Barriers to Reentry: Initial Borrowing Frictions, Refinancing, and Wealth Redistribution*

Heejin Yoon†

July 27, 2025

Latest Version

Abstract

This paper examines how frictions encountered during the *initial* purchase mortgage origination process shape borrowers' future refinancing behavior and contribute to wealth disparities. Leveraging variation in loan officer *workload* as a quasi-random source of lender-induced origination delays, I find that experiencing a 60+ day delay lowers quarterly refinancing rates by 16–24%. Minority borrowers, low-income households, and those with lower credit scores are more likely to encounter such frictions, with evidence pointing to lender bias as a potential driver of racial disparities. A simulation based on a quantitative model of refinancing behavior implies a present value loss of $6,641 per delayed borrower, which amounts to $2.8 billion in overpayments each year when scaled to the U.S. market. Importantly, these losses are not evenly distributed: conditional on the same delay event, minority borrowers incur greater financial losses than White borrowers, largely due to a lower baseline likelihood of acting on refinancing opportunities. Together, these findings demonstrate how subtle frictions in the origination process can lead to persistent financial disadvantages and entrench wealth inequality.

JEL codes: G21, G23, R30, R51

Keywords: Mortgage; Refinancing; Origination Frictions; Time-To-Close; Processing Delays

*First version: January 14, 2025. This version: July 27, 2025. A previous version of this paper was circulated under the title "Understanding Racial Disparities in Mortgage Refinancing." For helpful comments, suggestions, and discussions, I thank Gene Amromin, Allen Berger, Neil Bhutta, Anthony DeFusco, Yongheng Deng, Kris Gerardi, Congyan Han, Lu Han, Tim Hung (*discussant*), Poorya Kabir (*discussant*), You Suk Kim, Eyal Lahav (*discussant*), Jack Liebersohn (*discussant*), Erik Mayer, Philip Mulder (*discussant*), and Chris Timmins. I also thank participants at ABFER Poster Session (2025), AsianFA Annual Conference (2025), BFWG International Conference (2025), FIRS (2025), USC Marshall PhD Conference in Finance (2025), and WSB Summer Research Conference (2025). Financial support from the Institute for Humane Studies is gratefully acknowledged.

†Wisconsin School of Business, University of Wisconsin–Madison (e-mail: heejin.yoon@wisc.edu).

1. Introduction

Mortgage refinancing is a key decision in household finance, allowing borrowers to reduce interest costs and restructure debt. A growing body of research highlights its central role in both household wealth accumulation (Goodman and Mayer, 2018; Killewald et al., 2017) and the transmission of monetary policy (Beraja et al., 2019; Berger et al., 2021; Eichenbaum et al., 2022; Greenwald, 2018). Despite these benefits, both the likelihood and speed with which borrowers capitalize on refinancing opportunities vary widely, particularly along lines of race, income, age, and education (Andersen et al., 2020; Defusco and Mondragon, 2020; Deng and Gabriel, 2006; Firestone et al., 2007; Gerardi et al., 2023).

Much of the literature has attributed this puzzle—why borrowers respond so differently to refinancing opportunities—to borrower-side characteristics, such as behavioral biases and financial illiteracy (Agarwal et al., 2016; Keys et al., 2016). However, recent work has shifted attention to the supply side, highlighting that lenders also play an important role in shaping refinancing outcomes. For instance, borrower responsiveness is influenced by media exposure and lender advertising strategies (Grundl and Kim, 2019; Hu et al., 2024). Lender-side operational bottlenecks and labor market frictions can also constrain credit supply, disproportionately limiting refinancing access for marginal borrowers (Frazier and Goodstein, 2023; Fuster et al., 2024). These insights underscore the importance of supply-side factors, which remain relatively underexplored in the literature on refinancing heterogeneity.

Building on this perspective, I study a previously overlooked determinant of refinancing outcomes: frictions experienced during the *initial* mortgage origination process, which often stem from supply-side factors. Borrowers' *initial* mortgage experiences vary widely: while many navigate the process smoothly, others face significant challenges, often due to lender-side issues such as processing delays, excessive documentation requests, or unresponsive service. These frictions, though frequently overlooked, can shape borrowers' perceptions of the mortgage process in negative and lasting ways.[1] A large body of behavioral finance research shows that early personal experiences can significantly influence financial decisions, even among sophisticated individuals such as fund managers and corporate executives (Carvalho et al., 2023; Chernenko and Sunderam, 2016; Dittmar and Duchin, 2016; Gao et al., 2024; Koudijs and Voth, 2016; Malmendier et al., 2011, 2021). Drawing on these insights, I hypothesize that negative interactions with lenders at the origination stage can deter borrowers from future refinancing.

[1]This perspective is supported by Fannie Mae (2014), which documents that "*borrowers' perceived ease of obtaining a mortgage significantly influences their future intent to refinance.*"

To capture frictions in the *initial* borrowing process, I use *Time-To-Close*—the number of days taken to secure a mortgage—as a proxy. An extended period of *Time-To-Close* serves as a strong indicator of borrowing frictions and offers a useful lens for assessing their impact on *subsequent* refinancing behavior for several key reasons. First, as shown in Figure 1, delays in loan processing and closing are a major source of consumer dissatisfaction, accounting for 18–36% of mortgage-related complaints in the Consumer Financial Protection Bureau (CFPB) database, with substantial variation across borrowers.[2] Second, *Time-To-Close* is objectively measurable, providing a more reliable proxy than subjective aspects of service quality, such as borrower satisfaction ratings or perceived lender responsiveness.[3] Third, loan-level *Time-To-Close* values can be linked to future refinancing outcomes in my loan panel dataset, enabling a direct examination of how early borrowing frictions influence subsequent financial decisions.

While *Time-To-Close* serves as a reasonable proxy for *initial* borrowing frictions, two key identification challenges must be addressed. First, *omitted variable bias* may arise because longer mortgage processing times may correlate with unobserved borrower characteristics. For example, borrowers with lower financial literacy might take longer to complete applications due to difficulties navigating the mortgage process. If these unobserved traits also influence refinancing behavior, failing to control for them could bias the estimates. Second, *measurement error* poses another important challenge, as delays in *Time-To-Close* may conflate lender-side frictions with borrower- or seller-driven factors such as moving schedules or contract contingencies. Since these alternative sources of delay are unlikely to affect refinancing behaviors, inclusion of these non-lender-related components introduces noise, potentially attenuating the OLS estimates.

To overcome the empirical challenges, I employ an instrumental variable (IV) strategy that leverages time-varying capacity constraints (*Workload*) at the loan officer level as an exogenous source of variation in loan origination delays. The idea is straightforward: when a loan officer is handling a heavier pipeline of active (i.e., incomplete) applications at the time a borrower applies, the likelihood of a processing delay rises due to operational bottlenecks. Because borrowers cannot easily observe or influence officer workloads at the time of application, fluctuations in workload offer quasi-random variation in loan processing times. In addition, by instrumenting loan origination delays with officer workload, I isolate the compo-

[2] See panel (a) of Figure 2.

[3] Specifically, I focus on cases where *Time-To-Close* exceeds 60 days, as such delays are likely to be both salient and financially burdensome for borrowers. Loan origination delays longer than 60 days often surpass typical rate-lock periods of 30 to 60 days (for typical lock durations, see https://www.consumerfinance.gov/ask-cfpb/whats-a-lock-in-or-a-rate-lock-en-143/). When loan processing extends beyond the rate-lock period, borrowers face heightened uncertainty, including the risk of rate changes, additional lock-in fees, or even failure to close the transaction on time (Han and Hong, 2024).

nent of delay most likely to generate borrower dissatisfaction and discourage future refinancing. This approach helps mitigate endogeneity concerns arising from both omitted variable bias and measurement error.[4]

The IV estimates using the quarterly loan panel of the CoreLogic–MBS dataset show that experiencing a 60+ day origination delay lowers quarterly refinancing rates by 0.48 to 0.73 percentage points, representing a 15.8% to 24.2% decline relative to the mean refinancing rate of 3.02%. These estimates are substantially larger in magnitude than their OLS counterparts, which show a decline of 0.10 to 0.15 percentage points, underscoring the importance of addressing endogeneity issues.

To shed light on the mechanism behind this effect, I distinguish between two competing channels. One possibility is that delays erode trust in the original lender, deterring borrowers from refinancing with the same institution (*lender-specific deterrence*). Alternatively, delays may raise borrowers' perceived refinancing costs more broadly, suppressing refinancing with any lender (*generalized discouragement*). Testing these channels separately, I find that delays sharply reduce *recapture refinancing*—refinancing with the original lender—while the effect on *switching refinancing* to a new lender is small and statistically insignificant. This pattern supports the lender-specific deterrence channel, suggesting that the main effect of initial frictions operates primarily through the disruption of borrower–lender relationships.

Additional validation exercises and robustness checks reinforce the interpretation of the baseline estimates. First, the discouragement effect (i) intensifies with longer delays, (ii) fades as loans season—potentially reflecting the dilution of negative borrower memory over time, and (iii) selectively affects refinancing-related transactions (e.g., cash-out refinancing), while leaving prepayments associated with household moves unaffected. Second, the results are robust to alternative definitions of refinancing incentives based on the closed-form threshold from Agarwal et al. (2013) and to sample restrictions that rule out borrower-side constraints on refinancing eligibility. Finally, I validate my main findings using data from the National Survey of Mortgage Originations (NSMO), showing that borrower-reported delay experiences are significantly associated with reduced prepayment activity and heightened dissatisfaction across multiple dimensions of the mortgage process. Taken together, these patterns are consistent with the view that *initial* borrowing frictions create persistent barriers to borrower–lender re-engagement.

Having established that mortgage delays discourage future refinancing, I next examine which borrowers are most exposed to these origination frictions. I find that minority borrowers, as well as those

[4]Further discussion of the instrument's validity is provided in Section 3.3.1.

with lower incomes and lower credit scores, are significantly more likely to experience delays in mortgage origination, even after controlling for detailed borrower, loan, and lender characteristics.[5] For instance, even in the most stringent specification, minority borrowers are 1.84 percentage points more likely to encounter a 60+ day delay—a 18.6% increase relative to the baseline delay rate of 9.9%. Importantly, the evidence shows that racial gaps are larger in areas with heightened racial animus and weaker lending market competition, suggesting lender-side bias may be a contributing factor to the observed racial disparities.

As a final step, I quantify the financial burden arising from delays in *initial* mortgage origination. To do so, I use both a back-of-the-envelope calculation and a simulation-based quantification. The back-of-the-envelope calculation suggests that delayed borrowers incur an average overpayment of $586.5 per year due to missed refinancing opportunities. This estimate is calculated as the product of the average interest rate reduction from refinancing (87 basis points), the average loan amount ($279,288), and the estimated 24% decline in refinancing probability attributable to delays. Assuming that the refinancing opportunity occurs halfway through a 30-year loan term, this annual loss translates into a present value (PV) of $4,511.

While the back-of-the-envelope approach provides an intuitive benchmark, it rests on a simplifying assumption that borrowers face a single refinancing opportunity at a fixed point (e.g., year 15) in time, and it does not account for the *inertia* in refinancing behavior commonly observed even among non-delayed borrowers (Andersen et al., 2020). To address these limitations, I use a quantitative model of refinancing behavior to simulate plausible coupon rate trajectories for both delayed and non-delayed borrowers. The model incorporates the stochastic evolution of market mortgage rates and a borrower *responsiveness* to refinancing opportunities, with *responsiveness* for delayed borrowers scaled down by 24% based on reduced-form (2SLS) estimates. Here, *responsiveness* governs the likelihood that a borrower takes up a favorable refinancing opportunity during a given time period. Simulations from the quantitative model imply an average PV loss of $6,641 per delayed borrower, exceeding the back-of-the-envelope estimate. When scaled to the roughly 426,517 borrowers who experience delays annually, this translates into approximately $2.8 billion in excess mortgage payments per year.

Importantly, these losses are not evenly distributed: conditional on experiencing the same delay event,

[5] My findings are consistent with those in Wei and Zhao (2022) on racial disparities in loan processing times during the early 2000s.

borrowers from underserved groups face greater financial losses due to lower baseline *responsiveness*. For instance, Black borrowers lose $7,531 on average, more than $1,000 greater than the average loss for White borrowers ($6,504). Coupled with the fact that minority borrowers are more likely to experience origination delays in the first place, these findings illustrate how subtle frictions in the origination process create persistent financial disadvantages and serve as a channel of wealth redistribution from marginalized groups to others in the mortgage market.

Related Literature. This paper contributes to three key strands of literature. First, it adds to research on heterogeneous refinancing behavior and its distributional consequences. While much of the existing literature emphasizes demand-side explanations—such as behavioral biases, financial illiteracy, or liquidity constraints that hinder optimal refinancing (Clapp et al., 2001; Deng and Gabriel, 2006; Firestone et al., 2007; Gerardi et al., 2023; Keys et al., 2016; Andersen et al., 2020; Defusco and Mondragon, 2020)—this paper highlights the lasting effects of lender-side frictions in the *initial* mortgage process, specifically delays in mortgage originations. In doing so, it complements recent work on how supply-side factors—such as advertising strategies (Grundl and Kim, 2019), media exposure (Hu et al., 2024), and capacity constraints (Frazier and Goodstein, 2023; Huang et al., 2024)—shape refinancing outcomes. It also relates to emerging evidence that slower refinancers effectively subsidize faster ones, reinforcing inequality across borrower groups (Berger et al., 2024; Fisher et al., 2024; Zhang, 2024). By documenting both differential exposure to origination frictions and unequal financial losses from the same delay events, this paper highlights how lender-side inefficiencies can amplify racial and socioeconomic disparities in mortgage markets.

Second, this paper contributes to a growing consensus in research that small delays in financial and administrative processes can lead to large consequences. Recent studies show that modest disruptions—such as a 15-day lag in vendor payments (Barrot and Nanda, 2020), delayed patent approvals (Farre-Mensa et al., 2020), or a one-day increase in shipping time (Djankov et al., 2010)—can significantly reduce employment, sales, or trade flows. In consumer finance, Fuster et al. (2019) find that faster mortgage processing by FinTech lenders enables them to capture market share from slower traditional banks, while Doniger and Kay (2023) show that a 10-day delay in Paycheck Protection Program funding during the COVID-19 pandemic led to the loss of over two million jobs. In line with this literature, my findings show that delays in mortgage origination—though seemingly minor—can impose lasting financial costs by distorting household refinancing decisions.

Third, it builds on the growing body of research examining how past personal experiences shape financial decision-making. Prior studies document that even sophisticated financial professionals—such as fund managers (Chernenko and Sunderam, 2016), bank branch managers (Carvalho et al., 2023; Gao et al., 2024), firm executives (Dittmar and Duchin, 2016; Koudijs and Voth, 2016), and even central bankers (Malmendier et al., 2021)—form lasting financial beliefs based on their past experiences. My study contributes to this literature by demonstrating that borrowers' prior experiences with mortgage borrowing, particularly exposure to loan origination delays, influence their willingness to refinance in the future. This suggests that past interactions with lenders shape financial behavior in ways that have long-term implications for household wealth accumulation.

Outline. The remainder of the paper is organized as follows. Section 2 describes the data and key variables of interest, along with summary statistics. In Section 3, I empirically test the effect of experiencing initial mortgage delays on refinancing outcomes. Section 4 examines the heterogeneous exposure to borrowing frictions. Section 5 estimates the aggregate financial impact. Section 6 concludes.

2. Data

This study integrates CoreLogic (deeds and MLS) with the MBS Loan-Level Dataset provided by Fannie Mae, Freddie Mac, and Ginnie Mae for the empirical analysis. By matching the two primary datasets, I construct the quarterly loan panel originated between 2014–2021, which contains multiple observations for each mortgage until it terminates. Details of each dataset and the matching procedure are outlined below.

2.1. CoreLogic

I utilize two separate sources of information from CoreLogic for 18 U.S. states: (i) deeds and (ii) MLS datasets. The deeds data contain comprehensive information on all deed transfers in the U.S., including sale amounts, property types, and property addresses, acquired directly from county clerk and recorder offices. The deeds data also provide detailed information of mortgages recorded as liens on properties, such as mortgage amounts, lenders, conventional/FHA loan status, loan origination dates, and borrowers' first and last names. The MLS data contain information on property listings, including listing prices, listing dates, and the dates when sale contracts are signed and closed. I merge the deeds and MLS data using CoreLogic's unique parcel identification numbers and sale closing dates.

The analysis focuses on 30-year fixed-rate mortgages—the most common mortgage product in the U.S.—for single-family home purchases originated between 2014 and 2021. My sample is restricted to 18 U.S. states, where both deed and MLS data are consistently available and can be reliably matched, allowing me to construct a measure of *Time-To-Close*—the number of days taken to secure a mortgage.[6] Additionally, I exclude loans made to institutional buyers and those with unconventional features, such as interest-only payment structures, negative amortization loans, or contracts with teaser rates.

I limit the analysis period to 2014–2021 for several reasons. First, the loan officer NMLS ID information, crucial for my IV strategy discussed in Section 3.1, became consistently available in the CoreLogic dataset starting in 2014. Although full compliance of the Secure and Fair Enforcement for Mortgage Licensing (SAFE) Act[7] was mandated by 2011, consistent reporting of NMLS ID fields in CoreLogic did not begin until 2014. Second, the Ginnie Mae MBS Loan-Level Disclosure data, covering detailed information on FHA and other government-insured loans, has been publicly available since 2013. Finally, focusing on the period after 2013 helps avoid the complexities of the immediate post-crisis years (2009–2013), a period characterized by temporary policy interventions and regulatory reforms that could potentially influence refinancing behavior and confound the analysis.[8] Thus, starting the analysis in 2014 ensures reliable loan officer identification, comprehensive loan-level data coverage, and a focus on refinancing behavior under stabilized post-crisis market and regulatory conditions.

2.2. Fannie Mae/Freddie Mac/Ginnie Mae MBS Loan-Level Dataset

In addition to CoreLogic, I use datasets that provide detailed information on loans packaged into MBS and sold by Fannie Mae, Freddie Mac, or Ginnie Mae[9] from 2014 to 2021. The loan-level information of the MBS datasets includes loan amount, origination date, maturity, interest rate, credit score, loan-to-value

[6] Appendix A.1 details the rationale for selecting a subset of states, outlines the selection criteria, and lists the 18 selected states.

[7] The requirement for loan officers to obtain a unique NMLS identifier was introduced by the SAFE Act of 2008 and later reinforced through the Dodd-Frank Act of 2010.

[8] For example, federal programs such as the Home Affordable Refinance Program (HARP) and the Home Affordable Modification Program (HAMP), launched in response to the 2008 financial crisis, significantly altered refinancing incentives and borrower behavior during this period. See Agarwal et al. (2017, 2023) for discussions on these programs' impacts.

[9] In analyzing Ginnie Mae loans, I restrict the sample to FHA-insured mortgages, which represent the largest share of Ginnie Mae securitizations during the sample period.

(LTV) ratio, and debt-to-income (DTI) ratio, and property location. For each loan, I observe the monthly credit events such as prepayment, 90+ day delinquency, and foreclosure, until the loan is fully paid off. Thus, linking loan samples into MBS loan-level datasets allows me to incorporate several key variables essential to understanding refinancing behavior but absent from CoreLogic.[10]

Since there is no unique identifier for merging CoreLogic and the MBS datasets, I conduct matching based on loan characteristics. Specifically, after filtering both datasets to include only fixed-rate, 30-year purchase mortgages for single-family homes, I match loan records using: origination date, property location (3-digit ZIP code, CBSA code, and state), loan amount, occupancy status, and conventional/FHA loan indicator.[11] To ensure the matching accuracy, I remove duplicate observations and perform the matching without replacement. This process produces a quarterly loan performance panel with 5,883,962 observations.

To evaluate the representativeness of the matched dataset, Figure 3 compares key credit-related variables in the full CoreLogic sample with those in the matched sample from the 2015 snapshot. Panel (a) presents the combined GSE and FHA loan sample, while panels (b) and (c) separately show the GSE and FHA loan subsamples, respectively. The kernel densities are constructed using the actual origination volumes of GSE and FHA loans in each state as weights, accounting for variations in matching performance across states and loan types. Across all panels, the figure confirms that the variable distributions in the matched dataset closely resemble those in the population of loans. An exception may be the distribution of LTV ratios in the GSE sample in panel (b): loans with LTVs between 85% and 100% are somewhat overrepresented in the matched dataset, while those with LTVs below 75% are underrepresented. This slight imbalance arises because low-LTV loans often produce multiple potential matches, and the matching procedure discards such duplicates to prioritize accuracy. Nonetheless, the overall representativeness of the matched sample remains strong.

[10]For instance, credit score and interest rate variables are particularly critical. Borrowers with higher credit scores tend to refinance more frequently, and minority borrowers generally have lower credit scores than White borrowers (Gerardi et al., 2023). Moreover, Berger et al. (2021) highlight that refinancing decisions are strongly influenced by the *rate gap*, the difference between the original mortgage rate and the prevailing market rate for similar mortgages.

[11]The matching algorithm differs across dataset providers—Fannie Mae, Freddie Mac, and Ginnie Mae—due to variations in the available variables used for matching. For example, the Ginnie Mae MBS Loan-Level Disclosure dataset includes the exact origination date, whereas the Fannie Mae Single-Family Loan Performance Data and the Freddie Mac Single-Family Loan-Level Dataset provide only the origination year and month.

2.3. Supplementary Datasets

In addition to the CoreLogic–MBS matched dataset, I utilize InfoUSA and the National Survey of Mortgage Originations (NSMO) data to provide richer context and to conduct robustness checks for the empirical analysis.

InfoUSA. InfoUSA is a consumer database encompassing 120 million households and 292 million individuals. It is constructed from 29 billion records sourced from over 100 contributors, including census data, billing statements, telephone directories, and mail-order buyer or magazine subscription information. It provides exact home addresses alongside detailed household characteristics, such as the estimated age of the household head, family size, and the number of children. By linking InfoUSA to the CoreLogic–MBS dataset, additional borrower characteristics, e.g., fixed effects of the borrower age group, are incorporated into the regression analysis, enhancing control over borrower heterogeneity not captured in mortgage datasets alone.[12]

National Survey of Mortgage Originations (NSMO). NSMO is a mail-based quarterly survey jointly administered by the Federal Housing Finance Agency (FHFA) and the CFPB since 2014. It provides detailed information on borrowers' mortgage experiences, including delays during the origination process and satisfaction with each step of that process. The dataset also includes a rich set of borrower demographics (e.g., race, sex, income, presence of a co-borrower) and mortgage characteristics (e.g., loan type, loan amount category, loan purpose, initial FICO score, and LTV ratio), as well as quarterly updated credit information (e.g., FICO score, LTV ratio, and loan performance status). I use NSMO as an alternative dataset to externally validate key findings from the CoreLogic–MBS dataset.

2.4. Measuring *Time-To-Close* and *Rate Gap*

Time-To-Close. I measure initial mortgage borrowing frictions using *Time-To-Close*, defined as the number of days between the sale contract date (from CoreLogic MLS data) and the mortgage origination date (from CoreLogic deeds records). *Time-To-Close* closely corresponds to the "loan processing time" widely used in the mortgage literature (Choi et al., 2022; Fuster et al., 2019, 2024; Wei and Zhao, 2022), with the only distinction being the starting point: while conventional processing time is measured from the

[12]Since both CoreLogic and InfoUSA contain exact address information, I achieve almost a complete match for all observations.

loan application date, *Time-To-Close* begins at the sale contract date. In practice, this difference is minor because lenders typically require a signed purchase agreement before processing an application.[13]

The reliability of *Time-To-Close* is supported by several validations. Panel (b) of Figure 2 compares the median *Time-To-Close* in my data to the median loan processing time reported in Figure A.8 of Fuster et al. (2024), showing nearly identical time-series patterns. In addition, Appendix A.2 replicates racial disparities in average *Time-To-Close* for an earlier period (2001–2006), closely matching the patterns reported for loan processing time in panel (a) of Figure 3 in Wei and Zhao (2022). Together, these comparisons provide strong evidence that *Time-To-Close* is a reliable and valid measure of initial mortgage processing frictions for this study.

Rate Gap. To isolate the effect of initial loan delays on refinancing behavior, it is crucial to account for borrowers' refinancing incentives at each point in time. Specifically, I control for refinance incentive driven by fluctuations in the rate environment that could otherwise confound observed refinancing decisions. Following Berger et al. (2021), I measure *Rate Gap* as the difference between a loan's outstanding coupon rate (c_i) and the rate available for comparable mortgages at time t ($m_{i,t}$):

$$Rate\ Gap = c_i - m_{i,t}, \tag{1}$$

where the current available market rate ($m_{i,t}$) derived from the monthly average 30-year fixed-rate mortgage rate reported in the Freddie Mac Primary Mortgage Market Survey (PMMS). This rate is further adjusted by a loan-specific factor, modeled as a second-order polynomial function of the borrower's FICO score and the loan's quarterly updated LTV ratio.

Consistent with Berger et al. (2021) and Scharlemann and van Straelen (2024), I find that refinancing probabilities exhibit a distinct “step-like" nonlinear pattern across the distribution of *Rate Gap* values, as illustrated in Figure B1.

2.5. Summary Statistics

Table 1 presents summary statistics for the full sample of GSE and FHA loans, constructed from the matched CoreLogic–MBS dataset.

[13] Conversations with mortgage professionals confirm that loan applications are usually submitted on or immediately after the sale contract date.

Panel (a) provides descriptive statistics for the quarterly loan panel, where each loan appears multiple times over time. Detailed prepayment outcomes—including refinancing, cash-out refinancing, and prepayment due to selling or moving—are constructed using a matching algorithm developed for this study.[14] The average quarterly refinancing rate is 3.02%, with a standard deviation of 17.10%. The refinancing dummy is further classified into two types: *Recapture Refinance* and *Switching Refinance*. *Recapture Refinance* refers to borrowers refinancing with their original lender, while *Switching Refinance* captures refinancing with a different lender.[15] The quarterly mean values of *Recapture Refinance* and *Switching Refinance* are 0.95% and 2.07%, respectively, implying that 31.5% ($\frac{0.95}{3.02}$) of borrowers refinance with their original lender, while 68.5% switch lenders when refinancing.

For other prepayment types, *Cash-Out Refinance* occurs at an average quarterly rate of 1.19%, showing a similar recapture-to-switching ratio (29.4% recapture vs. 70.6% switching). *Prepayment Due to Selling and Moving* averages 1.47% per quarter, consistent with the monthly moving shock probability of 0.5% reported by Berger et al. (2021) and the annual moving rate of 7.52% in Fonseca and Liu (2024).

I define $\mathbb{1}(\textit{Time-To-Close} > 60 \text{ Days})$ as a dummy equal to one if *Time-To-Close* exceeds 60 days, which serves as the primary measure of frictional experiences in initial mortgage origination. As shown in Table 1, 11% of mortgages experienced such delays. The table also summarizes borrower and loan characteristics: 69% of borrowers are identified as White, 27% as minorities (7% Black and 20% Hispanic), and 4% as Asian.[16] Additionally, 33% of borrowers are female, and 48% have a co-borrower. First-time home buyers account for 54% of the sample, and 62% of loans are FHA-insured. The log of estimated monthly income—derived from loan amount, mortgage rate, and DTI ratio[17]—has a mean of 8.11, equivalent to $3,328. The mean log loan amount is 12.47, corresponding to $260,407. The average LTV at origination is 87.6%, and the average FICO score is 730.6.

Table 1 also documents time-varying loan characteristics. The *Current LTV*—calculated as the outstanding balance divided by the property's current market value (using ZIP code-level Zillow Home Value Index)—averages 73.1%. The *Rate Gap*, defined as the difference between the mortgage's coupon rate and the current market rate for comparable loans, averages −0.07 percentage points. Lastly, *Workload*

[14] The algorithm links loan records to subsequent mortgage originations and property transactions to classify prepayment types. Further details on the construction of these outcomes are provided in Appendix A.3.

[15] This classification is based on whether the refinancing mortgage in CoreLogic was originated by the same lender as the initial loan.

[16] Race and ethnicity are imputed using the Bayesian Improved First Name Surname Geocoding (BIFSG) method (Voicu, 2018), based on borrower names and location. Details are provided in Appendix A.4.

[17] Calculated as $\frac{\textit{Monthly Payment}}{\textit{DTI Ratio}} \times 100$, where $\textit{Monthly Payment} = \frac{\textit{Loan Amount} \times r/12 \times (1+r/12)^{360}}{(1+r/12)^{360}-1}$.

measures the number of active applications that were being managed by the loan officer at the time of each loan application. The median officer handles three concurrent applications, while those in the 75th percentile manage seven applications.

Panel (b) reports summary statistics for the cross-sectional loan-level dataset of 435,288 observations. *Time-To-Close* averages 40.2 days with a standard deviation of 21 days. Consistent with panel (a), about 10% of loans exceed 60 days to close. Borrower characteristics—including demographics, first-time home buyer status, FHA share, income, loan amount, LTV, and FICO score—all closely mirror those in quarterly loan panel in Table 1 panel (a).

3. The Impact of Initial Borrowing Frictions on Future Refinancing

3.1. OLS Specification

In this section, I now turn to the empirical analysis. I test whether delays in the initial loan origination impact future refinancing activities using the CoreLogic–MBS quarterly panel. Specifically, I estimate the following regression equation:

$$\begin{aligned} \mathit{Refinance}_{i,t} =& \alpha + \beta \cdot \mathbb{1}(\mathit{Time\text{-}To\text{-}Close} > 60 \text{ Days})_i + \delta \cdot X_{i,t} + \eta_{\mathit{age\ group}} \\ & + \eta_{\mathit{county}\times\mathit{origin\ year}} + \eta_{\mathit{year\text{-}quarter}} + \eta_{\mathit{lender}} + \epsilon_{i,t}, \end{aligned} \tag{2}$$

where $\mathit{Refinance}_{i,t}$ is an indicator equal to one if loan i is refinanced in quarter t, and zero otherwise. The key independent variable, $\mathbb{1}(\mathit{Time\text{-}To\text{-}Close} > 60 \text{ Days})_i$, is a dummy equal to one if loan i experienced an origination delay exceeding 60 days.

The regression controls for a list of borrower and loan-level characteristics, $X_{i,t}$, including borrower race/ethnicity, sex, presence of a co-borrower, first-time homebuyer status, income, loan amount, LTV ratio at origination, quarterly updated LTV ratio, FICO score, loan age, and the rate gap. To capture potential nonlinear effects, I also include the squared terms of these variables. Additionally, fixed effects for borrower age groups (sourced from InfoUSA), county-by-origination-year (or tract-by-origination-year), year-quarter, and lender (or loan officer) are included to account for unobserved heterogeneity across borrowers, geographic and temporal dimensions, and lender-specific factors.

Table 2 presents the OLS regression results. Columns (1)–(4) use the full sample of GSE and FHA loans, where I progressively tighten the specification to account for unobserved heterogeneity. In column (1), I control for a comprehensive set of borrower and loan characteristics and their square terms, along with fixed effects for borrower age group, county-by-origination-year, and year-quarter. The estimated coefficient on $\mathbb{1}(\textit{Time-To-Close} > 60 \text{ Days})$ is -0.147, indicating that loans with delays exceeding 60 days are 14.7 basis points less likely to be refinanced in a given quarter.

In column (2), I add lender fixed effects, making comparisons among borrowers who originated loans from the same lender. This inclusion slightly reduces the magnitude of the coefficient to -0.131, suggesting that part of the variation in refinancing behavior is attributable to lender-specific factors. In column (3), I replace lender fixed effects with loan officer fixed effects, providing a tighter control by comparing borrowers served by the same loan officer but with differing delay experiences. This further reduces the coefficient to -0.105, reflecting a 10.5 basis point decline in refinancing probability associated with origination delays. Finally, in column (4), I tighten geographic controls by replacing county-by-origin-year fixed effects with tract-by-origin-year fixed effects, and the change slightly increases the magnitude of the effect to -0.120. Overall, across these specifications, the estimated impact of initial delays ranges from -0.147 to -0.105, corresponding to a 3.5% to 4.9% reduction relative to the mean quarterly refinancing rate of 3.02%.[18]

In columns (5)–(8) of Table 2, I examine the impact of initial mortgage delays separately for the GSE and FHA loan subsamples, finding consistently negative and statistically significant effects. Columns (5) and (6) focus on the GSE sample. In Column (5), with full controls and fixed effects for borrower age group, county-by-origination-year, year-quarter, and loan officer, the coefficient on $\mathbb{1}(\textit{Time-To-Close} > 60 \text{ Days})$ is -0.129, indicating a 12.9 basis point reduction in quarterly refinancing probability. When county-by-origin-year fixed effects are replaced with tract-by-origin-year fixed effects in Column (6), the magnitude increases to -0.186.

Columns (7) and (8) report results for the FHA sample, where the coefficients range from -0.106 to -0.082. Although the absolute magnitude is smaller for FHA loans, the relative reduction compared to the mean refinancing rate is similar across both loan types—ranging from 3.8% to 5.4% for GSE loans and 4.0% to 5.2% for FHA loans.[19]

Overall, the results in Table 2 provide robust evidence of a negative association between delays in ini-

[18]Computed as *coefficient estimate* ÷ *quarterly mean refinancing rate*. For instance, $-0.147 \div 3.02 \approx -4.9\%$.

[19]For the GSE sample, the mean refinancing rate is 3.41%, and for the FHA sample, it is 2.03%.

tial borrowing and subsequent refinancing activity, consistent across detailed borrower, loan, geographic, and lender-specific controls, as well as across loan type subsamples.

3.2. Threats to Identification

While the OLS estimates presented in the previous section provide initial insights into the relationship between initial mortgage delays and subsequent refinancing behavior, two key identification challenges warrant careful attention.

Omitted Variable Bias. A primary concern is the potential endogeneity arising from unobserved borrower characteristics that may influence both the likelihood of experiencing origination delays and the propensity to refinance. For instance, borrowers with limited financial literacy or lower levels of sophistication might be more prone to delays during the mortgage origination process and simultaneously less inclined or able to navigate refinancing opportunities. The rich set of time-varying controls at the borrower- and loan-level, as well as tight fixed effects may help mitigate much of this concern. However, if such unobserved traits are not fully controlled for, the OLS estimates may overstate the true effect of origination delays on refinancing behavior. Conversely, if borrowers with higher prepayment risks face more stringent underwriting processes leading to longer origination times, the OLS estimates might understate the true effect. This possibility of bias in either direction presents the need for an identification strategy that isolates exogenous variation in origination delays, independent of borrower characteristics.

Measurement Error. Another—and perhaps more critical—challenge relates to measurement error in the key independent variable, $\mathbb{1}(\textit{Time-To-Close} > 60 \text{ Days})$. This binary indicator is intended to capture lender-induced delays, which may lead to borrower dissatisfaction and subsequently discourage interactions with the lenders for refinancing. However, it may also reflect postponements driven by borrowers or sellers for reasons unrelated to lender performance. For instance, borrowers might request extended closing periods due to personal financial planning or logistical needs, while sellers may delay transactions to accommodate their own schedules. These non-lender-related delays introduce noise into the measurement of lender-side frictions, potentially attenuating the estimated effect of origination delays on refinancing behavior. As a result, the observed delay indicator may imperfectly proxy the type of delay most relevant for influencing future refinancing decisions.

3.3. Instrumental Variable Approach

To address the identification challenges, I implement an IV strategy that leverages exogenous variation in loan officer-level processing capacity. In particular, I use the loan officer's workload at the time of application as an instrument for the likelihood of borrowers experiencing an origination delay. This approach directly tackles the empirical concerns outlined in the previous subsection in the following ways.

First, this approach mitigates the *omitted variable bias* by exploiting variation in delays driven by operational constraints that are plausibly unrelated to unobserved borrower characteristics. Conditional on applying to a given loan officer, it is unlikely that borrowers can anticipate or influence the officer's workload at the time of their application. Therefore, after controlling for detailed borrower, loan, geographic, and lender factors, fluctuations in loan officer workload provide a source of exogenous variation in processing delays that is independent of borrower traits.

Second, this IV strategy also addresses *measurement error* in the delay indicator. As discussed earlier, the observed variable, $\mathbb{1}(\textit{Time-To-Close} > 60 \text{ Days})$, may conflate lender-induced delays with those arising from borrower- or seller-driven factors. By instrumenting delays with loan officer workload—capturing variation in lender-side operational frictions—I isolate the component of delays most relevant to borrower dissatisfaction and subsequent refinancing behavior, thereby mitigating potential attenuation bias.

To my knowledge, this is the first study to exploit capacity constraints at the individual loan officer level as an instrument to identify the causal effect of lender-driven origination delays on borrower refinancing behavior. This strategy builds on prior research that leverages lender-side capacity constraints, which are known to predict mortgage origination delays (Choi et al., 2022; Fuster et al., 2024).[20] However, while previous studies measure capacity at broader levels—e.g., bank-level—I extend this approach by capturing time-varying constraints at the individual loan officer level.

I define *Workload* as the number of active (i.e., incomplete) loan applications a loan officer was handling at the time a new application was submitted.[21] I then estimate the following 2SLS specification,

[20] Choi et al. (2022) identify operational capacity constraints as a major bottleneck in purchase mortgage originations, while Fuster et al. (2024) show that these constraints lead to longer processing times and delays.

[21] This definition is conceptually similar to the bank-level operational capacity measure used in Choi et al. (2022), where capacity is proxied by the ratio of incomplete applications at the end of each quarter to total applications received.

adapted from Equation (2):

(First Stage)

$$\mathbb{1}(\textit{Time-To-Close} > 60 \text{ Days})_i = \alpha + \beta \cdot \textit{Workload}_i + \delta \cdot X_{i,t} + \eta_{\textit{age group}} + \eta_{\textit{county}\times\textit{origin year}} + \eta_{\textit{year-quarter}} + \eta_{\textit{loan officer}} + \epsilon_{i,t}, \tag{3}$$

(Second Stage)

$$\textit{Refinance}_{i,t} = \alpha + \beta \cdot \widehat{\mathbb{1}(\textit{Time-To-Close} > 60 \textit{ Days})_i} + \delta \cdot X_{i,t} + \eta_{\textit{age group}} + \eta_{\textit{county}\times\textit{origin year}} + \eta_{\textit{year-quarter}} + \eta_{\textit{loan officer}} + \epsilon_{i,t}. \tag{4}$$

where $\mathbb{1}(\textit{Time-To-Close} > 60 \text{ Days})_i$ is a dummy that equals one if loan i had a delay longer than 60 days until its closing; $\textit{Workload}_i$ measures the number of active (i.e., incomplete) loan applications an officer is managing at the time of each application; $\textit{Refinance}_{i,t}$ is an indicator variable whether loan i was refinanced in quarter t; $X_{i,t}$ include borrower and loan-level controls as in Equation (2); and $\eta_{\textit{age group}}$, $\eta_{\textit{county}\times\textit{origin year}}$, $\eta_{\textit{year-quarter}}$, and $\eta_{\textit{loan officer}}$ stand for borrower age groups, county-by-origination-year (or tract-by-origination-year), year-quarter, and loan officer fixed effects.

3.3.1. Validity of Instrument

Figure 4 visually illustrates the relationship between loan officer workload and the probability of experiencing an initial loan delay exceeding 60 days. Panel (a) presents a binned scatter plot using the raw values of *Workload* and the 60+ day delay indicator, and panel (b) shows the relationship after residualizing both variables by the full set of borrower and loan characteristics, squared terms, and fixed effects for borrower age group, county-by-origination-year, year-quarter, and loan officer. In both panels, there is a clear positive and monotonic relationship: as loan officer workload increases, the likelihood of origination delays rises.

The visual evidence in Figure 4 is formally tested in the first-stage regression results reported in columns (1) and (2) of Table 3. Consistent with the positive relationship observed in the binned scatter plots, *Workload* emerges as a strong and statistically significant predictor of delays exceeding 60 days. Column (1) controls for borrower and loan characteristics, their squared terms, and includes fixed effects for borrower age group, county-by-origination-year, year-quarter, and loan officer. In column (2), I tighten the geographic controls by replacing county-by-origination-year fixed effects with tract-by-origination-year fixed effects. Across both specifications, the coefficient on *Workload* remains stable and economically meaningful, with first-stage F-statistics well above the conventional threshold of 10, providing strong evidence in support of the instrument's relevance.

The validity of the exclusion restriction relies on the assumption that loan officer workload affects refinancing behavior only through its impact on origination delays, and not through any direct channel or correlation with borrower characteristics that independently influence refinancing outcomes. A potential concern is that borrowers who are inherently less likely to refinance—due to unobserved preferences or financial sophistication—might systematically apply during periods when loan officers are busier. Although this scenario is unlikely given that borrowers typically have limited visibility into loan officer workloads at the time of application, I provide indirect evidence to support this assumption through covariate balance tests.

Columns (3) and (4) of Table 3 examine whether *Workload* is systematically correlated with observable borrower characteristics. Across key variables—including race/ethnicity, sex, co-borrower status, first-time home buyer status, income, loan amount, LTV ratio, and FICO score—there are no statistically significant associations, with the exception of a few isolated cases (e.g., FHA loan status and loan amount). These results suggest that loan officer capacity constraints are largely orthogonal to borrower attributes that could independently drive refinancing behavior. Overall, this evidence supports the plausibility of the exclusion restriction by indicating that variation in *Workload* is not driven by borrower selection but reflects exogenous fluctuations in loan officer capacity.

3.3.2. 2SLS Results

Table 4 presents the 2SLS regression estimates of the impact of initial mortgage delays on refinancing behavior, using loan officer workload as an instrument for delays exceeding 60 days. Across specifications, the IV estimates are substantially larger than their OLS counterparts in Table 2, indicating a pronounced discouraging effect of lender-induced delays on subsequent refinancing.

Columns (1) and (2) report estimates for the full GSE and FHA sample. In column (1), controlling for borrower and loan characteristics, squared terms, and fixed effects for borrower age group, county-by-origination-year, year-quarter, and loan officer, the coefficient on $\mathbb{1}(\textit{Time-To-Close} > 60\text{Days})$ is -0.477. Tightening geographic controls in column (2) by replacing county-level with tract-level fixed effects increases the magnitude to −0.731. These estimates correspond to a 15.8% to 24.2% reduction relative to the mean quarterly refinancing rate of 3.02%.[22]

Columns (3)–(6) present separate estimates for the GSE and FHA subsamples. For GSE loans, the coefficients range from −0.569 to −0.980, implying a 16.7% to 28.7% reduction relative to the mean refinancing rate of 3.41%. For FHA loans, the effects range from −0.336 to −0.827, corresponding to a 16.6% to 40.7% reduction relative to the lower mean refinancing rate of 2.03%. Across both loan types, the estimates are economically meaningful and statistically significant, with larger magnitudes observed under tighter geographic controls.

Taken together, the 2SLS results reveal that loan officer capacity-driven origination delays substantially limit borrowers' refinancing opportunities. The sharp increase in magnitudes relative to the OLS estimates highlights the critical importance of addressing endogeneity issues when studying the consequences of lender-side frictions.

3.3.3. Mechanism

What drives the long-run discouragement effects? I consider two competing mechanisms: lender-specific deterrence and an increase in perceived general costs of refinancing.

Under lender-specific deterrence, borrowers who experience substantial delays develop dissatisfaction or distrust toward their original lender, leading them to avoid refinancing through that lender even when financially advantageous. This mechanism is consistent with findings from Johnson et al. (2019), who show that borrower suspicion significantly depresses refinancing uptake.

Alternatively, delays may raise borrowers' perceived general costs of refinancing. Borrowers could interpret a slow origination experience as a signal that refinancing, in general, would be burdensome—regardless of the lender—thus discouraging refinancing broadly. This interpretation aligns with evidence from Fannie Mae (2014), emphasizing the role of borrower expectations about transaction ease.

The structure of the CoreLogic–MBS dataset allows me to observe whether refinancing occurs with the original lender or a new one, providing a direct test of these two competing mechanisms. If lender-specific deterrence dominates, delays should substantially reduce refinancing with the original lender while leaving switching refinancing largely unaffected.

[22] $\frac{-0.477}{3.016} \approx -15.8\%$ and $\frac{-0.731}{3.016} \approx -24.2\%$.

The results are reported in Table 5. Columns (1) and (2) show that delays substantially reduce recapture refinancing, with coefficients ranging from −0.351 to −0.609. Given the mean recapture refinancing rate of 0.947%, these estimates imply declines of 37.1% to 64.3%.[23] In contrast, columns (3) and (4) report small and statistically insignificant effects on switching refinancing, indicating that borrowers remain willing to refinance when they can sever ties with the original lender.

Thus, the evidence points to a lender-specific trust erosion channel: initial mortgage delays create persistent barriers to borrower-lender re-engagement, rather than materially affecting refinancing willingness more broadly.

3.4. Robustness Checks

I next conduct a series of robustness and validation exercises. These exercises fall into three categories: (i) internal consistency and falsification tests, (ii) robustness to alternative definitions of refinancing incentives—i.e., the closed-form threshold from Agarwal et al. (2013)—and borrower sample restrictions following Keys et al. (2016), and (iii) external validation using the NSMO dataset.

3.4.1. Internal Consistency and Falsification Tests

Effect by Length of Delay. Panel (a) of Figure 5 presents estimates from a specification replacing the 60+ day delay indicator with dummies for varying delay lengths. The results show a clear monotonic pattern: as delays extend from 45+ to 120+ days, the negative impact on refinancing activity becomes larger. This gradient supports the interpretation that borrower discouragement is driven by the severity of lender-side frictions.

Variation Over Loan Age. Panel (b) of Figure 5 plots the effect of delays across subsamples defined by loan age. Consistent with expectations, the discouragement effect is larger for the first 3–4 years but attenuates thereafter, consistent with the trust erosion mechanism, in which the salience of negative lender interactions diminishes as borrowers gain distance from the original transaction.[24]

[23] $\frac{-0.351}{0.947} \approx -37.1\%$ and $\frac{-0.609}{0.947} \approx -64.3\%$.

[24] It is worth noting, however, that this pattern may partly reflect sample selection, as borrowers without initial delays are more likely to have refinanced earlier and thus are underrepresented in longer loan age subsamples. As a result, the observed attenuation should be interpreted with caution.

Falsification Test: Effects on Other Prepayment Events. Lastly, Table B1 reports IV estimates of the impact of initial mortgage delays on alternative prepayment outcomes: *Cash-Out Refinance* and *Prepayment Due to Selling and Moving*. The results for *Cash-Out Refinance* in column (1) show a statistically significant reduction at the 5% level, with a coefficient of -0.38, while the corresponding estimate in column (2) is smaller (-0.22) and statistically insignificant.[25] Meanwhile, the results for *Prepayment Due to Selling and Moving* in columns (3) and (4) are consistently small and statistically insignificant.

This divergence across prepayment types is informative: while both standard and cash-out refinancing require active borrower-lender interaction, selling and moving are often driven by external factors—such as job relocations or life events—that lead to prepayment independent of the origination experience. The absence of an effect on mobility-related prepayments suggests that delays specifically deter borrower-initiated interactions with lenders, not prepayment in general.

3.4.2. Robustness to Alternative Refinancing Incentive Measures and Sample Filters

I next assess whether the main findings are robust to alternative definitions of refinancing incentives and sample restrictions used in prior literature.

Controlling for Alternative Definition of In-the-Moneyness. Rather than relying on the observed interest rate gap, I construct a borrower-specific refinancing threshold using the closed-form solution proposed in Agarwal et al. (2013). This threshold is derived from an option pricing framework that incorporates loan size, interest rate volatility, closing costs, tax deductibility, and other loan-level characteristics to determine whether a borrower is truly "in-the-money." As shown in Table A2, re-estimating the 2SLS regressions while controlling for this alternative refinancing incentive—along with its squared term—yields comparable estimates. This result supports the robustness of the main findings to alternative definitions of refinancing incentives. Detailed construction of this measure is provided in Appendix A.5.

Restricting to Less Credit-Constrained Borrowers. Following Keys et al. (2016), I restrict the sample to borrowers who are unlikely to face binding constraints on refinancing eligibility. This approach helps

[25] When separately examining *recapture* and *switching* cash-out refinancing in Table B2, I find patterns consistent with the regular refinancing results reported in Table 5: the decline is concentrated in recapture refinancing, while switching remains largely unaffected.

isolate the effect of origination delays from borrower-side credit frictions. Although such constraints are less salient in my 2014–2021 sample period compared to the immediate post-crisis years examined in Keys et al. (2016), I nonetheless apply similar sample restrictions to conservatively address this possibility.

I conduct the analysis across progressively stricter subsamples. Table B3 reports the estimated effects of origination delays on refinancing for the following groups: (i) the full GSE sample (replicating column (3) of Table 4); (ii) GSE borrowers with *FICO > 680* and *LTV at Origination < 90%*; (iii) those with *FICO > 680*, *LTV at Origination < 90%*, and no missed payment history; and (iv) those with *FICO > 680*, *Current LTV < 90%*, and no missed payment history. Across all subsamples, the estimated effect of delays remains statistically and economically significant, reinforcing that the main results are not driven by borrower credit quality or refinancing eligibility.

3.4.3. External Validation Using the NSMO Data

To further validate my main findings using an independent data source, I turn to NSMO. The NSMO dataset offers two primary advantages that strengthen its value as external validation. First, it includes direct survey responses on borrower-reported delays during mortgage origination—specifically, delays in mortgage processing and closing. These self-reported measures allow me to clearly identify lender-driven frictions, significantly reducing measurement error concerns relative to possible non-lender-driven delays in the CoreLogic–MBS dataset. Second, NSMO provides detailed controls for borrower sophistication and non-financial characteristics—such as education level, employment type, English proficiency, and parental status—that help mitigate potential omitted variable bias and unobserved heterogeneity in refinancing decisions.

While this survey-based data has inherent limitations—such as the relatively small sample size and the absence of several key variables[26]—NSMO nonetheless offers a valuable opportunity to externally assess the validity of the main findings by examining whether borrower-reported delays during origination are associated with suppressed future prepayment behavior.[27]

[26]For example, NSMO does not explicitly distinguish refinancing from other prepayment types, requiring the use of a general prepayment dummy as the outcome variable. Additionally, the lack of lender and geographic identifiers precludes the inclusion of certain fixed effects.

[27]My analysis complements Bhutta and Doubinko (2025), who also use NSMO to study borrower experiences. While their analysis is cross-sectional and focuses on prepayment activity in 2020–2021, I leverage a loan-quarter panel spanning 2013–2021 to examine how self-reported delays causally affect subsequent prepayment behavior over time. Their findings are qualitatively consistent with mine, reinforcing the importance of origination experiences in shaping downstream mortgage refinancing decisions.

Effect of Experiencing Origination Delays on Future Prepayment. I construct a quarterly panel using the NSMO dataset, restricting my sample to borrowers who originated 30-year, fixed-rate, single-family home purchase mortgages between 2013 and 2021. This restriction yields a final sample of 14,585 unique loans, resulting in 241,048 loan-quarter observations. Summary statistics for key variables from the NSMO dataset are reported in Table B4.

Since NSMO does not distinguish refinancing from other forms of prepayment, I use a general prepayment indicator as a proxy for refinancing activity. I define two primary independent variables based on survey responses capturing borrower-reported origination delays:

i. ***Processing Delay***: An indicator equal to one if the borrower responds *yes* to the question, "In the process of getting this mortgage from your *mortgage lender/broker*, did you redo/refile paperwork due to processing delays?"

ii. ***Closing Delay***: An indicator equal to one if the borrower responds *yes* to the question, "In the process of getting this mortgage from your *mortgage lender/broker*, did you delay or postpone your closing date?"

Additionally, I construct a measure of the rate gap following the same methodology as described in Section 2.4. Specifically, I recover the initial mortgage rate by summing the reported rate spread and the Freddie Mac PMMS rate at origination, both available in the NSMO data.

Table 6 presents the regression results using this loan-quarter level panel, examining the impact of experiencing origination delays on future prepayment behavior. The findings indicate that self-reported delays—both in processing and closing—are significantly associated with reduced subsequent prepayment. Column (1) shows that borrowers reporting a processing delay exhibit a 0.33 percentage point decrease in prepayment probability, while column (2) finds a 0.41 percentage point reduction associated with closing delays. These results remain robust across loan-type subsamples: GSE loans (columns (3)–(4)) and FHA loans (columns (5)–(6)), while FHA borrowers have a larger negative response to closing delays relative to processing delays.

The magnitude of these estimates is broadly consistent, albeit slightly smaller, compared to the 2SLS results presented in Table 4. This smaller magnitude aligns logically with expectations, given that the dependent variable, *Prepaid*, in NSMO includes exogenous prepayment events unrelated to refinancing decisions (e.g., moving shocks) that would not be influenced by prior delay experiences. The consistency in findings across two independent datasets and different identification strategies provides strong external validation, reinforcing the causal interpretation that origination delays suppress future refinancing activity.

Mechanism: Initial Delay Experience and Borrower Satisfaction. Beyond documenting the effect of delays on prepayment, I also examine the underlying mechanism driving this behavior. In particular, I test whether respondent's reported delay experiences lead to lower subjective satisfaction with various aspects of the mortgage origination process—consistent with a trust erosion channel. To capture borrower satisfaction comprehensively, I construct six distinct outcome variables from survey responses: *Perceived Fair Treatment*, *Dissatisfied by Lender*, *Dissatisfied by Application*, *Dissatisfied by Documentation*, *Dissatisfied by Closing*, and *Dissatisfied by Overall*.[28]

As presented in Table 7, experiencing origination delays—whether processing or closing—is significantly associated with reduced perceptions of fair treatment by lenders (column (1)) and substantially higher dissatisfaction across multiple dimensions of the mortgage process (columns (2)–(6)). For example, column (2) of panel (a) shows that borrowers who encountered processing delays are nearly 295% more likely to report dissatisfaction with their lenders, relative to a 4% baseline among those without processing delays. These results suggest that the reduced refinancing propensity following origination delays is closely linked to borrower dissatisfaction and trust erosion, rather than to purely mechanical or financial barriers.

4. Who Is More Exposed to Initial Borrowing Frictions?

The preceding analysis shows that lender-induced delays in mortgage origination substantially deter future refinancing, primarily through lender-specific trust erosion. I now turn to the question of which borrower groups are most exposed to these frictions. Identifying the distribution of delays across demographic and financial characteristics is critical, as disproportionate exposure could exacerbate inequalities in credit access and long-term financial outcomes.

[28] *Perceived Fair Treatment* equals one if borrowers respond "yes" to the question, "most mortgage lenders generally treat borrowers well." *Dissatisfied by Lender/Application/Documentation/Closing* equal one if respondents answer "not at all" to the following questions: "Overall, how satisfied are you with the lender or mortgage broker you used?"; "Overall, how satisfied are you with the application process?"; "Overall, how satisfied are you with the documentation process required for the loan?"; and "Overall, how satisfied are you with the loan closing process?" Lastly, *Dissatisfied by Overall* equals one if any dissatisfaction indicator equals one.

Prior research, such as Wei and Zhao (2022), documents that minority borrowers faced longer processing times during the pre-crisis period. I extend this analysis to the post-crisis era, examining whether racial disparities persist and whether other vulnerable groups—such as low-income or lower-credit-score borrowers—also experience heightened exposure to origination delays.

To quantify these patterns, I estimate the following loan-level regression:

$$\begin{aligned}\mathbb{1}(\textit{Time-To-Close} > 60 \text{ Days}) =& \alpha + \beta_1 \cdot \textit{Minority}_i + \beta_2 \cdot \textit{Asian}_i + \beta_3 \cdot \textit{Other Race}_i + \delta \cdot X_i \\ &+ \eta_{\textit{age group}} + \eta_{\textit{county} \times \textit{origin year}} + \eta_{\textit{lender}} + \epsilon_i, \end{aligned} \tag{5}$$

where the dependent variable $\mathbb{1}(\textit{Time-To-Close} > 60 \text{ Days})$ is a binary indicator equal to one if the loan has taken more than 60 days for closing. The key variable of interest, *Minority*, is a dummy variable equal to 1 for Black and Hispanic borrowers. Additional race dummies, *Asian* and *Other Race*, are also included in the regression.

The regression controls a set of borrower- and loan-level characteristics at origination, denoted by X_i, which may influence loan origination durations. These controls include indicators for female borrowers and the presence of a co-borrower, first-time home buyer status, the logarithms of borrower income and loan amount, the origination LTV ratio, and the FICO score. Additionally, I include fixed effects for borrower age groups, county-by-origination-year, and lender (or loan officer) to account for unobserved heterogeneity across borrower demographics, time-varying local economic conditions, and lender-specific practices.

Table 8 presents the regression results, beginning with racial disparities. Consistent with Wei and Zhao (2022), I find that minority borrowers are significantly more likely to face origination delays. In the specification that incorporates only race indicators and county-by-year fixed effects (column (1)), minority borrowers are 3.68 percentage points more likely to experience delays. Adding lender fixed effects in column (2) reduces the gap to 3.13 percentage points, suggesting that part of the disparity stems from differences in lender selection.

Controlling for borrower demographics, income, loan amount, LTV, and FICO score in column (3) further reduces the estimated gap to 2.53 percentage points. Even after tightening identification by including loan officer fixed effects in column (4)—comparing borrowers served by the same officer—the disparity persists at 1.84 percentage points, representing an 18.6% increase relative to the baseline delay rate of 9.9%.

Patterns are similar when examining the GSE and FHA subsamples in columns (5) and (6). Minority borrowers in both markets continue to face significantly higher probabilities of delay, with coefficients of 1.49 and 1.87 percentage points, respectively. These results suggest that racial disparities in exposure to borrowing frictions persist across loan product types.

Beyond race, the results also show that lower-income borrowers and those with weaker credit scores are more exposed to delays. Across columns (3)–(6), higher income and FICO scores are consistently associated with shorter processing times. For instance, in column (4), a one percent increase in income reduces the probability of a delay by 0.71 percentage points, while a 100-point increase in FICO score lowers the probability by approximately 1.57 percentage points. These patterns also hold across both GSE and FHA subsamples.

4.1. Evidence of Lender Bias in Racial Gaps in Mortgage Delays

The previous analysis shows that race, income, and credit scores all influence exposure to mortgage origination delays. Longer *Time-To-Close* for lower-income or lower-credit-score borrowers may partly be justified by underwriting practices, where riskier applicants undergo more extensive review. However, race is not a factor in credit risk models or underwriting criteria. This raises the question of whether racial disparities in *Time-To-Close* reflect lender-side discrimination or differences in unobservable borrower risk. To address this question, I conduct a series of tests designed to detect patterns consistent with discriminatory behavior.

Variation by Racial Animus. I first test whether minority borrowers experience greater delays in areas with higher levels of racial animus. Following Stephens-Davidowitz (2014), I proxy for racial animus using the frequency of racially charged Google search terms at the metropolitan statistical area (MSA) level. Column (2) of Table 9 reports estimates from regressions interacting the minority indicator with a *High Race Animus* dummy, which equals one for MSAs above the median in racial animus.

The interaction coefficient is positive and statistically significant, indicating that minority borrowers in high-animus areas are substantially more likely to experience closing delays. The magnitude suggests that racial disparities in *Time-To-Close* are roughly three times larger in these regions compared to areas with lower animus. This finding is consistent with prior evidence that discriminatory behavior intensifies in regions with heightened racial bias across other markets, including auto lending, labor, and municipal finance (Butler et al., 2022; Charles and Guryan, 2008; Dougal et al., 2019).

Variation by Local Market Competition. Next, I test whether minority borrowers face larger delays in less competitive lending markets. When competition is limited, lenders may exercise greater discretion, allowing taste-based discrimination to persist (Berkovec et al., 1998). Column (3) of Table 9 includes an interaction between *Minority* and *Low Local Competition*, defined as counties within the top tercile of the top-four lender market share.

The positive and significant coefficients on this interaction suggest that minority borrowers are indeed more exposed to delays in less competitive markets. This reinforces the interpretation that lender-side preferences, rather than borrower credit risk, contribute to racial disparities in *Time-To-Close*.

Contrasting with Delinquency Outcomes. If racial disparities in closing delays were purely driven by unobserved borrower risk correlated with race, similar patterns would be expected in other credit outcomes, such as delinquency. To test this, I re-estimate the models using a dummy for 90+ days delinquency as the dependent variable (columns (4)–(6) of Table 9).

While the minority indicator remains positive and significant across all three columns[29], the interaction terms with racial animus and market concentration dummies are statistically insignificant and indeed negative. This divergence highlights a key distinction: while racial disparities in closing delays are amplified in environments conducive to discrimination, delinquency rates among minority borrowers do not exhibit similar sensitivity to local racial animus or market structure. This contrast strengthens the case that the observed racial disparities in *Time-To-Close* reflect lender-side bias rather than underlying borrower risk.

5. Measuring the Financial Consequences of Mortgage Origination Delays

The empirical results show that delays in mortgage origination significantly reduce the likelihood of refinancing, thereby increasing long-term borrowing costs. In this section, I quantify the financial consequence arising from these delays using two complementary approaches. First, I provide a back-of-the-

[29]This is consistent with established findings on higher delinquency rates among minority borrowers even after controlling for credit risk factors (Berkovec et al., 1997; Bayer et al., 2016; Kermani and Wong, 2024).

envelope calculation to estimate the average cost of missed refinancing opportunities due to origination delays. Second, I use a quantitative model of refinancing behavior to simulate how these initial origination delays affect a borrower's coupon rate trajectories over the life of the loan. While the back-of-the-envelope calculation offers a simple and intuitive benchmark, the simulation approach offers a more detailed assessment by realistically capturing both the timing of refinancing opportunities and the extent of borrower *responsiveness*, while also incorporating heterogeneity across racial and income groups.

5.1. A Back-of-the-Envelope Estimate

The reduced-form evidence in Section 3 suggests that experiencing a 60+ day closing delay reduces a borrower's quarterly refinancing probability by up to 24%—that is, by 73 basis points relative to the baseline quarterly refinancing rate of 3.02%. Based on the CoreLogic–MBS loan panel, the average realized interest rate reduction from refinancing is 87 basis points, and the average loan balance is \$279,288.[30] This implies that the missed annual interest cost savings for a delayed borrower can be approximated as:

$$\underbrace{87 \text{ bp}}_{\text{avg. realized savings}} \times \underbrace{\$279,288}_{\text{avg. loan size}} \times \underbrace{24\%}_{\text{drop in refi rate}} = \$586.5. \tag{6}$$

The PV of the cumulative annual overpayment of \$586.5 varies substantially depending on when the refinancing opportunity arises during the 30-year loan term. Figure 6 illustrates this relationship. If the opportunity arises immediately after origination, the discounted PV of forgone savings can be as high as \$11,500, as the delayed borrower incurs overpayments of \$586.5 annually from the second quarter through the 120th quarter. The forgone savings decline monotonically as the refinancing opportunity occurs later in the loan term, eventually converging to zero by the final quarter.

As a benchmark, suppose the refinancing opportunity arises exactly at the midpoint of 30 years of loan term (i.e., after 60 quarters). In this case, the borrower incurs quarterly overpayments of $\$586.5/4 = \$146.6$ from quarter 61 through 120. The discounted PV, using a 3% annual discount rate, is then:

$$\text{PV} = \sum_{t=61}^{120} \frac{\$586.5/4}{(1+0.03/4)^t} \approx \$4{,}511. \tag{7}$$

That is, even a quick calculation under simplifying assumptions suggests that a subtle frictional expe-

[30] Drawn from exp(12.54) in panel (b) of Table 1.

rience during mortgage origination can translate into several thousand dollars in lost savings by limiting refinancing opportunities.

5.2. Simulation-Based Quantification of Financial Losses

5.2.1. Motivation

The back-of-the-envelope calculation above provides a useful first-pass estimate but abstracts from two key dimensions that shape the financial consequences of missed refinancing opportunities.

First, it relies on a simplifying assumption that a single refinancing opportunity—resulting in an 87 basis point interest rate reduction—occurs at a fixed point in time (e.g., year 15). Indeed, the discounted PV of overpayment varies substantially with the timing of that event, as shown in Figure 6, making it difficult to precisely quantify the resulting financial loss. Moreover, in reality, borrowers may encounter multiple chances to refinance—or none at all—over the course of the loan. Accurately capturing the long-run financial impact of origination delays therefore requires simulating realistic mortgage rate paths and the arrival of such opportunities throughout the loan term.

Second, the back-of-the-envelope calculation does not account for refinancing *inertia*—the tendency of even non-delayed borrowers to refinance gradually rather than respond immediately to opportunities (Andersen et al., 2020). This matters because the financial burden of delays depends on a borrower's underlying *responsiveness*. If, for example, borrowers are generally inattentive or slow to act, delays may have little additional effect, since refinancing would have been unlikely regardless. In contrast, if borrowers are more responsive, delays can meaningfully reduce the likelihood of refinancing. Therefore, incorporating typical borrower *responsiveness*, as observed in the data, is critical for accurately quantifying the financial loss caused by origination delays.

Motivated by these concerns, I employ a quantitative model of refinancing behavior and simulate plausible coupon rate trajectories for both delayed and non-delayed borrowers. The model incorporates two key components: (i) a stochastically evolving market mortgage rate and (ii) a borrower *responsiveness* parameter, which governs the likelihood of taking up a favorable refinancing opportunity within a given time period. For delayed borrowers, *responsiveness* is scaled down relative to the baseline estimated from non-delayed borrowers. Using this framework, I repeatedly simulate coupon trajectories for delayed and non-delayed borrowers to quantify the overpayments caused by origination delays. As an extension, I estimate *responsiveness* separately by race and income—motivated by documented heterogeneity in refinancing behavior (Andersen et al., 2020; Gerardi et al., 2023)—to examine distributional implications in the resulting financial losses.

5.2.2. Model Setup

Consider a household with a fixed-rate mortgage carrying coupon $c_{i,t}$ at time t. The market mortgage rate m_t evolves stochastically and is common across all borrowers. That is, any borrower can refinance into a new mortgage at the prevailing market rate.[31]

Following the standard modeling framework that incorporates time-dependent inaction in mortgage refinancing (Andersen et al., 2020; Berger et al., 2024), I assume that a borrower refinances only when both of the following conditions are met:

i. **The refinancing opportunity is in-the-money:** The current market rate is below the borrower's coupon rate, i.e., $c_{i,t-1} - m_t > 0$.

ii. **The borrower becomes responsive:** Each quarter, borrowers randomly become ready to act on in-the-money refinancing opportunities. This form of time-dependent inaction is modeled as a Poisson process with intensity χ, borrower *responsiveness*.[32]

Importantly, I do not model explicit refinancing costs (e.g., up-front fees). This modeling choice allows the discouraging effect of initial delays to operate solely through a reduction in the *responsiveness* parameter χ, thereby simplifying the analysis and enhancing tractability. Moreover, recent research suggests that time-dependent inaction frictions are more influential than up-front costs in shaping refinancing behavior.[33]

[31] This abstraction of borrower-invariant pricing is justified by institutional features of the U.S. mortgage-backed securities (MBS) market, including loan pooling and the to-be-announced (TBA) trading convention. The uniform pricing assumption is standard in economics models of mortgage refinancing behavior (Berger et al., 2024; Eichenbaum et al., 2022; Zhang, 2024).

[32] This is often referred to as an *inattention* friction in the literature, though it more broadly captures any form of time-dependent inaction in refinancing behavior.

[33] For example, Berger et al. (2024, Forthcoming) show that households refinance sluggishly even when up-front costs are minimal. Zhang (2024) also documents that 68–83% of closing costs are effectively rolled into the loan via higher mortgage rates.

5.2.3. Estimating Household Refinancing *Responsiveness* Parameter

I construct a likelihood function based on observed refinancing behavior in the CoreLogic–MBS dataset. The primary objective is to recover a representative baseline estimate of the *responsiveness* parameter χ, focusing on borrowers whose behavior has not been influenced by prior origination delays. I then apply the 24% reduction in refinancing probability—estimated from the 2SLS analysis in Section 3.3—to the baseline χ estimate to calibrate the parameter for delayed borrowers.[34]

Specifically, conditional on being in-the-money (i.e., $c_{i,t-1} - m_t > 0$), whether a borrower refinances in a given quarter depends solely on the probability of being responsive to refinancing opportunities:

$$P(y_{i,t} = 1 \mid c_{i,t-1} - m_t > 0) = 1 - \exp(-\chi\, dt), \tag{8}$$

$$P(y_{i,t} = 0 \mid c_{i,t-1} - m_t > 0) = \exp(-\chi\, dt), \tag{9}$$

where $y_{i,t}$ is an indicator equal to 1 if household i refinances at time t, and 0 otherwise; and $dt = \frac{1}{4}$ denotes the length of a quarter.

The corresponding log-likelihood function is specified as follows:

$$\mathscr{L}(\chi) = \sum_i \sum_t \left[y_{i,t} \log\left(1 - e^{-\chi\, dt}\right) - (1 - y_{i,t})\, \chi\, dt \right] \mathbb{1}(c_{i,t-1} - m_t > 0). \tag{10}$$

I estimate χ by maximizing the likelihood function in Equation (10). To recover a representative baseline estimate, I restrict the CoreLogic–MBS quarterly loan panel to a subsample of loans that were not subject to origination delays and did not experience any credit event, such as prepayment and default, or missed payment. Additionally, because loans originated after 2020 have relatively short performance histories, I include only those originated prior to 2020. Finally, to exclude borrowers unlikely to refinance, I follow Keys et al. (2016) and retain only loans with initial FICO scores above 680 and current LTV ratios below 90%.

The resulting estimates are presented in Table 10. The first row reports the baseline estimate of $\chi = 0.405$, which implies a 90.4% probability[35] that a borrower will miss an in-the-money refinancing

[34] A complementary strategy is to estimate both the baseline refinancing *responsiveness* (χ) and the discouragement effect of origination delays (γ) simultaneously within the model. I implement this approach in Appendix A.6, which reveals even larger financial losses and distributional consequences.

[35] Computed as $\exp(-0.405 \times \frac{1}{4})$.

opportunity in a given quarter—closely matching the 87% *asleep* probability estimated by Andersen et al. (2020).

Table 10 also reports estimates by race and income subgroup. The results indicate that χ varies systematically across groups. By race, the *responsiveness* parameter is lowest for Black borrowers (0.247), followed by Hispanic (0.318), White (0.432), and Asian (0.489) borrowers. By income, low-income households—defined as those in the bottom income tercile—exhibit the lowest χ (0.293), followed by middle-income (0.419) and high-income (0.527) households. These subgroup estimates are consistent with prior findings (Gerardi et al., 2023), which show that historically underserved groups refinance more slowly even when unconstrained and when refinancing would be financially beneficial.[36]

5.2.4. Simulating Mortgage Rates and Refinancing Outcomes

Having outlined the model structure and estimated key parameters, I now use the framework to simulate borrower refinancing behavior and quantify the financial consequences of origination delays. I begin by generating market mortgage rate paths using an exogenous stochastic process, reflecting the assumption that borrowers are price-takers in the mortgage market. These simulated rate paths are then used to construct borrower coupon trajectories over a 30-year horizon for both delayed and non-delayed borrowers, with refinancing decisions governed by their *responsiveness* parameter χ. By conducting a large number of simulations, I compare outcomes between delayed and non-delayed borrowers to estimate the average cumulative overpayment attributable to origination delays.

Mortgage Rate Path Simulation. The exogenous mortgage rate, m_t, is modeled as the sum of a stochastic short rate process and a fixed spread. Specifically, the short rate r_t follows the Cox–Ingersoll–Ross (CIR) process (Cox et al., 1985), which captures mean reversion and ensures non-negativity in interest rate dynamics:

$$dr_t = \kappa(\mu - r_t)\,dt + \sigma\sqrt{r_t}\,dB_t, \tag{11}$$

where κ is the speed of mean reversion, μ is the long-run mean rate, σ is the volatility parameter, and dB_t is a standard Brownian motion. For the simulation, I adopt parameter values presented in Table 11,

[36] These estimates are based on borrowers who did not experience origination delays and are not credit-constrained, as described above.

closely following Berger et al. (2024): $\kappa = 0.13$, $\mu = 0.035$, and $\sigma = 0.06$. The initial short rate is set to 3%, and I simulate 240 quarters of rate paths, discarding the first 120 quarters as a burn-in period to mitigate sensitivity to initial conditions.

The market mortgage rate is then computed as $m_t = r_t + 0.0168$, where the fixed 168 basis point spread simplifies the simulation while remaining consistent with observed mortgage pricing patterns during the sample period. As shown in Figure 7, the spread between the Freddie Mac PMMS 30-year mortgage rate and the 30-day U.S. Treasury rate remained relatively stable from 2013 to 2021, particularly in the pre-COVID years.[37]

To capture a wide range of interest rate scenarios, I simulate 10,000 mortgage rate paths, which form the basis for analyzing refinancing behavior.

Refinancing Dynamics and Coupon Rate Paths. Using the 10,000 simulated mortgage rate paths, I generate borrower coupon trajectories—1,000 per path—over a 30-year horizon with quarterly time steps. Specifically, at each period t, refinancing occurs only if the mortgage is in-the-money ($c_{i,t-1} - m_t > 0$). Conditional on this, the probability that the borrower takes up the refinancing opportunity in that period is determined by a Poisson process with intensity χ.

Formally, the coupon rate evolves recursively as:

$$c_{i,t} = \begin{cases} m_t, & \text{if } c_{i,t-1} - m_t > 0 \quad \text{and} \quad dN_t^{\chi} = 1 \\ c_{i,t-1}, & \text{otherwise,} \end{cases} \tag{12}$$

where $dN_t^{\chi} = 1$ indicates a jump in the Poisson process—that is, the borrower becomes responsive and executes the refinancing, provided the opportunity is in-the-money.

Panel (a) of Figure 8 presents a representative example from the full set of 10 million[38] simulations, showing the evolution of coupon rates for four borrower types—*never refinancers*, *always refinancers*, (non-delayed) *average borrowers*, and *delayed borrowers*—alongside the corresponding market mortgage rate path. The *responsiveness* parameter χ is set to 0 for *never refinancers* and to a very high value (∞) for *always refinancers*, approximating immediate refinancing when an opportunity arises.[39] For *average*

[37] It is possible to extend the model to endogenize the mortgage spread by introducing investor-side pricing behavior. I implement this extension in Appendix A.7, which yields larger estimates of the financial consequences of origination delays.

[38] 10,000 mortgage rate paths, each paired with 1,000 borrower coupon trajectories.

[39] In practice, setting $\chi = 10{,}000$ yields near-instant refinancing once an in-the-money opportunity occurs.

borrowers, χ is set to the baseline estimate from Section 5.2.3, $\chi = 0.405$. For *delayed borrowers*, the parameter is reduced to $\chi = 0.308$, reflecting a 24% decline consistent with the empirically observed drop in refinancing probability due to origination delays, as discussed in Section 3.[40]

As shown in the figure, the coupon rate for *never refinancers* remains fixed throughout the loan term, entirely unaffected by fluctuations in the market rate. At the other extreme, *always refinancers* adjust their coupon rate almost immediately whenever an in-the-money opportunity arises, producing a trajectory that closely tracks the lower envelope of the market mortgage rate path.

In contrast to the two polar cases discussed above, *average* and *delayed borrowers* respond to refinancing opportunities intermittently, generating coupon paths that lie between those of *never* and *always refinancers*. However, due to their lower *responsiveness*, *delayed borrowers* refinance less frequently than *average borrowers*, resulting in flatter coupon trajectories and higher cumulative interest payments.

Quantifying Delay-Induced Overpayment. To translate refinancing patterns into monetary terms, I compute the quarterly overpayment made by *delayed borrowers* relative to *average borrowers* as the product of their coupon rate difference and the average loan amount ($279,288). Then, discounting these quarterly overpayments at an annual rate of 3% yields their PV over the life of the loan.

Panel (b) of Figure 8 illustrates this calculation for the representative simulation shown in panel (a). In this example, overpayments remain at zero until the 16th quarter, when the market rate drops below the initial coupon and only the average borrower refinances. In subsequent periods, overpayments increase as refinancing gaps widen, reaching a peak of $1,532 between quarters 26 and 28. The PV of cumulative overpayments, discounted at an annual rate of 3%, amounts to $19,991.

To generalize beyond the single illustrative example in Figure 8, I compute the PV of overpayments incurred by *delayed borrowers* relative to *average borrowers* across all 10 million simulated paths. Figure 9 plots the resulting distribution of PV values. The average PV of overpayment is $6,641—substantially higher than the back-of-the-envelope estimate of $4,511 in Section 5.1. This difference likely reflects the fact that the model-based simulation approach captures refinancing opportunities that arise earlier than the midpoint or occur multiple times.

It is worth noting that Figure 9 shows a mass point at zero. This corresponds to simulations in which

[40] A key advantage of the Poisson framework is its flexibility: the *responsiveness* parameter for delayed borrowers can be directly computed as $(1 - 0.24) \times \chi$, where χ is the baseline estimate.

the market mortgage rate never falls below the borrower's coupon rate over the entire 30-year horizon. In these cases, no in-the-money refinancing opportunities arise, and both delayed and non-delayed borrowers follow identical trajectories—resulting in zero overpayment.

5.2.5. Aggregate Financial Burden

The previous section quantifies the financial consequences of origination delays at the individual borrower level, showing that delayed borrowers pay an average of $6,641 more in PV terms than observationally similar borrowers who refinance at the baseline *responsiveness*. To assess the broader implications, I scale this estimate to the national level using data on annual mortgage originations and the incidence of closing delays.

Using HMDA data, I find that an average of 4,308,256 purchase mortgages were originated each year between 2014 and 2021. Based on my CoreLogic–MBS sample, approximately 9.9% of these loans experience a closing delay of over 60 days, implying roughly: $4,308,256 \times 9.9\% = 426,517$ delayed borrowers per year.

Multiplying the number of delayed borrowers each year by the average PV of overpayment, $6,641, yields the annual aggregate financial burden imposed by origination delays.

$$\text{Aggregate overpayment} = 426,517 \times \$6,641 \approx \$2.83 \text{ billion per year.} \tag{13}$$

Alternatively, using the back-of-the-envelope estimate of $4,511 per borrower from Section 5.1—derived from a simplified calculation—yields an aggregate estimate of:

$$426,517 \times \$4,511 \approx \$1.92 \text{ billion per year.} \tag{14}$$

Taken together, these estimates suggest that initial origination delays may impose aggregate losses of $1.9–$2.8 billion annually, demonstrating that seemingly modest frictions in the origination process can indeed result in large, persistent financial consequences.

5.2.6. Distributional Consequences by Race and Income

Does the same origination delay incidence impose the same financial burden on all borrowers? Likely not—because the cost of a delay depends critically on a borrower's baseline refinancing *responsiveness*. To quantify these distributional differences, I compute the average PV of overpayments due to origination delays across a wide range of baseline *responsiveness* parameters, χ. Specifically, starting from $\chi = 0$, I increment χ in steps of 0.01. For each χ value, I generate 1,000 coupon trajectories for each simulated mortgage rate path and calculate the average PV of overpayments across these trajectories.

Figure 10 visualizes the resulting relationship between χ and the average PV of delay-induced overpayment. The curve exhibits a distinct inverted-U shape: at one end, borrowers with extremely low *responsiveness*—i.e., *never refinancers* with χ near zero—incur little or no cost from delays because they are unlikely to refinance under any circumstance. At the other extreme, highly responsive borrowers (e.g., $\chi \approx 2.0$) also experience limited losses, as their strong propensity to act enables them to eventually take up refinancing opportunities even after a delay. The largest overpayments arise for borrowers with moderate but imperfect *responsiveness* (e.g., $\chi \approx 0.15$–0.20), with average losses approaching $8,000.

This nonlinear relationship suggests that the same origination delay can lead to systematically different financial consequences across demographic groups due to heterogeneity in their baseline refinancing *responsiveness*. As shown earlier in Section 5.2.3, estimated χ values vary meaningfully by race and income. Figure 11 illustrates the resulting disparities: by race, Black borrowers experience the highest average delay-induced overpayment at $7,531, followed by Hispanic ($7,132), White ($6,504), and Asian ($6,227) borrowers. That is, because minority borrowers' χ values lie near the range associated with peak overpayments in Figure 10, the financial consequences of delays are amplified for these groups. A similar pattern emerges across income levels, with average overpayments of $7,277 for low-income borrowers, $6,589 for middle-income borrowers, and $6,053 for high-income borrowers.

These results highlight a compounding disadvantage faced by underserved populations: minority and low-income borrowers are not only more likely to experience origination delays, but also incur greater financial losses when such delays occur. This dual burden may contribute to persistent disparities in financial outcomes and further exacerbate long-standing inequalities in wealth accumulation.

6. Conclusion

This paper examines how frictions in the initial mortgage borrowing process shape future refinancing behavior and contribute to wealth disparities. Using a matched dataset combining CoreLogic and the MBS Loan-Level Datasets from Fannie Mae, Freddie Mac, and Ginnie Mae, I show that extended loan origination times for purchase mortgages significantly reduce borrowers' likelihood of subsequent refinancings. To address the identification challenges, I employ an IV strategy that leverages variation in loan officer workload at the time of application. The results indicate that experiencing a 60+ day delay lowers quarterly refinancing rates by approximately 0.48 to 0.73 percentage points—equivalent to a 15.8% to 24.2% reduction relative to the mean refinancing rate of 3.02%.

Beyond the overall discouraging effect of initial borrowing frictions on refinancing activity, I further examine which borrower groups are more exposed to prolonged loan origination times. I find that minority borrowers—as well as those with lower incomes or lower FICO scores—are significantly more likely to experience delays. Notably, racial disparities in origination delays are most pronounced in areas with heightened racial animus and limited lending market competition, suggesting that lender-side bias, rather than unobserved credit risk of minority borrowers, plays a central role in driving these disparities.

The financial consequences of these frictions are substantial. A back-of-the-envelope calculation and simulation from a quantitative model suggest that missed refinancing opportunities due to initial loan delays results in a PV loss of \$4,511 to \$6,641 per delayed borrower, amounting to \$1.9–\$2.8 billion in aggregate annual costs. Moreover, minority and low-income borrowers incur greater financial losses conditional on experiencing the same delay, due to a lower baseline likelihood of acting on refinancing opportunities.

Overall, this study sheds light on an important but underexplored channel through which supply-side frictions in mortgage markets shape long-term household financial outcomes and contribute to the persistence of wealth inequality.

References

AGARWAL, S., G. AMROMIN, I. BEN-DAVID, S. CHOMSISENGPHET, T. PISKORSKI, AND A. SERU (2017): "Policy Intervention in Debt Renegotiation: Evidence from the Home Affordable Modification Program," *Journal of Political Economy*, 125, 654–712.

AGARWAL, S., G. AMROMIN, S. CHOMSISENGPHET, T. LANDVOIGT, T. PISKORSKI, A. SERU, AND V. YAO (2023): "Mortgage Refinancing, Consumer Spending, and Competition: Evidence from the Home Affordable Refinance Program," *Review of Economic Studies*, 90, 499–537.

AGARWAL, S., S. CHOMSISENGPHET, H. KIEFER, L. C. KIEFER, AND P. C. MEDINA (2024): "Refinancing Inequality During the COVID-19 Pandemic," *Journal of Financial and Quantitative Analysis*, 59, 2133–2163.

AGARWAL, S., J. C. DRISCOLL, AND D. I. LAIBSON (2013): "Optimal Mortgage Refinancing: A Closed-Form Solution," *Journal of Money, Credit and Banking*, 45, 591–622.

AGARWAL, S., R. J. ROSEN, AND V. YAO (2016): "Why Do Borrowers Make Mortgage Refinancing Mistakes?" *Management Science*, 62, 3494–3509.

AMBROSE, B. W., J. N. CONKLIN, AND L. A. LOPEZ (2021): "Does Borrower and Broker Race Affect the Cost of Mortgage Credit?" *Review of Financial Studies*, 34, 790–826.

ANDERSEN, S., J. Y. CAMPBELL, K. M. NIELSEN, AND T. RAMADORAI (2020): "Sources of Inaction in Household Finance: Evidence from the Danish Mortgage Market," *American Economic Review*, 110, 3184–3230.

BARROT, J.-N. AND R. NANDA (2020): "The Employment Effects of Faster Payment: Evidence from the Federal Quickpay Reform," *Journal of Finance*, 75, 3139–3173.

BAYER, P., F. FERREIRA, AND S. L. ROSS (2016): "The Vulnerability of Minority Homeowners in the Housing Boom and Bust," *American Economic Journal: Economic Policy*, 8, 1–27.

BERAJA, M., A. FUSTER, E. HURST, AND J. VAVRA (2019): "Regional Heterogeneity and the Refinancing Channel of Monetary Policy," *The Quarterly Journal of Economics*, 134, 109–183.

BERGER, D., K. MILBRADT, F. TOURRE, AND J. VAVRA (2021): "Mortgage Prepayment and Path-Dependent Effects of Monetary Policy," *American Economic Review*, 111, 2829–2878.

BERGER, D., K. MILBRADT, J. VAVRA, AND F. TOURRE (2024): "Refinancing Frictions, Mortgage Pricing and Redistribution," Working Paper.

——— (Forthcoming): "Optimal Mortgage Refinancing with Inattention," *American Economic Review: Insights*.

BERKOVEC, J. A., G. B. CANNER, S. A. GABRIEL, AND T. H. HANNAN (1997): "Mortgage Discrimination and FHA Loan Performance," in *Mortgage Lending, Racial Discrimination and Federal Policy*, Routledge, 1st edition ed.

——— (1998): "Discrimination, Competition, and Loan Performance in FHA Mortgage Lending," *Review of Economics and Statistics*, 80, 241–250.

BHUTTA, N. AND V. Z. DOUBINKO (2025): "Do Mortgage Borrowing Experiences Differ by Race and Ethnicity? Evidence from the National Survey of Mortgage Originations," CFI Special Report, FRB of Philadelphia.

BUTLER, A. W., E. J. MAYER, AND J. P. WESTON (2022): "Racial Disparities in the Auto Loan Market," *Review of Financial Studies*, 36, 1–41.

CARVALHO, D., J. GAO, AND P. MA (2023): "Loan Spreads and Credit Cycles: The Role of Lenders' Personal Economic Experiences," *Journal of Financial Economics*, 148, 118–149.

CHARLES, K. K. AND J. GURYAN (2008): "Prejudice and Wages: An Empirical Assessment of Becker's *The Economics of Discrimination*," *Journal of Political Economy*, 116, 773–809.

CHERNENKO, S. AND A. SUNDERAM (2016): "Liquidity Transformation in Asset Management: Evidence from the Cash Holdings of Mutual Funds," NBER Working Paper 22391, National Bureau of Economic Research.

CHOI, D. B., H.-S. CHOI, AND J.-E. KIM (2022): "Clogged Intermediation: Were Home Buyers Crowded Out?" *Journal of Money, Credit and Banking*, 54, 1065–1098.

CLAPP, J. M., G. M. GOLDBERG, J. P. HARDING, AND M. LACOUR-LITTLE (2001): "Movers and Shuckers: Interdependent Prepayment Decisions," *Real Estate Economics*, 29, 411–450.

COX, J. C., J. E. INGERSOLL, AND S. A. ROSS (1985): "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385–407.

DEFUSCO, A. A. AND J. MONDRAGON (2020): "No Job, No Money, No Refi: Frictions to Refinancing in a Recession," *Journal of Finance*, 75, 2327–2376.

DENG, Y. AND S. GABRIEL (2006): "Risk-Based Pricing and the Enhancement of Mortgage Credit Availability among Underserved and Higher Credit-Risk Populations," *Journal of Money, Credit and Banking*, 38, 1431–1460.

DITTMAR, A. AND R. DUCHIN (2016): "Looking in the Rearview Mirror: The Effect of Managers' Professional Experience on Corporate Financial Policy," *Review of Financial Studies*, 29, 565–602.

DJANKOV, S., C. FREUND, AND C. S. PHAM (2010): "Trading on Time," *Review of Economics and Statistics*, 92, 166–173.

DONIGER, C. L. AND B. KAY (2023): "Long-Lived Employment Effects of Delays in Emergency Financing for Small Businesses," *Journal of Monetary Economics*, 140, 78–91.

DOUGAL, C., P. GAO, W. J. MAYEW, AND C. A. PARSONS (2019): "What's in a (School) Name? Racial Discrimination in Higher Education Bond Markets," *Journal of Financial Economics*, 134, 570–590.

EICHENBAUM, M., S. REBELO, AND A. WONG (2022): "State-Dependent Effects of Monetary Policy: The Refinancing Channel," *American Economic Review*, 112, 721–761.

FANNIE MAE (2014): "What Motivates Borrowers to Refinance? Past Refinancing Behavior and Future Refinancing Intent," National Housing Survey Topic Analysis.

FARRE-MENSA, J., D. HEGDE, AND A. LJUNGQVIST (2020): "What Is a Patent Worth? Evidence from the U.S. Patent "Lottery"," *Journal of Finance*, 75, 639–682.

FIRESTONE, S., R. VAN ORDER, AND P. ZORN (2007): "The Performance of Low-Income and Minority Mortgages," *Real Estate Economics*, 35, 479–504.

FISHER, J., A. GAVAZZA, L. LIU, T. RAMADORAI, AND J. TRIPATHY (2024): "Refinancing Cross-Subsidies in the Mortgage Market," *Journal of Financial Economics*, 158, 103876.

FONSECA, J. AND L. LIU (2024): "Mortgage Lock-In, Mobility, and Labor Reallocation," *Journal of Finance*, 79, 3729–3772.

FRAME, W. S., R. HUANG, E. X. JIANG, Y. LEE, W. S. LIU, E. J. MAYER, AND A. SUNDERAM (forthcoming): "The Impact of Minority Representation at Mortgage Lenders," *Journal of Finance*.

FRAZIER, N. AND R. GOODSTEIN (2023): "Is There Crowd Out in Mortgage Refinance?" Working Paper.

FUSTER, A., A. HIZMO, L. LAMBIE-HANSON, J. VICKERY, AND P. WILLEN (2024): "How Resilient Is Mortgage Credit Supply? Evidence from the COVID-19 Pandemic," Working Paper.

FUSTER, A., M. PLOSSER, P. SCHNABL, AND J. VICKERY (2019): "The Role of Technology in Mortgage Lending," *Review of Financial Studies*, 32, 1854–1899.

GAO, J., Y. WU, AND W. ZHANG (2024): "Do Local Branches Shape Banks' Mortgage Lending Decisions?" Working Paper.

GERARDI, K., P. S. WILLEN, AND D. H. ZHANG (2023): "Mortgage Prepayment, Race, and Monetary Policy," *Journal of Financial Economics*, 147, 498–524.

GOODMAN, L. S. AND C. MAYER (2018): "Homeownership and the American Dream," *Journal of Economic Perspectives*, 32, 31–58.

GREENWALD, D. (2018): "The Mortgage Credit Channel of Macroeconomic Transmission," .

GRUNDL, S. AND Y. S. KIM (2019): "Consumer Mistakes and Advertising: The Case of Mortgage Refinancing," *Quantitative Marketing and Economics*, 17, 161–213.

HAN, L. AND S.-H. HONG (2024): "Cash Is King? Understanding Financing Risk in Housing Markets," *Review of Finance*, 28, 2083–2118.

HU, L., K. LI, P. T. H. NGO, AND D. SOSYURA (2024): "Media as a Money Doctor: Evidence from Refinancing Decisions," Working Paper.

HUANG, R., S. TITMAN, E. J. MAYER, AND D. X. XU (2024): "Human Capital and Local Credit Supply: Evidence from the Mortgage Industry," Working Paper.

JOHNSON, E. J., S. MEIER, AND O. TOUBIA (2019): "What's the Catch? Suspicion of Bank Motives and Sluggish Refinancing," *Review of Financial Studies*, 32, 467–495.

KERMANI, A. AND F. WONG (2024): "Racial Disparities in Housing Returns," Working Paper.

KEYS, B. J., D. G. POPE, AND J. C. POPE (2016): "Failure to Refinance," *Journal of Financial Economics*, 122, 482–499.

KILLEWALD, A., F. T. PFEFFER, AND J. N. SCHACHNER (2017): "Wealth Inequality and Accumulation," *Annual Review of Sociology*, 43, 379–404.

KOUDIJS, P. AND H.-J. VOTH (2016): "Leverage and Beliefs: Personal Experience and Risk-Taking in Margin Lending," *American Economic Review*, 106, 3367–3400.

MALMENDIER, U., S. NAGEL, AND Z. YAN (2021): "The Making of Hawks and Doves," *Journal of Monetary Economics*, 117, 19–42.

MALMENDIER, U., G. TATE, AND J. YAN (2011): "Overconfidence and Early-Life Experiences: The Effect of Managerial Traits on Corporate Financial Policies," *Journal of Finance*, 66, 1687–1733.

SCHARLEMANN, T. AND E. VAN STRAELEN (2024): "More Tax, Less Refi? The Mortgage Interest Deduction and Monetary Policy Pass-Through," Finance and Economics Discussion Series 2024-082, Board of Governors of the Federal Reserve System, Washington.

STEPHENS-DAVIDOWITZ, S. (2014): "The Cost of Racial Animus on a Black Candidate: Evidence Using Google Search Data," *Journal of Public Economics*, 118, 26–40.

VOICU, I. (2018): "Using First Name Information to Improve Race and Ethnicity Classification," *Statistics and Public Policy*, 5, 1–13.

WEI, B. AND F. ZHAO (2022): "Racial Disparities in Mortgage Lending: New Evidence Based on Processing Time," *Review of Corporate Finance Studies*, 11, 775–813.

ZHANG, D. (2024): "Closing Costs, Refinancing, and Inefficiencies in the Mortgage Market," Working Paper.

Figure 1. Breakdown of Sub-Issues in Mortgage-Related Complaints from the CFPB Consumer Complaint Database

This figure displays the distribution of sub-issues within mortgage application and mortgage closing complaints in the CFPB Consumer Complaint Database for 2024. The upper bar ("Original") represents the unadjusted share of each sub-issue. The lower bar ("Adjusted") reclassifies all complaints containing the keywords *delay*, *late*, or *postpone* under the "Delay" sub-issue.

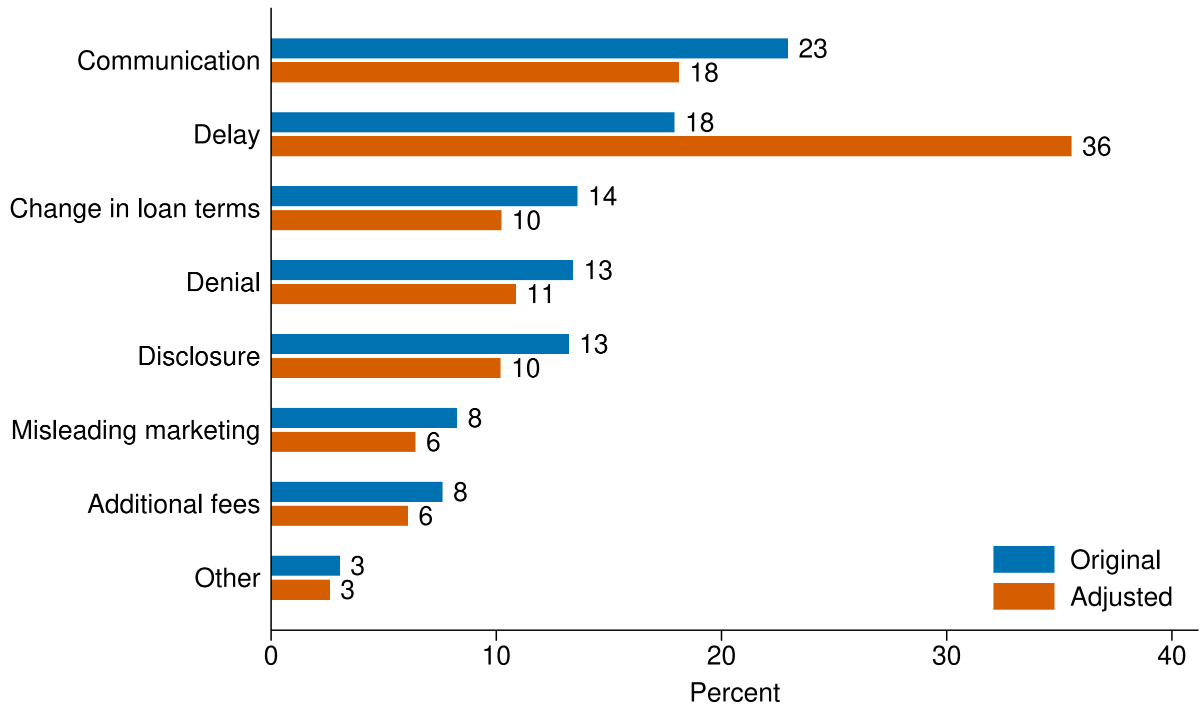
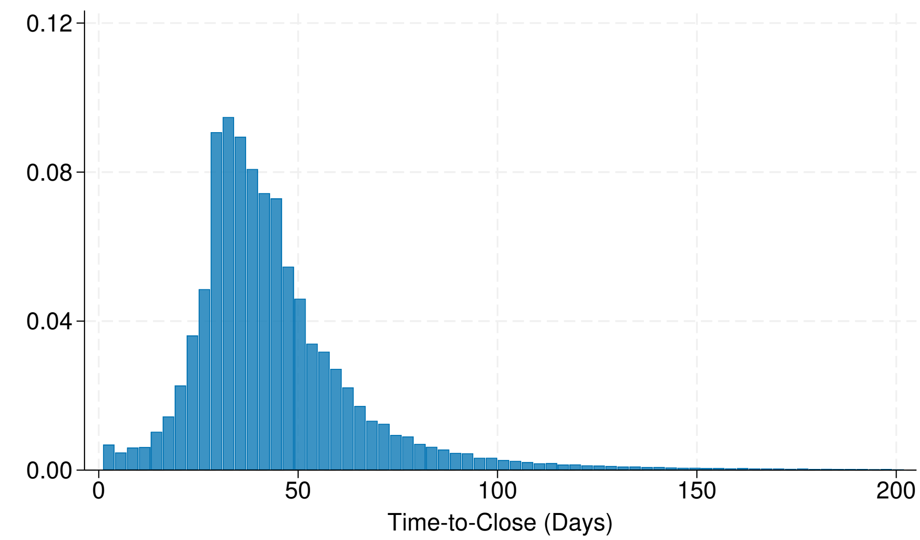
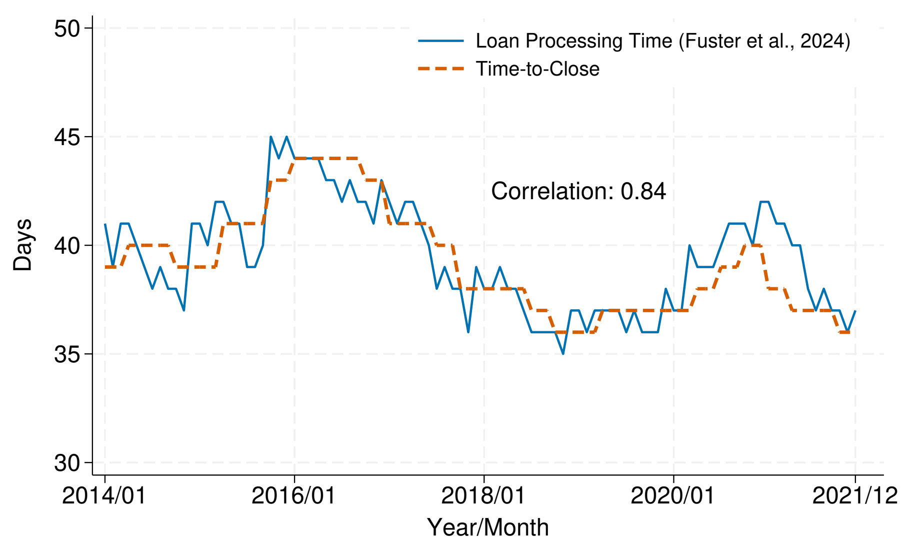


Figure 2. Cross-Sectional Distribution and Time-Series Trends of *Time-To-Close*

This figure illustrates the cross-sectional distribution and time-series trends of the *Time-To-Close* variable in the matched CoreLogic–MBS dataset. Panel (a) presents the cross-sectional distribution of *Time-To-Close*. Panel (b) depicts the quarterly time-series of the median *Time-To-Close* from 2014 to 2021, alongside the monthly median loan processing time for purchase mortgages as reported in Figure A.8 of Fuster et al. (2024).



(a) Cross-Sectional Distribution of *Time-To-Close*

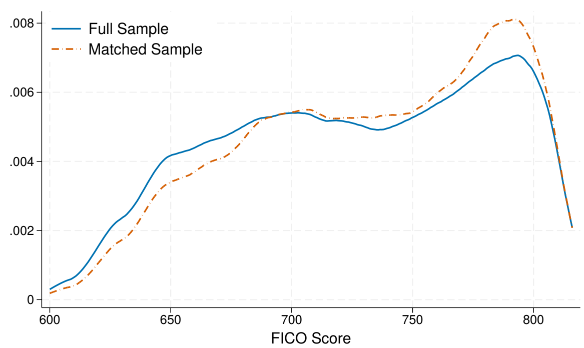


(b) Time-Series of Median *Loan Processing Time* (from Fuster et al. (2024)) and *Time-To-Close*

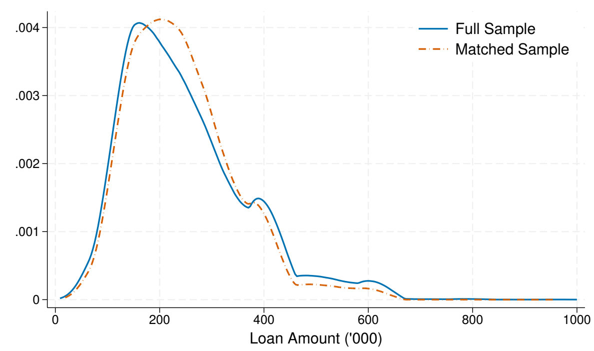
Figure 3. Kernel Density Plot of Key Variables

This figure compares the distributions of key variables—*FICO Score*, *Loan Amount*, *DTI Ratio*, and *LTV Ratio*—in the full sample with those in the matched CoreLogic–MBS dataset using kernel density plots from a 2015 snapshot. Panel (a) presents the combined GSE and FHA sample, while Panels (b) and (c) show the GSE and FHA samples separately.

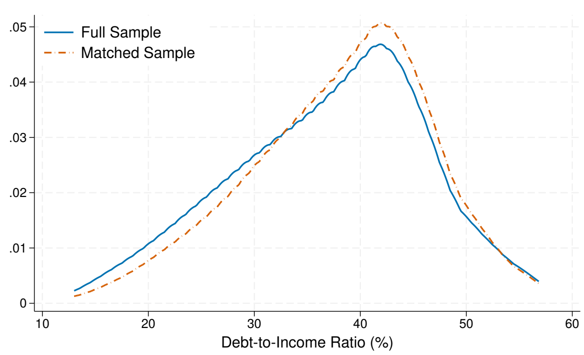
(a) GSE + FHA Sample



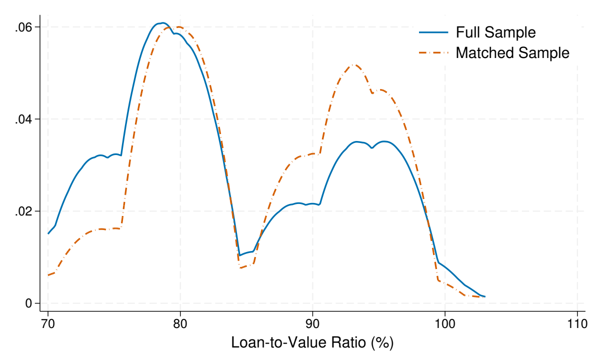
FICO Score



Loan Amount

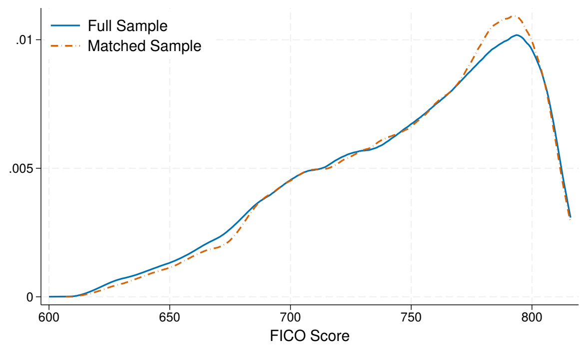


DTI Ratio

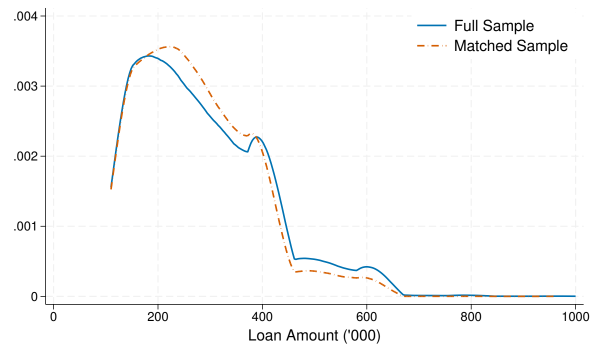


LTV Ratio

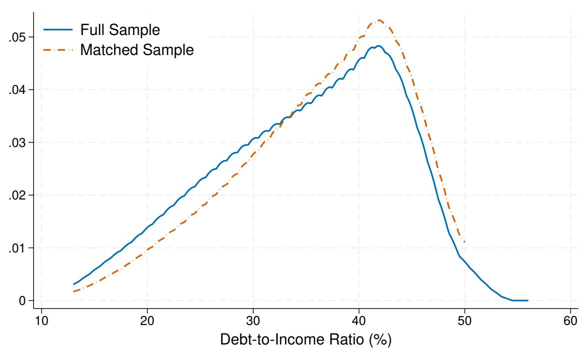
(b) GSE Sample



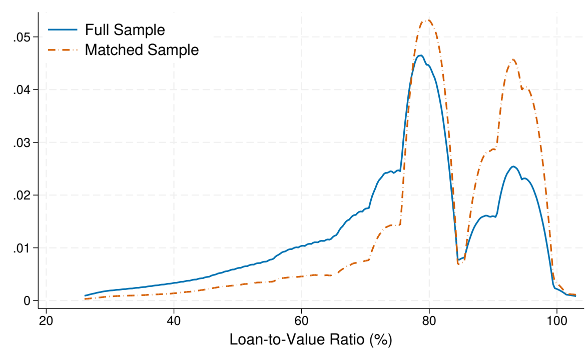
FICO Score



Loan Amount

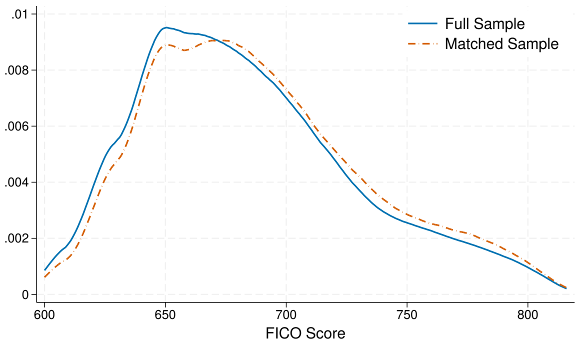


DTI Ratio

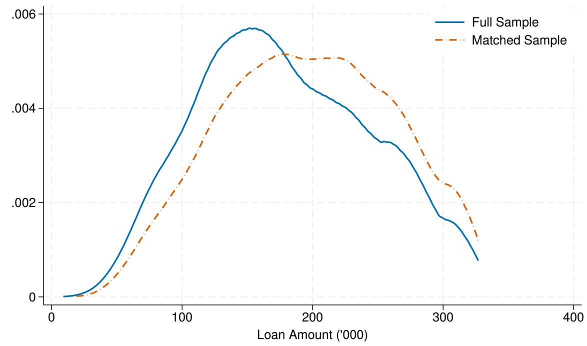


LTV Ratio

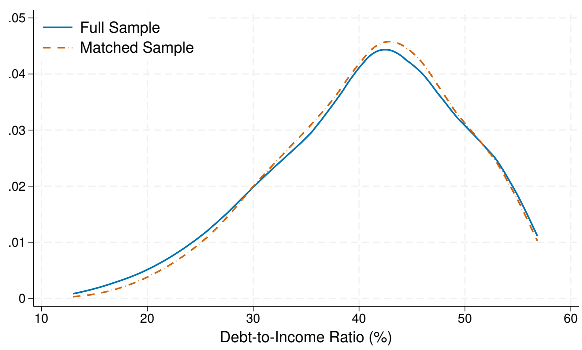
(c) FHA Sample



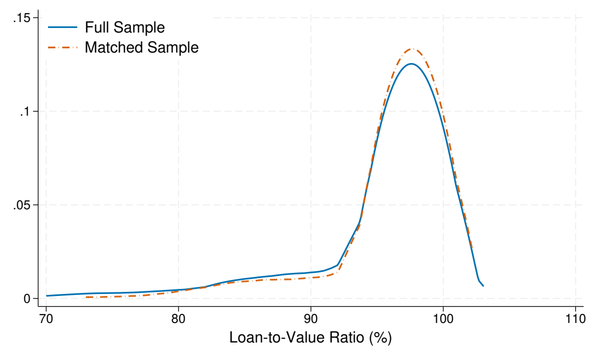
FICO Score



Loan Amount



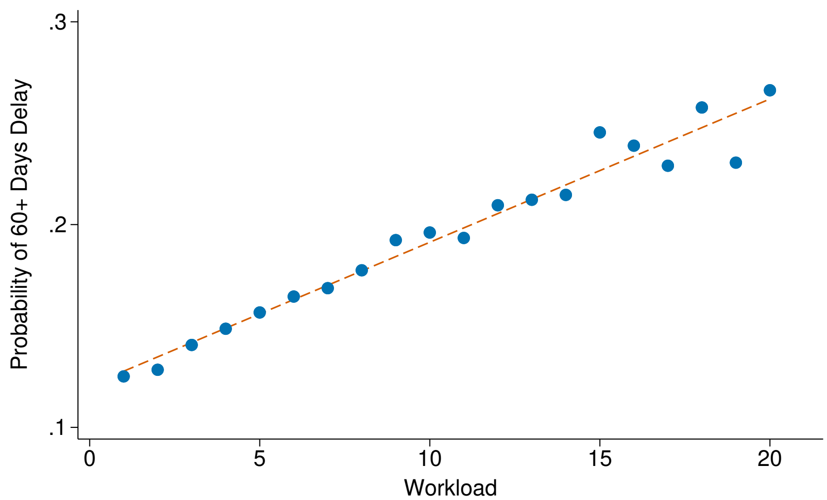
DTI Ratio



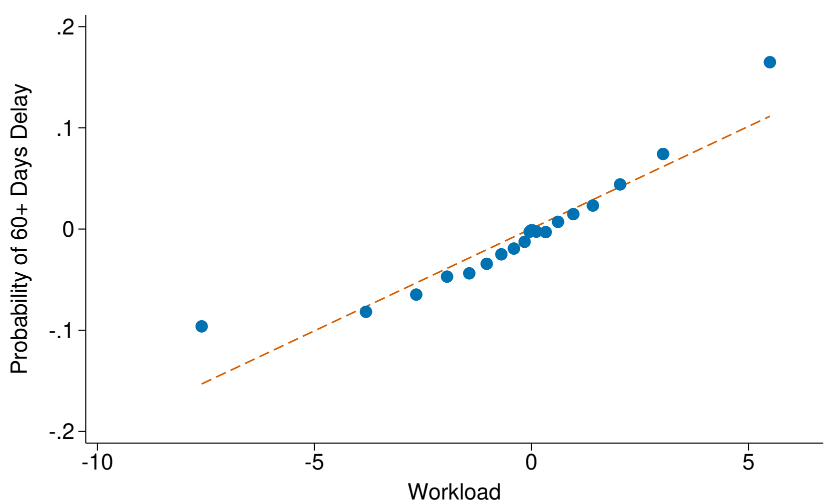
LTV Ratio

Figure 4. Loan Officer Workload and Probability of 60+ Day Loan Closing Delay

This figure presents a binned scatter plot of I(*Time-To-Close* > 60 Days) against loan officer *Workload*. Panel (a) shows a binned scatter plot using the raw values of I(*Time-To-Close* > 60 Days) and *Workload*. Panel (b) presents the relationship after residualizing both I(*Time-To-Close* > 60 Days) and *Workload* by the full set of borrower and loan characteristics, squared terms, and fixed effects for borrower age group, county-by-origination-year, year-quarter, and loan officer.



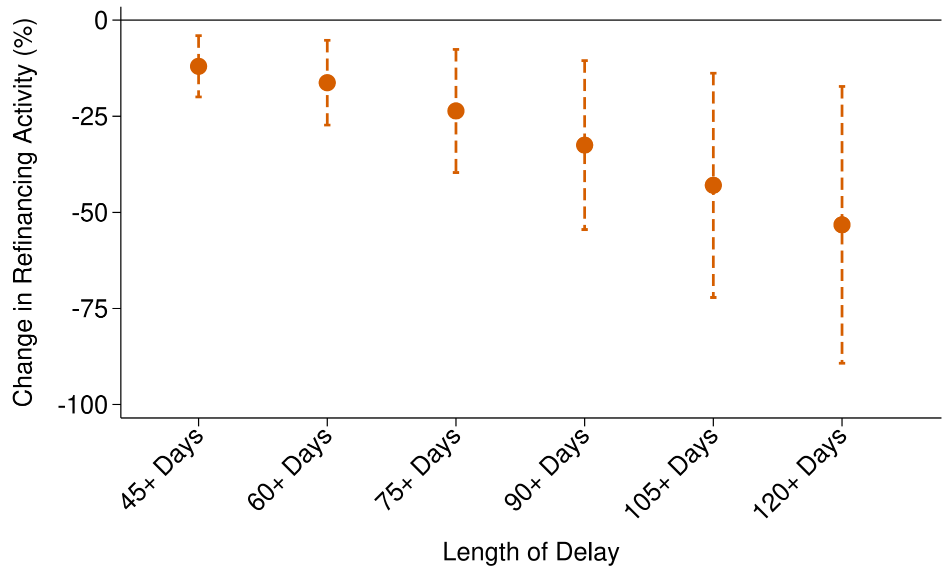
(a) Raw Binned Scatter Plot



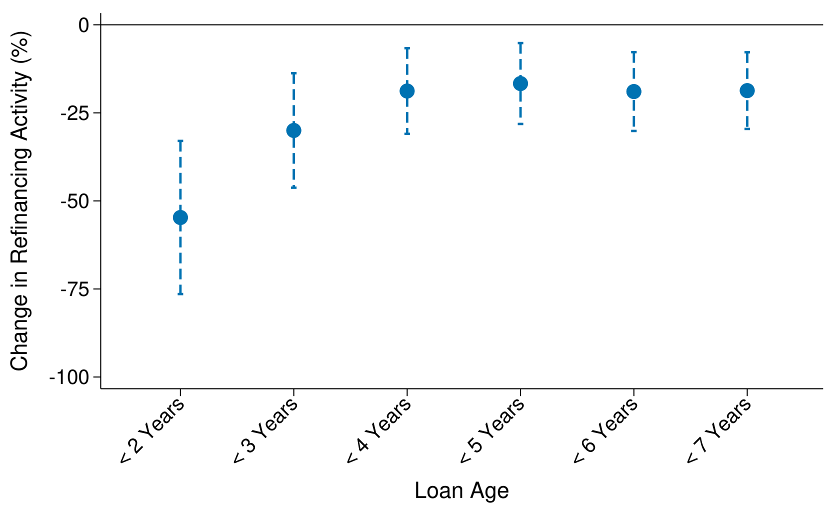
(b) Residualized Binned Scatter Plot

Figure 5. Percentage Change in Refinancing Rates by Length of Closing Delay and Loan Age Subgroup

Panel (a) presents the percentage change in refinancing rates from an IV regression of *Refinance* on various lengths of closing delays. Panel (b) displays the percentage changes from IV regressions across different loan age subgroups. Percentage changes are calculated by dividing the coefficient estimates by the mean quarterly refinancing rate. All specifications follow column (2) of Table 4.



(a) Coefficient Estimates by Length of Closing Delay



(b) Coefficient Estimates by Loan Age

Figure 6. Discounted PV of Overpayment by Timing of Refinancing Opportunity

This figure illustrates how the PV of overpayment varies depending on the timing of a missed refinancing opportunity over a 30-year loan term. The overpayment is based on a back-of-the-envelope estimate of $586.5 in annual losses due to delayed refinancing. The PV is computed under the assumption of quarterly payments and a 3% annual discount rate. Missed opportunities earlier in the loan term result in substantially higher discounted losses.

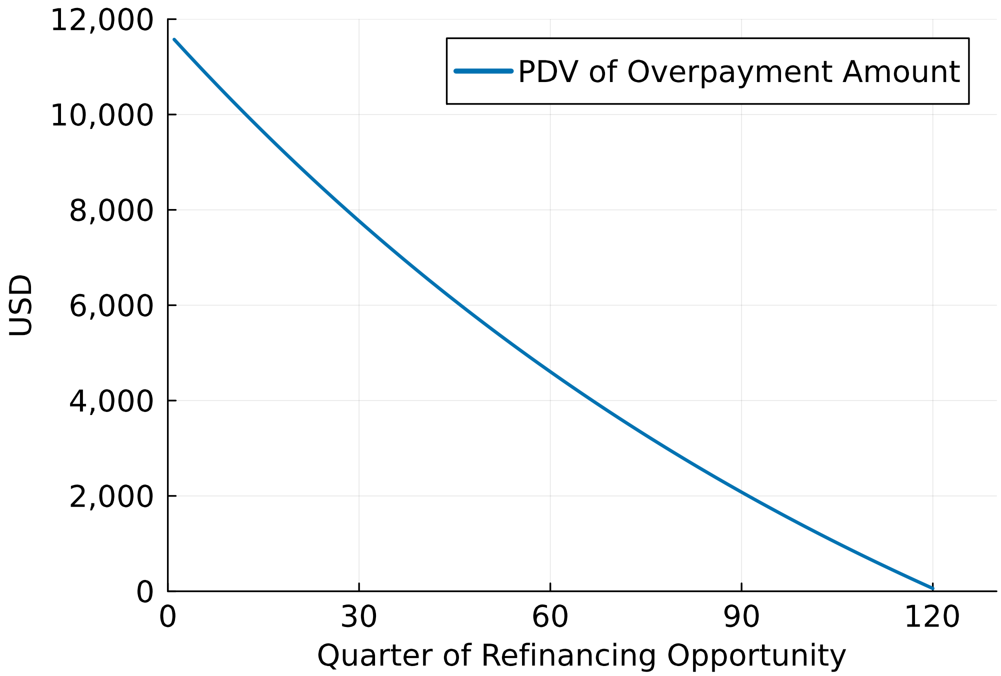


Figure 7. Mortgage Rate Spread Over Time

This figure displays the mortgage spread (i.e., the difference between the mortgage rate and short-term Treasury rate) over the period 2013–2021.

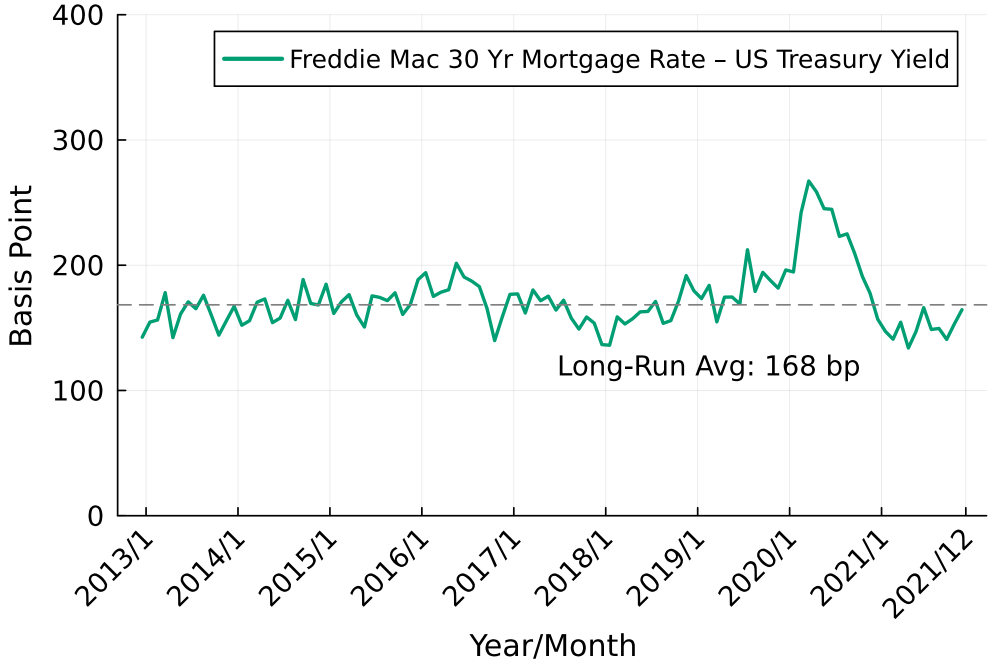
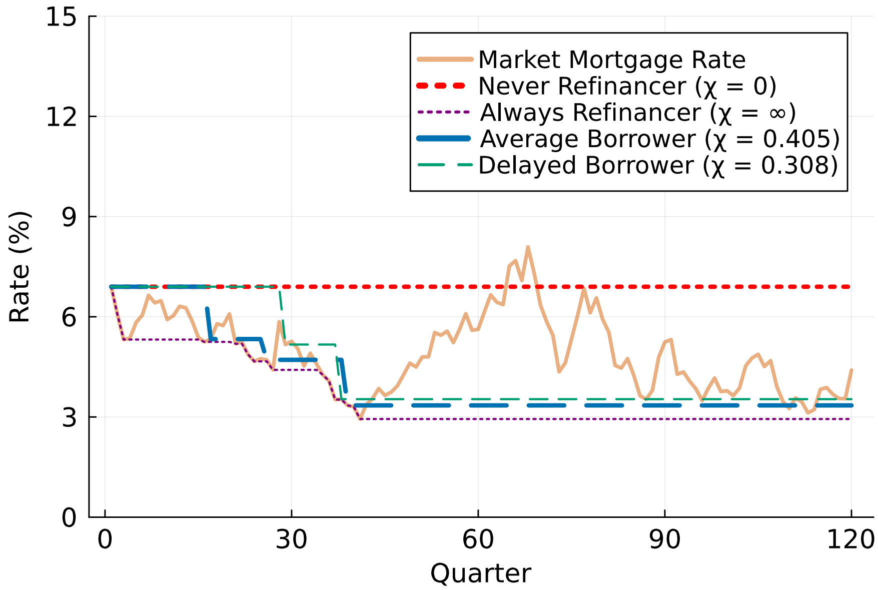
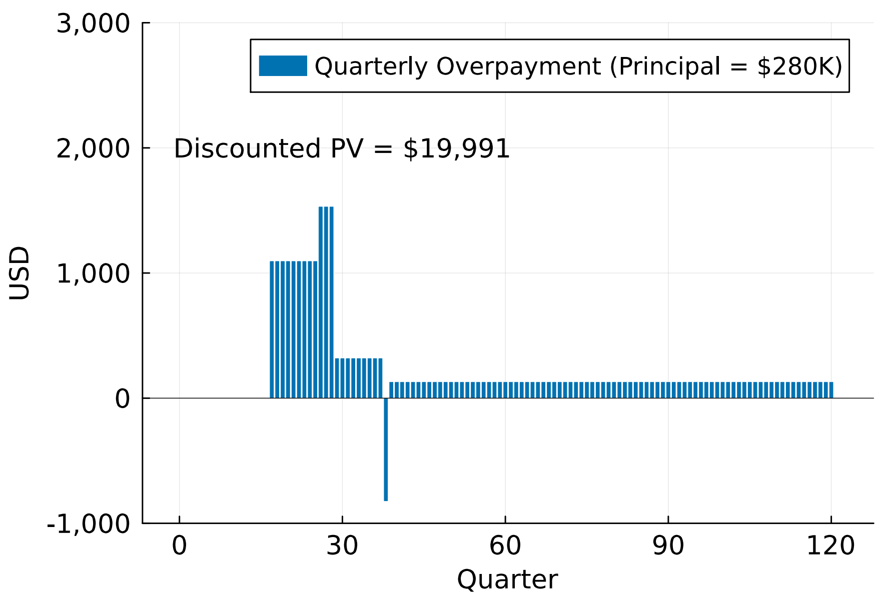


Figure 8. Simulated Coupon Paths and Overpayment Due to Delays

This figure presents an illustrative example of simulated mortgage rate dynamics. Panel (a) shows coupon rate trajectories for borrowers with different refinancing responsiveness parameters (χ), plotted alongside the simulated market mortgage rate, which is derived from a CIR short-rate process with a fixed spread. Panel (b) computes the overpayment incurred by delayed borrowers relative to baseline (non-delayed) borrowers by multiplying the coupon rate differential by the average loan size over time.



(a) Simulated Coupon Rate Paths and Market Mortgage Rates



(b) Overpayment of Delayed Borrowers Relative to Average (Non-Delayed) Borrowers

Figure 9. Distribution of PV of Overpayment

This figure presents the distribution of the PV of overpayment resulting from initial origination delays across simulation runs. Overpayment is calculated by comparing realized coupon payments for delayed borrowers to those of average borrowers, discounted at an annual rate of 3%.

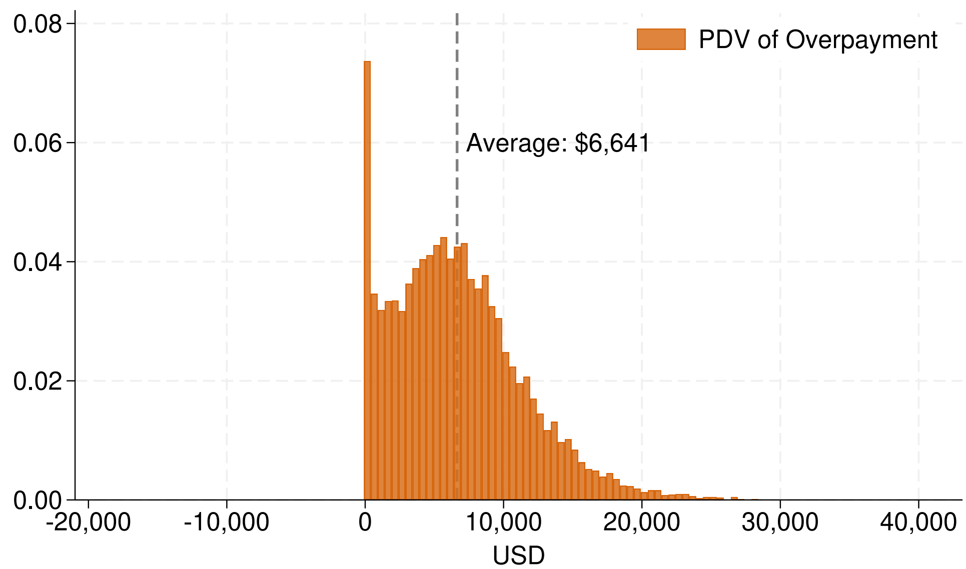


Figure 10. PV of Overpayment from Origination Delays Across χ Values

This figure shows how the average PV of overpayment resulting from initial origination delays varies with different refinancing friction parameter χ.

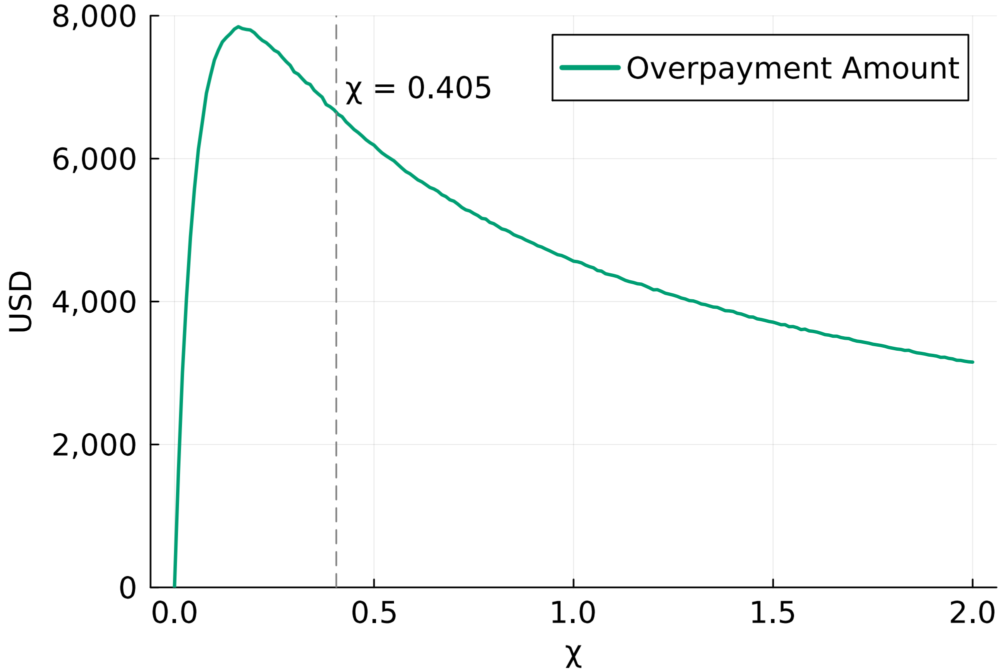
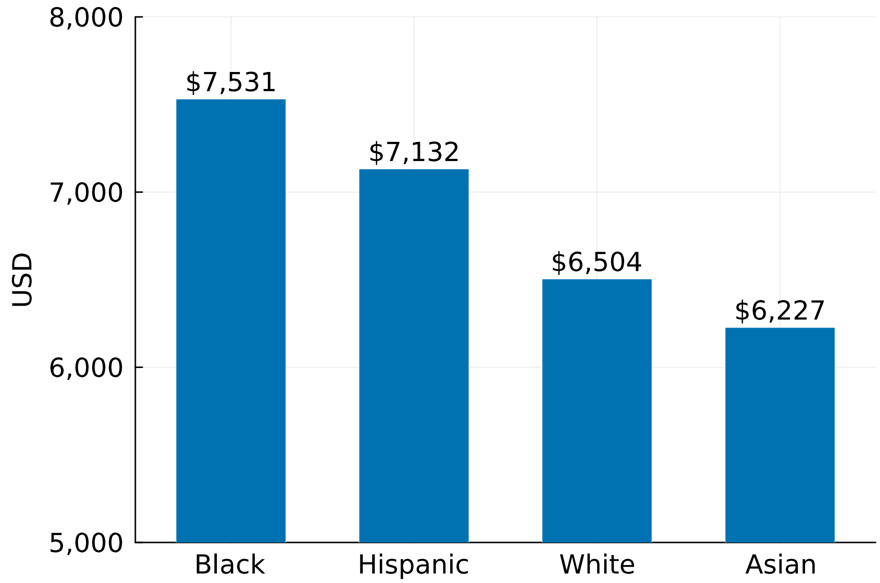
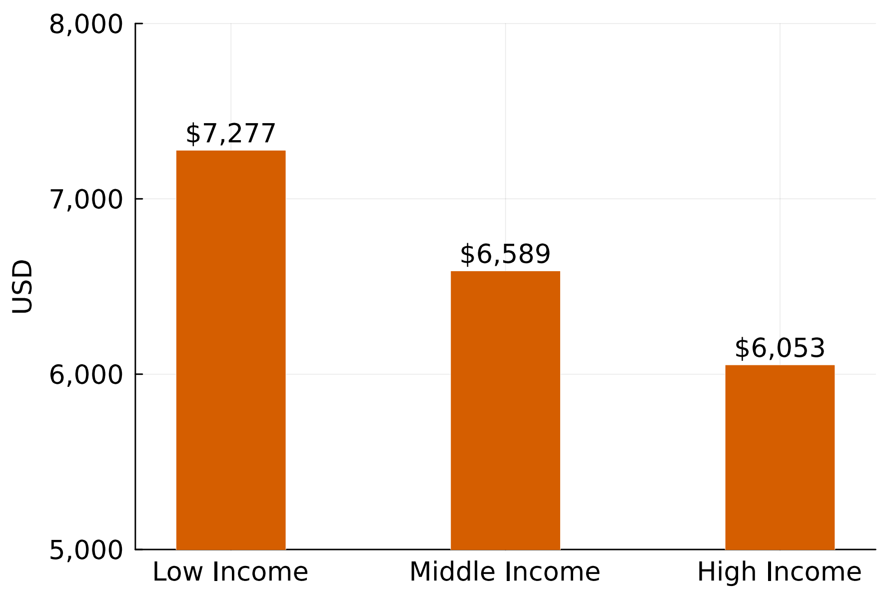


Figure 11. Average PV of Overpayment by Race and Income Groups

This figure reports the average PV of overpayment due to origination delays across borrower subgroups, using separately estimated refinancing responsiveness parameters (χ) from Table 10. Panel (a) shows average PV by race. Panel (b) shows average PV by income group.



(a) By Race



(b) By Income Group

Table 1. Summary Statistics

This table reports summary statistics for the matched panel dataset combining CoreLogic with Fannie Mae, Freddie Mac, and Ginnie Mae MBS Loan-Level Dataset. Panel (a) presents statistics from the quarterly loan panel, where each loan appears multiple times over time. Panel (b) provides loan-level summary statistics, with a single observation per loan at origination.

(a) Quarterly Loan Panel

	Obs.	Mean	S.D.	P25	P50	P75
Refinance	5,883,962	3.02	17.10	0.00	0.00	0.00
Recapture Refinance	5,883,962	0.95	9.68	0.00	0.00	0.00
Switching Refinance	5,883,962	2.07	14.24	0.00	0.00	0.00
Cash-Out Refinance	5,883,962	1.19	10.86	0.00	0.00	0.00
Recapture Cash-Out Refinance	5,883,962	0.35	5.87	0.00	0.00	0.00
Switching Cash-Out Refinance	5,883,962	0.85	9.17	0.00	0.00	0.00
Prepaid Due to Selling and Moving	5,883,962	1.47	12.03	0.00	0.00	0.00
$\mathbb{1}$(*Time-To-Close* $>$ 60 Days)	5,883,962	0.11	0.31	0.00	0.00	0.00
White	5,883,962	0.69	0.46	0.00	1.00	1.00
Minority	5,883,962	0.27	0.44	0.00	0.00	1.00
Black	5,883,962	0.07	0.25	0.00	0.00	0.00
Hispanic	5,883,962	0.20	0.40	0.00	0.00	0.00
Asian	5,883,962	0.04	0.21	0.00	0.00	0.00
Other Race	5,883,962	0.00	0.04	0.00	0.00	0.00
Female	5,883,962	0.33	0.47	0.00	0.00	1.00
Coborrower	5,883,962	0.48	0.50	0.00	0.00	1.00
First-Time Home Buyer	5,883,962	0.54	0.50	0.00	1.00	1.00
FHA	5,883,962	0.62	0.49	0.00	1.00	1.00
ln(*Income*)	5,883,962	8.11	0.55	7.74	8.15	8.52
ln(*Loan Amount*)	5,883,962	12.47	0.54	12.12	12.52	12.87
LTV at Origination (%)	5,883,962	87.59	13.55	80.00	92.00	97.00
Current LTV (%)	5,883,962	73.06	15.81	63.18	74.98	85.41
FICO	5,883,962	730.61	54.78	688.00	737.00	779.00
Loan Age	5,883,962	7.42	6.36	2.00	6.00	11.00
Rate Gap (%)	5,883,962	-0.07	1.00	-0.56	-0.03	0.54
Workload	5,883,962	5.32	6.65	1.00	3.00	7.00

(b) Loan-Level Dataset

	Obs.	Mean	S.D.	P25	P50	P75
Time-To-Close	435,288	40.20	21.01	30.00	37.00	46.00
$\mathbb{1}$(*Time-To-Close* > 60 Days)	435,288	0.10	0.30	0.00	0.00	0.00
White	435,288	0.69	0.46	0.00	1.00	1.00
Minority	435,288	0.26	0.44	0.00	0.00	1.00
Black	435,288	0.04	0.19	0.00	0.00	0.00
Hispanic	435,288	0.23	0.42	0.00	0.00	0.00
Asian	435,288	0.04	0.20	0.00	0.00	0.00
Other Race	435,288	0.00	0.04	0.00	0.00	0.00
Female	435,288	0.33	0.47	0.00	0.00	1.00
Coborrower	435,288	0.49	0.50	0.00	0.00	1.00
First-Time Home Buyer	435,288	0.53	0.50	0.00	1.00	1.00
FHA	435,288	0.26	0.44	0.00	0.00	1.00
ln(*Income*)	435,288	8.17	0.55	7.80	8.21	8.56
ln(*Loan Amount*)	435,288	12.54	0.53	12.20	12.59	12.92
LTV (%)	435,288	87.30	13.26	80.00	91.32	97.00
FICO	435,288	732.65	81.15	691.00	740.00	779.00

Table 2. OLS Regression Results: Impact of Initial Mortgage Delays on Refinancing Behavior

This table presents the OLS regression results examining the effect of initial mortgage delays on refinancing activities. The analysis is based on quarterly loan performance observations from the CoreLogic–MBS dataset, covering loans originated between 2014 and 2021. In columns (1)–(4), I use the full sample of GSE and FHA loans. In columns (5) and (6), I use the GSE loan subsample. In columns (7) and (8), I use the FHA loan subsample. The t-statistics are reported in parentheses and all standard errors are clustered at the county and year level. ***, **, and * denote the significance of the parameter estimates at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Refinance							
	GSE + FHA Sample				GSE Sample		FHA Sample	
$\mathbb{1}$(Time-To-Close > 60 Days)	-0.1469***	-0.1307***	-0.1048***	-0.1195**	-0.1291**	-0.1855**	-0.1061***	-0.0815**
	(-3.93)	(-3.90)	(-3.07)	(-2.54)	(-2.17)	(-2.01)	(-3.97)	(-2.20)
Minority	-0.5622***	-0.5329***	-0.4183***	-0.3872***	-0.4159***	-0.3548***	-0.3871***	-0.3698***
	(-8.28)	(-8.39)	(-10.28)	(-8.66)	(-5.52)	(-3.50)	(-7.11)	(-6.56)
Asian	0.6466***	0.3822***	0.3298***	0.0333	0.3919***	0.0583	-0.1355	-0.5168***
	(3.98)	(3.04)	(2.74)	(0.30)	(2.82)	(0.25)	(-1.19)	(-4.18)
Female	-0.0263	-0.0184	-0.0356	-0.0217	-0.0184	0.0590	-0.0699**	-0.1007***
	(-1.37)	(-0.91)	(-1.43)	(-0.52)	(-0.47)	(0.94)	(-2.46)	(-3.10)
Coborrower	0.2075***	0.2077***	0.2041***	0.1959***	0.2639***	0.3251***	0.0468**	-0.0154
	(4.56)	(4.96)	(7.11)	(6.69)	(6.62)	(5.29)	(2.12)	(-0.54)
First-Time Home Buyer	0.0763	0.0762	0.0166	-0.0580	0.2007***	0.2260**	-0.3306***	-0.4178***
	(1.47)	(1.45)	(0.47)	(-1.19)	(4.39)	(2.53)	(-9.24)	(-9.80)
ln(Income)	-4.6802***	-4.5629***	-5.8399***	-5.4183***	-4.0497***	-2.7294	-6.2155***	-5.9807***
	(-7.45)	(-7.97)	(-8.32)	(-5.78)	(-3.18)	(-1.44)	(-7.32)	(-5.17)
ln(Loan Amount)	-5.3441**	-6.5072***	-4.7784**	0.0852	-5.4025*	-7.2756*	-2.7830	4.5252*
	(-2.32)	(-2.84)	(-2.34)	(0.04)	(-1.70)	(-1.77)	(-1.46)	(1.79)
LTV at Origination	-0.1968***	-0.1932***	-0.1662***	-0.1633***	-0.3902***	-0.4339***	-0.0556**	-0.0804
	(-6.32)	(-5.91)	(-5.27)	(-3.97)	(-8.41)	(-6.83)	(-2.00)	(-1.59)
Current LTV	0.1094***	0.1030***	0.0586*	0.0249	0.3668***	0.4173***	0.1373***	0.1776***
	(4.19)	(3.93)	(1.93)	(0.57)	(8.63)	(6.66)	(5.02)	(5.31)
FICO	0.0170	0.0168*	0.0168**	0.0043	0.0661***	0.0451	0.0361***	0.0417***
	(1.62)	(1.86)	(2.04)	(0.56)	(3.72)	(1.41)	(7.86)	(6.51)
Loan Age	0.5946***	0.6195***	0.7651***	0.9424***	0.7876***	0.9202***	0.3890***	0.5172***
	(15.69)	(14.91)	(11.80)	(10.27)	(10.48)	(8.29)	(8.17)	(6.95)
Rate Gap	1.5575***	1.5543***	1.6480***	1.6780***	2.1060***	2.2191***	1.7035***	1.9400***
	(12.27)	(12.31)	(13.10)	(17.71)	(12.33)	(15.09)	(18.72)	(20.57)
FHA	-0.7448***	-0.8001***	-0.9270***	-1.1961***				
	(-5.07)	(-5.54)	(-7.66)	(-14.44)				
Square Terms of Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County × Origin. Year FE	Yes	Yes	Yes	-	Yes	-	Yes	-
Tract × Origin. Year FE	-	-	-	Yes	-	Yes	-	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lender FE	-	Yes	-	-	-	-	-	-
Loan Officer FE	-	-	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	3.016	3.016	3.016	3.016	3.413	3.412	2.030	2.030
R-Squared	0.051	0.054	0.081	0.115	0.102	0.147	0.054	0.090
Obs.	5,883,962	5,883,962	5,883,962	5,883,876	2,230,114	2,230,044	3,653,833	3,653,804

Table 3. Validation Tests for Instrumental Variable

This table presents regression results assessing the relevance and exclusion conditions of the instrument, *Workload*. Columns (1) and (2) report the first-stage regression results, demonstrating the relationship between *Workload* and the likelihood of loan closing delays. Columns (3) and (4) present covariate balance test results, where the dependent variable is *Workload*, and the independent variables include covariates used in the IV regressions. The t-statistics are reported in parentheses and all standard errors are clustered at the county and year level. ***, **, and * denote the significance of the parameter estimates at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
	$\mathbb{1}$(*Time-To-Close* > 60 Days)		*Workload*	
Workload	0.0165***	0.0164***		
	(18.37)	(18.84)		
Minority	0.0150***	0.0086***	0.1163	0.1105
	(6.09)	(3.92)	(1.59)	(1.39)
Asian	0.0092*	0.0133***	0.0662	0.1944
	(1.94)	(2.64)	(0.51)	(1.27)
Female	-0.0043**	-0.0024	-0.0506	0.0062
	(-2.28)	(-1.30)	(-1.03)	(0.13)
First-Time Home Buyer	-0.0082***	-0.0066***	-0.0389	-0.0128
	(-4.45)	(-3.60)	(-0.74)	(-0.26)
Coborrower	0.0015	0.0012	0.0157	0.0474
	(0.82)	(0.65)	(0.33)	(1.10)
ln(Income)	-0.2637***	-0.2842***	-0.0344	-0.0471
	(-4.71)	(-5.15)	(-0.83)	(-0.83)
ln(Loan Amount)	-0.1455	-0.2114	0.2073***	0.2740***
	(-1.26)	(-1.62)	(2.66)	(2.69)
LTV at Origination	-0.0007	-0.0008	0.0020	0.0018
	(-0.94)	(-1.00)	(0.97)	(0.68)
FICO	0.0366	0.0493	-0.0340	-0.0594
	(0.53)	(0.69)	(-0.89)	(-1.50)
FHA	0.0308***	0.0323***	0.1928***	0.1702***
	(6.91)	(7.77)	(3.83)	(2.97)
Current LTV, Loan Age, Rate Gap	Yes	Yes	-	-
Square Terms of Controls	Yes	Yes	-	-
Age Group FE	Yes	Yes	Yes	Yes
County × Origin. Year FE	Yes	-	Yes	-
Tract × Origin. Year FE	-	Yes	-	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Loan Officer FE	Yes	Yes	Yes	Yes
Dep. Var. Mean	0.114	0.114	4.643	5.014
R-Squared	0.520	0.578	0.798	0.898
First-Stage F-Statistics	27.06	27.78	-	-
Obs.	5,883,962	5,883,876	381,664	343,419

Table 4. 2SLS Regression Results: Impact of Initial Mortgage Delays on Refinancing Behavior

This table presents the 2SLS regression results examining the effect of initial mortgage delays on refinancing activities. I use *Workload* as an instrument for 60+ day loan closing delays. The analysis is based on quarterly loan performance observations from the CoreLogic–MBS dataset, covering loans originated between 2014 and 2021. In columns (1) and (2), I use the full sample of GSE and FHA loans. In columns (5) and (6), I use the GSE loan subsample. In columns (7) and (8), I use the FHA loan subsample. The t-statistics are reported in parentheses and all standard errors are clustered at the county and year level. ***, **, and * denote the significance of the parameter estimates at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
			Refinance			
	GSE + FHA Sample		GSE Sample		FHA Sample	
$\mathbb{1}$(Time-To-Close > 60 Days)	-0.4772***	-0.7312***	-0.5693*	-0.9803**	-0.3361**	-0.8271***
	(-2.71)	(-3.61)	(-1.87)	(-2.08)	(-2.23)	(-4.05)
Minority	-0.3979***	-0.4909***	-0.3492***	-0.5167***	-0.3993***	-0.3699***
	(-8.77)	(-9.52)	(-4.21)	(-3.44)	(-6.83)	(-6.13)
Asian	0.3129**	0.0604	0.3642***	-0.0837	-0.2810***	-0.3224**
	(2.25)	(0.45)	(2.89)	(-0.28)	(-2.70)	(-2.31)
Female	-0.0166	-0.0085	-0.0086	0.0918	-0.0473**	-0.0997***
	(-0.74)	(-0.30)	(-0.22)	(1.35)	(-2.11)	(-3.08)
Coborrower	0.2244***	0.2168***	0.2540***	0.3433***	0.0686***	0.0031
	(7.73)	(7.23)	(4.97)	(5.56)	(2.63)	(0.10)
First-Time Home Buyer	-0.0111	-0.1118**	0.2039***	0.1432**	-0.3602***	-0.4393***
	(-0.30)	(-2.29)	(4.34)	(2.00)	(-8.64)	(-11.64)
ln(Income)	-5.9505***	-5.3630***	-2.8463*	-2.5091	-6.6053***	-6.9011***
	(-6.97)	(-4.63)	(-1.92)	(-1.38)	(-7.86)	(-5.59)
ln(Loan Amount)	-4.1682	2.5663	-5.6503	-3.9925	-2.4717	5.1647*
	(-1.64)	(0.96)	(-1.44)	(-0.98)	(-1.29)	(1.94)
LTV at Origination	-0.0916***	-0.0398	-0.3573***	-0.3627***	-0.0265	-0.0510
	(-3.38)	(-0.93)	(-8.08)	(-8.80)	(-0.92)	(-0.77)
Current LTV	-0.0851**	-0.1561**	0.3419***	0.3743***	0.1186***	0.1465***
	(-2.41)	(-2.49)	(7.77)	(9.34)	(5.72)	(5.42)
FICO	0.0169**	0.0066	0.0582***	0.0420*	0.0412***	0.0512***
	(1.99)	(0.86)	(3.53)	(1.66)	(8.09)	(7.28)
Loan Age	0.8822***	1.0775***	0.8633***	0.9892***	0.4650***	0.5821***
	(11.34)	(9.44)	(9.60)	(12.16)	(7.58)	(6.83)
Rate Gap	1.5126***	1.4864***	2.1255***	2.1552***	1.7242***	1.9917***
	(16.67)	(22.12)	(12.46)	(15.77)	(20.12)	(24.32)
FHA	-1.0284***	-1.2470***				
	(-10.17)	(-14.93)				
Square Terms of Controls	Yes	Yes	Yes	Yes	Yes	Yes
Age Group FE	Yes	Yes	Yes	Yes	Yes	Yes
County × Origin. Year FE	Yes	-	Yes	-	Yes	-
Tract × Origin. Year FE	-	Yes	-	Yes	-	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Loan Officer FE	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	3.016	3.016	3.413	3.412	2.030	2.030
R-Squared	0.012	0.012	0.013	0.015	0.007	0.007
Obs.	5,883,962	5,883,876	2,230,114	2,230,044	3,653,833	3,653,804

Table 5. Heterogeneous Effects of Initial Mortgage Delays on Refinancing Outcomes: Recapture vs. Switching

This table presents the 2SLS regression results examining the effect of initial mortgage delays on recapture and switching refinancing activities. I use *Workload* as an instrument for loan closing delays exceeding 60 days. The analysis is based on quarterly loan performance observations from the CoreLogic–MBS dataset, covering loans originated between 2014 and 2021. In columns (1) and (2), the dependent variable is *Recapture Refinance*, which indicates refinancing by the original lender. In columns (3) and (4), the dependent variable is *Switching Refinance*, representing refinancing through a different lender. t-statistics are reported in parentheses, with standard errors clustered at the county and year level. ***, **, and * denote the significance of the parameter estimates at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
	Recapture Refinance		*Switching Refinance*	
$\mathbb{1}$(Time-To-Close > 60 Days)	-0.3512***	-0.6085***	-0.1260	-0.1227
	(-2.78)	(-4.24)	(-1.02)	(-0.87)
Minority	-0.1622***	-0.1543***	-0.2358***	-0.3366***
	(-4.99)	(-3.53)	(-6.89)	(-10.02)
Asian	-0.0626	-0.2035**	0.3755***	0.2639*
	(-1.30)	(-2.08)	(3.28)	(1.68)
Female	0.0231	0.0126	-0.0397*	-0.0211
	(1.40)	(0.68)	(-1.90)	(-0.88)
Coborrower	0.1193***	0.1279***	0.1052***	0.0889***
	(5.52)	(6.96)	(4.50)	(3.81)
First-Time Home Buyer	0.0377**	0.0116	-0.0488*	-0.1234***
	(2.08)	(0.56)	(-1.67)	(-3.18)
ln(Income)	-3.7419***	-3.9738***	-2.2086***	-1.3892
	(-10.05)	(-6.51)	(-2.76)	(-1.60)
ln(Loan Amount)	-0.0286	2.5408**	-4.1396**	0.0255
	(-0.03)	(2.17)	(-2.29)	(0.01)
LTV at Origination	-0.0851***	-0.0740***	-0.0065	0.0342
	(-8.40)	(-5.39)	(-0.26)	(1.00)
Current LTV	0.0090**	-0.0088	-0.0940***	-0.1473***
	(2.08)	(-0.87)	(-2.76)	(-2.69)
FICO	0.0115**	-0.0002	0.0054	0.0068
	(2.42)	(-0.04)	(0.85)	(1.04)
Loan Age	0.2745***	0.3493***	0.6077***	0.7282***
	(11.19)	(9.57)	(10.53)	(8.93)
Rate Gap	0.7449***	0.7469***	0.7677***	0.7394***
	(14.00)	(14.08)	(15.85)	(19.52)
FHA	-0.3752***	-0.4767***	-0.6532***	-0.7704***
	(-5.96)	(-8.39)	(-10.52)	(-10.27)
Square Terms of Controls	Yes	Yes	Yes	Yes
Age Group FE	Yes	Yes	Yes	Yes
County × Origin. Year FE	Yes	-	Yes	-
Tract × Origin. Year FE	-	Yes	-	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Loan Officer FE	Yes	Yes	Yes	Yes
Dep. Var. Mean	0.947	0.947	2.069	2.069
R-Squared	0.005	0.005	0.007	0.007
Obs.	5,883,962	5,883,876	5,883,962	5,883,876

Table 6. Effect of Origination Delays on Subsequent Prepayment (NSMO)

This table presents the OLS regressions results examining the effect of borrower-reported delay experiences on future prepayment outcomes. The analysis is based on quarterly panel data from the NSMO, covering loans originated between 2013 to 2021. The dependent variable is a binary indicator for whether the loan prepaid in a given quarter. Delay variable is either an indicator for reported delays in loan processing (*Delay in Loan Processing*) or closing (*Delay in Loan Closing*). In columns (1)–(2), I use the full sample of GSE and FHA loans. In columns (3)–(4), I use the GSE loan subsample. In columns (5)–(6), I use the FHA loan subsample. The t-statistics are reported in parentheses and all standard errors are clustered at the year level. ***, **, and * denote the significance of the parameter estimates at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
			Prepaid			
	GSE + FHA Sample		GSE Sample		FHA Sample	
Processing Delay	-0.3344**		-0.3570**		-0.1684	
	(-2.71)		(-2.66)		(-0.82)	
Closing Delay		-0.4134**		-0.3569*		-0.6554*
		(-2.63)		(-2.25)		(-2.18)
Minority	-0.4868***	-0.4761***	-0.3959**	-0.3851**	-0.7340**	-0.7184**
	(-4.04)	(-4.10)	(-2.68)	(-2.67)	(-2.77)	(-2.78)
Asian	0.3782	0.3940	0.4257	0.4390	-0.0473	-0.0040
	(1.30)	(1.34)	(1.44)	(1.47)	(-0.06)	(-0.00)
Female	-0.4766***	-0.4736***	-0.4782***	-0.4754***	-0.4322	-0.4391
	(-4.59)	(-4.61)	(-4.39)	(-4.42)	(-1.77)	(-1.73)
Coborrower	-0.1228	-0.1244	-0.0604	-0.0625	-0.4662**	-0.4577**
	(-1.54)	(-1.54)	(-0.78)	(-0.81)	(-3.23)	(-3.12)
First-Time Home Buyer	0.9078***	0.9135***	0.8462***	0.8493***	1.1338***	1.1497***
	(7.30)	(7.34)	(7.53)	(7.42)	(3.75)	(3.95)
College Degree	-0.0983	-0.0968	-0.0473	-0.0488	-0.2031	-0.1805
	(-0.85)	(-0.83)	(-0.40)	(-0.41)	(-0.94)	(-0.86)
Non-Native English	-0.2224	-0.2189	-0.2310	-0.2277	-0.0816	-0.0989
	(-1.13)	(-1.12)	(-1.70)	(-1.71)	(-0.13)	(-0.15)
Has Child Under 18	-0.0717	-0.0723	-0.0769	-0.0754	0.0227	0.0361
	(-0.70)	(-0.70)	(-0.50)	(-0.50)	(0.11)	(0.17)
Full-Time Employee	0.4180***	0.4233***	0.2676*	0.2744*	1.3744**	1.3575**
	(4.91)	(4.82)	(2.26)	(2.30)	(2.58)	(2.46)
LTV at Origination	-0.0107	-0.0104	-0.0195	-0.0186	0.0845	0.0722
	(-0.53)	(-0.52)	(-0.83)	(-0.79)	(0.94)	(0.78)
Current LTV	-0.0552**	-0.0545**	-0.0598***	-0.0592***	-0.0512	-0.0548
	(-2.79)	(-2.77)	(-3.71)	(-3.66)	(-0.54)	(-0.58)
Current FICO	0.0181**	0.0183**	0.0067	0.0069	0.0222	0.0227
	(2.52)	(2.52)	(0.51)	(0.53)	(1.32)	(1.33)
Loan Age	0.6436**	0.6417**	0.5968**	0.5938**	0.9440***	0.9472***
	(3.30)	(3.30)	(2.97)	(2.97)	(4.91)	(4.89)
Rate Gap	1.4549***	1.4485***	1.4645***	1.4573***	2.0194***	2.0349***
	(6.09)	(6.08)	(5.37)	(5.39)	(7.99)	(7.68)
FHA	1.1038***	1.1035***				
	(3.73)	(3.67)				
Square Terms of Controls	Yes	Yes	Yes	Yes	Yes	Yes
Age Group FE	Yes	Yes	Yes	Yes	Yes	Yes
Origin. Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Income Group FE	Yes	Yes	Yes	Yes	Yes	Yes
Loan Amount Group FE	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	4.040	4.040	4.058	4.058	3.962	3.962
R-Squared	0.035	0.035	0.037	0.037	0.032	0.032
Obs.	241,048	241,048	195,941	195,941	45,107	45,107

Table 7. Relationship Between Reported Origination Delays and Borrower Satisfaction (NSMO)

This table presents the OLS regression results examining the effect of borrower-reported delay experiences on borrower satisfaction. The analysis is based on loan-level data from the NSMO, covering loans originated between 2013 to 2021. Column (1) uses a binary indicator for perceived fair treatment by the lender. Columns (2)–(6) use indicators for dissatisfaction with specific aspects of the mortgage process: lender interaction, application, documentation, closing, and overall experience. All regressions include controls for borrower demographics, loan characteristics, and fixed effects for origination year, income group, loan amount group, and borrower age group. The t-statistics are reported in parentheses and all standard errors are clustered at the county and year level. ***, **, and * denote the significance of the parameter estimates at the 1%, 5%, and 10% levels, respectively.

(a) Processing Delay

	(1)	(2)	(3)	(4)	(5)	(6)
	Perceived Fair Treatment	*Dissatisfied by*:				
		Lender	*Application*	*Documentation*	*Closing*	*Overall*
Processing Delay	-0.1004***	0.1180***	0.1850***	0.1682***	0.1734***	0.2926***
	(-8.72)	(19.33)	(14.44)	(3.79)	(20.99)	(21.98)
Minority	-0.0518***	0.0012	0.0000	-0.0070	0.0069	0.0011
	(-5.43)	(0.19)	(0.00)	(-0.90)	(0.74)	(0.08)
Asian	-0.0719***	0.0081	-0.0102	-0.0008	0.0225	0.0144
	(-3.85)	(0.54)	(-0.92)	(-0.08)	(1.69)	(0.76)
Female	0.0160**	0.0027	-0.0104**	-0.0077**	-0.0039	-0.0152***
	(2.54)	(0.71)	(-2.33)	(-2.65)	(-1.03)	(-3.38)
Coborrower	0.0048	-0.0011	0.0006	0.0014	-0.0066	-0.0045
	(0.53)	(-0.31)	(0.11)	(0.32)	(-1.39)	(-0.68)
First-Time Home Buyer	0.0237*	0.0008	-0.0010	0.0033	-0.0032	-0.0015
	(2.05)	(0.25)	(-0.28)	(0.86)	(-0.79)	(-0.36)
College Degree	-0.0238**	0.0014	0.0031	0.0057	0.0001	0.0097
	(-2.77)	(0.46)	(0.60)	(1.13)	(0.02)	(1.49)
Non-Native English	-0.0347***	-0.0071	-0.0030	0.0088	-0.0034	0.0021
	(-4.31)	(-1.58)	(-0.46)	(1.22)	(-0.91)	(0.25)
Has Child Under 18	0.0295**	-0.0027	-0.0038	-0.0046	-0.0052	-0.0067
	(2.74)	(-1.53)	(-0.63)	(-0.81)	(-0.93)	(-0.94)
Full-Time Employee	-0.0151*	0.0001	0.0062	-0.0027	-0.0187*	-0.0095
	(-2.03)	(0.02)	(0.60)	(-0.31)	(-1.98)	(-0.96)
LTV at Origination	-0.0020	0.0001	0.0016	0.0011	0.0012	0.0023
	(-1.44)	(0.08)	(1.27)	(0.86)	(1.23)	(1.45)
FICO	0.0002	0.0001	0.0006	-0.0001	0.0004	0.0008
	(0.23)	(0.40)	(1.17)	(-0.08)	(0.81)	(1.38)
FHA	-0.0154	0.0102*	0.0129***	0.0121**	0.0024	0.0192***
	(-1.27)	(2.26)	(4.59)	(2.31)	(0.77)	(3.44)
Square Terms of Controls	Yes	Yes	Yes	Yes	Yes	Yes
Age Group FE	Yes	Yes	Yes	Yes	Yes	Yes
Origin. Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Income Group FE	Yes	Yes	Yes	Yes	Yes	Yes
Loan Amount Group FE	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	0.826	0.040	0.063	0.070	0.069	0.134
R-Squared	0.028	0.054	0.090	0.094	0.071	0.111
Obs.	14,585	14,585	14,585	14,585	14,585	14,585

(b) Closing Delay

	(1)	(2)	(3)	(4)	(5)	(6)
	Perceived Fair Treatment	*Dissatisfied by:*				
		Lender	*Application*	*Documentation*	*Closing*	*Overall*
Closing Delay	-0.0554***	0.0723***	0.1050***	0.1058***	0.1208***	0.1900***
	(-4.66)	(10.78)	(8.26)	(4.28)	(11.02)	(16.84)
Minority	-0.0533***	0.0025	0.0026	-0.0053	0.0078	0.0036
	(-5.90)	(0.39)	(0.40)	(-0.70)	(0.82)	(0.28)
Asian	-0.0716***	0.0075	-0.0108	-0.0018	0.0211	0.0125
	(-4.07)	(0.50)	(-0.93)	(-0.18)	(1.61)	(0.65)
Female	0.0164**	0.0021	-0.0112*	-0.0085**	-0.0050	-0.0168**
	(2.69)	(0.49)	(-2.30)	(-2.86)	(-1.16)	(-3.25)
Coborrower	0.0035	0.0005	0.0031	0.0036	-0.0044	-0.0007
	(0.38)	(0.15)	(0.63)	(0.86)	(-1.04)	(-0.12)
First-Time Home Buyer	0.0246*	-0.0003	-0.0028	0.0016	-0.0050	-0.0044
	(2.17)	(-0.10)	(-0.85)	(0.41)	(-1.19)	(-0.99)
College Degree	-0.0235**	0.0010	0.0025	0.0051	-0.0004	0.0087
	(-2.66)	(0.38)	(0.49)	(1.07)	(-0.08)	(1.32)
Non-Native English	-0.0347***	-0.0072	-0.0030	0.0086	-0.0037	0.0018
	(-4.42)	(-1.37)	(-0.43)	(1.04)	(-1.35)	(0.18)
Has Child Under 18	0.0303**	-0.0035	-0.0051	-0.0057	-0.0063	-0.0086
	(2.81)	(-1.81)	(-0.90)	(-1.05)	(-1.03)	(-1.20)
Full-Time Employee	-0.0150*	-0.0002	0.0058	-0.0031	-0.0192*	-0.0102
	(-2.16)	(-0.03)	(0.55)	(-0.34)	(-1.92)	(-0.94)
LTV at Origination	-0.0018	-0.0002	0.0012	0.0007	0.0008	0.0017
	(-1.27)	(-0.16)	(0.90)	(0.59)	(0.90)	(1.01)
FICO	0.0003	0.0000	0.0004	-0.0002	0.0002	0.0005
	(0.38)	(0.02)	(0.86)	(-0.30)	(0.38)	(0.68)
FHA	-0.0183	0.0133**	0.0180***	0.0165**	0.0066	0.0266***
	(-1.51)	(2.88)	(6.30)	(2.49)	(1.71)	(3.76)
Square Terms of Controls	Yes	Yes	Yes	Yes	Yes	Yes
Age Group FE	Yes	Yes	Yes	Yes	Yes	Yes
Origin. Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Income Group FE	Yes	Yes	Yes	Yes	Yes	Yes
Loan Amount Group FE	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	0.826	0.040	0.063	0.070	0.069	0.134
R-Squared	0.023	0.029	0.044	0.066	0.048	0.066
Obs.	14,585	14,585	14,585	14,585	14,585	14,585

Table 8. Borrower Characteristics and the Likelihood of Initial Loan Delays

This table presents the OLS regression results examining how borrower characteristics—including race, income, and credit scores—are associated with the likelihood of loan closing delays. The analysis uses loan-level observations from the CoreLogic–MBS dataset for loans originated between 2014 and 2021. The dependent variable is I(*Time-To-Close* > 60 Days), an indicator equal to one if *Time-To-Close* exceeds 60 days. The t-statistics are reported in parentheses and all standard errors are clustered at the county and year level. ***, **, and * denote the significance of the parameter estimates at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	$\mathbb{1}$(*Time-To-Close* > 60 Days)					
	GSE + FHA Sample				GSE Sample	FHA Sample
Minority	0.0368***	0.0313***	0.0253***	0.0184***	0.0149***	0.0187***
	(7.49)	(7.30)	(7.21)	(6.66)	(4.89)	(6.27)
Asian	0.0194***	0.0176***	0.0178***	0.0118***	0.0069	0.0211***
	(4.52)	(4.45)	(4.38)	(2.79)	(1.42)	(3.02)
Other Race	0.0081	0.0052	0.0042	-0.0041	0.0272	-0.0237
	(0.61)	(0.39)	(0.32)	(-0.27)	(1.04)	(-1.37)
Female			-0.0011	-0.0016	-0.0028	0.0003
			(-0.93)	(-1.11)	(-1.27)	(0.15)
ln(Income)			-0.0089***	-0.0071**	-0.0095***	-0.0050
			(-3.06)	(-2.39)	(-3.24)	(-1.07)
ln(Loan Amount)			-0.0068	-0.0058	0.0439***	-0.0503***
			(-1.03)	(-0.99)	(7.36)	(-6.22)
Coborrower			0.0054***	0.0059***	0.0030*	0.0081***
			(4.06)	(3.62)	(1.90)	(3.38)
First-Time Home Buyer			-0.0087***	-0.0101***	-0.0082***	-0.0125***
			(-5.93)	(-6.92)	(-3.95)	(-5.88)
FICO			-0.0164***	-0.0157***	-0.0179***	-0.0181***
			(-11.00)	(-10.07)	(-7.58)	(-8.41)
LTV			-0.0006***	-0.0006***	-0.0008***	-0.0017***
			(-5.92)	(-6.31)	(-7.85)	(-6.51)
FHA			0.0224***	0.0238***		
			(6.94)	(7.40)		
Age Group FE	-	-	Yes	Yes	Yes	Yes
County × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Lender FE	-	Yes	Yes	-	-	-
Loan Officer FE	-	-	-	Yes	Yes	Yes
Dep. Var. Mean	0.099	0.099	0.099	0.099	0.080	0.148
R-Squared	0.123	0.141	0.143	0.284	0.321	0.279
Obs.	435,288	435,288	435,288	435,288	159,477	258,583

Table 9. Indirect Test for Lender Discrimination in Initial Loan Delays

This table presents the OLS regression results examining the cross-sectional variations in the effect of borrower minority status on loan closing delays. The analysis uses loan-level observations from the CoreLogic–MBS dataset for loans originated between 2014 and 2021. In columns (1)–(3), the dependent variable is I(*Time-To-Close* > 60 Days), an indicator equal to one if *Time-To-Close* exceeds 60 days. In columns (2) and (5), I interact the *Minority* indicator with a dummy for high race animus areas. In columns (3) and (6), I interact *Minority* with an indicator for low local lending market competition. The t-statistics are reported in parentheses and all standard errors are clustered at the county and year level. ***, **, and * denote the significance of the parameter estimates at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	$\mathbb{1}$(*Time-To-Close* > 60 Days)			$\mathbb{1}$(*90+ Days Delinquent*)		
Minority	0.0184***	0.0095***	0.0109**	0.0085***	0.0090***	0.0140***
	(6.66)	(4.10)	(2.40)	(5.11)	(3.72)	(3.75)
Minority × High Race Animus		0.0173***			-0.0002	
		(4.95)			(-0.07)	
Minority × Low Local Competition			0.0039*			-0.0027
			(1.73)			(-1.42)
Borrower & Loan Controls	Yes	Yes	Yes	Yes	Yes	Yes
County × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Loan Officer FE	Yes	Yes	Yes	Yes	Yes	Yes
Age Group FE	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	0.099	0.099	0.099	0.057	0.058	0.057
R-Squared	0.284	0.284	0.284	0.209	0.210	0.209
Obs.	435,288	405,347	435,288	433,732	403,938	433,732

Table 10. Estimates of Refinancing Responsiveness Parameter ($\hat{\chi}$)

This table reports the maximum likelihood estimates of the refinancing responsiveness parameter ($\hat{\chi}$) from Equation (10), estimated for the full sample and separately by race and income group. Standard errors are computed using nonparametric bootstrapping with 200 replications and are shown in the third column.

Parameters	Estimates	Standard Errors
All Borrowers		
$\hat{\chi}$	0.4054	0.0016
By Race Group		
$\hat{\chi}_{\text{Asian}}$	0.4890	0.0109
$\hat{\chi}_{\text{Black}}$	0.2473	0.0061
$\hat{\chi}_{\text{Hispanic}}$	0.3179	0.0035
$\hat{\chi}_{\text{White}}$	0.4319	0.0019
By Income Group		
$\hat{\chi}_{\text{Low Income}}$	0.2930	0.0024
$\hat{\chi}_{\text{Middle Income}}$	0.4192	0.0030
$\hat{\chi}_{\text{High Income}}$	0.5268	0.0033

Table 11. Calibration Parameters for the CIR Short Rate Process

This table reports the calibration parameters for the CIR short rate process used in the structural model. The long-run mean (μ), mean reversion speed (κ), and volatility parameter (σ) are taken from Berger et al. (2024), who estimate them via maximum likelihood using monthly data on the 3-month U.S. Treasury rate from 1971 to 2021.

Parameters	Value	Description	Sources
κ	0.13	Mean reversion parameter	Berger et al. (2024)
μ	0.035	Long-run short rate mean	Berger et al. (2024)
σ	0.06	Volatility	Berger et al. (2024)

A.1. Selection of 18 States

While CoreLogic deed records provide near-universal coverage across the U.S., the MLS data vary significantly by region, with limited availability in some states (e.g., Alaska and Arkansas).[41] Table A1 summarizes the share of purchase mortgages in the deeds dataset that can be matched to MLS data. To ensure the reliability and representativeness of the analysis, I restrict the sample to 18 U.S. states where MLS matches account for more than 10% of purchase mortgage records. The selected states are Alabama, Arizona, California, Colorado, Delaware, the District of Columbia, Florida, Georgia, Illinois, Maryland, Minnesota, Mississippi, New Jersey, New York, Oregon, Pennsylvania, Rhode Island, and Virginia.

Table A1. State-Level Coverage of CoreLogic Mortgage–MLS Records

This table reports the share of purchase mortgage records in the CoreLogic deed dataset that can be matched to MLS records, by state. The matching is performed using a combination of borrower names, property addresses, and transaction dates.

State	Number of Observations		Ratio (B/A)	State	Number of Observations		Ratio (B/A)
	CoreLogic Mortgage (A)	CoreLogic MLS (B)			CoreLogic Mortgage (A)	CoreLogic MLS (B)	
AL	415,656	86,062	20.70%	MO	683,208	37,977	5.60%
AK	69,836	0	0.00%	MT	126,513	0	0.00%
AZ	1,200,998	410,039	34.10%	NE	226,719	11	0.00%
AR	289,890	2	0.00%	NV	513,799	48,700	9.50%
CA	3,661,569	650,317	17.80%	NH	156,101	0	0.00%
CO	1,038,854	355,293	34.20%	NJ	907,123	264,729	29.20%
CT	313,908	1,392	0.40%	NM	209,840	0	0.00%
DE	123,068	49,839	40.50%	NY	994,164	221,796	22.30%
DC	45,213	19,205	42.50%	NC	1,287,793	69,368	5.40%
FL	2,583,680	810,493	31.40%	ND	89,280	84	0.10%
GA	1,363,933	319,909	23.50%	OH	1,291,163	59,643	4.60%
HI	80,403	3,881	4.80%	OK	437,992	1,650	0.40%
ID	317,398	11	0.00%	OR	624,061	214,456	34.40%
IL	1,248,449	471,124	37.70%	PA	1,182,143	363,785	30.80%
IN	865,381	1,102	0.10%	RI	90,209	14,594	16.20%
IA	373,431	9,051	2.40%	SC	622,825	10,582	1.70%
KS	278,586	25,008	9.00%	SD	3,884	0	0.00%
KY	318,112	27,093	8.50%	TN	881,528	1,907	0.20%
LA	386,188	35,384	9.20%	TX	3,362,279	267,112	7.90%
ME	135,175	0	0.00%	UT	506,567	1	0.00%
MD	766,528	364,832	47.60%	VA	943,232	213,050	22.60%
MA	488,112	14,347	2.90%	WA	1,026,051	62,175	6.10%
MI	1,022,002	5	0.00%	WV	58,762	5,016	8.50%
MN	692,556	249,758	36.10%	WI	605,529	56,076	9.30%
MS	79,158	11,300	14.30%	WY	70,332	0	0.00%

[41] The CoreLogic MLS dataset is sourced from local MLS organizations, and its coverage depends on data-sharing agreements with these entities.

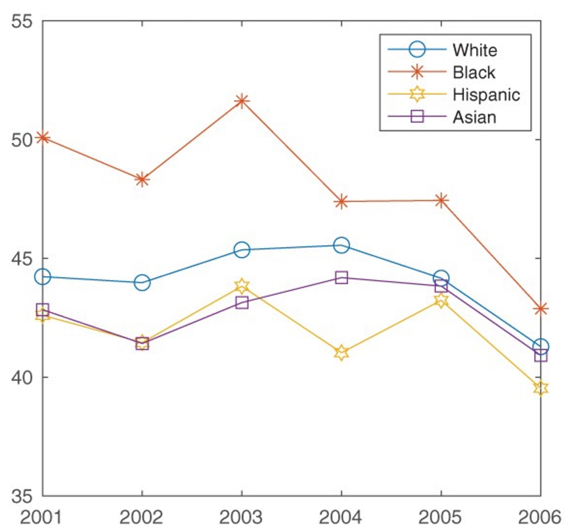
A.2. Validation of *Time-To-Close* Using Wei and Zhao (2022)

Figure A1 compares the average loan processing times by racial group for mortgages originated between 2001 and 2006, as reported in Wei and Zhao (2022), with the corresponding average values of *Time-To-Close* from my dataset. While my primary analysis focuses on the 2014–2021 period, I compute values for 2001–2006 specifically for this comparison.

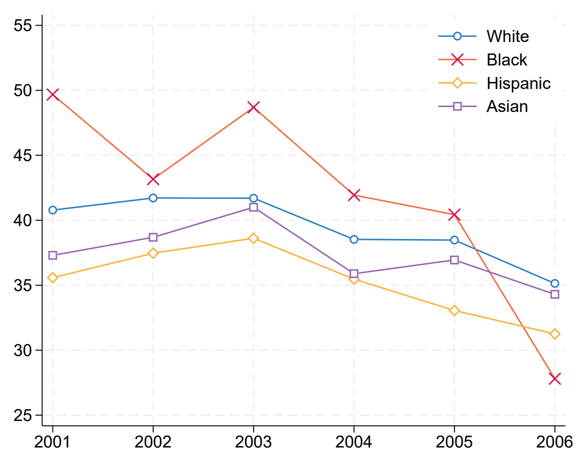
The trends in both panels of Figure A1 exhibit strong consistency. In both datasets, Black borrowers experience the longest average processing times, followed by white, Asian, and Hispanic borrowers. Additionally, the processing time for Black borrowers increases from 2002 to 2003 before declining over the next three years, with similar magnitudes in both datasets. This consistency reinforces the validity of the *Time-To-Close* variable used throughout this study.

Figure A1. Average *Loan Processing Time* and *Time-To-Close* Values by Racial Groups

This figure compares average mortgage processing times by race using two different data sources during the 2001–2006 period. Panel (a) reports average *loan processing times* from Wei and Zhao (2022), based on confidential HMDA. Panel (b) shows average *Time-To-Close* values constructed from the CoreLogic–MBS dataset used in this study.



Loan Processing Time **from Wei and Zhao (2022)**



Time-To-Close

A.3. Identifying Mortgage Outcomes in CoreLogic

CoreLogic does not directly provide loan performance information, but this information can be inferred by connecting mortgage records with subsequent property transactions. Detailed procedures are described as below.

Step 1: Identifying Prepayments For each mortgage record (the "old mortgage"), I identify the next mortgage ("new mortgage") originated against the same property. By analyzing the loan purpose of the new mortgage, I classify the outcome of the old mortgage as follows:

- Cash-out refinance: If the new loan is classified as a cash-out refinance, the old mortgage is marked as prepaid due to cash-out refinance, with the origination date of the new loan recorded as the outcome date.
- Rate-reduction refinance: If the new loan is a rate-reduction refinance, the old mortgage is labeled prepaid due to rate-reduction refinance, again using the new loan's origination date as the outcome date.
- Prepaid due to selling and moving: If the new loan is a purchase mortgage, the old mortgage is categorized as prepaid due to selling and moving, with the outcome date set to the origination date of the new loan.

To ensure accuracy, I verify whether the borrower identities are consistent. That is, for refinanced loans, the borrower names on both the old and new mortgages should match, while for sales, the borrower names should differ.

Step 2: Identifying Defaults If an old mortgage is classified as prepaid due to selling and moving, I further check transaction records for distress indicators. If the property was involved in a short sale, REO (Real Estate Owned), or foreclosure, I reclassify the loan as default since the transaction suggests financial distress.

Step 3: Detecting All-Cash Transactions To account for all-cash sales, I cross-reference mortgage records with property sales data. If the borrower name from the old loan matches the seller name in an all-cash transaction, I adjust the loan's outcome and outcome date accordingly.

Step 4: Verifying Unmatched Loans For loans that do not match with a new mortgage or an all-cash transaction, I determine whether they remain active. This is done by matching each loan with the most recent property record and checking if the borrower name still appears as the current owner.

A.4. Identifying Borrower Race/Ethnicity Using BIFSG

Borrower race and ethnicity are not directly observed in the CoreLogic dataset. Instead, I infer these attributes using borrower first and last names and location information through the Bayesian Improved First Name Surname Geocoding (BIFSG) method (Voicu, 2018). This method is increasingly used in the mortgage studies, such as Ambrose et al. (2021) and Frame et al. (forthcoming). The BIFSG method estimates the probability of an individual belonging to a specific racial/ethnic group (e.g., white, Black, Hispanic, Asian and Pacific Islander, American Indian and Alaskan Native, or Other) based on first names, last names, and ZIP codes of individuals. Specifically:

$$p(r|s,f,z) = \frac{p(r|s) \times p(f|r) \times p(z|r)}{\sum_{r' \in White,Black,Hispanic,Asian,Native,Other} p(r'|s) \times p(f|r') \times p(z|r')}, \tag{A1}$$

where $p(r|s,f,z)$ is the posterior probability of belonging to racial/ethnic group r; $p(r|s)$ is the probability of belonging to group r conditional on surname s; $p(f|r)$ is the probability of having first name f conditional on r; and $p(z|r)$ is the probability of residing in ZIP code z conditional on r. Upon obtaining the probability, I assign each borrower to the racial/ethnic group with the highest probability, following the approach used in Ambrose et al. (2021) and Frame et al. (forthcoming).

To validate the accuracy of BIFSG imputation results, I utilize the matched CoreLogic–HMDA dataset. Since HMDA provides reliable, self-reported borrower race/ethnicity information, this matched dataset allows me to assess the validity of the BIFSG predictions. Specifically, I compute the accuracy rate for each race r, defined as the number of BIFSG predictions for race r that align with HMDA-reported information, divided by the total number of BIFSG predictions for race r. The accuracy rates are notably high: 79.4% for whites, 91.1% for Black and Hispanic borrowers, and 98.1% for Asians.

A.5. Measuring Refinancing Incentives from the Closed-Form Solution of Agarwal et al. (2013)

Agarwal et al. (2013)—hereafter ADL—derive a closed-form solution for the optimal refinancing threshold that accounts for a range of borrower- and contract-level factors, including closing costs, loan size, tax deductibility, interest rate volatility, moving risk, principal amortization, and inflation.[42] Compared to simple rules of thumb—such as treating any positive rate gap as a refinancing opportunity—this framework offers a more theoretically grounded benchmark for evaluating whether a borrower stands to benefit from refinancing.

As a robustness check, I examine whether the 2SLS results in Table 4 are sensitive to an alternative, continuous measure of refinancing incentives: the difference between the observed rate gap and a borrower-specific refinancing threshold implied by the ADL model.

Specifically, I apply the square-root rule approximation proposed by ADL, which yields a closed-form expression for the refinancing threshold x^*:

$$
\begin{aligned}
x^* &= -\sqrt{\frac{\sigma\kappa}{M(1-\tau)}} \cdot \sqrt{2(\rho+\lambda)}, \quad \text{where} \qquad\qquad (\text{A2})\\
\lambda &= \mu + \frac{m_0}{\exp(m_0\Gamma)-1} + \pi,\\
\kappa &= F + fM\left[1 - \frac{\tau}{\theta+\rho+\pi}\left(\frac{1-\exp(-(\theta+\rho+\pi)N)}{N} \cdot \frac{\rho+\pi}{\theta+\rho+\pi} + \theta\right)\right].
\end{aligned}
$$

where σ is the mortgage rate volatility, τ is the marginal tax rate, ρ is the real discount rate, π is the inflation rate, μ is the hazard rate of exogenous mobility (e.g., relocation), m_0 is the current mortgage rate, Γ is the remaining loan term in years, M is the current outstanding loan balance, F is the fixed cost of refinancing, f is the refinancing point cost as a fraction of the loan balance, θ is the expected arrival rate of exogenous moving shocks, and N is the number of years of the new mortgage.

Adopting the parameter values[43] proposed by ADL—also used in subsequent studies (e.g., Agarwal et al., 2016, 2024; Gerardi et al., 2023; Keys et al., 2016), I simplify Equation (A2) to the following expression:

$$
x^* = \sqrt{\frac{0.0109\,(2{,}000+0.007905M)}{0.72M}\sqrt{2\left(0.18+\frac{m_0}{\exp(m_0\Gamma)-1}\right)}}. \qquad (\text{A3})
$$

[42] The solution is derived under several simplifying assumptions, including risk-neutral borrowers and a random walk for the real mortgage rate.

[43] Specifically, $\sigma = 0.0109$, $\tau = 0.28$, $\rho = 0.05$, $\mu = 0.1$, $\pi = 0.03$, $F = 2{,}000$, $f = 0.01$, $\theta = 0.2$, and $N = 30$.

The threshold, in its simplified form, is a function of three key inputs: the current mortgage rate (m_0), the remaining loan balance (M), and the remaining loan term in years (Γ), allowing for quarter-by-quarter calculation of a loan-specific refinancing threshold x^*.

Using the simplified expression in Equation (A3), I compute the refinancing threshold for each loan–quarter observation in the CoreLogic–MBS dataset. Figure A2 shows the distribution of ADL-implied thresholds across the loan panel. In most cases, the threshold falls between 100 and 200 basis points, indicating that, under the ADL framework, a sizable rate reduction is required for refinancing to be considered in-the-money.

To incorporate this alternative measure of refinancing incentives into the regression framework, I re-estimate the 2SLS specification from Table 4, replacing the baseline polynomial rate gap terms (*Rate Gap* and *Rate Gap*^{2}) with polynomial terms based on the difference between the observed rate gap and the ADL-implied threshold: *Rate Gap* − *ADL Threshold* and (*Rate Gap* − *ADL Threshold*)^2.

As shown in Table A2, the coefficient on the linear term remains statistically significant and is larger in magnitude than in the baseline specification, suggesting that the ADL-based measure captures meaningful variation in refinancing behavior that aligns with theoretical predictions.

Importantly, the main coefficient of interest—I(Time-To-Close > 60 Days)—remains virtually unchanged. This confirms that the estimated impact of origination delays is robust to an alternative, theoretically motivated definition of in-the-moneyness.

Figure A2. Distribution of ADL (2013) Refinancing Threshold

This figure shows the distribution of borrower-specific refinancing thresholds implied by the closed-form solution of Agarwal et al. (2013) (ADL). The thresholds are computed using quarterly loan-level data from the CoreLogic–MBS dataset. Following ADL, I apply the square-root rule to approximate the refinancing threshold.

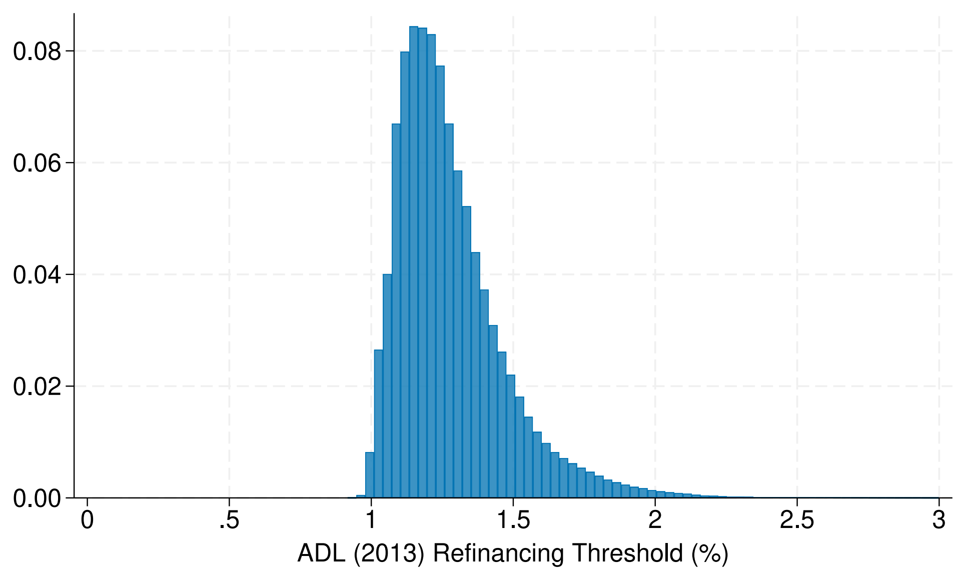


Table A2. 2SLS Regression Results Using ADL-Implied Refinancing Threshold

This table presents the 2SLS regression results examining the effect of initial mortgage delays on refinancing activities, controlling for the refinancing incentive measured as the excess of the observed rate gap over the loan-specific ADL-implied refinancing threshold. I use *Workload* as an instrument for 60+ day loan closing delays. The analysis is based on quarterly loan performance observations from the CoreLogic–MBS dataset, covering loans originated between 2014 and 2021. In columns (1) and (2), I use the full sample of GSE and FHA loans. In columns (3) and (4), I use the GSE loan subsample. In columns (5) and (6), I use the FHA loan subsample. The t-statistics are reported in parentheses and all standard errors are clustered at the county and year level. ***, **, and * denote the significance of the parameter estimates at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Refinance					
	GSE + FHA Sample		GSE Sample		FHA Sample	
I(Time-To-Close > 60 Days)	-0.4704***	-0.7158***	-0.5442*	-0.9534**	-0.3416**	-0.8302***
	(-2.71)	(-3.61)	(-1.81)	(-2.03)	(-2.27)	(-4.10)
Minority	-0.3975***	-0.4899***	-0.3489***	-0.5167***	-0.3995***	-0.3724***
	(-8.89)	(-9.57)	(-4.31)	(-3.46)	(-6.85)	(-6.19)
Asian	0.3007**	0.0386	0.3513***	-0.1193	-0.2793***	-0.3220**
	(2.19)	(0.30)	(2.79)	(-0.39)	(-2.67)	(-2.30)
Female	-0.0158	-0.0073	-0.0075	0.0915	-0.0476**	-0.1008***
	(-0.71)	(-0.26)	(-0.20)	(1.33)	(-2.12)	(-3.11)
Coborrower	0.2308***	0.2205***	0.2580***	0.3448***	0.0692***	0.0034
	(7.95)	(7.34)	(5.07)	(5.60)	(2.65)	(0.11)
First-Time Home Buyer	-0.0028	-0.1011**	0.1995***	0.1438**	-0.3596***	-0.4393***
	(-0.08)	(-2.13)	(4.25)	(2.01)	(-8.61)	(-11.63)
ln(Income)	-4.3788***	-3.9148***	-2.3442	-2.1683	-6.0277***	-6.3722***
	(-5.05)	(-3.32)	(-1.55)	(-1.20)	(-7.13)	(-5.18)
ln(Loan Amount)	-9.5856***	-2.6700	-11.9297***	-9.8565**	-8.7533***	-2.2993
	(-3.75)	(-1.05)	(-3.05)	(-2.54)	(-4.38)	(-0.87)
LTV at Origination	-0.0734***	-0.0210	-0.3407***	-0.3454***	-0.0197	-0.0444
	(-2.88)	(-0.51)	(-7.83)	(-8.50)	(-0.70)	(-0.68)
Current LTV	-0.0919***	-0.1638***	0.3309***	0.3599***	0.1047***	0.1315***
	(-2.76)	(-2.69)	(7.72)	(9.23)	(4.85)	(4.78)
FICO	0.0162*	0.0060	0.0568***	0.0401	0.0407***	0.0507***
	(1.91)	(0.80)	(3.42)	(1.59)	(8.17)	(7.27)
Loan Age	0.8388***	1.0317***	0.8123***	0.9397***	0.4540***	0.5734***
	(11.71)	(9.67)	(9.59)	(11.98)	(7.73)	(6.99)
Rate Gap − ADL Threshold	3.8315***	3.8828***	4.0680***	4.1496***	2.6624***	2.9106***
	(20.67)	(21.99)	(23.47)	(26.28)	(22.17)	(23.06)
$\left(\text{Rate Gap} - \text{ADL Threshold}\right)^2$	0.9364***	0.9538***	0.8742***	0.8775***	0.3464***	0.3351***
	(19.85)	(17.38)	(29.75)	(15.38)	(22.03)	(18.57)
FHA	-1.0139***	-1.2192***				
	(-10.19)	(-15.51)				
Square Terms of Controls	Yes	Yes	Yes	Yes	Yes	Yes
Age Group FE	Yes	Yes	Yes	Yes	Yes	Yes
County × Origin. Year FE	Yes	-	Yes	-	Yes	-
Tract × Origin. Year FE	-	Yes	-	Yes	-	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Loan Officer FE	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	3.016	3.016	3.413	3.412	2.030	2.030
R-Squared	0.014	0.014	0.015	0.016	0.008	0.008
Obs.	5,883,962	5,883,876	2,230,114	2,230,044	3,653,833	3,653,804

A.6. Joint Estimation of Refinancing Responsiveness and Delay Effect Parameters

In the baseline structural model presented in Section 5.2, I estimate a single refinancing responsiveness parameter, χ, using the sample of non-delayed borrowers. I then impose a 24% reduction in χ for delayed borrowers in the simulation of coupon rate trajectories.

In this section, I extend the model by modifying the log-likelihood function to jointly estimate an additional delay-induced discouragement parameter, γ. To incorporate γ, I allow refinancing responsiveness to vary by delay status. Specifically, I model the Poisson arrival rate of refinancing decisions for borrower i at time t as:

$$\chi_i = \begin{cases} \chi & \text{if} \quad \text{Delay}_i = 0, \\ \chi + \gamma & \text{if} \quad \text{Delay}_i = 1. \end{cases} \tag{A4}$$

Accordingly, the log-likelihood function in Equation (10) becomes:

$$\begin{aligned} \mathscr{L}(\chi,\gamma) = &\sum_{i:\,\text{Delay}_i=0} \sum_t \left[y_{i,t} \log\left(1-e^{-\chi\, dt}\right) - (1-y_{i,t})\chi\, dt\right] \mathbb{1}(c_{i,t-1} - m_t > 0) \quad + \\ &\sum_{i:\,\text{Delay}_i=1} \sum_t \left[y_{i,t} \log\left(1-e^{-(\chi+\gamma)\, dt}\right) - (1-y_{i,t})(\chi+\gamma)\, dt\right] \mathbb{1}(c_{i,t-1} - m_t > 0). \end{aligned} \tag{A5}$$

Relative to the baseline estimation in Section 5.2.3, the main change is that the estimation sample now includes both delayed and non-delayed borrowers. The same CoreLogic–MBS loan panel is used, with refinancing behavior observed quarterly. Delay status is defined as in Section 3.1, based on whether *Time-To-Close* exceeds 60 days. All other covariates and simulation procedures remain unchanged.

Table A3 reports the joint maximum likelihood estimates of $\hat{\chi}$ and $\hat{\gamma}$, both overall and by borrower subgroup. Across all specifications, the estimate of $\hat{\gamma}$ is negative and statistically significant, confirming that borrowers who experienced origination delays are less responsive to refinancing opportunities.

For the full sample, the baseline refinancing intensity is estimated at $\hat{\chi} = 0.4054$, which is nearly identical to the estimate obtained in the main specification using only non-delayed borrowers. The delay-induced discouragement effect is estimated at $\hat{\gamma} = -0.1566$, implying a 38% reduction in refinancing hazard $\left(\frac{0.1566}{0.4054}\right)$—a stronger effect than the 24% decline assumed in the baseline calibration.

At the subgroup level, the estimated $\hat{\chi}$ values, again, closely match those reported in Table 10, suggesting that the inclusion of delayed borrowers and estimation of $\hat{\gamma}$ does not affect baseline responsiveness. However, the estimated discouragement effects are particularly large for underserved groups.

For example, the proportional reduction in refinancing responsiveness—as measured by $\hat{\gamma}/\hat{\chi}$—is 52% for Hispanic borrowers (0.1659/0.3180), 51% for Black borrowers (0.1266/0.2473), and 44% for Asian borrowers (0.2164/0.4890), compared to 34% for White borrowers (0.1447/0.4319). A similar pattern emerges across income groups: the refinancing hazard is reduced by 48% for low-income borrowers (0.1409/0.2930), 41% for middle-income borrowers, and 31% for high-income borrowers.

Panel (a) of Figure A4 shows the distribution of the simulated PV of overpayment under the full model. The average PV rises to \$12,226—nearly double the baseline estimate of \$6,641 in Section 5.2.4, which assumed a fixed 24% reduction in refinancing probability. The increase reflects the larger estimated delay effect ($\hat{\gamma} = -0.1566$), which contributes to more persistent refinancing frictions and higher cumulative overpayment among delayed borrowers.

Panel (b) of Figure A4 illustrates the distributional consequences by race and income group. Compared to Figure 11, the estimated overpayment amounts are considerably higher across all groups, reflecting the stronger discouragement effect incorporated in this extended model. While overpayment is substantial across all subgroups, these financial burdens are disproportionately concentrated among underserved populations. For example, the average PV of overpayment is \$20,103 for Black borrowers and \$20,067 for Hispanic borrowers, compared to just \$9,910 for White borrowers. Notably, Asian borrowers—who typically exhibit high baseline responsiveness—now experience higher losses (\$12,648) than White borrowers, due to their larger estimated discouragement effect ($\hat{\gamma}/\hat{\chi}$). A similar pattern holds across income groups, with low-income borrowers facing considerably greater overpayment than their higher-income counterparts.

These results reinforce the baseline findings in Section 5.2 by showing that when the discouragement effect is jointly estimated and incorporated into the model, both the average PV of overpayment and its distributional consequences become more severe. In particular, the financial burden of refinancing delays falls disproportionately on historically underserved racial and income groups.

Table A3. Estimates of Refinancing Responsiveness ($\hat{\chi}$) and Delay Discouragement Effect ($\hat{\gamma}$)

This table reports the joint maximum likelihood estimates of the refinancing responsiveness parameter ($\hat{\chi}$) and the delay discouragement effect parameter ($\hat{\gamma}$) for the full sample and by race and income groups. The parameter $\hat{\chi}$ captures the baseline refinancing responsiveness, while $\hat{\gamma}$ measures the reduction in responsiveness associated with experiencing a delayed mortgage origination (i.e., *Time-To-Close* $>$ 60 days). Standard errors are computed using nonparametric bootstrapping with 200 replications and are shown in the third column.

Parameters	Estimates	Standard Errors
All Borrowers		
$\hat{\chi}$	0.4054	0.0015
$\hat{\gamma}$	−0.1566	0.0031
By Race Group		
$\hat{\chi}_{\text{Asian}}$	0.4890	0.0097
$\hat{\gamma}_{\text{Asian}}$	−0.2164	0.0182
$\hat{\chi}_{\text{Black}}$	0.2473	0.0056
$\hat{\gamma}_{\text{Black}}$	−0.1266	0.0093
$\hat{\chi}_{\text{Hispanic}}$	0.3180	0.0034
$\hat{\gamma}_{\text{Hispanic}}$	−0.1659	0.0062
$\hat{\chi}_{\text{White}}$	0.4319	0.0019
$\hat{\gamma}_{\text{White}}$	−0.1447	0.0040
By Income Group		
$\hat{\chi}_{\text{Low Income}}$	0.2930	0.0023
$\hat{\gamma}_{\text{Low Income}}$	−0.1409	0.0044
$\hat{\chi}_{\text{Middle Income}}$	0.4191	0.0027
$\hat{\gamma}_{\text{Middle Income}}$	−0.1723	0.0058
$\hat{\chi}_{\text{High Income}}$	0.5270	0.0033
$\hat{\gamma}_{\text{High Income}}$	−0.1636	0.0066

Figure A3. Distribution of PV of Overpayment: Joint Estimation of χ and γ

This figure shows the distribution of PV of overpayment across the loan panel under the extended model that jointly estimates both refinancing responsiveness ($\hat{\chi}$) and delay discouragement effect ($\hat{\gamma}$).

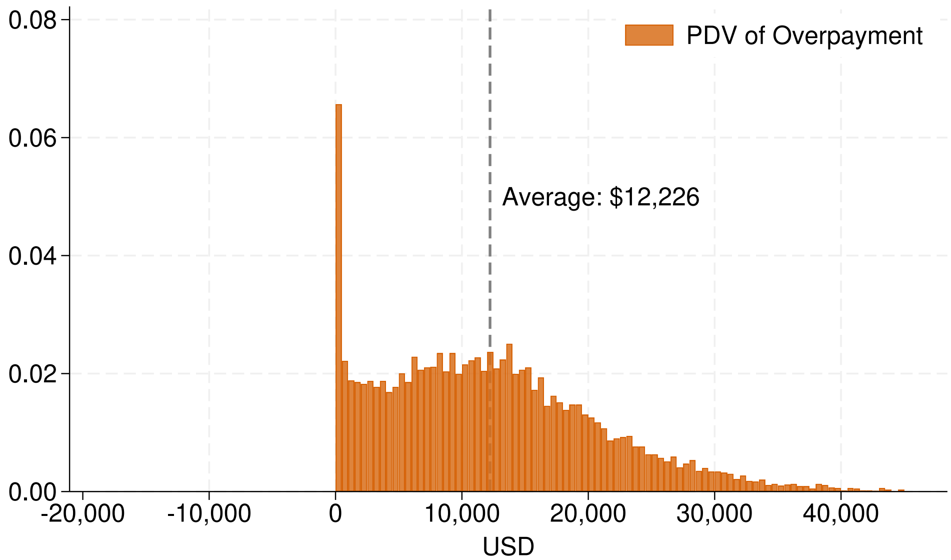
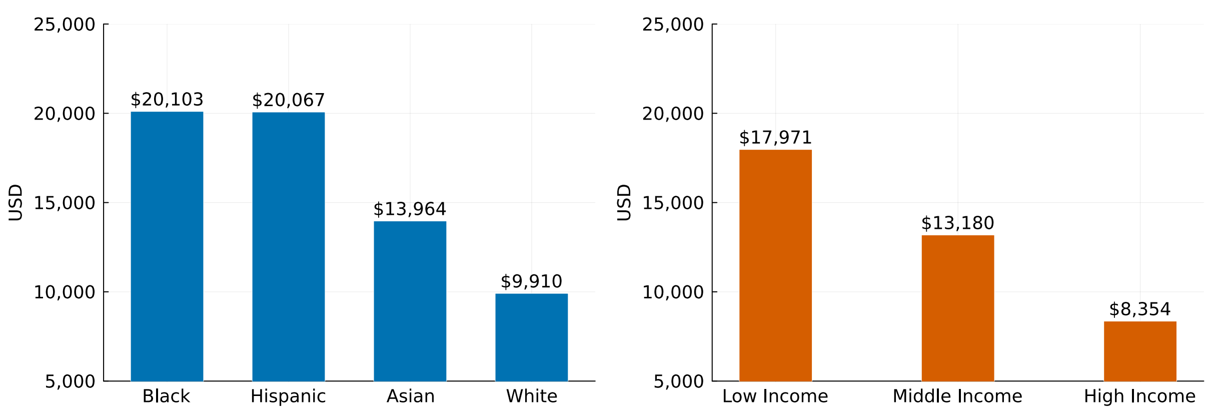


Figure A4. Average PV of Overpayment by Race and Income Groups: Joint Estimation of χ and γ

This figure plots the average PV of overpayment under the joint estimation model, separately by race (panel (a)) and income group (panel (b)).



(a) By Race **(b) By Income Group**

A.7. Extending the Model to Endogenous Mortgage Spreads

In Section 5.2, I treat the mortgage spread as exogenous—fixed at 168 basis points—and map it to simulated short rate process. This simplification facilitates a transparent evaluation of borrower-side frictions without requiring a full model of investor-side pricing behavior. Moreover, empirically, the mortgage spread had been stable over 2013–2019 period as shown in Figure 7.

Nonetheless, it is possible to endogenize mortgage spreads within the structural model by incorporating investor-side pricing behavior. Following the framework of Berger et al. (2024), I assume that investors value mortgage-backed securities (MBS) based on the expected PV of future cash flows, taking borrowers' refinancing behavior as given as in Section 5.2.2. Specifically, when a unit amount of mortgage with coupon c is originated at time 0 and the short rate is r, its value to secondary market investors is given by:

$$P(c,r \mid \chi) := \mathbb{E}\left[\int_0^\tau e^{-\int_0^t r_s ds}(c-f)dt + e^{-\int_0^\tau r_s ds} \mid \chi\right], \tag{A6}$$

where τ is the stochastic prepayment time, f is the servicing fee retained by the lender, and χ denotes the borrower's responsiveness to refinancing opportunities. The expectation is taken over the joint distribution of short rate paths and prepayment timing.

In equilibrium, mortgage lenders must break even by setting the coupon $m(r)$ such that the MBS price equals the loan principal plus a fixed gain-on-sale π:

$$P(m(r), r \mid \chi) = 1 + \pi. \tag{A7}$$

This condition implicitly defines the equilibrium mortgage rate as a function of the short rate r and borrower responsiveness χ. Given $\hat{\chi} = 0.4054$, the spread $m(r) - r$ is endogenously determined by the short rate r. Specifically, the mortgage spread is a decreasing function of r: a lower short rate increases borrowers' incentive to refinance, raising expected prepayment risk. This shortens the expected duration of the mortgage and reduces its value $P(c, r \mid \chi)$ to investors. To satisfy the break-even condition, lenders must raise the coupon $m(r)$, resulting in a wider mortgage spread. This mechanism generates a downward-sloping relationship between spreads and short rates.

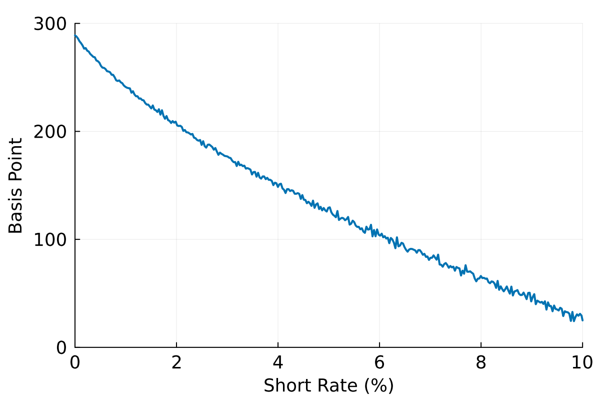
Panel (a) of Figure A5 plots the model-implied mortgage spread as a function of the short rate, based on the equilibrium condition in Equation (A7) and assuming a borrower responsiveness of $\hat{\chi} = 0.4054$.

The figure confirms the model's prediction: the spread declines with the short rate, reflecting lower prepayment risk in high-rate environments. In panel (b), I map the historical short rate series to model-implied spreads using the spread function from panel (a), and compare it to observed mortgage spreads over the same period. The two series exhibit similar dynamics, with a correlation of 0.59, indicating that the structural model, when combined with rational investor pricing, captures meaningful variation in spreads and supports the external validity of the pricing framework.

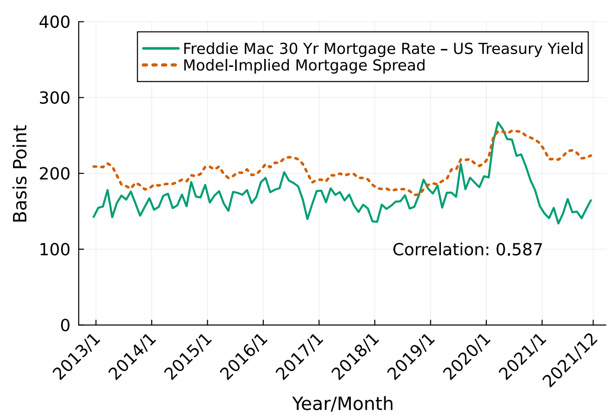
In the final step, I simulate the short rate process using the CIR model and apply the model-implied spread function (based on $\hat{\chi} = 0.4054$) to generate corresponding mortgage rates. I then reassess the PV of overpayment resulting from origination delays. Figure A6 shows the distribution of PV across the loan panel. The average overpayment rises to \$11,259—substantially higher than the baseline estimate using a flat spread. This increase reflects the fact that under the same interest rate paths, the mortgage rate is more frequently "in the money" when the spread is endogenously determined. The endogenous pricing mechanism amplifies refinancing incentives in low-rate environments, resulting in more missed opportunities for delayed borrowers and greater cumulative financial losses.

Figure A5. Model-Implied Mortgage Spread

Panel (a) plots the model-implied mortgage spread as a function of the short rate, holding the borrower refinancing parameter fixed at $\hat{\chi} = 0.4054$. Panel (b) maps the historical short rate between 2013 to 2021 to the model-implied spread using the function in panel (a), and compares it with actual market mortgage spreads over time.



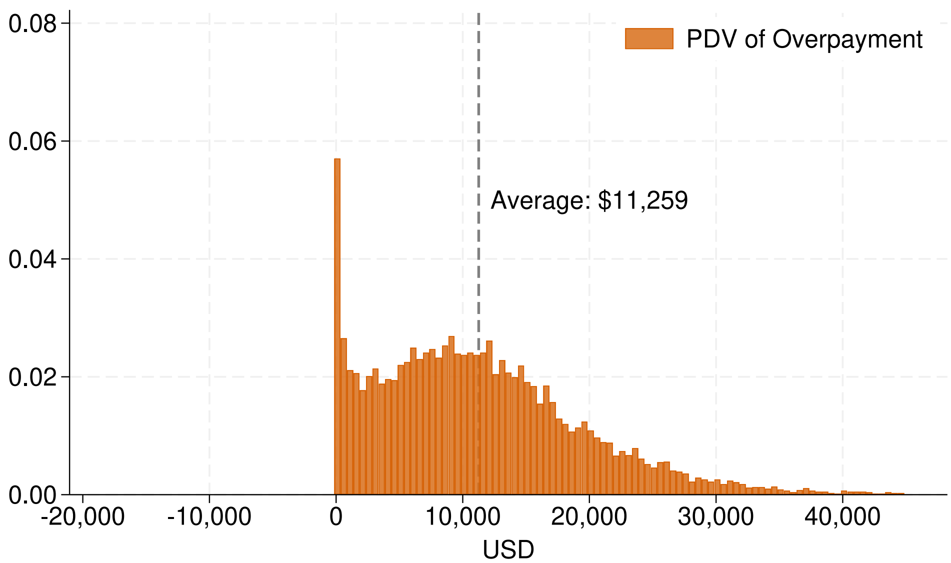
(a) Model-Implied Mortgage Spread by Short Rates



(b) Actual vs. Model-Implied Mortgage Spread

Figure A6. Distribution of PV of Overpayment: Endogenous Mortgage Spread

This figure shows the distribution of PV of overpayment by origination delays across the loan panel, under endogenous mortgage rate pricing. Mortgage rates are constructed by adding the model-implied spread to simulated CIR short rate paths.



B. Additional Figures and Tables

B.1. Figures

Figure B1. Quarterly Average Refinancing Rates By Rate Gaps

This figure shows the average quarterly refinancing rates categorized by ranges of rate gaps.

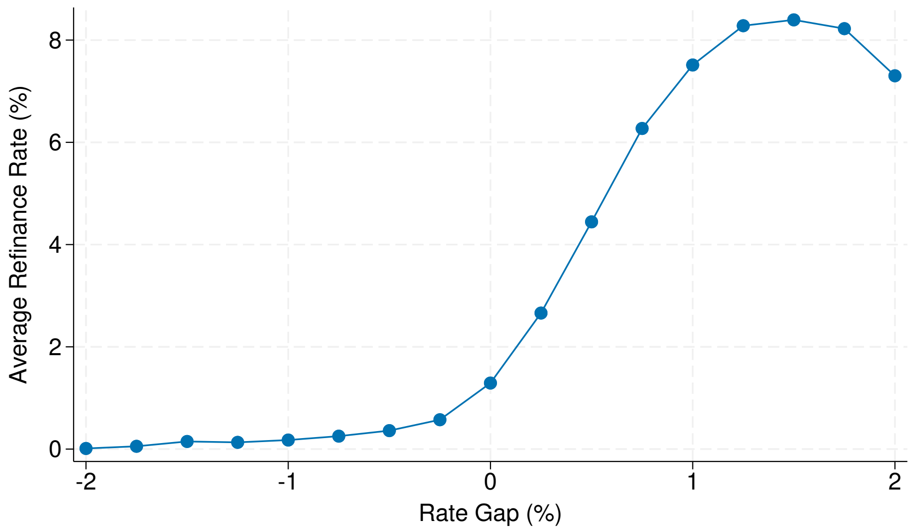


Table B1. Impact of Initial Mortgage Delays on Cash-Out Refinance and Prepayment Due to Moving and Selling

This table presents the 2SLS regression results examining the effect of delays for initial mortgages on quarterly cash-out refinancing and prepayment due to moving and selling, using *Workload* as an instrument. In columns (1) and (2), the dependent variable is *Cash-Out Refinance*, indicating loans cash-out refinanced during the quarter. In columns (7) and (8), the dependent variable is *Prepaid Due to Moving and selling*, indicating loans prepaid due to moving and selling during the quarter. The t-statistics are reported in parentheses and all standard errors are clustered at the county and year level. ***, **, and * denote the significance of the parameter estimates at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
	Cash-Out Refinance		*Prepaid Due to Selling and Moving*	
$\mathbb{1}$(Time-To-Close > 60 Days)	-0.3828**	-0.2196	-0.2373	-0.1698
	(-2.40)	(-1.17)	(-1.50)	(-0.98)
Minority	-0.2159***	-0.1588***	-0.4363***	-0.4004***
	(-5.39)	(-5.19)	(-18.79)	(-10.96)
Asian	-0.4756***	-0.5282***	-0.2482***	-0.2206**
	(-5.57)	(-7.03)	(-5.15)	(-2.17)
Female	-0.0378	-0.0489*	0.0394*	0.0508**
	(-1.63)	(-1.75)	(1.88)	(2.11)
Coborrower	0.0110	-0.0080	-0.0657***	-0.0576***
	(0.50)	(-0.33)	(-3.14)	(-2.59)
First-Time Home Buyer	-0.3638***	-0.4260***	-0.4605***	-0.5453***
	(-10.56)	(-11.30)	(-13.79)	(-19.84)
ln(Income)	0.1306	-0.7997	1.2186	1.1114
	(0.16)	(-0.84)	(1.48)	(1.45)
ln(Loan Amount)	8.0003***	14.4576***	3.9205	5.7470***
	(6.89)	(7.05)	(1.45)	(3.13)
LTV at Origination	0.0946***	0.1597***	0.1226***	0.1979***
	(4.83)	(5.80)	(3.79)	(9.11)
Current LTV	-0.1827***	-0.2860***	-0.0883**	-0.1682***
	(-8.13)	(-9.41)	(-2.49)	(-8.20)
FICO	0.0706***	0.0780***	0.0253***	0.0240***
	(12.72)	(10.26)	(5.52)	(4.51)
Loan Age	0.2115***	0.2879***	0.2582***	0.3195***
	(14.80)	(11.36)	(9.57)	(17.34)
Rate Gap	0.8195***	0.8140***	0.0593	0.0503
	(13.53)	(13.25)	(1.36)	(1.16)
FHA	-1.3051***	-1.3966***	-1.1752***	-1.3703***
	(-10.89)	(-13.76)	(-8.35)	(-19.55)
Square Terms of Controls	Yes	Yes	Yes	Yes
Age Group FE	Yes	Yes	Yes	Yes
County × Origin. Year FE	Yes	-	Yes	-
Tract × Origin. Year FE	-	Yes	-	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Loan Officer FE	Yes	Yes	Yes	Yes
Dep. Var. Mean	1.195	1.195	1.469	1.468
R-Squared	0.006	0.006	0.004	0.004
Obs.	5,884,007	5,883,910	5,884,007	5,883,910

Table B2. Heterogeneous Effects of Initial Mortgage Delays on Cash-Out Refinancing Outcomes: Recapture vs. Switching

This table presents the 2SLS regression results examining the effect of initial mortgage delays on recapture and switching cash-out refinancing activities. I use *Workload* as an instrument for loan closing delays exceeding 60 days. The analysis is based on quarterly loan performance observations from the CoreLogic–MBS dataset, covering loans originated between 2014 and 2021. In columns (1) and (2), the dependent variable is *Recapture Cash-Out Refinance*, which indicates cash-out refinancing by the original lender. In columns (3) and (4), the dependent variable is *Switching Cash-Out Refinance*, representing cash-out refinancing through a different lender. t-statistics are reported in parentheses, with standard errors clustered at the county and year level. ***, **, and * denote the significance of the parameter estimates at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
	Recapture Cash-Out Refinance		*Switching Cash-Out Refinance*	
$\mathbb{1}$(Time-To-Close > 60 Days)	-0.2241***	-0.2472*	-0.1588	0.0276
	(-3.01)	(-1.77)	(-1.16)	(0.20)
Minority	-0.0716***	-0.0564***	-0.1444***	-0.1024***
	(-4.26)	(-2.68)	(-4.59)	(-5.05)
Asian	-0.1880***	-0.2286***	-0.2876***	-0.2996***
	(-5.40)	(-5.38)	(-5.06)	(-3.95)
Female	-0.0017	-0.0119	-0.0360*	-0.0371
	(-0.14)	(-0.90)	(-1.94)	(-1.63)
Coborrower	0.0252**	0.0145	-0.0142	-0.0225
	(2.04)	(0.92)	(-0.83)	(-1.30)
First-Time Home Buyer	-0.1159***	-0.1343***	-0.2479***	-0.2917***
	(-7.50)	(-8.64)	(-9.18)	(-9.79)
ln(Income)	-0.2514	-0.8846	0.3820	0.0850
	(-0.62)	(-1.51)	(0.67)	(0.14)
ln(Loan Amount)	1.5929***	3.6941***	6.4074***	10.7635***
	(3.93)	(4.85)	(6.19)	(6.69)
LTV at Origination	0.0009	0.0129*	0.0937***	0.1468***
	(0.14)	(1.72)	(6.31)	(6.19)
Current LTV	-0.0275***	-0.0528***	-0.1552***	-0.2332***
	(-4.17)	(-5.97)	(-9.12)	(-10.25)
FICO	0.0187***	0.0242***	0.0519***	0.0538***
	(9.96)	(6.33)	(9.85)	(8.50)
Loan Age	0.0561***	0.0780***	0.1554***	0.2098***
	(7.55)	(6.89)	(16.18)	(12.83)
Rate Gap	0.2758***	0.2982***	0.3475***	0.3715***
	(9.35)	(8.83)	(7.71)	(7.22)
FHA	-0.3710***	-0.4215***	-0.9341***	-0.9750***
	(-8.10)	(-11.00)	(-11.56)	(-13.15)
Square Terms of Controls	Yes	Yes	Yes	Yes
Age Group FE	Yes	Yes	Yes	Yes
County × Origin. Year FE	Yes	-	Yes	-
Tract × Origin. Year FE	-	Yes	-	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Loan Officer FE	Yes	Yes	Yes	Yes
Dep. Var. Mean	0.346	0.346	0.849	0.849
R-Squared	0.002	0.002	0.004	0.004
Obs.	5,884,007	5,883,910	5,884,007	5,883,910

Table B3. Robustness Test: Restricting to Less Credit-Constrained Borrowers

This table presents 2SLS regression results examining the effect of initial mortgage delays on refinancing activity across progressively stricter borrower subsamples, following the sample restriction strategy of Keys et al. (2016). Column (1) replicates the baseline result using the full GSE sample (identical to column (3) of Table 4). Column (2) restricts the sample to borrowers with *FICO* above 680 and *LTV at Origination* below 90%. Column (3) adds an additional filter, excluding borrowers with any missed payment history. Column (4) further excludes loans with (quarterly updated) *Current LTV* above 90%. t-statistics are reported in parentheses and all standard errors are clustered at the county and year level. ***, **, and * denote the significance of the parameter estimates at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
	Refinance			
	GSE Sample	GSE Sample w/ *FICO* > 680, *LTV at Origination* < 90%	GSE Sample w/ *FICO* > 680, *LTV at Origination* < 90%, no missed payment	GSE Sample w/ *FICO* > 680, *Current LTV* < 90%, no missed payment
$\mathbb{1}$(Time-To-Close > 60 Days)	-0.5693*	-1.1126**	-1.3499**	-1.3268**
	(-1.87)	(-2.14)	(-2.04)	(-1.99)
Minority	-0.3492***	-0.5005***	-0.4283***	-0.4397***
	(-4.21)	(-3.61)	(-2.84)	(-2.94)
Asian	0.3642***	0.6520***	0.8271***	0.8294***
	(2.89)	(4.37)	(4.37)	(4.26)
Female	-0.0086	0.0071	-0.0487	-0.0595
	(-0.22)	(0.14)	(-0.67)	(-0.77)
Coborrower	0.2540***	0.3444***	0.2031**	0.2021**
	(4.97)	(4.69)	(2.50)	(2.37)
First-Time Home Buyer	0.2039***	0.4648***	0.4326***	0.4402***
	(4.34)	(6.79)	(4.91)	(5.05)
ln(Income)	-2.8463*	-4.1046*	-2.2224	-2.4853
	(-1.92)	(-1.86)	(-0.87)	(-0.96)
ln(Loan Amount)	-5.6503	-4.5402	-8.4799*	-8.9590*
	(-1.44)	(-0.86)	(-1.66)	(-1.69)
LTV at Origination	-0.3573***	-0.2757***	-0.3772***	-0.4847***
	(-8.08)	(-4.02)	(-4.24)	(-5.04)
Current LTV	0.3419***	0.2274***	0.2590***	0.3626***
	(7.77)	(7.35)	(6.73)	(8.38)
FICO	0.0582***	0.1301***	0.1005*	0.1009*
	(3.53)	(2.78)	(1.66)	(1.69)
Loan Age	0.8633***	0.9260***	1.0485***	1.0156***
	(9.60)	(10.58)	(10.01)	(9.70)
Rate Gap	2.1255***	2.1232***	2.7575***	2.8772***
	(12.46)	(10.63)	(11.25)	(10.93)
Square Terms of Controls	Yes	Yes	Yes	Yes
Age Group FE	Yes	Yes	Yes	Yes
County × Origin. Year FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Loan Officer FE	Yes	Yes	Yes	Yes
Dep. Var. Mean	3.413	3.130	3.390	3.414
R-Squared	0.013	0.012	0.014	0.014
Obs.	2,230,114	1,060,508	926,903	913,991

Table B4. Summary Statistics: NSMO Dataset

This table reports summary statistics from the NSMO dataset. Panel (a) summarizes the loan-quarter panel, where each loan contributes multiple observations over time. Panel (b) presents loan-level statistics, restricting to a single observation per loan at origination.

(a) Quarterly Loan Panel

	Obs.	Mean	S.D.	P25	P50	P75
Prepaid	241,048	4.04	19.69	0.00	0.00	0.00
Processing Delay	241,048	0.18	0.39	0.00	0.00	0.00
Closing Delay	241,048	0.26	0.44	0.00	0.00	1.00
White	241,048	0.79	0.41	1.00	1.00	1.00
Minority	241,048	0.14	0.35	0.00	0.00	0.00
Black	241,048	0.06	0.24	0.00	0.00	0.00
Hispanic	241,048	0.08	0.28	0.00	0.00	0.00
Asian	241,048	0.04	0.19	0.00	0.00	0.00
Other Race	241,048	0.03	0.16	0.00	0.00	0.00
Female	241,048	0.47	0.50	0.00	0.00	1.00
Coborrower	241,048	0.51	0.50	0.00	1.00	1.00
First-Time Home Buyer	241,048	0.63	0.48	0.00	1.00	1.00
College Degree	241,048	0.65	0.48	0.00	1.00	1.00
Non-Native English	241,048	0.07	0.26	0.00	0.00	0.00
Has Child Under 18	241,048	0.30	0.46	0.00	0.00	1.00
Full-Time Employee	241,048	0.04	0.20	0.00	0.00	0.00
FHA	241,048	0.19	0.39	0.00	0.00	0.00
LTV at Origination (%)	241,048	85.47	14.94	79.00	90.00	96.00
Current LTV (%)	241,048	69.86	18.74	58.00	72.00	84.00
FICO	241,048	742.45	62.29	699.00	753.00	795.00
Loan Age	241,048	11.49	8.49	5.00	9.00	17.00
Rate Gap (%)	241,048	-0.24	1.07	-0.74	-0.15	0.42

(b) Loan-Level Dataset

	Obs.	Mean	S.D.	P25	P50	P75
Perceived Fair Treatment	14,585	0.83	0.38	1.00	1.00	1.00
Dissatisfied by: Lender	14,585	0.04	0.20	0.00	0.00	0.00
Dissatisfied by: Application	14,585	0.06	0.24	0.00	0.00	0.00
Dissatisfied by: Documentation	14,585	0.07	0.26	0.00	0.00	0.00
Dissatisfied by: Closing	14,585	0.07	0.25	0.00	0.00	0.00
Dissatisfied by: Overall	14,585	0.13	0.34	0.00	0.00	0.00
Processing Delay	14,585	0.17	0.37	0.00	0.00	0.00
Closing Delay	14,585	0.25	0.43	0.00	0.00	0.00
White	14,585	0.79	0.41	1.00	1.00	1.00
Minority	14,585	0.14	0.35	0.00	0.00	0.00
Black	14,585	0.06	0.23	0.00	0.00	0.00
Hispanic	14,585	0.08	0.28	0.00	0.00	0.00
Asian	14,585	0.04	0.20	0.00	0.00	0.00
Other Race	14,585	0.03	0.16	0.00	0.00	0.00
Female	14,585	0.46	0.50	0.00	0.00	1.00
Coborrower	14,585	0.52	0.50	0.00	1.00	1.00
First-Time Home Buyer	14,585	0.65	0.48	0.00	1.00	1.00
College Degree	14,585	0.66	0.47	0.00	1.00	1.00
Non-Native English	14,585	0.09	0.29	0.00	0.00	0.00
Has Child Under 18	14,585	0.32	0.47	0.00	0.00	1.00
Full-Time Employee	14,585	0.04	0.20	0.00	0.00	0.00
LTV (%)	14,585	85.40	14.89	79.00	90.00	96.00
FICO	14,585	743.96	61.98	702.00	755.00	796.00
FHA	14,585	0.18	0.38	0.00	0.00	0.00