Does the Disclosure of Consumer Complaints Reduce Racial Disparities in the Mortgage Lending Market?*

Xiang Li†

November 18, 2023

Abstract

The Consumer Financial Protection Bureau (CFPB) publicly disclosed consumer complaint narratives in 2015. Utilizing a difference-in-differences design, I find that, following disclosure, CFPB-supervised banks whose complaint narratives are disclosed are less prone to discriminate against minority borrowers in the mortgage lending market. This reduces racial disparities in interest rates, default rates, and rejection rates. The disclosure saves minority borrowers $102 million in interest payments and aids over 14,000 minority households in securing loans annually, thereby narrowing the racial gap in homeownership. Stakeholders including consumers, peer banks, and stock market investors facilitate the disclosure's effects on reducing discrimination.

*I am indebted to my advisor, Thomas Chemmanur, for his numerous guidance, support, and encouragement during my dissertation research. I would like to express my sincere gratitude to my other dissertation committee members, Philip Strahan, Hassan Tehranian, and Rawley Heimer, for their guidance and support. I want to sincerely thank Ningzhe Zhou for providing excellent research assistance. I am grateful for the valuable feedback and insights gained from participants at the following conferences and seminars: AEA, FARS, Inter-Finance PhD seminar, AFBC, New Zealand Finance Meeting, FMA New Ideas Session, Haskell & White Academic Conference, Boca Corporate Finance and Governance Conference, AFA (Poster Session), Hawaii Accounting Research Conference, and AFFECT New Ideas sessions.

†Finance Department, Carroll School of Management, Boston College, Chestnut Hill, MA 02467, Email: xiang.li.5@bc.edu.

1 Introduction

This study finds that mandatory disclosure of consumer complaints against banks by the Consumer Financial Protection Bureau (CFPB) alleviates racial discrepancies in the mortgage lending market. The CFPB plays a crucial role in advocating for consumer rights and fostering fairness in the financial markets. Since its inception in 2011, it has been receiving complaints against financial institutions. On June 25, 2015, it took a further step by publicly disclosing complaint narratives and financial institutions' responses (CFPB (2015)). This transparency initiative aims to promote a comprehensive analysis of complaint data, facilitating a productive discourse between consumers and financial institutions. Despite the implementation of the disclosure policy, little empirical evidence exists supporting its effectiveness in protecting minority consumers or mitigating racial bias in lending markets. This study aims to address these critical questions: Does the disclosure of complaint narratives against banks lead to a reduction in discrimination against minority borrowers? What are the underlying mechanisms propelling this change? Which stakeholders play instrumental roles in influencing the effects of such disclosures?

However, the answers to the aforementioned questions remain uncertain. On one side, the release of complaint narratives by the CFPB may motivate banks to reduce discriminatory treatment. This is driven by stringent government oversight, reputation concerns, and the fear of significant litigation risks.[1] Moreover, the CFPB's disclosure enables minority consumers to compare loan options thoroughly. This helps them to avoid banks with a history of discriminatory complaints. Additionally, certain banks may have a significant number of discriminatory complaints in specific areas. Their competitors, aware of the disclosed narratives by the CFPB, may be drawn to enter that underserved market for a slice of the pie (Dou, Hung, She, and Wang (2023)). Lastly, the capital markets may use these narratives to avoid investing in banks with a high volume of discrimination complaints because such banks may be exposed to both ethical and litigation risks (Pan, Pikulina, Siegel, and Wang (2022)).

[1] Between 2012 and 2022, the CFPB imposed an approximate total of $66 million in fines on 21 banks for illegal discrimination against applicants, predominantly minorities. Moreover, these banks were mandated to contribute over $51 million towards loan subsidy programs. These programs intended to lessen the loan burden of minority groups, may comprise measures such as reductions in interest rates, assistance with closing costs, and support for down payments.

On the other side, consumers' limited search in the mortgage market (Woodward and Hall (2012)) may obstruct them from incorporating this data into their decision-making. Furthermore, minority consumers may have limited choices if only a few banks dominate the local financial market (Li (2023); Stanton, Walden, and Wallace (2014)). The disclosure activities may raise administrative workload and consumer privacy concerns, which may lead to resistance. Therefore, the influence of disclosure on racial disparities in the lending market remains a subject for empirical examination.

The CFPB discloses consumer complaint narratives covering a variety of financial products, such as credit cards, mortgages, auto loans, and student loans. My research primarily concentrates on the mortgage market for two key reasons. First, discrimination against minority consumers in the mortgage market has been consistently observed in both real court cases and academic studies. Despite the enactment of anti-discrimination laws like the Fair Housing Act of 1968 in the United States, discriminatory practices persist in the housing and finance sectors. The U.S. Department of Justice notably brought allegations against several major mortgage lenders for violating fair lending principles during the housing boom. The settlements exceeded $500 million.[2] Moreover, academic scholars consistently find evidence of unequal loan pricing, indicating the enduring existence of discriminatory behavior in the mortgage sector (Ambrose, Conklin, and Lopez (2021); Bartlett, Morse, Stanton, and Wallace (2022); Cheng, Lin, and Liu (2015); Ghent, Hernandez-Murillo, and Owyang (2014); Woodward and Hall (2012)). In particular, Bartlett et al. (2022) pinpoint discrepancies in interest rates between minority and White borrowers. These disparities cost minority borrowers an additional $450 million annually.

Second, the mortgage market is the most significant but opaque consumer financial market. As of 2023, mortgages account for 70.6% of consumer debt in the United States, and the total amount reaches $11.92 trillion. Approximately 83.4 million homeowners bear a mortgage loan.[3] However, minority consumers frequently face challenges due to insufficient knowledge about the quality and details of mortgage products and services. Gaining knowledge through experience is difficult for them since decisions involving choosing a mortgage are infrequent (Campbell, Jackson, Madrian, and Tufano (2011); Lusardi and

[2]Regarding the three most substantial settlements, Bank of America reached a settlement of $335 million in 2011, followed by Wells Fargo's settlement of $175 million in 2012, and JPMorgan Chase's settlement of $55 million in 2017.

[3]For more details, see: https://www.lendingtree.com/home/mortgage/u-s-mortgage-market-statistics/

Mitchell (2011); Murphy (2005)). Moreover, people often avoid discussing personal finance due to societal conventions, and financial advisors can sometimes provide biased advice to serve their own interests (Guiso, Pozzi, Tsoy, Gambacorta, and Mistrulli (2022)). Even with pertinent information, minority consumers may struggle to understand it compared to white consumers. This could be due to processing biases, lack of attention, and financial illiteracy (Gurun, Matvos, and Seru (2016); Keys, Pope, and Pope (2016)). Given these two reasons, it is important to investigate whether disclosure can reduce racial disparities in the mortgage lending market and to understand the underlying mechanisms.

Empirically studying this question involves addressing two crucial challenges. The first challenge is to accurately identify discrimination while addressing omitted variable bias. To deal with the first issue, I follow Bartlett et al. (2022) and use information from the government-sponsored enterprises (GSEs) to construct an 8×9 matrix known as the Loan-Level Price Adjustments (LLPAs). This matrix is an essential tool for adjusting credit-risk pricing based on credit score and Loan-to-Value (LTV) ratios. By including this grid, I can identify any differences in interest rates within it. These differences imply potential discriminatory practices, as they cannot be solely attributed to varying credit risks.

The second challenge is to identify exogenous variation in disclosure at the lender level. Regarding the source of exogenous variation, I utilize a disclosure policy implemented by the CFPB. Since June 2015, the public can access consumer complaint narratives. But this is limited to banks with total assets exceeding $10 billion. This policy shift serves as an exogenous shock, enabling the observation of varied bank behaviors post-disclosure. Taking into account the potential impact of bank size on the outcomes of interest, my sample primarily focuses on banks with total assets under $100 billion as of the first quarter of 2015. Using this sample, I employ a difference in differences method to study shifts in lending behavior and service outcomes for minority borrowers in the mortgage market. I focus on banks close to the asset threshold during both the pre and post-June 2015 periods.

Additionally, in my model, I consider lender $\times$ similar borrower fixed effects and lender $\times$ year and month fixed effects. By controlling for lender $\times$ similar borrower fixed effects, I can compare borrowers with similar characteristics at the same bank before and after the disclosure shock. Meanwhile, controlling for bank $\times$ year and month fixed effects absorbs any unobservable time-varying shocks at the bank level that could potentially influence lending practices and service outcomes.

I use the pricing grid from LLPAs within the HMDA-GSE merged dataset to test the effect of disclosure on racial disparities in interest rates. I find that following the disclosure of complaint narratives, the excess interest rates applied to minority borrowers compared with white borrowers decline by around 2.55 basis points. This decrease mitigates nearly 87% of the racial disparities in interest rates, which originally stand at 2.7 basis points. These findings suggest that the enactment of the disclosure policy effectively saves minority borrowers an average of $102 million in interest rates each year in the mortgage market. Moreover, the reduction in racial disparities becomes noticeable immediately after the disclosure and exhibits a persistent influence. Further analysis reveals that the disclosure has a more pronounced effect on mitigating racial disparities among banks that engage in discriminatory practices. The disciplinary impacts of disclosure intensify when lenders' discriminatory treatments receive increased social attention or the litigation risks associated with discrimination are high.

In addition, I use the "outcome test" approach (Becker, 1957, 1993) to isolate racial bias from omitted variable bias or statistical discrimination. I investigate the effects of disclosure on the default rate of minority borrowers.[4] If racial bias exists, marginal minority borrowers may exhibit a lower default rate compared to marginal White borrowers because minorities are held to unfairly high standards by banks. My findings substantiate this hypothesis. Additionally, they further reveal that compared to consumers from control groups, the unfair lower default rate among minority consumers in treated groups almost disappears following disclosure. The racial discrepancies I observe through the outcome test originate from banks imposing stricter approval criteria on consumers. The findings from analyzing the rejection rate in the HMDA origination dataset are consistent with the results obtained from the default rate analysis. Minority borrowers are 7.7 percentage points more likely to face loan rejections controlling for creditworthiness, implying that banks use higher standards to screen their loan applications. After disclosure, the initial 7.7 percentage points higher rejection rates for minorities decrease by 1.6 percentage points. This disclosure policy lessens the unjust accessibility standards imposed upon minorities and aids over 14,000

[4] Lenders engage in statistical discrimination when they utilize race as a substitute for unobservable creditworthiness aspects to maximize profits. In this study, I use the term "statistical discrimination" to refer to decisions that maximize profits. Conversely, I employ the term "racial bias" to denote biased lending decisions. This bias could stem from particular preferences or incorrect beliefs. A more detailed description is provided in Section 5.

minority households in obtaining mortgage loans annually.

I analyze the responses of different stakeholders to gain more insights into additional mechanisms that contribute to the effectiveness of the disclosure policy. First, I utilize an empirical strategy consistent with the baseline analysis to study consumers' reactions toward the disclosure. The results indicate that minority consumers learn from complaint narratives. They actively avoid applying for loans from banks with discrimination-related complaints. I also find that peer banks use insights from their competitors' disclosed weaknesses to make decisions about market entry. This is particularly evident in areas where customer service is deficient and discriminatory treatments are prevalent. Simultaneously, the stock market may interpret the disclosure policy as an adverse shock that provokes significant reactions. Considering potential litigation risks, investors may view the future performance of banks engaged in discriminatory practices with skepticism. Additionally, in line with the emphasis on socially responsible investing (SRI), investors may be reluctant to invest in firms with discriminatory practices. Through an event study method, I evaluate the market reaction subsequent to the disclosure policy. I find that banks supervised by the CFPB exhibit lower market reactions. This is characterized by a relative decrease in cumulative abnormal returns (CAR), with a more severe decrease in firm value among banks engaged in discriminatory practices.

The persistently low homeownership rate among minority residents is an issue of concern. My research finds that disclosure improves mortgage loan application rates for minority borrowers. Employing county-year level data on homeownership by race, I examine the real effect of the disclosure policy. I test whether improved credit access will increase homeownership for minority borrowers. These results imply that disclosure increases the homeownership rate among minority households, with little impact on white households. Thereby it significantly reduces racial gaps in homeownership.

Finally, I survey the status of financial consumer protection organizations in the main developed countries. Apart from the CFPB in the United States, nearly no financial regulatory agencies in other developed countries choose to disclose consumer complaint information publicly. My research shows that public exposure to customer complaints can mitigate discriminatory and unfair practices in the financial system through multiple avenues. From this perspective, these findings may provide practical insights for real-world financial practices. Other countries can look to this disclosure policy as a reference when crafting their strategies for handling and publicizing consumer complaints. The adoption of such disclosure policies may have considerable promise in reducing racial disparities at a global level.

Notably, to further aid minority borrowers in the United States in mitigating potential discrimination in mortgage lending, I develop an online inquiry system [5]. This permits the public to directly and transparently access information regarding local banks' discriminatory practices.

The rest of the paper is as follows. Section 2 describes the related background, existing research, and contributions of my research. Section 3 introduces the datasets used and the results of summary statistics. Section 4 reports the main analysis of the paper, the impact of disclosure on racial disparities in loan interest rates. Section 5 reports the outcome test for default rates. Section 6 reports the impact on rejection rates. Section 7 presents the reactions of various stakeholders to the disclosure policy. Section 8 examines the real effects of the disclosure policy on resident homeownership. Section 9 introduces the website ranking banks for discrimination at the zip code level. Section 10 concludes.

2 Background and Literature Review

2.1 Background of Complaints Disclosure

The Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 established the CFPB to safeguard consumers and promote fairness and competition in consumer financial markets. The CFPB began accepting consumer complaints regarding credit cards in July 2011. Then it expanded its scope to complaints about mortgages, bank accounts, credit reporting, and other financial products and services. After a pilot phase in 2012 and 2013, the CFPB gradually began to release information related to consumer complaints to the public.[6] On June 25, 2015, the CFPB publicly released the narratives of consumer complaints. The complaint database exclusively records complaints against banks supervised

[5]For more details, see: https://discriminationwarrior.com

[6]On March 28, 2013, the CFPB disclosed complaints about mortgages and other products. Since then, the public has been able to download information related to the financial product type involved in the complaint, the submitting consumer's ZIP Code, the submission date, the name of the implicated bank, and the bank's brief response.

by the CFPB with total assets surpassing \$10 billion. Complaints concerning depository institutions holding less than \$10 billion in assets are not included in the database. They are directed to the relevant safety and soundness regulators.

In fact, the CFPB solicited comments on the issue before deciding to make complaint narratives public. They received 137 distinct responses from consumer groups, trade organizations, businesses, and individuals, along with 30,000 identical letters from individuals. Banks and their associations object to the disclosure of these narratives. They argue that releasing "unverified" consumer narratives could unfairly damage companies' reputations. The CFPB took steps to address these objections. Firstly, the CFPB will confirm business relationships between the banks and complainants to prevent fraud. Secondly, the CFPB will make banks' responses to complaints public, allowing banks to contest the complaints.

In contrast, consumer advocates, civil rights groups, and proponents of open governance supported the inclusion of narratives. They highlighted three key benefits: (i) empowering consumers with timely and relevant information to make informed decisions before purchase to promote customer-oriented businesses and uncovering unfair or deceptive practices post-purchase; (ii) aiding the Bureau and other stakeholders in identifying harmful trends before they escalate into widespread damage; and (iii) encouraging greater use of the database as a practical tool to promote accumulating data on consumer experiences in the financial sector. These benefits align with the CFPB's original intentions for this measure (CFPB (2015)). Nevertheless, supporters also expressed privacy concerns for consumers who provide complaint narratives. To address these concerns, the CFPB committed to anonymizing identifiable information and disclosing only limited zip code data.

This study focuses on complaints related to discrimination. Textual analysis of complaint narratives reveals a relatively consistent annual reporting of 700-900 discrimination-related complaints in the mortgage market. This result is depicted in panel A of Figure 1. Additionally, panel B of Figure 1 demonstrates that 5-7% of mortgage market complaint narratives are related to discrimination each year. This stable proportion highlights the persistent issue of discrimination. Given the smaller fraction of minority consumers in the lending market,[7] the 5-7% representation is particularly significant. This emphasizes the

[7] In the HMDA dataset spanning from 2011 to 2019, used in this study, loan applications from minority consumers make up 12.24% of total applications. However, this proportion diminishes to 11.55% when considering loans that receive approval.

severity of discrimination problems.

[Insert Figure 1 about here]

Table 1 presents a consumer complaint example concerning "Applying for a mortgage or refinancing an existing mortgage," from CFPB's public dissemination. The components include "Date received", "Product", "Consumer complaint narrative", "Company", and "Company response to consumer", and others. Although personal data are de-identified in complaint narratives, claims of racial discrimination are evident, especially through statements like "I believe that I am being discriminated against because I disclosed my race as XXXX".[8] The company's reply to this particular situation is categorized as "closed with monetary relief," and the company does not dispute the complaint. These indicators suggest potential misconduct by the bank towards this consumer. Undoubtedly the aggrieved party received compensation for the unfair treatment with the presence of the CFPB's complaint system.

[Insert Table 1 about here]

2.2 Literature Review and Contribution

My research contributes to the extant literature on the effectiveness of the CFPB. Supporters highlight its success in curbing deceptive and predatory practices within credit markets. Fuster, Plosser, and Vickery (2021) emphasize its role in reducing lending and potentially preventing foreclosures. DeFusco, Johnson, and Mondragon (2020) find CFPB's Ability-to-Repay and Qualified Mortgage Rule restrains high-leverage mortgages and mitigates instability within the financial system.[9] In contrast, critics argue that the CFPB's oversight increases regulation and compliance costs and heightens legal liabilities. This could reduce the supply of consumer credit (U.S. Chamber of Commerce (2018); Neugebauer and Williams (2015); Fuster et al. (2021); DeFusco et al. (2020)). Unlike these research perspectives, my paper adopts a diversity, equity, and inclusion (DEI) lens to investigate the CFPB's contribution to addressing discrimination in the financial market. This substantiates its positive influence on advancing DEI.

[8] The race information is erased by the CFPB to protect consumer privacy.

[9] For more details, please refer to prepared remarks of CFPB director Richard Cordray at the LendIt USA Conference. This is the related link: https://www.consumerfinance.gov/about-us/newsroom/prepared-remarks-cfpb-director-richard-cordray-lendit-usa-conference/.

Among the various oversight strategies implemented by the CFPB, the distinctive and vital practice of consumer complaint disclosure has received limited research attention.[10] Previous studies like Dou and Roh (2023) evaluate the impact of disclosing total complaint counts on mortgage applications. They reveal a notable reduction in applications submitted to banks with a history of complaints after the disclosure. Dou et al. (2023) examine how unregulated competitor banks react to complaint disclosures compared to CFPB-regulated institutions. They observe an increase in mortgage approvals in areas with a high number of complaints. This indicates that banks effectively learn from the operational shortcomings of their counterparts. My study however focuses on the disclosure of complaint narratives instead of total complaint counts. These narrative disclosures offer more comprehensive information than mere complaint numbers. Therefore, it is possible to better understand whether banks' misconduct in the mortgage market is alleviated through the disclosure process.

This study is the first to examine the consequences of disclosure on DEI within the financial market. It aligns with the effects of disclosure on DEI observed in other markets. For example, Bennedsen, Simintzi, Tsoutsoura, and Wolfenzon (2022) draw from a 2006 Danish law that compels firms to disclose gender-specific wage data. They reveal its impact on the gender wage gap and firm performance. Pan et al. (2022) analyze market reactions to the initial disclosure of CEO-worker pay ratios by U.S. public firms in 2018. They observe unfavorable responses from firms with higher reported ratios. This is particularly noticeable among shareholders who are sensitive to inequality. In addition to analyzing the distinct mortgage market, my study also expands the potential scope of disclosure policies by considering them from the standpoint of racial inequalities.

Racial discrimination is a significant barrier to achieving DEI. My study contributes to the existing literature on discrimination in the financial market. It demonstrates that disclosure alleviates such discriminatory practices. Many studies emphasize the unequal treatment of minority consumers in the financial sector. However, pinpointing instances of discrimination remains a complex challenge for researchers. In previous research, Bartlett et al. (2022) find that Latinx/Black borrowers face higher risk-adjusted interest rates in

[10] I catalog 27 developed nations worldwide, of which 13 have financial regulatory agencies that establish dedicated channels for receiving consumer feedback. However, only the United States, through the CFPB, makes consumer complaints publicly available.

GSE-securitized and FHA-insured loans, accounting for borrowers' creditworthiness. This unjust treatment results in an annual financial burden exceeding $450 million. The outcome tests (Becker (1957)) suggest that marginal approved minority borrowers tend to have lower default rates. Based on this outcome test, Butler, Mayer, and Weston (2023) reveal that although minority borrowers face an extra 70 basis points in interest rates compared to white counterparts, they have lower default rates in the auto loan market. Moreover, Li (2023) introduces a textual analysis methodology to identify instances of discrimination within financial markets.[11] In this study, I demonstrate how disclosure reduces discrimination by using these three distinct methodologies to quantify discrimination. First, following Li (2023), I apply textual analysis on mortgage market complaints under CFPB supervision to reveal cases of racial discrimination. Second, I build on Bartlett et al. (2022)'s method to identify discrimination through the GSE pricing grid. My research indicates that the narrative disclosure of consumer complaints in the mortgage market results in an estimated annual saving of $100 million annually. Finally, I examine whether the disclosure of discriminatory practices benefits marginal minority borrowers by rectifying unjust approval standards. My findings affirm this proposition, in line with Butler et al. (2023)'s outcome test.

Previous research often addresses discrimination through the lens of market power. However, my study offers a distinct approach by examining it from the perspective of regulatory policy. Black and Strahan (2001) investigate the impact of bank competition on gender pay gaps among bank officers. They discover a decrease in male-biased wages and an increase in female managerial positions within a more competitive financial market. Similarly, Li (2023) explores how increased bank competition mitigates discrimination against minority borrowers. In contrast, my study reveals that disclosure serves as an effective regulatory instrument to alleviate discrimination. This supplements the current research by illustrating how disclosure reduces discrimination by promoting intensified competition within the lending sector.

[11] Currently, the adoption of textual analysis for scrutinizing discriminatory practices is widely used. As articulated in Hacamo (2022), over 10,000 instances of workplace racial bias are identified through textual analysis of 7 million job reviews on Indeed.com.

3 Data Description and Summary Statistics

3.1 Data Description

3.1.1 HMDA-GSE Merged Dataset

The main dataset employed in this study is the integrated HMDA-GSE dataset. The Home Mortgage Disclosure Act (HMDA) compliance surveys are the sole data source of loan-level information regarding the applicant's race and ethnicity. This dataset includes 90% of mortgage originations in the U.S. as reported by Engel and McCoy (2016). The HMDA data contain essential details such as the applicant's income, race, ethnicity, loan amount, lender name, and the census tract of the property. I use the Government-Sponsored Enterprises (GSE) dataset to complement the above dataset. It offers detailed loan information such as interest rate, default, LTV, credit score, loan product, loan purpose, and loan term. These factors are crucial for understanding the personal characteristics of borrowers.

Since there is no direct connection between these two datasets, I employ the "fuzzy data matching" methods in accordance with Law and Mislang (2022) to merge them. The detailed process is explained in Appendix B. This combined dataset comprises all approved loans securitized by the GSEs from 2011 to 2019.[12] Additionally, the handling of loans with diverse terms differs within the real market. Therefore, I apply filtering to this dataset following Bartlett et al. (2022). This filtering is based on variables like credit scores, LTV, and loan amount. Detailed specifics are provided in Appendix B. The outcome is a pooled cross-sectional dataset of loan-bank-years. I use this dataset to examine whether disclosing consumer complaint narratives affects interest rates and default rates among minority borrowers.

3.1.2 HMDA Origination Dataset

Another important dataset used in this study is the HMDA origination dataset. Unlike the HMDA-GSE merged dataset, the original HMDA dataset includes unapproved loans.

[12] Given the profound impact of the COVID-19 pandemic, which has ravaged the globe, including the United States, since 2020, my research excludes samples after 2019 to mitigate the interference from this shock on the financial markets.

This allows me to assess the tangible effects of disclosing consumer complaint narratives on the rejection rates of minority borrowers. However, the HMDA dataset lacks enough loan attributes. So I adopt a strategy from Bartlett et al. (2022) to mitigate potential bias arising from omitted variables. As the borrower's credit score or LTV cannot be observed directly, I employ a substitution approach in my analysis. I use the HMDA-GSE merged dataset to calculate the median credit score and LTV at the census-tract level. A comprehensive explanation of this process is in Appendix B.

3.1.3 CFPB Dataset

I apply textual analysis methods to CFPB's consumer complaint database to identify complaints related to racial discrimination. Further, I pair the annual number of discrimination-related complaints with both the HMDA-GSE merged dataset and the HMDA origination dataset. This merging process is based on ZIP Codes and bank names from the CFPB's database. [13] In the merging process, most complaints are associated with a singular county. If a ZIP Code encompasses multiple counties, following Dou and Roh (2023), I match it to the county with the highest population.

3.1.4 Additional Datasets

My research also incorporates several supplementary datasets and indicators from varied sources to identify the impacts of disclosure.

(i) I obtain stock market returns for publicly listed banks from Wharton Research Data Services (WRDS). This dataset is around June 25, 2015, both before and after the implementation of the disclosure policy. It includes market-adjusted cumulative abnormal returns (CAR) and CAR based on the CAPM. WRDS also offers financial performance data for these banks for the fourth quarter of 2014. I use financial performance as control variables in analyzing market responses.

(ii) I use the Call Reports dataset to get the total assets of each bank in the primary dataset for the first quarter of 2015. This indicator identifies banks influenced by disclosure. Additionally, I consider banks with total assets ranging from 0 to 10 billion as the control

[13] Due to privacy concerns, the narratives of complaints before June 25, 2015, are not made public. Therefore, the discrimination-related complaints identified in this study pertain to the period after the implementation of the disclosure policy.

group and banks in the 10-100 billion range as the treated group.[14] This filter facilitates the estimation of the causal effects of disclosure by ensuring comparability among banks of comparable size. The results are robust to different cutoffs.

(iii) I employ variables such as branch-level deposits from the Summary of Deposits (SOD) dataset to supplement the HMDA origination dataset. I use this dataset to examine whether disclosing discrimination complaints against local banks affects application numbers in local markets at the county level.

(iv) The American Community Survey (ACS) provides household-level characteristics aggregated at the county level. I select the homeownership-related variables from 2011 to 2019. Homeownership is measured as the percentage of housing units occupied by owners, considering various races. I employ this measurement to examine the real effect of disclosure policy on the racial gaps in homeownership.

3.2 Summary Statistics

Table 2 provides detailed summary statistics for the treated and control banks from the primary HMDA-GSE merged dataset. This dataset ranges from 2011 to May 2015, before the enactment of the CFPB's disclosure policy. The treated group includes banks with total assets ranging from 10 to 100 billion dollars, while control banks manage assets below 10 billion. I use this dataset to assess the influence of disclosing consumer complaint narratives on both loan interest rates and default rates.

In treated banks, 7% of loans come from minority consumers, while in the control group, it is 8%. Minority consumers here refer to Black or Hispanic individuals. [15] In terms of interest rates, treated banks exhibit an average rate of 4.11%, slightly lower than the 4.22% in control banks. A comparable pattern is evident in default rates. The mean values for LTV ratios and credit scores are also similar between treated and control banks. Loan

[14]The reasons for dropping big banks are as follows: following the implementation of the Dodd-Frank Act, bank holding companies (BHCs) with assets exceeding $100 billion have become the target of heightened supervision and are subjected to regulatory stress tests under the Comprehensive Capital Analysis and Review (CCAR) framework in 2011. Furthermore, larger banks, compared to smaller banks, may have distinct capital structures or exhibit economies of scale.

[15]As for other variables utilized in this study, their definitions can be found in Table A2. Particularly, the final three variables featured in Table 2: Cash-out Refinance, Purchase, and Refinance, represent three distinct categories of loan purpose. To provide meaningful values, individual explanations are given for each category.

borrowers in the treated group have an average annual income of $107,000. In the control group, the average income is $100,000. Additionally, the average loan amount for these groups is $239,000 and $247,000, respectively. This table presents similar characteristics across critical variables for the treated and control groups. This similarity in risk attributes before the shock indicates it is reasonable to categorize treated and control groups in this study.

[Insert Table 2 about here]

4 CFPB Disclosure and Racial Gaps in Interest Rates

4.1 Research Design

This paper examines how banks respond to the disclosure of complaint narratives during loan application processes. I account for three critical dimensions: racial disparities in interest rates, default rates, and loan rejections. In this section, I specifically focus on the influence of disclosure on racial differences in interest rates for two primary reasons. Firstly, loan interest rates are important for borrowers. Racial disparities in this area cost minority borrowers up to $450 million annually, as indicated by Bartlett et al. (2022). The presence of this unfair treatment also draws significant attention from regulatory bodies.[16] Secondly, the distinct loan pricing strategy in defaulted guaranteed GSE loans can effectively capture discriminatory practices within banks.

According to U.S. fair-lending legislation, judicial rulings have affirmed that lenders are allowed to consider specific borrower characteristics when determining loan prices.[17] This remains legitimate even if these proxy variables result in less favorable outcomes for minorities as long as the lender can demonstrate that these variables serve a legitimate business necessity. Courts have explicitly determined that when a lending practice leads to a disparate impact, the responsibility falls on the defendant-lender. Lenders need to prove that

[16] Senators Elizabeth Warren and Doug Jones express concerns to financial regulators about persistent lender discrimination. This assertion references the recent study by Bartlett et al. (2022), which identifies disparities in interest rates between minority and white borrowers. Their communication can be accessed using this link: https://www.warren.senate.gov/download/2019610-letter-to-regulators-on-fintech-final.

[17] U.S. fair-lending law encompasses both the Fair Housing Act and the Equal Credit Opportunity Act (ECOA), along with all related regulatory implementations and judicial interpretations about these acts.

any policy, procedure, or practice is associated with the applicant's creditworthiness.[18] In other words, variations in creditworthiness can justify differing lending outcomes. However, it is unjustifiable to charge higher rates to applicants in financially underserved areas or those with limited shopping tendencies with the aim of increasing profits.

I adopt a setting following Bartlett et al. (2022) to identify discrimination and address concerns about omitted variables. Using this method, all *legitimate business necessity* variables are observable. Within this context, the GSEs establish credit-risk pricing adjustments through a fee structure that relies solely on the borrower's position within an 8×9 matrix of LTV ratios and credit scores, known as loan-level price adjustments (LLPAs). Lenders are insured against credit risks by paying these LLPA fees. Figure 2 presents a typical Fannie Mae LLPA grid for single-family loans with a 360-month term. Freddie Mac, another primary entity, also employs a similar pricing model in the GSE loan sphere. Consequently, this can break down a borrower's interest rate into three parts. These three parts are base mortgage rate, credit risk (determined by the borrower's LTV and credit score), and a residual that reflects strategic lender pricing. My study centers on the racial disparities that emerge in this strategic lender pricing to highlight potential disparate stemming from banks.

I use the HMDA-GSE merged dataset to empirically analyze how the difference in interest rates between minority borrowers and white borrowers changes after the CFPB disclosures complaint narratives. The baseline regression model is presented as follows:

$$
\begin{aligned}
InterestRate_{ilt} &\\
&= \alpha + \beta_1 Treat_l \times Post_t \times Minority_i + \beta_2 Post_t \times Minority_i \\
&+ \beta_3 Treat_l \times Minority_i + \beta_4 Minority_i \\
&+ \mu_{Bucket \times LoanPurpose \times YearMonth} + \mu_{Lender \times YearMonth} \\
&+ \mu_{BorrowerCharacteristics} + \epsilon_{ilt}
\end{aligned} \tag{1}
$$

My primary empirical framework employs a triple-difference approach at the mortgage

[18] For more details, refer to A.B. & S. Auto Service, Inc. v. South Shore Bank of Chicago, 962 F. Supp. 1056 (N.D. Ill. 1997) which states: "[In a disparate impact claim under the ECOA], once the plaintiff has made the prima facie case, the defendant-lender must demonstrate that any policy, procedure, or practice has a manifest relationship to the creditworthiness of the applicant...". Also, see Lewis v. ACB Business Services, Inc., 135 F.3d 389, 406 (6th Cir. 1998) and Miller v. Countrywide Bank, NA, 571 F.Supp.2d 251, 258 (D. Mass 2008) for more insights.

loan level. I investigate how mortgage interest rates change among different racial groups after the disclosure. In Equation (1), i, l, and t index loan applicants, lenders, and time (year-month), respectively. $Treat_l$ categorizes lenders into treated and control groups based on whether lender l was under CFPB supervision in 2015 (i.e., covered by the disclosure policy). The regression sample only includes banks with total assets of less than $100 billion as of the first quarter of 2015. This mitigates confounding factors arising from variations in regulatory pressure, capital structures, or economies of scale between large and small banks. $Post_t$ equals one if it is June 2015 or later, marking the effective initiation of the CFPB disclosure policy. $Minority_i$ equals one if the loan applicant i is Black or Latinx. Accordingly, the coefficient β_1 in the triple-difference estimation signifies the impact of the disclosure on racial gaps in interest rates. It estimates the difference in changes in racial gaps between treated lenders and control lenders whose complaint narratives are not disclosed to the public.

As stated earlier, the pricing grid is critical in mitigating the influence of credit risk in my identification. Nonetheless, the interest rates within the pricing grid may change with time or loan purposes. Therefore, in Equation (1), I integrate dummy variables corresponding to the 8x9 grid levels (referred to as "bucket" in the equation). These variables interact with the loan's year, month, and cash-out refinance status in my regression model. Additionally, my regression model considers lender interactions with the year and month. These fixed effects control for differences in lender pricing and temporal variations. Furthermore, I incorporate fixed effects related to various loan borrower characteristics. These include the loan amount deciles, applicant income deciles, applicant gender, and whether the applicant has a co-applicant. This comprehensive inclusion further controls for potential factors that may impact loan pricing. Standard errors are double clustered at the lender and year levels.[19]

4.2 Baseline Results

Table 3 presents the outcomes of my baseline model. Column (1) includes only the first two fixed effects in Equation (1). In column (2), I incorporate fixed effects for borrower characteristics. Lenders may have different pricing strategies (i.e., different markups) when using the pricing grid. So I interact the lender fixed effects with the bucket-purpose-time fixed

[19]The findings hold steady when I cluster standard errors at the lender level, as shown in Table A3 in the Appendix.

effects in column (3). Furthermore, I construct a similar borrower identifier to approximate individual fixed effects.[20] I control for these fixed effects in column (4). This allows a direct comparison regarding the impacts of disclosure on borrowers with similar creditworthiness but different races. In the final column, I include joint lender-similar borrower fixed effects. Controlling for these joint fixed effects allows for a comprehensive analysis of the behaviors of nearly identical borrowers within the same lender before and after the implementation of the disclosure policy. It ensures that my estimation results are unaffected by the bank-consumer relationship.

[Insert Table 3 about here]

Table 3 displays the estimated coefficients on the *Minority* variable. These coefficients reveal that minority borrowers face interest rates that are 2.5 to 2.9 basis points higher compared to those of white borrowers before the disclosure policy. Since the inception of the Mortgage Bankers' Association Annual Performance Report in 2008, the average net production income has been 55 basis points in the principal year.[21] Thus, an increase of 2.7 basis points in the interest rate equates to a 1.88 basis points rise in annual principal repayment. This constitutes 3.4% of the bank's annual net profit.[22] However, this discrepancy changes after the implementation of the CFPB disclosure policy. The estimated coefficient on the triple-difference term, Treat×Post×Minority, indicates that after disclosure, the interest rate applied to minority borrowers is reduced by 2.3 to 2.8 basis points relative to white borrowers. This reduction mitigates approximately 87% of the existing racial disparities. Evidently, the enforcement of the disclosure policy effectively saves minority mortgage borrowers an average of $102 million annually, due to reduced

[20]The borrower fixed effect under scrutiny results from interplays among the following categorical variables: pricing grid, loan purpose, loan amount deciles, applicant income deciles, applicant gender, and co-applicant status. Leveraging these attributes allows me to segment loan applicants into extremely granular categories, thus identifying borrowers with remarkably similar traits. Given that borrower characteristic fixed effects are absorbed by similar borrower fixed effects due to collinearity, there is no necessity to control borrower characteristics in columns (4) and (5).

[21]For more details, see: https://www.mba.org/news-and-research/newsroom/news/2023/04/06/imb-production-profits-falls-to-series-low-in-2022

[22]Consider a mortgage loan where the home price is $200,000 and the down payment is 20%. With an annual interest rate of 4%, an increase to 4.025% would mean an additional annual repayment of approximately $36, representing 2.25 basis points of the principal (80% of $200,000 equals $160,000). The specific numbers used in this example do not drastically alter the final outcome.

discrimination.[23]

Table 3 also shows that the model effectively explains 76% to 82% of the variations in interest rates observed within our sample. The remaining unexplained variance might be attributed to strategic pricing. Further, to validate the applicability of the settings in this study, I use GSE loans from my dataset to replicate the primary results of Bartlett et al. (2022). These results are shown in Table A4 in the Appendix. I employ the same setting in Table 3 of Bartlett et al. (2022) to estimate the differences in purchase and refinance loan interest rates for minority borrowers. Remarkably, the sign and magnitude of the coefficients obtained are similar to the estimates of Bartlett et al. (2022), even though I use different matching methods to get the final dataset. This consistency indicates that my results are free of sample selection bias in the matching process.

Another concern could be the potential differences in discount points chosen by minority and white borrowers.[24] This discrepancy may explain the considerably higher interest rates paid by minority borrowers compared to their non-minority counterparts in Federal Housing Administration (FHA) loans, as suggested by Bhutta and Hizmo (2021).[25] However, this conclusion remains a subject of debate. Bartlett et al. (2022) identify continued discriminatory pricing within FHA loans even after accounting for discount points using a larger sample. In my study, I focus on GSE loans to address this debate, as both Bartlett et al. (2022) and Zhang and Willen (2021) identify instances of discriminatory pricing within the GSE mortgage market. Additionally, I demonstrate the robustness of my results by examining whether significant differences exist in racial gaps in discount points

[23] According to a report by the Federal Reserve of New York, the market size of housing debt in 2022 is approximately $12.26 trillion. The Survey of Consumer Finance highlights that the segment of African-American/Latinx contributors in the mortgage market approximates to about 13.3%. Data from the HMDA indicate that the loan amount processed by banks under the CFPB regulation constitutes 35.75% of the total loan volume post-2015. Back-of-the-envelope calculations suggest that if the interest rate charged to minority borrowers falls by an average of 2.55 basis points following the disclosure, it will result in an annual decrease in their repayment by an amount equivalent to 1.77 basis points of the principal. Consequently, the minority borrowers within the banks regulated by the CFPB in the total housing debt would benefit, resulting in an annual saving of approximately $102 million.

[24] In the mortgage market, borrowers have the option to pay discount points - an upfront lump sum to the lender in exchange for a reduced loan interest rate. Conversely, they can opt for negative points, receiving credits from the lender while agreeing to pay a higher loan interest rate. Hence, the observed racial differences in interest rates could result from endogenous discount point decisions if there is a difference in the preferences for discount points between minority and non-minority borrowers.

[25] This notion aligns with perspective in Zhang and Willen (2021) that no actual discrimination present in FHA loans.

between the treated and control groups. Given that the HMDA only provides information about discount points after 2018, I utilize the HMDA-GSE merged dataset for 2018-2019 to examine disparities in discount points within the same framework as Table 3 (adding the interest rate as a control variable). The findings are in Table A5 in the Appendix. I do not find significant differences in racial gaps in discount points between treated and control lenders. This outcome effectively addresses the concern that the observed racial gaps might be caused by omitted variables like discount points in baseline results.

Finally, I adopt an event study approach to examine the dynamic impacts of disclosure. An important concern about the triple difference analysis is whether the racial gaps in both treated and control groups follow a parallel trend before the CFPB disclosure. In this study, the treated and control groups are categorized based on bank size. The trends of racial gaps may differ among banks of different sizes.

Figure 3 displays the quarterly coefficient estimates for racial gaps in interest rates within treated lenders compared to control lenders. I examine results from five quarters preceding the implementation of the disclosure policy (effective in the second quarter of 2015) and ten quarters post-disclosure. Any sample periods beyond this selected range are incorporated into the beginning periods (before 2014) and end periods (after 2017). The baseline periods encompass the fifth quarter before the disclosure's start date and prior periods.

[Insert Figure 3 about here]

From Figure 3, I observe no substantial difference in racial gaps in interest rates between the treated and control groups before implementing the disclosure policy. However, there is a sudden decrease in the racial gaps in interest rates for treated lenders compared to control lenders from the first quarter after the disclosure. Although coefficient estimates are not significantly different from zero for some periods, the overall impact remains negative throughout the entire subsequent period (the tenth quarter and all subsequent quarters post-disclosure). This trend confirms that the racial gaps in interest rates in treated and control groups followed parallel trends before the policy's implementation. The treatment effects are only apparent after the enactment of the disclosure policy and last for an extended period.

Another concern is whether a different regulation coincides with the enforcement of the disclosure policy and could have concurrently influenced the racial disparities in interest rates within both the treated and control banks. Under these circumstances, parallel trends would still be satisfied. However, this confounding factor may lead to the primary outcomes rather than the disclosure policy. To alleviate this concern, I thoroughly examine changes in the CFPB's regulatory policies. I find no other concurrent regulations in 2015 that could plausibly influence my results.

4.3 Moderation Effects

In this section, I analyze the heterogeneous effects of the CFPB disclosure policy. I focus on how banks respond to the disclosure of complaint narratives differently across various scenarios. These scenarios include: (i) the complaints against banks involving discriminatory treatments, (ii) banks located in areas with considerable Google search attention towards the CFPB, and (iii) increased risk of discrimination litigation against banks.

In my initial series of examinations, I focus on discrimination complaints. The CFPB publicly discloses complaint narratives for all supervised banks. However, only a subset of these banks receives complaints related to racial discrimination. Through textual analysis, I divide the treated group into two subgroups based on whether they received complaints about racial discrimination in 2015.[26] One subgroup consists of banks under CFPB supervision with discrimination complaints. The other includes supervised banks without such complaints.

I conduct separate regression for both groups following the setting of Equation (1). The results are presented in the first two columns in panel A of Table 4. Column (1) of panel A reports the changes in racial gaps after disclosure between treated banks with discrimination complaints and control banks. It shows a significantly negative coefficient on *Treat×Post×Minority*. In contrast, column (2) represents the outcomes for treated banks without discrimination complaints. It reveals a negative but statistically insignificant coefficient regarding the triple interaction term. These findings underscore that disclosure more noticeably mitigates racial gaps among banks already engaged in discriminatory practices.

In the second test, I examine the moderation effect of social attention on banks' decision-making. This task is difficult due to the limited data on public interest in the disclosure. To address this challenge, I calculate the state-level variations in the Google Search Index for the term "CFPB" during the 12 months preceding the disclosure date. I assign a *High*

[26] For the method of identifying whether complaints contain discrimination, please refer to Li (2023). Within the HMDA-GSE merged dataset utilized for the analysis in this section, 32.61% of treated banks received complaints pertaining to discrimination in 2015.

indicator with a value of one for observations in states with values above the median, and a value of zero otherwise, following Dou et al. (2023).[27] Using this indicator, I divide the treated group into two subgroups according to the Google Search Index levels in the states where the banks are located.

I use the same setting as the analysis of discrimination complaints. The findings are presented in columns (3) and (4) in panel A of Table 4. The outcome in column (3) is statistically significant (column (3)), but the results in column (4) are not significant. These results imply that internet searches conducted by social groups, serving as an indicator of social attention, contribute to mitigating racial disparities in the mortgage market. Social awareness may prompt banks to prioritize their reputation and curtail misconduct.[28]

[Insert Table 4 about here]

Finally, I explore if disclosure has a greater impact on reducing racial disparities during periods with more litigation cases associated with racial discrimination. From 2011 to 2019, there were eight litigations concerning discrimination. The involved banks incurred various penalties. Some were mandated to invest in subsidy programs aimed at assisting minority borrowers.[29] To evaluate the efficacy of these potential litigation risks in mitigating racial disparities, I enhance Equation (1) by introducing a moderating variable, *Litigation*. It captures litigation intensity over time in terms of litigation number and penalty amount. These variables quantify the monthly occurrences of litigations and their corresponding penalties. While identifying this effect, including banks involved in litigation within the sample may cause endogeneity bias. This is because banks subject to legal penalties would have a strong motivation to minimize discriminatory actions. Hence, by observing the operation of their peers, I can assess the reactions of these peer banks to potential litigation threats.

After excluding samples involving litigated banks, I proceed to estimate this moderation

[27] Google monitors and compiles users' search volume by keywords for each state, then calculates the search volume index as the proportion of searches from a given state to those from the state with the highest search volume.

[28] Darian Dorsey, Chief of Staff of the CFPB, provides anecdotes about banks linking executive bonuses to their responsiveness in handling complaints. This anecdotal evidence suggests that bank executives are incentivized to address complaints, as a portion of their bonuses is tied to the efficacy of the banks' complaint resolution (Cortez (2015)).

[29] The data regarding these litigations are gathered from the CFPB website. For detailed information, see Table A6 in the Appendix.

effect. The outcomes are detailed in panel B of Table 4. The first two columns use *Litigation* to present the total number of CFPB enforcements. The last two use *Litigation* to denote the penalty amount charged to lenders. Columns (2) and (4) take advantage of variations within the individual level using similar borrower fixed effects. These are consistent with the specification used in column (4) in Table 3. Column (1) indicates that when the total number of litigations increases by one in a given month, the additional interest for minority borrowers drops by six basis points in the post-disclosure period. Column (3) states that when the penalty of litigations increases by one million dollars in a month, minority borrowers experience a decrease of 1.4 basis points in the additional interest in treated banks after disclosure. Both outcomes are statistically significant at the 5% level. These findings strongly suggest that the CFPB's action imposes litigation threats on banks that the CFPB does not sue. Banks are aware of this threat and might restrain illegal behavior to reduce their exposure to litigation risk. This finding holds significant policy implications. It indicates that the CFPB oversight can indirectly reduce discriminatory treatment in the lending market. To my best knowledge, my paper is the first to analyze the impact of the CFPB oversight on preventing discrimination in the mortgage market.

5 CFPB Disclosure and Outcome Tests: Evidence from Default Rates

5.1 Research Design

In this section, I focus on the default rate of loans. Default actions depend on many factors beyond the sole influence of banks. Thus I employ the "outcome test" approach to identify racial discrimination and the corresponding impact of disclosure policy on it (Becker, 1957, 1993). Outcome tests assess whether loans provided to marginal minority borrowers yield higher profits compared to those to marginal white borrowers. If marginal minority borrowers indeed generate higher profits, it signifies that lenders establish a higher screening threshold for minority applicants due to racial prejudices. The profit that banks earn from consumers depends on two main factors. One is the interest rate that directly affects the gains of banks. The other is the default rate which leads to losses for banks. In prior analyses, I demonstrate that minority borrowers face higher interest rates. This section primarily examines whether racial gaps also persist in default rates and whether the disclosure policy plays a mitigating role in this context. Detecting a substantial discrepancy in default rates between marginally qualified minority and White borrowers is critical. Such difference would suggest that omitted variables or statistical discrimination are not the main cause but rather racial biases within banks during the consumer selection procedure.[30]

However, conducting outcome tests accurately has two challenges at least (Butler et al. (2023)). Firstly, researchers cannot directly observe the literal marginal borrower. This is a situation known as the "infra-marginality problem" (Ayres (2002)). The definition of marginal borrower is neither standardized nor publicly available across banks. Thus researchers can only identify them by controlling covariates. However, inaccuracies can arise if these covariates fail to completely eliminate the creditworthiness disparity between minority borrowers and Whites. If so, the difference in the default rates might not exclusively result from racial bias. In addition, if the distribution of creditworthiness for minorities remains lower than that for whites even after controlling for covariates, marginal minority borrowers might exhibit higher default rates than Whites even without racial bias. As a result, I can observe lower default rates among marginal minority borrowers only when significant racial bias is present and capable of counterbalancing these confounding factors.

A second concern is that researchers cannot directly observe loan profitability. When focusing on the default rate, potential misestimations of the disparity between minority and White borrowers may occur. This can happen when other factors that influence banks' profitability assessments are overlooked. For example, the prepayment risk might be higher for White borrowers. Banks could decrease their evaluation of the profitability of

[30] Discrimination encountered by minority borrowers in the lending market may arise from two potential sources: banks' racial biases or statistical discrimination. Racial bias encompasses both taste-based discrimination and miscalibrated beliefs (Becker (1957); Bordalo, Coffman, Gennaioli, and Shleifer (2016); Arnold, Dobbie, and Yang (2018); Egan, Matvos, and Seru (2022); Ewens and Townsend (2020)). Taste-based discrimination implies banks' dissatisfaction with approving applications from minority borrowers or contentment with approving those from White borrowers. Miscalibrated beliefs involve banks harboring incorrect stereotypes about minority borrowers. Since these biases are typically implicit and can contribute to discrimination, I combine them into a single group. An alternative explanation pertains to the concept of statistical discrimination (Phelps (1972); Arrow (1974); Ewens, Tomlin, and Wang (2014)). In the context of this research, statistical discrimination manifests when banks treat minority and non-minority borrowers differently due to inherently higher risks associated with minority borrowers. As a result, from a profit-maximization perspective in business, it is plausible that banks might reject a larger proportion of applications from minority borrowers.

White borrowers. In such cases, banks might not require minority borrowers to maintain significantly lower default rates to generate profits higher than those from White borrowers. This makes it more difficult to observe lower default rates among minorities.

Conducting outcome tests by comparing default rates among marginal minority and White borrowers might not be perfect. However, the preceding analysis underscores the conservative nature of this approach. The outcome test can only detect a significantly lower default rate among marginal minority borrowers compared to marginal White borrowers when substantial racial bias exists in banks' consumer selection. Hence, any evidence of lower default rates for minorities should be regarded as compelling proof of discrimination. It indicates that banks use stricter approval standards for minority applicants.

The outcome test method has proven effective in the auto loan market (Butler et al. (2023)). Following Butler et al. (2023), my research investigates racial disparities in the mortgage market. I also examine the impact of the disclosure policy on racial gaps in default rates, particularly among marginal borrowers. If disparities in default rates between racial groups decrease after disclosure, specifically if marginal minority borrowers have higher default rates, it strongly suggests that disclosing complaint narratives encourages banks to mitigate their stringent criteria. It leads to a more equitable screening standard for both minority and White borrowers' loan applications.

5.2 Outcome Test Results

I define default rates following Butler et al. (2023). A loan is considered as defaulted when the borrower becomes delinquent for 90 or more days. I restrict the sample to individuals with credit scores below 660 to address the issue of identifying marginal borrowers. This approach significantly reduces the observation number.[31] Due to the distinct factors influencing defaults and interest rates, the fixed effects designs used here are slightly different from those in Section 4. Equation (2) shows the identification strategy for the "outcome test"

[31] The handling of the sample here aligns with Section 4, where observations from large banks (those with total assets exceeding $100 billion) are excluded.

approach:

$$
\begin{aligned}
&DefaultRate_{ilt} \\
&\qquad = \alpha + \beta_1 Treat_l \times Post_t \times Minority_i + \beta_2 Post_t \times Minority_i \\
&\qquad + \beta_3 Treat_l \times Minority_i + \beta_4 Minority_i + \gamma X_{ilt} + \mu_{Lender \times YearMonth} \\
&\qquad + \mu_{Zip \times Year} + \epsilon_{ilt}
\end{aligned} \tag{2}
$$

In the equation, the symbols i, l, and t still represent loan applicants, lenders, and time (year-month) respectively. Table 5 outlines tests using a triple-difference approach to identify the impact of disclosure on racial disparities in default rates. This approach considers loan characteristics such as interest rate, purpose, and the logarithm of the loan amount. Applicant characteristics such as gender, the logarithm of income, and the presence of a co-applicant are represented by X_{ilt} in Equation (2). In addition, I incorporate zip-by-year fixed effects to capture potential influences on loan default rates stemming from changes in local economic conditions. I include pricing grid-loan purpose fixed effects in column (3) to better account for borrower creditworthiness.

Columns (1)-(3) present the discriminatory treatment experienced by minority borrowers in the subprime sample and the impact of the disclosure policy on mitigating this inequitable treatment. The negative and statistically significant results on the *Treat×Minority* coefficient confirm the existence of discrimination against minorities in treated banks. The *Treat×Post×Minority* coefficient shows an average increase of 28.8 percentage points (70.1% of one standard deviation) in the default rates of minorities within the treated group compared to the control group after disclosure. This outcome reflects that racial disparities in default rates are almost eradicated after disclosure. It emphasizes the strong impact of the disclosure policy.[32]

[Insert Table 5 about here]

In the last column of Table 5, I do not apply the cutoff set used to identify marginal borrowers. This full sample outcome does not offer sufficient evidence of discrimination against minorities. The full sample includes many borrowers distant from the credit provision margin. This makes the test less sensitive to margin differences. It implies that applying

[32] To address potential concerns about the parallel trend assumption, I also adopt an event study methodology to scrutinize the dynamic impacts of disclosure. For more details, refer to Figure A1 in the Appendix.

the test to the subprime borrower sample who are closer to the margin is an appropriate identification approach.

6 CFPB Disclosure and Racial Gaps in Rejection Rates

6.1 Overall Trend Analysis

Previous examinations indicate that the disclosure policy nearly eliminates disparities in default rates among marginal borrowers of different races. These differences may result from banks applying stricter admission criteria to minority consumers. The implementation of the disclosure policy has the potential to reduce banks' propensity to maintain such strict and unfair standards. In this section, I focus on loan application denials. I aim to address the following two questions: (i) whether this biased screening standard exists before the disclosure, and (ii) whether the disclosure policy mitigates these disparities. Initially, I analyze the historical trend of loan application rejections across the full sample based on the HMDA origination dataset. It is presented in Figure 4.

Figure 4 depicts the trend of racial disparities in residualized rejection rates within banks under CFPB supervision and those except from such oversight. This analysis focuses on the period around the CFPB's disclosure in June 2015. I calculate differences in rejection rates between minority and non-minority groups for supervised and unsupervised banks separately. Using the racial gaps in rejection rates can avoid the influences of different loan screening criteria between CFPB-supervised and unsupervised banks. This establishes the groundwork for a comparative analysis of racial disparities between these groups of banks.[33] Further, the residualized rejection rates are devoid of specific observable borrower risk factors and the fixed effects arrangement outlined in Equation (3). I use the residualized rejection rates to control for borrower creditworthiness and plot changes in residualized rejection rates.

[33] Assume CFPB-supervised and unsupervised banks have distinct risk preferences. The former exhibits greater risk aversion. Then borrowers face higher rates of rejection in CFPB-supervised banks. Under such a scenario, comparing the disparity in rejection rates of minority borrowers between CFPB-supervised and unsupervised banks directly becomes untenable (given the divergent and stringent selection criteria of CFPB-supervised banks, which introduce differentiation in the traits of minority borrowers between the two types of banks). However, by initially conducting an intra-group subtraction (i.e., subtracting the rejection rate of White borrowers from that of minority borrowers), we effectively mitigate the disparity arising from non-discriminatory loan screening standards.

This approach allows for a precise assessment of the disclosure policy's influence on racial disparities after controlling for temporal effects and borrower traits.[34]

Figure 4 shows the trend of racial gaps in rejection rates. It highlights that the difference in rejection rates between minority and non-minority consumers is more pronounced in banks supervised by the CFPB. In other words, the racial gaps are more substantial among banks operating under CFPB supervision. Following the implementation of the disclosure in mid-2015, the decrease in excess rejection rates of minority borrowers is notably more apparent in banks under CFPB supervision compared to those without such oversight. Additionally, the parallel trends in rejection rate differences across different bank groups before disclosure validate the identification strategy used later in this study.

[Insert Figure 4 about here]

6.2 Empirical Analysis

In this section, I empirically assess whether the disclosure policy alleviates the discrimination in loan accessibility faced by minorities. This evaluation is based on the HMDA origination dataset. I exclude the influence of large banks determined by total assets. This exclusion is in line with previous analyses.[35] I employ the loan-level rejection dummy variable as the

[34]Directly observing the disparity trend between minority and non-minority groups can partially mitigate the disturbance caused by inter-bank heterogeneity. However, if notable characteristic differences exist between minority consumers attracted to CFPB-supervised and CFPB-unsupervised banks, the racial gaps within these groups remain incomparable. By incorporating residualized rejection rates, I can optimally mitigate interferences from both bank and consumer characteristics. This residualization process is designed to nullify statistical discrimination within loan scenarios (Phelps (1972); Arrow (1974); Ewens et al. (2014)). Statistical discrimination refers to banks' reasonable rejection or approval decisions based on the specific features of minority and non-minority borrowers to maximize profits without racial bias.

[35]Large lenders are excluded here because they might have their own credit risk models. These models may incorporate legitimate-business-necessity variables that are not visible to the GSE underwriter (and by extension, not to me as a researcher). My empirical model might inadvertently introduce credit risk into my estimates if these lenders use fundamental models to selectively retain loans in their portfolios (Bartlett et al. (2022)). Therefore, I exclude banks with total assets exceeding $100 billion to mitigate this concern.

dependent variable. The model is as follows:

$$
\begin{aligned}
&RecjectionRate_{ilt} \\
&\qquad = \alpha + \beta_1 Treat_l \times Post_t \times Minority_i + \beta_2 Post_t \times Minority_i \\
&\qquad + \beta_3 Treat_l \times Minority_i + \beta_4 Minority_i + \mu_{Lender \times Year} \\
&\qquad + \mu_{Lender \times Similar\ Borrower} + \mu_{Lender \times Census\ Tract} + \mu_{Census\ Tract \times Year} \\
&\qquad + \epsilon_{ilt}
\end{aligned}
\tag{3}
$$

In Equation (3), the symbols i, l still represent loan applicants and lenders, respectively. The symbol t has now been modified to indicate the year. Table 6 presents the results of these regressions. The first two columns show the baseline outcomes. I adjust for lender-by-year fixed effects, lender-similar borrower joint fixed effects, and lender-census tract joint fixed effects in column (1),[36] and I introduce census tract-by-year fixed effects in column (2).[37] My findings reveal that minority applicants are 3.5 percentage points less likely to get loan approval than White applicants in the control group. This number increases by 4.2 percentage points in the treated group (column (2)). It indicates that minority applicants in treated banks face a 7.7 percentage point decrease in loan approval likelihood compared to White applicants. A preliminary calculation implies that this equates to more than 70,000 minority applicants in treated banks failing to obtain loans annually.

[36]The declined loans are included within HMDA origination dataset. So there is a substantial increase in the total sample size. This increase allows for a more granular fixed effects setting. Additionally, in the mortgage market, a loan officer may discourage a prospective minority applicant from applying or guide non-minority applicants to improve their probability of loan approval. Suppose this situation results from the behavior prevalent across a branch by loan officers. In that case, the lender by census tract fixed effects will allow me to measure discrimination in rejection rates at the branch level.

[37]The "similar borrower" variable incorporated in this test aligns closely with those used in Section 4. The only difference is that I replace the credit score and LTV group from the HMDA-GSE merged dataset with the median credit score and LTV at the census-tract level. This difference is due to the lack of individual-level credit score and LTV data in the HMDA dataset. Thus the census tract median serves as an approximation. Bartlett et al. (2022) confirm that this approximation does not affect conclusions related to discrimination. For further clarification, please refer to Section 3.1.

These loans would have been granted if they were White.[38] Nonetheless, subsequent to the implementation of disclosure, the discrimination faced by minorities in accessing loans diminishes by 1.6 percentage points. This implies that annually, over 14,000 minority households will successfully apply for a mortgage loan due to the disclosure policy.[39]

[Insert Table 6 about here]

I further explore the disclosure impact by investigating whether banks respond differently to the disclosure of complaint narratives under the following conditions: (i) the complaints received by banks are related to discriminatory treatments, and (ii) banks operate in regions where the CFPB attracts substantial Google search attention.

Columns (3) to (6) in Table 6 illustrate the moderating effects of discrimination complaints and Google attention. The *Treat×Minority* coefficient is positive and significant in column (3). It indicates that within the treated group, minority borrowers face higher rejection rates in banks with noticeable discrimination. The coefficient on the same variable in column (4) has a decreased magnitude. It conveys that banks without discrimination complaints have lower additional rejection rates. These findings further validate my discrimination measurement. The coefficient of *Treat×Post×Minority* shows the moderating effects of discrimination complaints. The result in column (3) indicates that banks already engaged in discriminatory practices play a predominant role in driving the observed treatment effect following disclosure. Additionally, the results from columns (5) and (6) suggest that societal attention also enhances the positive effects of disclosure. Furthermore, I adopt an event study approach to examine the time-varying influences of disclosure, building upon the specification outlined in column (2). The findings are detailed in Figure A2 in the Appendix.

[38] I multiply the annual aggregate of minority mortgage loan applications by the change in their probability of approval to calculate these estimates. The approval rate may either decrease due to discrimination or increase owing to the disclosure policy. Data from the HMDA origination dataset indicates that after 2015, the annual average of mortgage loan applications totals approximately 19.16 million. 34.91% of these applications are received by banks regulated by the CFPB and 13.77% from minority borrowers. Based on these numbers, I roughly calculate that there are about 0.92 million (19.16 million × 34.91% × 13.77%) minority mortgage loan applications per year in CFPB-supervised banks. Therefore, roughly 70.92 thousand (0.92 million × 7.7%) applicants may be denied annually due to racial bias.

[39] Based on the previous framework, this calculation illuminates that the disclosure policy's implementation benefits an estimated 14.73 thousand (0.92 million × 1.6%) minority applicants annually. It is clear that my estimates in this subsection may be conservative as the HMDA origination dataset does not encompass all applications in the U.S. mortgage market.

7 Reaction of Other Stakeholders

7.1 Consumer's Reaction Towards the Disclosure

Prior analyses mainly focus on banks' reactions to disclosure. In reality, the responses of other stakeholders are equally significant. I will separately discuss the reactions of consumers, rival banks, and stock market investors to this policy. In this section, I concentrate on consumers' responses to the CFPB's disclosure policy. Existing literature substantiates that consumers make their purchasing decisions based on product or service reviews. Within lending markets, Dou and Roh (2023) find that consumers tend to avoid banks with a high number of complaints after the implementation of CFPB's 2013 disclosure policy. This avoidance leads to reduced loan applications. I further examine the effect of the disclosure of complaint narratives, particularly focusing on minority borrowers' reactions. If consumers are able to utilize the published complaint narratives, they may opt for banks that engage in fewer discriminatory practices when applying for mortgage loans. These banks signal fairer treatment for minority consumers.

I use data aggregated from the HMDA origination dataset to discuss consumers' reactions. The identification strategy is consistent with the previous sections. The regression model takes the following form:

$$\begin{aligned} &Application_{ilct} \\ &\quad = \alpha + \beta_1 Treat\_Complaints_l \times Post_t \times Minority_i \\ &\quad + \beta_2 Treat\_Non\ Complaints_l \times Post_t \times Minority_i + \delta T_{ilt} + \gamma X_{lct} \\ &\quad + \mu_{Lender \times Year} + \mu_{Lender \times County} + \mu_{County \times Year} + \epsilon_{ilct} \end{aligned} \tag{4}$$

Equation (4) uses data aggregated at the county-lender-minority-year level. This means that a lender, l, in a county, c, has two observations, i, representing the total applications from minority and non-minority applicants respectively. The dependent variable *Application* has two forms in my analysis: (i) the log of the number of mortgage applications from minority or white borrowers aggregated at a specific county-lender-year level, and (ii) the log of the dollar amount of mortgage applications from minority or white borrowers aggregated at a specific county-lender-year level. Importantly, to better capture consumers' perceptions of discriminatory practices, I categorize CFPB-supervised banks into two groups, *Treat_Complaints* and *Treat_Non Complaints*.[40] These two treated groups are compared with banks in the control groups. The coefficients represented by β_1 and β_2 in Equation (4) are of primary interest. They reflect whether consumers pay attention to and benefit from the disclosure of complaint narratives. In T_{ilt}, I control for other triple-difference-related items, such as interaction terms and single terms.

Following Dou and Roh (2023), control variables, X_{lct}, in Equation (4) include: (i) the proportion of mortgages that a lender approves in a specific county, as higher approval rates may lead to more applications; (ii) a dummy variable, assigned a value of one if the bank has a physical presence in the county for the given year; and (iii) the logarithm of the total deposits gathered by the bank's branches within the county during that year. Additionally, I control for lender-by-year and county-by-year fixed effects to absorb different time-varying unobservable shocks. I also use lender-county fixed effects as a proxy for branch-level fixed effects.

[Insert Table 7 about here]

Table 7 presents the primary estimated results. In columns (1) and (2), the total number of mortgage applications is the dependent variable. Columns (3) and (4) detail the results for the total dollar amount of mortgage applications. Columns (2) and (4) include lender-county fixed effects, but columns (1) and (3) do not include them. The different specification does not significantly impact the overall results. This indicates the robustness of the analysis here. Both core coefficients exhibit a negative trend regardless of the dependent variable employed. The *Treat_Complaints* group (as indicated by row *1(Complaint)*) shows a more significant and statistically distinct pattern compared to the *Treat_Non Complaints* group (as indicated by row *1(Non Complaint)*). The p-value for the Equality Test reflects the t-test results for the difference between these two sets of estimated coefficients. It is reported in the table. These findings suggest that consumers are likely to extract information from publicly available complaint narratives after disclosure. They subsequently reduce their engagement with banks that have potential discrimination risks.

[40] In this analysis, I categorize each county-lender into two treated groups based on the receipt of discrimination-related complaints during the years 2015-2019. Specifically, if a county-lender receives any discrimination-related complaints in a given year, it is classified into the *Treat_Complaints* group for that year. Otherwise, it is categorized into the *Treat_Non Complaints* group.

7.2 Rival Bank's Reaction Towards the Disclosure

In this section, I explore whether the implementation of the disclosure policy has increased competition among banks in the local market. Dou et al. (2023) find that after the disclosure of the aggregate complaint count against CFPB-regulated banks in 2013, rival banks learn about the service quality of CFPB banks from the total number of complaints. This leads them to enter areas with higher complaint density. Similarly, I investigate whether rival banks use complaint narratives to shape their market entry strategies. Specifically, if incumbent banks under CFPB supervision in a particular region are heavily involved in discriminatory practices, their rival banks may expand and gain market share by catering to underserved minority borrowers. To conduct this examination, I aggregate the HMDA origination dataset to the county-year level and calculate the density of banks at the county level. I use this as the dependent variable and estimate the subsequent model:

$$BankDensity_{it} = \alpha + \beta_1 Complaint\ Share_i \times Post_t + \gamma X_{it} + \mu_{State\times Year} + \mu_{County} + \epsilon_{it} \quad (5)$$

In Equation (5), symbols i and t denote county and year, respectively. The dependent variable, *BankDensity*, represents the total number of unique bank brands per ten thousand residents. The primary independent variable, *Share×Post*, captures the varying degrees of impact on market structures in different counties after disclosing. The metric *Complaint Share* is the total number of discrimination-related complaints received in a county in 2015 divided by the number of loans issued to minority applicants. I use this to measure the extent of the disclosure's impact across various counties. The variable *Post* equals one after the implementation of the disclosure policy. I interact *Complaint Share* with *Post* to identify the differential effects of disclosure across counties based on the specific proportions of discriminatory complaints. In this analysis, I incorporate X_{it} to represent the lagged average approval rate and the lagged average bank deposits (Dou et al. (2023)). Furthermore, I incorporate county-level fixed effects and state-by-year joint fixed effects to absorb the potential impact of local economic and temporal shocks. In the regressions, standard errors are clustered at the county level. These outcomes remain robust even when the standard error is clustered at both the county and year levels.

[Insert Table 8 about here]

Table 8 illustrates the response of rival banks to disclosure. Columns (1) to (3) demonstrate the heterogeneous effects of disclosure on the total number of banks per ten thousand residents in the current period, lagged one period, and lagged two periods. These effects are identified by the continuous treatment variable *Complaint Share*. The results imply that more rival banks enter the market after the disclosure, potentially intensifying market competition. In Table A7 in the appendix, I also investigate the impacts of the disclosure across different counties using an alternative proxy. Specifically, I define *Amount Share* as the market share held by treated banks within each county in 2015. The regression framework is in alignment with that of Table 8. I find the results are robust. This finding indicates that an increase of one standard deviation in *Amount Share* multiplied by *Post* within a county leads to the addition of 1.3 new unique bank brands for every ten thousand people in the following year. This amounts to 4.9% of one standard deviation.

These findings suggest that peer banks may enhance their market position by extracting insights from the disclosed weaknesses in their competitors' operations. These results are stronger in areas with inadequate customer service and instances of discriminatory practices. In other words, if a county is impacted by the disclosure, and the CFPB reveals evident inferior service quality (as reflected in an elevated number of discrimination complaints), rival banks may seize this opportunity to penetrate the local market. Thereby the competition will be heightened. Technological advancements can make it easier to collect and share these complaint narratives swiftly and precisely. These advancements are likely to increase the importance of such disclosures as a policy tool to encourage market competition. [41] The increased competition may help protect minority borrowers from discriminatory treatment in the mortgage lending market (Li (2023)).

7.3 Stock Market's Reaction Towards the Disclosure

In this section, I analyze how investors in the stock market react to the disclosure policy for CFPB-supervised banks. In 2013 the CFPB disclosed complaint counts. Dou and Roh (2023) confirm that the establishment of a complaint database helps spread new information discernible by the stock market. In 2015 the CFPB improved complaint disclosures by

[41] Banks are incentivized to use the complaint database in order to enhance the quality of their mortgage services. This is due to their frequent assessment of their performance against competitors. They use metrics derived from the database, as outlined in the Consumer Financial Protection Bureau (CFPB) report of 2013.

adding complaint narratives. This enhancement allows me to identify complaints linked to discriminatory practices. Consequently, I can closely monitor the stock market's response to banks engaged in discriminatory behaviors compared to those that are not.

I construct a sample of 341 listed banks to assess the market's reaction. I obtain this sample from the HMDA origination dataset. These institutions have no missing daily returns during event windows of three, five, or seven days. I use the U.S. Daily Event Study function available on WRDS. I calculate the cumulative abnormal returns (CAR) within event days [-1, +1], [-2, +2], and [-3, +3]. CAR is the difference between a bank's daily return and the value-weighted CRSP market return. To ensure robustness, I calculate the CAR based on the Capital Asset Pricing Model (CAPM) and the Market Adjusted Model (referred to as *Market* in this section). I then conduct cross-sectional regressions that establish connections between banks' cumulative abnormal returns and their engagement in receiving complaints associated with discrimination:

$$CAR_i = \alpha + \beta_1 Complaint_i + \beta_2 Non\text{-}Complaint_i + \gamma X_i + \epsilon_i \tag{6}$$

In Equation (6), *CAR* represents the abnormal returns for each bank. These returns are calculated using either the CAPM or the Market Adjusted Model across various event windows. I classify the banks in the sample into three distinct groups for identification: those supervised by the CFPB with discrimination complaints in 2015, termed as *Complaint*; those supervised by the CFPB without discrimination complaints in 2015, *Non-Complaint*; and banks not supervised by the CFPB serving as a control group for the aforementioned categories. In order to control for potential confounding variables, I introduce variables including the banks' total assets, equity-to-assets ratio, return on assets, and total deposits. They are established as of the end of 2014 (Dou and Roh (2023)).

[Insert Table 9 about here]

Table 9 presents the stock market's response to banks in the *Complaint* group (or *Non-Complaint group*) relative to those not supervised by the CFPB. The first three columns show the outcomes for the event windows of [-1, +1], [-2, +2], and [-3, +3] trading days, respectively. Furthermore, I conduct a t-test for the two coefficients estimated in every column to assess whether significant differences exist between them. The results show that CFPB-supervised banks experience a negative market reaction after the disclosure.

Moreover, the market response appears more adverse for banks involved in discrimination complaints. The estimated coefficients suggest that banks in the Complaint group experience a drop in CAR ranging from approximately 1.3 to 2.6 percentage points compared to banks not under supervision after the disclosure. The distinction between banks with discrimination complaints and those without is statistically significant, especially within shorter event windows (such as [-1, +1]). Columns (4) to (6) in Table 9 use cumulative abnormal returns based on the Market Adjusted Model. They yield consistent results.

These observations highlight the significant results of the disclosure. They show that the market interprets the release of such information as a negative shock that leads to substantial reactions. On the one hand, the enforcement against discriminatory practices in the U.S. market is stringent. This results in considerable economic and reputational costs once courts have ruled on unlawful discriminatory lending practices.[42] This situation might make investors worried about these banks' future performance. Thereby they will sell stocks associated with such possible legal risks. [43] On the other hand, investors might exhibit aversion to inequality, given the increasing significance of socially responsible investing (Hartzmark and Sussman (2019)). They might be reluctant to invest in banks considered socially irresponsible because these institutions appear to have treated minority consumers unfairly.[44]

8 Real Effects of Disclosure

In my analysis of the rejection rate in Section 6, I find that disclosing complaint narratives helps more than 14,000 minority households successfully apply for mortgage loans annually. To provide additional support for the real-world effectiveness of this disclosure, I examine this inference in this section to determine its actual existence. Specifically, I assess whether

[42] In Table A6 of the Appendix, I document litigations penalizing banks supervised by the CFPB for instances of racial discrimination. In addition, as displayed in footnote 6, over the past decades, the U.S. Department of Justice has taken legal action against numerous major mortgage lenders for violating fair lending principles during the housing boom, leading to settlements exceeding $500 million.

[43] For instance, Wells Fargo shares dropped by 1.3 percent on a down day due to litigation and fines related to discriminatory practices. Refer to https://www.reuters.com/article/us-wells-lending-settlement-idUSBRE86B0V220120712.

[44] Previous research indicates that concerns regarding inequality might have a significant influence within the financial markets (Pan et al. (2022)).

the disclosure policy contributes to reducing the racial gaps in homeownership. I acquire housing occupancy variables for different races at the county-year level from the American Community Survey (ACS). I match this with the county-year level data aggregated from the HMDA origination dataset. The final dataset allows for assessing the real effect of disclosure on homeownership. In fact, this dataset has a structural alignment with the one used in Section 7.2. The regression model I employ is outlined as follows:

$$
\begin{aligned}
&Homeownership_{it} \\
&\qquad = \alpha + \beta_1 Complaint\ Share_i \times Post_t + \gamma X_{it} + \mu_{State \times Year} + \mu_{County} \\
&\qquad + \epsilon_{it}
\end{aligned} \tag{7}
$$

Equation (7) corresponds to Equation (5). Symbols i and t represent county and year, respectively. It is important to highlight that I select three distinct dependent variables to identify racial gaps in the analysis: (i) homeownership among minority residents, specifically Black and Hispanic residents; (ii) homeownership among White residents; and (iii) the disparity between the two indicators.[45] In this analysis, I continue to use the metric *Complaint Share* to measure the magnitude of the disclosure's impact across various counties. To control for potential bias arising from unobservable macroeconomic variables, I incorporate the growth rate of income per capita, unemployment rates, and the logarithm of the population at the county level. I also include the county-level fixed effects and state-by-year joint fixed effects.

[Insert Table 10 about here]

Table 10 presents the estimated outcomes. Column (1) details the effect of the disclosure policy on minority residents. These findings indicate that a one standard deviation change in *Complaint Share* × *Post* results in a 0.32% increase in homeownership for minority residents within a county (corresponding to 2.67% of one standard deviation in homeownership). According to a report from the National Association of Realtors (NAR), the annual average growth rate of homeownership among Black Americans has been only 0.04% over the past

[45] I compute the variable by subtracting homeownership among White residents from homeownership among minority residents. This variable measures the extent to which homeownership among minority residents falls short compared to White residents. Its value is typically negative. Therefore, a positive β_1 signifies a reduction in racial gaps.

decade.[46] This contrast supports the possible strong effect of the disclosure policy in mitigating racial disparities.

Column (2) reports an insignificant impact of disclosure on homeownership rates for White residents. However, column (3) shows a notably positive effect of disclosure on the difference between minority and White homeownership rates. These results support my inference that a one standard deviation increase in the independent variable positively contributes to a 2.79% of one standard deviation improvement in racial gaps. Now, my findings demonstrate that the disclosure of complaint narratives has a substantial and noteworthy economic impact on racial disparities. Therefore, complaint disclosure can be an effective policy tool that can be integrated into the toolbox of financial regulatory entities.

9 Website Ranking Banks for Discrimination

To enhance support for minority borrowers in the United States and to address potential discrimination in mortgage lending practices, I establish an online platform. This website prominently features an analytical framework that ranks banks based on indicators of racial discrimination at the zip code level. Specifically, I ascertain whether minority borrowers are subjected to higher loan interest rates or face increased rejection probabilities by specific banks. Methodologically, the analysis is refined to zip code specificity by including only data pertaining to a single zip code in each regression model. The dataset is consistent with those employed in prior research concerning interest rates or rejection rates. The regression model I employ is outlined as follows:

$$Outcomes_{ilt} = \alpha + \beta Minority_i \times Bank\ id_l + FEs + \epsilon_{ilt} \tag{8}$$

In Equation (8), the symbols i, l, and t represent borrower, lender, and year, respectively. *Outcomes* represent the interest rates or rejection rates for minority borrowers i in bank l. *Minority* equals one if the borrower i is Black or Latinx. *Bank id* specifically identifies bank l. Accordingly, the coefficient β is instrumental in uncovering the disparities in interest rates or rejection rates between minority and non-minority borrowers within bank l. A β value

[46]For more details, please see: https://www.nar.realtor/newsroom/more-americans-own-their-homes-but-black-white-homeownership-rate-gap-is-biggest-in-a-decade-nar

greater than zero would suggest the presence of racial discrimination in bank l, manifesting as higher interest rates or elevated rejection probabilities for minority borrowers. The regression employs the same fixed effects setting consistent with previous examinations of interest rates or rejection rates.

I rank the banks by the degree of discrimination in each zip code, which is determined by the magnitude of the β coefficient in the above results. These rankings, alongside detailed discrimination levels, are made publicly accessible through a website I named Discrimination Warrior (https://discriminationwarrior.com). This platform, upon the input of a specific zip code, displays the extent of discrimination practiced by banks in that locality against minority borrowers, as evidenced by the analyzed interest and rejection rates. Figure 5 and Figure 6 present the results for zip code 02135, a region in Boston. This initiative aims to provide the public with direct and transparent access to local banks' discriminatory practices.

10 Conclusion

The topic of racial disparities remains an ongoing focal point of discussion among scholars and policymakers. In this study, I offer the first comprehensive investigation into the impacts of complaint narrative disclosure on DEI within the financial market. I utilize innovative identification methods such as using GSE's pricing grid and "outcome test" strategy. These approaches are designed to discern racial bias between minority and white borrowers in the mortgage market. They also control for potential differences in risk and creditworthiness. I analyze the interest rates, default rates, and rejection rates for minority borrowers. The results confirm that the implementation of the disclosure policy has significantly reduced previously existing racial disparities. After the implementation of the disclosure policy, minority borrowers in the mortgage save an average of \$102 million annually. Furthermore, the disclosure policy helps more than 14,000 minority households get mortgage loans each year. I also investigate whether the increased loan approval has a real effect on homeownership at the county level. I find that disclosure mitigates racial gaps in homeownership. This result indicates that information disclosure in the lending market significantly influences homeownership by reducing discrimination. Also, the reactions from other stakeholders support these outcomes and provide channels for the effective implementation of the disclosure policy.

Currently, approximately 93% of developed countries have regulatory entities able to acknowledge and resolve customer complaints about financial services. Half of these countries have dedicated financial regulatory bodies that prioritize protecting consumer rights. However, except for the CFPB in the United States, nearly no other financial regulatory agencies in other countries choose to share consumer complaint information publicly.[47] My research demonstrates that the public disclosure of customer complaints effectively alleviates discriminatory and unfair treatment of clients in the U.S. financial system via various channels. Other countries may draw insights from this disclosure policy and develop their own strategies for managing and disclosing consumer complaints. Implementing such measures has the potential to foster DEI globally.

Lastly, I establish an online query platform to help minority borrowers in the United States avoid potential bias in mortgage lending.[48] This system provides the public with direct and transparent access to data on banks' discriminatory practices at detailed location levels. I determine each bank's level of unfair treatment using the excess rejection rate and interest rate applied to minority borrowers. Subsequently, I rank banks based on degrees of unfair treatment. Moreover, I make accessible information related to discriminatory complaints filed with the CFPB on my website. Visitors to this website can easily acquire information related to discrimination levels exhibited by banks within their communities. They do not need to have advanced data analysis skills. My website reduces the barriers to accessing information about discrimination in mortgage lending markets. It presents details of discriminatory treatments in a user-friendly manner. Empowered by my website and advanced data analysis abilities, minority borrowers can enhance their financial literacy and make informed decisions to avoid engaging with discriminatory banks.

[47] Only a few countries provide limited citations or summaries of such data, in the form of case studies or annual reports, rather than full disclosure. For additional details regarding this information, refer to Table A7 in the Appendix.

[48] More information is available on the website: https://discriminationwarrior.com

References

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.

Arnold, David, Will Dobbie, and Crystal S Yang, 2018, Racial bias in bail decisions, *Quarterly Journal of Economics* 133, 1885–1932.

Arrow, Kenneth J., 1974, *The theory of discrimination* (Princeton University Press, Princeton).

Ayres, Ian, 2002, Outcome tests of racial disparities in police practices, *Justice Research and Policy* 4, 131–142.

Bartlett, Robert, Adair Morse, Richard Stanton, and Nancy Wallace, 2022, Consumer-lending discrimination in the fintech era, *Journal of Financial Economics* 143, 30–56.

Becker, Gary S., 1957, *The Economics of Discrimination*, second edition (Univ. Chicago Press, Chicago), 2d ed. 1971.

Becker, Gary S, 1993, Nobel lecture: The economic way of looking at behavior, *Journal of Political Economy* 101, 385–409.

Bennedsen, Morten, Elena Simintzi, Margarita Tsoutsoura, and Daniel Wolfenzon, 2022, Do firms respond to gender pay gap transparency?, *Journal of Finance* 77, 2051–2091.

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

Black, Sandra E, and Philip E Strahan, 2001, The division of spoils: rent-sharing and discrimination in a regulated industry, *American Economic Review* 91, 814–831.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer, 2016, Stereotypes, *Quarterly Journal of Economics* 131, 1753–1794.

Butler, Alexander W, Erik J Mayer, and James P Weston, 2023, Racial disparities in the auto loan market, *Review of Financial Studies* 36, 1–41.

Campbell, J, H Jackson, B Madrian, and P Tufano, 2011, Consumer financial protection, *Journal of Economic Perspectives* 25, 91–114.

CFPB, 2015, Disclosure of consumer complaint narrative data, *Federal Register* 80, 15572–15583, Docket No. CFPB-2014-0016. Final Policy Statement.

Cheng, P, Z Lin, and Y Liu, 2015, Racial discrepancy in mortgage interest rates, *Journal of Real Estate Finance and Economics* 51, 101–120.

Cortez, Nathan, 2015, Agency publicity in the internet era, Working paper, Southern Methodist University.

DeFusco, Anthony A, Stephanie Johnson, and John Mondragon, 2020, Regulating household leverage, *Review of Economic Studies* 87, 914–958.

Dou, Yiwei, Mingyi Hung, Guoman She, and Lynn Linghuan Wang, 2023, Learning from peers: Evidence from disclosure of consumer complaints, *Journal of Accounting and Economics* 101620.

Dou, Yiwei, and Yongoh Roh, 2023, Public disclosure and consumer financial protection, *Journal of Financial and Quantitative Analysis* 1–56.

Egan, Mark, Gregor Matvos, and Amit Seru, 2022, When harry fired sally: The double standard in punishing misconduct, *Journal of Political Economy* 130, 1184–1248.

Engel, Kathleen C, and Patricia A McCoy, 2016, *The subprime virus: Reckless credit, regulatory failure, and next steps* (Oxford University Press).

Ewens, Michael, Bryan Tomlin, and Liang Choon Wang, 2014, Statistical discrimination or prejudice? a large sample field experiment, *Review of Economics and Statistics* 96, 119–134.

Ewens, Michael, and Richard R Townsend, 2020, Are early stage investors biased against women?, *Journal of Financial Economics* 135, 653–677.

Fuster, Andreas, Matthew C Plosser, and James I Vickery, 2021, Does cfpb oversight crimp credit?, Working paper, Ecole Polytechnique Fédérale de Lausanne.

Ghent, Andra C, Ruben Hernandez-Murillo, and Michael T Owyang, 2014, Differences in subprime loan pricing across races and neighborhoods, *Regional Science and Urban Economics* 48, 199–215.

Guiso, Luigi, Andrea Pozzi, Anton Tsoy, Leonardo Gambacorta, and Paolo Emilio Mistrulli, 2022, The cost of steering in financial markets: Evidence from the mortgage market, *Journal of Financial Economics* 143, 1209–1226.

Gurun, Umit G, Gregor Matvos, and Amit Seru, 2016, Advertising expensive mortgages, *Journal of Finance* 71, 2371–2416.

Hacamo, Isaac, 2022, Racial prejudice in the workplace and firm boycotts, Working paper, Kelley School of Business, Indiana University.

Hartzmark, Samuel M, and Abigail B Sussman, 2019, Do investors value sustainability? a natural experiment examining ranking and fund flows, *Journal of Finance* 74, 2789–

2837.

Keys, Benjamin J, Devin G Pope, and Jaren C Pope, 2016, Failure to refinance, *Journal of Financial Economics* 122, 482–499.

Law, Kody, and Nathan Mislang, 2022, The heterogeneity of bank responses to the fintech challenge, Working paper, Cornell University.

Li, X, 2023, Bank competition and entrepreneurial gaps: evidence from bank deregulation, Working paper, Boston College.

Lusardi, A, and O S Mitchell, 2011, Financial literacy and planning: Implications for retirement wellbeing, Working paper, National Bureau of Economic Research.

Murphy, Angela J, 2005, Money, money, money: An exploratory study on the financial literacy of black college students, *College Student Journal* 39, 478–489.

Neugebauer, R., and R. Williams, 2015, Reform the cfpb to better protect consumers, Mimeo, U.S. Congress.

Pan, Yihui, Elena S Pikulina, Stephan Siegel, and Tracy Yue Wang, 2022, Do equity markets care about income inequality? evidence from pay ratio disclosure, *Journal of Finance* 77, 1371–1411.

Phelps, Edmund S, 1972, The statistical theory of racism and sexism, *American Economic Review* 62, 659–661.

Stanton, Richard, Johan Walden, and Nancy Wallace, 2014, The industrial organization of the us residential mortgage market, *Annual Review of Financial Economics* 6, 259–288.

U.S. Chamber of Commerce, 2018, Consumer financial protection bureau: Working towards fundamental reform, Center for Capital Markets Effectiveness.

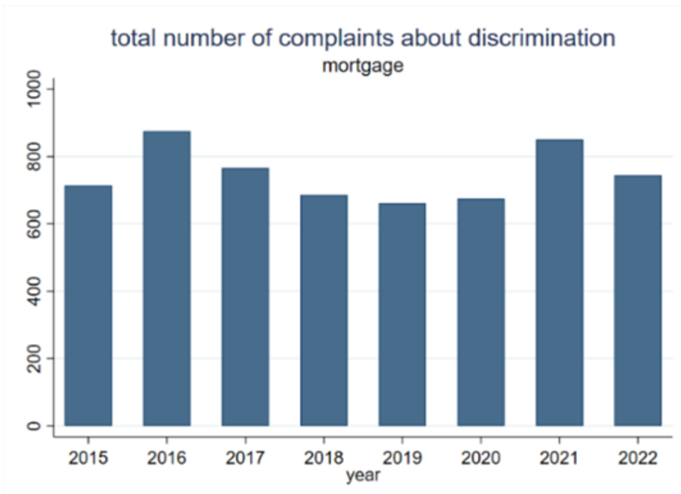
Woodward, Susan E, and Robert E Hall, 2012, Diagnosing consumer confusion and sub-optimal shopping effort: Theory and mortgage-market evidence, *American Economic Review* 102, 3249–3276.

Zhang, David Hao, and Paul S Willen, 2021, Do lenders still discriminate? a robust approach for assessing differences in menus, Working paper, National Bureau of Economic Research.

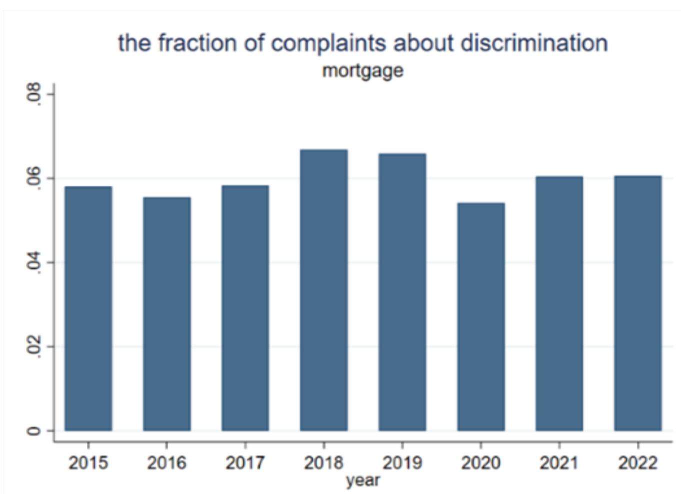
Figures and Tables

Figure 1. Number and Fraction of Complaints about Discrimination in Mortgage Market

These two graphs respectively depict the total number and proportion of consumer complaints related to discrimination in the mortgage market regulated by the CFPB from 2015 to 2022.



(a) Total number of complaints about discrimination



(b) The fraction of complaints about discrimination

Figure 2. Sample Representation of the GSE Grid

The figure illustrates the 2018 LLPA (loan-level price adjustment) grid of Fannie Mae, sourced from the Fannie Mae Selling Guide, published on June 5, 2018. Fannie Mae's LLPA grid has a corresponding grid at Freddie Mac, known as the Credit Fees in Price chart. These matrices dictate the additional g-fee (guarantee fee) that lenders owe the GSE for the mortgage guarantee, fully determined by credit score and LTV.

Table 1: All Eligible Mortgages – LLPA by Credit Score/LTV Ratio

Representative Credit Score	LTV Range Applicable for all mortgages with terms greater than 15 years									
	≤ 60.00%	60.01 – 70.00%	70.01 – 75.00%	75.01 – 80.00%	80.01 – 85.00%	85.01 – 90.00%	90.01 – 95.00%	95.01 – 97.00%	>97.00%	SFC
≥ 740	0.000%	0.250%	0.250%	0.500%	0.250%	0.250%	0.250%	0.750%	**0.750%**	N/A
720 – 739	0.000%	0.250%	0.500%	0.750%	0.500%	0.500%	0.500%	1.000%	**1.000%**	N/A
700 – 719	0.000%	0.500%	1.000%	1.250%	1.000%	1.000%	1.000%	1.500%	**1.500%**	N/A
680 – 699	0.000%	0.500%	1.250%	1.750%	1.500%	1.250%	1.250%	1.500%	**1.500%**	N/A
660 – 679	0.000%	1.000%	2.250%	2.750%	2.750%	2.250%	2.250%	2.250%	**2.250%**	N/A
640 – 659	0.500%	1.250%	2.750%	3.000%	3.250%	2.750%	2.750%	2.750%	**2.750%**	N/A
620 – 639	0.500%	1.500%	3.000%	3.000%	3.250%	3.250%	3.250%	3.500%	**3.500%**	N/A
< 620[1]	0.500%	1.500%	3.000%	3.000%	3.250%	3.250%	3.250%	3.750%	**3.750%**	N/A

Figure 3. Dynamic Trend of the Impact of Disclosure on Interest Rate

This figure represents the quarterly coefficient estimates for racial disparities in interest rates among treated lenders relative to control lenders. The illustration contains outcomes from five quarters before the second quarter of 2015 - the starting point of the disclosure policy - and continues through ten quarters after its application, that is, between 2014 and 2017. Samples outside of this chosen span are grouped into the initial and final periods. During the estimation, the fifth quarter prior to the disclosure and earlier periods are used as the baseline period. The data come from the HMDA-GSE merged dataset spanning 2011-2019. The Y-axis reflects the magnitude of the racial disparities in interest rates, while the X-axis signifies varying periods, with period zero corresponding to the disclosure policy's implementation. The regression specification used to generate this figure is consistent with that of column (2) in Table 3. Black circles symbolize the coefficient estimates for different periods, and the black vertical lines denote the 95% confidence intervals.

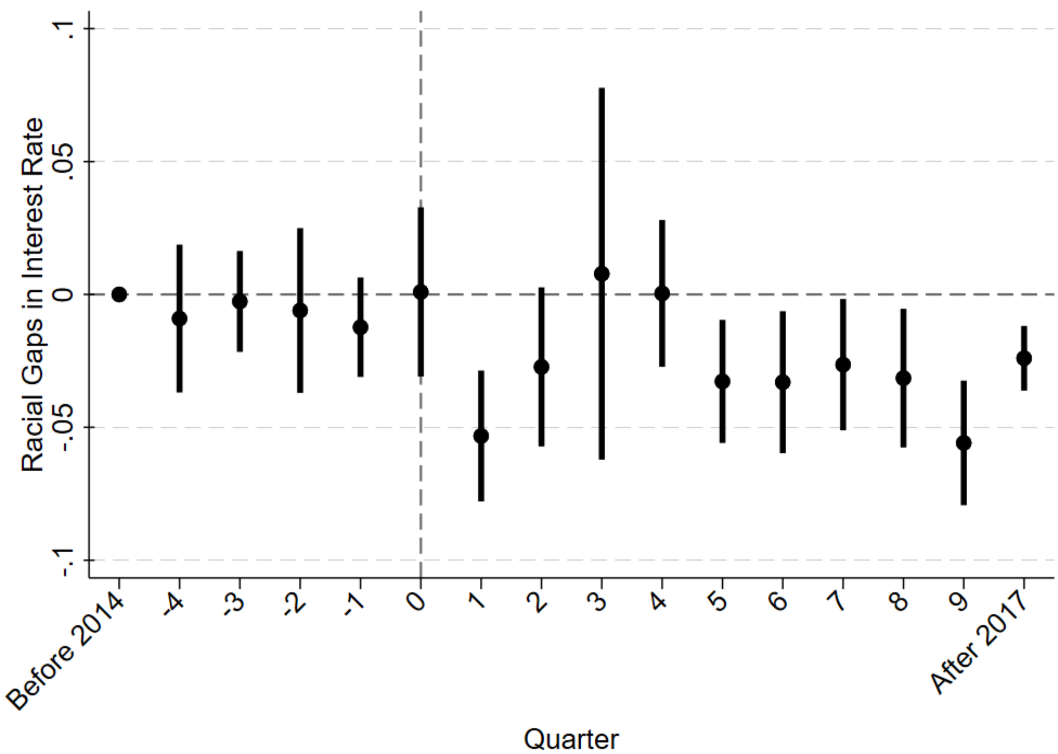


Figure 4. The Trend of Racial Disparities in Rejection Rates between CFPB-supervised and CFPB-unsupervised Banks

This figure shows the average excess rejection rates of minority borrowers relative to non-minority borrowers across both CFPB-supervised and CFPB-unsupervised banks from 2012 to 2019. The data for the diagram are drawn from the HMDA origination dataset. I utilize the residualized rejection rates after accounting for observable differences by applying lender-by-year fixed effects, lender-similar borrower fixed effects, lender-census tract fixed effects, and census tract-by-year fixed effects, as demonstrated in Equation (3). Furthermore, I utilize a 3-year moving average for a more accurate depiction of the actual trend. The figure features a vertical dashed line symbolizing mid-2015, the point in time when the CFPB disclosed the complaint narratives.

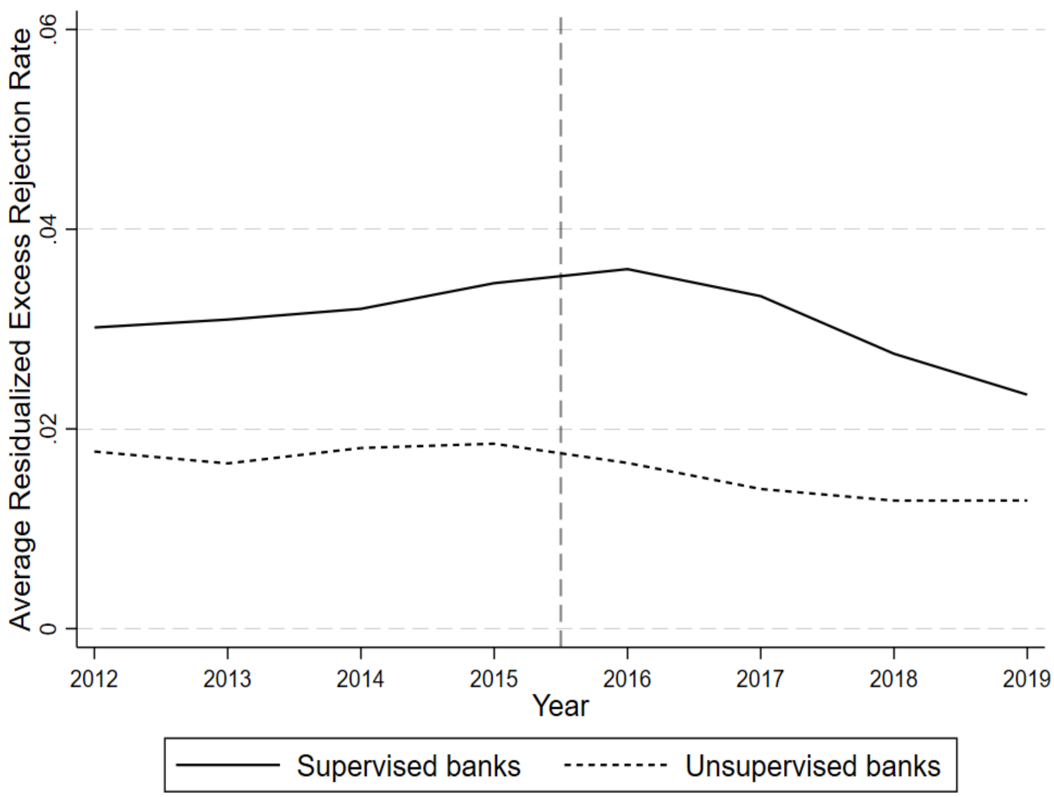


Figure 5. An Example of Rankings for Banks' Discrimination Based on Racial Disparities in Interest Rates

This figure displays discrimination in interest rates from the Discrimination Warrior website pertaining to zip code 02135, which is a region in Boston. Visitors to the site can select either interest rates or rejection rates and enter the relevant zip code. Upon entering, the website sequentially presents results showing the extent of discrimination against minorities by banks, ranked in order of severity.

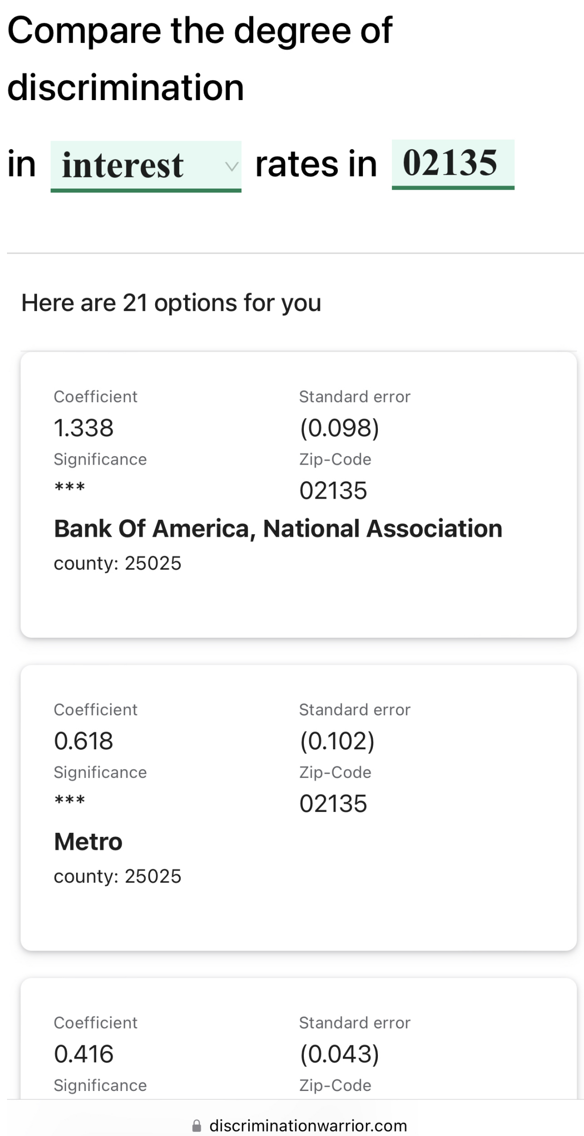


Figure 6. An Example of Rankings for Banks' Discrimination Based on Racial Disparities in Rejection Rates

This figure displays discrimination in rejection rates from the Discrimination Warrior website pertaining to zip code 02135, which is a region in Boston. Visitors to the site can select either interest rates or rejection rates and enter the relevant zip code. Upon entering, the website sequentially presents results showing the extent of discrimination against minorities by banks, ranked in order of severity.

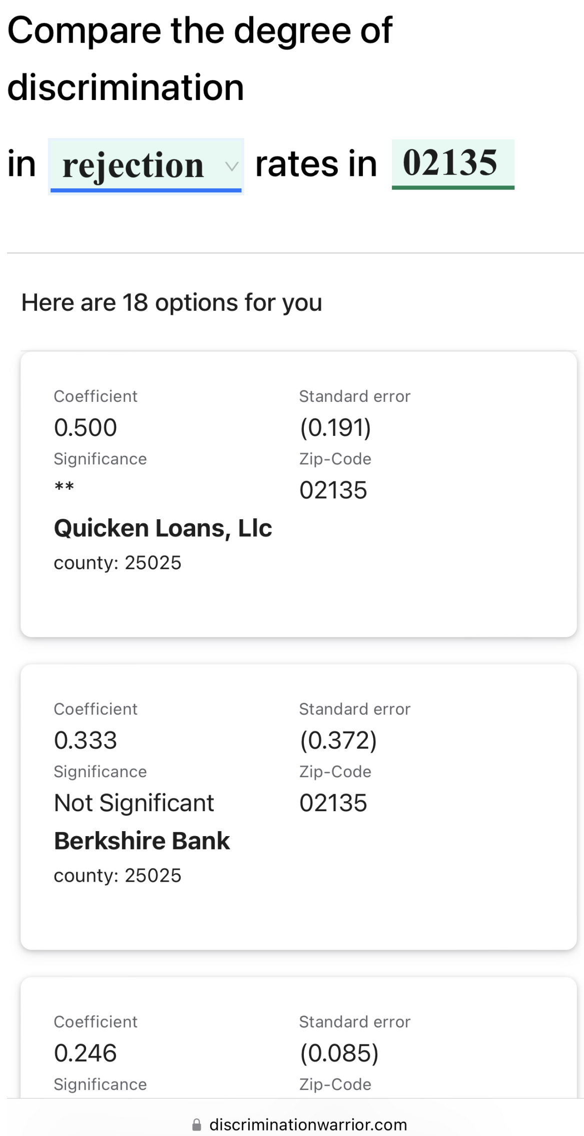


Table 1. Discrimination Complaint Example

This table displays a consumer complaint example concerning "Applying for a mortgage or refinancing an existing mortgage," as publicly disseminated by the CFPB. This includes components such as "Date received", "Product", "Consumer complaint narrative", "Company", and "Company response to consumer", among others.

Date received	2021/1/28
Product	Mortgage
Subproduct	Conventional home mortgage
Issue	Applying for a mortgage or refinancing an existing mortgage
Consumer complaint narrative	I was denied a mortgage loan from Bank of America for a property in XXXX XXXX, NJ on XX/XX/2021. I haven't received written confirmation yet, but the verbal reasoning is due to my employment history and employment gaps. The loan officer sounded very condescending when she told me that I was denied. It doesn't make sense to me to be denied for that reason alone as my employment history was stated on Day 1 and I was pre-qualified for the loan. To make matters worse, I was denied after having an appraisal done on the property so I was fairly far into the process with a refund unlikely for the $570.00 I was charged for the appraisal. I believe that I am being discriminated against because I disclosed my race as XXXX on Section X of the XXXX loan application. I would greatly appreciate it if this could be looked into to ensure that Bank of America didn't discriminate against me by showing that they also denied mortgage loans to people of other races, particularly XXXX people, with similar credit, income or debt-to-income ratio, savings, educational, and employment backgrounds as me. Quick summary of my background: I have excellent credit, my credit score is over XXXX. My 2 employment gaps greater than 30 days were related to school. I have a XXXX XXXX XXXX XXXX and currently in XXXX XXXX seeking a XXXX. I work full time as a mortgage loan advisor where I earn over $45000.00 annually. I have savings of $30000.00. The house I was looking to purchase cost $180000.00.
Company	BANK OF AMERICA, NATIONAL ASSOCIATION
State	PA
Submitted via	Web
Company response to consumer	Closed with monetary relief
Company disputed	No

Table 2. Summary Statistics of the HMDA-GSE Merged Dataset

This table presents summary statistics for the key variables derived from the HMDA-GSE merged dataset, which I use to identify the effects of the disclosure policy on interest rates and default rates when consumers apply for mortgage loans. The treated banks are those with total assets of 10 to 100 billion prior to the policy's implementation, while the control banks have total assets ranging from 0 to 10 billion. The variables are averaged from 2011 to May 2015, the pre-disclosure period. The table provides unconditional means, standard deviations, and p-values for the differences in means between the treated and control groups before treatment. The interest rate variable undergoes winsorization at the 1% level. For the t-test, standard errors are clustered at the bank level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All(Treated&Control)			*Treated*		*Control*		*t − Test*
	Observations	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*p − Value*
Minority	340,320	0.082	0.274	0.067	0.250	0.084	0.278	0.466
Interest Rate	340,320	4.203	0.473	4.112	0.467	4.217	0.472	0.096
Default Rate	340,320	0.065	0.246	0.059	0.236	0.066	0.248	0.633
LTV	340,320	77.200	14.210	76.310	14.300	77.340	14.190	0.443
Credit Score	340,285	752.700	44.660	759.200	41.880	751.700	45.000	0.019
Gender	340,320	0.727	0.445	0.729	0.444	0.727	0.446	0.867
Income	338,509	100.900	74.740	107.100	88.200	99.910	72.350	0.117
Loan Amount	340,320	246.100	127.400	239.100	127.600	247.200	127.300	0.606
Co-applicant	340,320	0.534	0.499	0.559	0.496	0.530	0.499	0.236
Cash-out Refinance	340,320	0.181	0.385	0.163	0.370	0.184	0.388	0.390
Purchase	340,320	0.492	0.500	0.514	0.500	0.489	0.500	0.712
Refinance	340,320	0.326	0.469	0.323	0.467	0.327	0.469	0.931

Table 3. The Impact of Disclosure on Interest Rate

This table reports results on the effect of complaint narrative disclosure on the racial gaps in interest rates. The data come from the HMDA-GSE merged dataset spanning 2011-2019. The dependent variable is the interest rate on originated fixed-rate mortgages. Each column provides estimated coefficients on four independent variables, with other coefficient groups (*Treat*, *Post*, and their interaction term *Treat×Post*) being absorbed by fixed effects due to collinearity. The key estimated coefficient, *Treat×Post×Minority*, denotes the effect of disclosure on racial gaps in interest rates. In column (1), fixed effects for the interaction of lender and time, and pricing grid-loan purpose-time joint fixed effects are controlled. Column (2) incorporates controls for borrower characteristics. Column (3) combines the two fixed effects from column (1) into lender-pricing grid-loan purpose-time joint fixed effects. Column (4), building on column (3), replaces borrower characteristics with the more stringent similar borrower fixed effects. Column (5) introduces joint fixed effects between lenders and similar borrowers, based on column (2). Standard errors, shown in parentheses, are clustered at the lender and year level. The symbols *, **, and *** correspond to significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
Dep. Var =	Interest Rate				
Treat×Post×Minority	-0.023***	-0.023***	-0.023***	-0.023***	-0.028**
	(0.006)	(0.006)	(0.005)	(0.005)	(0.009)
Post×Minority	-0.001	0.002	-0.001	-0.001	0.001
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)
Treat×Minority	-0.002	-0.009*	-0.004	-0.006	0.000
	(0.004)	(0.005)	(0.005)	(0.004)	(0.006)
Minority	0.028***	0.025***	0.028***	0.029***	0.028***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
Lender×Year-Month FE	YES	YES			YES
Bucket×Loan Purpose ×Year-Month FE	YES	YES			YES
Lender×Bucket×Loan-Purpose×Year-Month FE			YES	YES	
Borrower Characteristics FE		YES	YES		
Similar Borrower FE				YES	
Lender×Similar Borrower FE					YES
Observations	1,305,738	1,273,848	1,031,982	1,023,277	803,991
R-squared	0.756	0.772	0.799	0.808	0.816

Table 4. Moderation Effect of the Impact of Disclosure on Interest Rates

This table presents findings on the moderating effect of disclosure on racial disparities in interest rates, with data drawn from the HMDA-GSE merged dataset spanning the years 2011-2019. The dependent variable is the interest rate on originated fixed-rate mortgages. Panel A illustrates heterogeneous outcomes related to discrimination complaints and Google search attention by representing the contrasting outcomes between two subgroups of the treated group, in comparison to the same control group. The treated group in columns (1) and (2) is categorized based on the incidence of discrimination complaints, whereas the treated group in columns (3) and (4) is categorized based on Google search attention. The estimated coefficient on *Treat×Post×Minority* identifies the treatment effect in the subsample analysis. All four columns in panel A control for joint fixed effects of lender-pricing grid-loan purpose-time as well as borrower characteristics fixed effects. Panel B employs the continuous variable *Litigation* as the moderating variable, with the first two columns discussing the number of litigations and the last two discussing the penalty amounts (in one million dollars). Columns (1) and (3) control for the same fixed effects as in panel A, whereas columns (2) and (4) replace borrower characteristics with more stringent similar borrower fixed effects. Standard errors in this table, indicated in parentheses, are clustered at the lender and year level. The symbols *, **, and *** correspond to significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Discrimination Complaint and Search Attention

	(1)	(2)	(3)	(4)
	Complaint		Google Search Index	
	Yes	No	High	Low
Dep. Var =	Interest Rate			
Treat×Post×Minority	-0.028^{**}	-0.002	-0.036^{***}	-0.007
	(0.011)	(0.017)	(0.004)	(0.007)
Lender×Bucket×Loan Purpose×Year-Month FE	YES	YES	YES	YES
Borrower Characteristics FE	YES	YES	YES	YES
Observations	987,148	984,186	983,219	977,896
R-squared	0.796	0.796	0.797	0.796

Panel B. Litigation Risk				
	(1)	(2)	(3)	(4)
	Number		Penalty	
Dep. Var =	Interest Rate			
Treat×Post×Minority×Litigation	-0.059**	-0.052*	-0.014***	-0.012**
	(0.020)	(0.026)	(0.003)	(0.004)
Lender×Bucket×Loan Purpose×Year-Month FE	YES	YES	YES	YES
Borrower Characteristics FE	YES		YES	
Similar Borrower FE		YES		YES
Observations	972,084	963,400	972,084	963,400
R-squared	0.795	0.805	0.795	0.805

Table 5. Impact of Disclosure on Default Rates of Subprime Minority Borrowers

This table reports results on the effect of complaint narrative disclosure on the racial gaps in default rates. The data come from the HMDA-GSE merged dataset spanning 2011-2019. The dependent variable denotes whether the loan falls into the delinquency of 90 or more days. Columns (1)-(3) limit the sample to those with credit scores under 660, showcasing the results concerning subprime minority borrowers, while column (4) outlines results from the entire sample. Each column provides estimated coefficients on four independent variables, with other coefficient groups (*Treat*, *Post*, and their interaction term *Treat*×*Post*) being absorbed by fixed effects due to collinearity. The key estimated coefficient, *Treat*×*Post*×*Minority*, denotes the effect of disclosure on racial gaps in default rates. Column (1) solely controls for lender-time (year-month) fixed effects and zip-by-year fixed effects, and column (2) incorporates controls for loan traits such as interest rate, loan purpose, and the logarithm of the loan amount, applicant characteristics like gender, the logarithm of income, and the presence of a co-applicant. Column (3) extends column (2) by including pricing grid-loan purpose fixed effects. The model specification in column (4) is consistent with column (2). Standard errors, shown in parentheses, are clustered at the lender and year level. The symbols *, **, and *** correspond to significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)
	Marginal Borrowers			Full
Dep. Var =	Default rate			
Treat×*Post*×*Minority*	0.304***	0.288***	0.286***	0.012
	(0.035)	(0.051)	(0.046)	(0.009)
Post×*Minority*	-0.012	-0.008	-0.010	0.009**
	(0.043)	(0.034)	(0.036)	(0.004)
Treat×*Minority*	-0.289***	-0.289***	-0.290***	-0.009
	(0.025)	(0.024)	(0.021)	(0.009)
Minority	0.029	0.026	0.027	0.018***
	(0.042)	(0.034)	(0.036)	(0.003)
Control		YES	YES	YES
Lender×Year-Month FE	YES	YES	YES	YES
Zip×Year FE	YES	YES	YES	YES
Bucket×Loan Purpose FE			YES	
Observations	25,759	24,187	24,185	1,243,580
R-square	0.491	0.500	0.502	0.149

Table 6. The Impact of Disclosure on Rejection Rates

This table reports results on the effect of complaint narrative disclosure on the racial gaps in rejection rates. The data come from the HMDA origination dataset spanning 2011-2019. The dependent variable is the loan-level rejection status. Each column provides estimated coefficients on four independent variables, with other coefficient groups (*Treat*, *Post*, and their interaction term *Treat*×*Post*) being absorbed by fixed effects due to collinearity. The key estimated coefficient, *Treat*×*Post*×*Minority*, denotes the effect of disclosure on racial gaps in rejection rates. The first pair of columns present the fundamental results, where I account for lender-by-year fixed effects, lender-similar borrower combined fixed effects, and lender-census tract combined fixed effects in column (1), whereas I also include census tract-by-year fixed effects in column (2). Columns (3) through (6) display the moderation effects of discrimination complaints and Google attention, based on the specification detailed in column (2). Standard errors, shown in parentheses, are clustered at the lender and year level. The symbols *, **, and *** correspond to significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline		Complaint		Google Search Index	
			Yes	No	High	Low
Dep. Var =	Rejection Rate					
Treat×*Post* ×*Minority*	-0.018*	-0.016*	-0.033***	-0.005	-0.019*	-0.013
	(0.009)	(0.008)	(0.010)	(0.005)	(0.009)	(0.008)
Post×*Minority*	-0.002	-0.003	-0.003	-0.003	-0.003	-0.004
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Treat×*Minority*	0.043***	0.042***	0.066***	0.020***	0.046***	0.038***
	(0.009)	(0.008)	(0.008)	(0.006)	(0.010)	(0.008)
Minority	0.034***	0.035***	0.035***	0.035***	0.035***	0.035***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Lender×Year FE	YES	YES	YES	YES	YES	YES
Lender×Similar-Borrower FE	YES	YES	YES	YES	YES	YES
Lender×Census-Tract FE	YES	YES	YES	YES	YES	YES
Census Tract×Year FE		YES	YES	YES	YES	YES
Observations	17,383,935	17,377,535	15,822,650	15,951,820	15,853,196	15,585,715
R-squared	0.459	0.477	0.481	0.482	0.484	0.486

Table 7. Consumer Reaction to the Disclosure

This table presents the impact of complaint narrative disclosure on the application behavior of consumers, drawing data from the HMDA origination dataset for the period 2011 to 2019 and organizing it at the county-lender-minority-year level. The first dependent variable, referred to as *Application Number*, is the log of the number of mortgage applications from minority or White borrowers within at a specific county-lender-year level, while the second, known as *Application Amount*, is the log of the dollar value of mortgage applications from such borrowers within the same context. The table highlights the estimated results for two main independent variables, *1(Complaint)×Post×Minority* and *1(Non Complaint)×Post×Minority*, both of which are triple-difference estimated coefficients. In creating these variables, I divide CFPB-supervised banks into two groups based on discrimination-related complaints during 2015-2019 and then multiply each with the product of *Post* and *Minority*, resulting in two triple terms. The estimated coefficient for *1(Complaint)×Post×Minority* indicates racial gaps in banks with discriminatory practices under CFPB supervision, compared to the control group of CFPB-unsupervised banks, as influenced by disclosure. Similarly, *1(Non Complaint)×Post×Minority* shows the changes in racial gaps in CFPB-supervised banks without such discriminatory complaints. Across all columns, the considerations include (i) the percentage of mortgages that a lender approves in a specific county; (ii) an indicator variable valued at one, denoting the bank's physical presence in the county for that year; and (iii) the logarithm of the total deposits collected by the bank's branches within the county during that year. Additional controls include lender-by-year and county-by-year fixed effects in columns (1) and (3), with the introduction of lender-county fixed effects in columns (2) and (4). Standard errors, shown in parentheses, are clustered at the county and year level. The symbols *, **, and *** correspond to significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)
Dep. Var =	Application Number		Application Amount	
1(Complaint)×Post×Minority	-0.251**	-0.245*	-0.238**	-0.230**
	(0.082)	(0.115)	(0.082)	(0.084)
1(Non Complaint)×Post×Minority	-0.028	-0.019	-0.072*	-0.068*
	(0.031)	(0.090)	(0.032)	(0.036)
P-value for Equality Test	0.011	0.256	0.037	0.065
Control	YES	YES	YES	YES
Lender×Year FE	YES	YES	YES	YES
Lender×County FE	YES	YES	YES	YES
County×Year FE		YES		YES
Observations	1,307,477	1,307,285	1,307,477	1,307,285
R-squared	0.824	0.832	0.817	0.826

Table 8. Rival Bank's Reaction to the Disclosure

This table reports results on the effect of complaint narrative disclosure on the entry of rival banks. The data, sourced and consolidated from the HMDA origination dataset for the years 2011 to 2019, are aggregated at the county-year level. The dependent variable, *BankDensity*, is calculated as the number of unique bank brands per ten thousand individuals at the county level. The table showcases the estimated results for the *Complaint Share*$\times$*Post* independent variable, with other coefficient groups (*Complaint Share* and *Post*) being subsumed by fixed effects due to collinearity. This coefficient identifies changes in the unique bank brand number per ten thousand people in counties affected by the disclosure. I employ the measure *Complaint Share*, defined as the total number of discrimination-related complaints received per county in 2015 divided by the number of loans issued to minority applicants, to assess the extent of exposure to the effects of disclosure. In all columns, I account for the growth rate of income per capita and unemployment rates and control for county-level fixed effects and state-by-year joint fixed effects. Columns (1) to (3) display variations in the quantity of unique banks per ten thousand individuals in the present period (t), the next year (t+1), and two years later (t+2). Standard errors, shown in parentheses, are clustered at the county level. The symbols *, **, and *** correspond to significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)
	T	T+1	T+2
Dep. Var =		Bank Density	
Complaint Share$\times$*Post*	12.291***	14.062***	13.873***
	(3.481)	(3.702)	(3.968)
Control	YES	YES	YES
State$\times$Year FE	YES	YES	YES
County FE	YES	YES	YES
Observations	21,004	17,547	15,469
R-squared	0.756	0.780	0.790

Table 9. Market Reaction to the Disclosure

This table reports results on the effect of complaint narrative disclosure on the behavior of investors in the stock market. The data for analysis come from all listed banks within the HMDA origination dataset. Using the U.S. Daily Event Study function offered by WRDS, I estimate their cumulative abnormal return (CAR) based on the disclosure date of June 25, 2015. The dependent variables represent the CAR calculated based on the Capital Asset Pricing Model (CAPM) and the Market Adjusted Model (referred to as *Market*), during the three event windows of [-1, +1], [-2, +2], and [-3, +3] trading days. The table primarily showcases two independent variables, namely a dummy representing treated banks that have received discrimination complaints in 2015, labeled as *Complaint*, and a dummy symbolizing treated banks that have not received discrimination complaints in 2015, termed *Non-Complaint*. In all columns, I control for the banks' total assets, equity-to-assets ratio, return on assets, and total deposits, and execute a t-test to compare the coefficients of two primary independent variables. The robust standard errors are reported in parentheses. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Dep.Var =	CAR (CAPM)			CAR (Market)		
	[-1, +1]	[-2, +2]	[-3, +3]	[-1, +1]	[-2, +2]	[-3, +3]
Complaint	-0.013**	-0.026***	-0.018*	-0.013**	-0.026***	-0.019*
	(0.006)	(0.009)	(0.010)	(0.006)	(0.009)	(0.010)
Non-Complaint	-0.004	-0.015**	-0.013	-0.004	-0.015**	-0.013
	(0.004)	(0.007)	(0.011)	(0.004)	(0.007)	(0.011)
P-value for Equality Test	0.041	0.183	0.550	0.040	0.184	0.506
Control	YES	YES	YES	YES	YES	YES
Observations	341	341	341	341	341	341
R-squared	0.033	0.095	0.027	0.033	0.101	0.031

Table 10. The Impact of Disclosure on Homeownership

This table reports results on the real effect of complaint narrative disclosure on homeownership. The data, sourced and consolidated from the HMDA origination dataset for the years 2011 to 2019, are aggregated at the county-year level. The dependent variable, *Homeownership Ratio*, is extracted from the ACS datasets and represents housing homeownership across different races at the county-year level. The table showcases the estimated results for the *Complaint Share*$\times$*Post* independent variable, with other coefficient groups (*Complaint Share* and *Post*) being subsumed by fixed effects due to collinearity. This coefficient identifies changes in the homeownership of different races and racial gaps in homeownership in counties affected by the disclosure. I employ the measure *Complaint Share*, defined as the total number of discrimination-related complaints received per county in 2015 divided by the number of loans issued to minority applicants, to assess the extent of exposure to the effects of disclosure. In all columns, I account for the growth rate of income per capita, unemployment rates, and logarithm of population, and control for county-level fixed effects and state-by-year joint fixed effects. Columns (1) to (3) display variations in (i) homeownership among minority residents, specifically Black and Hispanic residents; (ii) homeownership among White residents; and (iii) the difference between the two indicators. Standard errors, shown in parentheses, are clustered at the county level. The symbols *, **, and *** correspond to significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)
	Minority	White	Diff
Dep.Var =	Homeownership Ratio		
Complaint Share$\times$*Post*	0.545**	0.070	0.489**
	(0.257)	(0.073)	(0.245)
Control	YES	YES	YES
State$\times$Year FE	YES	YES	YES
County FE	YES	YES	YES
Observations	4,523	4,523	4,523
R-squared	0.852	0.959	0.806

Internet Appendix for
Does the Disclosure of Consumer Complaints Reduce Racial Disparities in the Mortgage Lending Market?

Contents

Appendix A. Supplementary Figures and Tables

Figure A1. Dynamic Trend of the Impact of Disclosure on Default Rates

This figure represents the coefficient estimates for racial disparities in default rates among treated lenders relative to control lenders. The illustration contains outcomes from five quarters before the second quarter of 2015 - the starting point of the disclosure policy - and continues through eleven quarters after its application. Samples outside of this chosen span are grouped into the initial and final periods. During the estimation, the fifth quarter prior to the disclosure and earlier periods are used as the baseline period. For brevity, I combine three quarters into one bin. The data come from the HMDA-GSE merged dataset spanning 2011-2019. The Y-axis reflects the magnitude of the racial disparities in default rates, while the X-axis signifies varying periods, with period zero corresponding to the disclosure policy's implementation. The regression specification used to generate this figure is consistent with that of column (2) in Table 5. Black circles symbolize the coefficient estimates for different periods, and the black vertical lines denote the 95% confidence intervals.

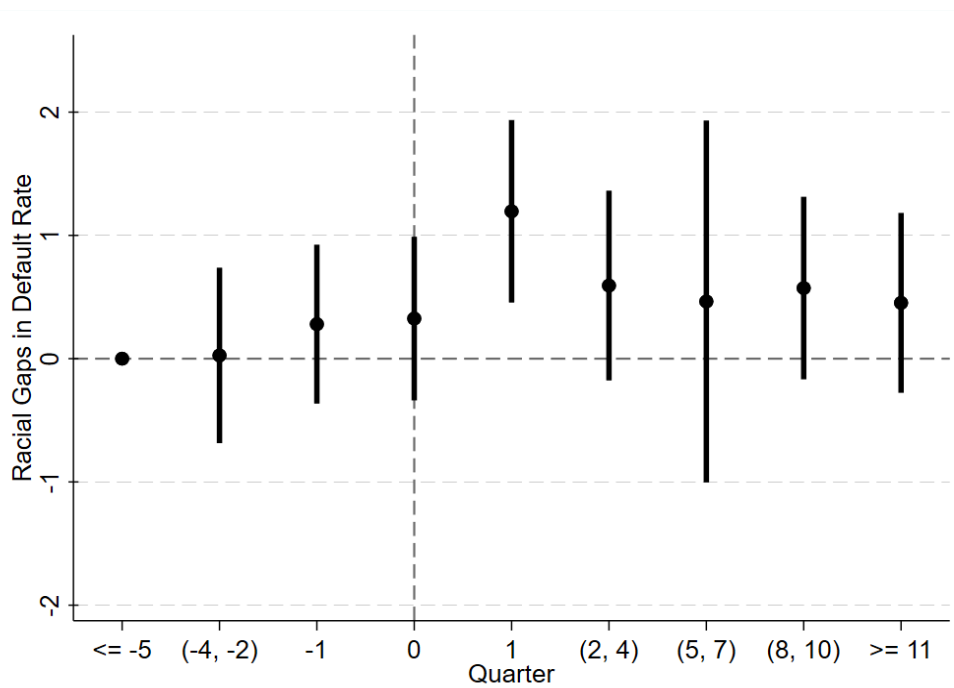


Figure A2. Dynamic Trend of the Impact of Disclosure on Rejection Rate

This figure represents the yearly coefficient estimates for racial disparities in rejection rates among treated lenders relative to control lenders. The illustration contains outcomes from three years before 2015 - the starting year of the disclosure policy - and continues through three years after its application, that is, between 2012 and 2018. Samples outside of this chosen span are grouped into the initial and final periods. During the estimation, the first year prior to the disclosure is used as the baseline period. The data come from the HMDA-GSE merged dataset spanning 2011-2019. The Y-axis reflects the magnitude of the racial disparities in rejection rates, while the X-axis signifies varying periods, with period zero corresponding to the disclosure policy's implementation. The regression specification used to generate this figure is consistent with that of column (2) in Table 6. Black circles symbolize the coefficient estimates for different periods, and the black vertical lines denote the 95% confidence intervals.

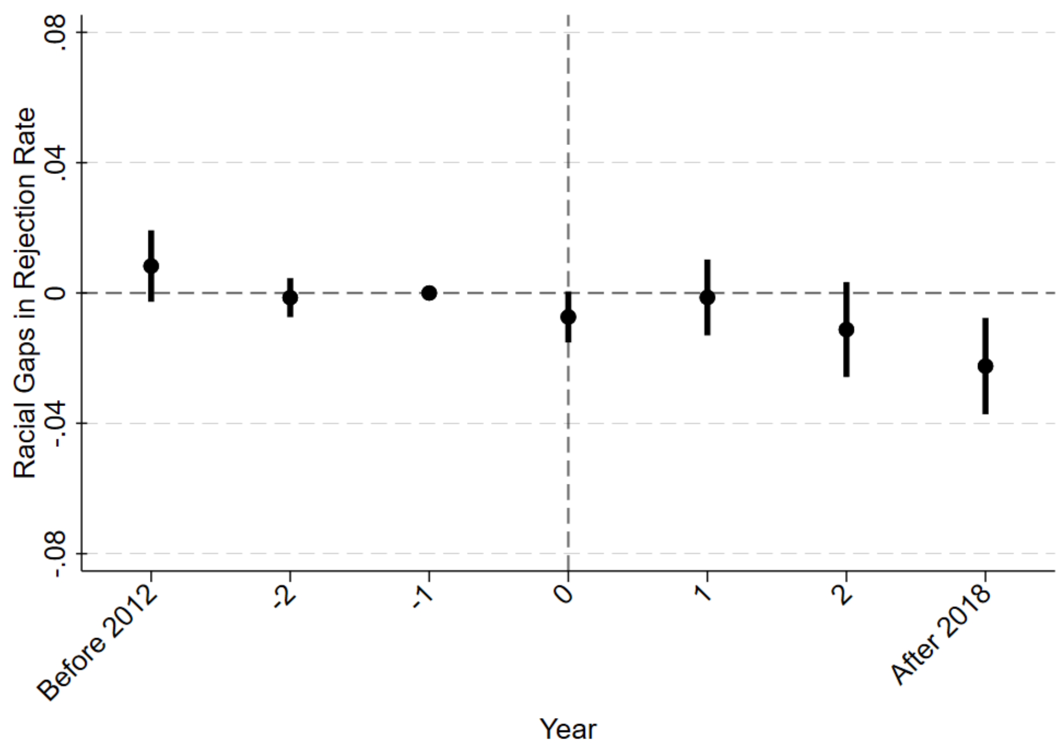


Table A1. Variable Definitions

Variable	Definition
HMDA LARS 2011-2019	
Rejection	An indicator variable indicating whether a loan was approved at application, marked as 1 for approval, and 0 otherwise.
Application Number	A continuous variable indicating the number of applications received by a ZIP code in different years.
Minority	An indicator variable indicating if the loan applicant is a minority, marked as 1 for Latinx or African American, and 0 for Asian or White.
Gender	An indicator variable indicating if the loan applicant is a man, marked as 1 for male, and 0 for female.
Loan Amount	A continuous variable indicating the amount of a particular loan.
Income	A continuous variable indicating the income level of the loan applicant.
Co-applicant	An indicator variable indicating whether the loan has a co-applicant or not.
Census Tract	A categorical variable indicating the 11-digit code of the Census Tract of a loan.
Zip Code	A categorical variable indicating the 5-digit zip code of the bank's location.
County	A categorical variable indicating the 5-digit county code of the bank's location.
State	A categorical variable indicating the 2-digit state code where the bank is situated.
GSE 2011-2019	
Interest Rate	A continuous variable indicating the interest rate of a loan

Continued on next page

Table A1 – *Continued from previous page*

Variable	Definition
Default	An indicator variable represents if a loan has not been repaid within a designated term. It is marked as 1 if the duration of default exceeds 90 days, and 0 otherwise.
LTV Group	An indicator variable categorizing the continuous Loan-To-Value (LTV) into eight divisions (effectively seven). This classification follows the method of Bartlett et al. (2022).
Credit Score Group	An indicator variable categorizing the continuous Credit Score into eight sections. The categorization is in line with Bartlett et al. (2022).
Loan Purpose	An indicator variable indicating the category of a loan, with the primary types being Cash-out Refinance, Purchase, and Refinance in this study.
Other Sources	
Complaints Number	A continuous variable indicating the number of complaints a bank received in 2015, with data gathered from the CFPB dataset.
Penalty	A continuous variable indicating the amounts of fines related to discrimination lawsuits a bank face in a given year, with data gathered from the CFPB website.
Google Index Diff	A continuous variable indicating the increase in the Google Index for a state in 2015 relative to 2014, where Google Index refers to the search volume for the CFPB.
Population	A continuous variable indicating the population of a county in different years, sourced from the Census dataset.
Total Deposits	A continuous variable measuring the total deposits of a bank in different years, sourced from SOD.
Total Assets	A continuous variable indicating the total assets of a bank sourced from WRDS.

Continued on next page

Table A1 – *Continued from previous page*

Variable	**Definition**
Equity	A continuous variable indicating the equity of a listed bank sourced from WRDS.
Net Profit	A continuous variable indicating the net profit of a listed bank sourced from WRDS.
CAR (CAPM)	A continuous variable indicating the cumulative abnormal returns of a listed bank, sourced from WRDS.
CAR (Market)	A continuous variable indicating the market-adjusted returns of a listed bank, sourced from WRDS.

Table A2. Summary Statistics for the Main Datasets

This table presents summary statistics for the key variables drawn from the main datasets utilized in this paper. Panel A shows the summary statistics results for the HMDA-GSE merged dataset, while Panel B provides the summary statistics for the HMDA origination dataset. These datasets enable me to identify the effects of the disclosure policy on interest rates, default rates, and rejection rates, along with other extended analyses. Within both panels, I report the number of observations, mean, standard deviation, minimum, and maximum values for the main variables. The time frame for these samples spans from 2011 to 2019, with a restriction that banks in the samples have total assets of less than 100 billion dollars, consistent with the treatment in the main text.

Panel A: HMDA-GSE Merged Dataset

	(1)	(2)	(3)	(4)	(5)
Variables	N	Mean	SD	Min	Max
Minority	1,330,621	0.111	0.314	0	1
Interest Rate	1,330,621	4.299	0.499	3.375	5.625
Default Rate	1,330,621	0.081	0.273	0	1
LTV	1,330,621	77.855	14.414	30	95
Credit Score	1,330,482	747.941	45.054	634	816
Gender	1,330,621	0.679	0.467	0	1
Income	1,297,045	101.048	69.420	0	5,602
Loan Amount	1,330,621	259.205	128.733	40	1,375
Co-applicant	1,330,621	0.478	0.500	0	1
Cash-out Refinance	1,330,621	0.197	0.398	0	1
Purchase	1,330,621	0.569	0.495	0	1
Refinance	1,330,621	0.234	0.423	0	1

Panel B: HMDA Origination Dataset

	(1)	(2)	(3)	(4)	(5)
Variables	N	Mean	SD	Min	Max
LTV (Census Tract)	31,634,238	78.348	9.935	42	95
Credit Score (Census Tract)	31,634,238	754.214	31.365	659	810
Gender	31,634,238	0.698	0.459	0	1
Income	31,066,586	108.162	80.602	16	539
Loan Amount	31,632,905	250.134	159.282	29	1,000
Co-applicant	31,634,238	0.498	0.500	0	1
Purchase	31,634,238	0.427	0.495	0	1
Refinance	31,634,238	0.573	0.495	0	1

Table A3. The Impact of Disclosure on Interest Rate (Clustered at the Lender Level)

This table reports results on the effect of complaint narrative disclosure on the racial gaps in interest rates, with standard errors clustered at the level of the lender. The data come from the HMDA-GSE merged dataset spanning 2011-2019. The dependent variable is the interest rate on originated fixed-rate mortgages. Each column provides estimated coefficients on four independent variables, with other coefficient groups (*Treat*, *post*, and their interaction term *Treat*×*Post*) being absorbed by fixed effects due to collinearity. The key estimated coefficient, *Treat*×*Post*×*Minority*, denotes the effect of disclosure on racial gaps in interest rates. In column (1), fixed effects for the interaction of lender and time, and pricing grid-loan purpose-time joint fixed effects are controlled. Column (2) incorporates controls for borrower characteristics. Column (3) combines the two fixed effects from column (1) into lender-pricing grid-loan purpose-time joint fixed effects. Column (4), building on column (3), replaces borrower characteristics with the more stringent similar borrower fixed effects. Column (5) introduces joint fixed effects between lenders and similar borrowers, based on column (2). Standard errors, shown in parentheses, are clustered at the lender level. The symbols *, **, and *** correspond to significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
Dep. Var =	Interest Rate				
Treat×*Post*×*Minority*	-0.023**	-0.023**	-0.023**	-0.023**	-0.028**
	(0.010)	(0.011)	(0.011)	(0.012)	(0.012)
Post×*Minority*	-0.001	0.002	-0.001	-0.001	0.001
	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
Treat×*Minority*	-0.002	-0.009	-0.004	-0.006	-0.001
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Minority	0.028***	0.025***	0.028***	0.029***	0.028***
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
Lender×Year-Month FE	YES	YES			YES
Bucket×Loan-Purpose×Year-Month FE	YES	YES			YES
Lender×Bucket×Loan-Purpose×Year-Month FE			YES	YES	
Borrower Characteristics FE		YES	YES		
Similar Borrower FE				YES	
Lender×Similar Borrower FE					YES
Observations	1,305,738	1,273,848	1,031,982	1,023,277	803,991
R-squared	0.756	0.772	0.799	0.808	0.816

Table A4. Replication - Interest Rate Differentials

This table replicates Table 3 in Bartlett et al. (2022). The utilized data are sourced from the HMDA-GSE merged dataset for the years 2011-2019, with the same model specification as in Bartlett et al. (2022). The dependent variable is the interest rate associated with originated fixed-rate mortgages. The independent variable is assigned a value of 1 if the borrower is Black or Latinx, and 0 in other cases. Standard errors, shown in parentheses, are clustered at the lender level. The symbols *, **, and *** correspond to significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)
	Purchase	Refinance
Dep. Var =	Interest Rate	
Minority Borrower	0.036***	0.014***
	(0.004)	(0.003)
Lender×Year-Month FE	YES	YES
Bucket×Loan Purpose×Year-Month FE	YES	YES
Amount decile FE	YES	YES
Observations	1,034,355	792,849
R-squared	0.755	0.762

Table A5. Racial Disparities in Discount Points

This table examines the difference in racial disparities concerning discount points between the treated and control groups. The data used are derived from the HMDA-GSE merged dataset for the years 2018-2019. The dependent variable is the discount points chosen by borrowers. The calculated outcome of *Treat×Minority* in the table tests whether minority borrowers in the treat and control groups exhibit different preferences for discount points. In column (1), fixed effects for the interaction of lender and time, and pricing grid-loan purpose-time joint fixed effects are controlled. Column (2) incorporates controls for borrower characteristics. Column (3) combines the two fixed effects from column (1) into lender-pricing grid-loan purpose-time joint fixed effects. Column (4), building on column (3), replaces borrower characteristics with the more stringent similar borrower fixed effects. Column (5) introduces joint fixed effects between lenders and similar borrowers, based on column (2). Standard errors, shown in parentheses, are clustered at the lender and year level. The symbols *, **, and *** correspond to significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
Dep. Var =	Discount Point				
Treat×Minority	-31.368	6.687	28.633	6.591	12.368
	(42.806)	(20.827)	(29.835)	(33.724)	(41.193)
Control	YES	YES	YES	YES	YES
Lender×Year-Month FE	YES	YES			YES
Bucket×Loan-Purpose×Year-Month FE	YES	YES			YES
Lender×Bucket×Loan-Purpose×Year-Month FE			YES	YES	
Borrower Characteristics FE		YES	YES		
Similar Borrower FE				YES	
Lender×Similar Borrower FE					YES
Observations	260,131	249,639	206,534	195,851	121,219
R-squared	0.360	0.447	0.517	0.575	0.630

Table A6. Summary of CFPB Litigation Penalty

Bank Name	Date Filed	Civil Money Penalty	Loan Subsidy Program Investment
Washington Federal	9/10/2013	$34,000	-
Mortgage Master, Inc.	9/10/2013	$425,000	-
National City Bank	23/12/2013	-	-
Provident Funding Associates, L.P.	28/5/2015	-	-
Hudson City Savings Bank, F.S.B.	24/9/2015	$5.5 million	$25 million
BancorpSouth Bank	29/6/2016	$3 million	$4 million
Nationstar Mortgage LLC	15/3/2017	$1.75 million	-
Freedom Mortgage Corporation	5/6/2019	$1.75 million	-
Townstone Financial, Inc. and Barry Sturner	15/7/2020	-	-

Table A7. Rival Bank Reaction to the Disclosure - Another *Share* Measurement

This table reports results on the effect of complaint narrative disclosure on the entry of rival banks, utilizing an independent variable that differs from the one used in Table 8. The data, sourced and consolidated from the HMDA origination dataset for the years 2011 to 2019, are aggregated at the county-year level. The dependent variable, *BankDensity*, is calculated as the number of unique bank brands per ten thousand individuals at the county level. The table showcases the estimated results for the $Amount\ Share \times Post$ independent variable, with other coefficient groups (*Amount Share* and *Post*) being subsumed by fixed effects due to collinearity. This coefficient identifies changes in the unique bank brand number per ten thousand people in counties affected by the disclosure. Unlike Table 8, I use the measure *Amount Share*, defined as the market share held by treated banks within each county in 2015, to evaluate the extent of exposure to the effects of disclosure. In all columns, I account for the growth rate of income per capita and unemployment rates and control for county-level fixed effects and state-by-year joint fixed effects. Based on the *Amount Share*, columns (1) to (3) display variations in the quantity of unique banks per ten thousand individuals in the present period (t), the next year (t+1), and two years later (t+2). Standard errors, shown in parentheses, are clustered at the county level. The symbols *, **, and *** correspond to significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)
	T	T+1	T+2
Dep. Var =		Bank Density	
$Amount\ Share \times Post$	6.151***	9.374***	10.890***
	(2.341)	(2.783)	(3.059)
Control	YES	YES	YES
State×Year FE	YES	YES	YES
County FE	YES	YES	YES
Observations	20,921	17,495	15,426
R-squared	0.757	0.780	0.791

Table A8. Financial Regulatory Authorities in Developed Countries

Country	Financial Regulatory Authority	Specifically Established for Consumer Protection	Formed Year	Receiving Complaints	Starting Year of Receiving Complaints	Complaints Disclosure
Australia	Australian Financial Complaints Authority	Yes	2018	Yes	2018	No
Austria	Austrian Financial Market Authority	No	2002	Yes	2002	No
Belgium	Financial Services and Markets Authority	Yes	2011	Yes	2011	No
Canada	The Financial Consumer Agency of Canada	Yes	2001	Yes	2001	No
Croatia	Croatian Financial Services Supervisory Agency	No	2005	Yes	2005	No
Czech Republic	Czech National Bank	No	1993	Yes	2021	No
Denmark	Danish Financial Complaint Board	Yes	2015	Yes	2015	No
Finland	Finnish Financial Ombudsman Bureau	Yes	2009	Yes	2009	No
France	Autorité de Contrôle Prudentiel et de Résolution	No	2010	No	-	-
Germany	Federal Financial Supervisory Authority	No	2002	Yes	2002	No
Hungary	Hungarian National Bank	No	1924	Yes	2020	No

Continued on next page

Table A8 – *Continued from previous page*

Country	Financial Regulatory Authority	Specifically Established for Consumer Protection	Formed Year	Receiving Complaints	Starting Year of Receiving Complaints	Complaints Disclosure
Iceland	Complaints Committee on Transactions with Financial Firms	Yes	2022	Yes	2022	No
Ireland	Financial Services and Pensions Ombudsman	Yes	2017	Yes	2018	No
Italy	Institute for the Supervision of Insurance	No	2013	Yes	2013	No
Japan	Supervision Bureau	No	2000	No	-	-
Korea	Financial Supervisory Service	No	1998	Yes	2011	No
Netherlands	Netherlands Authority for the Financial Markets	No	2002	Yes	2002	No
New Zealand	Financial Markets Authority	No	2011	Yes	2011	No
Norway	Norwegian Financial Services Complaints Board	Yes	2014	Yes	2014	No
Portugal	Bank of Portugal	No	1846	Yes	2020	No
Singapore	Financial Industry Disputes Resolution Centre	Yes	2005	Yes	2005	No
Slovakia	National Bank of Slovakia	No	1993	Yes	2020	No
Slovenia	Financial Consumer Protection Department	Yes	2015	Yes	2015	No

Continued on next page

Table A8 – *Continued from previous page*

Country	Financial Regulatory Authority	Specifically Established for Consumer Protection	Formed Year	Receiving Complaints	Starting Year of Receiving Complaints	Complaints Disclosure
Spain	Bank of Spain	No	1782	Yes	2004	No
Switzerland	Swiss Banking Ombudsman	Yes	1993	Yes	1993	No
United Kingdom	Financial Ombudsman Service	Yes	2001	Yes	2001	No
United States	**Consumer Financial Protection Bureau**	**Yes**	**2011**	**Yes**	**2011**	**Yes**

Appendix B. Clarifications on the Process of Merging Datasets

HMDA-GSE Merged Dataset: Linkage Process

My research utilized two datasets, namely the Home Mortgage Disclosure Act (HMDA) and Government-Sponsored Enterprises (GSE), which contain valuable information about loan origination. Unfortunately, there is currently no commonly available mapping file that connects these two datasets. Following Law and Mislang (2022), I adopt the "fuzzy data matching" methods to overcome this challenge. These methods leverage the substantial overlap in variables shared by both datasets. When the information in these overlapping fields demonstrates consistent patterns between the two datasets, my matching process becomes efficient.

To establish a connection between the HMDA and GSE datasets, I make use of the tools provided by the *Fedmatch* R package developed by Cohen et al. (2021). This specific matching approach involves separating each dataset based on geography and lender ID. Subsequently, I apply the matching algorithm at the loan level.

Regarding the initial step, the GSE dataset includes lender names, which I integrated into the HMDA dataset using the Robert Avery file. However, for small lenders whose names present challenges in recognition within the GSE dataset, I employed accurate name matching techniques based on Jaccard string similarity and manual proofreading. Once I successfully match the lender IDs, I discuss the loan matching process.

To enhance the accuracy of the matching process, I divide each dataset into smaller partitions based on lender ID and geographic information. However, it's important to note that there is no exact correspondence between the two datasets in terms of geography. I use 3-digit zip codes to correspond to HMDA's census tracts and counties. In the "grid" formed by lender ID and geographic information, I match each loan based on observable characteristics at the time of loan origination. Furthermore, it's worth mentioning that certain variables in the GSE dataset have undergone changes since 2018. As a result, I have several additional variables available for use as matching references, including interest rate, manufactured property status, and loan purpose.

HMDA-GSE Merged Dataset: Filtering Process

To omit loans not conforming to standard lending criteria, I implement filtering on this dataset according to Bartlett et al. (2022), resulting in a final sample of 1,869,345 mortgage loans. During the filtering process, I remove GSE loans with credit scores below 620 (10,310 observations), LTVs under 0.3 (76,779 observations), and above 0.95 (121,430 observations). I also remove loans with interest rates below 2.75% or higher than 8% (10,273 observations), loans amounting to less than $40,000 (14,207 observations), and loans whose term is not equal to 360 months (769,736 observations). Thus, the final sample includes 56.53% (1,056,807 observations) of purchase mortgages and 43.47% (812,538 observations) of refinance loans.

Please note that for the actual analysis, I utilize the subsample from this final collection, where total assets range from $0 to $100 billion to perform the primary analysis. This subsample consists of 1,330,621 loans, accounting for 71.18% of the total sample. This refined sample allows for a more accurate and comprehensive examination of the mortgage market, thereby aiding in comprehending borrower and lender attributes, loan conditions, as well as the impact of the disclosure on pricing and approval rates.

HMDA Origination Dataset: Variable Supplementation Process

Besides dictating the loan interest rate, the lender also possesses the authority to either endorse or dismiss a loan application entirely. After the 2008 mortgage crisis, numerous lenders have imposed their personalized, more rigorous, approval stipulations over and above those stipulated by the GSE. Hence, although a loan application may garner creditworthiness approval within the GSE underwriting system, the lender retains the privilege to refuse the application.

A significant caveat is that I lack loan-level data on several crucial underwriting variables, such as the borrower's credit score or LTV ratio, since rejected loans are never issued. These data constraints imply that, unlike the interest-rate analysis in the main paper, any analysis of rejection rates is invariably subject to the omitted variable problem addressed in the introduction part. I cannot be completely confident that any discrepancies in rejection rates are attributable to discrimination rather than differences in unobservable variables. Nonetheless, the HMDA data permit me to control for loan-level lender, year, borrower income, and loan amount. In addition, I employ proxies for these in my analysis using the median credit scores and LTVs of the census tract, which I estimate using the merged HMDA-GSE data. Controlling for these variables, along with census tract level median credit scores and LTVs, mitigates the omitted-variable problem (Bartlett et al. (2022)). I also apply the same filter to the loan amount and loan term that is used for variables in the HMDA-GSE merged dataset.[49]

Moving towards the matching process, I carry out it in three stages. In the first stage, I retain the census tract information, time information (i.e., year), LTV, and credit score information in the HMDA-GSE dataset. Consequently, I obtain the applicants' LTV and credit scores for multiple loan records per census tract per year. I then calculate the median of LTV and credit scores at the census tract level (Bartlett et al. (2022)). I maintain the two median variables produced here along with the census tract and year, and I can then easily aggregate the data to the census tract-year level.

Subsequently, I match the LTV and credit score indicators at the census tract-year level back to the original HMDA using the census tract and year. Since LTV and credit score

[49]Before 2018, the HMDA origination datasets do not include the loan term variable, so I employ the filter for loan term only in the years after 2018.

information primarily come from the GSE, and this database only includes records of loan applications that were not rejected, I cannot find corresponding values in the original HMDA dataset for about 21.65% of the samples in this step.

Finally, to discuss the impact of complaints disclosure in more detail, I match the number of bank complaints published by the CFPB in 2015 and the total bank assets values from the Call Report database for the first quarter of 2015 (Fuster et al. (2021)) into the aforementioned dataset using the arid (an identification number of banks) and year, to obtain the revised HMDA dataset to analyze the influence of disclosure on rejection. The final dataset comprises 42,198,230 loans, with 31,634,238 observations (accounting for 74.97% of the full sample) belonging to lenders whose total assets range from \$0 to \$100 billion.

References

Bartlett, Robert, Adair Morse, Richard Stanton, and Nancy Wallace, 2022, Consumer-lending discrimination in the fintech era, *Journal of Financial Economics* 143, 30–56.

Cohen, Gregory J, Jacob Dice, Melanie Friedrichs, Kamran Gupta, William Hayes, Isabel Kitschelt, Seung Jung Lee, W Blake Marsh, Nathan Mislang, Maya Shaton, et al., 2021, The us syndicated loan market: Matching data, *Journal of Financial Research* 44, 695–723.

Fuster, Andreas, Matthew C Plosser, and James I Vickery, 2021, Does cfpb oversight crimp credit?, Working paper, Ecole Polytechnique Fédérale de Lausanne.

Law, Kody, and Nathan Mislang, 2022, The heterogeneity of bank responses to the fintech challenge, Working paper, Cornell University.