Understanding Racial Disparities in Mortgage Refinancing

Heejin Yoon*

January 14, 2025

Abstract

This paper investigates whether discriminatory practices during the initial mortgage origination process contribute to racial disparities in refinancing. Using a matched dataset from CoreLogic and the GSE Single-Family Loan-Level Dataset, we find that minority borrowers face disproportionately longer loan processing times, which significantly discourage future refinancing by undermining trust in lenders. Leveraging an instrumental variable (IV) approach that uses loan officer workload as an instrument for processing delays, we estimate that these delay experiences reduce quarterly refinancing rates by 1.18–1.28 percentage points. Missed refinancing opportunities during the 2009–2021 period led to an aggregate overpayment of $594 million by minorities over the remaining life of their loans. Our findings highlight how subtle forms of discriminatory practices in mortgage lending can impose substantial financial burdens on minority borrowers.

JEL codes: G21, G23, R30, R51

*University of Wisconsin–Madison (e-mail: heejin.yoon@wisc.edu)

1. Introduction

Refinancing behavior among US borrowers varies widely. Some borrowers act quickly to take advantage of lower interest rates, while others delay or never refinance, missing out on significant savings.[1] This heterogeneity creates a significant cross-subsidy from slow to fast borrowers (Berger et al., 2024; Fisher et al., 2024; Zhang, 2024). In addition, the variation in refinancing behavior is not entirely random; instead, it is strongly correlated with race. Black and Hispanic borrowers are less likely than white borrowers to refinance their mortgages, even when other factors that influence refinancing decisions are accounted.[2] The racial gap in refinancing propensity exacerbates mortgage rate inequality, disproportionately channeling the benefits of expansionary monetary policy toward white borrowers and deepening existing wealth gaps (Gerardi et al., 2023).

Why do there exist racial disparities in mortgage refinancing? Despite its significant implications for household wealth, the underlying causes of racial disparities in refinancing behavior remain largely unexplored in the literature. This lack of investigation may stem from the difficulty in pinpointing the sources, as these disparities persist even after controlling for observables (e.g., credit scores), which rules out differences in such variables as explanations. Another—perhaps more critical—reason for this silence may be the risk of inadvertently reinforcing harmful stereotypes. Refinancing decisions are complex, requiring borrowers to consider various factors, such as potential savings from lower mortgage rates, refinancing costs, loan size, marginal tax rates, the likelihood of moving, and even personal discount rates (Agarwal et al., 2013). Consequently, acting quickly to refinance opportunities is often seen as a sign of being financially savvy, while missing the opportunity as irrational, often described in terms like *failure* (Keys et al., 2016), *mistake* (Agarwal et al., 2016), and *woodhead behavior* (Deng and Quigley, 2012). Due to this way of framing, exploring the causes of racial disparities risks suggesting that minorities inherently are less financially sophisticated, less informed, or more prone to behavioral biases.

This paper addresses this research gap by linking racial disparities in refinancing behaviors to discriminatory practices in the lending market. Borrowers' refinancing decisions may be influenced by their experiences during the initial mortgage origination process, which can vary significantly

[1] Agarwal et al. (2016); Andersen et al. (2020); Archer and Ling (1993); Campbell (2006); Gerardi et al. (2023); Johnson et al. (2019); Keys et al. (2016); Mcconnell and Singh (1994); Schwartz and Torous (1989), and Stanton (1995).

[2] Earlier works include Clapp et al. (2001); Deng and Gabriel (2006); Firestone et al. (2007), and Kelly (1995). For more recent work, see Gerardi et al. (2023).

depending on how they were treated. For some, this experience is a positive milestone, symbolizing progress toward homeownership and financial stability. For others—particularly those who faced obstacles or hostility—it can be a stressful ordeal, discouraging them from repeating the process. These contrasting experiences during the significant event of buying a home may shape borrowers' willingness to pursue refinancing opportunities. If unequal treatment exists in the loan origination process across racial groups—e.g., loan officers ignoring minorities' inquiries or applying excessive scrutiny to minority applicants (Hanson et al., 2016; Korver-Glenn, 2018)[3]—lenders may effectively discourage minorities from seeking future refinancing opportunities.

To quantify unpleasant and potentially discriminatory experiences during the initial mortgage origination process, we measure loan processing times. Extended processing times are a strong candidate for explaining racial disparities in refinancing behavior through the negative prior experience channel for the following reasons. First, delays in loan processing are directly linked to consumer dissatisfaction.[4] Events exceeding 60 days for loan processing account for roughly 10% of loan originations, and typically surpass the standard mortgage rate lock period, which is commonly 30 to 60 days.[5] When processing extends beyond this period, borrowers face increased uncertainty, including risks of rate changes, additional lock-in costs, or even the failure to close the transaction on time (Han and Hong, 2024). Thus, by using both a continuous measure of processing time (in days) and a binary measure for delays of 60+ days, we capture a friction that likely shapes borrowers' perceptions of the mortgage process and affects their willingness to refinance.[6]

Second, traditional measures of discrimination in the lending market—such as disparities in approval rates or mortgage costs—are less suitable for examining this channel. On the extensive margin, refinancing decisions are conditional on borrowers already being mortgagors, meaning our sample includes only those who were previously approved. On the intensive margin, higher mortgage rates could either increase or decrease the incentive to refinance: they represent a negative experience

[3]An anonymous interview with a loan officer in Korver-Glenn (2018) highlights discriminatory practices in the mortgage application process: "*... a friend of mine at another mortgage company, who took an application from, uh, an African American couple. And he took the application, um, the income and everything was good—I mean, it was ... sort of taken as, like, 'This is crazy—there's no way they're doing this, there's no way they're doing that,' so they got scrutinized to death, um through a lot of due diligence ...*"

[4]The Consumer Financial Protection Bureau (CFPB) Consumer Complaint Database indicates that over 15% of complaints are related to delays in loan processing.

[5]https://www.consumerfinance.gov/ask-cfpb/whats-a-lock-in-or-a-rate-lock-en-143/

[6]We verify that our results are not sensitive to the choice of different thresholds for loan processing delays, such as those exceeding 45 days or 90 days.

per se, but they also offer greater potential savings from future refinancing.

Lastly, loan processing time provides the advantage of being an objectively measurable variable compared to other aspects of service quality, e.g., providing inadequate information or being disrespectful. While service quality can be assessed through survey data (e.g., the National Survey of Mortgage Originations) or the Consumer Complaint Database from the CFPB, these approaches do not allow for tracking future refinancing activities for the affected loans. In contrast, loan processing time is straightforward to calculate as the number of days between the loan application and origination dates, and it can be linked to subsequent loan performance.

Using a matched dataset of CoreLogic and the Single-Family Loan-Level Dataset for 18 states, we find that minority borrowers (Black and Hispanic) experience significantly longer mortgage processing times compared to white borrowers. Our estimates indicate that being a minority increases the likelihood of 60+ day processing delays by two percentage points—a 17% increase from the baseline average of 11.9%. This disparity remains robust across various specifications, including alternative race identification methods and different samples, such as the CoreLogic-only dataset and the matched dataset of CoreLogic and Home Mortgage Disclosure Act (HMDA).

Further additional evidence points to lender-side discrimination as the driver of these disparities. For instance, racial disparities in processing times are more pronounced in regions with heightened racial animus and in markets with lower competition, where taste-based discrimination is more likely to persist. Conversely, we find that FinTech lenders, which rely less on human involvement, and periods following CFPB actions against discriminatory lending[7], significantly reduce these disparities. These findings suggest that human biases and lender-side discrimination contribute to the observed inequities.

Having established that minority borrowers face longer processing times for initial mortgages, we next examine how these delays—disproportionately affecting minorities—discourage future refinancing. A key challenge in this analysis is addressing endogeneity. For example, unobserved borrower characteristics may introduce attenuation bias if lenders impose stricter scrutiny on borrowers with higher *ex ante* prepayment risk, thereby creating a positive correlation between processing delays and refinancing likelihood. Conversely, if borrower financial sophistication is associated with shorter

[7] We define the years following 2012 as the post-anti-discriminatory regulation period. For further details, refer to Appendix A4.

processing times and a greater propensity to refinance, the OLS estimates may overstate the true effect. These factors complicate the interpretation of the OLS estimates, highlighting the need for careful identification strategies.

To address this, we employ an instrumental variable (IV) strategy, using loan officer workload at the time of each loan's application as an instrument for processing delays. Higher workloads increase the likelihood of delays due to capacity constraints of individual officers but are plausibly unrelated to borrower characteristics, satisfying both the relevance and exogeneity conditions. The IV results using the quarterly loan performance panel of the CoreLogic–GSE matched dataset show that processing delays for the initial loans reduce quarterly refinancing rates by 1.3 percentage points. These findings indicate that initial loan origination experiences significantly influence borrowers' refinancing behavior, reducing refinancing activities by approximately one-third from the baseline quarterly average refinancing rate of 3.3%.

To better understand the mechanism, we decompose refinancing events into two categories: "recapture refinancing," where borrowers refinance with the same lender, and "switching refinancing," where borrowers refinance with a new lender. Our analysis reveals that the discouraging effect of delays is much more pronounced in recapture refinancing. Specifically, the coefficient estimate for recapture refinancing is -0.9, indicating that initial loan delays nearly eliminate refinancing with the current lender. In contrast, the coefficient estimate for switching refinancing is -0.3, which translates to a 13% decrease in such activities, and is statistically insignificant. This finding supports the interpretation that negative experiences during the initial mortgage process likely erode trust in the original lender responsible for the delay, discouraging future interactions with them. Collectively, our results present how discriminatory lending practices can create persistent barriers, effectively locking borrowers into less favorable loan terms and disproportionately burdening minority groups.

As a final step, we estimate the financial consequences of our findings for minority borrowers. Our back-of-the-envelope calculations indicate that, during our sample period, an additional 159,639 minority borrowers experienced delays of 60 days or more due to their racial status. These delays discouraged 17,414 minority borrowers from refinancing to lower-rate mortgages, resulting in an estimated annual aggregate overpayment of $36 million. If these borrowers remain locked into their existing loans for the remaining loan term, as assumed by Keys et al. (2016), the total overpayment would amount to $594 million, calculated using the average mortgage rate for 2009–2021 as the discount rate.

Related Literature We contribute to the extensive literature on mortgage refinancing, particularly studies documenting racial disparities in refinancing propensities (Clapp et al., 2001; Deng and Gabriel, 2006; Firestone et al., 2007; Gerardi et al., 2023; Kelly, 1995). Our paper advances this field by exploring the root causes of these disparities, moving beyond the view of race as an inherent trait. Specifically, by providing evidence that negative experiences during the initial mortgage process discourage subsequent refinancing, our builds on prior research examining why borrowers forgo refinancing opportunities despite potential cost savings (Keys et al., 2016; Johnson et al., 2019; Andersen et al., 2020; Defusco and Mondragon, 2020). In particular, our findings align closely with Johnson et al. (2019), who claim that past behavior by financial institutions damages borrower relationships, contributing to sluggish refinancing. Unlike their survey-based approach, we utilize detailed loan- and performance-level datasets, offering a more granular and quantifiable perspective.

Methodologically, we contribute to the refinancing literature by distinguishing prepayment events into refinancing, cash-out refinancing, and prepayments due to selling and moving. This distinction is essential, as prepayments driven by moving shocks, refinances motivated by rate reductions, and refnances for cash extractions are fundamentally different. Our approach is similar to Gerardi et al. (2023); Lambie-Hanson and Reid (2018), who differentiate refinancing from prepayments due to sales through borrower address changes. We extend this framework by identifying whether refinancing is for rate reduction or cash-out, using the loan purpose of new loans associated with the same borrower and address. More importantly, we differentiate between "recapture refinancing" (refinancing with the same lender) and "switching refinancing" (refinancing with a different lender). To our knowledge, this distinction is novel and opens new research avenues, such as exploring how lender promotional efforts shape borrowers' refinancing decisions.

Lastly, our findings are highly relevant to the growing body of research on racial disparities in the lending market (Ambrose et al., 2021; Bartlett et al., 2022; Bhutta et al., forthcoming; Butler et al., 2022; Frame et al., forthcoming; Munnell et al., 1996). While most studies focus on lending *outcomes*, such as unequal access to credit or disparities in credit costs, relatively few (Begley and Purnanandam, 2021; Wei and Zhao, 2022) examine subtler forms of discrimination in lending *process*. Recent research suggests that racial discrimination has declined due to decades of fair lending efforts

(Bhutta et al., forthcoming; Susin, 2024). However, we highlight that lenders may still engage in subtler forms of discrimination that are harder to detect. Importantly, subtle discrimination in the lending *process* are equally harmful by imposing significant financial costs through discouraged refinance activities.

The remainder of the paper is organized as follows. In Section 2, we explain the data and the key variables of interest with summary statistics. In Section 3, we empirically test our hypothesis and estimate the aggregate financial impact. Section 4 concludes.

2. Data

This study integrates two primary datasets forour empirical analysis: CoreLogic (deeds and MLS datasets) and the GSE Single-Family Loan-Level Dataset. By matching CoreLogic and the GSE data, we construct (1) a cross-sectional loan-level dataset and (2) a quarterly loan-performance panel (i.e., multiple observations for each loan). Details of each dataset and the matching procedure are outlined below.

2.1. CoreLogic

The CoreLogic deeds data contain comprehensive information on all deed transfers in the US, including sale amounts, property types, and property addresses, sourced 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, FHA/conventional loan status, loan origination dates, and borrowers' first and last names. The CoreLogic MLS data contain granular information on property listings, such as listing price, listing date, and the dates when sale contracts are signed and closed.

The deeds and MLS data are merged using CoreLogic's unique parcel identification numbers and sale closing dates. While the deeds data offer near-universal coverage across the US, the coverage of the MLS data varies by region.[8] To ensure a reliable and representative dataset, we restrict our analysis to 18 states where the MLS dataset covers more than 10% of mortgage records in the deeds

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

data. Appendix A1 lists the selected states and provides the coverage ratio of the MLS dataset in each state.

Our analyses focus on fixed-rate, 30-year purchase mortgages for single-family homes originated between 2009 and 2021. We also exclude loans made to institutional buyers and those with missing or unconventional features, resulting in the final CoreLogic Mortgage–MLS sample of 3.5 million observations. In addition to the explicitly provided information, we derive two key variables using the CoreLogic data: (1) borrowers' race/ethnicity and (2) loan processing time.

2.1.1. Identifying Borrower Race/Ethnicity

Borrower race and ethnicity are not directly observed in the CoreLogic dataset. Instead, we infer these attributes using borrower first and last names and location information through the Bayesian Improved First Name Surname Geocoding (BIFSG) method (Voicu, 2018).[9] 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:

$$p(r|s,f,z) = \frac{p(r|s) \times p(f|r) \times p(z|r)}{\sum_{r' \in White, Black, Hispanic, Asian, Native, Other} p(r'|s) \times p(f|r') \times p(z|r')}, \tag{1}$$

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.

We assign each borrower to the racial/ethnic group with the highest probability, following the approach of Ambrose et al. (2021) and Frame et al. (forthcoming). We corroborate the BIFSG race/ethnicity assignment using the imputed race from InfoUSA and self-reported race and ethnicity variable from the HMDA in Section 3.1.

[9]This method is increasingly used in the mortgage studies, such as Ambrose et al. (2021) and Frame et al. (forthcoming).

2.1.2. Measuring Loan Processing Time

We calculate *Loan Processing Time* as the number of days between the sale contract date and the mortgage origination date. While it might be ideal to use the loan application date as the starting point (Choi et al., 2022; Fuster et al., 2019; Wei and Zhao, 2022), we use the sale contract date as CoreLogic lacks application dates. This alternative serves as an appropriate proxy since lenders typically require a signed purchase agreement to initiate the underwriting process.

For validation of our *Loan Processing Time* variable, we compare it with the loan processing times by Wei and Zhao (2022) who use confidential HMDA data to measure the number of days between the application and origination dates. Panel A of Figure 1 displays the average processing times of Wei and Zhao (2022) by race for GSE loans during the 2001–2006 period, and panel B presents the corresponding averages of our *Loan Processing Time* variable.[10] The patterns observed in both panels are notably consistent. For instance, Black borrowers face the longest average processing times followed by white, Asian, and Hispanic borrowers. In addition, Black borrowers' average processing time increased from 2002 to 2003 and then decreased for the following three years with similar magnitudes in both panels. These similarities supports the accuracy of our *Loan Processing Time* measure.

2.2. GSE Single-Family Loan-Level Dataset

The GSE Single-Family Loan-Level Dataset provides information on mortgages purchased by Fannie Mae and Freddie Mac.[11] The GSE dataset complements CoreLogic by adding a few variables that are essential for analyzing refinancing behaviors but absent from CoreLogic, such as credit scores and mortgage interest rates.[12]

The variables provided by the GSE dataset include loan amount, origination date, maturity, interest rate, credit score, loan-to-value (LTV) ratio, and debt-to-income (DTI) ratio. In addition, it tracks monthly loan performances, documenting key credit events for each loan, such as prepayment,

[10] While our primary analysis period spans 2009–2021, we compute values for the 2001–2006 period specifically for comparison with Wei and Zhao (2022).

[11] According to HMDA, 59.2% of conventional purchase mortgages that are not sold to financial institutions (e.g., commercial banks) are securitized through either Fannie Mae or Freddie Mac.

[12] These two variables are particularly critical. Borrowers with higher credit scores tend to refinance at faster rates, and minority borrowers generally have lower credit scores compared to white borrowers (Gerardi et al., 2023). Additionally, 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.

90+ day delinquency, and foreclosure.

2.3. Matching CoreLogic and GSE Single-Family Loan-Level Dataset

We conduct matching between CoreLogic and the GSE dataset based on the loan characteristics, as there is no unique identifier. In particular, we filter both datasets to include only fixed-rate, 30-year purchase mortgages for single-family houses and match the loan records using the following variables: origination date, 3-digit ZIP code, CBSA code, loan amount, termination status (i.e., prepaid, foreclosed, or current), and termination date.[13] To ensure accuracy, we exclude duplicate observations and perform the matching without replacement. This process yields 264,789 uniquely matched loan-level dataset[14], which we can expand to a quarterly loan-performance panel with 3,952,418 observations.

To assess the representativeness of the matched dataset, Figure 2 compares key credit-related variables between the full population of the GSE dataset and the matched sample, using a snapshot from 2014. The figure shows that the distribution of the variables in the matched dataset closely resembles those in the population of the GSE loans.

2.4. Supplementary Datasets

In addition to the CoreLogic–GSE matched dataset, we utilize InfoUSA and HMDA data to provide richer context and robustness for our 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. The database provides exact home addresses and detailed household characteristics, such as the estimated age of the household head, estimated household income, and imputed race/ethnicity. By linking the CoreLogic–GSE dataset to InfoUSA[15], we integrate additional variables (e.g., the borrower's age group) into our regression analysis.

[13] Although CoreLogic does not explicitly provides loan performance information, we can identify this by our own algorithm. Details are provided in Appendix A2.

[14] The matching rate is approximately 18%, as the CoreLogic dataset contains 1,539,482 unique GSE loan observations identified by the matching variables.

[15] They can be matched using the exact address information.

HMDA The HMDA dataset offers information on the entire landscape of US mortgage applications, including loan application outcomes (e.g., approved or denied), applicant characteristics (e.g., income and self-reported race/ethnicity), and loan-level details (e.g., loan type, purpose, amount, and census tract location). By connecting HMDA data with CoreLogic[16], we can perform additional robustness tests. For example, we could validate our BIFSG imputation method for borrower race/ethnicity, leveraging HMDA's accurate, self-reported racial and ethnic classifications.

2.5. Summary Statistics

Table 1 presents summary statistics. Panel A provides descriptive statistics for the loan-level cross-section consisting of 264,789 observations of fixed-rate, 30-year purchase mortgages for single-family homes, originated between 2009 and 2021 across 18 states. The variable *Loan Processing Time* has a mean of 41.5 days and a standard deviation of 21.8 days. Approximately 11.9% of loans required more than 60 days for origination, as indicated by I(*Loan Processing Time* > *60 Days*). Borrower racial composition, imputed using the BIFSG method, reveals that 85% of borrowers are identified as white, 11.3% as minorities[17] (2.1% Black and 9.2% Hispanic), and 3.6% as Asian. 31.4% of borrowers are female and 51.1% of borrower have a co-borrower. The logarithm of estimated borrower monthly income, backed-out using loan amount, mortgage rate, and DTI ratio[18], has a mean of 8.17, which is equivalent to $3,533. The logarithm of loan amounts has a mean of 12.41, corresponding to an average loan amount of $245,242. The average LTV ratio at origination is 83.3% and the average FICO score is 754.8.

Panel B reports summary statistics for the quarterly loan performance panel, which is also constructed from the CoreLogic–GSE match. The average quarterly refinancing rate in our sample is 2.58%, with a standard deviation of 15.86%. Borrower refinancing behavior can be further categorized into two types: recapture (i.e., refinancing with the current lender) and switch (i.e., refinancing with a new lender).[19] The quarterly mean values of *Refinance (Recapture)* and *Refinance (Switch)*

[16] The matching procedure between HMDA and CoreLogic are discussed in Appendix A3.

[17] The minority share is comparable to HMDA records, where minority borrowers account for 11.5% (3.2% Black and 8.3% Hispanic) of the GSE purchase mortgage sample in the same period.

[18] Borrower monthly income is calculated as: $\frac{\textit{Monthly Payment}}{\textit{DTI Ratio}} \times 100$, where $\textit{Monthly Payment} = \frac{\textit{Loan Amount} \times r/12 \times (1+r/12)^{360}}{(1+r/12)^{360}-1}$.

[19] This information is identified from CoreLogic by determining whether the refinancing mortgage is originated by the same lender as the current one.

are 0.87% and 1.71%, respectively, indicating that approximately 33.7% ($\frac{0.87}{2.58}$) of borrowers refinance with their initial lenders. Borrower characteristics, such as racial composition, sex, the presence of a co-borrower, log-transformed monthly income and loan amount, LTV ratio at origination, and FICO score, are comparable to those reported in panel A. *Current LTV*, calculated as the remaining loan balance in each quarter divided by the current market value of the home (adjusted using a ZIP code-level Zillow Home Value Index), averages 68.9%. Loans have an average age of 10.2 quarters. *Rate Gap*, defined as the difference between the outstanding mortgage's coupon rate and the current available rate for similar mortgages[20], has a mean of 0.12 percentage points. Lastly, *Workload* is defined as the number of active loan applications the loan officer was managing when each loan was applied. The median loan officer handles three other applications, while those in the top 75th percentile manage seven applications concurrently. The sample size for *Workload* is smaller because the loan officer identifier (i.e., loan officer NMLS ID) has only been available in CoreLogic data since 2014.

3. Empirical Results

3.1. Racial Disparities in Loan Processing Times

We begin our empirical analysis by examining whether race influences the processing time of loan applications. Using the loan-level CoreLogic–GSE dataset, we estimate the following regression equation:

$$\begin{aligned} Y_i = &\alpha + \beta_1 \cdot Minority_i + \beta_2 \cdot Asian_i + \beta_2 \cdot Other\ Race_i + \delta \cdot X_i + \eta_{age\ group} \\ &+ \eta_{county \times origin\ year} + \eta_{lender} + \epsilon_i, \end{aligned} \tag{2}$$

where the dependent variable, Y, is either *Loan Processing Time*, the number of days required to obtain a mortgage, or I(*Loan Processing Time* > *60 Days*), a binary indicator equal to 1 if the processing time exceeded 60 days. 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

[20]Following Berger et al. (2021), the current available market rate is calculated as the monthly average 30-year fixed-rate mortgage rate from the Freddie Mac Primary Mortgage Market Survey, adjusted by an estimated loan-specific factor that is a quadratic function of the borrower's FICO score and the loan's current LTV ratio.

included in the regression.

The regression controls a comprehensive set of borrower- and loan-level characteristics at origination, denoted by X_i, which may influence processing times. These controls include dummies for female and presence of a co-borrower, the logarithms of borrower income and loan amount, origination LTV ratio, and FICO score. Fixed effects for borrower age groups obtained from InfoUSA ($\eta_{age\ group}$), county by origin year ($\eta_{county \times origin\ year}$), and lender ($\eta_{lender}$) are also included to account for unobserved heterogeneity across borrower demographics, time-varying local economic conditions, and lender-specific practices.

Table 2 presents the regression results. We find that minority borrowers spend significantly longer times in securing their mortgages. In column (1), we regress *Loan Processing Time* on *Minority*, and the coefficient for *Minority* is 1.83. This indicates that Black and Hispanic borrowers wait an additional 1.83 days for mortgage approval compared to white borrowers. With a mean loan processing time of 41.52 days, this corresponds to a 4.4% increase ($\frac{1.82}{41.52}$).

While statistically significant, an extra 1.82 days for loan processing may appear economically modest. However, the economic implications become more evident when we focus on tail events. From column (2) of Table 2, we use I(*Loan Processing Time* > *60 Days*)—an indicator for the event where loan processing time exceeds 60 days—as the dependent variable. In column (2), the coefficient for *Minority* is 0.0206, suggesting that minority borrowers are 2.06 percentage points more likely to face a processing delay of 60+ days. Given that, on average, 11.9% of borrowers experience such delays, this represents a 17.3% increase ($\frac{2.06\%}{11.9\%}$), highlighting significant disparities in occurrence of extended processing times.

In column (3), we separately estimate the effects for Black and Hispanic borrowers. Black borrowers face a 3.01 percentage point increase in the likelihood of a 60+ day delay, while Hispanic borrowers experience an increase of 1.82 percentage points. Both estimates are statistically significant at the 1% level, and the difference test for the coefficient estimates is significant at the 5% level with t-statistics of 2.31. These findings indicate that even among minority groups, Black borrowers face particularly severe delays.

In column (4), we address concerns about measurement error in the BIFSG race assignment using the imputed race information in the InfoUSA dataset.[21] While the magnitude of the coefficient for

[21]InfoUSA does not disclose the methodology used to impute race data, which limits transparency and necessitates

Minority becomes smaller, the results remain consistently positive, corroborating our findings and race measurement of the BIFSG algorithm.

Although our CoreLogic–GSE matched dataset appears to represent the broader population as shown in Figure 2, concerns may arise regarding the relatively small sample size of the dataset in Table 2. To mitigate this concern, in Table 3, we assess the robustness of our results using two alternative samples.

In the first two columns of Table 3, we replicate the regression from Table 2 using only the CoreLogic dataset. The CoreLogic-only sample has the advantage of a larger sample size, which includes loans beyond GSE loans, such as FHA-insured mortgages. However, a key limitation of this dataset is the absence of borrower FICO scores, which may introduce omitted variable bias if FICO scores are correlated with both race and loan processing times.

Bearing these advantages and limitations in mind, the results are compelling. In column (1), the estimated coefficient for *Minority* is 0.0246, with a t-statistic exceeding 20.[22] In column (2), focusing specifically on FHA loans within the CoreLogic sample[23], we find even a higher coefficient of 0.0275. The fact that our results remain consistent across loan types reinforces that minority borrowers are disproportionately exposed to delays in loan processing.

The next two columns leverage the matched CoreLogic–HMDA dataset.[24] While still missing FICO scores, HMDA provides reliable, self-reported borrower race/ethnicity information, enabling us to check the accuracy of the BIFSG race imputation.[25] In column (3), the coefficient for *Minority (HMDA)*, representing borrowers who reported themselves as Black or Hispanic in HMDA, is 0.0189, and statistically significant at the 1% level, which is comparable to our previous estimates.

Finally, in column (4), we address potential measurement concerns regarding *Loan Processing Time*. A longer *Loan Processing Time* might be a result of borrowers engaging in multiple applications to secure better terms.[26] If this were the case, our earlier findings may reflect differences in shopping

cautious interpretation. However, this information could be valuable for validating our BISFG method.

[22] The larger magnitude compared to the matched dataset may reflect unobserved differences in FICO scores across race groups.

[23] FHA loans are more commonly utilized by minority borrowers. In our FHA sample, for example, minorities represent 33.0% of borrowers (8.2% Black and 24.8% Hispanic), and this proportion is significantly higher than in the GSE loan sample (11.3%).

[24] Further details on the matching procedure can be found in Appendix A3.

[25] The accuracy rates (i.e., $\frac{\text{Correct Prediction}}{\text{Total Observation}}$) of the BIFSG imputation are 79.4% for whites, 91.1% for minorities, and 98.1% for Asian.

[26] This could be a valid concern, particularly because we calculate *Loan Processing time* as the difference between the sale contract date and the mortgage origination date.

behaviors rather than lender-driven delays. To rule out this possibility, we exclude all the loan observations that are potentially applied multiple times. The results using the subsample that involves only one application attempt remain consistent. The coefficient for *Minority (HMDA)* slightly increases to 0.0192, suggesting that minorities face significant loan processing delays even after accounting for potential confounding factors such as shopping behaviors.

3.2. Evidence of Lender Discrimination in Loan Processing Times

We further investigate whether the observed racial disparities in loan processing times are associated with lending market discrimination. To do so, we conduct four indirect tests that exploit geographical and temporal variations to identify whether racial disparities are more pronounced in settings with *ex ante* greater discrimination.

First, we examine whether racial disparities are larger in regions with higher levels of racial animus. We use the racial animus measure based on the frequency of racially charged Google search terms in metropolitan statistical areas (MSAs) constructed by by Stephens-Davidowitz (2014). Column (1) of Table 4 presents the results, showing a positive and statistically significant coefficient on the interaction between *Minority* and an indicator for areas with above-median racial animus. In particular, the coefficient estimates indicate that the probability of 60+ day processing delays for minority borrowers is approximately 1.8 times higher in regions with greater racial hostility. This finding aligns with prior research demonstrating that discriminatory practices are more pronounced in areas with heightened racial animus in the consumer credit, labor, and tax-exmpt bond issuance markets (Butler et al., 2022; Charles and Guryan, 2008; Dougal et al., 2019).

Next, we explore whether racial disparities are larger in less competitive lending markets. Markets dominated by a few large players may allow taste-based discriminatory practices to persist due to a lack of competitive pressure (Berkovec et al., 1998). In column (2) of Table 4, we include an interaction term between *Minority* and an indicator for low market competition, defined as counties in the top tercile of the top-4 lender market share. The positive and significant coefficient on this interaction term suggests that minorities are more likely to face delays in less competitive markets, providing further evidence that lender-side racial preferences contribute to disparities.

Third, we leverage the rise of FinTech lenders during our sample period to assess whether minimizing human involvement in loan processing mitigates racial disparities, possibly reducing discriminatory lending practices by biased loan officers (Howell et al., 2024). Column (3) of Table 4 presents the result, where we include an interaction term between *Minority* and a FinTech lender dummy, as defined by Fuster et al. (2019). The coefficient of *Minority* × *FinTech Lender* is significantly negative, suggesting that FinTech lenders significantly reduce racial disparities in loan processing times. Our finding thus supports the hypothesis that human biases are the driver of racial disparities in lending processes.

Lastly, we test whether racial disparities have declined following the implementation of stricter anti-discrimination regulations. Around 2012–2013, the CFPB implemented enhanced oversight to curb discriminatory lending practices.[27] In column (4) of Table 4, we include an interaction term between *Minority* and a post-2013 dummy to find the negative and significant coefficient on this interaction term. A reduction in racial disparities in loan processing times after these regulatory actions further supports that racial disparities in loan processing time may be driven by discriminatory lending practices.

Collectively, these indirect test results provide compelling evidence that racial disparities in loan processing times are associated with lender-side discrimination. To confirm the robustness of these findings, we repeat the regressions in Table 4 using the larger CoreLogic-only dataset. The results, reported in Table A2, remain consistent across all four specifications.[28] Additionally, in Appendix A5, we test whether lender-specific overlays are positively correlated with the additional delays experienced by minority borrowers, following the methodology of Bhutta et al. (forthcoming). The findings from this test further reinforce lender-side discrimination as a key driver of racial disparities in loan processing times.

3.3. Initial Mortgage Delays and Refinancing Behavior

Having established that minority borrowers experience longer processing times for initial mortgages, we now examine whether these delay experiences discourage subsequent refinancing activities. To address this question, we analyze quarterly loan-performance data from the matched CoreLogic–GSE

[27] See Appendix A4 for more details.

[28] While the CoreLogic-only sample lacks borrower FICO scores, it is implausible that an omitted variable is systematically correlated with all four contextual factors—heightened racial animus, lower local market competition, FinTech lending, and the post-CFPB regulatory action period—further strengthening our conclusion that lender-side discrimination contributes to these disparities.

dataset. Specifically, we estimate the following regression equation:

$$\begin{aligned} \textit{Refinance}_{i,t} =& \alpha + \beta \cdot \mathrm{I}(\textit{Loan Processing Time} > \textit{60 Days})_i + \delta \cdot X_{i,t} + \eta_{\textit{age group}} \\ &+ \eta_{\textit{county} \times \textit{origin year}} + \eta_{\textit{year-quarter}} + \eta_{\textit{lender}} + \epsilon_{i,t}, \end{aligned} \tag{3}$$

where $\textit{Refinance}_{i,t}$ is an indicator variable whether loan i was refinanced in quarter t. The key independent variable, $\mathrm{I}(\textit{Loan Processing Time} > \textit{60 Days})_i$, is a dummy that equals 1 if loan i had a delay longer than 60 days for its origination process.

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, income, loan amount, origination and quarterly updated LTV ratios, FICO scores, loan age, and the refinancing incentive measured by the rate gap.[29] The rate gap at time t is calculated as the difference between the coupon rate on the outstanding loan and the current available mortgage rate to the borrower if he refinances at t, where the latter value is estimated from a quadratic function of the borrower's FICO score and updated LTV ratio.[30] Fixed effects are included for borrower age groups, county by origin year, year-quarter, and lender to control for unobserved heterogeneity across borrowers, locations, time periods, and lenders.

Columns (1) of Table 5 present the estimation results of Equation (3). We find a pattern consistent with the story that the prior delay experience is associated with reduction in refinancing activity. The coefficient of $\mathrm{I}(\textit{Loan Processing Time} > \textit{60 Days})$ is -0.103, meaning that the delay experience reduces quarterly refinancing activities by 0.1 percentage points. Borrowers with bad memories from the initial loan delays may be disincentivized from going through the loan application process again. The control variables also exhibit expected patterns: for instance, larger rate gaps, which reflect stronger financial incentives for refinancing, significantly enhance refinancing activities.

3.3.1. Instrumental Variable Strategy

While the OLS estimates in column (1) provide valuable insights, the specification in Equation (3) may suffer from endogeneity concerns which could underestimate the true effects mainly in two ways. First, our delay indicator, $\mathrm{I}(\textit{Loan Processing Time} > \textit{60 Days})$, may partially capture

[29] To capture nonlinearities, we also include square terms for these controls.

[30] Consistent with Berger et al. (2021) and Scharlemann and van Straelen (2024), we observe that refinancing probabilities exhibit a "step-like" nonlinear pattern, as shown in Figure 3.

borrower- or seller-driven delays unrelated to lending practices. For instance, borrowers might request extended closing periods due to financial or logistical constraints, or sellers might delay transactions to accommodate their own moving schedules. These types of delays are unlikely to affect refinancing behavior, and thus the estimated average effect from the OLS would be attenuated.

Second, unobserved borrower characteristics may simultaneously influence both initial processing delays and refinancing propensity. For example, borrowers with higher prepayment risk—those more likely to refinance—might face stricter scrutiny from lenders during the initial loan process, resulting in longer processing times.[31] Since prepayment risk is positively correlated with both delays and refinancing likelihood, the OLS estimates would attenuate the true discouraging effect of loan processing delays on refinancing behavior.

To address these concerns, we employ an IV strategy, using the loan officer's workload at the time of application as an instrument for processing delays. We quantify *Workload* as the number of active (i.e., incomplete) loan applications the officer was managing when a new application was submitted. Using the instrument, we estimate the following 2SLS specification, adapted from Equation (3):

(First Stage)

$$\begin{aligned} \mathrm{I}(\textit{Loan Processing Time} > \textit{60 Days})_i = \alpha + \beta \cdot \textit{Workload}_i + \delta \cdot X_{i,t} + \eta_{\textit{age group}} + \eta_{\textit{county}\times\textit{origin year}} \\ + \eta_{\textit{year-quarter}} + \eta_{\textit{lender}} + \eta_{\textit{processing speed quintile}} + \epsilon_{i,t}, \end{aligned} \tag{4}$$

(Second Stage)

$$\begin{aligned} \textit{Refinance}_{i,t} = \alpha + \beta \cdot \mathrm{I}(\widehat{\textit{Loan Processing Time}} > \textit{60 Days})_i + \delta \cdot X_{i,t} + \eta_{\textit{age group}} \\ + \eta_{\textit{county}\times\textit{origin year}} + \eta_{\textit{year-quarter}} + \eta_{\textit{lender}} + \eta_{\textit{processing speed quintile}} + \epsilon_{i,t}. \end{aligned} \tag{5}$$

The first stage leverages the idea that capacity constraints of loan officers, captured by *Workload*, predict processing delays. This builds on the idea of Choi et al. (2022), who demonstrate that capacity-constrained banks reduce mortgage originations to new home buyers, particularly during periods of high demand.[32] Our *Workload* variable captures similar but time-varying capacity constraints at the loan officer-level, which makes processing delays more likely when officers are busy

[31] Lenders often retain servicing rights after securitization, which incentivizes them to screen against borrowers with high prepayment risks (Buttimer and Lin, 2005; Mayock and Shi, 2022).

[32] Choi et al. (2022) measure bank operating capacity by the ratio of incomplete applications at the end of each quarter to the total number of applications received in that quarter.

handling other applications.

Columns (1) and (2) of Table 6 present the first-stage results. Note that column (2) additionally includes fixed effects for the average processing speed quintile of individual loan officers. By controlling for the loan officer's average processing speed, we isolate the workload effect from officer-specific characteristics such as efficiency or scrutiny. In both columns, *Workload* strongly predicts 60+ day delays, with first-stage F-statistics exceeding 15.

The exclusion restriction of the IV specification assumes that heavy loan officer workload affect refinancing activities solely through their impact on loan processing delays. This assumption may be violated if borrowers with certain characteristics—such as minorities or females—were systematically assigned to the busiest loan officers. Although this assumption is not directly testable, columns (3) and (4) of Table 6 provide indirect evidence. None of the observed borrower characteristics, race/ethnicity, sex, presence of a co-borrower, income, loan amount, LTV ratio, and FICO score, shows significant associations with *Worload*, suggesting that being assigned to capacity-constrained loan officers is a random event to borrowers.

Columns (2) and (3) of Table 5 present the IV estimates. We find that processing delays reduce refinancing activity by 1.18 to 1.28 percentage points every quarter. These estimated effects are substantially larger than the OLS estimates in column (1), likely due to the correction of attenuation bias in the IV specification. This underscores the importance of addressing endogeneity concerns and isolating lender-driven delays from borrower- or seller-related factors.

To better understand the mechanisms underlying these effects, we decompose refinancing into two categories: "recapture" refinancing, where the borrower refinances with the same lender, and "switching" refinancing, where the borrower moves to a new lender. Then, we re-run our IV regressions using *Refinance (Recapture)* and *Refinance (Switch)* as the dependent variables, respectively. Since delays during the initial mortgage process are tied more to unpleasant experiences with the current lender, rather than the whole mortgage lending industry, we expect recapture refinancing should be affected more than switching refinancing.

The last four columns of Table 5 presents the IV results for these two subcategories of refinancing. The significant negative coefficients in columns (4) and (5) indicate that delays reduce recapture refinancing by 0.88 to 0.94 percentage points, suggesting that borrowers are less likely to refinance with lenders responsible for their initial delays. While the absolute magnitudes of the estimates are smaller than those in columns (2) and (3), the relative effect sizes are larger when we consider the means of the dependent variables (3.333% for *Refinance* and 1.081% for *Refinance (Recapture)*).

In contrast, the coefficients in columns (6) and (7) for switching refinancing are small and statistically insignificant, indicating that delays during the initial mortgage process do not reduce, nor increase, borrowers' decisions to refinance with a new lender. This asymmetry highlights that disutility from the initial processing delay is the main channel that leads borrowers to refinancing less frequently or not refinancing at all, locking them into less favorable loan terms.[33]

3.4. Estimating Aggregate Financial Impact: Back-of-the-Envelope Calculations

So far, we have demonstrated that (i) minority borrowers are disproportionately exposed to delays in mortgage processing and that (ii) these delays discourage borrowers from future refinancing. We now shift our focus to understanding the broader financial implications of our findings: the aggregate impact of the loan processing delays on minority borrowers. To quantify this, we conduct a back-of-the-envelope calculation to estimate the extent of missed refinancing opportunities among minority borrowers through this channel.

To begin, we calculate the number of loans that experienced processing delays attributable to borrowers being minorities. According to HMDA data, 7,749,470 purchase mortgages were originated by Black and Hispanic borrowers during our sample period (2009–2021). Column (1) of Table 2 indicates that Black and Hispanic borrowers are 2.06 percentage points more likely to experience a delay of 60+ days in obtaining their initial mortgage. This implies that approximately 159,639 borrowers (i.e., $7,749,470 \times 0.0206$) face additional delays simply due to their minority status.

Next, we estimate the number of loans that are discouraged from refinancing due to these delays. We begin by calculating the number of refinances that would have occurred if delays had no discouraging effect. The average quarterly refinance rate of 2.58% implies a survival rate of $100\% - 2.58\% = 97.42\%$ of loans per quarter. Using this, we estimate the number of loans that would have been refinanced within the first 10.2 quarters—the average loan age in our sample—as

[33] Table A3 presents additional IV estimation results for *Cash-Out Refinance*, *Cash-Out Refinance (Recapture)*, *Cash-Out Refinance (Switch)*, and *Prepaid Due to Move*. The findings for cash-out refinances align with the refinance results in Table 5, showing a statistically significant effect for recaptured cash-out refinances but an insignificant effect for switching cash-out refinances. Additionally, the effect of initial delays on moves is negative but not statistically significant.

follows:

$$\underbrace{159,639}_{\text{loans originated}} - \underbrace{159,639 \times (0.9742)^{10.2}}_{\text{loans survived}} = \underbrace{37,361.}_{\text{loans refinanced}} \tag{6}$$

We then account for the discouraging effect of initial delays. Column (2) of Table 5 shows that borrowers who experience a delay of 60+ days are 1.28 percentage points less likely to refinance each quarter. This reduces the quarterly refinance rate to $2.58\% - 1.28\% = 1.30\%$, or a survival rate of $100\% - 1.30\% = 98.7\%$. Using this adjusted rate, the number of refinanced loans decreases to:

$$\underbrace{159,639}_{\text{loans originated}} - \underbrace{159,639 \times (0.9870)^{10.2}}_{\text{loans survived}} = \underbrace{19,946.}_{\text{loans refinanced}} \tag{7}$$

The difference between these two numbers, $37,361 - 19,946 = 17,414$, represents the number of minority borrowers who forgo refinancing due to initial processing delays during our sample period.

Discussion This calculation highlights the significant aggregate impact of processing delays, especially for minority borrowers. First, the financial consequences for minority borrowers are substantial. With the average loan balance of \$239,276 and the average rate gap of 0.87% when refinancing, the annual savings per borrower from refinancing would be:

$$\$239,276 \times 0.87\% = \$2,082. \tag{8}$$

For 17,414 borrowers who missed the opportunity to refinance, this translates into an aggregate annual overpayment of \$36 million (i.e., $\$2,082 \times 17,414$). If these borrowers remain locked into their existing loan terms for the remainder of the loan period, as assumed by Keys et al. (2016), the total overpayment would reach \$594 million, calculated using the average mortgage rate for 2009–2021, 4%, as the discount rate.[34]

Second, these discouraged refinances can be understood as an indirect form of loan rejection for minority borrowers. Bhutta et al. (forthcoming) report a 1.4 percentage point racial gap[35] in refinance

[34] $\$36 \text{ million} \times \frac{1-(1+0.0685)^{-27.5}}{0.04}$.

[35] This is calculated as the weighted average of the coefficients in column (2) of Table A.4 for *Black* (0.023) and *Hispanic* (0.009) in Bhutta et al. (forthcoming), weighted by their respective shares.

approval rates between minorities and white borrowers. They caution that their estimates could mislead the true effect if there is a substantial amount of indirect discouragement—e.g., lenders not responding to minority inquiries—and our findings clearly illustrate this phenomenon. Specifically, during our sample period, approximately 4.5 million refinance mortgages were originated to Black and Hispanic borrowers. Our estimate of 17,414 discouraged refinances translates to an additional 0.39 percentage point gap (i.e., $\frac{17{,}414}{17{,}414+4.5\text{ million}}$) in refinance approval rates, which represents more than 27% increase in the racial gap from 1.4 percentage points by Bhutta et al. (forthcoming).

4. Conclusion

This paper explores a critical yet underexamined topic in the refinancing literature: the drivers of racial disparities in mortgage refinancing. By analyzing loan processing times using a matched dataset from CoreLogic and the Single-Family Loan-Level Dataset, we demonstrate that minority borrowers face significantly longer delays during initial mortgage origination. These disparities are particularly pronounced in regions with higher racial animus and less competitive markets, while they are less evident in FinTech lending and during periods of stronger regulatory oversight. This pattern suggests that the racial gap in loan processing times is linked to discriminatory practices in the lending market.

Our findings show that these delays discourage refinancing activity, particularly "recapture refinancing," indicating that negative experiences erode trust in lenders and deter future interactions. Using an IV approach to address endogeneity concerns, we provide robust evidence that these delays significantly reduce refinancing rates.

The financial implications are substantial. Minority borrowers discouraged by delays miss refinancing opportunities, resulting in considerable lifetime overpayments. Our analysis estimates an aggregate cost of $594 million for minority borrowers due to these missed opportunities. These findings highlight the significant amount of financial burdens imposed by subtle forms of discrimination, even as overt racial disparities in mortgage approval rates have declined in recent decades.

References

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.

Archer, W. R., and D. C. Ling. 1993. Pricing Mortgage-Backed Securities: Integrating Optimal Call and Empirical Models of Prepayment. *Real Estate Economics* 21:373–404.

Bartlett, R., A. Morse, R. Stanton, and N. Wallace. 2022. Consumer-Lending Discrimination in the FinTech Era. *Journal of Financial Economics* 143:30–56.

Begley, T. A., and A. Purnanandam. 2021. Color and Credit: Race, Regulation, and the Quality of Financial Services. *Journal of Financial Economics* 141:48–65.

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.

Berkovec, J. A., G. B. Canner, S. A. Gabriel, and T. H. Hannan. 1998. Discrimination, Competition, and Loan Performance in FHA Mortgage Lending. *Review of Economics and Statistics* 80:241–250.

Bhutta, N., A. Hizmo, and D. Ringo. forthcoming. How Much Does Racial Bias Affect Mortgage Lending? Evidence from Human and Algorithmic Credit Decisions. *Journal of Finance* .

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.

Buttimer, R. J., and C.-C. Lin. 2005. Valuing US and Canadian Mortgage Servicing Rights with Default and Prepayment. *Journal of Housing Economics* 14:194–211.

Campbell, J. Y. 2006. Household Finance. *Journal of Finance* 61:1553–1604.

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.

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.

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.

Deng, Y., and J. M. Quigley. 2012. Woodhead Behavior and the Pricing of Residential Mortgages. NUS Institute of Real Estate Studies Working Paper Series IRES2012-025.

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.

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.

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* .

Fuster, A., M. Plosser, P. Schnabl, and J. Vickery. 2019. The Role of Technology in Mortgage Lending. *Review of Financial Studies* 32:1854–1899.

Gerardi, K., P. S. Willen, and D. H. Zhang. 2023. Mortgage Prepayment, Race, and Monetary Policy. *Journal of Financial Economics* 147:498–524.

Han, L., and S.-H. Hong. 2024. Cash Is King? Understanding Financing Risk in Housing Markets. *Review of Finance* 28:2083–2118.

Hanson, A., Z. Hawley, H. Martin, and B. Liu. 2016. Discrimination in Mortgage Lending: Evidence from a Correspondence Experiment. *Journal of Urban Economics* 92:48–65.

Howell, S. T., T. Kuchler, D. Snitkof, J. Stroebel, and J. Wong. 2024. Lender Automation and Racial Disparities in Credit Access. *Journal of Finance* 79:1457–1512.

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.

Kelly, A. 1995. Racial and Ethnic Disparities in Mortgage Prepayment. *Journal of Housing Economics* 4:350–372.

Keys, B. J., D. G. Pope, and J. C. Pope. 2016. Failure to Refinance. *Journal of Financial Economics* 122:482–499.

Korver-Glenn, E. 2018. Compounding Inequalities: How Racial Stereotypes and Discrimination Accumulate across the Stages of Housing Exchange. *American Sociological Review* 83:627–656.

Lambie-Hanson, L., and C. Reid. 2018. Stuck in Subprime? Examining the Barriers to Refinancing Mortgage Debt. *Housing Policy Debate* 28:770–796.

Mayock, T., and L. Shi. 2022. Adverse Selection in the Market for Mortgage Servicing Rights. *Journal of Housing Economics* 58:101858.

Mcconnell, J. J., and M. Singh. 1994. Rational Prepayments and the Valuation of Collateralized Mortgage Obligations. *Journal of Finance* 49:891–921.

Munnell, A. H., G. M. B. Tootell, L. E. Browne, and J. McEneaney. 1996. Mortgage Lending in Boston: Interpreting HMDA Data. *American Economic Review* 86:25–53.

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.

Schwartz, E. S., and W. N. Torous. 1989. Prepayment and the Valuation of Mortgage-Backed Securities. *Journal of Finance* 44:375–392.

Stanton, R. 1995. Rational Prepayment and the Valuation of Mortgage-Backed Securities. *Review of Financial Studies* 8:677–708.

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.

Susin, S. 2024. Racial and Ethnic Mortgage Rate Disparities, 2000-2023. FHFA Insights.

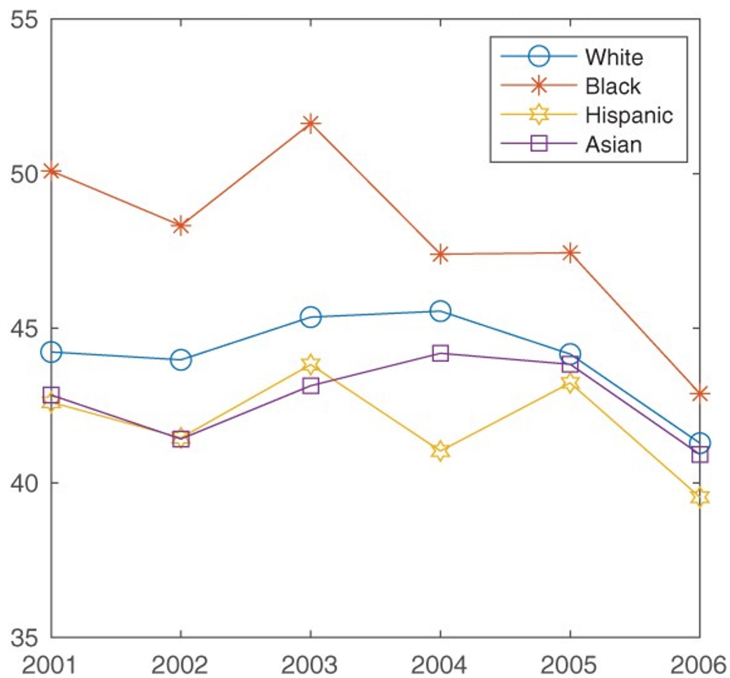
Voicu, I. 2018. Using First Name Information to Improve Race and Ethnicity Classification. *Statistics and Public Policy* 5:1–13.

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.

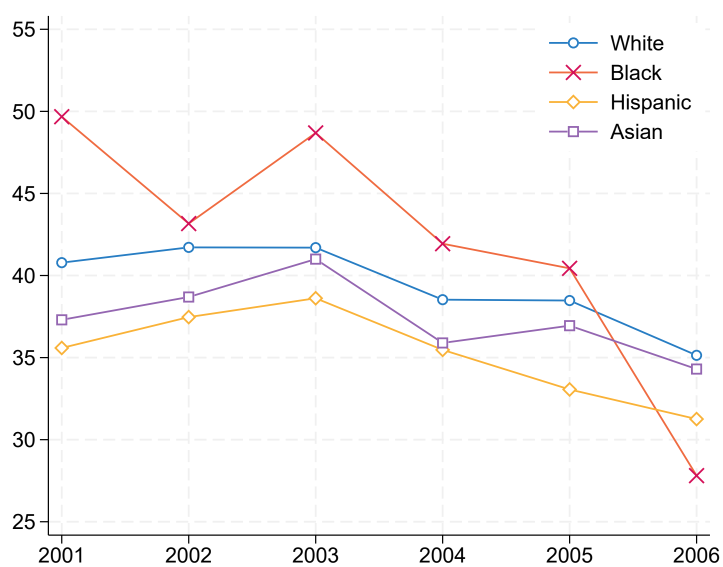
Zhang, D. 2024. Closing Costs, Refinancing, and Inefficiencies in the Mortgage Market. Working Paper.

Figure 1. Comparison of *Loan Processing Time* Measure with Wei and Zhao (2022)

This figure compares loan processing times by racial groups for loans originated between 2001 and 2006, as reported in Wei and Zhao (2022) and based on our own calculations.



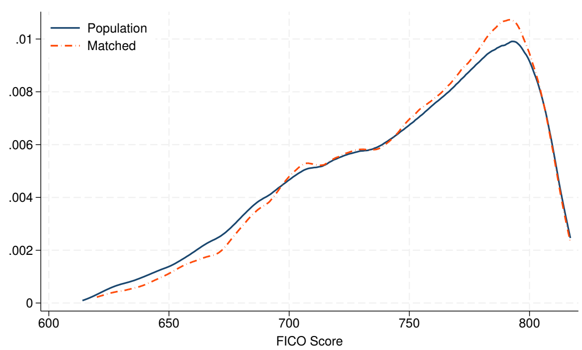
(a) Calculation by Wei and Zhao (2022)



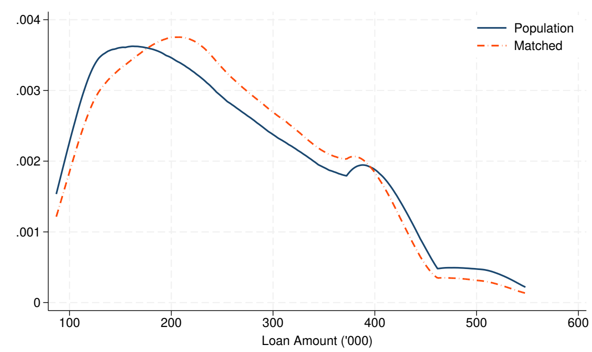
(b) Our Own Calculation

Figure 2. Kernel Density Plot of Key Variables

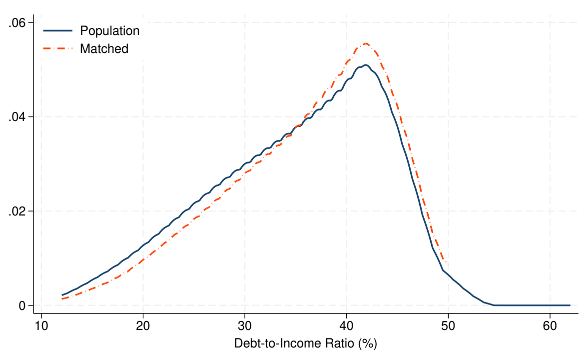
This figure compares distributions in the population GSE dataset with those in the matched CoreLogic–GSE dataset by drawing kernel density plots for key variables, including *FICO Score*, *Loan Amount*, *DTI Ratio*, and *LTV Ratio*, using a 2014 snapshot.



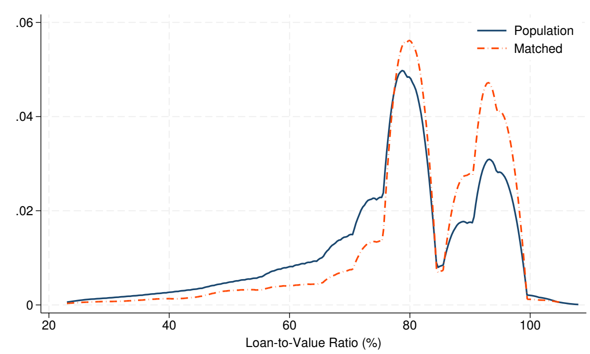
(a) FICO Score



(b) Loan Amount



(c) DTI Ratio



(d) LTV Ratio

Figure 3. Average Quarterly Refinancing Rates By Rate Gaps

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

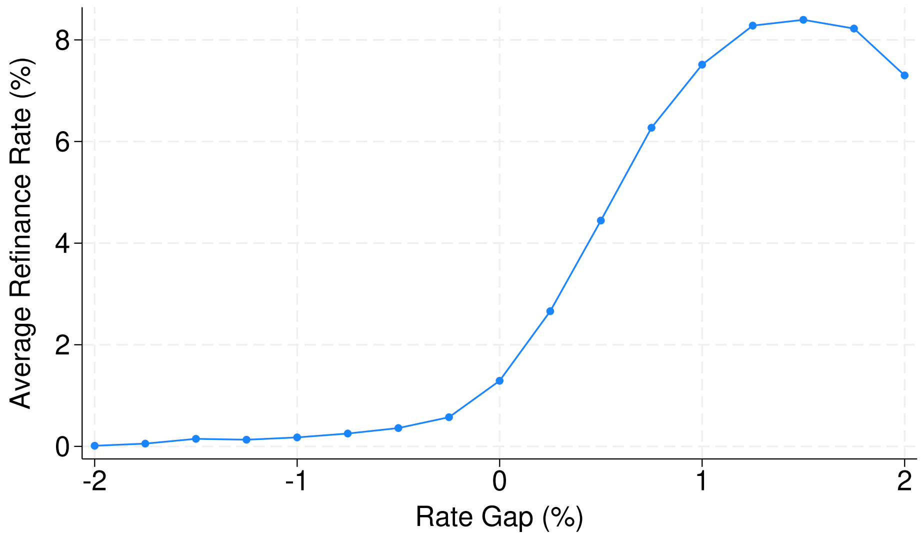


Table 1. Summary Statistics

This table presents summary statistics for the matched CoreLogic and GSE Single-Family Loan-Level Dataset. Panel A reports descriptive statistics for the loan-level cross-sectional dataset, and Panel B summarizes key variables for the quarterly loan performance dataset.

	Obs.	Mean	S.D.	P25	P50	P75
Panel A: Loan-Level Cross-Sectional Dataset						
Loan Processing Time (Days)	264,789	41.516	21.762	30.000	38.000	48.000
I(Loan Processing Time > 60 Days)	264,789	0.119	0.324	0.000	0.000	0.000
White	264,789	0.850	0.357	1.000	1.000	1.000
Minority	264,789	0.113	0.317	0.000	0.000	0.000
Black	264,789	0.021	0.144	0.000	0.000	0.000
Hispanic	264,789	0.092	0.289	0.000	0.000	0.000
Asian	264,789	0.036	0.186	0.000	0.000	0.000
Other Race	264,789	0.001	0.037	0.000	0.000	0.000
Female	264,789	0.314	0.464	0.000	0.000	1.000
Coborrower	264,789	0.511	0.500	0.000	1.000	1.000
ln(Income)	264,789	8.171	0.544	7.827	8.208	8.542
ln(Loan Amount)	264,789	12.408	0.515	12.096	12.468	12.789
LTV (%)	264,789	83.274	12.666	80.000	80.000	95.000
FICO (100 pts)	264,789	7.548	0.814	7.260	7.630	7.890
Panel B: Quarterly Loan-Performance Panel Dataset						
Refinance	3,952,418	2.581	15.858	0.000	0.000	0.000
Refinance (Recapture)	3,952,418	0.870	9.286	0.000	0.000	0.000
Refinance (Switch)	3,952,418	1.711	12.970	0.000	0.000	0.000
I(Loan Processing Time > 60 Days)	3,952,418	0.130	0.336	0.000	0.000	0.000
White	3,952,418	0.856	0.351	1.000	1.000	1.000
Minority	3,952,418	0.108	0.311	0.000	0.000	0.000
Black	3,952,418	0.022	0.148	0.000	0.000	0.000
Hispanic	3,952,418	0.086	0.281	0.000	0.000	0.000
Asian	3,952,418	0.034	0.182	0.000	0.000	0.000
Other Race	3,952,418	0.001	0.036	0.000	0.000	0.000
Female	3,952,418	0.313	0.464	0.000	0.000	1.000
Coborrower	3,952,418	0.505	0.500	0.000	1.000	1.000
ln(Income)	3,952,418	8.104	0.558	7.747	8.141	8.488
ln(Loan Amount)	3,952,418	12.325	0.529	11.995	12.382	12.719
LTV at Origination (%)	3,952,418	83.249	12.614	80.000	80.000	95.000
Current LTV (%)	3,952,418	68.970	17.030	58.195	71.584	81.198
FICO (100 pts)	3,952,418	7.550	0.423	7.260	7.640	7.900
Loan Age	3,952,418	10.225	9.033	3.000	8.000	15.000
Rate Gap (%)	3,952,418	0.124	0.805	-0.358	0.108	0.627
Workload	2,294,265	5.682	6.703	2.000	3.000	7.000

Table 2. Baseline Estimates of Racial Disparities in Loan Processing Times

This table presents regression results examining the effect of borrower minority status on loan processing times. The analysis uses loan-level observations from the CoreLogic–GSE dataset for loans originated between 2009 and 2021. In column (1), the dependent variable is *Loan Processing Time*, measured as the number of days between the sale contract date and the mortgage origination date. In columns (2)–(4), the dependent variable is I(*Loan Processing Time* > *60 Days*), an indicator equal to 1 if *Loan Processing Time* exceeds 60 days. The key independent variable is an indicator for the borrower being a minority, identified using the BIFSG method. Column (3) estimates separate effects for Black and Hispanic borrowers, while column (4) uses a minority indicator imputed by InfoUSA. 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)
	Loan Processing Time	I(*Loan Processing Time* > *60 Days*)		
Minority	1.8262***	0.0206***		
	(11.22)	(10.35)		
Black			0.0301***	
			(6.36)	
Hispanic			0.0182***	
			(8.53)	
Asian	1.2103***	0.0155***	0.0153***	
	(4.53)	(4.35)	(4.28)	
Minority (InfoUSA)				0.0065***
				(3.44)
Asian (InfoUSA)				0.0034
				(1.00)
Female	-0.1520*	-0.0001	-0.0001	-0.0001
	(-1.72)	(-0.08)	(-0.10)	(-0.09)
Coborrower	0.3862***	0.0031**	0.0031**	0.0029**
	(4.70)	(2.52)	(2.55)	(2.39)
ln(Income)	-0.8734***	-0.0113***	-0.0113***	-0.0119***
	(-5.98)	(-5.27)	(-5.30)	(-5.55)
ln(Loan Amount)	2.6222***	0.0410***	0.0411***	0.0400***
	(12.39)	(13.40)	(13.44)	(13.15)
LTV	-0.0324***	-0.0007***	-0.0007***	-0.0007***
	(-7.93)	(-12.52)	(-12.52)	(-12.06)
FICO	-1.6336***	-0.0215***	-0.0215***	-0.0222***
	(-15.43)	(-13.52)	(-13.54)	(-13.92)
Age Group FE	Yes	Yes	Yes	Yes
County × Year FE	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes
Dep. Var. Mean	41.516	0.119	0.119	0.119
R-Squared	0.185	0.160	0.160	0.160
Obs.	264,789	264,789	264,789	264,789

Table 3. Robustness Checks for Racial Disparities in Loan Processing Times Using Alternative Samples

This table presents regression results examining the effect of borrower minority status on loan processing times using alternative samples. The analysis uses loan-level observations for loans originated between 2009 and 2021. The dependent variable is I(*Loan Processing Time* > *60 Days*), an indicator equal to 1 if *Loan Processing Time* exceeds 60 days. Columns (1) and (2) use the CoreLogic-only sample, with column (2) focusing exclusively on FHA-insured loans. Columns (3) and (4) use the CoreLogic–HMDA sample, with column (4) excluding applications that potentially involve multiple formal mortgage applications. In columns (1) and (2), the key independent variable is an indicator for the borrower being a minority, identified using the BIFSG method. In columns (3) and (4), the key independent variable is an indicator for the borrower being a minority, as self-reported in HMDA. 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)
	I(*Loan Processing Time* > *60 Days*)			
	CoreLogic		CoreLogic–HMDA	
	Full Sample	FHA	Full Sample	Sample w/ 1 Application
Minority	0.0246***	0.0275***		
	(20.77)	(20.23)		
Minority (HMDA)			0.0189***	0.0192***
			(12.94)	(8.45)
Asian	0.0100***	0.0182***		
	(7.65)	(5.75)		
Asian (HMDA)			0.0028**	0.0017
			(2.07)	(0.93)
ln(Income)	0.0054***	0.0092***	-0.0062***	-0.0058***
	(3.54)	(4.37)	(-7.48)	(-5.82)
ln(Loan Amount)	-0.0227***	-0.0921***	0.0279***	0.0437***
	(-5.13)	(-16.76)	(10.81)	(10.58)
LTV	-0.0001**	-0.0007***	-0.0007***	-0.0007***
	(-2.01)	(-5.42)	(-10.02)	(-7.20)
FHA	-0.0254***		0.0093**	0.0207***
	(-5.93)		(2.07)	(5.14)
Age Group FE	Yes	Yes	Yes	Yes
County × Year FE	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes
Dep. Var. Mean	0.118	0.149	0.093	0.094
R-Squared	0.128	0.114	0.145	0.196
Obs.	3,464,724	990,620	1,030,414	326,989

Table 4. Indirect Test for Lender Discrimination in Racial Disparities in Loan Processing Times

This table presents regression results examining the heterogeneous effect of borrower minority status on loan processing times across cross-sectional and temporal variations. The analysis uses loan-level observations from the CoreLogic–GSE dataset for loans originated between 2009 and 2021. The dependent variable is I(*Loan Processing Time* > *60 Days*), an indicator equal to 1 if *Loan Processing Time* exceeds 60 days. Control variables include dummies for other race categories (*Asian* and *Other Race*), as well as *Female*, *Coborrower*, ln(*Income*), ln(*Loan Amount*), *LTV*, and *FICO*. 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)
	I(*Loan Processing Time* > *60 Days*)			
Minority	0.0154***	0.0111**	0.0214***	0.0303***
	(5.75)	(2.28)	(10.26)	(4.10)
Minority × High Race Animus	0.0119***			
	(3.08)			
High Race Animus	0.0388			
	(0.28)			
Minority × Low Local Competition		0.0050**		
		(2.12)		
Minority × FinTech Lender			-0.0111*	
			(-1.67)	
Minority × Post 2013				-0.0136*
				(-1.80)
Controls	Yes	Yes	Yes	Yes
Age Group FE	Yes	Yes	Yes	Yes
County × Year FE	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes
Dep. Var. Mean	0.119	0.119	0.119	0.119
R-Squared	0.160	0.160	0.160	0.160
Obs.	243,723	264,789	264,789	264,789

Table 5. Impact of Initial Mortgage Processing Delays on Refinancing Behavior

This table presents the OLS and IV regression results examining the effect of processing delays for initial mortgages on quarterly refinancing activity. The analysis uses quarterly loan performance-level observations from the CoreLogic–GSE dataset for loans originated between 2009 and 2021. Column (1) reports the OLS regression results, while columns (2)–(7) present the IV regression results, using *Workload* as an instrument for processing delays. In columns (1)–(3), the dependent variable is *Refinance*, an indicator equal to 1 if the loan was refinanced during the quarter. In columns (4) and (5), the dependent variable is *Refinance (Recapture)*, indicating loans refinanced by the original lender in the quarter. In columns (6) and (7), the dependent variable is *Refinance (Switch)*, indicating loans refinanced by a new lender in 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)	(5)	(6)	(7)
	Refinance			*Refinance (Recapture)*		*Refinance (Switch)*	
	OLS	IV					
I(Loan Processing Time > 60 Days)	-0.1033***	-1.2798***	-1.1803***	-0.9415***	-0.8803***	-0.3384	-0.3001
	(-2.89)	(-2.62)	(-2.59)	(-2.70)	(-2.61)	(-0.84)	(-0.80)
Minority	-0.4248***	-0.4918***	-0.4816***	-0.2496***	-0.2493***	-0.2422***	-0.2323***
	(-8.19)	(-6.65)	(-6.86)	(-6.42)	(-6.14)	(-4.18)	(-4.16)
Asian	0.4485***	0.4452**	0.4748**	-0.1113*	-0.0603	0.5565***	0.5351***
	(3.41)	(2.48)	(2.49)	(-1.69)	(-0.85)	(3.92)	(3.61)
Female	-0.0180	-0.0117	-0.0082	0.0073	0.0093	-0.0191	-0.0175
	(-0.64)	(-0.32)	(-0.22)	(0.33)	(0.43)	(-0.64)	(-0.59)
Coborrower	0.2311***	0.2767***	0.2673***	0.1167***	0.1108***	0.1599***	0.1565***
	(8.21)	(7.57)	(6.92)	(6.44)	(5.65)	(5.02)	(4.90)
ln(Income)	-1.4737*	-1.5695	-1.7553	-1.5885***	-1.7228***	0.0190	-0.0326
	(-1.86)	(-1.27)	(-1.49)	(-2.58)	(-2.66)	(0.02)	(-0.04)
ln(Loan Amount)	-11.7761***	-8.2926***	-7.8540***	-1.5653	-1.6935	-6.7273***	-6.1605***
	(-6.63)	(-3.06)	(-2.98)	(-1.36)	(-1.44)	(-3.22)	(-2.99)
LTV at Origination	-0.2505***	-0.2774***	-0.2721***	-0.1009***	-0.0982***	-0.1765***	-0.1739***
	(-15.44)	(-14.07)	(-13.80)	(-9.21)	(-8.83)	(-11.18)	(-10.71)
Current LTV	0.2356***	0.2600***	0.2642***	0.0874***	0.0878***	0.1727***	0.1764***
	(15.55)	(14.05)	(13.96)	(9.59)	(9.28)	(12.54)	(12.57)
FICO	5.8583***	4.1572***	4.3030***	0.3930	0.7439	3.7642***	3.5591***
	(5.58)	(3.13)	(3.24)	(0.57)	(1.06)	(3.51)	(3.43)
Loan Age	0.3118***	0.6450***	0.6429***	0.2140***	0.2153***	0.4310***	0.4276***
	(8.43)	(13.12)	(13.14)	(10.09)	(10.38)	(12.75)	(12.51)
Rate Gap	1.8022***	1.6612***	1.6566***	0.7142***	0.7163***	0.9469***	0.9403***
	(14.32)	(9.91)	(10.34)	(10.78)	(11.30)	(8.43)	(8.68)
Square Terms of Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County × Origin. Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Avg Loan Processing Time Quintile FE	-	-	Yes	-	Yes	-	Yes
Dep. Var. Mean	2.581	3.333	3.332	1.081	1.090	2.252	2.242
R-Squared	0.056	0.011	0.011	0.004	0.004	0.007	0.007
Obs.	3,952,418	2,294,265	2,193,977	2,294,265	2,193,977	2,294,265	2,193,977

Table 6. 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 processing 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)
	I(*Loan Processing Time > 60 Days*)		*Workload*	
Workload	0.0051***	0.0053***		
	(12.03)	(12.32)		
Minority	0.0219***	0.0190***	0.0800	0.1196
	(7.76)	(6.63)	(0.86)	(1.32)
Asian	0.0084	0.0123**	0.0029	0.0056
	(1.49)	(2.07)	(1.05)	(1.13)
Other Race	0.0425**	0.0318	0.0548	0.0899
	(2.09)	(1.42)	(0.20)	(0.34)
Female	0.0011	-0.0004	0.0138	0.0007
	(0.59)	(-0.17)	(0.37)	(0.02)
Coborrower	0.0017	0.0008	0.0231	0.0241
	(0.90)	(0.41)	(0.69)	(0.72)
ln(Income)	-0.1115*	-0.1526***	0.0361	0.0131
	(-1.89)	(-2.94)	(1.19)	(0.41)
ln(Loan Amount)	-0.0802	-0.0534	0.0669	0.1024
	(-0.61)	(-0.35)	(0.76)	(1.14)
LTV at Origination	0.0015**	0.0016**	0.0047*	0.0042
	(2.15)	(2.16)	(1.74)	(1.54)
FICO	0.0366	0.0493	-0.0340	-0.0594
	(0.53)	(0.69)	(-0.89)	(-1.50)
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	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes
Avg Loan Processing Time Quintile FE	-	Yes	-	Yes
Dep. Var. Mean	0.119	0.119	0.111	0.111
R-Squared	0.168	0.550	0.214	0.537
First-Stage F-Statistics	17.67	17.01	-	-
Obs.	2,294,265	2,193,977	2,294,265	2,193,977

A1. Selection of 18 States

As shown in Table A1, the coverage of the MLS data is very limited in some states. To ensure the reliability and representativeness of our CoreLogic Mortgage–MLS dataset, we restrict our analysis to states where the MLS dataset provides adequate coverage. Specifically, we include only states where the MLS data, after excluding records with missing sale contract dates, covers more than 10% of purchase mortgage records in the deeds data. The selected states are Alabama, Arizona, California, Colorado, Delaware, District of Columbia, Florida, Georgia, Illinois, Maryland, Minnesota, Mississippi, New Jersey, New York, Oregon, Pennsylvania, Rhode Island, and Virginia.

Table A1. CoreLogic MLS Data Coverage Across States

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%

A2. Identifying Mortgage Outcomes with CoreLogic Data

The CoreLogic deeds data does not directly provide loan performance information; however, we can infer it through a systematic approach. To track the performance of mortgages, I have developed a set of steps that leverage the relationships between mortgage records and associated property transactions.

First, for each mortgage record (referred to as the "old mortgage"), I match it with the "next mortgage" that is originated against the same collateral. Once matched, I identify the next mortgage's loan purpose to classify the outcome of the old loan.

If the loan purpose is cash-out refinance, the old loan's outcome is recorded as *prepaid due to cash-out refi*, with the outcome date being the origination date of the new loan. If the loan purpose is rate-reduction refinance, the old loan's outcome is categorized as *prepaid due to rate-reduction refi*, with the outcome date, again, being the origination date of the new loan. If the loan purpose is purchase, the old loan's outcome is identified as *repaid due to selling and moving*, with the outcome date corresponding to the new loan's origination date. To ensure the accuracy of these classifications, I check for potential errors in the CoreLogic records. Specifically, if a loan is refinanced (whether cash-out or rate-reduction), the borrower names on the old and new loans should be the same. Conversely, if a loan is *prepaid due to selling and moving*, the borrower names on the old and new loans should be different.

For old loans classified as *prepaid due to selling and moving*, I further match them with the transaction records of the new loans. If these transaction records indicate a short sale, REO, or foreclosure, I revise the old loan's outcome from *prepaid due to selling and moving* to *default*, because the property transaction is associated with foreclosure. Additionally, I examine each old loan–new loan pair to identify whether an all-cash transaction occurred between the origination dates of the two loans. If the borrower name of the old loan matches the seller name in the all-cash transaction record, I adjust the old loan's outcome and outcome date accordingly.

Finally, for loan observations not matched with a new loan or an all-cash transaction record, I verify whether the loans are still current. To do so, I match the loan with the most recent property information and confirm whether the borrower name on the old loan aligns with the owner name in the latest property record.

This process enables us to infer loan performance outcomes from the CoreLogic data, filling a critical gap in the information provided by the dataset.

A3. CoreLogic–HMDA Match

To link HMDA loans with CoreLogic, we begin by obtaining geocoded census tract information for CoreLogic property addresses. Mortgages are then matched based on census tract, government agency involvement, loan purpose, and rounded mortgage amounts. When applicable, additional criteria such as initial interest rates and loan terms are used to refine the matches.

A4. The Timeline of CFPB's Actions Around 2012–2013

- On October 1, 2012, the CFPB released the examination procedures of Unfair, Deceptive, or Abusive Acts or Practices (UDAAPs) in the CFPB Supervision and Examination Manual (https://www.consumerfinance.gov/compliance/supervision-examinations/unfair-deceptive-or-abusive-acts-or-practices-udaaps-examination-procedures/).
- On December 6, 2012, the CFPB and the Department of Justice (DOJ) signed an agreement facilitating strong coordination between the two agencies on fair lending enforcement, including joint investigations (https://www.justice.gov/opa/pr/justice-department-and-consumer-financial-protection-bureau-pledge-work-together-protect/).
- On December 23, 2013, the CFPB and the DOJ filed the first joint complaint against National City Bank for their discriminatory lending practices (https://www.consumerfinance.gov/about-us/newsroom/cfpb-and-doj-take-action-against-national-city-bank-for-discriminatory-mortgage-pricing/).

A5. Lender-Specific Overlays and Minorities' Loan Processing Delay

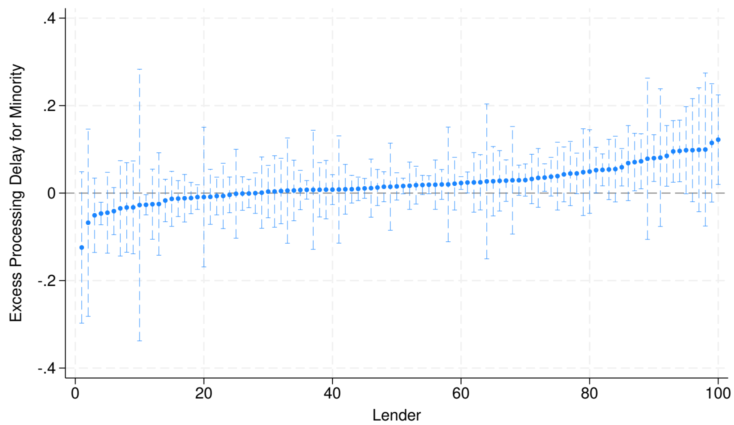
We investigate whether lender-specific "strictness" (i.e., overlays) accounts for the excess loan processing delays experienced by minority borrowers. To show this, we implement a two-step analysis, following Bhutta et al. (forthcoming).

First, we construct a lender-specific measure of strictness for the 100 largest lenders. This measure captures each lender's deviation from the market average loan processing time for White borrowers, controlling for borrower and loan-level characteristics. Specifically, we estimate a regression similar to column (1) of Table 2 using only the White borrower sample and excluding race-related variables. The coefficients on lender fixed effects are extracted as the measure of lender strictness, reflecting how processing times for White borrowers vary across lenders.

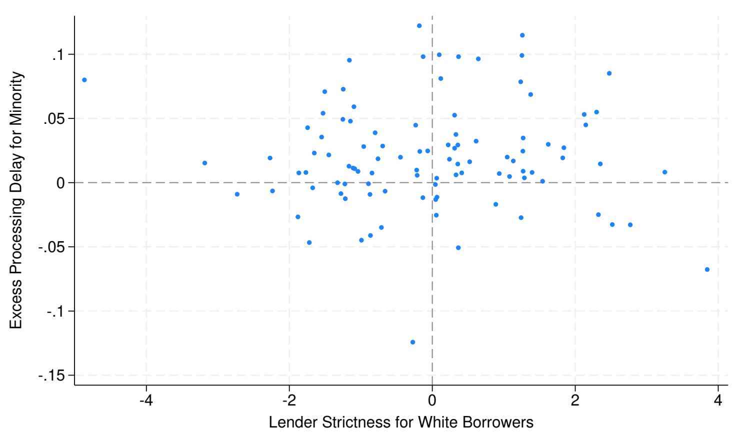
Second, we quantify lender-specific excess delays for minority borrowers by allowing the *Minority* coefficient to vary across lenders in a regression that includes our full set of controls, again similar to column (1) of Table 2. This step provides a lender-level estimate of the additional processing time experienced by minority applicants.

Figure A1 presents a scatterplot comparing the lender-specific strictness measure for White borrowers to the excess delays for minorities. If unobserved borrower risk factors were responsible for the disparities, we would expect a positive correlation between these two measures. However, the scatterplot reveals no such relationship, with the correlation coefficient being extremely small and even negative (-0.0094). This lack of association suggests that unobserved risk factors are unlikely to explain the racial disparities, instead pointing toward lender-side discrimination as a key driver.

Figure A1. Lender-Specific Excess Delay for Minority Borrowers and Lender Strictness for Whites



(a) Lender-Specific Excess Delay for Minorty Borrowers



(b) Lender-Specific Strictness for White Borrowers and Excess Delay for Minorty Borrowers

Table A2. Indirect Evidence of Lender Discrimination in Loan Processing Times Using the CoreLogic-Only Sample

This table presents regression results examining the heterogeneous effect of borrower minority status on loan processing times across cross-sectional and temporal variations using an alternative sample. The analysis uses loan-level observations from the CoreLogic dataset for loans originated between 2009 and 2021. The dependent variable is I(*Loan Processing Time* > *60 Days*), an indicator equal to 1 if *Loan Processing Time* exceeds 60 days. Control variables include dummies for other race categories (*Asian* and *Other Race*), as well as *Female*, *Coborrower*, ln(*Income*), ln(*Loan Amount*), and *LTV*. 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)
	I(Loan Processing Time > 60 Days)			
Minority	0.0165***	0.0181***	0.0259***	0.0546***
	(4.67)	(9.49)	(10.61)	(10.92)
Minority × High Race Animus	0.0138***			
	(2.64)			
High Race Animus	-0.0378			
	(-1.29)			
Minority × Low Local Competition		0.0247***		
		(5.02)		
Minority × FinTech Lender			-0.0108***	
			(-5.03)	
Minority × Post 2013				-0.0349***
				(-7.31)
Controls	Yes	Yes	Yes	Yes
Age Group FE	Yes	Yes	Yes	Yes
County × Year FE	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes
Dep. Var. Mean	0.119	0.119	0.119	0.119
R-Squared	0.134	0.132	0.132	0.132
Obs.	3,361,559	3,534,651	3,534,651	3,534,651

Table A3. Impact of Initial Mortgage Processing Delays on Cash-Out Refinance and Moving

This table presents IV regression results examining the effect of processing delays for initial mortgages on quarterly cash-out refinancing and moving, using *Workload* as an instrument. The analysis uses quarterly loan performance-level observations from the CoreLogic–GSE dataset for loans originated between 2009 and 2021. In columns (1) and (2), the dependent variable is *Cash-Out Refinance*, indicating loans cash-out refinanced during the quarter. Columns (3) and (4) focus on *Cash-Out Refinance (Recapture)*, while columns (5) and (6) focus on *Cash-Out Refinance (Switch)*. In columns (7) and (8), the dependent variable is *Prepaid Due to Move*, indicating loans prepaid due to selling and moving 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)	(5)	(6)	(7)	(8)
	Cash-Out Refinance		*Cash-Out Refinance (Recapture)*		*Cash-Out Refinance (Switch)*		*Prepaid Due to Move*	
I(Loan Processing Time > 60 Days)	-1.0840***	-0.9546**	-0.6866***	-0.6574***	-0.3974	-0.2973	-0.3887	-0.3980
	(-2.58)	(-2.42)	(-3.29)	(-3.36)	(-1.17)	(-0.92)	(-1.08)	(-1.14)
Minority	-0.1841***	-0.2141***	-0.0586***	-0.0668***	-0.1255***	-0.1473***	-0.5299***	-0.5421***
	(-3.08)	(-3.53)	(-2.66)	(-2.91)	(-2.72)	(-3.13)	(-13.14)	(-13.85)
Asian	-0.6219***	-0.6032***	-0.1948***	-0.1841***	-0.4271***	-0.4191***	-0.3220***	-0.3144***
	(-9.15)	(-9.32)	(-6.43)	(-5.96)	(-7.30)	(-7.47)	(-5.45)	(-5.43)
Female	-0.0481	-0.0520*	0.0026	0.0017	-0.0507**	-0.0537**	0.0325	0.0426*
	(-1.63)	(-1.66)	(0.16)	(0.11)	(-2.18)	(-2.18)	(1.24)	(1.66)
Coborrower	0.0323	0.0376	0.0281*	0.0286*	0.0041	0.0090	-0.1237***	-0.1176***
	(1.22)	(1.33)	(1.94)	(1.93)	(0.18)	(0.38)	(-4.97)	(-4.67)
ln(Income)	2.1077***	1.7677**	0.5679	0.3974	1.5398***	1.3703**	-2.2625***	-2.0851***
	(2.81)	(2.39)	(1.60)	(1.08)	(2.60)	(2.35)	(-3.10)	(-2.84)
ln(Loan Amount)	9.8809***	10.4614***	2.0362***	2.2599***	7.8447***	8.2015***	11.0075***	11.1775***
	(7.35)	(7.72)	(3.41)	(3.70)	(7.55)	(7.86)	(7.27)	(7.22)
LTV at Origination	0.0695***	0.0750***	0.0110**	0.0140**	0.0585***	0.0610***	0.0134	0.0151
	(5.24)	(5.50)	(2.03)	(2.51)	(5.22)	(5.38)	(0.75)	(0.86)
Current LTV	-0.0155	-0.0192	0.0022	0.0014	-0.0177*	-0.0206**	0.0909***	0.0927***
	(-1.24)	(-1.54)	(0.49)	(0.33)	(-1.75)	(-1.99)	(5.88)	(5.91)
FICO	0.0357***	0.0359***	0.0063	0.0068	0.0293***	0.0291***	0.0525***	0.0558***
	(4.20)	(4.32)	(1.25)	(1.30)	(3.56)	(3.66)	(6.59)	(6.75)
Loan Age	0.1481***	0.1466***	0.0474***	0.0454***	0.1007***	0.1011***	0.1491***	0.1463***
	(7.61)	(7.49)	(5.59)	(5.57)	(6.54)	(6.31)	(8.36)	(7.83)
Rate Gap	0.8195***	0.8140***	0.3386***	0.3411***	0.4809***	0.4730***	0.0593	0.0503
	(13.53)	(13.25)	(11.82)	(11.68)	(11.17)	(10.78)	(1.36)	(1.16)
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	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Avg Loan Processing Time Quintile FE	-	Yes	-	Yes	-	Yes	-	Yes
Dep. Var. Mean	1.446	1.453	0.421	0.422	1.025	1.031	1.762	1.779
R-Squared	0.003	0.003	0.000	0.000	0.002	0.002	0.002	0.002
Obs.	2,294,265	2,193,977	2,294,265	2,193,977	2,294,265	2,193,977	2,294,265	2,193,977