KNOW THYSELF: FREE CREDIT REPORTS AND THE RETAIL MORTGAGE MARKET*

Amit Kumar†

[This Version: May 2021] [Click here for the latest version]

Abstract

Under imprecise creditworthiness information, borrowers may make erroneous credit decisions. Credit reports—which record one's creditworthiness—became free in the entire U.S. in 2005, while they already had been free in seven states. Exploiting this in a difference-in-differences setting, this paper shows that cheaper credit reports to consumers changed mortgage market outcomes that are indicative of improvement in borrower pool. Approval ratios and applications increased, whereas defaults decreased. Low-income quartile borrowers and *ex-ante* high-creditworthy areas saw larger increase in approvals. Additional findings, including increased interest rates, suggest a demand-driven pool improvement, as consumers receive precise creditworthiness signal from their reports.

JEL Codes: D12, D83, G21, G28, L51

Keywords: Credit Reports, Information Provision to Consumers, Household Finance, Mortgages, Regulation of Credit Information

---

*I am profoundly grateful to Utpal Bhattacharya, Emilio Bisetti and John G. F. Nash for their guidance in shaping this paper. For their valuable feedback during different doctoral symposiums, I express my sincere gratitude to Renée Adams and Anjan Thakor. I am also indebted to Deniz Okat, Abhiroop Mukherjee, Arkodipta Sarkar, and Alminas Zaldokas for various consultations and to Andreas Fuster, Daniel Ringo, Shashwat Alok, Francesco D'Acunto, Vimal Balasubramaniam and Sebastian Hillenbrand for conference discussions. I also want to sincerely thank the discussants and participants at the following meetings for helpful suggestions: HKUST brownbag seminar, American Real Estate and Urban Economics Association (AREUEA) 2021 natioanl conference, Asian Bureau of Finance and Economic Research (ABFER) 8th annual conference, 2021 Annual meeting of the Swiss Society for Financial Market Research (SGF 2021), PhD symposium at McCombs Business School (UT Austin), Financial Management Association (FMA) 2020 annual meeting, 33rd Australasian Finance and Banking Conference (AFBC 2020), 11th Emerging markets conference, Southern Denmark University 5th Finance Workshop, PhD mentoring day at 28th Finance Forum (AEFIN 2020), 15th Annual Conference on Asia Pacific Financial Markets (CAFM), and 2020 Greater China Area Finance Conference.

†Hong Kong University of Science and Technology. All errors are my own. Email: amit.kumar@connect.ust.hk

Introduction

Borrowers may make mistakes in credit decisions when they have imprecise information of their creditworthiness. Credit applications of those who overestimate their creditworthiness may get rejected, a notable example of which comes from the U.S. mortgage market in which about 34% of the applications were rejected from 2000 to 2008. The overestimation is reflected in the fact that rejections due to credit history were twice more than those due to debt-to-income (DTI) ratio. As the rejections are costly, the decisions behind these applications appear sub-optimal. Not only a rejection likely increases future rejection probability and interest rates on *any* formal credit application, but also the borrower loses the application fees (~$300 to $400 for a mortgage application). With correct self-assessment of creditworthiness, the borrowers with bad credit history could save the rejection costs by not applying.

At the same time, those who underestimate their creditworthiness may decide not to apply for credit and end up paying the opportunity cost of lack of credit (the *discouraged borrowers*). This tendency to do so is also, surprisingly, prevalent. Survey of Consumer Finances (1998–2007) (SCF) reveals that a staggering ~15% of the U.S. households report as being discouraged when deciding on applying for credit (all types). Similarly, relating to mortgage credit, Survey of Consumer Expectations (2013–2020) (SCE) shows that among those who were unlikely to apply for mortgage or refinance, 13% were discouraged.

This paper examines the retail mortgage market outcomes when economic cost of accessing creditworthiness information is reduced. Specifically, the paper evaluates the causal effect of lowering consumers' economic costs of accessing *credit reports*—an authoritative information source of consumer creditworthiness—on the mortgage market outcomes. Lowering the economic costs resulted in three main effects. First, mortgage demand and approval ratio increased, and subprime population fraction decreased. Second, good quality borrowers seem to be behind the changes: origination was higher in more creditworthy areas and among prime consumers, and also new mortgages were less likely to be defaulted upon. Finally, credit-history rejections decrease slightly, but DTI ratio rejections do not.

This paper proposes that credit reports affect the market outcomes through a consumer *self-learning mechanism*; here the reports aid consumers self-assess their creditworthiness. As the reports contain crucial financial information on consumers e.g., their creditworthiness, credit history and borrowing capacity (Figure I), lenders utilize those in deciding on mortgage applications. Thus, before applying for credit, consumers may learn their creditworthiness signaled by the report and decide on the application. Those with good signal may stay-in/enter the credit market, while those with bad may search for a suitable (subprime) lender, or do not apply (exit the market); therefore, the borrower pool improves due to better consumer sorting and results in higher approval ratio. Whether the demand for credit increase depends on prior distribution of over- and under-estimators and of those who are unaware of the role of credit reports in a mortgage application.

It is when the economic costs of accessing credit reports are high that reducing those would change the market outcomes. Despite the monetary cost of the reports being trivially small (historically just ~$8), and despite the consequences of imprecise creditworthiness information, the woefully low usage of reports among consumers does suggest that the costs are high.[1] Out of approximately 1 billion reports generated annually in the U.S. in early 2000's, a mere 1.6% were requested by consumers (Avery, Calem, & Canner, 2004). If we assume that all these reports were requested by potential mortgage applicants, less than 5% of the applications in 2004 could potentially have come from those who had checked their credit history. Similarly, even in 2020, almost 12% of respondents in the SCE survey do not know their credit score, and 20% have either never checked their credit reports, or checked it more than two years ago.

Many factors potentially contribute to the economic costs of accessing credit reports. Search costs (where to obtain the reports from), unawareness costs (what role do the reports play in credit/mortgage applications) and illiteracy costs (what does the information in the reports mean and how to use it in credit decisions) seem to be the main factors.[2]

A key challenge in examining the link between consumers' economic costs of credit reports and the credit market outcomes is to establish causality. This paper addresses it using a natural experiment in the U.S.—the enactment of the federal *Fair and Accurate Transaction Act of 2003* (FACTA)—that led to a close-to-exogenous reduction in consumers' economic costs of accessing credit reports. The FACTA made access to three credit reports free annually from 2005 for all consumers, through the website www.annualcreditreport.com. However, seven states—Colorado, Georgia, Maine, Maryland, Massachusetts, New Jersey, and Vermont—had already

[1] Imprecise information on creditworthiness may lead to: wrong self-assessment of creditworthiness (Perry, 2008), debt underestimation (Brown, Haughwout, Lee, & Van Der Klaauw, 2011), and worse financial outcomes (Courchane, Gailey, & Zorn, 2008).

[2] In 2000, the three consumer reporting agencies (CRAs)—Equifax, Experian, and Transunion—blocked calls of millions of consumers who wanted to discuss the content of their credit reports, and ended up settling a lawsuit for this (Federal Trade Commission, January 13, 2000).

been allowing their residents free credit reports before the FACTA. So, consumers from all states except these seven (the pre-FACTA states) saw a reduction in cost of the reports in 2005. More importantly, now consumers could access reports with a few clicks, whereas earlier it could take a week for the report to arrive after having successfully registered a request for the report. Then, the law also potentially raised general awareness about the reports. All in all, the experiment reduced the monetary, awareness and search cost of accessing the reports. Consumer interest in free credit reports, measured using Google Search Interest for the keyphrase *"Free Credit Reports"*, heightened in the treated states relative to the control, suggesting awareness as well as consumers demand for reports.

This paper employs the aforementioned reduction in economic cost of credit reports in a difference-in-differences (DID) design to draw causal conclusions. The control group consists of the seven pre-FACTA states, and the treatment the states bordering the control states. The event (treatment) year is 2005. Further, by restricting the focus to a narrow geographic area encompassing *only* the counties at the border between the treated and control states, the empirical strategy aims to control for confounding effect of local economic conditions.

This empirical design mitigates the endogeneity in the treatment and control assignment. Recall that the treatment is assigned not by states' local laws, but by the federal law, FACTA, that was binding on all the states. So, the treated states did not *choose* to become treated. At the same time, the "control" is assigned to the pre-FACTA states due to their local laws that were already in force at the time of the event. In spite of this reassuring assignment scheme, the enactment of FACTA in 2003 could still be argued to be an endogenous response to the prevailing conditions. The circumstances of the enactment, however, suggest otherwise. As it happened, most of the FACTA provisions were not new, but were consolidated from another existing federal law, the *Fair Credit Reporting Act of 1970* (*FCRA*), to which FACTA added the free credit report provision. Owing to a sunset clause added to the FCRA in 1996, it was bound to expire in 2003. Hence, the upcoming expiration of FCRA in 2003, not the contemporary economic conditions, brought FACTA to fruition (Nott & Welborn, 2003). Note also that since most of the FACTA provisions, other than the free credit report provision, were already in place under the FCRA, the concern that they may create confounding effects is limited.

The outcome variables are analyzed at the census tracts level—a sub-county geographic area roughly encompassing a population of about 4,000. Comparing the outcomes in census tracts in the control and treated counties, the DID estimates suggest that lower economic costs resulted in an increase of 1–2 percentage points in the approval ratio and 13.8%–16.0% in mortgage applications. In dollar terms, increase in origination due to higher approval was about $5.5 billion, and due to the demand increase, about $38.1 billion, aggregated *only* across the treated bordering counties (not all counties from treated states). The increase in approval ratio is consistent with an improvement in borrower pool under a self-learning mechanism. The increase in applications indicates that creditworthiness under-estimators are more than over-estimators, though this interpretation is masked by other factors that are explored later.

Were the investment-seeking borrowers behind the increased demand, as was the case during the financial crisis that unfolded just three years after the event? Tests focusing on the owner-occupied and non-owner-occupied mortgages suggest otherwise. While the non-owner-occupied mortgages saw slight uptick of 1 percentage point as the fraction of total and successful applications, the bulk of increase seems to have come from occupancy-seeking borrowers.

If the increase in mortgage origination and approval ratio was a result of improved borrower pool, the defaults on mortgages should decrease, or it should at least not increase. Over a six-year period since inception, the event-year mortgages from the treated areas were less likely to be defaulted upon than pre-event year ones, after accounting for the trends from the control areas. Besides, this superior performance persisted through the 2008 financial crisis.

Heterogeneity in effects across consumer characteristics further helps to characterize the role economic costs of credit reports play in the mortgage market. The heterogeneity across consumer creditworthiness shows that approval ratio and mortgage applications increased more in the *ex-ante* high creditworthiness areas. This accords with the idea that the reports aid consumers in assessing their creditworthiness. Heterogeneity across borrower income reveals that while approval ratio increased for the low-income quartile borrowers, the number of applications did not. As lower income is associated with overestimation of one's creditworthiness (Perry, 2008), low-income quartile consumers are likely to correct for it by exiting, resulting in no significant increase in applications, but in higher approval ratio.

Finally, evidence in the support of self-learning channel is confirmed in multiple tests. *First,* while the credit-history denials decreased in the treated areas by 0.3 percentage points (though significant only in the *ex-ante* high rejection areas), debt-to-income denials did not. *Second,* withdrawal of in-process applications dropped by 0.9 percentage points in the treated areas. This accords with the idea that borrowers with the knowledge of their creditworthiness likely face less uncertainty over application acceptance, apply to fewer lenders (confident pre-application search), and save the cost of multiple applications. *Finally,* among the originated mortgages, the fraction of first-time homebuyers increased by 1 percentage point in the treated areas, clearly showing *entry* of borrowers.

Finally, a supply-driven increase in origination seems unlikely because in the treated areas mortgage interest rates increased, high-lenders-density areas did not see more origination or approvals vis-à-vis low-lender-density areas, private securitization did not increase, and subprime origination was 30 times smaller than prime.

All in all, imprecise creditworthiness information among consumers likely play a role in frequent mortgage denials due to credit history and in *discouraging* potential borrowers from applying for mortgage. Reducing consumers' economic costs of credit reports improves consumer outcomes on both these fronts, leading to improved borrower pool and higher demand for mortgage. This paper draws conclusions using the mortgage decisions, yet the findings may apply to all types of credit decisions. Moreover, as these findings are causal in nature, a policy intervention aimed at educating consumers of their creditworthiness may yield similar results.

This paper primarily relates to the literature on effects of information provision on credit market participants. This is the first paper to show that consumers' economic costs of credit reports, when lowered, leads to improved mortgage market outcomes in a way consistent with improved borrower pool. In a field experiment, Homonoff, O'Brien, and Sussman (2019) find that borrowers are less likely to default when provided with information on their FICO® scores. Similarly, bank customers who enroll in free FICO scores program are less likely to default, reduce credit utilization, and increase credit card spending (Mikhed, 2015). Moreover, when financial product disclosure is increased, sophisticated borrowers default less, and when financial product is standardized, unsophisticated borrowers default less (Kulkarni, Truffa, & Iberti, 2018). Also, consumers increase mortgage borrowing after bankruptcy flags from their credit reports are removed (Dobbie, Goldsmith-Pinkham, Mahoney, & Song, 2016). It also leads to aggregate welfare loss, cheaper credit for poorer defaulters, and expensive for poorer non-defaulters (Liberman, Neilson, Opazo, & Zimmerman, 2018).

This paper also relates to the extensive literature on financial literacy. Low financial literacy leads to detrimental economic outcomes, such as high mortgage delinquency and foreclosure (Gerardi, Goette, & Meier, 2010), poor mortgage choice (Moore, 2003), and large debt (Lusardi & Tufano, 2009; Stango & Zinman, 2009). Further, field experiments reveal that less financially literate distressed borrowers benefit less from loan-modification contracts (Hundtofte, 2017), and educational intervention improves consumers financial product purchases (Balakina, Balasubramaniam, Dimri, & Sane, 2020). This paper shows that providing free credit reports results in increased mortgage demand and lower defaults.

The rest of the paper is organized as follows. Section (1) describes the U.S. laws related to consumers' access to credit reports, Section (2) presents the research design, and Section (3) describes the data this paper uses. Section (4) discusses the main results, and Section (5) contains supplementary results that aid interpretation of the main findings. Finally, Section (6) concludes the paper.

1 U.S. Laws Governing Consumers' Access to Credit Reports

The FCRA governed consumer credit information-related laws before FACTA. Even under the FCRA, consumers had the right to see the contents of their credit reports, except for the credit score, under specific and restrictive provisions. For example, a consumer could receive a free report if he/she made a request within 60 days after receiving a notice of an *adverse action* taken against him or her on the basis of the information in the credit report (Avery, Calem, Canner, & Bostic, 2003).[3] The 1992 amendment to the FCRA mandated that the cost of disclosure of credit information should be reasonable, while that in 1996 capped the cost at $8. The latter also provisioned the law to lapse in 2003.

Even though the FCRA allowed free credit reports at the federal level under specific circumstances, consumers rarely proactively requested their credit report for own use. Out of approximately 1 billion credit reports generated annually, only 1.6% were disclosed to consumers (Avery et al., 2004). Of these 1.6%, only 5.25% were proactively requested by consumers, while 94.75% were disclosed to consumers under the FCRA provisions mentioned earlier (Nott & Welborn, 2003).[4] Thus, only 0.084% of all credit reports generated were disclosed to consumers as a result of their own request.

The pre-FACTA states had enacted local laws to allow the residents free credit reports (Panel

[3] An adverse action notice can be sent to a consumer by the *user* of a consumer report (e.g. banks, financial institutions, insurance firms) or a debt collection agency affiliated with the CRA stating that the consumer's credit rating may be or has been adversely affected. Consumers can receive credit report free of charge once in 12 months if he or she makes a request to the CRA for the credit report and certifies that: (A) she/he is unemployed and intends to apply for employment in the 60 day period beginning on the date on which the certification is made; (B) she/he is a recipient of public welfare assistance; (C) she/he has reason to believe that the file on the consumer at the agency contains inaccurate information due to fraud.

[4] Breakdown of the 94.75% credit reports disclosed under FCRA provisions: 84% due to *adverse action*; 11.5% due to fraud claim; 0.4% due to unemployment, 0.1% due to consumer being on public assistance.

A of Figure II). For example, through The State of Colorado SENATE BILL 133, the state enacted its free credit report law on April 21, 1997. Section 4, paragraph (E) of this bill added the following to Title 12, Article 14.3-104 of the Colorado Statute:

> *(E): Each consumer reporting agency shall, upon request of a consumer, provide the consumer with one disclosure copy of his or her file per year at no charge whether or not the consumer has made the request in response to the notification required in paragraph (a) of this subsection.*

The 1996 amendment to the FCRA added the sunset clause to it, mandating it to expire in 2003. In order to make its provisions permanent, the FACTA was enacted on December 4, 2003, with a new key provision added: free annual disclosure of credit reports to consumers by each of the three national credit reporting agencies.

2 Empirical Research Design

As discussed, this paper uses a DID setting in which the seven pre-FACTA states constitute the control group, and the states bordering these states the treatment group. Further, the final sample consists of only the counties lying at the border of the control states and surrounding treatment states. Panel A of Figure (II) depicts the pre-FACTA states and the years they adopted local free credit report laws, and Panel B shows the treatment and control states. Figure (III) shows the counties at the border of these states. The event is year 2005 when www.annualcreditreport.com was established to distribute the free credit reports.[5]

Main regressions use the following two-way fixed-effects estimator equation:

$$Y_{icsjt} = \beta_0 + \beta_1 \text{Treatment}_{icsj} \times \text{Post}_t + \delta \times \text{Economic controls} + \alpha_i + \gamma_{jt} + \varepsilon_{icsjt} \qquad (1)$$

where Y_{icsjt} is the outcome variable for a census tract i from a county c lying at the border between treatment state s and control state j. Recall that there are seven control states, thus j ranges from one to seven. t indexes years 2000–2008 and $Post_t$ takes value 0 for year $t < 2005$ and value 1 for year $t \geq 2005$. $Treatment_{icsj}$ is 0 for all the census tracts i in counties c from control states j, and is 1 for those from treatment states s. Standard errors are clustered at the county level to provide for correlation in error terms for the observations from census tracts belonging

[5] The website was rolled-out in four phases from Dec 2004 to Jan 2005. Phase I rollout was on Dec 1, 2004 in 13 states: *AK, AZ, CA, CO, HI, ID, MT, NV, NM, OR, UT, WA, and WY*. Phase II rollout was on Jan 3, 2005 in 12 states: *IL, IN, IA, KS, MI, MN, MO, NE, ND, OH, SD, and WI*. Phase III rollout was on Jan 6, 2005 in 11 states: *AL, AR, FL, GA, KY, LA, MS, OK, SC, TN, and TX*. Phase IV rollout was on Jan 9, 2005 in the remaining 14 states and DC.

to the same county.

Economic controls include a host of time-varying, county- and state-level variables that capture local economic and credit conditions. These are—annual growth rate of county's income per capita, county's aggregate employment and state's gross domestic product (GDP), and number of mortgage lenders (in log) in a census tract.

α_i represents "*Census Tract*" fixed effects, the first of the two-way fixed effects. These account for any time-invariant differences across census tracts at a highly granular geographic area encompassing just about 4,000 population. As a census tract covers a smaller geographic area than a county or a state does, these fixed effects also control for any differences in state policies, including recourse or non-recourse status, electoral landscape etc.

$\gamma_{j,t}$ represents "*Border* $\times$ *Year*" fixed effects, the second of two-way fixed effects. Here j refers to the border of control state j. Owing to the interaction of border identifier with year, these allow for a flexible and robust way to account for any time-varying regional economic shocks that may affect bordering states across different years.[6] Thus, together with the time-varying economic controls, these fixed effects are able to reasonably account for the confounding effects of other local economic trends on the mortgage market outcomes, and allowing us to cleanly estimate the desired treatment effect.

The key assumption for the DID estimate in Equation (1) to be causal is that the treated states would have had similar trends as the control states in the absence of the treatment (parallel-trends assumption). Though it is unverifiable, Panel A of Figure (IV) plots the trend of mean approval ratio across the two groups before the event, and they seem to be parallel.

Furthermore, Panel B of Figure (IV) plots the coefficients (β_k) from regression of *Approval Ratio* according to the following specification:

$$\text{Y}_{icsjt} = \beta_0 + \sum_{k=T-3}^{T-1} \beta_k \,\text{Treatment}_{icsj} \times \text{Event}_k + \sum_{k=T+1}^{T+4} \beta_k \,\text{Treatment}_{icsj} \times \text{Event}_k + \alpha_i + \gamma_{j,t} + \varepsilon_{icsjt}$$

where $\text{Event}_k = 1$ if $t = T - k$. $\text{Event}_k = 0$ if $t \neq T - k, k = \{-3, 4\}$. T = Event year 2005. These coefficients represent the difference in approval ratio for the two groups over the years relative to the pre-event year (2004). We see from the plot in Panel B that for the most part no significant difference exists between the treated and control census tracts before the event, but the differ-

[6] Consider a control state CO. All census tracts from the counties at the border between CO and the surrounding states—WY, UT, AZ, NM, OK, KS, and NE—take the same value (j), thus are grouped as one unit. Thus, the fixed effects only utilize the variation in the outcome variables *within* each such contiguous geographic areas.

ence becomes significant afterwards. Overall, the two plots in Figure (IV) together provide reasonable assurance that the parallel trend assumption is satisfied in the current setting.

Another underlying assumption in this setting is that higher cost of credit reports leads to its lower usage. Given that a report costs only around USD 8, one may wonder whether this is large enough to keep consumers from obtaining their credit reports. Note that the economic costs of accessing credit reports may still be high for consumers, especially for financially less knowledgeable ones, as economic costs include the knowledge whether such reports exist, where to obtain these from, and what role do they play in a mortgage process.

The credit reports usage data do support the above assumption. The usage were low in general, and were even lower in the treatment states before FACTA was enacted. *First*, recall that only 0.084% of about 1 billion credit reports issued annually are consumer-requested. Assuming each of these reports to be requested by a separate consumer, 0.84 million consumers accessed their credit reports, while the annual mortgage applications around 2004 numbered about 16.8 million. So, even if all these reports were requested by mortgage-seeking borrowers, *less than 5%* of the mortgage applicants had checked their credit reports.[7] *Second*, the data from testimony in the U.S. senate hearing confirms the stark difference in the usage of credit reports in the pre-FACTA states and the rest. Relative to the national average, its usage was 250% higher in GA, 204% higher in MD, 153% higher in CO, 35% higher in NJ, and 25% higher in MA (U.S. Senate. 108th Congress, 2004a).[8]. *Third*, the SCE data suggest that even a decade after credit reports became free, consumers who report they are unaware of their credit score is ~12%, and those who checked their report either never or at least more than two years ago is ~20% (see the Baseline Results on survey evidence for details).

We further need to ensure that the experiment—the establishment of the website www.annualcreditreport.com under FACTA—was a salient event, and consumers showed interest in free credit reports around the experiment. The examination of the Search Interest data from Google Trends supports this. *First*, the search interest for the key phrase *Free Credit Report* heightened in Jan 2005, coinciding perfectly with the establishment of the

[7] This may be why the Federal Reserve Board actively encourage consumers to check their credit reports—*It can be especially helpful to see a copy of your credit report before you apply for, say a car loan, a mortgage, or a credit card*(Federal Reserve Bank of Philadelphia, 2015).

[8] Not only did the pre-FACTA states have higher usage of credit reports, but they also seem to have enjoyed better consumer credit environments: the rate of consumer bankruptcies was the lowest (second lowest) nationally in Vermont (Massachusetts) in 2002, and the interest rate on a conventional mortgage in Vermont and Massachusetts was below the national median (U.S. Senate. 108th Congress, 2004b).

website (Panel A of Figure (V)).[9] *Second,* the plot of differential search interest across the treated and control states using the Interest-by-subregion Google Trends data suggests that consumer interest in free credit reports heightened in the treated states in the year of website's establishment. Panel B of Figure (V) shows the mean of the popularity rank for the two groups each year from 2003 to 2008. We see that the keyphrase was equally popular in both the treatment and control states in the pre-event year 2004, but it became more popular in the treatment states in 2005.[10] Also, some anecdotal evidence suggest that the website issued about 52 million credit reports to consumers in the first two years (Wikipedia, n.d.).

Finally, it is worth pointing out that even though the credit reports issued under FACTA do not contain the numerical credit score, consumers are not left in the dark about it. In fact, the website actively asks consumers if they wish to retrieve their scores from any of the three CRAs, and provides a link to corresponding CRA's website for further steps. It is reasonable to expect that the economic costs of accessing credit reports and credit scores dramatically reduced to just a few clicks, which in the earlier system involved calling a CRA requesting for reports, and waiting for it to arrive through mail. In fact, in 2000, the three CRA's settled a lawsuit by Federal Trade Commission (FTC) for blocking calls of millions of consumers who wanted to discuss the content of their credit reports (Federal Trade Commission, January 13, 2000).

In this empirical strategy, it may appear to be a concern that the control group is already treated. However, a DID setting only requires that during the sample period the control group remain free from interventions that may affect the outcome of interest. Hence, the specification in Equation (1) estimates the treatment effects as intended.

3 Data

The key data used in this paper come from the U.S. mortgage data available under the *Home Mortgage Disclosure Act of 1975* (HMDA). HMDA is the most comprehensive source of mortgages application level data in the U.S.. HMDA data provide application-level details on appli-

[9] Search interest, provided by Google, is a standardized index representing the degree of searches for the keyword(s) on Google at any time relative to the highest point during the period of the analysis, over a given region(U.S. in the present case). In the time series, a value of 100 is the peak popularity for the term. A value of 50 means that the term is half as popular. In the cross-section, a value of 100 represents the location with the highest popularity of the keyword as a fraction of total searches in that location. A value of 50 indicates a location that is half as popular. A score of 0 means there were not enough data for this term. Google Trends data start from January 2004.

[10] From this plot, it may appear that control states' popularity after the event decreased. However, this is the artifact of the way Google calculates this popularity measure. This is essentially a rank measure assigning ranks from 0 to 100 to states every year, 100 being the most popular. So, increase in rank of some states decreases rank of others.

cants' race and gender, income, loan amount, the financial institution handling the mortgage application, outcome of the application, and geographic location of the property at the census tract level.[11] The period of the study is from 2000 to 2008. Since the experiment occurs at the beginning of 2005, extending the sample until 2008 allows for enough post-experiment observations. The data contain 190.4 million mortgage applications over the sample period.[12]

The application-level data from HMDA are aggregated to "$CensusTract \times Year$" panel in several steps. First, all observations that have state, county or census tract information missing or "NA", or state Federal Information Processing Standard (FIPS) code as "0", "00" or "0 " are dropped (2.5% of the observations), leaving 185.6 million mortgages with an identifiable county. Then, observations on three action types are removed: the covered loan purchased by the financial institutions from other institutions (18.80%), as these are not borrower initiated; pre-approval requests denied by financial institutions (0.01%), as these data were included in HMDA reporting only from 2004; and the pre-approval requests approved by the financial institutions but not accepted by the applicants, as this data, too, were included in HMDA only from 2004, and its reporting is not mandatory (0.025%). This leaves 150.7 million applications belonging to 77,526 unique census tracts (603,849 "$Census\ Tract \times Year$" observations). Finally, Census Bureau (n.d.) data facilitate the identification of tracts belonging to the bordering counties of the treated and control states (Figure III). These 12,044 census tracts—7,054 treated and 4,990 control—constitute the final sample of 90,353 "$Census\ Tract \times Year$" observations.

Though the coverage of mortgages in the HMDA data is the largest, these lack some key application-level information, such as credit score. The data from the two GSEs—the Federal National Mortgage Agency (Fannie) and the Federal National Home Loan Mortgage Corporation (Freddie)—contain mortgage pricing related information including debt-to-income ratio, credit score, first-time homebuyer flag, investment purpose and more. The GSE data pertain to the 30-year fixed rate single family mortgages, the most popular mortgage type in the U.S.. Over the sample period, this data contain 33 million observations. Unfortunately, the property locations in this data are available at the coarser 3-digit zip code (henceforth, zip3) and state level. The crosswalk files provided by the U.S. Department of Housing facilitate mapping the

[11] Until 2003, the census tracts in HMDA are from the Census 1990 definition, while those from the 2004 onward are from Census 2000 definition. To make the geographic area comparable across these two time-periods, I scale the census tract-level variables from 2000 to 2003 using the ratio of population residing in the 1990 definition of the census tract to that in the 2000 definition of the census tract using data from Census Bureau (2006).

[12] The sample includes mortgages for three purposes—home purchase, refinance, and home improvement, and all loan types (conventional loans, loans guaranteed by Veteran Administration (VA) and Farm Service Agency (FSA)/Rural Housing Administration (RHS), and loans insured by Federal Housing Administration (FHA)).

zip3-state level information to the bordering counties included in the sample.[13] Aggregation these individual mortgages to zip3-state level yields 225 unique zip3-states and 7,711 "$Zip3$-$State \times Quarter$" observations.

The third important piece of data regarding the financial information of commercial banks come from "Call Reports" (FFIEC Forms 031/041). Matching the mortgage lenders in the HMDA data with the RSSD ID of commercial banks in the Call Reports requires several steps. For each mortgage application, HMDA data provide an agency code (lender's regulator) and a respondent ID, the combination of which is utilized to identify respective banks.[14] As some mortgage lenders are not banks, but the affiliates of commercial banks, they are linked using their parent entities (available in the HMDA Ultimate Panel data). If both a HMDA reporter and its parent entity yielded a successful match, the parent's match is kept. HMDA started providing RSSD ID for the reporters since 2004, so instead of agency code and respondent ID, this identifier is used for matching of data from subsequent years.

This paper relies on a few more data sources. The SCE Credit Access Survey, a triannual internet-based rotating panel survey administered since 2013 by the Federal Reserve Bank of New York, provides rich detail on consumers' credit-related expectations (Survey of Consumer Expectations, 2013–2020). FRBNY and Equifax (n.d.) provide data on subprime population in a county. The annual survey of County Business Patterns (CBP) provides information on county-level employment Census Bureau (2000–2008). Also, the zip code level variables of CBP are mapped to census tracts level using (Missouri Census Data Center, 2010). Finally, data on state level economic conditions come from the Bureau of Economic Analysis, and population characteristics data at census-tract level are from Census 2000 (Manson, Schroeder, Van Riper, & Ruggles, 2019).

The key outcome variables of interest are the number of mortgage applications per 1000 adults in a census tract (*scaled applications*) and approval ratio. Approval ratio is the ratio of the number of successful applications (identified by action type "1" in the HMDA dataset) to the number of total applications in a census tract. Other variables of interest are the fraction of total applications denied for credit history or debt-to-income ratio and the fraction of total

[13] Areas delimited by 3-digit zip codes do not align with the county borders. Hence, to identify the 3-digit zip codes that lie along the county borders, I use the 2010 Q1 version of the 5-digit zip codes-to-county crosswalk file from Office of Policy Development and Research (n.d.). I remove all such 3-digit zip codes for which none of the underlying 5-digit zip codes lie within the bordering counties.

[14] A respondent ID equals the Federal Deposit Insurance Corporation (FDIC) Certificate ID if the lender's regulator is the FDIC and the Office of the Comptroller of the Currency (OCC) charter number if the OCC.

applications withdrawn by applicants while still under processing.

Summary Statistics

Panel A of Table (I) shows the summary statistics for the key variables over the sample period. We see that the treated census tracts have fewer scaled applications, lower mortgage approval ratio, and higher denials related to credit history and debt-to-income ratio.[15]

Panel B of Table (I) shows the comparison of the treatment and control groups in the pre-treatment period using t-test for difference in mean, the p-value for which are also shown. Results from the t-tests suggest that the control and treated census tracts differ in pre-treatment years in terms of mortgage-related variables, but are *similar* in the state- and county-level economic characteristics. The similarities in economic characteristics of treated and control areas support the comparison of outcomes across the two groups, whereas the differences in mortgage-related outcomes raise the concern that these groups may also differ on some unobserved characteristics, potentially causing an endogeneity issue. However, as a DID setting can accommodate pre-existing differences between the treatment and control subjects so long as they satisfy the *parallel-trends* assumption, the concern stands mitigated.

4 Results

This section first shows the findings on credit reports usage and discouraged borrowers. The discussion of baseline results regarding effects of free credit reports on mortgage demand, approval ratio, house prices, and default rates follows. Then, results highlighting the self-learning mechanism and the role of supply- and demand-side factors are discussed.

§A Baseline Results

Survey evidence on the Credit Reports Usage and Discouraged Borrowers

The SCE Credit Access Survey contains questions related to usage of credit reports and scores, and also includes questions related to *discouraged borrowers* phenomenon. It is possible to make robust inferences about the prevalence of these issues by utilizing the rotating panel design of

[15] We see that the four ratios—the approval ratio, the two denial ratios, and the withdrawal ratio—do not sum to one. There are three reasons for this. First, the reporting of the reason for denial is not mandatory under HMDA regulations; hence an application may be recorded as denied without any stated reason (70.81% of denied applications have at least one stated denial reason). Second, denial reasons could be other than credit history or debt-to-income ratio. Third, an application might be denied for multiple reasons.

the survey. Specifically, "Year-month" level fixed effects and standard errors clustered at the same level account for any year-specific shocks. Also, sampling weights of the survey makes it feasible to draw conclusions about the U.S. population.

Table (II) the extent of these issues in the U.S. population using the SCE survey data. Column (1) shows the regression on a constant of the dummy indicator of whether a respondent has never checked or requested credit report; an estimated 8% of the population falls into this category. Similarly, the indicator variable in column (2) is 1 if respondent has either never checked, or checked it at least more than two years ago (infrequent checkers); a staggering 20% of the population belong to this category. Finally, in column (3), the indicator variable is 1 if a respondent reports that he/she does not know his/her credit score; these account for 12% of the population.

The extent of *discouraged borrowers* can be estimated by focusing on those who respond "I don't think I would get approved" as the reason for the respondent is not likely to apply for mortgage/home based loan or refinance. Column (4) shows that among those who were in the coming 12 months very/somewhat unlikely to apply for mortgage/home-based loans or refinance, or those who assigned less than 10% probability to these actions, the proportion of discouraged borrowers is 13%.

The survey also provides insights into the link between usage of credit reports and discouraged borrowers phenomenon. In column (5) and (6), the indicator for discouraged borrowers is regressed on the credit report usage or unawareness about credit score. The results suggest that infrequent checkers and those unaware of their credit score are, respectively, 3% and 5% more likely to report as being discouraged.

Taken together, the analysis of the SCE data highlights the issue surrounding imprecise information about one's own creditworthiness. A significant proportion of consumers: (i) do not check their reports; (ii) are unaware of their credit score; and (iii) those who are infrequent checkers or unaware of their credit reports are *more likely* to be discouraged from applying for mortgage.

Effect on mortgage approval ratio, applications, and house prices

The results on the effects of lower economic costs of credit reports on mortgage approval ratio, number of (scaled) applications and house prices follow next. Each outcome variable is measured at the census tract level. The regression specification used here is from Equation (1); the coefficient of interest is "$Treat \times Post$", which estimates the change in the outcome variable in the treated areas relative to the control after the event. All specifications include *Census Tract* and "$Border \times Year$" fixed effects.

Columns (1) and (2) of Table (III) show the regression results for approval ratio; the former is plain DID without any control variable, the latter is with controls for local economic conditions included. The controls are annual growth rate of county income per capita, county aggregate employment, and state GDP. Coefficients on "$Treat \times Post$" suggest that the ratio increased by about 1–2 percentage points in the treated tracts. In real terms, keeping the number of applications in the treated areas at the pre-event level, a 2 percentage points increase in approval ratio corresponds to about \$5.5 billion more successful mortgages, aggregated across the treated bordering counties. The amount of increase may seem trivial at first, as the mortgage approval ratios are commonly believed to be high, in the upwards of 80%, but in the counties of interest here, the average ratio in the pre-event period is just 52%.[16]

Recall that lower economic costs of accessing credit reports does not affect the information lenders have on borrowers. Their access to the reports did not change due to the FACTA, neither the extent and scope of information contained in these reports. In fact, any other factors bearing relevance to mortgage decisions, such as borrower's income, employment, collateralizability of their assets etc., are not the subject of the act. The most likely change the event may bring is that borrowers become more aware of their credit history and other information that lenders use to evaluate them on. This may lead to new entry by borrowers (creditworthy borrowers to prime lender and/or less-creditworthy borrowers to sub-prime lenders) and better matching with lenders among existing borrowers; all of which predict that approval ratio would increase. The later sections evaluate some of these claims rigorously.

Columns (3) and (4) show the results for scaled applications. The estimates suggest that the scaled applications increased in the treated tracts by 13.3–15.4, a 13.8–16.0% increase over the pre-treatment average of 96.6.[17] In real terms, keeping the approval ratio in the treated areas at the pre-event level, the increase in applications roughly translates to \$38.1 billion increase

[16] A 2 percentage point increase in approval ratio is equivalent to ~5.22 more successful applications per treated tract (96.6 applications per 1000 adults in the pre-treatment period $\times$ 0.02 $\times$ 2.7 thousand adults per treated tract), about 36,827 more successful applications across the treated bordering counties (2.62 applications $\times$ 7,054 treated tracts), or a ~\$5.5 billion increase in mortgage origination across all bordering treated tracts (36,827 $\times$ \$150,597 average mortgage amount per application).

[17] The average mortgage size in treated tracts in the pre-treatment period was about \$150,597. Thus the demand for mortgage credit increased by about \$2.0 million per 1000 adults per census tract (\$150,597 $\times$ 13.29), by about \$5.4 million per treated census tract (\$2 million $\times$ 2.7 thousand adults per census tract), or, by about \$38.1 billion across the treated border counties (\$5.4 million $\times$ 7,054 treated tracts).

in mortgages, aggregated across the treated bordering counties. The increase in applications suggests that on average consumers tend to underestimate their creditworthiness when it comes to mortgage borrowing, a finding in contrast to other financial decision-making settings (Perry, 2008).

Coefficients on "$Treat \times Post$" in columns (5) and (6) quantify the changes in the growth rate of house prices. The regressions use census tract-level house price index from Bogin, Doerner, and Larson (2016), the index that starts in 2000. The coefficients suggest that the growth rate of house prices in the treated areas increased relative to the control by 1.7–1.8 percentage points after the event, though the estimates are statistically significant only at the 10% level. This finding is in line with that of Di Maggio and Kermani (2017): in their sample, house price growth rates increased by 3.3 percentage points following a 10% increase in mortgage origination.

Having discussed the changes in the mortgage approval ratio, applications and house prices in response to lower economic costs of credit reports, it is worthwhile to point out a lingering concern that the post-event sample period of this study is 2005–2008. The housing market was volatile during this period, and mortgage supply in the U.S. had started to shrink since 2005. Given the timing of the natural experiment, it would be injudicious to claim that the effect size estimated above are immune to these changes. A respite here is that the difference-in-differences design ameliorates the issue to the extent that the market-wide forces evenly affect the neighboring counties across states. In the robustness section, this issue is further explored.

It is also understood that one reason for the financial crisis was excessive mortgage borrowing by borrowers without means, often for investment motives vis-à-vis occupancy motives. Whether such borrowers are behind the increased origination in the current setting is important to examine. Table (IV) examines this by focusing on the changes in owner-occupied and non-owner-occupied mortgage category using the same DID specification. The dependent variable in columns (1) and (2) is the number of mortgage applications for the former category, and columns (3) and (4), for the latter. We see that the applications increased dramatically and significantly only for the owner-occupied category in the treated areas vis-à-vis the control, but not for the latter category. Investigating further, columns (5) and (6) examine the composition of non-owner-occupied category as a fraction of total applications, and columns (7) and (8) as a fraction of successful applications. The coefficients in these four columns suggest a modest 1 percentage point increase in non-occupancy mortgages at both the application and origination stage. All in all, the investment-motivated demand does show a slight uptick, but does not appear to be a dominant reason behind the robust 15% increase in the mortgage applications.

Mortgage defaults

We saw that lower economic cost of credit reports resulted in higher mortgage origination. The question then is, whether this would follow heightened, or diminished, mortgage defaults? If the rise in origination was a result of an improved borrower pool, mortgage defaults would decrease, or at least not increase. If the creditworthiness in the borrower pool deteriorated on average, but origination still increased due to subprime lending, the defaults would increase.

The GSE data, which contain a subset of the HMDA mortgages, allow us to compare the mortgage default rates. To compare the default rates, the paper first defines *monthly default rate* for mortgages originated in a given vintage year, 2004 or 2005, as a function of its age, measured in months since origination, denoted respective as $\text{Def}_{2004,age}$ and $\text{Def}_{2005,age}$. The monthly default rate at a given age is the ratio of number of mortgages that misses a scheduled payment by 30–59 days for the first time at that age to all the mortgages originated in the given vintage year. This rate is separately calculated for the treated and control areas. *Adjusted default rate*, defined below, then measures the differential probability to default at given age:

$$\text{Adjusted default rate}_{age} = (\text{Def}_{2005,age} - \text{Def}_{2004,age})_{treated} - (\text{Def}_{2005,age} - \text{Def}_{2004,age})_{control} \quad (2)$$

A negative adjusted default rate would imply that the mortgages from the treated areas are less likely to be defaulted upon than those from the control areas at a given age. The plot of adjusted default rate with age in Figure (VII) reveals that for most of the months within six years after origination, mortgages in the treated areas were *less likely* to be defaulted upon than those from the control areas. The mean value of the adjusted default rate is -0.012 percentage points (*p-value = 0.000*). Surprisingly, around the age of 48 months, which corresponds to the bust years after the financial crisis (48 months after 2005 is 2009), it is even more negative. This suggests that the performance of the treated mortgages was superior even during and after the crisis.[18] All in all, just as higher approval ratio is consistent with an improved *ex-ante* borrower pool, so too is lower *ex-post* defaults.

[18] It may seem counter-intuitive at first that the newly treated areas could perform even better (less default) than the nearby areas which already had free credit reports. This is plausible provided that the event brings out changes in the borrower pool in the treated areas, but not in the control areas. Improved performance in control areas is coming from improved pool in the post-event period relative to the worst pool from pre-event period, whereas in control areas, there is no improvement—(good) pool in post-event period is the same as in the pre-event period.

§B Characterizing the Effect

Characterization of the consumers, and the areas, that are more likely to benefit from easier access to credit reports allows more precise understanding of who suffers from the information frictions on creditworthiness, and the conclusions may offer insights that could be used for policy targeting. Examining the effect heterogeneity across pre-event characteristics of consumer and areas allows for such characterization. Specifically, the heterogeneity across consumer creditworthiness and income are examined.

§B.1 Heterogeneous effects by creditworthiness of borrowers

Given that creditworthy borrowers are more likely to be granted mortgage credit, if credit reports aid consumers in assessing creditworthiness, easier access to the reports should lead to *more* increase in applications and approvals in areas where fraction of creditworthy consumers was higher before the event. This is because under the self-learning mechanism, the exit would be larger and the entry smaller in the subprime areas vis-à-vis the prime areas.

To test this, a county is classified as having high creditworthiness if its subprime population fraction is less than the *regional mean* before the event.[19] The year 1999 is chosen as the classification year following Mian and Sufi (2009), who suggest that such classifications should be done at a time well before the start of the housing boom as creditworthiness of an area endogenously evolves with the housing boom. The earliest year the data on county subprime fraction, (FRBNY & Equifax, n.d.), is publicly available is 1999.

Panel A of Table (V) shows the results of regressing scaled applications and approval ratio separately using regression Equation (1) for counties with high and low creditworthiness. Columns (1) and (2) show that scaled application increased by 16.7–18.1 (17.3–18.7%), while columns (3) and (4) show that approval ratio increased by 2 percentage points in treated *ex-ante* high-creditworthiness counties relative to control counties with similar creditworthiness. The contrast emerges in the coefficients in columns (5) through (8), which estimate the effects in *ex-ante* low-creditworthiness treated areas with respect to control areas of similar creditworthiness. The increase is far smaller and barely statistically significant: 8.5–10.6 (8.8%–11%)

[19] The steps to calculate *regional mean* are as follows. A region is defined as the area encompassing a control (pre-FACTA) state and all the surrounding states. Consider the control state Colorado (CO) and all the surrounding treatment states. Regional mean for this region is the average rejection rate for the census tracts in all the counties at the border between CO and WY, UT, AZ, NM, OK, KS and NE. Regional means of rejection rates for all seven control states are calculated in this way, and a census tract is then classified as a "High rejection tract" if its rejection rate is more than the regional mean in 2004.

for scaled application and 1 percentage point for approval ratio. The results support the self-learning mechanism and suggest that creditworthy borrowers are more likely to benefit from easier access to credit reports.

As an alternative to the county-level creditworthiness measure, which could be noisy, number of payday lenders could be used as a geographically more precise proxy. Such lenders tend to operate in subprime areas (Prager, 2009), and their locations are available at the 5-digit zip-code level from Census Bureau (2000–2008), which can be scaled to the census-tract level using crosswalk files. However, among the states in the sample, only Colorado and all the bordering states allow unrestricted payday lending activities, while the rest restrict it to varying degrees (Prager (2009) and Bhutta (2014)). The next set of regressions reexamine the heterogeneity along consumer creditworthiness using this proxy, keeping only the counties at the border between CO and the surrounding treated states (WY, UT, AZ, NM, OK, KS and NE). Creditworthiness of a census tract is high if the number of payday lenders in the pre-event year 2004 is *less* than the mean across all census tracts within these counties.

Panel B of Table (V) shows the regression results. The difference across *ex-ante* high- and low-creditworthiness census tracts is sharper vis-à-vis the previous estimates for both the dependent variables: 68.4–72.8 (*p-value*$<$*.001*) vs. 43.3–43.7 (*p-value*$<$*.001*) for scaled applications and 6 percentage points (*p-value*$<$*.001*) vs. 2 (not statistically significant) for approval ratio. These results confirm the previous conclusions, only more strongly now.

§B.2 Heterogeneous effects by income level of borrowers

The effect of easier access to credit reports on number of applications for borrowers across the income levels may be different, because the consequences of a mortgage rejection are more severe for low-income borrowers. So upon learning creditworthiness more precisely, the likelihood of exiting the credit market or gravitating to subprime lenders vis-à-vis entering is higher for low income borrowers. In other words, more exits would occur among low income consumers than high income ones.

The above conclusion regarding exit can also be reached when we analyze the effect from the lens of over- and under-estimation. Perry (2008) finds that lower income is associated with higher propensity to overestimate one's creditworthiness. Thus to the extent that credit reports aid consumers in self-assessing their creditworthiness correctly, the downward revision of creditworthiness is more likely to occur for low-income consumers, making it rational for them to exit. Though exits are not directly observable in data, it can be inferred from differences in increase in applications across different income groups.

The approval ratios should increase for borrowers in all income groups, though it's magnitude may differ. As the marginal propensity to lend to high-income consumers is more (Agarwal, Chomsisengphet, Mahoney, & Stroebel, 2018), the opportunity for borrower pool to improve is smaller for them than for low-income groups. Thus the ratio is more likely to be high for low-income consumers.

These predictions are tested next. Income quartile cut-offs are calculated in-sample each year, and then applications in each quartile are aggregated at the census-tract level. Panel A of Table (VI) shows the results of regressing scaled applications for the income quartiles using Equation (1). The scaled applications did not increase significantly for lowest quartile, but it increased for other three, and the magnitude of increase is larger for higher quartiles. This finding is in line with the prediction; as exits are more likely to be a rational choice for low-income consumers, the rise in number of applications would be small for them. Coefficients in Panel B of the table estimate the changes in the approval ratios. We see that it increased statistically significantly only for the lowest income-quartile, again consistent with the intuition.

Increase in demand for mortgage after the economic costs of credit reports were reduced suggests that mortgage borrowers tend to underestimate their creditworthiness. This is in contrast with commonly understood belief that consumers tend to overestimate creditworthiness (Perry, 2008). The analysis here suggests that low-income consumers indeed tend to overestimate and the high-income consumers tend to underestimate.

§C Mechanism: Consumer Self-learning Channel

This section shows the results that speak to the self-learning channel.

§C.1 Reduction in subprime population fraction

An implication of better self-assessment of creditworthiness among consumers is that new prime consumers may enter the market and/or some (just-) subprime consumers may become prime. Thus, the subprime population fraction in the treated counties would decrease relative to the control counties. The Equifax data on subprime (credit score $\leq$660) population fraction at the county level facilitate this analysis.

Panel A of Figure (VI) plots the difference between the median of the subprime population percentage across treated counties and control counties calculated yearly from 2000 to 2009. We see that before the free credit report law, subprime population in the treated counties was increasing with respect to the control counties, but one year after the event, it starts to decrease. Using mean instead of median results in a similar patter, as Panel B of Figure (VI) shows. As previously mentioned, availability of credit reports by itself do not alter financial situation or creditworthiness of consumers, unless consumers use the reports and take actions. Thus, a decrease in subprime population in the treated counties directly points to improvements in the borrower pool consistent with increased consumer self-learning.

§C.2 Contraction in rejections due to credit history

Even though the information in the credit reports are relevant to the mortgage decision, just the fact that consumers could access the reports easily after the FACTA would not influence the approval ratios unless consumers access the reports, learn about their credit history and use the information in their credit decision. In other words, given that self-learning among consumers about their creditworthiness improves, the probability of rejection due to credit history should decrease and due to debt-to-income ratio should not.

The above predictions are tested by regressing the fraction of total applications rejected due to the respective reasons using the regression Equation (1). These regressions are estimated for full sample and over a sub-sample of the census tracts that had rejection rates higher than the *regional mean* in the pre-event year 2004, designated as high rejection areas (*regional mean* is defined in Footnote (19)). The rationale for testing over this sub-sample is that the areas where consumers were more often denied mortgages before the event are more likely to respond to lower economic costs of credit reports and value of information in the reports is high for them.

Table (VII) shows the results. In columns (1) through (4) we see that the fraction of applications denied due to credit history decreased by 0.3 percentage points in the treated tracts relative to the control, statistically significant only in the *ex-ante* high rejection rate areas (columns 3 and 4). The coefficients in columns (5) through (8) show that the debt-to-income ratio denials did not decrease statistically significantly. These results, despite being statistically weak, indicate that the rejection owing to different reasons changed in a manner consistent with consumers learning more about their credit history. A reason we could detect reduction in credit-history-related rejections only in the *ex-ante* high-rejection areas is that consumers who had faced more rejections are more likely to value the information. Besides, we cannot estimate a reduction in rejections due to a particular reason if mortgages are not likely to be rejected in the first place.

However, these results come with a caveat. HMDA does not mandate lenders to report rejection reasons, so if their reporting incentives changed with the event, we would incorrectly attributed lender-induced change to consumer self-learning. However, it seems unlikely as lenders reported a rejection reason in over 70.81% of the rejections. Also, incentives to report rejection reasons would need to change differently in a particular manner between treated and control tracts in the event year to bias the above results, an unlikely scenario.

§C.3 Confident search (Drop in in-process application withdrawals)

The propensity to withdraw an in-process application withdrawal too sheds light on the self-learning mechanism. Borrowers commonly initiate multiple formal mortgage applications at several lenders owing to the uncertainty over success of their applications and in a bid to secure better terms.[20] In essence, while they incur non-refundable application costs for each applications, in the end they finalize the mortgage with one lender—the one with higher certainty and/or better terms—and withdraw other applications. About 12% of the applications over the 2000–2008 period were withdrawn while being assessed by lenders (in-process withdrawals).[21]

If easier access to credit reports leads to better creditworthiness knowledge among consumers, they would be more certain over the approval probability and mortgage terms (confident applicant) and would apply at fewer lenders *ex-ante*, saving the costs of multiple applications. In aggregate, the fraction of in-process withdrawn applications should decrease.

Table (VIII) shows the result of regressing fraction of mortgages withdrawn while in process (*withdrawal ratio*), using Equation (1). The estimates show that it dropped by 0.9–0.11 percentage points in the treated tracts relative to the control. This decrease lends support to the self-learning mechanism.[22] A distinct advantage of inferring borrowers decision process from application withdrawals is that this action is purely a consumer decision, mostly outside of lender influence.

---

[20] Credit reporting agencies do not penalize multiple applications if they are made within a short window. Equifax (n.d.) puts it this way, "If you're shopping for a new auto or mortgage loan or a new utility provider, the multiple inquiries are generally counted as one inquiry for a given period of time. The length of this period may vary depending on the credit scoring model used, but it's typically from 14 to 45 days. This allows you to check at different lenders."

[21] Anecdotal evidence suggest that consumers tend to withdraw application when they find a better offer from other lenders (Reddit Forum, n.d.).

[22] In economic terms, the drop is equivalent to ~2.34 fewer in-process withdrawals per treated tract or ~16,513 fewer withdrawn applications aggregated over the treated border counties. At an average cost of ~$400 per withdrawn application, this represents ~U.S. $6.6 million saving in upfront mortgage application fees.

§C.4 New entry: Increase in first-time homebuyers

About 15% of households in the SCF survey of early 2000's and 13% of the respondents in the recent SCE surveys are *discouraged*. Self-learning mechanism predicts that to the extent consumers overestimate rejections and do not apply for credit, there would be entry of creditworthy borrowers into the credit markets. In the mortgage market, this prediction can be tested by examining the proportion of first-time homebuyers: a relative increase in their proportion in the treated areas would suggest increase in *entry*.

The GSE data contain information on first-time homebuyer, but the HMDA data do not. With reasonable approximations (see Footnote 13 for details), we can infer proportion of first-time homebuyers in the bordering counties of the states in the sample. A minor data issue is that about 6.71% of the observations in the GSE data pertaining to the bordering counties have first-time homebuyer status missing. Hence the outcome variable is defined in two ways: number of first-time homebuyers as the fraction of all mortgages or only the mortgages with valid first-time homebuyers information. The regression equation for this test is specified at the zip3-state level, different from the census tract-level specification in the earlier regressions:

$$Y_{zsjt} = \beta_0 + \beta_1 \text{Treatment}_{zsj} \times \text{Post}_t + \delta \times \text{Economic controls} + \alpha_{zs} + \gamma_{jt} + \varepsilon_{zsjt} \qquad (3)$$

Here, z indexes the areas delineated by a 3-digit zip code at the border of treated state s and control state j. α_{zs} is zip3-state fixed effects. γ_{jt} is the "Border×Quarter" fixed effects, similar to that in (1). The sample is limited to the zip3-state areas that come under the border counties of treated and control states.

Table (IX) shows the regression results. Coefficients in columns (1) through (4) show that the percentage of first-time homebuyers increased by 1 percentage point in the treatment areas relative to the control areas, in line with prediction that there should be new entry.[23]

[23] A concern is that the mortgage sample used in this result consists of those selected by the GSEs. However, as argued before, this selection would be an issue if GSEs' incentives to purchase first-time homebuyer mortgages relative to their overall purchase from the treated counties increased relative to the control in the year 2005. Such a time- and location-specific change seems improbable.

5 Supplementary Discussion

§A Is it the demand-side or supply-side factors that drive the results?

We saw that the treated areas saw increased mortgage demand and origination after the event. This paper argues that it is due to a demand-driven effect coming from improved self-learning among consumers. However, a supply-driven explanation is also plausible. In the natural experiment, lenders too were exposed to the knowledge that consumers' access to credit reports became easy and free, and they could increase the mortgage supply in response.

Several evidences we saw so far favor the demand-driven mechanism. Increase in applications and a decrease in in-process application withdrawals indicate a demand-driven mechanism as these decisions are consumer determined and are mostly independent of lenders' influence. Furthermore, under a demand-driven mechanism, the effects should be heterogeneous in relevant consumer characteristics, as was the case for creditworthiness and income. Besides, recall that propensity of lenders to extend credit to low-income borrowers is low (Agarwal et al., 2018), yet in the current setting we see that the approval ratio increased significantly for them vis-à-vis the high-income borrowers, making lender-driven effect unlikely.

Two further tests—equilibrium interest rates and heterogeneous effect by lender's density—may aid us in understanding the extent of the supply-side role behind the increased origination.

(i) Mortgage interest rates: The price at which lenders can sell *conforming* mortgages to the GSEs materially vary only across three dimensions, credit score, loan-to-value ratio (LTV), and loan type (Scharfstein & Sunderam, 2016,?).[24] Since the GSE sample in this paper include only 30-year fixed rate mortgages, only the first two are relevant. The residuals in a regression of interest rate on the first two attributes roughly reflect lenders pricing schedule independent of borrower risk and mortgage characteristics. Thus, including the "*Treat* $\times$ *Post*" in the above regression allows testing whether lenders lowered their pricing in the treated areas vis-à-vis control areas in a bid to increase mortgage origination, in which case the sign on the associated coefficient would be negative. Specifically, the prediction is tested using the following regression (this equation is similar to Equation (3), but it is specified at the loan level i):

$$\text{Interest Rate}_{izsjt} = \beta_0 + \beta_1 \text{Treatment}_{izsj} \times \text{Post}_t + \delta \times \text{Controls} + \alpha_{zs} + \gamma_{jt} + \varepsilon_{izsjt} \quad (4)$$

[24] The pricing schedule published of Fannie Mae is available at https://www.fanniemae.com/content/pricing/llpamatrix.pdf

Table (X) shows the results of the regression. Controls in column (1) are the two relevant pricing variables, credit score and CLTV (combined loan-to-value, it is loan-to-value ratio inclusive of all loans secured by a mortgaged property). In column (2), the following controls are added to make the specification more rigorous: debt-to-income ratio, number of units comprising the mortgaged property, and percentage of mortgage insurance coverage. The coefficient on "$Treat \times Post$" is 0.9–1.1 basis points, positive and significant. These estimates suggest that, if anything, lenders increased the risk-adjusted mortgage interest rates in the treated areas rather than lowering it, contradicting the idea of a supply-driven increase.[25]

(ii) Heterogeneous effects by density of lenders: If the increase in mortgage origination were driven by lenders, increase in mortgage origination and/or approval ratio would be greater in areas where the density of lenders is high. To examine this, a census tract is classified as having high lender density if the number of HMDA lenders per adult in the pre-event year 2004 in a census tract is more than the *regional mean* (Footnote (19) defines the mean).

Columns (1) through (4) of Table (XI) show the results of separately regressing dollar origination volume (in 1000 USD) per adult for low and high lender density areas. The regression specification is the same as Equation (1). The estimates are smaller in magnitude and have weaker statistical significance for high-density areas (columns 2 and 4) vis-à-vis low-density areas (specifications 1 and 3, respectively). Thus, high-lender-density treated areas saw smaller increase than low-lender-density treated areas after controlling for concurrent changes in the control areas with comparable lender density. t-test for the difference in the coefficient of "$Treat \times Post$" in high- and low- lender-density areas ($High - Low$) shows no statistical difference.

Then, columns (4) through (8) of the table repeats the analysis with approval ratio as the outcome variable. The results are similar—there is no statistical difference in the increase in the approval ratio in areas with a high or low lender density. Overall, these findings are inconsistent with the explanation that mortgage origination and approval ratio could have increased solely due to lenders increasing the supply of mortgages.

[25] The magnitude of the increase in the rates is small, potentially because of two reasons. First, interest rates on conforming (GSE-purchased) loans do not vary across regions or with dimensions other than FICO scores, loan-to-value ratio, and loan type (Hurst, Keys, Seru, & Vavra, 2016). Second, the supply of mortgages in the U.S. is highly elastic because of the large-scale purchases by the GSEs of conforming mortgages in the secondary market.

§B Did origination increase due to rise in private securitization?

An alternative explanation for increased mortgage origination is that higher commissions from private (non-government) securitization led lenders to increase the mortgage supply (Keys, Mukherjee, Seru, & Vig, 2010). If increased approval in the current context were due to private securitization, the fraction of originated mortgages being sold to non-government (private securitization) entities would increase in the treated areas.

Table (XII) examines the above prediction by employing Equation (1). The outcome variables are the fraction of total applications that lenders originated and (1) sold to non-government entities, (2) sold to the four GSEs (Fannie Mae, Freddie Mac, Ginnie Mae, and Farmer Mac), and (3) did not sell. Columns (1) and (2) show that there is no change (increase) in the fraction of mortgages sold to the private entities (non-GSEs), columns (3) and (4) show that the fraction of mortgages sold to the GSEs increased, and column (5) and (6) show that the fraction of unsold mortgages did not change either. Thus, there is no evidence of an increase in private securitization in the treated areas.

§C Did origination increase due to subprime lending? Credit score-based evidence

It may be argued that increased mortgage origination is due to an increase in the subprime credit (Mian & Sufi, 2009). Employing the comprehensive HMDA data and location-based proxies of creditworthiness, Table (V) already suggest that effect of free credit reports was stronger in the prime counties/census tracts than in the subprime. These proxies are informative and widely used (Di Maggio & Kermani, 2017; Mian & Sufi, 2009), but are imprecise. By restricting ourselves to the GSE sample, we can use precise application-level credit scores.

Table (XIII) shows the results of regressing separately the number of prime (credit score ≥ 620) and subprime *originated* mortgages in zip3-state areas using Equation (3). Columns (1) and (2) show that the number of prime mortgages increased by 308–312 in the treated zip3-state areas relative to similar control zip3-state areas, whereas columns (3) and (4) imply that subprime mortgages increased only by ~10 applications, which is 30 times smaller. In conclusion, the increased origination did not disproportionately go to subprime consumers. These estimates, however, are not directly comparable to the previous as the observation unit here is zip3-state, not census tracts which was the unit in the previous regressions.

These results come with the same selection issue that applied to the previous results utilizing the GSE data. Same argument as before allays this concern. In addition, before 2007 the GSEs sought to buy more subprime, not prime, mortgages to combat the housing bust (Elul, Gupta, & Musto, 2020), thus their changing incentives appears not to be a critical concern here.

§D Effect on banks

So far we focused on evaluating the effects on borrowers. The effect on banks is evaluated next, though with some caveats. First, commercial banks are not the dominant mortgage originators. Despite being 80% of mortgage lenders by number, banks accounted for just 37% of the mortgage lending in 2005, thus the conclusions drawn from studying banks may not be same for all lenders (Avery, Brevoort, & Canner, 2007). Second, many banks operate across states, so their treatment and control status in this natural experiment is continuous, rather than binary. The treatment intensity is proportional to bank's *ex-ante* mortgage activity in the treated and control states. Nonetheless, we can examine this question by classifying a bank as treated if in the pre-event year 2004 the ratio of mortgage it originated in control states to that in treated *and* control states combined is larger than the cross-sectional average. The regression equation is:

$$Y_{bt} = \beta_0 + \beta_1 \text{Treatment}_b \times \text{Post}_t + \delta \times \text{Bank controls}_{bt} + \alpha_l + \gamma_t + \varepsilon_{bt} \tag{5}$$

where Y_{bt} is the outcome variable (net interest margin (NIM), return on equity (RoE), and return on assets (RoA)). b indexes the banks; $Treatment$ is 1 if a bank is treated; $Post_t$ is 1 if year $\geq$ 2005; year t is year-quarter; α_i is bank fixed effects; γ_t year-quarter fixed effect; and $Bank\ controls$ include banks' log total assets, share of liquid assets to total assets, and cost of deposit.[26]

Regression results in Table (XIV) show that treated banks saw a 6 basis points increase in NIM (columns 1 and 2), a 0.74 percentage points increase in RoE (columns 3 and 4), and a 0.07 percentage points increase in RoA (columns 5 and 6). Therefore, keeping in mind the caveats described above, the outcomes from free credit reports seems positive for the banks as well.

[26] NIM is the ratio of net interest income (sum of RIAD4074 and RIAD4301) to earning assets. I use the definition of earning assets from St. Louis Fred: it is the sum of RCFD0071, RCFD1350, RCFD2122, RCFD3545, RCFD1754, and RCFD1772 (https://fred.stlouisfed.org/series/USNIM). RoE is the ratio of net income (RIAD4340) to book value of equity. RoA is the ratio of net income to book value of total assets. Liquid assets is the sum of RCFD1754, RCFD1773, RCFD3545, RCFD1754, RCFD3545, and RCFD1350. Cost of deposit is the ratio of RIAD4073 to earning assets.

§E An alternative mechanism based on information asymmetry

An alternative mechanism based on asymmetric information is plausible in which borrowers *privately know* their true creditworthiness type, but do not know what lenders know about them. Using free credit reports, borrowers learn that the information on them that lenders have is proportional to their true type. Hence, under the non-trivial search/application cost, bad borrowers self-select out. The borrower pool now improves relative to the situation in which borrowers do not know that lender has information about their true type, and optimistically expect that the information is better than what is warranted by their credit reports. Note that the improvement occurs here due to self-selecting-out by bad borrowers, but not by self-selecting-in by good borrowers as all borrowers *privately know* their true type. However, under the self-learning mechanism, borrowers themselves have imperfect information of their true type, thus both selecting-in by good borrowers selecting-out by bad borrowers contribute to pool improvement after credit reports become free.

The empirical findings are consistent primarily with the self-learning mechanism. We saw that in the treated areas both the mortgage applications and the first-time homebuyers fraction increased, not decreased. Both these findings provide evidence of selecting-in by borrowers, which is plausible only under the *self-learning* mechanism.

Another valid concern is that in assessing mortgage applications, together with the credit reports, lenders use private information such as those accumulated through relationship lending. This attenuates the effects of free credit reports. The concern is partially alleviated by the fact that lenders necessarily look at credit reports and scores when assessing borrowers.[27]

§F Robustness

The natural experiment utilized in this paper takes place in the year 2005, and the sample period chosen is 2000–2008 to allow for enough post-experiment observations. As the experiment is close to the financial crisis of 2008, it is crucial to ensure that the results are not caused by the unique lending environment that existed in 2007–2008. To this end, all the regressions when re-estimated by excluding the observations for years 2007 and 2008 yield similar results despite having only two post-experiment observations. These results are unreported for brevity.

[27] Experian (n.d.) explains: "Not all lenders think the same way, and they may have different ways of making their decisions. But all of them will look at some key factors to help them decide. These include: information on your credit report including your credit history and public record data."

6 Conclusion

A non-trivial proportion of consumers do not check their credit reports regularly, and do not know their credit scores. Various data suggest that consumers err in credit decisions in a manner consistent with them having imperfect information of their creditworthiness. Credit reports contain crucial creditworthiness information, and can aid them in credit-related decisions.

This paper examines the effect of lowering the consumers' economic cost of credit reports on the mortgage market outcomes. The enactment of the federal *Fair and Accurate Transactions Act of 2003* (FACTA) allowed all U.S. consumers to access three free credit reports annually from 2005, while seven states already had local laws permitting their residents to obtain the reports for free. The causal link is established by deploying this close-to-exogenous reduction in the cost of the reports in a difference-in-differences setting. Here, the border counties of the early-adopting states constitute the control group, and those of the surrounding states the treatment.

The key finding is that reducing consumers' economic costs of the reports improves mortgage market outcomes in a way consistent with improvements in the borrower pool, and benefits both consumers and lenders. Specifically, free credit reports resulted in an increase in mortgage demand and approval ratio, more credit to creditworthy borrowers, a reduction in defaults and subprime population proportion, more first-time homeowners, and better financial performance of lenders.

Though the findings pertain to the mortgage decisions of consumers, they broadly hold true for all consumer credit decisions when they have imperfect information of their creditworthiness. Furthermore, the causal nature of these findings implies that a policy intervention aimed at educating consumers of their creditworthiness may yield similar results.

References

Agarwal, S., Chomsisengphet, S., Mahoney, N., & Stroebel, J. (2018). Do Banks Pass Through Credit Expansions to Consumers Who Want to Borrow? *The Quarterly Journal of Economics*, *133*(1), 129–190.

Avery, R. B., Brevoort, K., & Canner, G. (2007). Opportunities and Issues in Using HMDA Data. *Journal of Real Estate Research*, *29*(4), 351–380.

Avery, R. B., Calem, P. S., & Canner, G. B. (2004). Credit Report Accuracy and Access to Credit. *Federal Reserve Bulletin*, *90*, 297.

Avery, R. B., Calem, P. S., Canner, G. B., & Bostic, R. W. (2003). An Overview of Consumer Data and Credit Reporting. *Federal Reserve Bulletin*, *89*, 47.

Balakina, O., Balasubramaniam, V., Dimri, A., & Sane, R. (2020). The Effect of Information Unshrouding on Financial Product Purchase Decision [Working Paper]. *Unpublished*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3519845

Bhutta, N. (2014). Payday Loans and Consumer Financial Health. *Journal of Banking & Finance*, *47*, 230–242.

Bogin, A., Doerner, W., & Larson, W. (2016). Local House Price Dynamics: New Indices and Stylized Facts [Working Paper]. *Unpublished*. Retrieved from https://www.fhfa.gov/PolicyProgramsResearch/Research/Pages/wp1601.aspx

Brown, M., Haughwout, A., Lee, D., & Van Der Klaauw, W. (2011). Do We Know What We Owe? A Comparison of Borrower-and Lender-reported Consumer Debt. *Unpublished*(523).

Census Bureau. (n.d.). *County Adjacency File* [Database]. Retrieved from https://www.census.gov/geographies/reference-files/2010/geo/county-adjacency.html

Census Bureau. (2000–2008). *County Business Patterns (CBP)* [Database]. Retrieved from https://www.census.gov/programs-surveys/cbp.html

Census Bureau. (2006). *Census of Population and Housing, 2000 [United States]: Census Tract Relationship Files (CTRF)* [Database]. Inter-university Consortium for Political and Social Research. Retrieved from https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/13287

Courchane, M., Gailey, A., & Zorn, P. (2008). Consumer Credit Literacy: What Price Perception? *Journal of Economics and Business*, *60*(1-2), 125–138.

Di Maggio, M., & Kermani, A. (2017). Credit-induced Boom and Bust. *The Review of Financial Studies*, *30*(11), 3711–3758.

Dobbie, W., Goldsmith-Pinkham, P., Mahoney, N., & Song, J. (2016). Bad Credit, No Problem? Credit and Labor Market Consequences of Bad Credit Reports. *The Journal of Finance*.

Elul, R., Gupta, D., & Musto, D. K. (2020). Concentration in Mortgage Markets: GSE Exposure and Risk-Taking in Uncertain Times [Working Paper]. *Unpublished*. Retrieved from https://www.philadelphiafed.org/-/media/research-and-data/publications/working-papers/2020/wp20-04.pdf

Equifax. (n.d.). *Understanding Hard Inquiries on Your Credit Report.* Retrieved 2020-06-10, from https://www.equifax.com/personal/education/credit/report/understanding-hard-inquiries-on-your-credit-report/

Experian. (n.d.). *What credit score do I need for a mortgage?* Retrieved 2020-06-10, from https://www.experian.co.uk/consumer/mortgages/guides/credit-and-mortgages.html

Federal Reserve Bank of Philadelphia. (2015). *What Your Credit Report Says About You.* Retrieved 2020-06-10, from https://www.philadelphiafed.org/consumer-resources/publications/what-your-credit-report-says

Federal Trade Commission. (January 13, 2000). *Nation's Big Three Consumer Reporting Agencies Agree To Pay $2.5 Million To Settle FTC Charges of Violating Fair Credit Reporting Act.* Retrieved 2021-06-15, from https://www.ftc.gov/news-events/press-releases/2000/01/nations-big-three-consumer-reporting-agencies-agree-pay-25

FRBNY, & Equifax. (n.d.). *Equifax Subprime Credit Population* [Database]. retrieved from FRED, Federal Reserve Bank of St. Louis.

Gerardi, K., Goette, L., & Meier, S. (2010). Financial Literacy and Subprime Mortgage Delinquency: Evidence from a Survey Matched to Administrative Data [Working Paper]. *Unpublished*. Retrieved from https://www.frbatlanta.org/-/media/documents/research/publications/wp/2010/wp1010.pdf

Homonoff, T., O'Brien, R., & Sussman, A. B. (2019). Does Knowing Your FICO Score Change Financial Behavior? Evidence from a Field Experiment with Student Loan Borrowers. *Review of Economics and Statistics*, 1–45.

Hundtofte, S. (2017). No Such Thing as a Free Option? Offers of Debt Forgiveness Under Imprecise Borrower Beliefs [Working Paper]. *Unpublished*. Retrieved from https://drive.google.com/file/d/1R1ii1F5WuCj6Qr3HhpmuWhmt_PNbPNSL

Hurst, E., Keys, B. J., Seru, A., & Vavra, J. (2016). Regional Redistribution through the US Mortgage Market. *American Economic Review*, *106*(10), 2982–3028.

Keys, B. J., Mukherjee, T., Seru, A., & Vig, V. (2010). Did Securitization Lead to Lax Screening? Evidence from Subprime Loans. *The Quarterly Journal of Economics*, *125*(1), 307–362.

Kulkarni, S., Truffa, S., & Iberti, G. (2018). Removing the Fine Print: Standardization, Disclosure, and Consumer Loan Outcomes [Working Paper]. *Unpublished*. Retrieved from https://static1.squarespace.com/static/58b5e6e15016e1efa0bfd0a5/t/5bd88062f4e1fc38159faefa/1540915300523/informational_frictions_chile.pdf

Liberman, A., Neilson, C., Opazo, L., & Zimmerman, S. (2018, September). The Equilibrium Effects of Information Deletion: Evidence from Consumer Credit Markets [Working Paper]. *Unpublished*. Retrieved from http://www.nber.org/papers/w25097

Lusardi, A., & Tufano, P. (2009, March). Debt Literacy, Financial Experiences, and Overindebtedness [Working Paper]. *Unpublished*. Retrieved from http://www.nber.org/papers/w14808

Manson, S., Schroeder, J., Van Riper, D., & Ruggles, S. (2019). *Ipums national historical geographic information system: Version 14.0* [Database]. IPUMS NHGIS.

Mian, A., & Sufi, A. (2009). The Consequences of Mortgage Credit Expansion: Evidence From the Us Mortgage Default Crisis. *The Quarterly Journal of Economics*, *124*(4), 1449–1496.

Mikhed, V. (2015). Can Credit Cards with Access to Complimentary Credit Score Information Benefit Consumers and Lenders? [Working Paper]. *Unpublished*. Retrieved from https://www.philadelphiafed.org/-/media/frbp/assets/consumer-finance/discussion-papers/dp15-03.pdf

Missouri Census Data Center. (2010). *Geocorr 2000: Geographic Correspondence Engine. Version 1.3.3 (August, 2010) with Census 2000 Geography* [Database]. Retrieved from http://mcdc.missouri.edu/applications/geocorr2000.html

Moore, D. L. (2003). *Survey of Financial Literacy in Washington State: Knowledge, Behavior, Attitudes, and Experiences*. Washington State Department of Financial Institutions.

Nott, L., & Welborn, A. (2003). A Consumers Access to a Free Credit Report: A Legal and Economic Analysis. *Congressional Research Service, Order Code RL32008*.

Office of Policy Development and Research. (n.d.). *HUD USPS Zip Code Crosswalk Files* [Database]. U.S. Department of Housing (HUD). Retrieved from https://www.huduser.gov/portal/datasets/usps_crosswalk.html

Perry, V. G. (2008). Is Ignorance Bliss? Consumer Accuracy in Judgments About Credit Ratings. *Journal of Consumer Affairs*, *42*(2), 189–205.

Prager, R. A. (2009). Determinants of the Locations of Payday Lenders, Pawnshops and Check-cashing Outlets [Working Paper]. *Unpublished*. Retrieved from https://www.federalreserve.gov/pubs/feds/2009/200933/200933pap.pdf

Reddit Forum. (n.d.). *Withdrawing a Mortgage Application.* Retrieved 2020-06-10, from https://www.reddit.com/r/personalfinance/comments/38k115/withdrawing_a_mortgage_application/

Scharfstein, D., & Sunderam, A. (2016). Market Power in Mortgage Lending and the Transmission of Monetary Policy [Working Paper]. *Unpublished*. Retrieved from https://www.hbs.edu/faculty/Pages/item.aspx?num=44239

Stango, V., & Zinman, J. (2009). Exponential Growth Bias and Household Finance. *The Journal of Finance*, *64*(6), 2807–2849.

Survey of Consumer Expectations. (2013–2020). *Survey of Consumer Expectations Credit Access Survey* [Database]. Federal Reserve Bank of New York. Retrieved from https://www.newyorkfed.org/microeconomics/sce

Survey of Consumer Finances. (1998–2007). *Survey of Consumer Finances* [Database]. Federal Reserve Board. Retrieved from https://www.federalreserve.gov/econres/aboutscf.htm

U.S. Senate. 108th Congress. (2004a). *The Fair Credit Reporting Act and Issues Presented by Reauthorization of the Expiring Preemption Provisions: Hearings Before the Committee on Banking, Housing, and Urban Affairs United States Senate*. U.S. Government Printing Office. (S. Hrg. 108-579. Pages: 375–376)

U.S. Senate. 108th Congress. (2004b). *The Fair Credit Reporting Act and Issues Presented by Reauthorization of the Expiring Preemption Provisions: Hearings Before the Committee on Banking, Housing, and Urban Affairs United States Senate*. U.S. Government Printing Office. (S. Hrg. 108-579. Prepared statement of Joel R. Reidenberg. Pages: 247–251)

Wikipedia. (n.d.). *AnnualCreditReport.com.* Retrieved 2020-12-29, from https://en.wikipedia.org/wiki/AnnualCreditReport.com

Figure I: A Sample Credit Report

This figure shows the summary page of a credit report obtained from the website www.annualcreditreport.com for free under the Fair and Accurate Transaction Act of 2003. The specific credit history-related details are not shown. The report contains, among other things, the details of the consumer's active accounts, debt-to-credit ratio, and an indication of the available borrowing capacity.

1. Summary

Review this summary for a quick view of key information contained in your Equifax Credit Report.

Report Date	Apr 14, 2020
Credit File Status	No fraud indicator on file
Alert Contacts	0 Records Found
Average Account Age	5 Months
Length of Credit History	8 Months
Accounts with Negative Information	0
Oldest Account	DISCOVER BANK (Opened Aug 29, 2019)
Most Recent Account	AMERICAN EXPRESS (Opened Jan 10, 2020)

Credit Accounts

Your credit report includes information about activity on your credit accounts that may affect your credit score and rating.

Account Type	Open	With Balance	Total Balance	Available	Credit Limit	Debt-to-Credit	Payment
Revolving	2	2	$606	$11,044	$11,650	5.0%	$70
Mortgage							
Installment							
Other							
Total	2	2	$606	$11,044	$11,650	5.0%	$70

Other Items

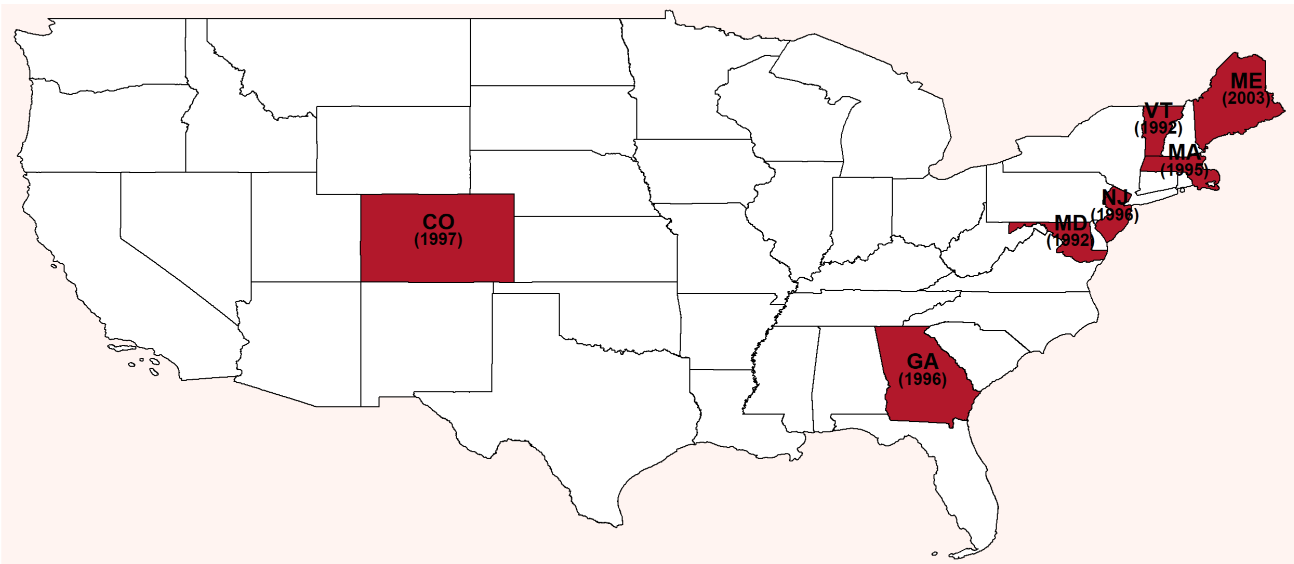
Your credit report includes your Personal Information and, if applicable, Consumer Statements, and could include other items that may affect your credit score and rating.

Consumer Statements		0 Statements Found
Personal Information		3 Items Found
Inquiries		2 Inquiries Found
Most Recent Inquiry	DISCOVER BANKAug 27, 2019	
Public Records		0 Records Found
Collections		0 Collections Found

Figure II: Differences in Access to Free Credit Reports Across US States

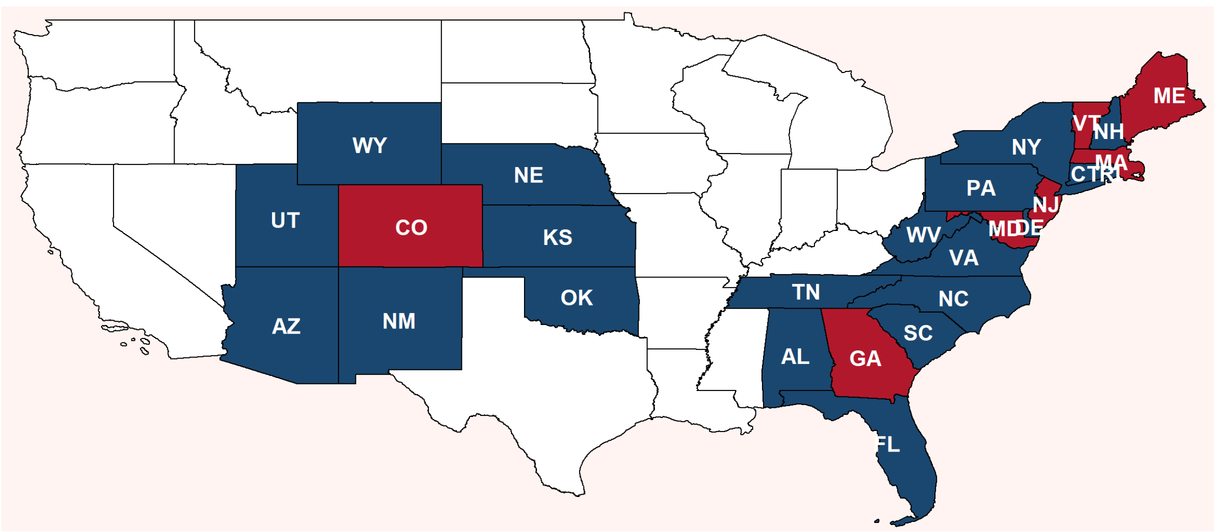
This figure illustrates the empirical setting of the paper. **Panel A** shows the pre-FACTA states — the states that had local laws allowing free credit reports before the enactment of FACTA in 2004 — and the respective years in which those states enacted free credit report laws. **Panel B** illustrates the difference-in-differences setting. The seven pre-FACTA states constitute the control states and the 26 states that surround them constitute the treated states.

Panel A: States Providing Free Credit Reports prior to FACTA (pre-FACTA states)



Pre-FACTA States

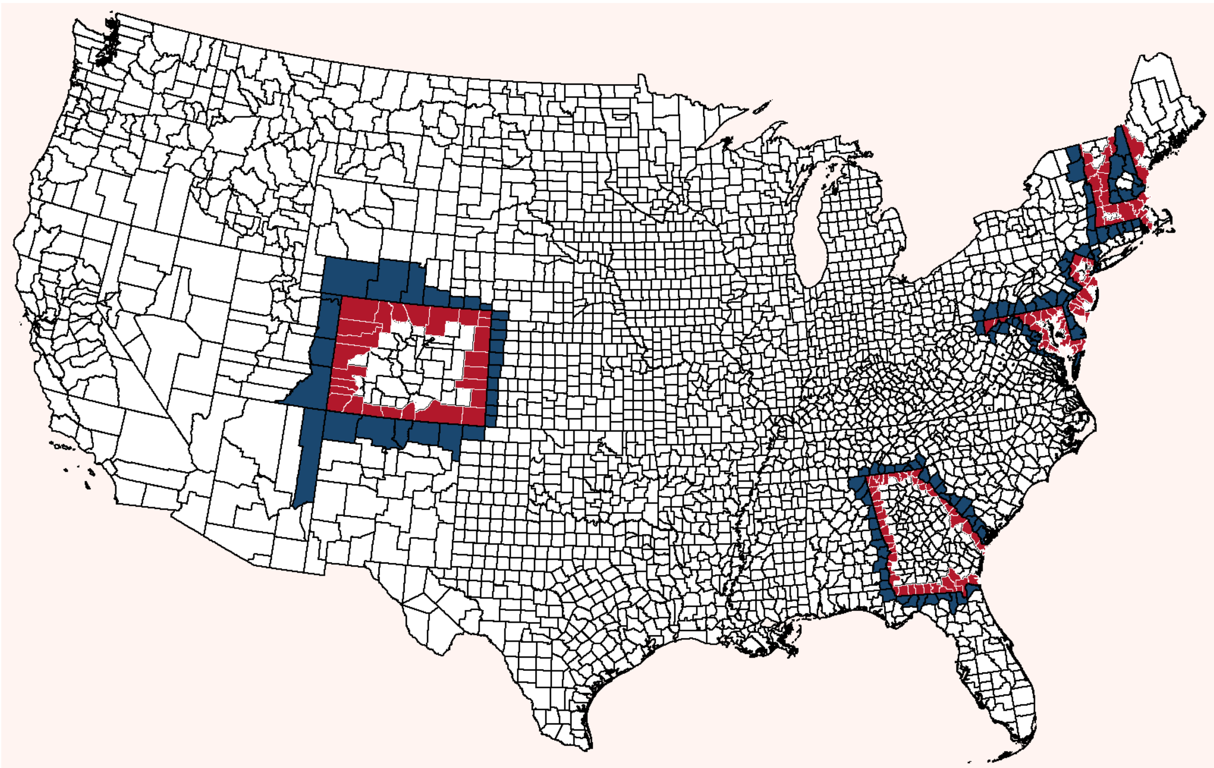
Panel B: Empirical Setting: Control and Treatment States



Control: Pre-FACTA States
Treatment: States surrounding the pre-FACTA states

Figure III: Control and Treatment Counties

This figure shows the treatment and control counties. All the counties that lie at the border of the seven pre-FACTA states constitute the control states. All the counties from the states surrounding the pre-FACTA states and lying at the border between them constitute the treatment counties.



Control: Counties at the border of pre-FACTA States

Treatment: Counties which belong to the bordering states and are adjacent to the control counties

Figure IV: Examining the Parallel Trends

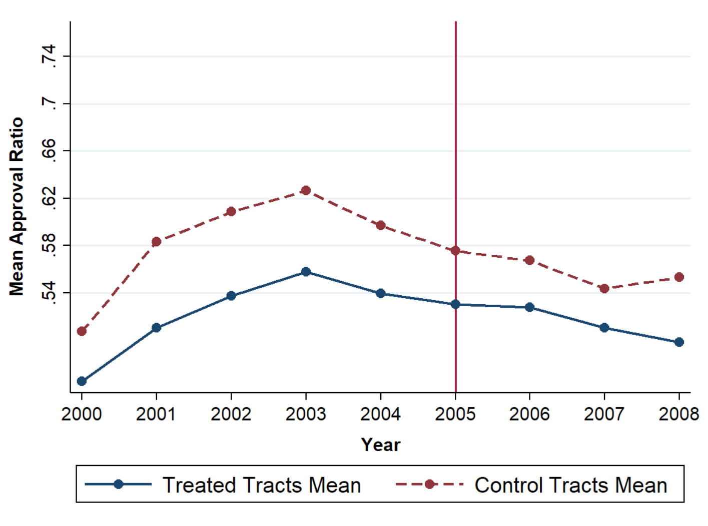
Panel A of this figure shows the mean approval ratio in the treated and control census tracts. **Panel B** of this figure shows the coefficients β_k from regressing *Approval Ratio* using the specification:

$$Y_{icsjt} = \beta_0 + \sum_{k=T-3}^{T-1} \beta_k \text{ Treatment}_{icsj} \times \text{Event}_k + \sum_{k=T+1}^{T+4} \beta_k \text{ Treatment}_{icsj} \times \text{Event}_k + \alpha_i + \gamma_{j,t} + \varepsilon_{icsjt}$$

where $\text{Event}_k = 1$ if $t = T - k$. $\text{Event}_k = 0$ if $t \neq T - k, k = \{-3, 4\}$. $T =$ Event year 2005.

Coefficients are estimated with respect to the base year 2004 ($j = 0$). The x-axis shows year relative to the pre-event year 2004, i.e., T = +1 is the first treated year 2005. The y-axis shows the coefficients β_k. The 95% confidence interval of β_k are also shown. The regression includes "*Border* $\times$ *Year*" and "*Census Tract*" fixed effects. Other terms in the equation are the same as those in Equation 1, and are described in Section 2. Standard errors are clustered by county.

Panel A: Mean Approval Ratio in Treated and Control Areas



Panel B: Coefficient Estimates of Approval Ratio by Years to Treatment

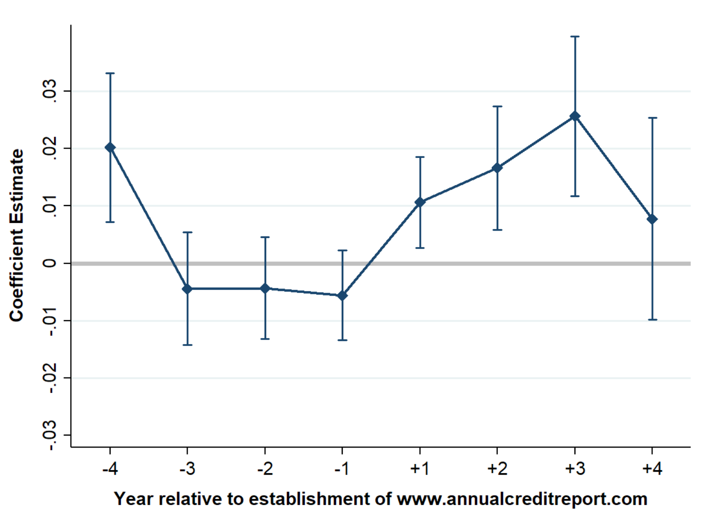
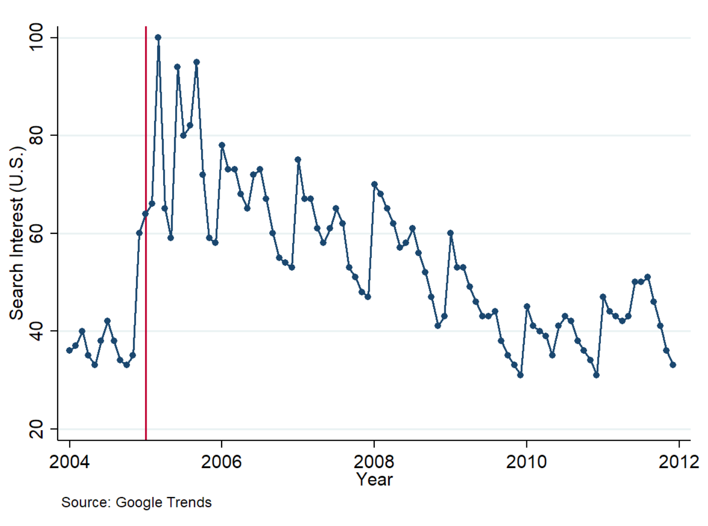


Figure V: Consumer Interest in Free Credit Reports: Google Trends

This figure shows the consumers interest in free credit report using Google Trends data. **Panel A** of this figure shows the plot of *Search Interest* for the keyphrase *Free Credit Report* in the US from Jan 1, 2004 till Dec 31, 2011. Numbers on the vertical axis represent search interest relative to the highest point on the chart during this period. A value of 100 (50) represents the peak popularity (half of the peak popularity) for the keyphrase. A value of 0 means there was not enough data. **Panel B** of this figure shows the difference in mean popularity rank of treatment and control states for the same keyphrase from 2004 to 2008. The popularity score of each state ranges from 0 to 100 and is calculated every year. A value of 100 represents the location with the highest popularity of the search term as a fraction of total searches in that location. A value of 50 indicates a location where it is half as popular.

Panel A: Google Search Interest for the Term "Free Credit Report" in the U.S.



Panel B: Relative Popularity of the keyphrase *Free Credit Report*

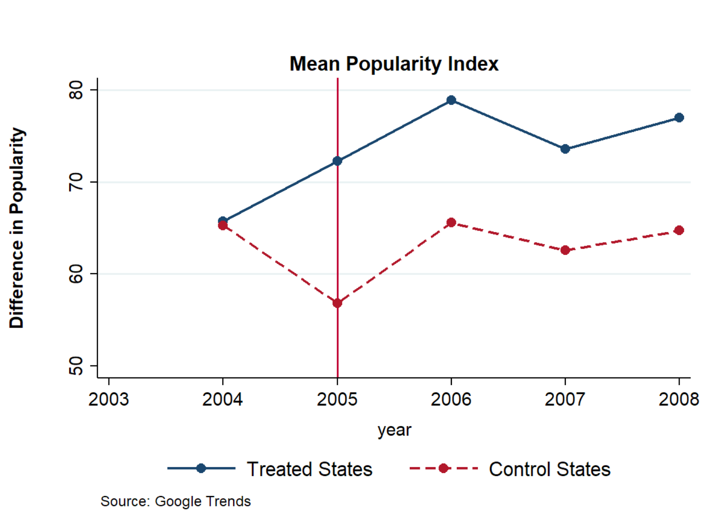
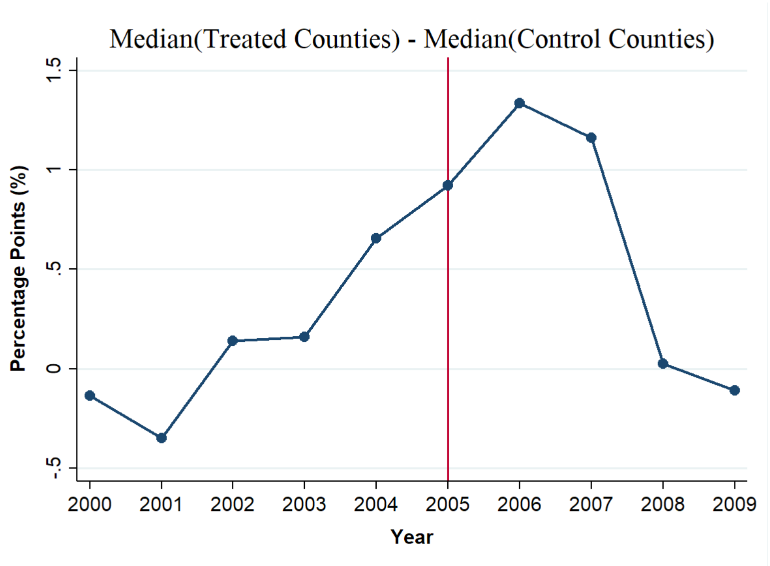


Figure VI: Subprime Population over Time

This figure plots the difference in subprime population fraction in treated and control counties. **Panel A** shows the difference in mean subprime population percentage calculated yearly over the treated counties and over the control counties. **Panel B** shows the difference in mean subprime population percentage calculated yearly over the treated counties and over the control counties.

Panel A: Difference in Subprime Population Fraction (using Median)



Panel B: Difference in Subprime Population Fraction (using Mean)

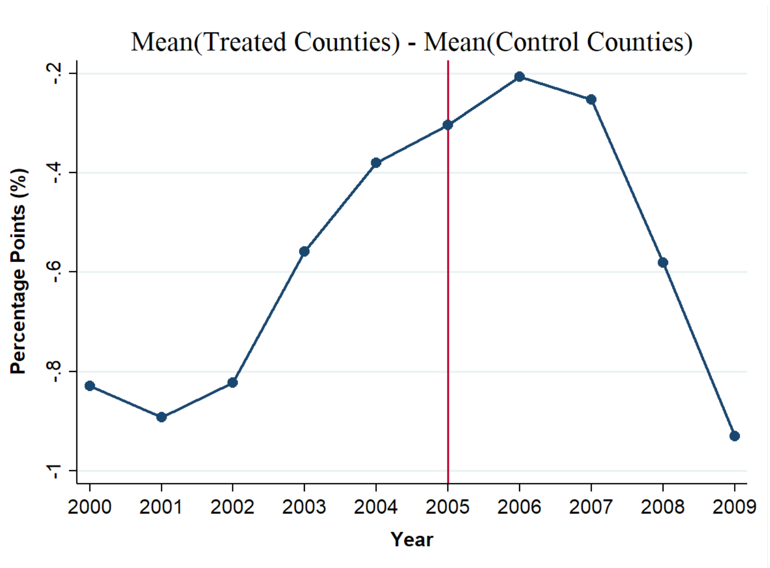


Figure VII: Effect of Free Credit Report on Mortgage Defaults

This figure shows the adjusted default rate for the sample of 30-year fixed-rate mortgages purchased by Fannie Mae and Freddie Mac. I separately calculate the percentage of total mortgages originated in the pre-event year 2004 and the post-event year 2005 which went into default in a month post-origination, $\text{Def}_{2005,age}$ and $\text{Def}_{2004,age}$, for the treated and control zip3-state areas respectively. A mortgage is in default when the scheduled payment is delayed by 30–59 days for the first time. I then calculate the adjusted default rate as:

$$\text{Adjusted Default Rate}_{age} = (\text{Def}_{2005,age} - \text{Def}_{2004,age})_{treated} - (\text{Def}_{2005,age} - \text{Def}_{2004,age})_{control}$$

where *age* represents months since origination.

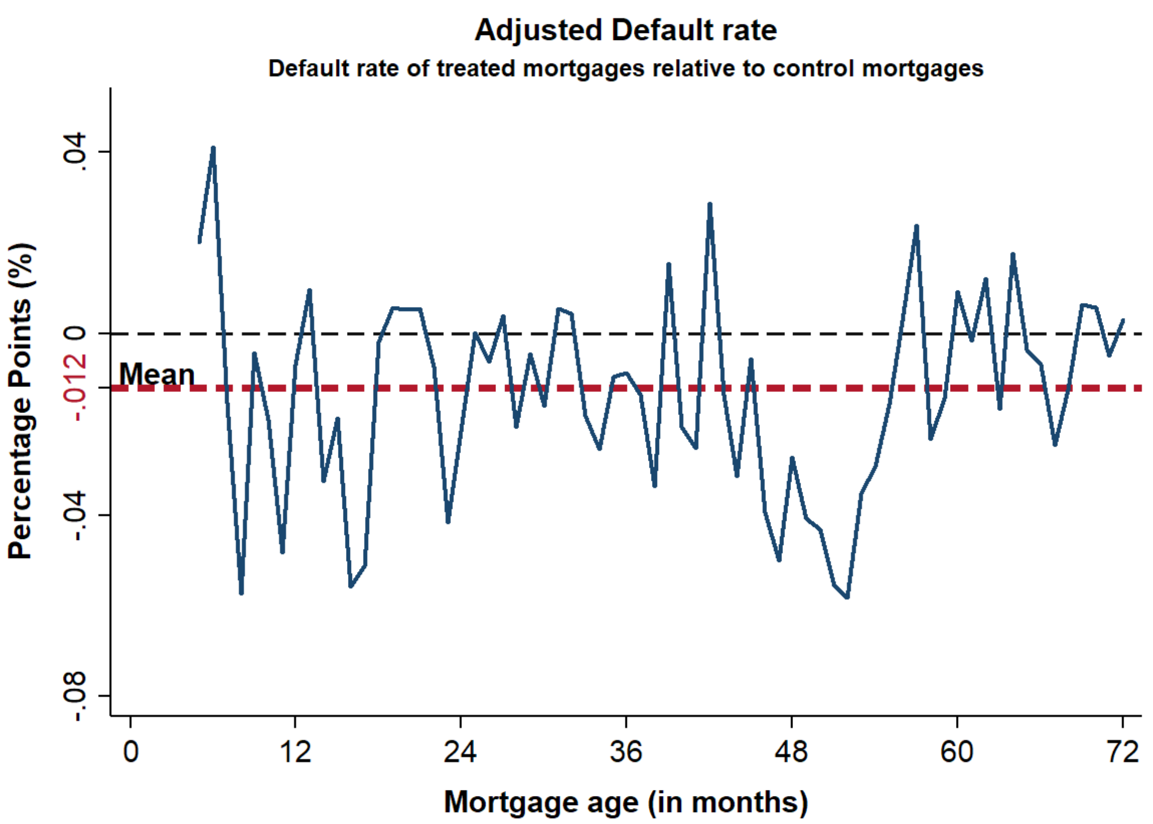


Table I: Summary Statistics

Panel A shows the statistics for the full sample time period (2000–2008). Panel B shows the statistics for the pre-treatment period (2000–2004) and the p-values for the t-test for difference in the control and treatment group. *Scaled applications,* (*N*) is the number of mortgage applications in a census tract scaled by the population aged 18 to 64 years in the tract. *Approval ratio* (*Aprv.*) is the ratio of the number of successful applications (identified by action type "1" in the HMDA dataset) to the number of total applications in a census tract. *Deny Credit Hist Ratio* and *Deny Debt-to-inc Ratio* are the ratio of applications denied due to credit history and debt-to-income ratio, respectively, to the number of total applications in a census tract. *Withdrawal Ratio* is the ratio of applications expressly withdrawn by the applicant to the number of total applications in the census tract.

The following four variables constitute *Economic Controls*. *Num. Lenders* (*log*) is the number of unique mortgage lenders in a census tracts, expressed in log. Δ *Inc per capita* is the annual growth rate of income per capita at the county level, Δ *Emp.* is the annual growth rate of the employment by all establishments at the county level, and Δ *State GDP* is the annual growth rate of the state gross domestic product.

Panel A: Full Sample (2000 – 2008)

	Full Sample				Control Group (C)				Treatment Group (T)			
	N	Mean	SD	Med.	N	Mean	SD	Med.	N	Mean	SD	Med.
Scaled Applications (N)	86822	83.40	74.74	66.39	36497	98.39	77.56	77.73	50325	72.52	70.66	56.48
Approval Ratio (Aprv.)	82717	0.54	0.13	0.55	35881	0.57	0.12	0.58	46836	0.52	0.14	0.53
Deny Credit Hist Ratio	82717	0.06	0.04	0.05	35881	0.05	0.04	0.04	46836	0.06	0.05	0.05
Deny Debt-to-inc Ratio	82717	0.03	0.03	0.03	35881	0.03	0.02	0.03	46836	0.03	0.03	0.03
Withdrawl Ratio	82717	0.12	0.05	0.12	35881	0.12	0.04	0.11	46836	0.12	0.06	0.12
Num. Lenders (log)	83230	3.16	0.78	3.30	34425	3.37	0.60	3.42	48805	3.01	0.85	3.20
Δ Inc per capita	2295	0.04	0.06	0.04	1143	0.04	0.05	0.04	1152	0.05	0.07	0.04
Δ Emp	2298	0.01	0.09	0.01	1138	0.01	0.09	0.01	1160	0.01	0.09	0.01
Δ State GDP	78	0.04	0.03	0.04	32	0.04	0.02	0.04	46	0.05	0.04	0.05

Panel B: Pre - Treatment Sample (2000 – 2004)

	Full Sample				Control Group (C)				Treatment Group (T)				(C-T)
	N	Mean	SD	Med.	N	Mean	SD	Med.	N	Mean	SD	Med.	p-val
Scaled applications (N)	48368	110.48	83.54	93.46	20290	129.66	86.13	108.65	28078	96.61	78.77	83.19	0.000
Approval Ratio (Aprv.)	47029	0.55	0.14	0.56	20073	0.58	0.13	0.60	26956	0.52	0.14	0.53	0.000
Deny Credit Hist Ratio	47029	0.06	0.04	0.05	20073	0.06	0.04	0.05	26956	0.07	0.05	0.06	0.000
Deny Debt-to-inc Ratio	47029	0.03	0.02	0.03	20073	0.03	0.02	0.03	26956	0.03	0.02	0.03	0.000
Withdrawl Ratio	47029	0.12	0.05	0.11	20073	0.12	0.04	0.11	26956	0.13	0.05	0.12	0.000
Num. Lenders (log)	44776	3.36	0.73	3.48	18218	3.53	0.60	3.59	26558	3.24	0.78	3.39	0.000
Δ Inc per capita	1275	0.04	0.06	0.04	635	0.04	0.05	0.04	640	0.04	0.07	0.04	0.614
Δ Emp	1274	0.01	0.09	0.01	632	0.01	0.09	0.01	642	0.00	0.09	0.01	0.311
Δ State GDP	50	0.05	0.02	0.05	19	0.05	0.02	0.05	31	0.05	0.02	0.05	0.619

Table II: Survey Evidence on the Credit Reports Usage and Discouraged Borrowers

This table reports the regression results from the SCE Credit Access survey. *Never* is one if a respondent has never checked his/her credit score (Q. *N23*). *Infrequently* is one if respondent has never checked it or last checked it more than 2 years ago (Q. *N23*). *Unaware* is one if respondent don't now his/her credit score (Q. *N22*). *Dscrgd* is one if respondent said "I do not think I would get approved" in *Q. N19*. Note that this question (*Q. N19*) is a conditional question in the survey. Hence the observations in specifications (4–6) include only the responses in which (i) for *Q. N17A*, respondent selected *very unlikely* or *somewhat unlikely* to apply for mortgage/home-based loan, or refinance, or (ii) for *Q. N17B*, mentioned the probability to apply for mortgage or to refinance as less than 10%. All regressions include $Year \times Month$ fixed effects (FE). Standard errors are clustered by survey's Year×Month. p-values are reported below the coefficients in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Check Credit Report		Know Credit Score	Mortgage-discouraged Borrowers		
	(1)	(2)	(3)	(4)	(5)	(6)
	Never	Infrequently	Unaware	Dscrgd	Dscrgd	Dscrgd
Check Infrequently					0.03**	
					(0.05)	
Unaware						0.05*
						(0.06)
Constant	0.08***	0.20***	0.12***	0.13***	0.13***	0.13***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Cluster (Year-Month)	Yes	Yes	Yes	Yes	Yes	Yes
FE (Year-Month)	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.007	0.007	0.007	0.003	0.004	0.005
Observations	19231	19231	20275	9059	9058	9058

Table III: Mortgage Applications, Approval Ratio, and House Prices

This table reports the estimates of the treatment effect of free credit reports on the number of mortgage applications, approval ratio, and growth in house prices. The regression specification is:

$Y_{icsjt} = \beta_0 + \beta_1 \text{Treatment}_{icsj} \times \text{Post}_t + \delta \times \text{Economic Controls} + \alpha_i + \gamma_{jt} + \varepsilon_{icsjt}$. See Eq. (1). *N*, *Aprv.*, and Δ*HPI* are the number of applications per 1000 adults, the approval ratio in a census tract and growth in house prices at census tract level, respectively. *Economic Controls* are number of mortgage lenders in a census tract and annual growth rate of county income per capita, county aggregate employment, and state gross domestic product (GDP). The coefficient associated with the $Treat \times Post$ interaction term captures the change in the dependent variable in the treated census tracts relative to the control census tracts. All regressions include the $Border \times Year$ fixed effects (FE) and the *Census Tract* FE. All variables are defined in Table (I). Standard errors are clustered by county. t-statistics are reported below the coefficients in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Aprv.	Aprv.	N	N	ΔHPI	ΔHPI
Treat $\times$ Post	0.01**	0.02***	13.30***	16.32***	1.74*	1.90*
	(2.41)	(2.72)	(2.95)	(3.75)	(1.83)	(1.87)
Economic Controls	No	Yes	No	Yes	No	Yes
Census Tract FE	Yes	Yes	Yes	Yes	Yes	Yes
Border $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster (County)	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (Adj.)	0.748	0.740	0.806	0.817	0.683	0.693
Observations	82669	77181	86810	81292	25383	24233

Table IV: Owner-occupied and Non-owner-occupied Mortgages

This table examines the changes in (i) owner-occupied mortgages applications, (ii) non-owner-occupied applications, (iii) the fraction of non-owner-occupied mortgages as total applications, and (iv) as a fraction of successful applications. The regression specification is:

$$Y_{icsjt} = \beta_0 + \beta_1 \text{Treatment}_{icsj} \times \text{Post}_t + \delta \times \text{Economic Controls} + \alpha_i + \gamma_{jt} + \varepsilon_{icsjt}. \quad \text{See Eq. (1).}$$

The dependent variable in columns (1) through (4) is the number of applications per 1000 adults in a census tract, *N*. In columns (1) and (2), *N* measures owner-occupied category mortgage applications only; in columns (3) and (4), non-owner-occupied only. The dependent variable in columns (5) and (6) is the non-owner-occupied mortgage as a fraction of total applications, and in columns (7) and (8), as the fraction of originated applications. *Economic Controls* are number of mortgage lenders in a census tract and annual growth rate of county income per capita, county aggregate employment, and state gross domestic product (GDP). The coefficient associated with the *Treat* $\times$ *Post* interaction term captures the change in the dependent variable in the treated census tracts relative to the control census tracts. All regressions include *Border* $\times$ *Year* fixed effects (FE) and *Census Tract* FE. All variables are defined in Table (I). Standard errors are clustered by county. t-statistics are reported below the coefficients in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Owner		Non-owner		Non-owner, % of all appl.		Non-owner, % of succ. appl.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	N	N	N	N	%	%	%	%
Treat $\times$ Post	12.80***	15.70***	0.79	0.97*	0.01**	0.01*	0.01**	0.01*
	(2.89)	(3.71)	(1.63)	(1.75)	(1.99)	(1.74)	(2.02)	(1.91)
Economic Controls	No	Yes	No	Yes	No	Yes	No	Yes
Census Tract FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Border $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster (County)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (Adj.)	0.808	0.819	0.756	0.758	0.088	0.081	0.087	0.080
Observations	86810	81292	86810	81292	82669	77181	82583	77093

Table V: Characterizing the Effect: Heterogeneity by Consumer Creditworthiness

This table reports the estimates of the treatment effect of free credit reports on the number of mortgage applications per 1000 adult (N) and the approval ratio ($Aprv.$) in *ex-ante* low and high creditworthiness areas. Panel A uses county-level subprime population fraction to classify areas by creditworthiness; a county is "subprime" if its subprime population fraction is more than the *regional mean* subprime population fraction in 1999. In Panel B, the classification benchmark is the average number of payday lenders in census tracts in counties at the border of Colorado (CO) and surrounding states in 2004. The regression specification is:

$$Y_{icsjt} = \beta_0 + \beta_1 \text{Treatment}_{icsj} \times \text{Post}_t + \delta \times \text{Economic Controls} + \alpha_i + \gamma_{jt} + \varepsilon_{icsjt}. \quad \text{See Eq. (1).}$$

Economic Controls are number of mortgage lenders in a census tract and annual growth rate of county income per capita, county aggregate employment, and state gross domestic product (GDP). The coefficient associated with the *Treat* × *Post* captures the change in the dependent variable in the treated census tracts relative to the control census tracts. All regressions include *Border* × *Year* fixed effects (FE) and *Census Tract* FE. All variables are defined in Table (I). Standard errors are clustered by county. t-statistics are reported below the coefficients in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A: County-level Creditworthiness Measure

	Ex-ante High Creditworthiness (Prime Counties)				*Ex-ante* Low Creditworthiness (Subprime Counties)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	N	N	Aprv.	Aprv.	N	N	Aprv.	Aprv.
Treat × Post	16.75**	18.77***	0.02**	0.02**	8.53*	11.48**	0.01	0.01*
	(2.30)	(2.66)	(2.32)	(2.61)	(1.66)	(2.40)	(1.57)	(1.73)
Economic Controls	No	Yes	No	Yes	No	Yes	No	Yes
Census Tract FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Border × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster (County)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (Adj.)	0.802	0.822	0.774	0.769	0.826	0.828	0.698	0.691
Observations	39264	35895	38185	34833	47251	45129	44384	42275

Panel B: Census tract-level Creditworthiness Measure

	High Creditworthiness (# Payday Lenders - Low)				Low Creditworthiness (# Payday Lenders - High)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	N	N	Aprv.	Aprv.	N	N	Aprv.	Aprv.
Treat × Post	68.43***	68.66***	0.06***	0.05***	43.72***	41.68***	0.02	0.02
	(5.31)	(4.95)	(3.52)	(3.08)	(3.82)	(3.76)	(0.94)	(0.77)
Economic Controls	No	Yes	No	Yes	No	Yes	No	Yes
Census Tract FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Border × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster (County)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (Adj.)	0.790	0.793	0.732	0.699	0.816	0.835	0.794	0.778
Observations	1452	1229	1395	1177	872	777	865	771

Table VI: Characterizing the Effect: Heterogeneity by Income Level of Consumers

This table reports estimates of the treatment effect of free credit reports on the number of mortgage applications and the approval ratio for each income quartile. The regression specification is:

$$Y_{icsjt} = \beta_0 + \beta_1 \text{Treatment}_{icsj} \times \text{Post}_t + \delta \times \text{Economic Controls} + \alpha_i + \gamma_{jt} + \varepsilon_{icsjt}. \quad \text{See Eq. (1).}$$

N and *Aprv.* are the number of applications per 1000 adults and the approval ratio in a census tract, respectively. Income quartiles are calculated every year for a given census tract. *Economic Controls* are number of mortgage lenders in a census tract and annual growth rate of county income per capita, county aggregate employment, and state gross domestic product (GDP). The coefficient associated with the *Treat* × *Post* interaction term captures the change in the dependent variable in the treated census tracts relative to the control census tracts. All regressions include *Border* × *Year* fixed effects (FE) and *Census Tract* FE. All variables are defined in Table (I). Standard errors are clustered by county. t-statistics are reported below the coefficients in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A: Number of Applications per 1000 adults

	Income Quartile 1		Income quartile 2		Income Quartile 3		Income quartile 4	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	N	N	N	N	N	N	N	N
Treat × Post	0.06	0.30	1.83**	2.06***	2.62**	2.91***	4.33**	5.17***
	(0.04)	(0.23)	(2.32)	(2.85)	(2.44)	(3.19)	(2.00)	(2.70)
Economic Controls	No	Yes	No	Yes	No	Yes	No	Yes
Census Tract FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Border × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster (County)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (Adj.)	0.762	0.763	0.776	0.776	0.745	0.747	0.660	0.671
Observations	88286	81292	88286	81292	88286	81292	88286	81292

Panel B: Approval Ratio

	Income Quartile 1		Income quartile 2		Income Quartile 3		Income quartile 4	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Aprv.	Aprv.	Aprv.	Aprv.	Aprv.	Aprv.	Aprv.	Aprv.
Treat × Post	0.01**	0.01**	0.01	0.01	0.01	0.01	0.01	0.01
	(2.04)	(2.43)	(0.92)	(1.06)	(0.86)	(0.82)	(1.31)	(1.31)
Economic Controls	No	Yes	No	Yes	No	Yes	No	Yes
Census Tract FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Border × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster (County)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (Adj.)	0.344	0.335	0.391	0.378	0.404	0.389	0.363	0.349
Observations	71891	66663	72440	67196	72560	67308	72007	66768

Table VII: Contraction in Rejections due to Credit History

This table reports the estimates of the treatment effect on the fraction of mortgage applications denied because of credit history and debt-to-income ratio, estimated separately. The regression specification is:

$\text{Interest Rate}_{icsjt} = \beta_0 + \beta_1 \text{Treatment}_{icsj} \times \text{Post}_t + \delta \times \text{Controls} + \alpha_i + \gamma_{jt} + \varepsilon_{icsjt}$. See Eq. (1). *%C.Hist* (*%DTI*) is calculated as the ratio of the number of denied applications due to credit history (debt-to-income ratio) to the total number of mortgage applications in a census tract. *High Denial Areas* are the census tracts where denial per capita in the pre-event year 2004 was more than the regional mean of denials across the census tracts. *Economic Controls* are number of mortgage lenders in a census tract and annual growth rate of county income per capita, county aggregate employment, and state gross domestic product (GDP). The coefficient associated with the *Treat* × *Post* interaction term captures the change in the fraction of mortgage applications denied due to a given reason in the treated census tracts relative to the control census tracts. All regressions include *Border* × *Year* fixed effects (FE) and *Census Tract* FE. All variables are defined in Table (I). Standard errors are clustered by county. t-statistics are reported below the coefficients in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	All Areas		High Denial Areas		All Areas		High Denial Areas	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	% C.Hist	% C.Hist	% C.Hist	% C.Hist	% DTI	% DTI	% DTI	% DTI
Treat × Post	-0.003	-0.003	-0.003**	-0.003*	-0.002	-0.002	-0.002	-0.002
	(-1.49)	(-1.55)	(-2.04)	(-1.85)	(-1.08)	(-1.21)	(-1.49)	(-1.39)
Economic Controls	No	Yes	No	Yes	No	Yes	No	Yes
Census Tract FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Border × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster (County)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (Adj.)	0.541	0.532	0.575	0.568	0.267	0.264	0.319	0.320
Observations	82669	77181	39075	36802	82669	77181	39075	36802

Table VIII: Confident Search

This table reports the estimates of the treatment effect on the fraction of mortgage applications withdrawn. The regression specification is:

$Y_{icsjt} = \beta_0 + \beta_1 \text{Treatment}_{icsj} \times \text{Post}_t + \delta \times \text{Economic Controls} + \alpha_i + \gamma_{jt} + \varepsilon_{icsjt}$. See Eq. (1). *Withdrawal Ratio* is the ratio of applications expressly withdrawn by consumers to the number of applications in a census tract. *Economic Controls* are number of mortgage lenders in a census tract and annual growth rate of county income per capita, county aggregate employment, and state gross domestic product (GDP). The coefficient associated with the *Treat* × *Post* interaction term captures the change in fraction of applications expressly withdrawn by applicants in the treated census tracts relative to the control census tracts. All regressions include *Border* × *Year* fixed effects (FE) and *Census Tract* FE. All variables are defined in Table (I). Standard errors are clustered by county. t-statistics are reported below the coefficients in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)
	Withdrawal Ratio	Withdrawal Ratio
Treat × Post	-0.009***	-0.010***
	(-2.80)	(-3.80)
Economic Controls	No	Yes
Census Tract FE	Yes	Yes
Border × Year FE	Yes	Yes
Cluster (County)	Yes	Yes
R^2 (Adj.)	0.340	0.338
Observations	82669	77181

Table IX: Increase in First-time Homebuyers

This table reports the estimates of the treatment effect on the fraction of first-time homebuyers. The regression specification is:

$$Y_{zsjt} = \beta_0 + \beta_1 \text{Treatment}_{zsj} \times \text{Post}_t + \delta \times \text{Economic controls} + \alpha_{zs} + \gamma_{jt} + \varepsilon_{zsjt}. \quad \text{See Eq. (3).}$$

The mortgage data used in this table are from the Fannie Mae and Freddie Mac combined single-family loan dataset (GSE data). The dependent variable in column 1 (2) is the ratio of the number of first-time homebuyers to the total number of mortgages (total number of mortgages for which the information on the first-time homebuyer is not missing) in a given zip3-state area. *Economic Controls* are number of mortgage lenders in a census tract and annual growth rate of county income per capita, county aggregate employment, and state gross domestic product (GDP). The coefficient associated with the $Treat \times Post$ interaction term captures the change in proportion of *First-time homebuyers* in the treated 3-digit zipcode-state areas relative to the control 3 digit zipcode areas. All regressions include *Zip3-State* fixed effects (FE) and $Border \times Quarter$ FE. All variables are defined in Table (I). Standard errors are clustered by county. t-statistics are reported below the coefficients in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Denominator - Applications with Known Status		Denominator - All Applications	
	(1)	(2)	(3)	(4)
	% First-time	% First-time	% First-time	% First-time
Treat × Post	0.01**	0.01**	0.01**	0.01*
	(2.55)	(2.31)	(2.00)	(1.78)
Economic Controls	No	Yes	No	Yes
Zip3-State FE	Yes	Yes	Yes	Yes
Border × Qtr FE	Yes	Yes	Yes	Yes
Cluster Zip3-State	Yes	Yes	Yes	Yes
R^2 (Adj.)	0.691	0.692	0.691	0.691
Observations	7706	7706	7711	7711

Table X: Equilibrium Interest Rates on GSE-purchased Mortgages

This table reports the estimates of the treatment effect of free credit reports on interest rates at zip3-state level geographic aggregation. The regression specification is:

$\text{Interest Rate}_{izsjt} = \beta_0 + \beta_1 \text{Treatment}_{izsj} \times \text{Post}_t + \delta \times \text{Controls} + \alpha_{zs} + \gamma_{jt} + \varepsilon_{izsjt}$. See Eq. (4). In column (1), controls included are *credit score* and *combined loan-to-value* (*CLTV*). In column (2) following controls are further added: *debt-to-income ratio*, *number of units in the property*, and *mortgage insurance percentage*. The coefficient associated with the *Treat* × *Post* interaction term captures the change in the dependent variable in the treated zip3-state areas relative to the control zip3-state areas. All regressions include *Zip3-State* fixed effects and *Border* × *Quarter* fixed effects (FE). All variables are defined in Table (I). Standard errors are clustered by county. t-statistics are reported below the coefficients in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Interest Rate (%)	
	(1)	(2)
	%	%
Treat × Post	0.009***	0.011***
	(14.37)	(12.60)
Controls	Yes	Yes
Zip3-State FE	Yes	Yes
Border × Qtr FE	Yes	Yes
Cluster Zip3-State	Yes	Yes
R^2 (Adj.)	0.731	0.758
Observations	7657115	3548884

Table XI: Effect Heterogeneity by Lenders Density

This table reports the estimates of the treatment effect on the origination volume (in 1000 USD) per adult and the approval ratio, estimated separately for census tracts having a high and low density of mortgage lenders per capita in 2004. The regression specification is:

$Y_{icsjt} = \beta_0 + \beta_1 \text{Treatment}_{icsj} \times \text{Post}_t + \delta \times \text{Economic Controls} + \alpha_i + \gamma_{jt} + \varepsilon_{icsjt}$. See Eq. (1). *Low* (*High*) identifies a census tract having a lower (higher) number of HMDA lenders than the *regional mean* number of HMDA lenders (per census tract) within the bordering counties between the given control state and all the treatment states surrounding it in 2004 (See Footnote 19). *Difference* [*High* - *Low*] shows the result of the t-test for the difference in coefficients of $Treat \times Post$ in specifications *High* and *Low*. The dependent variable in columns 1 through 4 is volume of mortgages originated (in 1000 USD) per adult in a census tract. The dependent variable in columns 4 through 8 is the approval ratio of mortgage applications at census tract-level. *Economic Controls* are number of mortgage lenders in a census tract and annual growth rate of county income per capita, county aggregate employment, and state gross domestic product (GDP). The coefficient associated with the $Treat \times Post$ interaction term captures the change in the dependent variable in the treated census tracts relative to the control census tracts. All variables are defined in Table (I). Standard errors are clustered by county. t-statistics are reported below the coefficients in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Volume (in 1000 USD) per Adult				Approval Ratio			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Low	High	Low	High	Low	High	Low	High
Treat × Post	0.002**	0.001	0.003***	0.002	0.016**	0.011*	0.017***	0.011**
	(2.20)	(1.15)	(3.06)	(1.64)	(2.51)	(1.93)	(2.83)	(2.03)
Difference [High - Low]		-0.001		-0.001		-0.005		-0.006
p-value		(0.594)		(0.493)		(0.610)		(0.536)
Economic Controls	No	No	Yes	Yes	No	No	Yes	Yes
Census Tract FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Border × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster (County)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (Adj.)	0.643	0.571	0.637	0.615	0.758	0.728	0.749	0.721
Observations	60708	25807	56642	24382	57632	24937	53589	23519

Table XII: Did Origination Increase due to Rise in Private Securitization?

This table reports the estimates of the treatment effect on the approval ratio estimated separately for mortgages sold to Non-GSEs, sold to GSEs, and not sold. The regression specification is:

$Y_{icsjt} = \beta_0 + \beta_1 \text{Treatment}_{icsj} \times \text{Post}_t + \delta \times \text{Economic Controls} + \alpha_i + \gamma_{jt} + \varepsilon_{icsjt}$. See Eq. (1).

The dependent variables are the fraction of total mortgage applications originated and sold to the non-GSEs (columns 1 and 2); originated and sold to the GSEs (columns 3 and 4); approved and not sold by the lending institution (columns 5 and 6). All the dependent variables are calculated at the census tract level. *Economic Controls* are number of mortgage lenders in a census tract and annual growth rate of county income per capita, county aggregate employment, and state gross domestic product (GDP). The coefficient associated with the *Treat* × *Post* interaction term captures the change in the dependent variable in the treated census tracts relative to the control census tracts. All variables are defined in Table (I). Standard errors are clustered by county. t-statistics are reported below the coefficients in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Sold to Non-GSE		Sold to GSE		Not Sold	
	(1)	(2)	(3)	(4)	(5)	(6)
	Fraction	Fraction	Fraction	Fraction	Fraction	Fraction
Treat × Post	-0.004	0.001	0.047**	0.046***	0.000	0.002
	(-0.35)	(0.06)	(2.55)	(2.77)	(0.05)	(0.50)
Economic Controls	No	Yes	No	Yes	No	Yes
Census Tract FE	Yes	Yes	Yes	Yes	Yes	Yes
Border × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster (County)	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (Adj.)	0.009	-0.003	0.004	-0.003	0.055	0.028
Observations	82669	77181	82669	77181	82669	77181

Table XIII: Did Origination Increase due to Subprime Lending? Credit Score-based Evidence

This table reports the estimates of the treatment effect on the number of mortgages originated to prime and subprime borrowers by government sponsored enterprises (GSEs) Fannie Mae and Freddie Mac. The regression specification is:

$Y_{zsjt} = \beta_0 + \beta_1 \text{Treatment}_{zsj} \times \text{Post}_t + \delta \times \text{Economic controls} + \alpha_{zs} + \gamma_{jt} + \varepsilon_{zsjt}$. See Eq. (3).

The dependent variable in column 1 is *Number of mortgages originated to Prime Borrowers* (credit score≥620) in a given zip3-state area. The dependent variable in column 2 is *Number of applications to subprime borrowers* (credit score<620) in a given zip3-state area. *Economic Controls* are number of mortgage lenders in a census tract and annual growth rate of county income per capita, county aggregate employment, and state gross domestic product (GDP). The coefficient associated with the *Treat* × *Post* interaction term captures the change in the dependent variable in the treated zip3-state areas relative to the control zip3-state areas. All regressions include *Zip3-State* fixed effects (FE) and *Border* × *Quarter* FE. All variables are defined in Table (I). Standard errors are clustered by county. t-statistics are reported below the coefficients in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)
	N-Prime	N-Prime	N-Subprime	N-Subprime
Treat × Post	308.58***	312.51***	10.48**	10.78**
	(3.39)	(3.33)	(2.12)	(2.16)
Economic Controls	No	Yes	No	Yes
Zip3-State FE	Yes	Yes	Yes	Yes
Border × Qtr FE	Yes	Yes	Yes	Yes
Cluster Zip3-State	Yes	Yes	Yes	Yes
R^2 (Adj.)	0.757	0.758	0.792	0.792
Observations	7711	7711	7711	7711

Table XIV: Effect of Free Credit Reports on Banks

This table reports the estimates of the treatment effect on banks. The regression specification is:

$$Y_{bt} = \beta_0 + \beta_1 \text{Treatment}_b \times \text{Post}_t + \delta \times \text{Bank Controls}_{bt} + \alpha_l + \gamma_t + \varepsilon_{bt}. \quad \text{See Eq. (5).}$$

NIM is Net Interest Margin. It is the ratio of net interest income to earning assets, expressed in percentage. *RoE* is Return on Equity. It is the ratio of net income to book value of equity, expressed in percentage. *RoA* is Return on Assest. It is the ratio of net income to book value of total assets, expressed in percentage. Bank Controls include: natural log of the total assets expressed in USD 1000; share of liquid assets in total assets, expressed in percentage; and cost of deposit (ratio of total interest expense to total earning assets, expressed in percentage). *Economic Controls* are number of mortgage lenders in a census tract and annual growth rate of county income per capita, county aggregate employment, and state gross domestic product (GDP). The coefficient associated with the $Treat \times Post$ interaction term captures the change in the dependent variable in the treated census tracts relative to the control census tracts. All regressions include $Border \times Year$ fixed effects (FE) and *Census Tract* FE. Standard errors are clustered by county. t-statistics are reported below the coefficients in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	NIM (%)	NIM (%)	RoE (%)	RoE (%)	RoA (%)	RoA (%)
Treat × Post	0.06***	0.06***	0.74***	0.74***	0.07***	0.07***
	(5.55)	(6.00)	(5.05)	(5.25)	(5.17)	(5.53)
Bank Controls	No	Yes	No	Yes	No	Yes
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Qtr FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster (Bank)	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (Adj.)	0.807	0.814	0.586	0.597	0.556	0.573
Observations	86323	86323	86323	86323	86323	86323