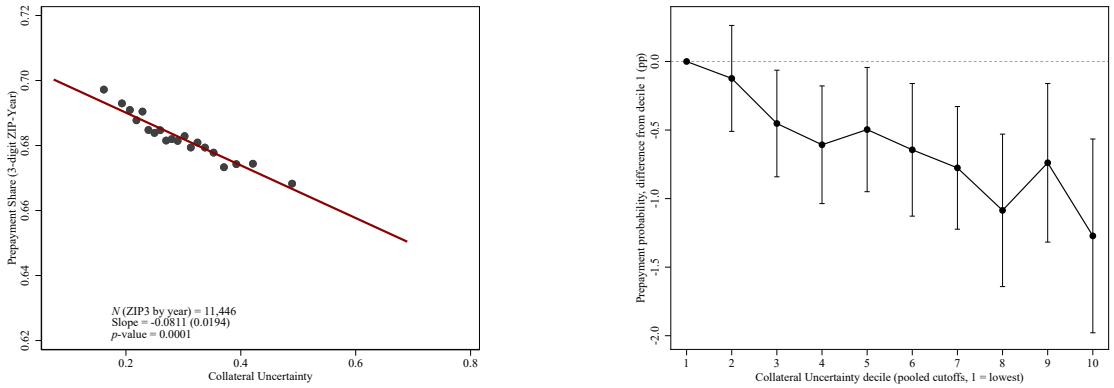
3.1. Collateral uncertainty reduces mortgage prepayment

We begin by testing whether collateral value uncertainty reduces realized refinancing. Since the Freddie Mac performance data do not directly identify refinancing, we use voluntary prepayment as our proxy. This measure is imperfect, as some prepayments reflect mobility or other non-refinancing exits. Still, in the agency fixed-rate mortgage market, voluntary prepayment is the main observed margin through which refinancing occurs. We therefore interpret the results in this section as evidence on realized refinancing behavior. We next validate this interpretation using first linked ZTRAX property and mortgage records. Second, we examine subsequent prepayment among loans with positive estimated net gains from refinancing. We then turn to HMDA refinancing denials to examine the underwriting channel more directly.

Panel (a) of Figure 1 plots a binned scatter of the prepayment share at the 3-digit ZIP code $\times$ origination year level against collateral uncertainty. Each point corresponds to one of 20 equally spaced bins of collateral uncertainty after partialling out commuting zone $\times$ year fixed effects. The relationship is strongly negative: ZIP codes with greater collateral uncertainty exhibit systematically lower prepayment rates. This pattern is economically meaningful and

Figure 1: *Collateral uncertainty and mortage prepayment (2000-2024)*

(a) *ZIP code mortgage prepayment Rate* **(b)** *Loan-level prepayment likelihood*



Note: Panel (a) plots a binned scatter plot of mortgage prepayment rates against collateral uncertainty at the three-digit ZIP code×origination-year level. The estimates control for commuting zone×origination-year fixed effects. The solid red line shows the corresponding weighted linear fit. The reported slope, standard error (in parentheses) and p-value are obtained from the weighted regression with standard errors clustered by commuting zone. Panel (b) adjusted prepayment-probability differences in percentage points relative to decile 1, normalized to zero. Bars show 95% confidence intervals based on standard errors clustered at commuting zones. The estimates are based on the decile equivalent specification of Equation 6, using the same controls and fixed effects. Both panels use fixed-rate purchase mortgages on single-family homes originated between 2000 and 2024 in Freddie Mac's Single-Family Loan-Level Dataset (SFLLD). Collateral uncertainty, denoted $\sigma_{z3,t}$, is constructed from generalized additive hedonic model (GAM) dispersion measures using ZTRAX transactions and aggregated from ZIP5-year to ZIP3-year using single-family transaction counts as weights.

already suggests that borrowers in harder-to-value housing markets refinance less, even before conditioning on detailed loan characteristics. Panel (b) of Figure 1 shows the corresponding relationship using loan-level regression estimates across collateral-uncertainty deciles. Relative to the lowest-uncertainty decile, controlling for origination characteristics and absorbing commuting-zone-by-origination-quarter and seller-by-origination-quarter fixed effects. Loans in higher-uncertainty areas are less likely to prepay than those in the lowest-uncertainty decile.

We then replace the decile indicators with a continuous measure of collateral uncertainty and estimate the following loan-level specification:

$$\text{Prepay}_{ibzt} = \beta\sigma_{zt} + X_i'\Gamma + \alpha_{c(z)\times t} + \alpha_{b\times y(t)} + \varepsilon_{ibzt}, \tag{6}$$

where Prepay_{ibzt} is an indicator equal to one if loan i is prepaid early, σ_{zt} is collateral uncertainty in 3-digit ZIP code z and period t, X_i is a vector of origination-loan characteristics, $\alpha_{c(z)\times t}$ are commuting zone $\times$ year fixed effects, and $\alpha_{b\times y(t)}$ are lender $\times$ year fixed effects. Standard errors

are clustered at the commuting-zone level to allow for spatial correlation in mortgage outcomes within local housing markets.

The vector X_i includes the borrower's FICO credit score and its squared term, the log of the original unpaid principal balance, the log of the loan maturity, and the original contract interest rate. Loan-to-value and Debt-to-income ratios enter as categorical fixed effects, with LTV binned into five groups ($\leq$50%, 50–80%, 80–85%, 85–97%, $>$97%) and DTI into four groups ($\leq$36%, 36–45%, 45–50%, $>$50%). Additional loan-level fixed effects absorb occupancy status, origination channel (retail, broker, or correspondent), conforming versus super-conforming loan status, mortgage insurance coverage tier, the number of property units (one to four), first-time homebuyer status, and a multiple-borrower indicator[9]. These controls absorb variation in refinancing incentives and borrower risk at origination, while the fixed effects absorb broad local credit-market conditions and lender-specific shocks. In particular, the commuting zone $\times$ year fixed effects ensure that identification does not come from aggregate differences in local housing or macroeconomic conditions across regions at a point in time. Once the lender $\times$ year fixed effects are added, identification comes from comparing observationally similar loans originated by the same lender in the same year and commuting zone across 3-digit ZIP codes with different levels of collateral uncertainty.

Table 4 reports the results. Columns (1) and (2) show that the OLS estimate of β is negative and statistically significant across specifications. The point estimates are very stable: adding lender $\times$ year fixed effects barely changes the estimate, suggesting that the result is not driven by time-varying differences in lender business models, pricing, or refinancing strategies.

Columns (3) and (4) turn to instrumental-variables specifications in which collateral uncertainty is instrumented with residualized measures of local housing-stock atypicality. The IV coefficients remain negative and are larger in magnitude, which is consistent with measurement error in the collateral uncertainty proxy. However, the first-stage F-statistics are approximately 6.7, indicating limited first-stage strength. Aggregation from five-digit to three-digit ZIP codes removes within-ZIP3 variation in housing-stock characteristics and may contribute to the limited first-stage strength. We therefore interpret these estimates cautiously: their direction accords with the OLS results. The economic magnitudes are large; moving from the bottom to the top

[9]We use the same set of controls and fixed effects in our flexible decile comparison specification.

Table 4: *Collateral Uncertainty and mortgage prepayment (2000-2021)*

Dep. var: Loan Prepaid Indicator (= 1 if loan prepaid early)				
	OLS		IV	
	(1) Base FE	(2) + Seller FE	(3) Base FE	(4) + Seller FE
Collateral Uncertainty	-0.0436***	-0.0424***	-0.0818**	-0.0805**
	(0.0114)	(0.0113)	(0.0364)	(0.0364)
Loan Controls	✓	✓	✓	✓
Borrower and Property Controls	✓	✓	✓	✓
CZ-Year Qtr FE	✓	✓	✓	✓
Seller-Year FE		✓		✓
Clustered SE	Commuting Zone	Commuting Zone	Commuting Zone	Commuting Zone
First-Stage F			6.66	6.67
# of CZ	317	317	317	317
# of ZIP3	680	680	680	680
Observations	8,048,722	8,048,722	8,047,431	8,047,431

Note: This table reports loan-level estimates of the probability that a loan is prepaid early for the period 2000-2024. The sample consists of fixed-rate, single-family purchase loans acquired by Freddie Mac and originated between 2000 and 2024, drawn from the Single-Family Loan-Level Dataset (SFLLD). Loans with a prepayment penalty, under special or affordable housing programs or HARP, interest-only loans, and loans with FICO scores outside [300, 850] are excluded. The outcome equals one if the loan is ever prepaid and zero otherwise; payoffs at maturity are coded as zero. Collateral Uncertainty is measured from ZTRAX at the 3-digit ZIP code and origination year, as the average of 5-digit ZIP code values weighted by the number of single-family transactions. Columns 1 and 2 report OLS estimates; columns 3 and 4 report IV (2SLS) estimates in which Collateral Uncertainty SFH is instrumented with residualized ZTRAX property characteristics (location, log area, year built, bathrooms) at the 3-digit ZIP-year level. All columns include loan controls, loan-characteristic fixed effects, and commuting zone by origination year-quarter fixed effects; columns 2 and 4 add seller by year-quarter fixed effects. Standard errors are clustered by commuting zone. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

decile of collateral uncertainty is related to a lower prepayment probability of 2.4 percentage points, or 3.8 per cent relative to the sample mean.

These results are difficult to reconcile with a purely demand-side interpretation. After comparing loans facing the same lender-year conditions and the same commuting-zone-by-time conditions, borrowers in high-uncertainty locations are still less likely to prepay. This is consistent with a lender-side refinancing friction: when collateral is harder to value, appraisal-based underwriting becomes tighter, and refinancing becomes less likely to be completed. We further confirm in Appendix D.1 that this result is not sensitive to the loan-type sample restriction: re-estimating the same specifications on the unrestricted Fannie Mae and Freddie Mac origination sample yields nearly identical coefficients (Table D.1).

One caveat is that voluntary prepayment is an indirect proxy for refinancing, not a direct observation of refinance originations. The next subsection addresses this limitation using

property-level deed and mortgage records that separate subsequent mortgage recordings from property sales.

3.2. Property-level evidence

The Freddie Mac result could partly reflect geographic differences in mobility: a sale pays off the outstanding mortgage and therefore appears as a prepayment even when the borrower does not refinance. We address this concern using ZTRAX to link each cleaned single-family purchase to later mortgage documents on the same parcel. We follow the property from 90 days after purchase until its next recorded sale or the end of the available records and define an event from the first qualifying mortgage recording within that spell. Thus, unlike the Freddie Mac outcome, a sale cannot itself generate the measured refinancing event. As described in Section 2, the broad measure combines records explicitly coded as refinances with mortgage recordings that satisfy conservative timing, amount, and loan-type restrictions.

We estimate linear probability models on 24,983,923 purchase spells from the 2000–2020 cohorts. Columns (1) and (2) of Table 5 use state and purchase-quarter fixed effects and measure collateral uncertainty at the purchase ZIP or property, respectively. Column (3), our preferred ZTRAX specification, uses property-level uncertainty and ZIP-by-purchase-quarter fixed effects. It therefore compares properties purchased in the same ZIP code and quarter while controlling for purchase price, square footage, building age, property type, bathrooms, occupancy, land use, location, and the length of the observable spell. Standard errors are clustered by ZIP code.

The estimates are negative and precisely estimated across all three specifications. At the ZIP level, a 0.10 increase in collateral uncertainty is associated with a 1.78 percentage-point reduction in the probability of a qualifying refinancing event. In the preferred within-ZIP-by-quarter specification, the coefficient is -0.0519 (standard error 0.0022), implying that a 0.10 increase in property-level uncertainty is associated with a 0.52 percentage-point lower refinancing probability, or about 1.1 percent of the 49.1 percent sample mean. The within-ZIP estimate is smaller in magnitude and is identified from differences in predicted valuation uncertainty across properties in the same place and purchase quarter, after removing common local-market and cohort variation.

Table 5: *Collateral uncertainty and mortgage recordings consistent with refinancing*

Dep. var: Mortgage Recordings Refinancing(= 1 if refinanced)			
	(1) ZIP CU	(2) Property CU	(3) Property CU
ZIP collateral uncertainty at purchase	-0.178^{***}		
	(0.00763)		
Property collateral uncertainty at purchase		-0.101^{***}	-0.0519^{***}
		(0.00337)	(0.00221)
Property controls	Yes	Yes	Yes
Exposure-duration bins	Yes	Yes	Yes
State fixed effects	Yes	Yes	No
Origination-quarter fixed effects	Yes	Yes	No
ZIP $\times$ origination-quarter fixed effects	No	No	Yes
Clustered standard errors	ZIP	ZIP	ZIP
Observations	24,983,923	24,983,923	24,856,503
Adjusted R^2	0.180	0.179	0.215

Note: This table reports linear probability models for a pooled ZTRAX sample of single-family purchase spells in 46 jurisdictions. The outcome equals one if an explicitly refinance-coded mortgage or a qualifying mortgage recording consistent with refinancing occurs before the next recorded sale or administrative censoring. For records without an explicit refinance code, qualifying events must occur at least 180 days after purchase, have a loan amount of at least $50,000 and at least 25 percent of the purchase price, and exclude purchase-money, home-equity, second or subordinate, modification, reverse, future-advance, and open-ended records. The outcome mean is 0.491. Property controls include log purchase price, log square footage, building age, attached-property status, bathrooms, occupancy, land use, latitude, longitude, and their interaction. All columns include exposure-duration bins. Origination quarter refers to the quarter of the purchase transaction. Standard errors, in parentheses, are clustered at the ZIP-code level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

The ZTRAX magnitude is comparable to the GSE result once the estimates are placed on a common scale. The within-lender GSE OLS coefficient in Table 4 implies that a 0.10 increase in collateral uncertainty is associated with 0.77 percentage points less prepayment, or 1.2 percent of the 63.1 percent GSE mean. The preferred within-ZIP ZTRAX estimate implies a somewhat smaller 0.52 percentage-point decline, but this is 1.1 percent of the 49.1 percent ZTRAX refinancing-event mean. Thus, although the absolute ZTRAX estimate is about one-third smaller, the two mean-scaled effects are nearly identical. The 2.4 percentage-point GSE magnitude reported above is larger because it combines the IV estimate with a bottom-to-top-decile change in uncertainty; it is therefore not directly comparable to the 0.10-unit OLS calculation used here. The ZIP-level ZTRAX estimate, which uses broader geographic variation, implies a larger 1.78 percentage-point effect for the same 0.10-unit change. Overall,

the estimates differ with the outcome and identifying variation but are of the same economic order of magnitude.

These results independently confirm that collateral uncertainty is associated with less realized refinancing outside the GSE performance data and are not driven by treating sale-related mortgage payoff as refinancing. The ZTRAX measure is still imperfect because loan-purpose coding is incomplete and some events are inferred from qualifying mortgage records. We therefore treat it as complementary validation rather than a replacement for the Freddie Mac analysis.

3.3. Refinancing opportunities and subsequent prepayment

To further assess whether the prepayment relationship reflects refinancing opportunities rather than geographic differences in mobility, we focus on outstanding mortgages for which refinancing offers positive modelled net savings. These borrowers have a financial incentive to refinance even without selling their property. This exercise complements the sale-censored property-level evidence by examining whether the prepayment gap persists among loans with a positive refinancing opportunity.

We construct annual January snapshots of active mortgages in the Freddie Mac data. For each loan, we calculate the payment reduction from refinancing the outstanding balance at the January Primary Mortgage Market Survey (PMMS) benchmark rate, holding remaining maturity fixed. We discount these payment reductions over the remaining contractual term and subtract refinancing costs equal to 2% of the outstanding balance (Keys et al., 2016). The estimation sample retains loans with positive net present discounted savings. Our baseline outcome equals one if the loan prepays in a reporting month after the January snapshot and within the following 36 and 60 months, and zero otherwise. The resulting sample covers January snapshots from 2001 through 2020. The unit of observation is a loan–January-snapshot pair, so an outstanding loan can contribute observations in more than one year.

We estimate a linear probability model similar to Equation 6 relating subsequent prepayment to collateral uncertainty in the property's three-digit ZIP code and snapshot year.[10] The baseline

[10]We construct a panel of outstanding mortgages observed each January. Each January observation provides a snapshot of the loan's outstanding balance, interest rate, and remaining contractual maturity, which we use to evaluate its refinancing opportunity. We match each loan observed in January of year t to collateral uncertainty in its three-digit ZIP code for calendar year t. We refer to this measure as snapshot-year collateral uncertainty,

specification includes origination loan controls, loan-characteristic fixed effects, commuting-zone-by-snapshot-year fixed effects, and origination-seller-by-snapshot-year fixed effects. We cluster standard errors at the commuting-zone level. We then add controls for the size of the refinancing opportunity, measured by bins of net refinancing savings relative to the outstanding balance. A further specification replaces the commuting-zone-by-snapshot-year fixed effects with commuting-zone-by-snapshot-year-by-origination-year fixed effects, comparing loans from the same origination vintage within the same local market and snapshot year.

Table 6 reports the results. Collateral uncertainty remains negatively associated with subsequent prepayment among loans with positive modelled refinancing gains. Focusing on the prepaid probability within the following 60 months, the baseline coefficient is -0.1163, implying that a 0.10 increase in collateral uncertainty (roughly equal to the observed standard deviation) is associated with a 1.16-percentage-point reduction in the probability of prepayment over the next 60 months, relative to a sample mean of 66.2%. The estimate changes little when we control for incentive size (-0.1166) or absorb origination-cohort differences within local market–snapshot-year cells (-0.1168). Thus, the absence of modelled refinancing gains at the snapshot, or the measured size of those gains, does not explain the relationship.

Figure 2 examines the relationship across collateral-uncertainty deciles at horizons of 36 and 60 months. The plotted coefficients measure adjusted differences in prepayment probabilities relative to the lowest-uncertainty decile. Using snapshot-year uncertainty, the highest-decile lower likelihood is -2.99 percentage points at 36 months and -3.30 percentage points at 60 months.

Taken together, these results show that higher collateral uncertainty is associated with lower subsequent prepayment even among loans with positive modelled net refinancing gains at the January snapshot. This exercise measures an opportunity at the starting date, as interest rates and borrower circumstances can change, and an individual borrower may not want to refinance at all. Moreover, prepayment remains a proxy for refinancing.

Having established the realized refinancing gap in two distinct data sources, we next turn to applications and denials to determine whether lender approval decisions contribute to that gap. This exercise is important as it allows us to distinguish clearly between borrowers' and lenders'

distinguishing it from uncertainty measured in the loan's origination year.

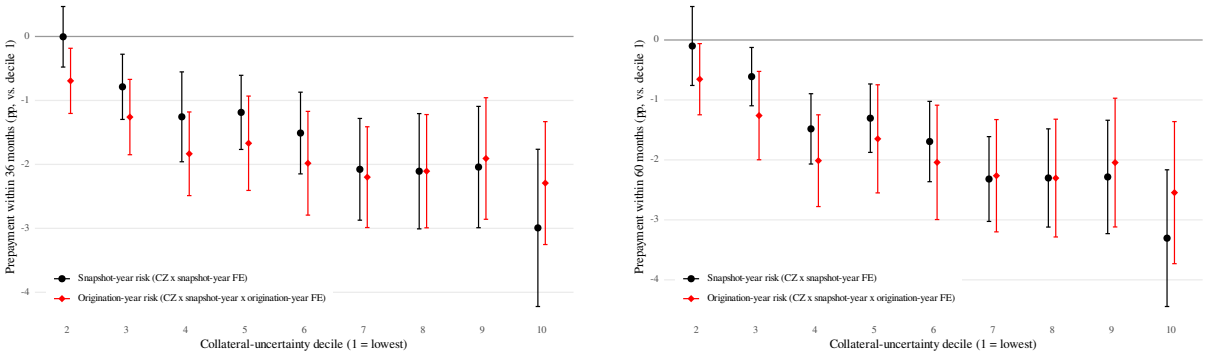
Table 6: *Collateral uncertainty and prepayment likelihood for loans with positive savings*

	Prepaid within 36 months			Prepaid within 60 months		
	(1) Main	(2) Incentive-Size	(3) + Cohort FE	(4) Main	(5) Incentive-Size	(6) + Cohort FE
Collateral Uncertainty	-0.1038***	-0.1044***	-0.1040***	-0.1163***	-0.1166***	-0.1168***
	(0.0213)	(0.0213)	(0.0201)	(0.0206)	(0.0206)	(0.0192)
Loan Controls	✓	✓	✓	✓	✓	✓
Loan Characteristic FEs	✓	✓	✓	✓	✓	✓
CZ-Snapshot Year FE	✓	✓		✓	✓	
CZ-Snapshot Year-Origination Year FE			✓			✓
Seller-Snapshot Year FE	✓	✓	✓	✓	✓	✓
Incentive-Size Controls		✓			✓	
# of CZs	315	315	315	313	313	313
# of ZIP codes	669	669	669	664	664	664
Clustered SE	Commuting Zone	Commuting Zone	Commuting Zone	Commuting Zone	Commuting Zone	Commuting Zone
Mean of outcome	0.4987	0.4987	0.4987	0.6620	0.6620	0.6620
Observations	16,011,827	16,011,827	16,011,064	12,878,791	12,878,791	12,878,203

Note: The sample consists of loan-year observations of fixed-rate, single-family Freddie Mac purchase loans from the Single-Family Loan-Level Dataset (SFLLD), observed each January between 2001 and 2022 (columns 1 to 3) or 2020 (columns 4 to 6) while current and in the money, that is, when refinancing at the Primary Mortgage Market Survey rate yields a positive present value net of a 2% cost. The dependent variable equals one if the loan is prepaid within 36 or 60 months of the snapshot. Collateral Uncertainty is the GAM-based hedonic residual dispersion measure of the loan's 3-digit ZIP code in the snapshot year, computed from ZTRAX records. All columns are OLS with loan controls, loan characteristic fixed effects, and seller by snapshot year and commuting zone by snapshot year fixed effects. Columns 2 and 5 add bins of the refinancing incentive; columns 3 and 6 interact the commuting zone fixed effects with origination year. Standard errors are clustered at the commuting zone level.

Figure 2: *Prepayment of in-the-money mortgages by collateral uncertainty decile (2000-2024)*

(a) *Within 36 months of the January snapshot* **(b)** *Within 60 months of the January snapshot*



Note: This figure plots OLS coefficients, in percentage points, on deciles 2 to 10 of Collateral Uncertainty relative to decile 1. The outcome indicates that the loan is prepaid within 36 months (Figure 1a) or 60 months (Figure 2b) after the January snapshot; the decile-1 prepayment rate is 55% and 72%. The sample is a stacked panel of January snapshots of current fixed-rate, single-family purchase loans from Freddie Mac's Single-Family Loan-Level Dataset that are in the money net of refinancing costs. Collateral Uncertainty is measured at the ZIP3 level from ZTRAX transactions. Black circles use the snapshot-year decile with commuting zone (CZ) by snapshot-year fixed effects; red diamonds use the origination-year decile with CZ by snapshot-year by origination-year fixed effects. Both include borrower and loan controls and seller by snapshot-year fixed effects. Bars are 95% confidence intervals clustered by CZ.

incentives.

3.4. Collateral uncertainty raises refinancing denials

We next use HMDA refinancing applications to examine whether lender approval decisions contribute to the refinancing gap. Focusing on applications explicitly submitted for refinancing removes the ambiguity in the prepayment data: these borrowers are seeking to refinance an existing mortgage, and the measured outcome cannot be generated by a sale-related mortgage payoff. Moreover, observing applications and their disposition separates the borrower's decision to seek refinancing from the lender's decision to approve it. Applicants have already expressed demand for refinancing; denials reveal whether lenders are willing to provide it.

We therefore ask whether applications backed by more uncertain collateral are more likely to be denied, conditional on borrower and loan characteristics and common lender-year and county-year conditions. We then examine the reasons lenders report for these decisions. Conditional on denial, we ask whether collateral is more likely to be cited as the primary reason in high-uncertainty locations. We find both higher denial rates and a greater prevalence of collateral-related reasons among denied applications. Together, these findings support a collateral-based underwriting channel through which valuation uncertainty limits refinancing access.

Our baseline specification is:

$$\text{Denied}_{ibzt} = \beta\sigma_{zt} + X_i'\Gamma + \alpha_{c(z)\times y(t)} + \alpha_{b\times y(t)} + \varepsilon_{ibzt}, \tag{7}$$

where Denied_{ibzt} is an indicator equal to one if the refinance application i submitted to lender b in ZIP code z and year t is denied, σ_{zt} is collateral uncertainty in the applicant's ZIP code, X_i is a vector of application-level controls, $\alpha_{c(z)\times y(t)}$ are county $\times$ year fixed effects, and $\alpha_{b\times y(t)}$ are lender $\times$ year fixed effects. Standard errors are clustered at the county level to allow for spatial correlation in underwriting outcomes within local mortgage markets.

The vector X_i includes the standard HMDA controls used throughout the paper: log applicant income, log loan amount, the loan-to-income ratio and its square, and fixed effects for applicant sex, ethnicity, race, owner-occupancy status, property type, loan type, and loan purpose. These controls absorb observable differences in borrower risk and application composition. The county $\times$ year fixed effects absorb local housing-market conditions, local labor-market conditions, any

county-level market condition such as lender concentration, and county-level credit-supply shifts that could drive the main results. Once the lender × year fixed effects are added, identification comes from comparing refinance applications handled by the same lender in the same year across ZIP codes within the same county that differ in collateral uncertainty. This is the relevant variation for the paper's mechanism: holding fixed lender-level underwriting policy and broad county-year conditions, does the same lender reject more refinance applications in locations where collateral is harder to value?

Table 7: *Collateral uncertainty and likelihood of mortgage refinancing rejections (2007-2024)*

	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
	Across Lenders	Within Lender	Within Lender	Across Lenders	Within Lender	Within Lender
Collateral Uncertainty	0.1170***	0.1055***	0.1124***	0.1950***	0.1820***	0.1906***
	(0.0137)	(0.0072)	(0.0073)	(0.0357)	(0.0267)	(0.0259)
CLL Sample			✓			✓
Loan Controls	✓	✓	✓	✓	✓	✓
County-Year FE	✓	✓	✓	✓	✓	✓
Institution-Year FE		✓	✓		✓	✓
# of Counties	1626	1626	1626	1629	1629	1629
# of ZIP codes	12,885	12,885	12,884	12,894	12,894	12,893
First-stage F				32.23	34.88	37.33
Clustered SE	County	County	County	County	County	County
Observations	55,419,738	55,417,364	53,132,600	55,394,105	55,391,733	53,107,473

Note: The sample consists of mortgage refinancing applications from the Home Mortgage Disclosure Act (HMDA) database between 2007 and 2024, originated or denied by lenders matched to an RSSD identifier. The outcome equals one for denied applications. Collateral Uncertainty is computed at the ZIP-code-by-year level from ZTRAX transaction and assessment records. Columns 1 to 3 report OLS estimates. Columns 4 to 6 report 2SLS estimates for Collateral Uncertainty (SFH), instrumented with residualized ZTRAX property characteristics; the first-stage F is the Kleibergen-Paap statistic. All columns include county×year fixed effects and borrower and loan controls; columns 2, 3, 5, and 6 add institution×year fixed effects, and columns 3 and 6 keep loans at or below the conforming loan limit. Standard errors are clustered at the county level. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

Table 7 reports the results. The OLS estimates in columns (1)–(3) are positive and highly significant. Once the lender × year fixed effects are included, the magnitude of the coefficient increases, and it remains essentially unchanged when the sample is restricted to conforming-loan-limit (CLL) applications only. The stability of the result across columns (2) and (3) is important. It indicates that the denial effect is not confined to applications outside the conforming segment, where balance-sheet lending or nonstandard underwriting might be more prominent. Instead, the relationship extends to the part of the refinance market where underwriting is otherwise relatively standardized.

Columns (4)–(6) report instrumental-variables estimates, where collateral uncertainty is instrumented using residualized measures of local housing-stock atypicality. The IV coefficients are again positive and statistically significant. The first stage is strong in this setting, with first-stage F-statistics between 32.33 and 37.33. These estimates imply that collateral uncertainty has a quantitatively large effect on refinance denials. Expressed in standardized terms, a one-standard-deviation increase in collateral uncertainty raises the probability of denial by 2.34 percentage points, or 9.62 percent relative to the sample mean denial rate.

A potential concern with the baseline is that smaller lenders tend to operate in geographically concentrated markets, which limits within-lender variation in collateral uncertainty and could make identification less reliable. In Table D.3 we report estimates of Equation 7 restricted to the 30 banks subject to Dodd-Frank Act stress testing (DFAST) for the 2018-2021 period. This group of large banks operates at a national scale and serves a broad range of local housing markets, so within-lender variation in collateral uncertainty is not mechanically constrained by geographic specialization. The estimates in Table D.3 are stable in sign and comparable or larger in magnitude relative to the baseline, which is inconsistent with the concern that our results are driven by lenders' geographic self-selection rather than collateral uncertainty.

Denials are observed at the application stage, so they directly capture lender approval decisions rather than borrower behavior. The positive coefficient in Table 7 therefore, points to a supply-side underwriting channel: when collateral is harder to value, lenders are less willing to approve refinance applications, even after conditioning on a rich set of borrower and loan characteristics and comparing applications within the same lender-year and county-year cells. This is difficult to reconcile with a purely borrower-side story based on inattention or inertia.

To probe the mechanism more directly, we next examine denial reasons. If collateral value uncertainty matters because it makes appraisal-based LTV constraints more likely to bind, then denied applications in high-uncertainty locations should be more likely to list collateral as the primary reason for rejection. We therefore estimate

$$\text{CollateralDenial}_{ibzt} = \delta\sigma_{zt} + X_i'\Pi + \alpha_{c(z)\times y(t)} + \alpha_{b\times y(t)} + u_{ibzt}, \tag{8}$$

on the sample of denied refinance applications only, where $\text{CollateralDenial}_{ibzt}$ equals one if

collateral is reported as the primary denial reason.

Table 8 reports the estimates. Across all specifications the coefficient on collateral uncertainty is positive and highly significant. This pattern is central for interpretation. Collateral uncertainty does not merely predict a higher overall denial rate; it predicts a reallocation of denial reasons toward collateral itself. That is precisely what one would expect if noisier valuations increase the probability that appraisal-based LTV thresholds bind at refinancing.

Taken together, Tables 7 and 8 provide direct evidence that collateral uncertainty operates through lender underwriting rather than only through borrower demand. The same locations in which realized prepayment is lower are also locations in which refinance applications are more likely to be denied, and conditional on denial, more likely to be denied for collateral-related reasons. This combination of facts is difficult to generate with a pure inattention or inertia mechanism. Instead, it is consistent with the paper's central claim: collateral value uncertainty creates a location-specific refinancing friction by making appraisal-based underwriting tighter and less predictable.

The next subsection studies this mechanism more directly by examining why collateral uncertainty should matter particularly through the LTV margin, and by providing additional evidence that the effect is strongest where valuation noise is most likely to push borrowers across underwriting cutoffs.

3.5. Mechanism

At refinancing, lenders do not observe the market value of the property directly and therefore have to estimate it. In practice, this estimate is based on the appraisal, which then enters the lender's assessment of whether the requested refinance loan can be approved. U.S. mortgage underwriting places substantial weight on verified repayment capacity, including income, employment, debt obligations, and credit history (Consumer Financial Protection Bureau, 2013). But in the refinancing context the lender also observes realized payment performance on the existing mortgage, so approval often depends importantly on whether the borrower has sufficient collateral to satisfy appraisal-based LTV requirements. This was especially evident during the HARP period, when many otherwise performing borrowers were unable to refinance through standard

Table 8: *Collateral uncertainty and collateral-based refinancing rejections (2007-2024)*

Dep. var: Collateral Denial Indicator (= 1 if application denied due to collateral)						
	OLS			IV		
	(1) Across Lenders	(2) Within Lender	(3) Within Lender	(4) Across Lenders	(5) Within Lender	(6) Within Lender
Collateral Uncertainty	0.2579***	0.1950***	0.2012***	0.1272***	0.1281***	0.1473***
	(0.0100)	(0.0084)	(0.0083)	(0.0200)	(0.0190)	(0.0178)
CLL Sample			✓			✓
Loan Controls	✓	✓	✓	✓	✓	✓
County-Year FE	✓	✓	✓	✓	✓	✓
Lender-Year FE		✓	✓		✓	✓
# of Counties	1269	1269	1269	1270	1270	1270
# of ZIP codes	12,778	12,778	12,776	12,779	12,779	12,777
First-stage F				37.15	37.09	40.56
Clustered SE	County	County	County	County	County	County
Observations	13,416,159	13,415,433	12,769,311	13,404,656	13,403,929	12,758,816

Note: The sample consists of home mortgage refinancing applications denied between 2007 and 2024 by lenders matched to an RSSD identifier, drawn from the Home Mortgage Disclosure Act (HMDA) database. The dependent variable equals one if collateral is the primary reported denial reason. Collateral Uncertainty is computed at the ZIP-code-by-year level from ZTRAX transaction and assessment records. Columns 1 to 3 report OLS estimates; columns 4 to 6 report 2SLS estimates using Collateral Uncertainty SFH, instrumented with residualized ZTRAX property characteristics, and the Kleibergen-Paap first-stage F. All columns include county×year fixed effects and borrower and loan controls and fixed effects; columns 2, 3, 5, and 6 add institution×year fixed effects, and columns 3 and 6 restrict the sample to loans at or below the conforming loan limit. Standard errors are clustered at the county level. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

channels because declines in house prices had pushed their LTV ratios above conventional eligibility limits. Policies such as HARP were designed precisely to relax these equity-related barriers for performing borrowers (Abel and Fuster, 2021). As a result, uncertainty about collateral values can be especially important for refinance approval decisions.

Collateral uncertainty can affect the lender's refinancing decision through two conceptually distinct channels. The first is the *appraisal channel*. When collateral values are harder to assess, appraisals become noisier, and a sufficiently low appraisal can push the measured refinance LTV above the relevant underwriting threshold, leading to denial or a cash-in requirement. The second is the *collateral recovery channel*. Even holding fixed the appraisal used at origination, higher collateral uncertainty reduces the lender's expected recovery in default, making refinancing less attractive ex ante. In the spirit of Jiang and Zhang (2025), the first channel works through appraisal-based underwriting constraints, while the second works through the value of the collateral in downside states. In this section we provide evidence on both channels.

We begin with the appraisal channel. To examine whether higher collateral uncertainty is

associated with noisier and less informative appraisals, we use the FHFA Enterprise Uniform Appraisal Dataset (UAD) Appraisal-Level Public Use File, which contains appraisal-level information on appraised values and the comparable properties used to support purchase and refinance appraisals.[11] Figure 3 presents four binscatter plots, each partialling out year fixed effects. The first panel shows that the absolute appraisal gap is larger in high-uncertainty locations, indicating that appraisals deviate more from contract prices where collateral values are harder to assess. The second and third panels show that refinance appraisals in these same locations rely on fewer comparable properties and on comparables that are farther away. The fourth panel shows that, conditional on under-appraisal, the severity of the under-appraisal is also greater in high-uncertainty areas. Taken together, these facts indicate that high-collateral-uncertainty markets have thinner and less informative comparable sets and exhibit more severe downside appraisal realizations. This is the pattern implied by the appraisal channel: when collateral is harder to value, the appraisal used to underwrite the refinance becomes less reliable, making the measured refinance LTV more uncertain and increasing the likelihood of denial.

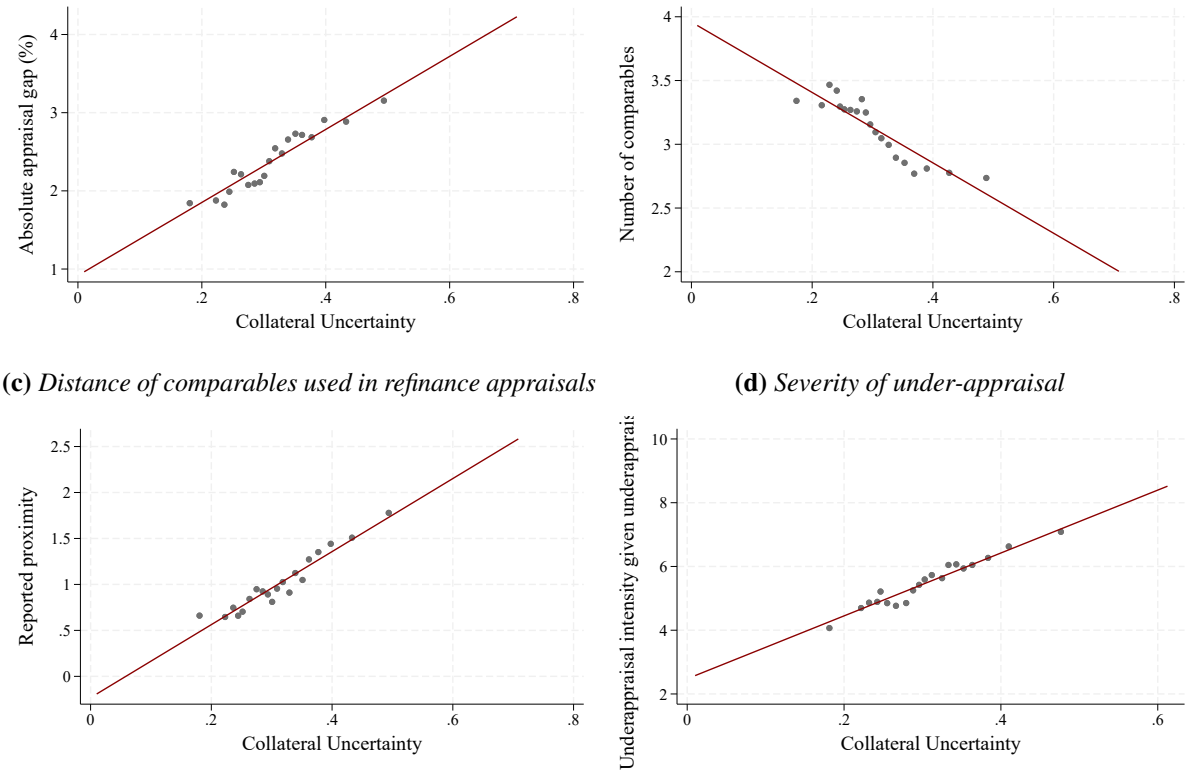
A further implication of the appraisal channel is that the effect of collateral uncertainty on collateral-related refinance denial should be stronger for cash-out refinances than for non-cash-out refinances. For a given property value and outstanding mortgage balance, cash-out refinancing increases the amount borrowed relative to rate-and-term refinancing and therefore leaves less margin below the relevant appraisal-based LTV cutoff. As a result, when collateral values are harder to assess and appraisals are noisier, a downward appraisal realization is more likely to push the measured refinance LTV above the underwriting threshold for cash-out loans. We test this prediction on the sample of 2018-2024 HMDA loan-level applications, for which we can distinguish cash-out from standard rate-and-term refinancing. Consistent with the outlined mechanism, Table 9 shows that the effect of collateral uncertainty on collateral-related denial is substantially larger for cash-out refinances than for non-cash-out refinances. This pattern provides additional support for the appraisal-based underwriting channel, as it is precisely in the applications with the least LTV slack that appraisal noise should matter most.

We next turn to the collateral recovery channel. If higher collateral uncertainty lowers the value lenders can expect to recover in default, then distressed sales should be associated with

[11] Appendix **??** provides additional details on the data and variable construction.

Figure 3: *Collateral uncertainty and appraisal uncertainty (2012-2021)*

(a) *Absolute appraisal gap* **(b)** *Number of comparables used in refinance appraisals*



(c) *Distance of comparables used in refinance appraisals* **(d)** *Severity of under-appraisal*

Note: The figure plots binscatter relationships between county-level collateral value uncertainty and four appraisal-based outcomes, after partialling out year fixed effects. Panel A plots the absolute percentage gap between appraised value and contract price. Panels B and C plot, respectively, the number and average distance of comparable properties used in refinance appraisals. Panel D plots the severity of under-appraisal, conditional on the appraisal coming in below the contract price. Panels A and D are based on purchase appraisals, where contract prices are observed in the public-use file; Panels B and C are based on refinance appraisals. Collateral uncertainty is measured by $\sigma_{c,t}$, the fitted residual variance from the county-level GAM-based hedonic model. Appendix **??** reports the regression output underlying these figures.

larger foreclosure discounts in high-uncertainty locations. To examine this prediction, we use CoreLogic repeat-sales data and estimate county-level foreclosure discounts from regressions of holding-period-adjusted excess returns on an indicator for whether the second sale is distressed. The dependent variable compares the realized resale price to the price implied by local county-level house price appreciation, so the estimated coefficient captures how much lower returns are when the second sale is a foreclosure rather than a normal resale. We then relate the estimated county foreclosure discount to the county average of the single-family collateral uncertainty measure. Figure 4 shows that foreclosure discounts are larger in counties with higher collateral uncertainty. This pattern is consistent with lower expected recovery values in high-uncertainty

Table 9: *Collateral uncertainty and collateral-related denials – Cash-out refinancing*

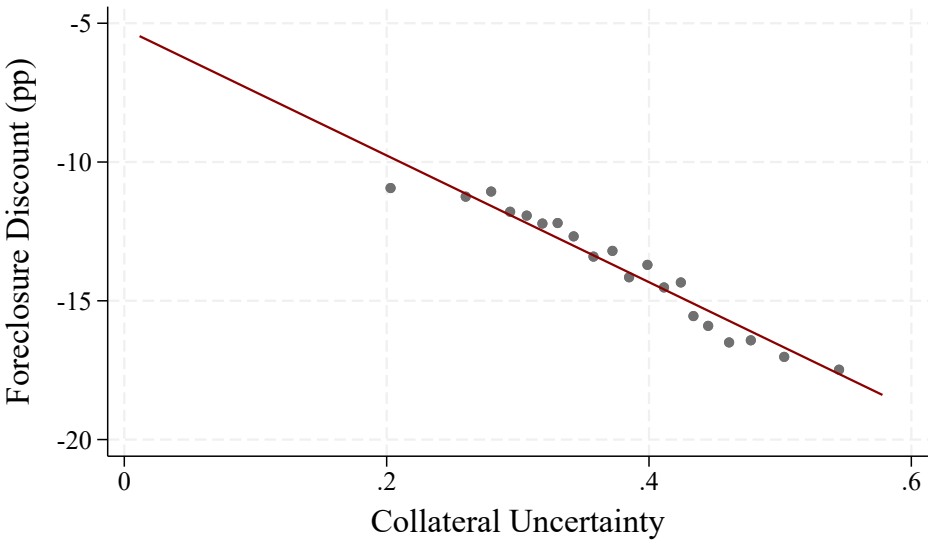
Dep. var: $\mathbb{1}$_Rej.				
	Cash-Out Refinancing		Rate-and-Term Refinancing	
	(1) OLS	(2) IV (2SLS)	(3) OLS	(4) IV (2SLS)
Collateral Uncertainty	0.0836***	0.1646***	0.0665***	0.0981***
	(0.0066)	(0.0202)	(0.0049)	(0.0167)
CLL Sample	✓	✓	✓	✓
Loan Controls	✓	✓	✓	✓
County-Year FE	✓	✓	✓	✓
Institution-Year FE	✓	✓	✓	✓
DTI/LTV Bins	✓	✓	✓	✓
# of Counties	1,497	1,498	1,498	1,499
# of ZIP codes	12,126	12,138	12,137	12,149
First-stage F		26.16		15.58
Clustered SE	County	County	County	County
Observations	3,419,406	3,417,307	6,462,022	6,459,675

Note: This table reports loan-level estimates for refinancing applications over 2018-2024. The outcome equals one if the application is denied for collateral-related reasons and zero otherwise. The data combine HMDA with ZTRAX ZIP-code-level collateral uncertainty, measured at the property's ZIP code and application year. In OLS specifications, collateral uncertainty is the transaction-weighted average across 1-4 unit properties and condos; in IV specifications, it is measured using 1-4 family properties only. Columns 1 and 3 report OLS estimates; columns 2 and 4 report IV (2SLS) estimates in which collateral uncertainty is instrumented with residualized ZTRAX property characteristics (geographic location, log area, year built, and bathrooms) at the ZIP-year level. All columns include county×year and institution×year fixed effects and restrict the sample to conforming-loan-limit applications. All specifications include controls for log applicant income, log loan amount, loan-to-income ratio and its square, DTI and LTV bins, and fixed effects for applicant sex, ethnicity, race, owner-occupancy, property type, loan type, and loan purpose. Standard errors are clustered at the county level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

areas and therefore supports the collateral recovery channel emphasized in the model.

Taken together, the evidence points to both channels operating in the same direction. Higher collateral uncertainty makes appraisals less informative and increases downside appraisal risk, which raises the probability that refinance underwriting becomes more restrictive. At the same time, higher collateral uncertainty lowers the value lenders can expect to recover in foreclosure. Both mechanisms imply that it is rational for lenders to refinance less in high-uncertainty locations. In that sense, the lower refinancing rates documented in the previous subsections are consistent not only with tighter appraisal-based underwriting, but also with weaker collateral values in downside states.

Figure 4: *Foreclosure discount and collateral uncertainty (2000-2017)*



Note: This figure plots the county-level relationship between foreclosure discounts and collateral uncertainty. The unit of observation is a county. The y-axis reports the estimated county foreclosure discount, measured as the coefficient on a distressed-sale indicator from county-specific repeat-sales regressions and expressed in percentage points. The x-axis reports the county average of the single-family-house collateral uncertainty measure over 1999–2017. The figure partials out state fixed effects using the binsreg procedure. More details on the construction of the foreclosure discount measure and the collateral uncertainty variable are provided in the Appendix D.4.

4. Economic consequences of supply driven refinancing frictions

A natural implication of our results is that collateral uncertainty generates household costs by preventing borrowers from refinancing into cheaper mortgage contracts. When refinancing is suppressed in high-uncertainty locations, borrowers remain locked into higher interest rates and forgo payment reductions that would otherwise be available as shown above. This logic is closely related to the "failure to refinance" literature, which documents large unrealized gains from non-refinancing even when interest-rate incentives are strong (Keys et al., 2016; Abel and Fuster, 2021; Eichenbaum et al., 2022). Our contribution is to show that part of these foregone savings is not merely the result of borrower inaction, but instead reflects a location-specific refinancing friction created by collateral uncertainty.

We next examine the economic consequences of these frictions. We study outstanding refinancing opportunities, lenders' retention decisions and borrowers' switching opportunities, and the response of prepayment to mortgage-rate incentives, before quantifying foregone borrower savings.

4.1. Unexercised refinancing opportunities

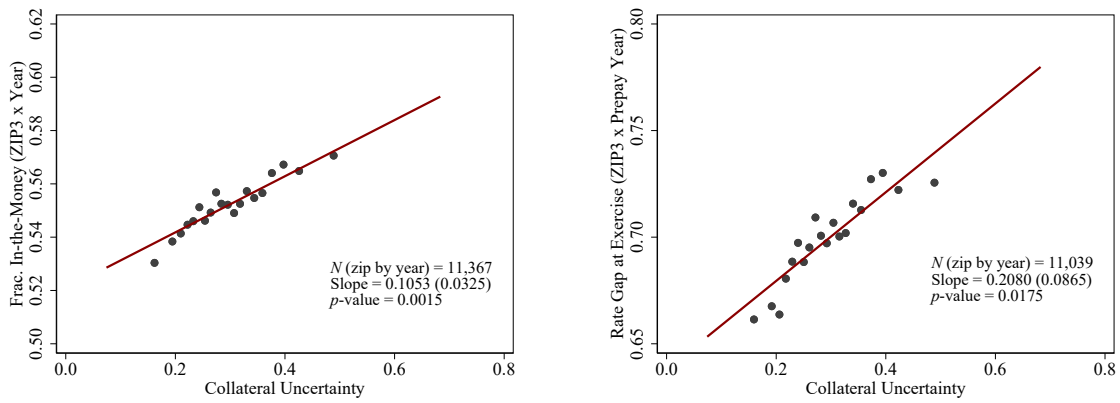
If collateral uncertainty constrains refinancing, borrowers in high-uncertainty locations should be more likely to remain in above-market contracts. That is to say, we should expect a larger fraction of borrowers with a positive spread between the loan rate and the prevailing loan market rate as well as a higher rate gap when exercised the refinancing option. Figure 5 examines this implication using the Freddie Mac data. Panel (a) plots the share of the constructed outstanding loan stock with a note rate above the prevailing mortgage rate, measured by the annual average Freddie Mac Primary Mortgage Market Survey (PMMS). Panel (b) plots the average note-rate gap in the year of prepayment. Both panels absorb commuting zone $\times$ year fixed effects.

Higher collateral uncertainty is associated with a larger share of above-market loans and a larger rate gap at payoff. These patterns complement the loan-level evidence: refinancing opportunities accumulate where collateral is harder to value, and loans that eventually prepay do so at larger observed rate gaps. The latter comparison conditions on prepayment and therefore

captures both exercise behaviour and selection into payoff. Our results are consistent with lenders' frictions affecting borrowers' ability to refinance.

Figure 5: *Collateral uncertainty and unexercised refinancing opportunities (2000-2024)*

(a) *Above-market-rate loan share* **(b)** *Rate gap at prepayment*



Note: The figure plots refinancing-incentive outcomes against collateral uncertainty at the three-digit ZIP code (ZIP3)×year level. Each dot represents the weighted mean of the outcome within one of 20 quantile bins of Collateral Uncertainty, adjusted for commuting zone×year fixed effects; the solid red line shows the weighted linear fit. The reported slope, standard error (in parentheses) and p-value come from the weighted cell regression with standard errors clustered by commuting zone. Loan data are from Freddie Mac's Single-Family Loan-Level Dataset (SFLLD) and ZIP3-year collateral uncertainty from ZTRAX over 2000–2024. Panel (a) reports the original-balance-weighted share of the constructed outstanding stock of loans in year t that have neither prepaid nor reached scheduled maturity with a note rate above the 30-year Primary Mortgage Market Survey (PMMS) rate. Cells require at least 50 such loans and are weighted by their number. Panel (b) reports the original-balance-weighted note-rate gap relative to PMMS among loans prepaid before term. Cells require at least 50 outstanding loans or 20 prepayments, respectively and are weighted by their count; outcomes are winsorized at the 1st and 99th percentiles. Each dot represents one of 20 quantile bins after absorbing commuting zone × year fixed effects. Cell regressions use outstanding-loan counts in Panel (a) and prepayment counts in Panel (b).

4.2. Lender retention and prepayment exposure

Why does collateral uncertainty matter even when mortgages are eligible for sale to the GSEs? Eligibility does not guarantee that a mortgage satisfies purchasers' valuation and delivery requirements or can be sold profitably. Amaral et al. (2026) document greater secondary-market frictions in high-uncertainty locations, including lower use of appraisal alternatives, higher repurchase risk, and longer times to sale. As a result, selling a mortgage may remain costly even when the loan falls below the conforming loan limit.

These frictions can also impede refinancing. Servicing an existing mortgage does not

require a new underwriting decision, whereas refinancing requires another lender to originate a replacement loan. Our earlier evidence shows that refinance denials are more common in high-uncertainty areas. Secondary-market frictions can reinforce this effect, particularly for lenders that rely on loan sales. Consequently, borrowers may face greater difficulty replacing an existing mortgage even when refinancing would lower borrowing costs.

For holders of outstanding mortgages, lower refinancing activity implies lower prepayment risk. When a mortgage carries an above-market coupon, reduced prepayment allows lenders or investors to receive those cash flows for longer. Holding funding costs and credit risk fixed, lower prepayment therefore increases the value of retention.

We evaluate this mechanism using two complementary tests. First, we examine whether collateral uncertainty reduces switching across originators among completed refinancings. Second, we examine whether lenders are more likely to retain rather than sell mortgages backed by more uncertain collateral. The first test captures refinancing frictions on the borrower margin; the second captures lenders' responses to those frictions.

In Table 10 we estimate the likelihood that a property that obtained a successful refinancing was originated by the same lender that first provided the mortgage[12] In our most restrictive specification, we control for county x year fixed effects and property-level controls. The results show a strong negative relationship between collateral uncertainty and the likelihood to switch across lenders. The results indicate that collateral uncertainty increases persistence in lender-

Table 10: *Collateral uncertainty and likelihood to switch lenders (2000-2024)*

	(1)	(2)	(3)
	Year FE	County-Year FE	County-Year FE + Property controls
Collateral uncertainty	-0.0523***	-0.0546***	-0.0317***
	(0.0062)	(0.0055)	(0.0043)
Observations	10,847,377	10,846,811	10,846,811
Adjusted R-squared	0.0309	0.0517	0.0556

Note:

borrower relationships.

As lenders internalize the lower probability of early repayment, they may find retention more attractive. To examine lenders' likelihood to sell or retain an approved mortgage, we use

[12]The property-level data do not allow us to track applications, hence we restrict to successful refinancing status to measure the ability of borrowers to switch lenders.

HMDA-originated loan data. Our baseline specification is

$$\text{Retain}_{ibzt} = \beta\sigma_{zt} + X_i'\Gamma + \alpha_{c(z)\times y(t)} + \alpha_{b\times y(t)} + \varepsilon_{ibzt}, \tag{9}$$

where Retain_{ibzt} is an indicator equal to one if originated loan i is retained by lender b rather than sold within the calendar year, σ_{zt} is collateral uncertainty in the property's ZIP code, X_i is a vector of borrower and loan controls, $\alpha_{c(z)\times y(t)}$ are county $\times$ year fixed effects, and $\alpha_{b\times y(t)}$ are lender $\times$ year fixed effects. Standard errors are clustered at the county level. The vector X_i includes log applicant income, log loan amount, the loan-to-income ratio and its square, debt-to-income and loan-to-value bins, and fixed effects for applicant sex, ethnicity, race, owner-occupancy status, property type, loan type, and loan purpose. The county $\times$ year fixed effects absorb local mortgage-market conditions and county-level credit-supply shifts, while the lender $\times$ year fixed effects absorb time-varying differences in lender funding conditions, business models, and securitization intensity. Once these fixed effects are included, identification comes from comparing otherwise similar loans originated by the same lender in the same year across ZIP codes within the same county that differ in collateral uncertainty.

Table 11 shows that collateral uncertainty increases retention rates in higher collateral uncertainty areas. The relationship remains positive and statistically significant after the inclusion of lender $\times$ year fixed effects, implying that it is not driven by differences in retention strategies across lenders. The estimates are also similar when the sample is restricted to conforming loans, indicating that the pattern is not confined to nonconforming products.

These findings suggest that lenders internalize the refinancing friction created by collateral uncertainty. In locations where uncertainty reduces refinancing and raises denial rates, originated mortgages are also more likely to be retained. This pattern is difficult to reconcile with a purely borrower-demand explanation. Instead, it implies that lenders recognize that mortgages backed by more uncertain collateral face lower prepayment risk and therefore generate more persistent cash flows.

Table 11: *Collateral uncertainty and lenders' loan retention likelihood (2007-2024)*

	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
	Across Lenders	Within Lender	Within Lender	Across Lenders	Within Lender	Within Lender
Collateral Uncertainty	0.0574***	0.0485***	0.0434***	0.2259***	0.1159***	0.1069***
	(0.0104)	(0.0050)	(0.0050)	(0.0240)	(0.0110)	(0.0116)
Loan Controls	✓	✓	✓	✓	✓	✓
County-Year FE	✓	✓	✓	✓	✓	✓
Lender-Year FE		✓	✓		✓	✓
CLL Sample			✓			✓
# of Counties	1626	1626	1626	1626	1626	1626
# of ZIP codes	12,887	12,887	12,886	12,887	12,887	12,886
First-stage F				33.28	35.53	38.95
Clustered SE	County	County	County	County	County	County
Observations	73,726,100	73,724,289	70,208,975	73,726,100	73,724,289	70,208,975

Note: This table reports loan-level estimates of the probability that a loan is retained on the lender's balance sheet for the period 2007-2024. The unit of observation is an individual home mortgage-originated loan; the outcome equals one if the loan is retained by the originating lender and zero otherwise. The data combines the Home Mortgage Disclosure Act (HMDA) and ZTRAX ZIP-code-level (5-digit) collateral uncertainty measures. In OLS specifications, collateral uncertainty is the transaction-weighted average across 1-4 unit properties and condos; in IV specifications, it is measured using 1-4 family properties only. Columns 1 to 3 report OLS estimates; columns 4 to 6 report IV (2SLS) estimates in which collateral uncertainty is instrumented with CoreLogic residualized property characteristics (geographic location, log area, year built, and bathrooms) at the ZIP-year level. Columns 1 and 4 include county×year fixed effects; columns 2-3 and 5-6 additionally include lender×year fixed effects. Columns 3 and 6 restrict the sample to conforming loan limit (CLL) loans only. All specifications include controls for log applicant income, log loan amount, loan-to-income ratio and its square, DTI and LTV bins, and fixed effects for applicant sex, ethnicity, race, owner-occupancy, property type, loan type, and loan purpose. Standard errors are clustered at the county level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

4.3. Prepayment responses to mortgage-rate incentives

Refinancing is a channel through which lower mortgage rates reach existing borrowers and influence monetary policy transmission (Eichenbaum et al., 2022). If collateral uncertainty restricts refinancing access, borrowers in high-uncertainty locations may be less able to take advantage of declining rates. We examine this implication using two complementary outcomes: state dependence in voluntary prepayment of outstanding mortgages and lender denials of refinancing applications.

We first construct a quarterly panel of outstanding loans from the Freddie Mac monthly performance records. At the start of each quarter, we identify outstanding loans that are current on their payments, unmodified, have positive balances and valid mortgage rates. We follow these loans through the quarter and record whether they prepay before contractual maturity. We then group eligible loans by three-digit ZIP code, origination year, and 50-basis-point band of

the original note rate. For each cell and quarter, we calculate the fraction of eligible loans that prepay, the average rate spread, and the composition of borrower and loan characteristics. The spread is the loan's mortgage rate at quarter start minus the latest Primary Mortgage Market Survey (PMMS) rate observed before that quarter, measured in percentage points.

This construction groups mortgages from the same local market and origination vintage with similar initial coupons, and follows their prepayment as refinancing incentives change. A decline in the market rate widens the spread for an outstanding fixed-rate mortgage and increases the potential benefit from refinancing. By allowing the spread coefficient to differ across collateral-uncertainty deciles, we assess whether a given increase in this incentive is associated with a smaller prepayment response where collateral is harder to value.

We estimate the following grouped linear probability model of the quarterly prepayment hazard:

$$h_{cq} = \alpha_{\mathrm{CZ}(c),q,v(c)} + \beta S_{cq} + \sum_{d=2}^{10} \gamma_d D_{cq}^{(d)} + \sum_{d=2}^{10} \delta_d \left(D_{cq}^{(d)} S_{cq} \right) + \overline{\mathbf{X}}_{cq}' \boldsymbol{\theta} + \varepsilon_{cq}. \tag{10}$$

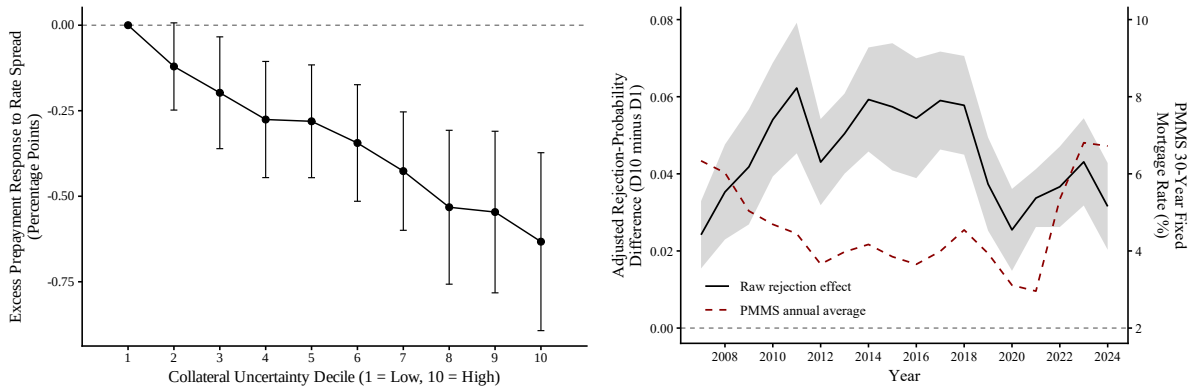
Here, c identifies a ZIP3 $\times$ origination-year $\times$ original-note-rate-band cell, q denotes calendar quarter, and $v(c)$ is the origination year. The outcome h_{cq} is the number of early prepayments during the quarter divided by the number of eligible loans at its start, and S_{cq} is the cell-average rate spread. The indicator $D_{cq}^{(d)}$ assigns the cell to decile d of observation-year ZIP3 collateral uncertainty, using fixed cutoffs from the pooled ZIP3-year distribution; decile 1 is the reference group. The vector $\overline{\mathbf{X}}_{cq}$ contains cell means of borrower and loan controls and categorical shares, including seller shares. Regressions weight cells by their eligible loan counts and absorb commuting zone $\times$ calendar-quarter $\times$ origination-year fixed effects. Standard errors are clustered by commuting zone and calendar quarter. The appendix provides the full risk-set and aggregation rules, including the treatment of maturity and other loan exits.

Panel (a) of Figure 6 plots the scaled percentage point response δ_d, the difference from decile 1 in the quarterly prepayment response to a one-percentage-point wider spread. These differences become more negative as collateral uncertainty rises. In decile 10, the response is 0.63 percentage points smaller than in decile 1. Thus, stronger refinancing incentives are associated with a smaller increase in voluntary prepayment in high-uncertainty locations. The

continuous-uncertainty specification and the decile-level coefficients γ_d, are reported in the appendix Table E.2 and in Figure E.1.

Figure 6: *Collateral uncertainty and responses to rate incentives*

(a) *Prepayment response to rate incentives (2000-2024)* **(b)** *Refinancing denial rates and rates paths (2007-2024)*



Note: Panel (a) plots, for each decile of collateral uncertainty, the difference in the response of the quarterly prepayment probability (percentage points) to a one-percentage-point rate spread, defined as the loan's current rate minus the latest 30-year Primary Mortgage Market Survey (PMMS) rate, relative to the lowest decile. Prepayment is early payoff. The sample is fixed-rate purchase loans from the Freddie Mac Single-Family Loan-Level Dataset (SFLLD), aggregated into ZIP3 by origination year by note-rate band by quarter cells weighted by loans, with loan-characteristic and seller controls and commuting zone by quarter by origination year fixed effects. Deciles use pooled ZIP3-year cutoffs for 2000–2024. Panel (b) plots the annual adjusted difference in the denial probability of Home Mortgage Disclosure Act (HMDA) refinancing applications between the top and bottom within-year deciles of ZIP-level Collateral Uncertainty, controlling for income and loan size, applicant and loan category fixed effects, and state by year and lender by year fixed effects, alongside the annual average PMMS rate (right axis). Collateral Uncertainty is measured from the Zillow Transaction and Assessment Dataset (ZTRAX). Intervals are 95%, clustered by commuting zone and quarter in (a) and by county in (b).

Our next exercise examines whether lender-side refinancing frictions are more pronounced when refinancing is particularly attractive to borrowers. Lower mortgage rates increase the potential savings from refinancing outstanding loans with higher coupons. An increase in the uncertainty-related denial gap during these periods would therefore indicate that the underwriting friction becomes more restrictive precisely when borrowers have more to gain from refinancing.

To examine this implication, Panel (b) of Figure 6 plots the annual adjusted difference in refinancing-denial probabilities between the highest and lowest deciles of ZIP5 collateral uncertainty, together with the annual average PMMS mortgage rate. We estimate a time varying OLS specification similar to Equation 8. The denial contrasts use within-year uncertainty deciles and control for borrower and loan characteristics, state-year fixed effects, and lender-year fixed effects. The excess denial gap widens during periods of low mortgage rates. This pattern is

consistent with collateral uncertainty restricting refinancing access precisely when favourable market rates create valuable opportunities to reduce borrowing costs. Together with the weaker prepayment response in Panel (a), this evidence suggests that lender-side frictions limit the ability of borrowers in high-uncertainty locations to benefit from lower mortgage rates.

Taken together, the two panels capture distinct margins of the refinancing process. Panel (a) shows that prepayment responds less to a given rate incentive where collateral uncertainty is higher; Panel (b) shows that an underwriting gap persists during low-rate periods. Together, these patterns are consistent with lender-side refinancing frictions limiting the transmission of lower mortgage rates to borrowers in high-uncertainty locations. They do not identify the causal response to a monetary-policy shock.

4.4. Foregone borrower savings

We finally study how large the lender-side friction affects household rejection rates and foregone dollar savings. We proceed in two steps. We first estimate the refinance denials that collateral uncertainty accounts for. We then value each of these denials with the savings the borrower forgoes when the refinancing does not take place.

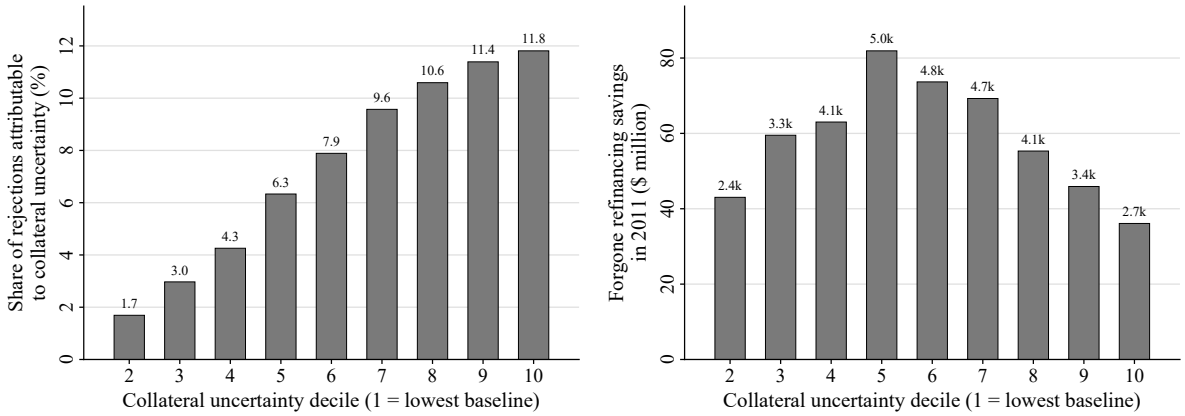
Excess denials: we re-estimate Equation (7), replacing the continuous measure with indicators for within-year deciles of ZIP-code collateral uncertainty using the first decile as reference. The estimated coefficient β_d is the excess denial probability of an application in decile d relative to an otherwise similar application handled by the same lender, in the same county and year, in the lowest-uncertainty decile.

We scale each coefficient by the decile's denial rate, $\beta_d/\bar{r}_d$. This ratio has a counterfactual reading. If applications in decile d were denied at the rate of otherwise similar applications in the bottom decile, the denial rate in decile d would fall from $\bar{r}_d$ to $\bar{r}_d - \beta_d$. The ratio $\beta_d/\bar{r}_d$ is therefore the relative reduction in denials that decile d would experience if it where to face the same denial probability as the lowest decile, or equivalently, the share of its denials attributable to collateral uncertainty.

In Figure E.2 we show that denial probabilities rise monotonically across deciles . Applications in the top decile are 3.5 percentage points more likely to be denied than those in the bottom

Figure 7: *Rejection and forgone saving attributable to collateral uncertainty*

(a) *Excess rejection rate (2007-2024)* **(b)** *Forgone refinancing savings (2011)*



Note: Deciles of ZIP code collateral uncertainty, formed within each year; decile 1 is the baseline. Panel A plots, for each decile, the share of its refinance denials attributable to collateral uncertainty, $\beta_d/\bar{r}_d$: the decile's coefficient from a linear probability model of denial divided by its denial rate. Sample: Home Mortgage Disclosure Act (HMDA) refinancing applications, 2007–2024. Panel B multiplies the extra denials in 2011 (labels, in thousands: 2011 applications times β_d) by the present value of a missed refinancing in each bin computed from Freddie Mac loans active in January 2011, net of a 2 percent refinancing cost. Labels report excess denials (thousands); bar heights report foregone savings ($ million).

decile, where the denial rate is 22.5 percent. When we compare these estimates with the rejection rate in each decile in Panel (a) of Figure 7, we show that the relative reduction in denials rises from 1.7 percent in the second decile to 11.8 percent in the top decile: had top-decile applications faced the denial rate of the bottom decile, about one in eight of their denials would not have occurred. We calculate that applying these reductions to the observed denials in each decile, collateral uncertainty accounts for about 578,000 refinance denials between 2007 and 2024, or 4.3 percent of all refinance denials in our sample. Defining deciles on the pooled 2007–2024 distribution gives nearly identical figures (573,000 denials, 4.25 percent).

The value of a missed refinancing: we next estimate the dollar value of each denial attributable to collateral uncertainty using the Freddie Mac loans outstanding in January 2011. For each fixed-rate loan, we compute the reduction in the monthly payment from refinancing the remaining balance at the prevailing PMMS rate, holding remaining maturity fixed. We discount this reduction over the remaining term and subtract a refinancing cost of 2 percent of the balance, following Keys et al. (2016). We average this net present value over loans with positive net savings in each decile. We denote this average by V_d and refer to it as the value of a missed

refinancing in decile d, Appendix E.4 reports the sample restrictions, the discount rate, and the exact formulas. We focus on 2011 because mortgage rates had fallen well below the coupons of most outstanding loans, so that the gains from refinancing were large, and because the uncertainty denial gap was wide in that year (Panel (b), Figure 6). A missed refinancing is worth between $13,400 and $18,300 in present value, depending on the decile, in line with the $11,500 estimate of Keys et al. (2016) for households that fail to refinance.

Panel (b) of Figure 7 combines the decile estimates from Equation (7) and the estimated V_d to obtain the dollar cost of these denials in 2011. The calculation has two steps. First, we compute the number of excess denials in each decile in 2011 as $\beta_d N_d$, where N_d is the number of refinance applications in decile d in 2011 and β_d is the coefficient from Equation (7). This is the number of denials that would not have occurred had these applications faced the denial rate of the bottom decile. Collateral uncertainty accounts for about 34,500 of the 762,000 refinance denials in 2011. Second, we multiply the excess denials in each decile by the value of a missed refinancing in that decile, so that foregone savings in decile d equal $\beta_d N_d V_d$. Summing across deciles, borrowers whose refinance applications were denied because of collateral uncertainty forgo about $528 million in present value. The bar labels in panel (b) report the excess denials and the bar heights the foregone savings. This is the cost of a single year of applications.

The cost in dollars is distributed differently from the share of denials. Panel (a) rises monotonically, but panel (b) peaks in the middle deciles. Two forces explain the difference. First, high-uncertainty ZIP codes have fewer refinance applications: in 2011, about 612,000 applications came from the second decile and 77,000 from the top decile. Second, loans in these locations are smaller (Jiang and Zhang, 2025), so a missed refinancing is worth about a quarter less in the top decile than in the second. The friction is therefore most severe where collateral is hardest to value, but its aggregate cost is not confined to the tail: most of the foregone savings accrue in moderately uncertain markets that account for a large share of refinancing activity.

These estimates should be taken as an order of magnitude rather than a precise figure, and some of our assumptions affect the estimates. On the one hand, they are likely to understate the true cost. We only count borrowers who actually applied. Borrowers who expect a low appraisal may not apply at all, and the lower number of applications in high-uncertainty markets is consistent with this. Moreover, we apply the average denial gap over 2007–2024 to 2011,

a year in which the gap was larger (Panel (b), Figure 6). On the other hand, we treat every excess denial as a refinancing that never happens and that would have lasted until maturity. In practice, some denied borrowers apply again and succeed, and some would have sold their house or refinanced again before the loan matured. Finally, our benchmark is a relative one with respect to the lowest decile. That is to say, we measure only the cost that borrowers bear in excess of what they would face in the easiest markets to value.

Taken together, the estimates translate the lender-side friction documented in Section 3 into household costs. The same decline in mortgage rates that lowered payments for borrowers in low-uncertainty markets left a sizable share of applicants elsewhere in above-market contracts, adding up to roughly half a billion dollars in foregone savings in a single year.

5. Conclusion

This paper shows that the difficulty of valuing a house restricts the ability of its owner to refinance. Using idiosyncratic price dispersion in housing transactions to measure collateral value uncertainty, we find that borrowers in high-uncertainty locations prepay less, including those with a clear financial incentive to refinance and in property records where a home sale cannot explain the payoff. We show that this gap reflects a lender-side friction. The same lender, in the same year, is more likely to deny refinance applications backed by more uncertain collateral, and more often cites collateral as the reason. Appraisal and foreclosure data point to two channels: appraisals are noisier where comparable sales are scarce, and expected recoveries in default are lower. Lenders respond to this friction. Mortgages backed by more uncertain collateral prepay less and, when refinanced, are more likely to stay with the same lender, and lenders are more likely to keep them on their balance sheets. Where collateral is harder to value, prepayment also responds less to lower mortgage rates, by about one-third in the most uncertain markets relative to the least uncertain ones. Collateral uncertainty accounts for about 578,000 refinance denials between 2007 and 2024, and the resulting foregone savings amount to about $528 million in 2011 alone. Lenders respond to this friction: mortgages backed by more uncertain collateral are more likely to be retained, and borrowers who do refinance are less likely to switch lender.

These findings have two broader implications for mortgage markets. First, they imply that the pass-through of lower interest rates to household balance sheets is spatially uneven. In markets where collateral is easier to value, declines in mortgage rates translate more readily into completed refinancings and lower monthly payments. In markets where collateral is harder to value, the same decline in rates produces less refinancing and weaker relief for borrowers. Second, our results suggest that standardization in secondary mortgage markets can miss local refinancing frictions. Appraisal uncertainty is highly local and depends on the informativeness of nearby comparable transactions, yet standardized underwriting and pricing frameworks are not designed to fully incorporate this type of place-specific collateral risk. These two implications are closely related: when secondary-market structures do not fully absorb local valuation frictions, those frictions remain embedded in primary-market refinancing decisions (Amaral et al., 2026), and interest-rate pass-through becomes uneven across space.

Our results also point to a natural policy implication. If collateral uncertainty blocks otherwise beneficial refinancing, then improving the precision of collateral valuation should ease an economically meaningful friction in mortgage markets. Policies that improve appraisal quality, expand access to better property-level data, or reduce comparable-scarcity problems in thin and heterogeneous housing markets could improve refinancing access precisely where current underwriting is most restrictive. In that sense, the paper highlights a channel through which mortgage-market design and appraisal infrastructure can shape the distributional effects of lower interest rates.

References

Abel, J. and Fuster, A. (2021), 'How do mortgage refinances affect debt, default, and spending? evidence from harp', *American Economic Journal: Macroeconomics* **13**(2), 254–291.

Agarwal, S., Rosen, R. J. and Yao, V. (2016), 'Why do borrowers make mortgage refinancing mistakes?', *Management Science* **62**(12), 3494–3509.

Amaral, F., Blunier, J. and Ruzzier, G. (2026), 'Collateral uncertainty and the limits of shadow-bank mortgage lending'.

Andersen, S., Campbell, J. Y., Nielsen, K. M. and Ramadorai, T. (2020), 'Sources of inaction in household finance: Evidence from the danish mortgage market', *American Economic Review* **110**(10), 3184–3230.

Belgibayeva, A., Bono, T., Bracke, P., Clara, N., Cocco, J. F. and Majer, T. (2025), 'When discounted rates end: The costs of taking action in the mortgage market', *The Review of Financial Studies* p. hhaf098.

Beraja, M., Fuster, A., Hurst, E. and Vavra, J. (2019), 'Regional heterogeneity and the refinancing channel of monetary policy', *The Quarterly Journal of Economics* **134**(1), 109–183.

Berger, D., Milbradt, K., Tourre, F. and Vavra, J. (2025), 'Optimal mortgage refinancing with inattention', *American Economic Review: Insights* **7**(4), 497–515.

Consumer Financial Protection Bureau (2013), 'Ability-to-repay and qualified mortgage standards under the truth in lending act (regulation z)', Final Rule. Final rule issued January 10, 2013.
URL: *https://files.consumerfinance.gov/f/201301*$_{c}fpb_{f}inal - rule_{a}bility - to - repay.pdf$

DeFusco, A. A. and Mondragon, J. (2020), 'No job, no money, no refi: Frictions to refinancing in a recession', *The Journal of Finance* **75**(5), 2327–2376.

Eichenbaum, M., Rebelo, S. and Wong, A. (2022), 'State-Dependent Effects of Monetary Policy: The Refinancing Channel', *American Economic Review* **112**(3), 721–761.

Jiang, E. X. and Zhang, A. L. (2025), 'Collateral value uncertainty and mortgage credit provision', *Journal of Financial Economics* **169**, 104054.

Keys, B. J., Pope, D. G. and Pope, J. C. (2016), 'Failure to refinance', *Journal of Financial Economics* **122**(3), 482–499.

Supplemental Appendix for "Valuation Uncertainty and Lender Frictions in Mortgage Refinancing"

A. Data and sample construction

This appendix provides additional information on the data sources and samples introduced in Section 2. Appendix A.1 describes the Freddie Mac loan-level data, Appendix A.2 the ZTRAX property records and the construction of refinancing events, Appendix A.3 the HMDA application data, Appendix A.4 the appraisal data, and Appendix A.5 the remaining sources.

A.1. Freddie Mac Single-Family Loan-Level Dataset

We use the Freddie Mac Single-Family Loan-Level Dataset (SFLLD), which contains origination characteristics and monthly performance records for fixed-rate mortgages acquired by Freddie Mac. Our sample consists of fixed-rate, single-family purchase loans originated between 2000 and 2024. We apply the following exclusions, in this order:

1. loans with a prepayment penalty;
2. loans originated under the Home Possible, HFA Advantage, or Refi Possible affordable-housing programs (program codes H, F, and R; we keep code 9, "not applicable");
3. Relief Refinance loans, including HARP, identified by the Relief Refinance indicator;
4. interest-only loans;
5. loans with a FICO score that is missing or outside the range [300, 850];
6. loans whose 3-digit ZIP code and origination year have no collateral uncertainty value.

Because the sample contains purchase loans only, the two refinance-program exclusions remove no additional loans. The Standard dataset contains Home Possible loans only from March 2015 onward. The resulting sample contains 8,048,748 loans; Table 1 reports summary statistics.

Prepayment identification: The performance files record a zero-balance code and the month in which the balance reaches zero. We code a loan as prepaid if its zero-balance code indicates a voluntary payoff (code 01) and the payoff occurs before the contractual maturity date, defined as the first payment date plus the original term minus one month. Payoffs on or after the maturity date are coded as zero. All other terminations (third-party sales, short sales and charge-offs, REO dispositions, whole-loan and reperforming-loan sales, and repurchases; codes 02, 03, 09, 15, 16 and 96) are coded as zero in the quarter in which they occur. The loan stays in that quarter's risk set and leaves the panel afterwards.

The SFLLD reports the property's 3-digit ZIP code (ZIP3). We aggregate 5-digit ZIP-code collateral uncertainty to the ZIP3-by-year level as a mean weighted by the number of single-family transactions. We assign each ZIP3 to a single commuting zone in two steps. We first link 5-digit ZIP codes to counties using the HUD-USPS ZIP-county crosswalk (2020Q4), and then link counties to commuting zones. Each ZIP3 goes to the commuting zone that holds the largest share of its 5-digit ZIP codes. A ZIP code that spans several counties is split by the share of its residential addresses in each county. Commuting zones define the fixed effects and the clustering unit in all Freddie Mac specifications.

A.2. ZTRAX property records and refinancing events

The Zillow Transaction and Assessment Dataset (ZTRAX) combines deed and mortgage records from county recorders with parcel characteristics from county assessors. We use ZTRAX for two purposes: to estimate collateral uncertainty (Appendix B.1) and to construct property-level refinancing events (Section 3.2).

Transaction cleaning: We keep transactions for single-family properties and drop transactions such as sales between relatives. We drop transactions that involve more than one parcel, because the price of the individual property is not identified; exact duplicates and repeated transactions of the same property on the same day; transactions with a missing price, date, year built, or living area; and transactions financed by the Federal Housing Administration, the Department of Veterans Affairs, or the Farm Service Agency or Rural Housing Service. To limit the influence of outliers, we winsorize transaction prices and living area at the 1st and 99th

percentiles within each year. We exclude Alaska, Hawaii, and the U.S. territories because of limited data coverage.

We link each cleaned single-family purchase to the subsequent mortgage documents recorded on the same parcel. A purchase spell begins 90 days after the purchase date and ends at the next recorded sale of the property or at the end of the available records, whichever comes first. Because the spell ends at the next sale, a property sale cannot itself generate a refinancing event. The analysis sample contains 24,983,923 purchase spells from the 2000–2020 purchase cohorts in 46 jurisdictions.

Refinancing events: The outcome equals one if, within the spell, the parcel receives either (i) a mortgage explicitly coded as a refinance, or (ii) a qualifying mortgage recording consistent with refinancing. A recording qualifies if it occurs at least 180 days after the purchase, has a loan amount of at least $50,000 and at least 25 percent of the purchase price, and is not coded as a purchase-money, home-equity, second or subordinate, modification, reverse, future-advance, or open-ended loan. The outcome mean is 0.491. Because loan-purpose coding is incomplete in the recorder data, part of the events is inferred from category (ii); we therefore refer to the outcome as a mortgage recording consistent with refinancing.

The property controls in Table 5 are log purchase price, log square footage, building age, attached-property status, number of bathrooms, owner occupancy, land use, latitude, longitude, and their interaction. All specifications include bins of the length of the observable spell, so that properties observed for longer are not mechanically more likely to record an event.

A.3. HMDA application data

We use the Home Mortgage Disclosure Act (HMDA) loan application register for 2007–2024. For 2007–2013 we use the historical files from the National Archives catalog (series 2456161); for 2014–2016 the FFIEC HMDA flat files; and from 2017 onward the national loan-level datasets published by the FFIEC.

We keep conventional loans (i.e., loans not insured or guaranteed by the FHA, VA, or FSA/RHS) on one-to-four-family properties, with loan purpose home purchase, home improvement, or refinancing (including cash-out refinancing). We keep applications that were either

originated or denied. We drop applications with a loan amount below $50,000 or above the 99th percentile of the annual loan-amount distribution, applications with applicant income below $1,000, and applications for properties outside the 50 states and the District of Columbia. We also drop applications with a missing or invalid state or county code.

HMDA reports the property's census tract. We assign each tract to a 5-digit ZIP code using the HUD-USPS tract-to-ZIP crosswalk, taking the fourth-quarter vintage closest to the application year (the 2010 vintage for 2007–2009). When a tract spans several ZIP codes, we assign it to the ZIP code containing the largest share of the tract's residential addresses. For 2022 onward, HMDA reports 2020 tract definitions; tracts not found in the crosswalk are first mapped to the 2010 tract that contains the largest share of their area (at least half) using the Census tract relationship file. We keep applications whose assigned ZIP code has a ZTRAX collateral uncertainty measure for the application year.

Lender identifiers: we map each HMDA respondent to its Federal Reserve RSSD identifier using the HMDA–RSSD match constructed by Robert Avery.[13] The denial and retention samples are restricted to lenders matched to an RSSD identifier. The lender$\times$year fixed effects are defined on this identifier.

The denial indicator equals one if the application is denied and zero if it is originated. The collateral-denial indicator, defined on denied applications only, equals one if collateral is reported as the primary denial reason. The retention indicator, defined on originated loans only, equals one if the loan is not sold within the calendar year of origination or is sold to an affiliated institution. The cash-out refinancing indicator is available from 2018 onward. From 2018 onward we also observe the debt-to-income (DTI) and combined loan-to-value (LTV) ratios, which enter the 2018–2024 specifications as binned fixed effects.

Conforming loans and large banks identification: We classify a loan as conforming-loan-limit (CLL) eligible if its amount is at or below the FHFA conforming loan limit for its county and year. The large-bank samples are restricted to banks subject to Dodd-Frank Act stress testing (DFAST) in 2018–2024, mapped to the RSSD identifier of their lead bank. The list contains 34 banks.

[13] The file is available at Neil Bhutta's web page. It matches every lender that has filed an HMDA report to the corresponding entity in the Federal Reserve's National Information Center (NIC). If the filer is a subsidiary of a bank holding company, it is matched to the lead (largest) bank of the holding company.

A.4. Enterprise Uniform Appraisal Dataset

We use the FHFA Enterprise Uniform Appraisal Dataset (UAD) Appraisal-Level Public Use File, which provides appraisal-level information for a nationally representative sample of Enterprise appraisals of one-unit single-family properties. The file contains purchase and refinance appraisals and reports appraised values, contract prices for purchase transactions, and characteristics of the comparable properties used in the appraisal.

We use the purchase appraisals for outcomes that require the contract price: the absolute appraisal gap, $100 \times |\text{appraised value} - \text{contract price}|/\text{contract price}$, and under-appraisal intensity, $100 \times (\text{contract price} - \text{appraised value})/\text{contract price}$, among appraisals below the contract price. We use the refinance appraisals for outcomes that depend on the comparable set: the number of comparable properties in the appraisal report and their reported distance from the subject property. Because our collateral uncertainty measure is available at the county-by-year level, we merge appraisals to it by county and year. The sample covers 2013–2022. For 2018–2022, we assign each county its average collateral uncertainty over 2014–2017. Figure 3 reports binned scatter plots after partialling out year fixed effects, and Appendix D.3 reports the underlying regressions.

A.5. Other data sources

- **Mortgage rates.** Freddie Mac Primary Mortgage Market Survey (PMMS), 30-year fixed-rate mortgage, weekly series averaged to months, quarters, or years as required.
- **Discount rate.** 10-year Treasury constant-maturity yield (FRED series DGS10), used to discount refinancing savings (Appendix E.4).
- **Foreclosure discounts.** CoreLogic repeat-sales pairs, 2000–2017, used to estimate county foreclosure discounts (Appendix D.4).
- **Housing-market liquidity.** Redfin county-level median days on market and average sale-to-list price ratio (Appendix B.3).
- **Conforming loan limits.** FHFA county-level conforming loan limits by year.

- **Large banks.** Federal Reserve lists of DFAST participants, 2018–2024.
- **Geographic crosswalks.** Census-tract-to-ZIP (HMDA), ZIP3-to-commuting-zone (Freddie Mac), and county-to-commuting-zone crosswalks.

B. The collateral uncertainty measure

This appendix complements Section 2. Appendix B.1 details the estimation of collateral uncertainty, Appendix B.2 describes its distribution and persistence, and Appendix B.3 relates it to housing-market liquidity.

B.1. Estimation of collateral uncertainty

Collateral uncertainty captures how precisely the market value of a property can be inferred from observable characteristics and nearby transactions. In thick markets with a homogeneous housing stock, observably similar homes sell at similar prices, and a lender or appraiser can value a property with little error. In thin markets, or where properties are atypical, comparable sales are scarce and observably similar homes sell at widely dispersed prices. Following Jiang and Zhang (2025), we measure this valuation difficulty as the predicted idiosyncratic dispersion of transaction prices. We construct the measure in three stages, estimated separately for each county on the cleaned ZTRAX transactions described in Appendix A.2. Estimating county by county allows each characteristic, and its dispersion, to differ across local markets.

Hedonic model: For each county c, we estimate a hedonic model of the log transaction price with a generalized additive model (GAM),

$$\ln(p_{it}) = \mu_{m(t)} + s_1(\text{lat}_i, \text{lon}_i, t) + s_2(\text{sqft}_i, t) + s_3(\text{yearbuilt}_i, t) + s_4(\text{bed}_i, t) + s_5(\text{bath}_i, t) + u_{it}, \quad \text{(B.1)}$$

where p_{it} is the price of property i sold at date t. Each term absorbs a distinct source of predictable price variation. $\mu_{m(t)}$ are year-month fixed effects, which capture high-frequency movements in the county price level, so that the residual is not driven by when a property sells.

$s_1(\text{lat}_i, \text{lon}_i, t)$ is a smooth surface over location that is allowed to shift over time. It plays the role of fine-grained neighbourhood effects and captures local trends such as gentrification, without imposing administrative boundaries such as ZIP codes. s_2 to s_5 are smooth functions of living area, year built, bedrooms, and bathrooms, each allowed to change over time, so that, for example, the premium for additional space or for newer construction can evolve differently across periods.

We use a GAM rather than a fixed polynomial in the characteristics because it lets the data determine the shape of each price gradient and of the location surface, which reduces the risk that misspecification of the hedonic function is mistaken for dispersion. The residual u_{it} is the part of the transaction price that observable characteristics, location, and timing cannot explain. Its square, u_{it}^2, is therefore an ex post measure of the idiosyncratic price dispersion of transaction i.

Predicted dispersion: the realized residual is observed only after the transaction and contains the realized price shock of the specific sale. What matters for a lender who underwrites a mortgage is the dispersion it can expect, given the information available before the sale. We therefore project the squared residuals on the same observables in a second county-specific GAM, corresponding to Equation (3) in the main text. The fitted value $\hat{g}_c(\cdot)$ is the expected squared pricing error for a property with these characteristics, at this location and date. It is high for property types and neighbourhoods whose prices are hard to predict, for example unusually large or old homes, or areas with few recent comparable sales. Because a linear smoother can produce small negative fitted variances, we winsorize the fitted variance, set negative values to zero, and take the square root to obtain the property-level predicted standard deviation of the pricing error (Equation (4)). Expressed in log points, $\hat{\sigma}_{it}$ approximates the expected percentage error with which the value of property i can be inferred from observables.

Aggregation: we average the property-level measure over the $N_{z,t}$ transactions in ZIP code z and year t, as in Equation (5). $\sigma_{z,t}$ measures the average predicted price dispersion of properties transacted in a ZIP code and year, which we interpret as collateral value uncertainty. The panel covers 2000–2024; we exclude the partial 2025 vintage.

We construct two variants. The baseline measure, *Collateral Uncertainty*, is estimated on transactions of one-to-four-unit properties and condominiums and enters all OLS specifications.

Collateral Uncertainty (SFH) is estimated on single-family homes only, the property type on which the instruments of Appendix C are constructed, and is the endogenous variable in all IV specifications. Each measure is trimmed at the 0.8th and 99.5th percentiles of its 2000–2024 distribution across ZIP-code-years. Table 3 reports summary statistics, and Appendix B.2 describes the distribution and persistence of the measure.

Other analyses use coarser geographies. Because the Freddie Mac data report only the 3-digit ZIP code, we aggregate $\sigma_{z,t}$ to the ZIP3-by-year level as a mean weighted by the number of single-family transactions. Because the appraisal and repeat-sales data identify only the county, the appraisal and foreclosure analyses use county-level averages.

By construction, $\sigma_{z,t}$ measures the dispersion of prices around their predicted value, not the level of prices or housing quality: the hedonic model absorbs differences in average prices across locations, periods, and property types. Appendices B.3 and D.3 show that the measure is higher where housing markets are less liquid and where appraisals are noisier and rely on fewer and more distant comparables, consistent with its interpretation as valuation difficulty.

B.2. Distribution and persistence

Figure B.1 plots the distribution of ZIP-code collateral uncertainty over 2000–2024 across the sample and over five-year periods. Across the 334,768 ZIP-code-years in the sample, the mean is 0.354, the standard deviation is 0.124, and the interquartile range runs from 0.260 to 0.441. The median rises from 2000–2004 to 2010–2014 and falls in 2020–2024, but within every period the dispersion across ZIP codes is large relative to these shifts. Some local housing markets are therefore systematically much harder to value than others, rather than fluctuating around a common national level.

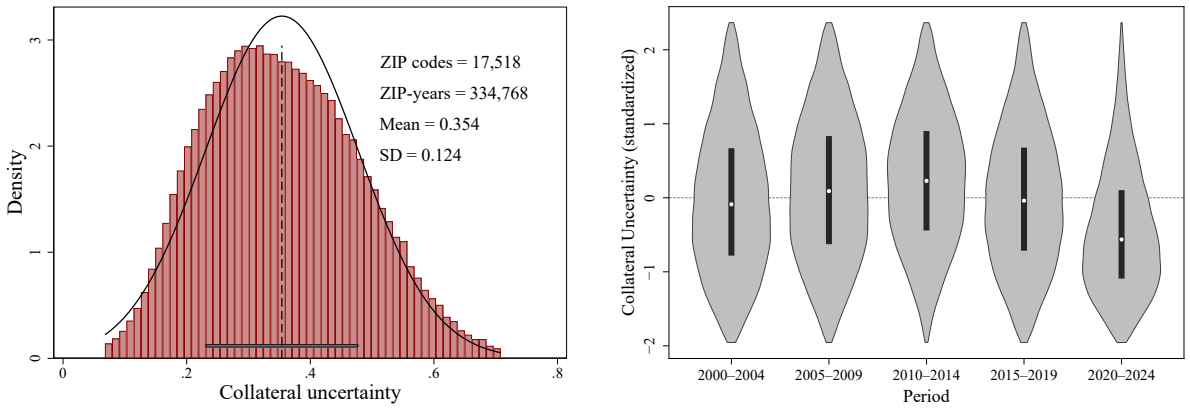
Panel (a) in Figure B.2 maps the average collateral uncertainty of each ZIP code over 2000–2024. Uncertainty varies widely both across and within states, and neighbouring ZIP codes often fall in different quintiles. They are concentrated in states that do not require public disclosure of sale prices, such as Texas, Utah, and Kansas, where ZTRAX records few transaction prices.

Panel (b) in Figure B.2 plots ZIP-code collateral uncertainty against its value five years earlier. The relationship lies close to the 45-degree line: ZIP codes with high uncertainty at

the beginning of a five-year window remain high-uncertainty ZIP codes at its end, and low-uncertainty ZIP codes remain comparatively easy to value. Collateral uncertainty therefore reflects durable features of local housing markets, such as the heterogeneity of the housing stock and the frequency of comparable transactions, rather than short-run shocks.

Figure B.1: *Distribution of collateral uncertainty (2000–2024)*

(a) *Distribution* **(b)** *Five-year distribution*



Note: Panel (a) plots the distribution of Collateral Uncertainty across 5 digits ZIP codes and years between 2000 and 2024. Panel (b) shows violin plots of ZIP-code-by-year collateral uncertainty by five-year period, standardized over the pooled 2000–2024 panel. The black bar marks the interquartile range and the white dot the median. Collateral uncertainty is estimated from ZTRAX transactions.

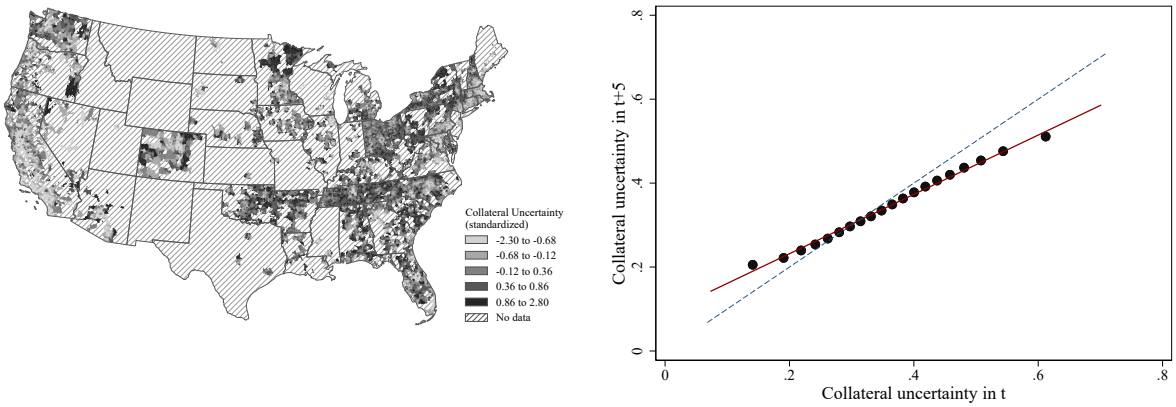
B.3. Collateral uncertainty and housing-market liquidity

We further validate the collateral uncertainty measure by relating it to housing-market liquidity. We use county-level data from Redfin on the median time on the market and the average ratio of sale price to list price.[14] Figure B.3 shows that counties with higher collateral uncertainty have homes that stay on the market longer (panel a) and sell at lower prices relative to their list price (panel b). These patterns are consistent with our measure being higher in thinner and less liquid housing markets, where valuation is more difficult and comparable transactions are less informative.

[14]Redfin is a real estate brokerage that compiles housing-market statistics from local multiple listing services. We use Redfin rather than Zillow because Zillow does not provide time-on-market data at a comparably granular geographic level. See https://www.redfin.com/news/data-center/.

Figure B.2: *Geography and persistency in collateral uncertainty (2000–2024)*

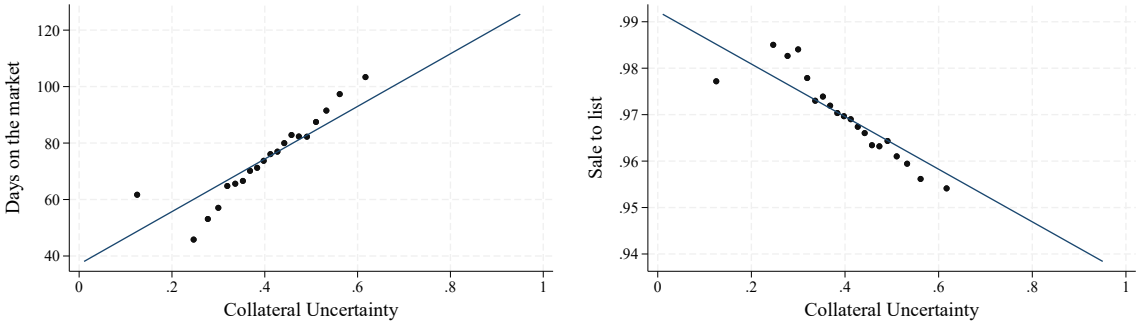
(a) *Geography* **(b)** *Five-year persistence*



Note: Panel (a) plots the 5-digit ZIP-code average collateral uncertainty over the years in 2000–2024. Shades are quintiles across mapped areas. Legend reports the quintile bounds in standard deviations. ZIP codes are matched to 2020 Census ZIP Code Tabulation Areas (ZCTAs), and ZIP codes that share a ZCTA are averaged. Panel (b) is a binned scatter plot of ZIP-code collateral uncertainty in year $t + 5$ against its value in year t; the dashed line is the 45-degree line. Both panels use the standardized measure over the pooled ZIP-code-year panel. Collateral uncertainty is estimated from ZTRAX transactions.

Figure B.3: *Collateral uncertainty and housing-market liquidity, 2012–2017*

(a) *Time on market* **(b)** *Sale-to-list price ratio*



Note: Binned scatter plots of county-level collateral uncertainty against two Redfin housing-market liquidity measures. Panel (a): median days on market. Panel (b): average sale-to-list price ratio.

C. Instrumental variable: housing-stock atypicality

We describe the construction of the instruments for ZIP-code-by-year collateral uncertainty, $\sigma_{z,t}$, introduced in Section 2.3.1 and discuss the exclusion restriction.

C.1. Construction

The main empirical concern is that $\sigma_{z,t}$ may be correlated with unobserved local conditions that also affect mortgage outcomes, such as borrower composition, local credit demand, or sorting. Following Jiang and Zhang (2025), we instrument $\sigma_{z,t}$ with measures of the *atypicality* of the local housing stock. When the housing stock is more heterogeneous, close comparables are scarcer, appraisals rely on fewer and more distant comparables, and prices of observably similar homes are more dispersed. We measure atypicality relative to each ZIP code's own housing stock in a given year, so the instruments do not capture whether homes are large, new, or of high quality on average. We construct the instruments from the cleaned ZTRAX transactions in three steps.

Distance to the typical property: for each ZIP code z and year t, we compute the mean of each housing feature,

$$\bar{X}_{m,z,t} = \frac{1}{N_{z,t}} \sum_{i \in (z,t)} X_{m,i}, \qquad m \in \{\text{geo}, \text{area}, \text{age}, \text{bath}\}, \tag{C.1}$$

where the features are geographic location, log living area, year built, and the number of bathrooms. For each transacted property i in ZIP code z and year t, we define its squared distance from the ZIP-code-year mean along each feature,

$$Z_{m,i} = \left(X_{m,i} - \bar{X}_{m,z,t}\right)^2. \tag{C.2}$$

For location, $Z_{\text{geo},i} = (\text{lat}_i - \overline{\text{lat}}_{c,t})^2 + (\text{lon}_i - \overline{\text{lon}}_{c,t})^2$ is the squared distance of the property's coordinates from the county-year mean coordinates. $Z_{m,i}$ is large when property i is unusual relative to the local stock, for example a very small home in a county dominated by larger homes.

residualization: distances can be mechanically correlated with feature levels, because

housing characteristics are not symmetrically distributed and extreme values (for example, very large homes) may be associated with the socioeconomic characteristics of owners. We therefore residualize each $Z_{m,i}$ on the levels of all housing features and transaction-date fixed effects,

$$Z_{m,i} = \sum_k \alpha_{m,k} X_{k,i} + \mu_{d(i)} + \epsilon_{m,i}, \tag{C.3}$$

where $\mu_{d(i)}$ are transaction-month fixed effects. The property-level instruments are the residuals, $\widetilde{Z}_{m,i} = \hat{\epsilon}_{m,i}$. They measure how atypical a property is relative to its local market, net of feature levels and common time effects.

Step 3: aggregation. We average the residualized distances to the ZIP-code-year level,

$$\widetilde{Z}_{m,z,t} = \frac{1}{N_{z,t}} \sum_{i \in (z,t)} \widetilde{Z}_{m,i}. \tag{C.4}$$

ZIP codes with higher $\widetilde{Z}_{m,z,t}$ have more heterogeneous housing stocks and, consequently, thinner comparable sets. For the Freddie Mac analysis, the instruments are aggregated to the ZIP3-by-year level in the same way as collateral uncertainty.

C.2. Exclusion restriction

The identifying assumption is that, conditional on the controls and fixed effects, housing-stock atypicality affects mortgage outcomes only through collateral uncertainty, and not through borrower risk, credit demand, or sorting. Two features of the construction support this assumption. First, atypicality is defined within local markets, so the instruments capture dispersion around the local housing stock rather than its average quality. Second, the residualization in Equation (C.3) removes the linear relationship with the levels of the observed housing characteristics and with common transaction-date shocks.

Jiang and Zhang (2025) use the same instruments in the context of mortgage lending. They show that the instruments are not systematically related to observable ex ante borrower characteristics and do not predict ex post mortgage delinquency. Because we apply the same design to refinancing outcomes, this evidence is informative for the exclusion restriction here as

well.

D. Refinancing and denials: additional results

This appendix reports additional results for Section **??**: prepayment in a broader sample of agency loans (Appendix D.1), robustness of the denial results (Appendix D.2), the regressions underlying the appraisal evidence (Appendix D.3), and the construction of foreclosure discounts (Appendix D.4).

D.1. Prepayment in the full agency sample

The baseline prepayment sample in Table 4 contains fixed-rate purchase loans without prepayment penalties or special-program features. Table D.1 re-estimates the OLS and IV specifications on a broader sample that pools Fannie Mae and Freddie Mac loans and includes both purchase and refinance originations. The coefficient on collateral uncertainty remains negative and statistically significant in all specifications, so the baseline result does not reflect selection into particular loan types. This sample covers originations in 2007–2021; the coefficients are therefore not directly comparable in magnitude with the baseline.

D.2. Refinancing denials: sample periods and large banks

HMDA reporting changed in 2018: from that year on, lenders report the applicant's debt-to-income (DTI) and combined loan-to-value (LTV) ratios. Table D.2 estimates the OLS specifications of Equation (7) separately for 2007–2017 and 2018–2024, adding DTI and LTV bins in the second period. The coefficient on collateral uncertainty is positive and significant in both periods and under each set of fixed effects. That it remains positive and significant in 2018–2024, when DTI and LTV bins are included, mitigates concerns that the baseline results reflect unobserved loan-level risk, particularly equity, rather than collateral uncertainty.

Large banks. Smaller lenders tend to operate in geographically concentrated markets, which limits within-lender variation in collateral uncertainty. Table D.3 restricts the sample to banks subject to Dodd-Frank Act stress testing (DFAST) over 2018–2024. These banks lend across a

Table D.1: *Collateral uncertainty and mortgage prepayment: Fannie Mae and Freddie Mac loans (2007–2021)*

Dep. var: Loan Prepaid Indicator (= 1 if loan prepaid early)				
	OLS		IV	
	(1) All Loans	(2) + Seller FE	(3) All Loans	(4) + Seller FE
Collateral Uncertainty	-0.0528***	-0.0521***	-0.1661***	-0.1613***
	(0.0158)	(0.0155)	(0.0447)	(0.0430)
Loan Controls	✓	✓	✓	✓
Loan Type FE	✓	✓	✓	✓
CZ-Year Qtr FE	✓	✓	✓	✓
Seller-Year FE		✓		✓
Clustered SE	Commuting Zone	Commuting Zone	Commuting Zone	Commuting Zone
First-Stage F			4.63	4.65
# of CZ	328	328	328	328
# of ZIP3	678	678	678	678
Observations	59,177,942	59,177,910	59,169,822	59,169,790

Note: Loan-level linear probability estimates on refinancing likelihood for Fannie Mae and Freddie Mac single-family purchase and refinance loans originated in 2007–2024. The outcome equals one if the loan is prepaid before maturity. Collateral uncertainty is measured at the 3-digit ZIP code and origination year; loans originated after 2017 are assigned the ZIP3 average over 2014–2017. Columns 1 and 2 report OLS estimates; columns 3 and 4 report 2SLS estimates, in which collateral uncertainty for single-family homes is instrumented with residualized property characteristics. All columns include loan controls and commuting zone×year-quarter fixed effects; columns 2 and 4 add seller×year fixed effects. Standard errors are clustered by commuting zone. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

Table D.2: *Collateral uncertainty and refinancing denials by sample period*

	2007-2017			2018-2024		
	(1) Across Lenders	(2) Within Lender	(3) Within Lender	(4) Across Lenders	(5) Within Lender	(6) Within Lender
Collateral Uncertainty	0.0398***	0.1109***	0.1185***	0.0807***	0.0718***	0.0755***
	(0.0143)	(0.0092)	(0.0091)	(0.0063)	(0.0050)	(0.0053)
CLL Sample			✓			✓
Loan Controls	✓	✓	✓	✓	✓	✓
County-Year FE	✓	✓	✓	✓	✓	✓
Institution-Year FE		✓	✓		✓	✓
# of Counties	1596	1596	1596	1509	1509	1508
# of ZIP codes	12,678	12,678	12,678	12,162	12,162	12,159
DTI/LTV Bins				✓	✓	✓
Clustered SE	County	County	County	County	County	County
Observations	37,026,985	37,025,535	35,484,682	10,330,404	10,329,695	9,882,765

Note: This table reports OLS estimates for refinancing applications in the Home Mortgage Disclosure Act (HMDA) data that were either originated or denied; the outcome equals one if the application was denied. Collateral Uncertainty is measured at the ZIP-code-by-year level from ZTRAX transaction records. Columns 4 to 6 add DTI and LTV bins, which are available only from 2018 onward. The CLL sample keeps loans at or below the county conforming loan limit. Loan controls are log applicant income, log loan amount, and the loan-to-income ratio and its square; all columns also include fixed effects for applicant and loan characteristics. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

broad range of local housing markets, so within-lender variation in collateral uncertainty is not constrained by geographic specialization.

Table D.3: *Collateral uncertainty and refinancing denials: large banks (2018–2024)*

	OLS			IV		
	(1) Across Lenders	(2) Within Lender	(3) Within Lender	(4) Across Lenders	(5) Within Lender	(6) Within Lender
Collateral Uncertainty	0.0594***	0.0719***	0.0822***	0.1817***	0.1612***	0.1834***
	(0.0085)	(0.0062)	(0.0070)	(0.0330)	(0.0308)	(0.0314)
CLL Sample			✓			✓
Loan Controls	✓	✓	✓	✓	✓	✓
County-Year FE	✓	✓	✓	✓	✓	✓
Institution-Year FE		✓	✓		✓	✓
# of Counties	1430	1430	1429	1430	1430	1429
# of ZIP codes	11,879	11,879	11,875	11,888	11,888	11,884
First-stage F				14.02	14.21	15.15
Clustered SE	County	County	County	County	County	County
Observations	1,645,814	1,645,814	1,488,541	1,647,041	1,647,041	1,488,985

Note: Loan-level estimates of the probability that a refinancing application is denied (one if denied, zero if originated). The sample covers applications to the lead banks of holding companies subject to Dodd-Frank Act stress tests (DFAST) over 2018–2024. Collateral uncertainty is measured at the ZIP-code level from ZTRAX. Columns 1–3 report OLS estimates; columns 4–6 report IV estimates that instrument Collateral Uncertainty (SFH) with local property characteristics, and report the Kleibergen-Paap first-stage F. Columns 2, 3, 5, and 6 add institution×year fixed effects; columns 3 and 6 keep CLL loans only. All columns include county×year fixed effects, borrower and loan controls, and DTI and LTV bins. Standard errors are clustered by county. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

D.3. Appraisal outcomes

This appendix reports the regressions underlying the binned scatter plots in Figure 3. We estimate appraisal-level regressions

$$Y_{ict} = \beta\, \sigma_{ct} + \alpha_t + \delta_{s(c)} + \varepsilon_{ict}, \tag{D.1}$$

where Y_{ict} is an outcome of appraisal i in county c and year t, σ_{ct} is county-level collateral uncertainty, α_t are year fixed effects, and $\delta_{s(c)}$ are state fixed effects, included in the richer specification. We use the appraisal weights from the public-use file, and we cluster standard errors at the county level. The data and outcomes are described in Appendix A.4.

Table D.4 reports the purchase-appraisal results. Higher collateral uncertainty is associated with a larger absolute gap between the appraised value and the contract price, and the coefficient is stable when state fixed effects are added. Among appraisals below the contract price, the shortfall is larger in higher-uncertainty counties. Table D.5 reports the refinance-appraisal results. Higher

collateral uncertainty is associated with fewer comparable properties and with comparables located farther from the subject property. The relation with the number of comparables weakens once state fixed effects are included but remains statistically significant, while the relation with comparable distance is stable.

Table D.4: *Collateral uncertainty and purchase appraisal outcomes (2013–2022)*

	Absolute Appraisal Gap (%)		Underappraisal Intensity Given Underappraisal (%)	
	(1) Year FE Only	(2) State + Year FE	(3) Year FE Only	(4) State + Year FE
Collateral Uncertainty	4.671***	4.686***	9.865***	7.548***
	(0.457)	(0.448)	(0.654)	(0.628)
Year FE	✓	✓	✓	✓
State FE		✓		✓
# of Counties	1217	1217	1067	1067
Clustered SE	County	County	County	County
Weights used	Yes	Yes	Yes	Yes
Observations	624969	624969	37950	37950
R2	0.002	0.003	0.037	0.048

Note: Estimates of Equation (D.1) for purchase appraisals in the FHFA Enterprise Uniform Appraisal Dataset Public Use File. Columns (1)–(2) use the absolute appraisal gap; columns (3)–(4) use under-appraisal intensity among appraisals below the contract price. Collateral uncertainty is the county-level single-family measure. Odd columns include year fixed effects; even columns add state fixed effects. Regressions use appraisal weights; standard errors are clustered by county. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

D.4. Foreclosure discounts

This appendix describes the county-level foreclosure discounts used in Figure 4. We use ZTRAX repeat-sales pairs over 2000–2017. Each observation is a pair of an initial transaction and a subsequent resale of the same property. For each pair we observe the county and ZIP code, the year and month of both sales, the first-sale price, the resale price, whether the first sale was a cash purchase, and distress indicators for the resale. We classify the resale as distressed if it is a foreclosure (REO) sale, a foreclosure (REO) transfer, or a short sale.

Let P_{i1} and P_{i2} denote the first- and second-sale prices of property i, and HP_i the holding period in years. To remove local house-price appreciation, we compute county appreciation over the holding period, $R_i^{loc} = \text{HPI}_{c,2}/\text{HPI}_{c,1}$, where $\text{HPI}_{c,1}$ and $\text{HPI}_{c,2}$ are the county hedonic house

Table D.5: *Collateral uncertainty and comparable selection in refinance appraisals (2013–2022)*

	Number of Comparables		Reported Distance of Comparables	
	(1) Year FE Only	(2) State + Year FE	(3) Year FE Only	(4) State + Year FE
Collateral Uncertainty	-2.761***	-0.404***	5.872***	6.017***
	(0.311)	(0.148)	(0.724)	(0.699)
Year FE	✓	✓	✓	✓
State FE		✓		✓
# of Counties	1218	1218	1218	1218
Clustered SE	County	County	County	County
Weights used	Yes	Yes	Yes	Yes
Observations	607775	607775	607131	607131
R2	0.054	0.126	0.064	0.117

Note: Estimates of Equation (D.1) for refinance appraisals in the FHFA Enterprise Uniform Appraisal Dataset (UAD) Public Use File, 2013–2022. The dependent variables are the number of comparable properties (columns 1–2) and their reported distance from the subject property (columns 3–4). Collateral Uncertainty is the county-by-year single-family measure from ZTRAX; values for 2018–2022 are set to the county's 2014–2017 average. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

price index at the two sale dates. The local-market-excess return is

$$lme_i = \frac{P_{i2}/R_i^{loc} - P_{i1}}{P_{i1}}, \tag{D.2}$$

and its holding-period-adjusted version is $\widetilde{lme}_i = \log(1 + lme_i)/\sqrt{HP_i}$. For each county c with at least 50 distressed repeat sales, we estimate

$$\widetilde{lme}_i = \beta_c \,\text{Distressed}_i + \gamma \log(P_{i1}) + \alpha_{y_1} + \alpha_{y_2} + \alpha_{m_1} + \alpha_{m_2} + \alpha_{zip} + \alpha_{cash} + \varepsilon_i, \tag{D.3}$$

where α_{y_1}, α_{y_2} and α_{m_1}, α_{m_2} are year and month fixed effects of the first and second sale, α_{zip} are ZIP-code fixed effects, and α_{cash} are fixed effects for a cash first sale. Standard errors are clustered at the ZIP-code level. The coefficient β_c is the average foreclosure-related return discount in county c; more negative values indicate larger discounts. Figure 4 relates β_c to the county average of collateral uncertainty.

E. Economic consequences: additional results

This appendix reports additional results for Section 4: retention among large banks (Appendix E.1), the construction and additional estimates of the prepayment hazard (Appendix E.2), denials by uncertainty decile (Appendix E.3), and the valuation of a missed refinancing (Appendix E.4).

E.1. Loan retention among large banks

Table E.1 re-estimates Equation (9) on loans originated by banks subject to Dodd-Frank Act stress testing (DFAST) over 2018–2024. Because these banks lend across many local markets, within-lender variation in collateral uncertainty is not constrained by geographic specialization. Table 11, reports the estimates for the largest bank and shows that the main results hold within the sample of DFAST banks.

Table E.1: *Collateral uncertainty and loan retention: large banks (2018–2024)*

Dep. var: Loan Retention Indicator (= 1 if loan retained on lender's balance sheet)						
	OLS			IV		
	(1) Across Lenders	(2) Within Lender	(3) Within Lender	(4) Across Lenders	(5) Within Lender	(6) Within Lender
Collateral Uncertainty	0.0678***	0.0423***	0.0418**	0.2301***	0.2054***	0.2022***
	(0.0174)	(0.0141)	(0.0172)	(0.0776)	(0.0648)	(0.0719)
DFAST Banks	✓	✓	✓	✓	✓	✓
Loan Controls	✓	✓	✓	✓	✓	✓
County-Year FE	✓	✓	✓	✓	✓	✓
Lender-Year FE		✓	✓		✓	✓
CLL Only			✓			✓
# of Counties	1244	1244	1244	1244	1244	1244
# of ZIP codes	12,065	12,065	12,065	12,062	12,062	12,062
First-stage F				18.26	18.25	20.54
Clustered SE	County	County	County	County	County	County
Observations	3,766,217	3,766,217	3,335,077	3,766,227	3,766,227	3,334,617

Note: Loan-level estimates of the probability that a loan originated in 2018–2024 by the lead bank of a DFAST stress-tested holding company is retained on balance sheet (not sold within the calendar year of origination). Collateral uncertainty is measured at the ZIP code and origination year. Columns 1–3 are OLS; columns 4–6 are 2SLS, instrumenting Collateral Uncertainty (SFH) with residualized ZTRAX property characteristics. All columns include county×year fixed effects and the loan controls and fixed effects of the main specification, including DTI and LTV bins; columns 2, 3, 5, and 6 add lender×year fixed effects, and columns 3 and 6 keep CLL loans only. Standard errors are clustered by county. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

E.2. Prepayment hazard: construction and additional estimates

We build a quarterly panel from the Freddie Mac monthly performance records of the fixed-rate purchase loans in Appendix A.1, for origination cohorts 2000–2024 (the table header of Table E.2). A loan enters the risk set at the start of quarter q if, at that date, it is current on its payments, has not been modified, has a positive unpaid balance and a valid note rate, and has not reached its contractual maturity. A loan that becomes delinquent leaves the risk set and re-enters once it is current again. Within the quarter, the loan is coded as prepaid if it is paid off (zero-balance code 01) before its contractual maturity. Payoffs at maturity, default-related terminations, and administrative exits are not counted as prepayments; such loans leave the risk set at their exit date.

The rate spread S is the loan's note rate at the start of the quarter minus the latest 30-year PMMS rate published before the quarter begins, in percentage points.

Aggregation: we group eligible loans into cells defined by 3-digit ZIP code, origination year, and 50-basis-point band of the original note rate. For each cell and quarter, the outcome h_{cq} is the number of prepayments during the quarter divided by the number of eligible loans at its start, and the regressors are cell averages of the spread and of the borrower and loan controls (including seller shares). Collateral uncertainty is measured in the observation year at the ZIP3 level. Regressions weight cells by their number of eligible loans.

Table E.2 reports the continuous counterpart of the decile specification in Section 4,

$$h_{cq} = \alpha_{\mathrm{CZ}(c),q,v(c)} + \beta S_{cq} + \theta \sigma_{cq} + \delta \left(\sigma_{cq} \times S_{cq}\right) + \overline{\mathbf{X}}_{cq}' \boldsymbol{\psi} + \varepsilon_{cq}, \tag{E.1}$$

where σ_{cq} is collateral uncertainty. The interaction coefficient δ is negative in the baseline (column 1): a wider spread raises prepayment less where collateral uncertainty is higher. The estimate remains negative when cells are weighted equally (column 2), when the sample is restricted to cells with a positive spread (column 3) or with a positive net present value of refinancing (column 4), and when collateral uncertainty is measured in the origination year rather than the observation year (column 5).

Deciles estimates: Figure E.1 reports the decile coefficients γ_d of the decile specification in

Table E.2: *Prepayment responses to rate incentives: continuous collateral uncertainty*

	(1) Loan weights	(2) Equal weights	(3) Origination-year risk
Dependent variable:	Quarterly prepayment rate		
Collateral Uncertainty	-0.0125^{***}	-0.0019	-0.0088^{**}
	(0.0043)	(0.0044)	(0.0036)
Rate spread	0.0211^{***}	0.0110^{*}	0.0285^{***}
	(0.0064)	(0.0060)	(0.0057)
Collateral Uncertainty × Rate spread	-0.0207^{***}	-0.0166^{***}	-0.0157^{***}
	(0.0041)	(0.0030)	(0.0034)
Cell weights	Loans	Equal	Loans
Risk timing	Observation year	Observation year	Origination year
CZ × quarter × origination-year FE	Yes	Yes	Yes
Loan controls and seller shares	Yes	Yes	Yes
Observations (loan-quarters)	131,618,233	131,618,233	126,304,399

Note: Grouped linear probability models of Equation (E.1), estimated on cells (3-digit ZIP code × origination year × 50-basis-point note-rate band) by quarter, built from purchase loans originated between 1999 and 2024 in the Freddie Mac Single-Family Loan-Level Dataset. The rate spread is the note rate minus the 30-year PMMS rate, in percentage points. Collateral Uncertainty is measured in the observation year (column 3: origination year). Standard errors are two-way clustered by commuting zone and quarter. *** p<0.01, ** p<0.05, * p<0.1.

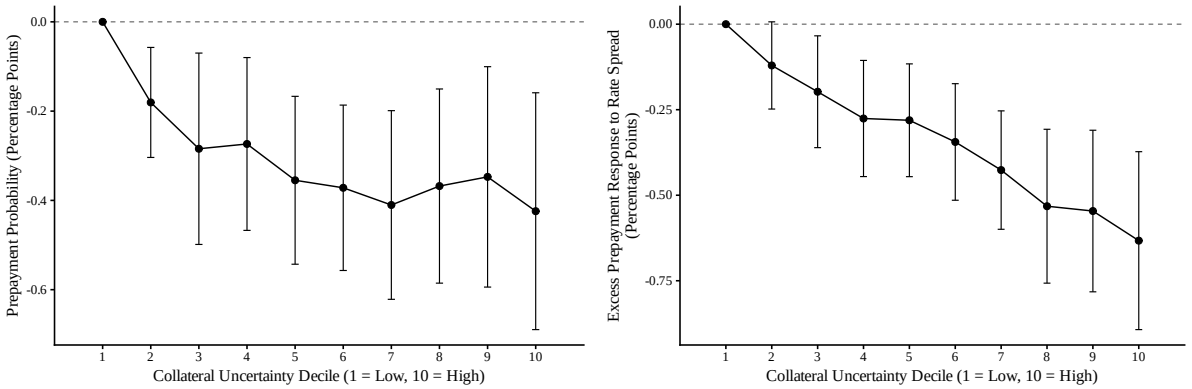
Section 4, that is, the difference in the quarterly prepayment probability relative to decile 1 at a zero spread. Both panels decline with collateral uncertainty: γ_d falls from about -0.18 to -0.42 percentage points between decile 2 and decile 10, while the response to a one-percentage-point rate spread in Panel (b) falls steadily from 1.77 percentage points in decile 1 to 1.14 in decile 10, so higher collateral uncertainty is associated with both lower prepayment and a markedly dampened response to refinancing incentives.

E.3. Refinancing denials by collateral uncertainty decile

Figure E.2 reports the decile coefficients β_d used in Section 4.4. They come from Equation (7) with the continuous measure replaced by indicators for within-year deciles of ZIP-code collateral uncertainty, with decile 1 as the reference. Panel (a) shows that denial probabilities rise monotonically across deciles, reaching 3.5 percentage points in the top decile against a 22.5 percent denial rate in the bottom decile. Panel (b) shows the share of each decile's denials attributable to collateral uncertainty when deciles are instead defined on the pooled 2007–2024 distribution; the implied total is 573,000 denials, or 4.25 percent of all refinance denials, compared with 578,000

Figure E.1: *Prepayment response to rate incentives (2000-2024)*

(a) *Unconditional effect (γ_d)* **(b)** *Conditional effect (δ_d)*

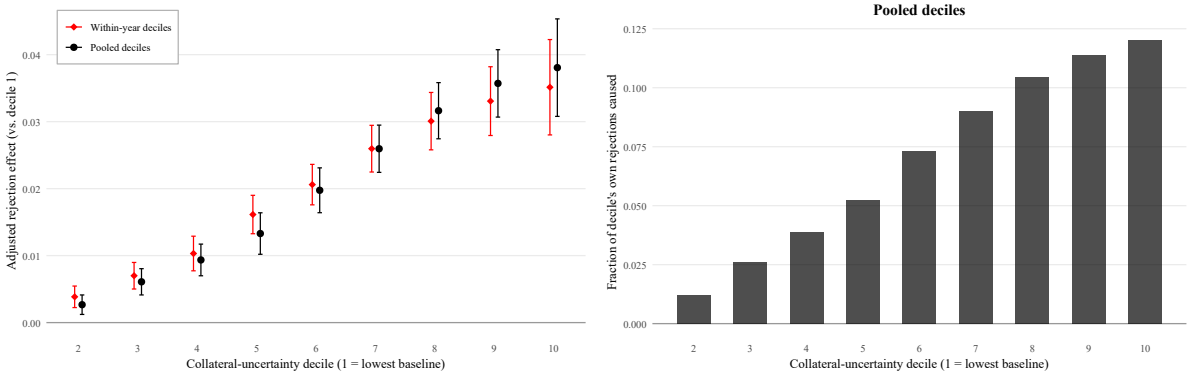


Note: Panel (a) shows the difference in the quarterly prepayment probability at a zero rate spread (β_d). Panel (b) shows the difference in its response to a one-percentage-point rate spread (δ_d), defined as the loan's current rate minus the latest 30-year Primary Mortgage Market Survey (PMMS) rate. Prepayment is any early payoff. The sample is fixed-rate purchase loans from the Freddie Mac Single-Family Loan-Level Dataset (SFLLD) observed between 2000 and 2024, with loan and seller controls and commuting zone by quarter by origination year fixed effects. Collateral Uncertainty is measured at the ZIP3-year level from the Zillow Transaction and Assessment Dataset (ZTRAX). Intervals are 95%, with standard errors clustered by commuting zone and quarter.

(4.3 percent) under within-year deciles.

Figure E.2: *Refinancing denials by collateral uncertainty decile (2007–2024)*

(a) *Excess denial probability, β_d* **(b)** *Share of denials attributable, pooled deciles*



Note: Panel (a) plots the coefficients β_d on deciles 2 to 10 of within-year ZIP-code collateral uncertainty (decile 1 omitted) from a linear probability model of denial for HMDA refinancing applications, 2007–2024, with the controls and county×year and lender×year fixed effects of Equation (7); bars are 95% confidence intervals clustered by county. Panel (b) plots, for each decile of the pooled 2007–2024 distribution, the share of its denials attributable to collateral uncertainty, $\beta_d/\bar{r}_d$. Collateral uncertainty is measured from ZTRAX.

E.4. The value of a missed refinancing

This appendix describes the computation of V_d, the present value of a missed refinancing in decile d, used in Section 4.4.

We take the Freddie Mac loans active in January 2011 from the fixed-rate purchase sample in Appendix A.1: loans that are current, have a positive unpaid balance, and have not reached maturity. Loans whose commuting zone by FICO-bin cell has no match in the 2011Q1 purchase-origination rate schedule used to construct the counterfactual refinancing rate are also excluded from the sample. We also drop loans that are modified and require a non-missing, valid zero-balance-code/interest rate and a valid 3 digit ZIP code. We assign each loan to a decile of 2011 ZIP3 collateral uncertainty.

Payment reduction: for loan i with unpaid balance B_i, note rate r_i, and N_i remaining months, the current monthly payment is

$$PMT_i^{old} = B_i \frac{r_i/12}{1-(1+r_i/12)^{-N_i}}. \tag{E.2}$$

The counterfactual payment after refinancing the remaining balance at the January 2011 PMMS 30-year rate, r^{refi}, holding the remaining maturity fixed, is

$$PMT_i^{refi} = B_i \frac{r^{refi}/12}{1-(1+r^{refi}/12)^{-N_i}}, \tag{E.3}$$

and the monthly payment reduction is $\Delta PMT_i = PMT_i^{old} - PMT_i^{refi}$. We set r^{refi} equal to 4.76 percent (PMMS January 2011 average).

Net present value: We discount the payment reduction over the remaining term at the January 2011 average 10-year Treasury yield, $\rho = 3.39$ percent per year ($\rho/12$ per month), and subtract a refinancing cost equal to 2 percent of the balance, following Keys et al. (2016):

$$PDV_i = \Delta PMT_i \frac{1-(1+\rho/12)^{-N_i}}{\rho/12} - 0.02\, B_i. \tag{E.4}$$

Value by decile. A loan has a profitable refinancing opportunity if $PDV_i > 0$. The value of a missed refinancing in decile d is the average net present value among these loans,

$$V_d = \frac{1}{|\mathcal{I}_d|} \sum_{i \in \mathcal{I}_d} PDV_i, \qquad \mathcal{I}_d = \{i \in d : PDV_i > 0\}. \tag{E.5}$$

Foregone savings: Foregone savings in decile d in 2011 are $\beta_d N_d V_d$, where β_d is the excess denial probability from Appendix E.3 and N_d is the number of 2011 HMDA refinancing applications in within-year decile d. The HMDA deciles (ZIP5, within-year) and the Freddie Mac deciles (ZIP3, 2011) are different partitions of the uncertainty distribution, which we match by decile number. The calculation assumes that every excess denial is a refinancing that does not take place and that would otherwise have lasted until maturity, and it applies the 2007–2024 coefficients β_d to 2011. Section 4.4 discusses the direction of these assumptions.