Racial Disparities in the U.S. Mortgage Market: The Evidence of Privacy Legislation*

[PRELIMINARY AND INCOMPLETE. COMMENTS ARE WELCOME.]

Xiangyu Lin, Sarah S. Zhang†

This version: June 2024

Abstract

We assess the impact of privacy protection regulations on bank lending, focusing on racial disparities in mortgage lending decisions. Utilizing a Triple Difference (DDD) framework, our analysis reveals that after the enactment of privacy protection laws, interest rates for minority borrowers decreased by 14%, and rejection rates decreased by 18%. Specifically, these changes have increased reliance on hard information and decreased dependence on soft information, reducing discriminatory practices, particularly in areas with high minority density. Concurrently, we have found evidence that FinTechs, shadow banks, and API banks play a significant role in mitigating racial disparities in mortgage lending.

JEL Classification: G21, G28, G31

Keywords: Privacy Legislation; Mortgage lending; Diversity Equity and Inclusion; Fintech

* We would like to thank Ian Garrett, Stuart Hyde, Olga Kolokolova, Yifan Li, Chi-Yang Tsou, Mostafa Hussien, Zeppenfeld Jennifer (discussant) for helpful comments, and conference participants at Academy of Sustainable Finance, Accounting, Accountability & Governance (2024) for useful comments and suggestions. All errors are ours.
† Alliance Manchester Business School, The University of Manchester, Manchester M15 6PB, UK. Email address: xiangyu.lin@postgrad.manchester.ac.uk (X. Lin), sarah.zhang@manchester.ac.uk (S. S. Zhang)

1. Introduction

> "*Diversity, equity, and inclusion (DEI) in organizations and economies is increasingly important for sustainable growth, resilience and fairer economic outcomes.*
>
> *Embracing DEI is not just a moral imperative, but also a strategic one that promotes sustainable growth*"
>
> ——World Economic Forum 2024

The availability and quality of borrowers' personal data play a critical role in the results of mortgage applications, influencing the success or denial of a mortgage and the interest rates offered. While the digitalization of personal financial data has made data sharing easier, consumers have started expressing significant concern and reluctance to share their personal information, perceiving the current privacy-related environment negatively (Goldfarb and Tucker, 2012). Consumers frequently fear price discrimination, misuse of consumer data, and the potential for financial abuse (Chen et al., 2021; Lin, 2022). Assessing the balance between data privacy preferences and data sharing to assess credit risk has become a central issue for data privacy regulation.

While more data could help banks and financial service providers, such as FinTechs and shadow banks, provide more accurate decisions regarding mortgage decisions and assessment of the risk of defaults, discrimination remains prevalent in financial markets, including the mortgage sector.[3] In particular, disparities in mortgage approval rates and interest rates between minority and non-minority borrowers persist (Black et al.,1978; Courchane and Nickerson,1997; Munnell et al., 1996).

To alleviate consumer concerns about privacy protection and enhance data security oversight, several U.S. states have enacted legislation since 2018 aimed at safeguarding residents' rights. A notable example is the California Consumer Privacy Act (CCPA), which grants consumers proactive control over how companies use their personal information. The CCPA imposes strict regulations on companies to ensure the security and integrity of collected data. Additionally, it introduces new rights for California

[3] Research by Bartlett et al. (2022) reveals that Latinx/Black borrowers, given similar risks, pay substantially higher rates for GSE securitized and FHA-insured loans, particularly in communities with high minority populations. Similarly, Bayer et al. (2018) found that after controlling for credit scores and other risk factors, African American and Hispanic borrowers were 103% and 78% more likely, respectively, to receive high-cost mortgage loans.

residents, including the ability to access and delete personal information and the option to opt-out of selling personal information to third parties.

This study investigates whether privacy legislation reduces racial disparities in the U.S. mortgage market. Utilizing mortgage application data from 2018 to 2023, we construct three sample groups: neighboring covered counties (belonging to states that enacted privacy legislation) and uncovered counties (not belonging to states that enacted privacy legislation) combined with mortgage market application, rejection rates, and racial origin data. Our neighboring county sample comprises 8,764,155 mortgage application observations, while the adjacent state sample includes 31,588,968 observations. Additionally, for robustness tests, we include an extra sample of adjacent counties from states where privacy legislation took effect, totalling 3,148,677 mortgage application observations.

Our findings suggest that privacy legislation narrows racial disparities: specifically, it reduced minority interest rates by 2% and decreased the difference between the annual percentage rate (APR) and the average prime offer rate (APOR) by 14%. Furthermore, the legislation reduced the rejection rate for minority borrowers by 18%.

In our mechanism analysis, we posit that the rules for data transparency and security required by privacy legislation cause the loan adjudication process to rely more on hard information and standard credit models, reducing reliance on soft information, such as face-to-face interactions and observable applicant racial information (Bartlett et al., 2022). On the other hand, the strict requirements for a right to non-discrimination significantly reduce discriminatory practices in lending.

Additionally, our findings show that institutions other than traditional banks, such as Fintechs (Howell et al., 2022), shadow banks, and API banks (following the identification of API banks by Lin et al., 2024) have significantly contributed to reducing racial disparities in lending. These institutions often utilize advanced technologies, including machine learning and artificial intelligence, to assess creditworthiness and other loan-related metrics. This technology can process vast amounts of data more accurately without the biases that might affect human decision-makers. This can lead to more objective loan decisions. Moreover, API banks automate many processes, reducing human error and bias. Automation ensures that policies and decisions are uniformly applied across all applicants, which helps reduce discriminatory practices that might otherwise affect loan approval rates and terms for different demographic groups. These institutions typically offer greater transparency in their operations, building trust among their clients. Better understanding of how decisions are made and what factors are considered can also help clients manage their financial situations more effectively.

Additionally, our analysis of borrower characteristics, such as gender features and income levels, suggests that privacy legislation has limited mitigating effects on gender disparities; it appears to have no direct impact on women's mortgage interest rates but does reduce the mortgage rejection rate by 3%. However, the privacy regulations seem to have no substantial difference across different income levels of minority borrowers, featuring broader financial inclusion.

Our paper makes significant contributions to the literature on data security and privacy protection, providing new evidence for studies related to privacy legislation (Doerr et al., 2023; Gupta et al., 2023). It also offers new insights into racial disparities in the mortgage market, enriching the diversity, equity, and inclusion discourse.

The remainder of this paper is organized as follows: Section 2 provides background on privacy legislation and outlines our hypotheses. Section 3 describes the data and research methodology. Section 4 presents the main empirical results. In Section 5, we delve deeper into mechanism analysis. Section 6 studies the differing impacts of various types of lending institutions. Section 7 conducts additional analyses on lenders' characteristics such as gender and income levels. Finally, Section 8 summarizes the findings.

2. Related literature and hypothesis development

2.1 Institutional Background

Since the late 19th century, the United States has enacted several privacy laws to enhance the protection of personal information. In 1890, Warren and Brandeis published "The Right to Privacy" in the Harvard Law Review, marking the theoretical origin of privacy rights in America, advocating for legal protection of individuals' "right to privacy." Throughout the 1930s to the 1960s, legislation such as the Communications Act was established, followed by regulations addressing communication privacy, and as computers became ubiquitous, there was growing concern over data privacy.

The Civil Rights Act of 1964, a landmark anti-discrimination law, notably in its Title VII, prohibited discrimination in employment based on race, colour, religion, sex, or national origin. In 1974, the U.S. Congress passed the Privacy Act of 1974, a federal law aimed at regulating the collection and use of personal data by governmental agencies. In the digital era, the importance of privacy protection became paramount, and in 2003, California enacted the Online Privacy Protection Act, mandating websites to disclose their privacy policies, marking the first such state-level law.

American privacy laws have continually evolved, reflecting technological advancements and increasing demands for privacy protections. Since 2018, various U.S. states have introduced and enacted consumer privacy regulations. California led the way with the California Consumer Privacy Act (CCPA), which was voted on in early 2018 and signed into law on June 12 of the same year.

The CCPA, specifically designed to protect the privacy rights and consumer rights of California residents, targets "businesses" or "enterprises." It mandates data protection and privacy obligations for businesses that collect, process, or sell the personal information of California residents. These privacy laws emphasize three key aspects: enhancing information transparency, granting consumers control over their data, and ensuring data security and integrity.

Firstly, privacy laws enhance transparency regarding how businesses handle personal information. They inform consumers about the types of information collected, its purposes, and whether it will be sold or shared. Secondly, consumers have gained more control over their personal information, as emphasized in the CCPA[4] through rights such as the right to know, the right to delete, the right to opt-out, and the right to non-discrimination. Thirdly, these privacy laws impose strict regulations on businesses to ensure the security and integrity of data post-collection, requiring businesses to protect customer data and prevent data breaches/security incidents.

By the end of 2023, twelve states had signed privacy legislation, continuously addressing data breaches and privacy protection issues. These laws mandate companies to develop and implement

[4] Additionally, the CCPA applies to businesses meeting any of the following criteria: (1) enterprises with global annual revenues exceeding $25 million; (2) businesses that annually buy, receive, sell, or share the personal information of at least 50,000 California residents or households for commercial purposes; (3) entities that derive more than 50% of their annual revenues from personal information. Privacy regulations announced by different states have nuanced differences, with some states not allowing any form of exemption. The CCPA covers the vast majority of companies, including many small and medium-sized enterprises. See here: https://iapp.org/news/a/new-california-privacy-law-to-affect-more-than-half-a-million-us-companies

mandatory protective tools to safeguard private information, significantly strengthening national data security (Huang et al., 2024). Figure 1 shows the states that have enacted privacy legislation: darker blue indicates earlier enactment, while lighter blue represents more recent enactments. Neighboring states are indicated in grey.

[Insert Figure 1 here]

2.2 Hypothesis Development

The recent wave of privacy legislation mandates that companies significantly enhance their customer data protection capabilities. Data privacy has a substantial impact on economic outcomes; concealing specific information about borrowers can lead to inefficiency while sharing data can reduce information asymmetry and moral hazard (Posner, 1981). Laws such as the California Consumer Privacy Act (CCPA) require entities, including banks, to ensure the security and integrity of customer data. Consequently, firms face demands and motivations to protect consumer privacy and reduce discrimination, thereby minimizing default actions and penalties.

On one hand, compliance with these privacy regulations increases operational costs. As identified by Black et al. (1977), compliance regulations can raise the costs of housing finance, potentially prompting lending institutions to allocate loans more frequently to areas less affected by these regulations. The enhancement of consumer privacy rights escalates compliance costs, encompassing proactive defense and passive compliance mechanisms, and updating technological adoption to upgrade data security and protection systems. Huang et al. (2024) demonstrated significant investments in banking IT infrastructure following the introduction of privacy regulations, with associated legal expenses and fixed costs related to data handling also rising (Doerr et al., 2023). In the mortgage market, financial institutions may offset these increased compliance costs by raising prices and reducing loan supply (Gupta et al., 2023). Simultaneously, privacy laws encourage the use of data anonymization or pseudonymization techniques to reduce the risk of privacy breaches. In machine learning projects, these techniques enable researchers and developers to utilize data without exposing individual identities. However, complex machine learning models processing large volumes of data sometimes indirectly consider variables closely related to race, potentially leading to so-called "proxy discrimination"—

where ostensibly neutral variables indirectly incorporate racial bias (Fuster et al., 2020). Moreover, as many researchers have noted, even after controlling for borrower characteristics, minority applicants experience a significantly higher rejection rate than their white counterparts (Bayer et al., 2018; Munnell et al., 1996).

On the other hand, privacy legislation enhances user control over their data and increases transparency, and accountability, which can alleviate the trade-offs associated with privacy issues and increase the use of non-traditional data (Doerr et al., 2023). It also reduces observable biases, and new algorithmic decision-making processes in lending can minimize overt biases, such as face-to-face discrimination, making financial assessments more objective (Barocas and Selbst, 2016; Bartlett et al., 2022). This is increasingly beneficial for equity and diversity objectives. Importantly, state-level privacy regulations mandate that businesses cannot discriminate against consumers exercising their privacy rights, ensuring fairness in all user interactions (Gupta et al., 2023).

Based on the above discussion, we propose the following hypotheses:

H1a: Privacy legislation will increase the rejection and interest rates for minority groups.

H1b: Privacy legislation will decrease the rejection and interest rates for minority groups.

3. Data and Methodology

3.1 Sample Selection

To investigate the impact of U.S. privacy legislation on racial disparities in the mortgage market, we identified a clear exogenous shock—state-level privacy protection laws enacted between 2018 and 2023. We accessed comprehensive mortgage application data from 2018 to 2023 from the HMDA database, which included detailed information on borrowers, and the properties mortgaged (down to the census tract, county, and state levels).

3.1.1 Privacy Legislation

Since 2018, various U.S. states have enacted privacy protection laws, emphasizing the importance of protecting consumer personal data in the banking sector. We gathered information on consumer privacy legislation enacted in the U.S. since 2018 by the International Association of Privacy Professionals (IAPP).

In our empirical analysis, we divided mortgage applications into two groups—treatment and control—to assess the impact of consumer privacy legislation on racial disparities in the mortgage market. Applications in states that enacted consumer privacy legislation constituted our treatment group, whereas those in states without such legislation formed our control group. To ensure a clearer identification and distinction of the research question, we further restricted our sample to include only counties on the comparative borders of states with and without privacy protection agreements. This approach helps eliminate potential confounding variables by maintaining relatively constant economic environments between the treatment and control groups. Figure 2 illustrates this setup.

[Insert Figure 2 here]

3.1.2 Mortgage Lending

We obtained data on mortgage applications and issuances from the Home Mortgage Disclosure Act (HMDA) database. This dataset covers 90% of mortgage originations in the U.S. and provides detailed information on loan rates, spreads (relative to the average prime offer rate for comparable mortgages), costs (expressed as a ratio of total settlement costs to the loan amount), purpose (for purchase or refinancing), loan amount, term (in months until maturity), loan-to-value ratio (LTV), potential property value, property location by census tract, qualifying loan status, lien status, and originator identity. Additionally, specific borrower information such as gender, race, and age is included.

We followed common data-cleaning procedures in mortgage lending studies (Buchak et al., 2018), including only loans that exceed GSE mortgage limits and retaining loans aimed at home purchases, improvements, refinancing, and cash-out refinancing, excluding reverse mortgages.

3.1.3 Minority

Following prior research, we defined borrowers of Latinx or Black descent as minorities, historically the most discriminated against groups in housing and financial sectors. We excluded Asian Americans from this classification due to their generally higher wealth levels compared to other racial groups, with a median net worth significantly higher than that of White, Black, and Hispanic households.

3.1.4 Fintech, Shadow bank and API bank

We matched the lender names and LEI IDs with RSSD IDs using HMDA data and followed the methodology of Buchak et al. (2018) to classify lending institutions into three mutually exclusive categories: (1) traditional banks, (2) non-fintech shadow banks, (3) fintech shadow banks. We differentiated banks from non-bank institutions by checking whether the financial institution offers deposits and distinguished between fintech companies and shadow banks based on whether loan operations are entirely online with no human intervention. We also conducted cross-verification to overcome subjectivity in categorization.

Furthermore, to distinguish between traditional and API banks, we followed the classification of Lin et al. (2024) by manually collecting API information from banks' official developer sites and integration platforms such as Programmable Web, APIdashboard, Openbanking tracker, Platformable, and APItracker. Our API adoption dataset includes publicly accessible APIs used by financial institutions for external communication (with third-party providers and partners). If a bank had at least one active API during the sample period, it was defined as an API bank. Otherwise, it was considered a traditional bank.

3.2 Empirical Methods

To examine the impact of the Privacy Protection Act on racial disparities in the mortgage market, we combined samples from neighboring covered counties (belonging to states with enacted privacy protection laws) and uncovered counties (not under such laws) with data on mortgage market applications, rejection rates, and racial origins. This specification compared the mortgage market outcomes of minority applicants in treatment counties with those in control counties against their counterparts from 2018-2023 in all border county pairs. This approach takes advantage of the similarity between applicants in neighboring counties over randomly selected households across counties.

Moreover, this method accounts for changes over time between counties and races, such as local economic conditions or credit market situations.

We primarily estimate the following triple difference (RDD) model as shown in Equation (1):

$$y_{i,l,t,c} = \alpha + \beta Treat_c \times Post_t \times Minority_i + \theta Treat_c \times Post_t + \mu Minority_i + \gamma Control_{i,t,c} + \varphi_l + \omega_t + \tau_c + \epsilon_{i,l,t,c} \quad (1)$$

where i, l, t and c represent the loan application, lender, year, and county, respectively. The coefficients of interest pertain to the sequence of estimates associated with the three-way interaction terms β. The independent dummy variable $Treat_c$ indicates whether the mortgage application was in a county (state) where the privacy legislation has been signed, while $Post_t$ indicates whether the mortgage application is submitted after the signing of privacy legislation. $Minority_i$ is a binary variable equal to one if the mortgage applicant belongs to a minority group and zero otherwise.

We employ three specific metrics to measure the dependent variable $y_{i,l,t,c}$: the loan interest rate, the differential between the loan's Annual Percentage Rate (APR) and the Average Prime Offer Rate (APOR) for a comparable transaction at the time the interest rate is set, and the scenario in which the mortgage application is denied by lenders.

We control for borrower characteristics such as Credit Model, Debt-to-Income Ratio, Loan-to-Value Ratio, Loan-to-Income Ratio, Ln Loan Amount, Observed Race, and Age, as these factors could affect the loan interest rates and rejection rates. Table A1 delineates all control variables used in our analysis. We include lender, year, and county $\varphi_l, \omega_t, \tau_c$ fixed effects ($\varphi_l, \omega_t, \tau_c$). Given that the privacy legislation is enacted at the state level, standard errors are clustered at the state level to account for potential heteroskedasticity.

4. Empirical Results

4.1 Summary Statistics

Table 1 presents descriptive statistics for the variables used in our analysis. We have winsorized all continuous variables at the 1st and 99th percentiles to mitigate the impact of potential outliers. Our sample includes 8,764,155 mortgage application observations from adjacent counties, 31,588,968 observations from adjacent states, and 3,148,677 observations from counties adjacent to states where privacy legislation has taken effect. This refinement ensures clearer identification in our analysis. The average value of the Treat × Post variable is 0.194, indicating that 19.4% of mortgage applications observed are located in counties within states that have enacted significant privacy legislation. The average for minority applications is 0.182, suggesting that 18.2% of loan applications are made by minority individuals.

[Insert Table 1 here]

4.2 Baseline Results

To examine whether privacy legislation reduces racial disparities in the U.S. mortgage market, we compare samples from neighboring counties within states that have enacted privacy legislation (covered counties) and those that have not (uncovered counties) in Table 2's Panel A. Panel B uses samples from states with and without enacted privacy laws.

[Insert Table 2 here]

Our baseline results indicate that, after controlling for borrower information, minority borrowers face higher loan interest rates compared to non-minorities, as shown in columns two and four of Table 2. The β coefficient illustrates the differential impact on minority versus non-minority mortgage interest rates before and after privacy legislation enactment. Consistently, the Treat×Post×Minority interaction shows a significant negative coefficient across all models, suggesting that privacy legislation beneficially influences the loan costs for minorities. Adding borrower-specific controls, commonly used in mortgage lending literature such as Credit Model, Debt-to-Income, Loan-to-Value, Loan-to-Income, Ln_Loan_Amount, Race Observed, and Age in columns (2) and (4) preserves the economic and statistical significance of the coefficients, indicating that privacy legislation decreases minority interest rates by 2% and reduces the differential from the standard interest rate by 14%.

Furthermore, in Table 3, we assess whether privacy laws decrease the rejection rates for minority loan applicants using a research design consistent with our baseline approach, focusing particularly on samples from states with neighboring states that have enacted privacy laws.

[Insert Table 3 here]

In the first column without any control variables and the second with borrower-specific controls consistent with Table 1, results uniformly display a significant negative coefficient for *Treat×Post×Minority*, showing that privacy legislation positively impacts the likelihood of loan approval for minorities, thus enhancing equity and inclusion in the mortgage market.

4.3 Dynamic Effects

To alleviate concerns about reverse causality and to validate the parallel trends assumption required for the Difference-in-Differences-in-Differences (DDD) approach, this study draws on the methodology of Bertrand and Mullainathan (2003) to examine the dynamic effects of privacy legislation on racial disparities in mortgage lending. If our results were influenced by reverse causality, or if there were pre-treatment trends (thus violating the parallel trends assumption), we would expect to observe significant changes in interest rates and rejection rates before the implementation of the privacy legislation.

In Table 4, we introduce a series of time dummy variables *Pre_5*, *Pre_4*, *Pre_3*, *Pre_2*, *Pre_1*, *Post_0*, *Post_1*, *Post_2*, *Post_3*, *Post_4*, and *Post_5* to replace the traditional *Post* × *Treat* interaction term, allowing for a more precise examination of the effects before and after the enactment of privacy legislation. These indicators correspond to the five years before through to the five years after the legislation's implementation year.

[Insert Figure 4 and Table 4 here]

The analysis indicates that the interaction terms on *Pre_4* × *Minority*, *Pre_3* × *Minority*, *Pre_2* × *Minority* and *Pre_1* × *Minority* related to minority applicants are statistically insignificant in both the interest rate and loan rejection regressions, thus supporting the validity of the parallel trends assumption. This suggests that reverse causality is unlikely to be driving our main findings. Conversely, the coefficients for *Post_0* × *Minority*, *Post_1* × *Minority*, *Post_2* × *Minority*, *Post_3* × *Minority*, *Post_4* × *Minority* and *Post_5* × *Minority* are significantly negative, consistent with our baseline results. This demonstrates that the implementation of privacy legislation has a sustained effect on reducing interest rates and rejection rates for minority applicants, indicating that privacy legislations help to decrease the costs for minorities in the mortgage market and increase their loan approval rates.

4.4 Effects of Effective Legislation

To address concerns regarding the enactment and effective dates of privacy protection legislation, we refined our sample to include mortgage applications from states where the privacy laws were already in effect, as compared to the signing dates utilized in Sections 4.2 and 4.3. Following the baseline methodology in Section 4.2, Table 5 illustrates the impact on loan rates and interest rate spreads in columns 1-4, and columns (5) and (6) detail the effects on loan rejection rates. Columns one, three, and five include no control variables, whereas columns two, four, and six incorporate a series of controls consistent with the baseline.

[Insert Table 5 here]

Consistent with baseline results, all models consistently demonstrate a significant negative coefficient for *Treat*×*Post*×*Minority*, indicating that privacy legislation positively affects the loan costs and approval rates for minorities, significantly reducing interest rates and rejection rates for minority applicants post-legislation enactment, thus enhancing equity and financial inclusion.

4.5 Other Robustness Checks

Following our baseline specification, we conducted various additional tests to validate the robustness of our findings, including: (i) different combinations of fixed effects (Table A2); and (ii) excluding samples from the Covid period, specifically 2020 to 2021 (Table A3); (iii) excluding samples from states where the mortgage was signed within the same year of privacy legislation enactment (Table A4); (iv) examining the lagged effects of privacy legislation (Table A5); and (v) excluding samples from the ten largest mortgage lenders (Table A6).

[Insert Tables A2 to A6 here]

These robustness checks use alternative evidence to mitigate concerns about business conditions. Consistently, all models show significant negative coefficients for *Treat*×*Post*×*Minority*, affirming that privacy legislation effectively reduces loan rates and rejection rates for minorities, thus fostering equality and financial inclusion.

5. Mechanisms

5.1 Transparency and information friction

We explored potential reasons why privacy legislation might reduce racial disparities in the mortgage market, likely due to increased reliance on hard information (such as standardized credit models) and reduced reliance on soft information (such as face-to-face interactions or subjective observations), enhancing transparency and reducing information friction. We measured reliance on soft information by assessing whether the applicant's race could be observed through facial or surname recognition, while reliance on hard information is based on whether loan decisions relied on standardized credit models.

[Insert Table 6 here]

Comparing coefficients across different subsamples in Panels A and B, we find that results are more pronounced when mortgage decisions rely less on soft information and more on hard information. This suggests that privacy laws may reduce racial disparities by increasing transparency in the decision-making process, validating our mechanism of transparency and information friction.

5.2 Discrimination and Bias

On the other hand, we sought to verify whether the mitigation of racial disparities by privacy legislation was due to a reduction in racial discrimination, as these laws explicitly mandate reduced racial discrimination. We measured racial bias by the concentration of minorities in an area, hypothesizing that areas with higher concentrations of minorities might exhibit more entrenched racial stereotypes.

[Insert Table 7 here]

Our findings, as presented in Table 7, indicate that in areas with higher concentrations of minorities—where racial stereotypes are potentially more pronounced—the effects of privacy legislation are more evident, confirming our mechanism of reducing discrimination. This demonstrates that privacy protection laws significantly mitigate bias in mortgage lending, contributing to the broader goals of equity and diversity in financial practices.

6. Bank and non-bank classification

6.1 Fintech and Shadow Bank

To investigate whether different mortgage providers responded differently to privacy protection laws, we adopted the classification method of Buchak et al. (2018), differentiating lenders into banks and non-bank financial institutions, and further distinguishing non-bank institutions into fintech companies and shadow banks based on whether their lending processes are entirely online.

[Insert Table 8 here]

Compared to traditional banks, fintech and shadow banks exhibited a more pronounced effect in mitigating racial disparities under privacy protection legislation. These institutions leverage advanced technologies, including machine learning and artificial intelligence, to assess creditworthiness and other loan-related metrics. Such technologies process vast amounts of data more accurately and are less likely to harbour biases that might affect human decision-makers. This leads to more objective loan decisions. Additionally, FinTechs automate many processes, reducing human error and bias. Automation ensures consistent application of policies and decisions across all applicants, which helps reduce discriminatory practices that could otherwise affect loan approval rates and terms for different demographic groups. Furthermore, FinTechs and shadow banks place a strong emphasis on customer experience, often making their services more accessible to a broader audience. They tend to offer more personalized products and services to meet diverse customer needs and circumstances, potentially reducing disparities in service access and financial product availability.

6.2 API banks

We also examined whether API banks differ from traditional banks in their impact, following the methodology of Lin et al. (2024) to categorize banks based on the presence of active public APIs on their official websites.

[Insert Table 9 here]

API banks, which provide channels for data communication and sharing with third-party providers like fintech companies, often show a more significant mitigation effect. Typically, API banks offer higher transparency in their operations, building trust among customers. Understanding how decisions are made and what factors are considered can help customers manage their financial situations more effectively.

7. Further analysis

7.1 Effects on Gender Disparities

We tested for differences in gender impact by replacing the Minority variable in our baseline model with Female, indicating whether the mortgage applicant was female.

[Insert Table 10 here]

However, the impact of privacy legislation on gender disparities appears limited; it seems only to reduce rejection rates for mortgage applications by women, without effectively influencing interest rates. This might be due to the less pronounced correlation between gender and credit scores compared to race. Women's credit scores may not be subject to the same systemic biases as those related to race. Additionally, the challenges women face in the loan approval process may relate to factors independent of credit scores, such as income disparities or job stability, which are not directly addressed by privacy legislation.

7.2 Borrower characteristics

Further subsample tests were conducted by dividing the samples into groups based on the borrowers' income and loan amounts: High or Low borrowers' income and High or Low Loan amount borrowers.

[Insert Table 11 here]

Yet, we found no significant differences based on borrowers' income types or loan amount categories, suggesting that privacy regulations do not inherently differ in their impact across different income levels of minority borrowers, maintaining broad financial inclusion.

8. Conclusion

We employed a Triple-Difference (DDD) model to examine whether privacy legislation helps reduce racial disparities in the mortgage market. The findings demonstrate that privacy laws indeed reduce interest rates and rejection rates, particularly enhancing reliance on hard information and reducing dependence on soft information while diminishing discriminatory practices, especially in areas with high minority density. Our research significantly contributes to the literature on data security and privacy protection, providing new evidence for the effects of privacy laws (Doerr et al., 2023; Gupta et al., 2023). Additionally, our findings contribute new insights into the ongoing discussions on financial inclusion and racial disparities in the mortgage market (Bartlett et al., 2022; Bayer et al., 2018; Black et al.,1978; Courchane and Nickerson,1997; Munnell et al., 1996) and offer valuable perspectives for diversity, equity, and inclusion initiatives. Our study holds strong policy implications for both financial institutions and regulatory bodies, demonstrating that the enactment of data privacy protection laws significantly alleviates consumer concerns about privacy and enhances data security controls, ensuring the security and integrity of collected data. can have beneficial effects on financial inclusion and reduce human biases.

Reference

Barocas, S., and A. Selbst, 2016, Big Data's disparate impact, *California Law Review*, 104, 671-732.

Bartlett, R., Morse, A., Stanton, R., & Wallace, N.,2022. Consumer-lending discrimination in the FinTech era. *Journal of Financial Economics*, 143(3), 1209-1226.

Bayer, P., F. Ferreira, and S. L. Ross, 2018, What drives racial and ethnic differences in high-cost mortgages? The role of high-risk lenders, *Review of Financial Studies*, 31, 175-205.

Bertrand M, Mullainathan S, 2003, Enjoying the quiet life? Corporate governance and managerial preferences. *Journal of Political Economy* 111(5):1043-1075.

Black, H. A., and R. Schweitzer, 1977, A canonical analysis of mortgage lending terms: Testing for lending discrimination at a commercial bank, *Urban Studies* 22, 13–19.

Buchak, G. et al., 2018, Fintech, regulatory arbitrage, and the rise of Shadow Banks, *Journal of Financial Economics*, 130(3), pp. 453–483. doi:10.1016/j.jfineco.2018.03.011.

Chen, L., Y. Huang, S. Ouyang, and W. Xiong (2021) "The data privacy paradox and digital demands", *NBER Working Paper.*

Courchane, M., and D. Nickerson, 1997, Discrimination resulting from overage practices, *Journal of Financial Services Research*, 11, 133-151.

Doerr, Sebastian and Gambacorta, Leonardo and Guiso, Luigi and Guiso, Luigi and Sanchez del Villar, Marina, 2023. Privacy Regulation and Fintech Lending

Fuster, A., P. Goldsmith-Pinkham, T. Ramadorai, and A. Walther, 2020, Predictably unequal? The effects of machine learning on credit markets, *Journal of Finance*, Federal Reserve Bank of New York.

Goldfarb, A., and C. Tucker, 2012, Shifts in privacy concerns, *American Economic Review*, 102 (3), 349–353.

Guiso, L., A. Pozzi, A. Tsoy, L. Gambacorta, and P. E. Mistrulli, 2022, The cost of steering in financial markets: Evidence from the mortgage market, *Journal of Financial Economics*, 143, 1209–1226.

Gupta, Manish and McGowan, Danny and Ongena, Steven R. G., 2023. The Cost of Privacy. The Impact of the California Consumer Protection Act on Mortgage Markets.

Howell, S., Kuchler, T., Snitkof, D., Stroebel, J., Wong, J., 2022, Automation in small business lending can reduce racial disparities: evidence from the paycheck protection program. *NBER Working paper*

Huang, J., Lin, X., Shi X., 2024. Market Pressure or Regulatory Pressure? U.S. small bank Pre-emptive IT investment to privacy regulations

Lin, T. (2022) "Valuing intrinsic and instrumental preferences for privacy", *Marketing Science*.

Munnell, A. H., L. Browne, J. Eneaney, and G. Tootel, 1996, Mortgage lending in Boston: interpreting HMDA data, *American Economic Review*, 86, 25-54.

Woodward, S. E., and R. E. Hall, 2012, Diagnosing consumer confusion and suboptimal shopping effort: Theory and mortgage-market evidence, *American Economic Review*, 102, 3249–3276.

Posner, R. A., 1981. The economics of privacy. *American Economic Review* 71, 405–409. 4

Figures and Tables

Figure 1. Coverage of sign into law legislation in states and neighboring states

This figure shows the privacy legislation signing dates across various states in the United States from 2018 to 2023. Different shades of blue represent the signing dates, with darker shades indicating earlier signing dates. Neighboring states that did not sign the legislation are marked in gray, while non-neighboring states are marked in white.

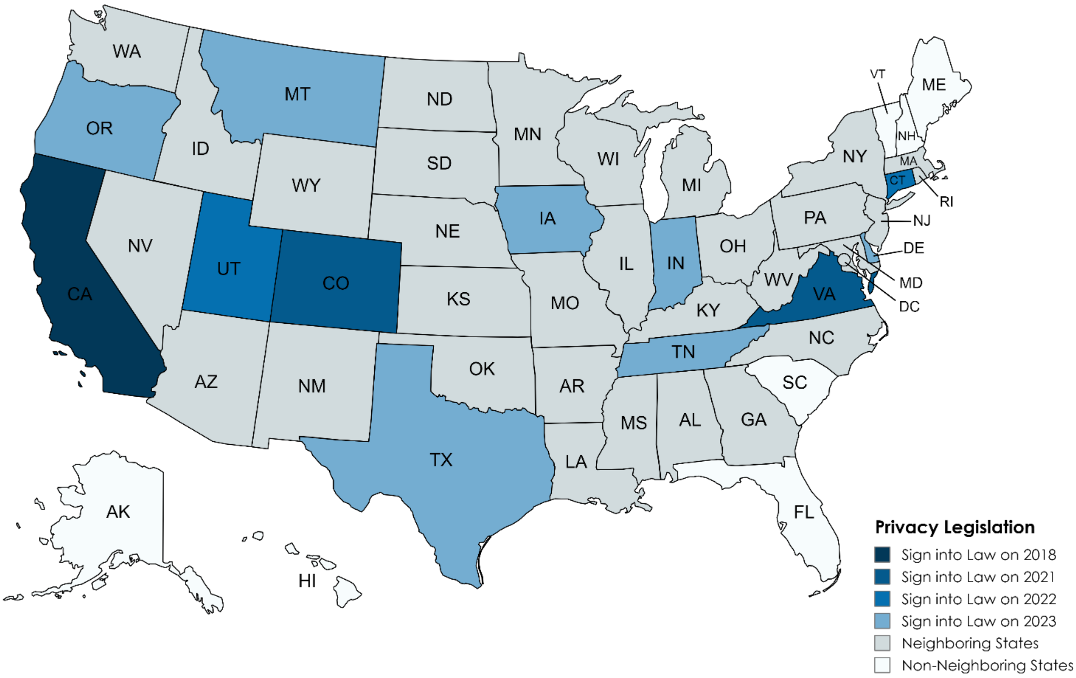

Figure 2. Coverage of only counties adjacent to privacy legislation state borders

This figure shows the signing dates of privacy legislation across various states in the United States from 2018 to 2023, focusing specifically on counties that are adjacent to other states. Different shades of blue represent the signing dates, with darker shades indicating earlier signing dates of the privacy legislation in the respective counties' states. Counties in neighboring states are marked in gray, while non-neighboring counties are marked in white.

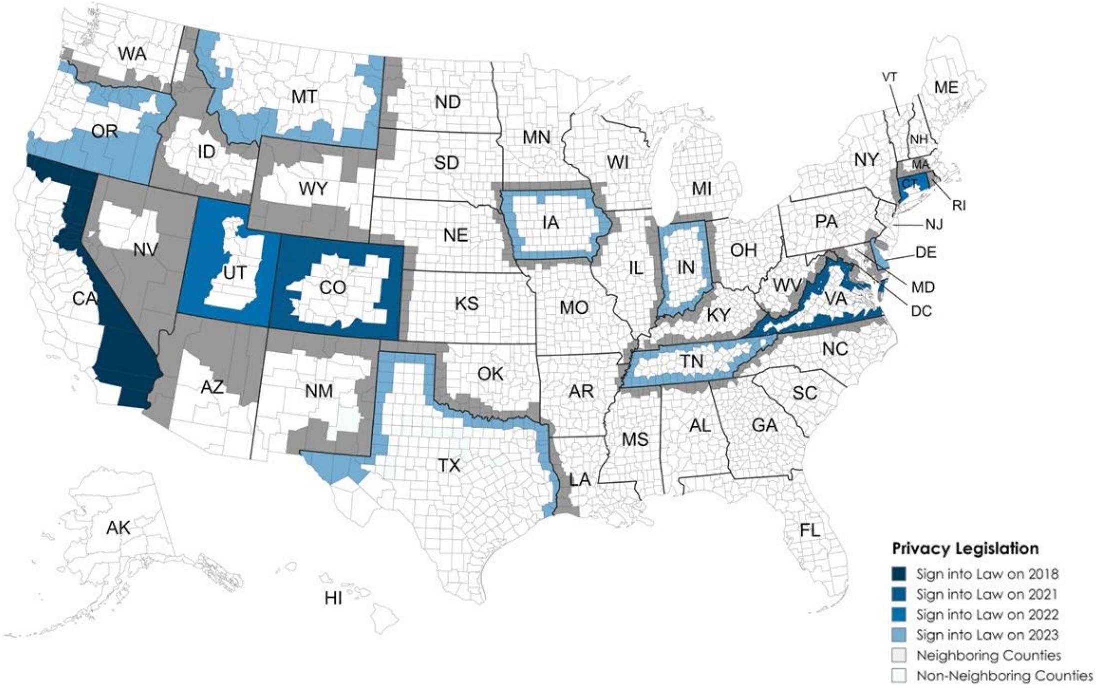

Figure 3. Coverage of Effective legislation states and neighboring states

This figure shows the privacy legislation effective dates across various states in the United States from 2018 to 2023. Different shades of blue represent the effective dates, with darker shades indicating earlier effective dates. Neighboring states that do not have the legislation in effect are marked in gray, while non-neighboring states are marked in white.

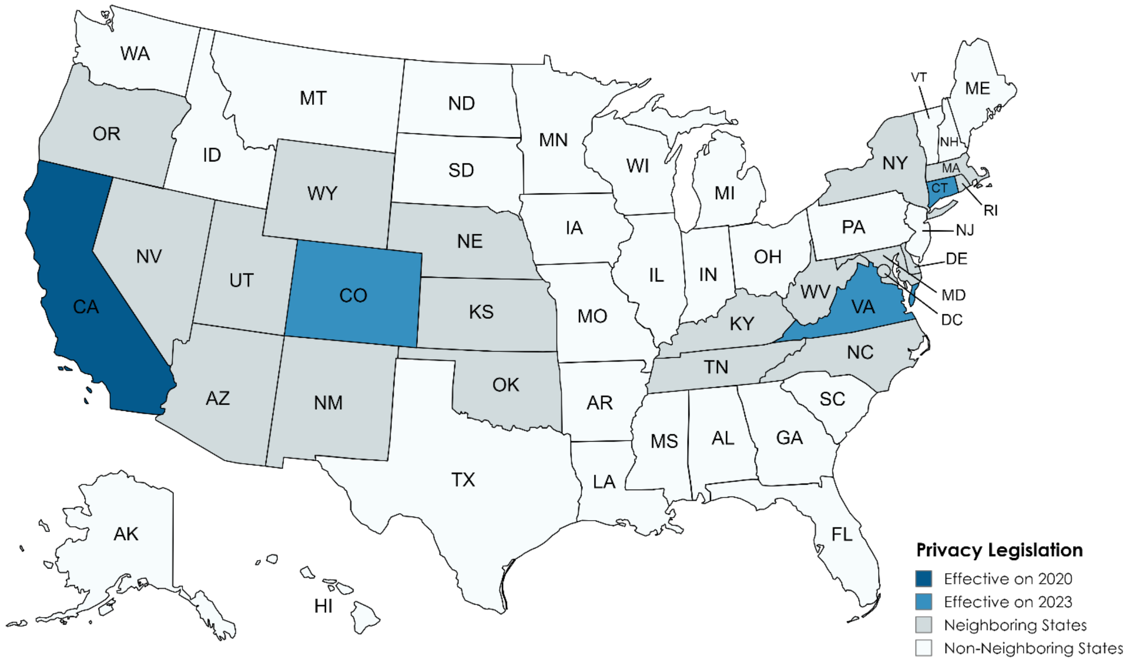

Figure 4. Dynamic Effects of Privacy Legislation on Mortgage Racial Gap

This figure shows the dynamic effects of privacy legislation on interest rates and denial rates, around the signing date of the privacy legislation into law. The horizontal axis presents the year relative to the passage of the privacy legislation, which split the event into 11 bins: *Pre_5*, *Pre_4*, *Pre_3*, *Pre_2*, *Pre_1*, *Post_0*, *Post_1*, *Post_2*, *Post_3*, *Post_4*, and *Post_5*. While the vertical axis presents the estimated coefficient of around the signing of privacy legislation for treatments counties for minority borrowers. The dependent variable are interest rate and racial rate. We decompose the periods before and after We included all control variables consistent with the baseline model. For detailed information on the definition and construction of all control variables, please refer to Appendix Table A1. Continuous variables, except macroeconomic variables, are weighted at the first and 99th percentiles. Standard errors are clustered at the bank level (statistics in parentheses). The bar chart represents a 95% confidence interval.

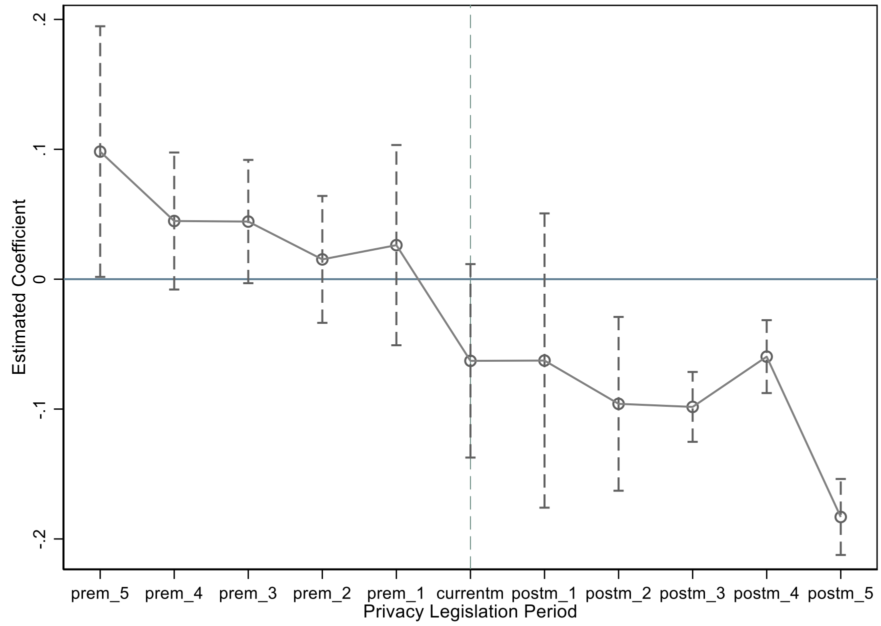

Figure 3a. Effect on interest rates

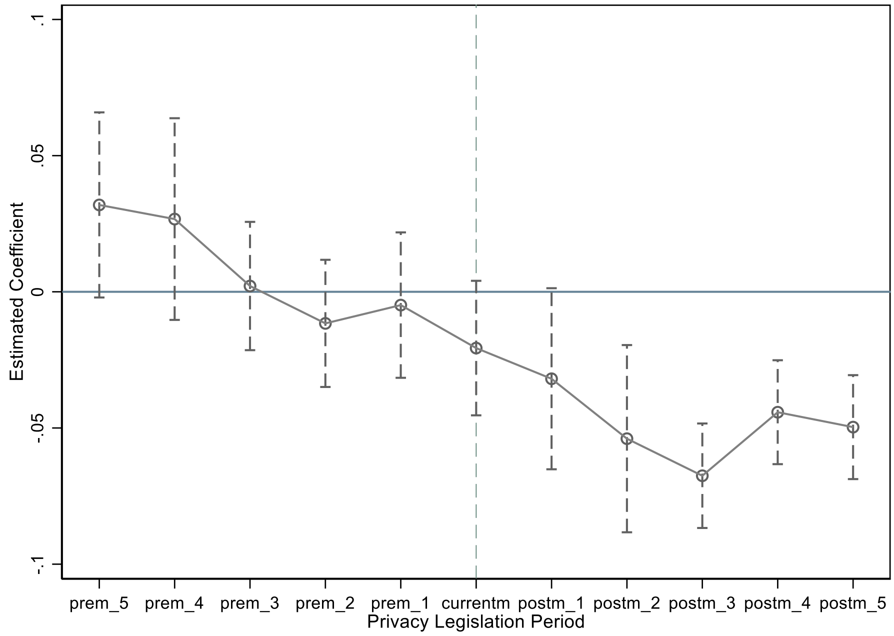

Figure 3b. Effect on Denial rates

Table 1. Summary Statistics

The table reports descriptive statistics for all variables in our main empirical analysis. The sample consists of U.S. mortgage data from 2010 to 2023. The sample used for baseline results includes 8,764,155 application-level observations. For detailed information on the definition and construction of all variables, please refer to the online Appendix Table A1. Continuous variables, except macroeconomic variables, are weighted at the first and 99th percentiles.

Variables	N	Mean	Std. Dev.	1st Perc.	Median	99th Perc.
Treat × Post	8,764,155	0.194	0.396	0	0	1
Privacy_Effective	3,148,677	0.258	0.437	0	0	1
Minority	7,679,759	0.182	0.386	0	0	1
Female	8,764,155	0.240	0.427	0	0	1
Interest Rate	7,103,757	4.143	1.610	1.990	3.625	9.500
Interest Rate Spread	6,724,514	0.435	0.969	-2.113	0.292	4.220
Denial	8,764,155	0.186	0.389	0	0	1
Credit Model	8,764,155	0.891	0.311	0	1	1
Debt_to_Income	8,764,155	0.625	0.484	0	1	1
LTV	8,764,155	0.609	0.276	0.051	0.685	1.000
Loan_to_Income	8,764,155	2.634	5.379	0.167	2.218	8.750
Ln_Loan_Amount	8,764,155	12.116	0.897	9.616	12.278	13.897
Age	8,764,155	3.835	1.417	1	4	7
Sex Observed	8,758,322	0.040	0.195	0	0	1
Race Observed	8,764,155	0.039	0.193	0	0	1
API Bank	8,764,155	0.181	0.385	0	0	1
Bank	8,761,956	0.371	0.483	0	0	1

Table 2. Effect of Privacy Legislation on Mortgage Racial Interest Gap

This table presents the regression results and t-statistics (in parentheses) for the impact of privacy legislation on interest rate and interest rate spread among minorities in the United States from 2018 to 2023. *Treat* indicates whether the mortgage application was in a state where the privacy legislation was signed, *Post* indicates whether the mortgage application was submitted after the enactment of privacy legislation in the state, and *Minority* indicates whether the applicant belongs to an ethnic minority. For the dependent variable, we primarily use *Interest Rate* to represent the cost of mortgage loans and *Interest Rate Spread* refers to the difference between the loan's Annual Percentage Rate (APR) and the Average Prime Offer Rate (APOR). For detailed information on the definition and construction of all control variables, please refer to Appendix Table A1. Continuous variables, except macroeconomic variables, are winsorized at the first and 99th percentiles. Panel A includes only mortgage application samples from *counties* adjacent to states where privacy legislation has been enacted versus counties adjacent to states where it has not been enacted. Panel B includes mortgage application samples from *states* both with and without enacted privacy legislation. Standard errors are clustered at the state level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A Neighboring County				
	Interest Rate		Interest Rate Spread	
	(1)	(2)	(3)	(4)
Treat × Post	0.010	0.016	0.008	0.014
	(0.34)	(0.51)	(0.28)	(0.50)
Minority	0.113***	0.046***	0.146***	0.065***
	(10.51)	(3.66)	(15.36)	(7.04)
Treat × Post× Minority	-0.091***	-0.073***	-0.089***	-0.062***
	(-3.78)	(-2.79)	(-4.54)	(-3.54)
Credit Model		-0.567***		-0.333***
		(-8.48)		(-5.43)
Debt_to_Income		0.077***		0.046***
		(21.05)		(13.94)
LTV		0.182***		0.408***
		(4.66)		(8.62)
Loan_to_Income		-0.005***		-0.006***
		(-6.16)		(-7.48)
Ln_Loan_Amount		-0.252***		-0.302***
		(-10.91)		(-13.85)
Race Observed		0.074***		-0.017
		(4.16)		(-1.01)
Age		-0.013***		-0.007
		(-3.08)		(-1.63)
Constant	4.163***	7.659***	0.433***	4.183***
	(702.74)	(23.64)	(69.72)	(15.10)
Lender FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
County FE	Y	Y	Y	Y
Observations	6,236,958	6,236,958	6,982,601	5,902,077
R-squared	0.651	0.659	0.294	0.317

Table 2. Effect of Privacy Legislation on Mortgage Racial Interest Gap (cont'd)

Panel B Neighboring State				
	Interest Rate		Interest Rate Spread	
	(1)	(2)	(3)	(4)
Treat × Post	-0.055	-0.047	-0.034	-0.025
	(-0.95)	(-0.88)	(-0.81)	(-0.66)
Minority	0.134***	0.062***	0.169***	0.084***
	(15.58)	(7.69)	(21.15)	(11.76)
Treat × Post× Minority	-0.086***	-0.088***	-0.079***	-0.078***
	(-3.73)	(-4.23)	(-5.05)	(-5.54)
Credit Model		-0.463***		-0.282***
		(-5.68)		(-3.47)
Debt_to_Income		0.079***		0.047***
		(27.26)		(20.09)
Loan_to_Value		0.204***		0.416***
		(4.84)		(9.56)
Loan_to_Income		-0.003***		-0.004***
		(-4.79)		(-5.69)
Ln_Loan_Amount		-0.278***		-0.320***
		(-18.19)		(-18.64)
Race Observed		0.078***		-0.030*
		(5.98)		(-1.80)
Age		-0.013***		-0.008
		(-3.02)		(-1.67)
Constant	4.152***	7.860***	0.411***	4.337***
	(374.33)	(34.44)	(52.54)	(18.81)
Lender FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
County FE	Y	Y	Y	Y
Observations	31,588,968	31,588,968	29,870,062	29,870,062
R-squared	0.643	0.661	0.298	0.337

Table 3. Effect of Privacy Legislation on Mortgage Racial Denial Gap

This table presents the regression results and t-statistics (in parentheses) for the impact of privacy legislation on mortgage denial rates among minorities in the United States from 2018 to 2023. *Treat* indicates whether the mortgage application was in a state where the privacy legislation was signed, *Post* indicates whether the mortgage application was submitted after the enactment of privacy legislation in the state, and *Minority* indicates whether the applicant belongs to an ethnic minority. For the dependent variable, we primarily use dummy *variable Denial* to represent whether this application has been rejected or not. Standard errors are clustered at the state level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Denial	
	(1)	(2)
Treat × Post	0.007**	0.007***
	(2.17)	(2.74)
Minority	0.111***	0.081***
	(14.79)	(14.25)
Treat × Post× Minority	-0.042***	-0.034***
	(-3.19)	(-3.25)
Credit Model		-0.167***
		(-13.48)
Debt_to_Income		0.087***
		(30.12)
Loan_to_Value		0.017
		(1.53)
Loan_to_Income		0.005***
		(13.60)
Ln_Loan_Amount		-0.059***
		(-22.69)
Race Observed		0.007
		(1.36)
Age		0.002*
		(1.97)
Constant	0.169***	0.942***
	(114.45)	(35.32)
Lender FE	Y	Y
Year FE	Y	Y
County FE	Y	Y
Observations	7,679,546	7,679,546
R-squared	0.174	0.226

Table 4. Dynamic Effect of Privacy Legislation

The table presents results of the dynamic regression model that analyzes effects of privacy legislation signed in t=0 on interest rate and denial rates for minority mortgage borrowers, based on dummy variables *Pre_5*, *Pre_4*, *Pre_3*, *Pre_2*, *Pre_1*, *Post_0*, *Post_1*, *Post_2*, *Post_3*, *Post_4*, and *Post_5*. T-statistics are presented in parentheses. Standard errors are clustered at the state level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Interest Rate	Denial
VARIABLES	(1)	(2)
Pre_5 × Minority	0.098*	0.032
	(1.75)	(1.45)
Pre_4 × Minority	0.045	0.027
	(1.61)	(1.45)
Pre_3 × Minority	0.044	0.002
	(1.59)	(0.18)
Pre_2 × Minority	0.015	-0.012
	(0.63)	(-1.00)
Pre_1 × Minority	0.026	-0.005
	(0.69)	(-0.37)
Post_0 × Minority	-0.063*	-0.021*
	(-1.70)	(-1.69)
Post_1 × Minority	-0.063	-0.032*
	(-1.12)	(-1.94)
Post_2 × Minority	-0.096***	-0.054***
	(-2.89)	(-3.16)
Post_3 × Minority	-0.098***	-0.068***
	(-7.38)	(-7.10)
Post_4 × Minority	-0.060***	-0.044***
	(-4.29)	(-4.67)
Post_5 × Minority	-0.183***	-0.050***
	(-12.63)	(-5.25)
Minority	0.105***	0.110***
	(8.15)	(11.67)
Pre and Post	Y	Y
Lender FE	Y	Y
Year FE	Y	Y
County FE	Y	Y
Observations	7,891,053	9,771,243
R-squared	0.651	0.174

Table 5. Effects of Effective Privacy Legislation on Mortgage Racial Gap

This table presents the effects of privacy legislation on interest rates, interest rate spread and denial rates, based on a subsample of counties in states where privacy protection laws are already in effect. Specifically, the treatment group consists of states where the privacy legislation has been enacted, while the control group includes neighboring states without such legislation. The $Privacy\_Effective_t$ variable represents the year t when the privacy legislation became effective, as opposed to the signing dates used in Tables 2 – Table 4. Standard errors are clustered at the state level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Interest Rate		Interest Rate Spread		Denial	
	(1)	(2)	(3)	(4)	(5)	(6)
Privacy_Effective	0.051**	0.047**	0.071***	0.065***	0.001	0.001
	(2.42)	(2.14)	(2.99)	(2.89)	(0.25)	(0.55)
Minority	0.119***	0.051***	0.150***	0.073***	0.101***	0.077***
	(12.05)	(5.90)	(19.37)	(12.68)	(18.41)	(16.36)
Privacy_Effective × Minority	-0.076***	-0.077***	-0.067***	-0.065***	-0.042***	-0.040***
	(-4.92)	(-4.10)	(-6.63)	(-4.88)	(-7.28)	(-7.75)
Control Variables	N	Y	N	Y	N	Y
Lender FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Observations	15,867,976	15,867,976	14,965,273	14,965,273	19,372,134	19,372,134
R-squared	0.646	0.663	0.290	0.327	0.173	0.214

Table 6. Mechanism-Transparency and Information Frictions

We analyse potential mechanisms of privacy legislation's effects on mortgage interest rates and denial by splitting our sample into subsamples. Panel A shows results based on the benchmark model in Equation (1) for a subsample of mortgage applications where race has been observed (Columns (1) and (2)) and unobserved (Columns (3) and (4)), reflecting the reliance on soft information when race is observed. Panel B shows that results based on a subsample where a credit model has not been applied (Columns (1) and (2)) and where a credit model has been applied (Columns (3) and (4)), reflecting the reliance on hard information produced by standard credit models. T-statistics are shown in parentheses. Standard errors are clustered at the state level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Race been observed or not

	Race observed		Race unobserved	
	Interest Rate	Denial	Interest Rate	Denial
	(1)	(2)	(3)	(4)
Treat × Post	0.035	0.001	0.012	0.007***
	(0.19)	(0.08)	(0.45)	(2.74)
Minority	0.109***	0.119***	0.043***	0.079***
	(5.30)	(18.17)	(3.54)	(14.00)
Treat × Post × Minority	-0.108	-0.035**	-0.070***	-0.033***
	(-1.57)	(-2.04)	(-2.77)	(-3.27)
Control Variables	Y	Y	Y	Y
Lender FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
County FE	Y	Y	Y	Y
Observations	257,933	340,947	5,978,667	7,338,252
R-squared	0.622	0.217	0.661	0.228

Panel B. Credit Model Usage

	No Credit model		Using Credit model	
	Interest Rate	Denial	Interest Rate	Denial
	(1)	(2)	(3)	(4)
Treat × Post	-0.003	0.012*	0.014	0.006**
	(-0.04)	(2.01)	(0.42)	(2.21)
Minority	0.116***	0.129***	0.038***	0.071***
	(10.94)	(16.49)	(3.36)	(14.33)
Treat × Post × Minority	-0.038	-0.041***	-0.067**	-0.029***
	(-1.21)	(-3.15)	(-2.63)	(-3.26)
Control Variables	Y	Y	Y	Y
Lender FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
County FE	Y	Y	Y	Y
Observations	462,782	865,375	5,774,032	6,814,030
R-squared	0.679	0.263	0.663	0.169

Table 7. Mechanism-Discrimination decrease

This table explores differences in mortgage decisions before and after privacy legislation in areas with low intensity of minorities and lower rates of racial bias and discrimination (Columns (1) and (2)) and areas with high intensity of minorities with higher rates of racial bias and discrimination (Columns (3) and (4)). T-statistics are shown in parentheses. Standard errors are clustered at the state level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Minority non-intensity area		Minority intensity area	
	Interest Rate	Denial	Interest Rate	Denial
	(1)	(2)	(3)	(4)
Treat × Post	-0.008	0.003	0.031	0.009**
	(-0.53)	(1.17)	(0.75)	(2.47)
Minority	0.055***	0.065***	0.051***	0.080***
	(8.02)	(21.64)	(3.62)	(11.98)
Treat × Post × Minority	-0.023	-0.009	-0.070***	-0.032***
	(-1.11)	(-1.49)	(-2.89)	(-2.94)
Control Variables	Y	Y	Y	Y
Lender FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
County FE	Y	Y	Y	Y
Observations	2,915,070	3,468,545	3,321,641	4,210,767
R-squared	0.552	0.212	0.692	0.239

Table 8. Cross-sectional effects for banks and non-banks

This table presents results for the benchmark model in Equation 1 based on subsamples of mortgage applications received by traditional banks (Columns (1) - (3)) and non-banks (Columns (4) - (6)). T-statistics are shown in parentheses. Standard errors are clustered at the state level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Bank			Non-Bank		
	Interest Rate	Interest Rate Spread	Denial	Interest Rate	Interest Rate Spread	Denial
	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post	0.006	0.014	0.006*	0.019	0.026	-0.003
	(0.14)	(0.31)	(1.82)	(0.71)	(1.09)	(-1.31)
Minority	0.057***	0.066***	0.110***	0.035***	0.062***	0.045***
	(5.35)	(6.89)	(19.28)	(2.78)	(4.79)	(11.52)
Treat × Post × Minority	-0.056**	-0.032*	-0.027***	-0.070***	-0.074***	-0.015*
	(-2.07)	(-1.80)	(-3.59)	(-2.74)	(-3.46)	(-2.00)
Control Variables	Y	Y	Y	Y	Y	Y
Lender FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Observations	3,107,391	2,967,565	4,038,455	1,854,481	1,723,984	2,175,097
R-squared	0.634	0.287	0.217	0.724	0.520	0.292

Table 9. Cross-sectional effects for API banks and non-API banks/non-banks

This table presents results for the benchmark model in Equation 1 based on subsamples of mortgage applications received by banks that have adopted APIs (Columns (1) - (3)) and banks that have not adopted APIs as well as non-banks (Columns (4) - (6)). T-statistics are shown in parentheses. Standard errors are clustered at the state level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	API-Bank			Non-API Bank and Non-Bank		
	Interest Rate	Interest Rate Spread	Denial	Interest Rate	Interest Rate Spread	Denial
	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post	0.096	0.096**	0.020*	0.009	0.005	0.002
	(1.25)	(2.08)	(1.95)	(0.33)	(0.21)	(1.08)
Minority	0.029	0.050***	0.122***	0.048***	0.066***	0.068***
	(1.62)	(3.18)	(16.65)	(4.01)	(7.60)	(13.85)
Treat × Post × Minority	-0.078***	-0.048*	-0.035***	-0.072***	-0.062***	-0.030***
	(-2.94)	(-2.00)	(-3.86)	(-2.72)	(-3.29)	(-2.72)
Control Variables	Y	Y	Y	Y	Y	Y
Lender FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Observations	1,006,207	984,220	1,427,639	5,230,746	4,917,851	6,251,903
R-squared	0.649	0.376	0.211	0.668	0.325	0.226

Table 10. Effect of Privacy Legislation on Mortgage Gender Gap

This table presents the effect of privacy legislation signing (Panel A) and effective dates (Panel B) on interest rates, interest rate spread, and denial rates on mortgage applications filed by female applicants (*Female*). T-statistics are shown in parentheses. Standard errors are clustered at the state level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Effect of privacy legislation signing on gender disparities						
	Interest Rate		Interest Rate Spread		Denial	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post	-0.004	0.008	0.001	0.006	0.002	0.002
	(-0.15)	(0.25)	(0.04)	(0.23)	(0.81)	(1.32)
Female	0.038***	-0.016***	0.044***	-0.020***	0.048***	0.008***
	(8.33)	(-3.75)	(9.40)	(-5.10)	(29.93)	(5.01)
Treat × Post × Female	0.007	-0.009	-0.024***	-0.005	-0.008***	-0.006***
	(0.58)	(-0.96)	(-3.74)	(-0.84)	(-3.78)	(-4.05)
Control Variables	N	Y	N	Y	N	Y
Lender FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Observations	9,907,903	9,907,903	8,762,743	7,943,328	7,101,984	7,101,984
R-squared	0.652	0.652	0.216	0.292	0.660	0.155

Table 10. Effect of Privacy Legislation on Mortgage Gender Gap (cont'd)

Panel B Female Effects of Effective Privacy Legislation on Mortgage Racial and Gender Gap

	Interest Rate		Interest Rate Spread		Denial	
	(1)	(2)	(3)	(4)	(5)	(6)
Privacy_Effective	0.038*	0.031	0.060**	0.051**	-0.006**	-0.006**
	(1.74)	(1.41)	(2.52)	(2.28)	(-2.41)	(-2.37)
Female	0.031***	-0.022***	0.039***	-0.022***	0.031***	0.011***
	(12.01)	(-4.67)	(13.97)	(-4.22)	(19.58)	(7.45)
Privacy_Effective× Female	-0.009	-0.000	-0.010	0.003	-0.012***	-0.010***
	(-1.00)	(-0.02)	(-1.22)	(0.46)	(-7.87)	(-5.89)
Control Variables	N	Y	N	Y	N	Y
Lender FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Observations	15,859,373	15,859,373	14,957,447	14,957,447	19,362,883	19,362,883
R-squared	0.646	0.663	0.288	0.326	0.167	0.210

Table 11. Cross-sectional borrower's characteristics

This table further presents the effects of privacy legislation signing based on subsamples for borrowers with high and low income (Panel A) and borrowers with high and low loan amounts (Panel B), based on the regression model in Equation (1). T-statistics are shown in parentheses. Standard errors are clustered at the state level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Borrowers' income subsample analysis						
	High-income borrowers			Low-income borrowers		
	Interest Rate	Interest Rate Spread	Denial	Interest Rate	Interest Rate Spread	Denial
	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post	0.015	0.014	0.007***	0.033	0.028	0.004
	(0.73)	(0.64)	(3.39)	(0.89)	(0.85)	(0.94)
Minority	0.112***	0.115***	0.080***	0.026***	0.047***	0.073***
	(10.38)	(13.32)	(19.35)	(2.90)	(6.19)	(12.02)
Treat × Post × Minority	-0.075**	-0.064***	-0.039***	-0.071***	-0.056***	-0.028**
	(-2.52)	(-2.97)	(-4.37)	(-3.50)	(-4.16)	(-2.44)
Control Variables	Y	Y	Y	Y	Y	Y
Lender FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Observations	3,325,740	3,116,885	3,841,099	2,910,965	2,784,924	3,838,189
R-squared	0.687	0.306	0.169	0.637	0.339	0.256

Table 11. Cross-sectional borrower's characteristics (cont'd)

Panel B. Borrowers' loan amount subsample analysis						
	High-Loan amount borrowers			Low-Loan amount borrowers		
	Interest Rate	Interest Rate Spread	Denial	Interest Rate	Interest Rate Spread	Denial
	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post	0.011	0.013	0.002	-0.030	-0.016	0.004
	(0.78)	(0.98)	(1.60)	(-0.43)	(-0.27)	(0.89)
Minority	0.090***	0.123***	0.055***	0.074***	0.096***	0.109***
	(9.77)	(14.88)	(14.66)	(3.94)	(6.57)	(14.65)
Treat × Post × Minority	-0.065***	-0.061***	-0.021***	-0.139***	-0.118***	-0.033**
	(-5.02)	(-6.24)	(-3.50)	(-2.75)	(-3.05)	(-2.35)
Control Variables	Y	Y	Y	Y	Y	Y
Lender FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Observations	3,357,698	3,220,560	3,770,901	2,878,977	2,681,237	3,908,359
R-squared	0.729	0.228	0.127	0.621	0.328	0.243

Online Appendix
Table A1. Variable definition

Variable	Definition
Treat × Post	*Treat* indicates whether the mortgage application was in a state where the privacy legislation was signed, *Post* indicates whether the mortgage application was submitted after the privacy legislation has been signed into law in the state
Privacy_Effective	Equals 1 if the application i in year t at state s's privacy legislation is effective, 0 otherwise
Minority	Equals 1 if the applicant's ethnicity is Latino, or their race is Black, 0 otherwise.
Female	Equals 1 if the applicant is female, 0 otherwise.
Interest Rate	The cost of mortgage loans in percent
Interest Rate Spread	The difference between the loan's Annual Percentage Rate (APR) and the Average Prime Offer Rate (APOR) for a comparable transaction at the time the interest rate is set.
Denial	Whether lenders deny this mortgage application
Credit Model	Whether a credit scoring model is used to generate the credit score, or scores, relied on in making the credit decision
Debt_to_Income	The ratio, as a percentage, of the applicant's or borrower's total monthly debt to the total monthly income relied on in making the credit decision
LTV	Loan_to_Value, the ratio of the total amount of debt secured by the property to the value of the property relied on in making the credit decision
Loan_to_Income	The ratio, as a percentage, of the applicant's or borrower's current amount of lending to the total monthly income relied on in making the credit decision
Ln_Loan_Amount	The amount of the covered loan, or the amount applied for
Age	The age of the applicant is categorized into seven groups using numeric values from 1 to 7: age<25, 25-34, 35-44, 45-54, 55-64, 65-74, >74
Sex Observed	Whether the sex of the applicant or borrower was collected based on visual observation or surname
Race Observed	Whether the race of the applicant or borrower was collected based on visual observation or surname
Fintech	Whether the lending progress is online only or not (Buchak et al., 2018)
API Bank	Whether the lender is an API bank or not
Bank	Whether the lender is a bank or not

Table A2. Robustness_ fixed effects

This table presents the regression results and t-statistics (in parentheses) for the impact of privacy legislation on interest rate, interest rate spread and denial among minorities in the United States from 2018 to 2023. *Treat* indicates whether the mortgage application was in a state where the privacy legislation was signed, *Post* indicates whether the mortgage application was submitted after the enactment of the privacy legislation in the state, and *Minority* indicates whether the applicant belongs to an ethnic minority. For the dependent variable, we primarily use *Interest Rate* to represent the cost of mortgage loans and *Interest Rate Spread* refers to the difference between the loan's Annual Percentage Rate (APR) and the Average Prime Offer Rate (APOR), while we use dummy *variable Denial* to represent whether this application has been rejected or not. For detailed information on the definition and construction of all control variables, please refer to Appendix Table A1. Continuous variables, except macroeconomic variables, are winsorized at the first and 99th percentiles. Standard errors are clustered at the state level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Interest Rate			Interest Rate Spread			Denial		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treat × Post	0.555**	0.106***	0.033**	0.012	0.092***	0.008	0.011	0.018*	0.010
	(2.03)	(6.33)	(2.13)	(0.94)	(5.34)	(0.65)	(1.63)	(1.77)	(1.64)
Minority	0.152***	0.085***	0.050***	0.063***	0.116***	0.067***	0.087***	0.106***	0.087***
	(6.71)	(3.38)	(3.60)	(5.08)	(5.00)	(6.04)	(16.39)	(15.41)	(16.36)
Treat × Post × Minority	-0.317**	-0.108***	-0.076***	-0.061***	-0.095***	-0.063***	-0.035***	-0.042***	-0.035***
	(-2.68)	(-3.00)	(-3.09)	(-3.54)	(-2.99)	(-3.91)	(-3.49)	(-2.98)	(-3.58)
Control Variables	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lender FE	Y	N	Y	Y	N	Y	Y	N	Y
Year FE	N	Y	Y	N	Y	Y	N	Y	Y
County FE	N	N	N	N	N	N	N	N	N
Observations	6,236,958	6,237,172	6,236,958	5,902,077	5,902,290	5,902,077	7,679,546	7,679,759	7,679,546
R-squared	0.230	0.572	0.658	0.310	0.106	0.315	0.224	0.125	0.225

Table A3. Robustness_ Exclude Covid period 2020-2021

This table presents the regression results and t-statistics (in parentheses) for the impact of privacy legislation on interest rate, interest rate spread and denial among minorities in the United States from 2018 to 2023 but excludes the Covid period (2020 and 2021). *Treat* indicates whether the mortgage application was in a state where the privacy legislation was signed, *Post* indicates whether the mortgage application was submitted after the enactment of the privacy legislation in the state, and *Minority* indicates whether the applicant belongs to an ethnic minority. For the dependent variable, we primarily use *Interest Rate* to represent the cost of mortgage loans and *Interest Rate Spread* refers to the difference between the loan's Annual Percentage Rate (APR) and the Average Prime Offer Rate (APOR), while we use dummy *variable Denial* to represent whether this application has been rejected or not. For detailed information on the definition and construction of all control variables, please refer to Appendix Table A1. Continuous variables, except macroeconomic variables, are winsorized at the first and 99th percentiles. Standard errors are clustered at the state level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A Minority			
	Interest Rate	Interest Rate Spread	Denial
	(1)	(2)	(3)
Treat × Post	0.003	0.003	0.009***
	(0.06)	(0.06)	(3.15)
Minority	0.050**	0.062***	0.091***
	(2.55)	(3.80)	(13.94)
Treat × Post × Minority	-0.094**	-0.072**	-0.033***
	(-2.61)	(-2.66)	(-3.19)
Control Variables	Y	Y	Y
Lender FE	Y	Y	Y
Year FE	Y	Y	Y
County FE	Y	Y	Y
Observations	3,116,054	2,926,161	4,019,358
R-squared	0.504	0.301	0.237

Panel B Female			
	Interest Rate	Interest Rate Spread	Denial
	(1)	(2)	(3)
Treat × Post	-0.007	-0.008	0.005**
	(-0.13)	(-0.18)	(2.52)
Female	-0.022***	-0.026***	0.015***
	(-3.65)	(-4.73)	(7.53)
Treat × Post × Female	-0.011	0.001	-0.007***
	(-0.93)	(0.08)	(-3.72)
Control Variables	Y	Y	Y
Lender FE	Y	Y	Y
Year FE	Y	Y	Y
County FE	Y	Y	Y
Observations	3,113,977	2,924,239	4,017,122
R-squared	0.504	0.301	0.232

Table A4. Robustness_ Exclude Privacy legislation current year

This table presents the regression results and t-statistics (in parentheses) for the impact of privacy legislation on interest rate, interest rate spread and denial among minorities in the United States from 2018 to 2023 but excludes the current year of the privacy legislation signed into law. *Treat* indicates whether the mortgage application was in a state where the privacy legislation was signed, *Post* indicates whether the mortgage application was submitted after the enactment of the privacy legislation in the state, and *Minority* indicates whether the applicant belongs to an ethnic minority. For the dependent variable, we primarily use *Interest Rate* to represent the cost of mortgage loans and *Interest Rate Spread* refers to the difference between the loan's Annual Percentage Rate (APR) and the Average Prime Offer Rate (APOR), while we use dummy *variable Denial* to represent whether this application has been rejected or not. For detailed information on the definition and construction of all control variables, please refer to Appendix Table A1. Continuous variables, except macroeconomic variables, are winsorized at the first and 99th percentiles. Standard errors are clustered at the state level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Minority			
	Interest Rate	Interest Rate Spread	Denial
	(1)	(2)	(3)
Treat × Post	0.058	0.059	0.009***
	(1.18)	(1.23)	(3.82)
Minority	0.041***	0.061***	0.081***
	(3.43)	(7.09)	(14.27)
Treat × Post × Minority	-0.068**	-0.055***	-0.042***
	(-2.60)	(-2.80)	(-4.38)
Control Variables	Y	Y	Y
Lender FE	Y	Y	Y
Year FE	Y	Y	Y
County FE	Y	Y	Y
Observations	5,827,713	5,517,165	7,169,108
R-squared	0.642	0.320	0.226

Panel B. Female			
	Interest Rate	Interest Rate Spread	Denial
	(1)	(2)	(3)
Treat × Post	0.046	0.048	0.003
	(0.93)	(0.99)	(1.35)
Female	-0.021***	-0.023***	0.011***
	(-4.81)	(-6.34)	(6.36)
Treat × Post × Female	0.005	0.007*	-0.007***
	(0.98)	(2.01)	(-4.25)
Control Variables	Y	Y	Y
Lender FE	Y	Y	Y
Year FE	Y	Y	Y
County FE	Y	Y	Y
Observations	5,824,963	5,514,681	7,166,137
R-squared	0.642	0.320	0.221

Table A5. Robustness_ Lag effect of privacy legislation

This table presents the regression results and t-statistics (in parentheses) for the impact of privacy legislation on interest rate, interest rate spread and denial among minorities in the United States from 2018 to 2023. *Treat* indicates whether the mortgage application was in a state where the privacy legislation was signed, and *Post* indicates whether the mortgage application was submitted after the enactment of the privacy legislation in the state, we lagged the Post in these regressions. *Minority* indicates whether the applicant belongs to an ethnic minority. For the dependent variable, we primarily use *Interest Rate* to represent the cost of mortgage loans and *Interest Rate Spread* refers to the difference between the loan's Annual Percentage Rate (APR) and the Average Prime Offer Rate (APOR), while we use dummy *variable Denial* to represent whether this application has been rejected or not. For detailed information on the definition and construction of all control variables, please refer to Appendix Table A1. Continuous variables, except macroeconomic variables, are winsorized at the first and 99th percentiles. Standard errors are clustered at the state level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Interest Rate	Interest Spread	Denial
VARIABLES	(1)	(2)	(3)
Treat × Post-1	0.018	0.037	0.003
	(0.41)	(1.03)	(1.13)
Minority	0.041***	0.059***	0.080***
	(3.71)	(7.61)	(15.18)
Treat × Post-1 × Minority	-0.072***	-0.050***	-0.042***
	(-3.02)	(-2.98)	(-4.59)
Control Variables	Y	Y	Y
Lender FE	Y	Y	Y
Year FE	Y	Y	Y
County FE	Y	Y	Y
Observations	6,236,958	5,902,077	7,679,546
R-squared	0.659	0.317	0.226

Table A6. Robustness_ Exclude the 10 largest lenders

This table presents the regression results and t-statistics (in parentheses) for the impact of privacy legislation on interest rate, interest rate spread and denial among minorities in the United States from 2018 to 2023. We exclude the top ten lenders (which have received the most applications). *Treat* indicates whether the mortgage application was in a state where the privacy legislation was signed, *Post* indicates whether the mortgage application was submitted after the enactment of the privacy legislation in the state, and *Minority* indicates whether the applicant belongs to an ethnic minority. For the dependent variable, we primarily use *Interest Rate* to represent the cost of mortgage loans and *Interest Rate Spread* refers to the difference between the loan's Annual Percentage Rate (APR) and the Average Prime Offer Rate (APOR), while we use dummy *variable Denial* to represent whether this application has been rejected or not. For detailed information on the definition and construction of all control variables, please refer to Appendix Table A1. Continuous variables, except macroeconomic variables, are winsorized at the first and 99th percentiles. Standard errors are clustered at the state level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Minority			
	Interest Rate	Interest Rate Spread	Denial
	(1)	(2)	(3)
Treat × Post	0.029	0.030	0.002
	(0.90)	(1.16)	(0.72)
Minority	0.053***	0.072***	0.072***
	(4.89)	(7.76)	(12.84)
Treat × Post × Minority	-0.097***	-0.089***	-0.031**
	(-3.34)	(-4.15)	(-2.58)
Control Variables	Y	Y	Y
Lender FE	Y	Y	Y
Year FE	Y	Y	Y
County FE	Y	Y	Y
Observations	4,516,186	4,234,844	5,447,896
R-squared	0.656	0.340	0.244

Panel B. Female			
	Interest Rate	Interest Rate Spread	Denial
	(1)	(2)	(3)
Treat × Post	0.017	0.018	-0.002
	(0.51)	(0.69)	(-1.11)
Female	-0.015***	-0.016***	0.012***
	(-2.88)	(-2.97)	(6.22)
Treat × Post × Female	-0.003	-0.001	-0.006**
	(-0.29)	(-0.32)	(-2.23)
Control Variables	Y	Y	Y
Lender FE	Y	Y	Y
Year FE	Y	Y	Y
County FE	Y	Y	Y
Observations	4,513,368	4,232,297	5,444,845
R-squared	0.656	0.340	0.241

Table A7. Event study

Panel A: Event dates for privacy legislation			
	Event date	State	Description
(1)	Jun 28, 2018	California	The California Consumer Privacy Act (CCPA) was signed into law on June 28, 2018, and went into effect on January 1, 2020. The CCPA applies to any business, including any for-profit entity that collects consumers' data, which does business in California
(2)	Jul 7, 2021	Colorado	Colorado Senate Bill 190 (SB 190), also known as the Colorado Privacy Act (CPA), was introduced and signed into law on July 7, 2021, and its effective Date is July 1, 2023
(3)	May 10, 2022	Connecticut	Connecticut Data Privacy Act (CTDPA), it has been signed into Law by Governor Ned Lamont on May 10, 2022, and its effective Date is July 1, 2023.
(4)	Mar 24, 2022	Utah	Utah Senate Bill 227 (SB 227), also known as the Utah Consumer Privacy Act (UCPA), was signed into law by Governor Spencer Cox on March 24, 2022. The law is set to take effect on December 31, 2023.
(5)	Mar 2, 2021	Virginia	Virginia Senate Bill 1392 (SB 1392), known as the Virginia Consumer Data Protection Act (VCDPA), was signed into law by Governor Ralph Northam on March 2, 2021. The law is set to take effect on January 1, 2023.
(6)	Sep 11, 2023	Delaware	Delaware Personal Data Privacy Act (DPDPA) was signed into law on September 11, 2023, and it will take effect on January 1, 2025.
(7)	May 1, 2023	Indiana	Indiana Consumer Data Protection Act (ICDPA) was signed into law on May 1, 2023, and it will take effect on January 1, 2026.
(8)	Mar 28, 2023	Iowa	Iowa Consumer Data Protection Act (ICDPA) was signed into law on March 28, 2023, and it will take effect on January 1, 2025.
(9)	May 19, 2023	Montana	Montana Consumer Data Privacy Act (MCDPA) was signed into law on May 19, 2023, and it will take effect on October 1, 2024.
(10)	Jun 22, 2023	Oregon	Oregon Consumer Privacy Act (OCPA) was signed into law on June 22, 2023, and it will take effect on July 1, 2024.
(11)	May 11, 2023	Tennessee	Tennessee Information Protection Act (TIPA) was signed into law on May 11, 2023, and it will take effect on July 1, 2025.
(12)	Jun 18, 2023	Texas	Texas Data Privacy and Security Act (TDPSA) was signed into law on June 18, 2023, it will take effect on July 1, 2024.

Panel B Average abnormal return per event, estimated over 12 events combined				
Event Windows		Cumulative abnormal return (Mean)	StdCsect	Rank Test
(1)	Day (-2,2)	0.0043	1.813*	1.491
			(0.070)	(0.137)
(2)	Day (-3,3)	0.0061	2.315**	1.827*
			(0.021)	(0.069)
(3)	Day (-5,5)	0.0076	2.263**	2.043**
			(0.024)	(0.042)