Know Thyself: Free Credit Reports and the Retail Mortgage Market*

Amit Kumar†

This Version: Jan 2021
Click here for the latest version

Abstract

Under imprecise creditworthiness information, borrowers may take erroneous credit decisions. Credit reports—which record one's creditworthiness—became free in entire US in 2005, while they already had been free in seven states. Exploiting this in a difference-in-differences setting, I show that cheaper credit reports to consumers improved mortgage market outcomes. It resulted in increased mortgage demand and approvals, more origination to creditworthy borrowers, and reduced defaults and subprime population fraction. Also, lenders' financial performance improved, and more consumers became first-time homeowners. Additional findings, including increased interest rates, suggest a demand-driven channel in which borrowers learn their creditworthiness from credit 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. I express my sincere gratitude to Professor Renée Adams and Professor Anjan Thakor for their valuable feedback during doctoral sessions. I am also indebted to Deniz Okat, Abhiroop Mukherjee, Arkodipta Sarkar, and Alminas Zaldokas for various consultations and to 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, PhD Symposium at McCombs Business School (UT Austin), Financial Management Association (FMA) 2020 annual meeting, 33rd Australasian Finance and Banking Conference, Annual Meeting of the Swiss Society for Financial Market Research (SGF Conference), 11th Emerging Markets Conference, Southern Denmark University 5th Finance Workshop, PhD Mentoring Day at 28th Finance Forum (AEFIN), Southern Finance Association(SFA) 2020 annual meeting, 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: akumarac@ust.hk

Introduction

Consumers primarily take out mortgages to finance housing. US consumers made about 150 million mortgage applications between 2000 and 2008, and interestingly, one fifth of these were rejected. Underneath these rejections lies an interesting pattern: the most frequent rejection reason was consumers' credit history, not debt-to-income ratio. Specifically, 28% of all rejections, or 39.4% of all rejections that mentioned a denial reason, were due to consumers' credit history (8.34 million such rejections). Debt-to-income ratio, which captures borrowers repayment ability, accounted for just half as many rejections. Given high mortgage rejection costs that include $300–400 in application fees and increase in future rejection probability and interest rate, it is surprising that more consumers were rejected for having bad credit history than for having low repayment ability. Do retail consumers have imperfect information about their creditworthiness? If they had perfect information, those with bad credit history would not have applied, and would have saved the application and rejection costs. Hence, imperfect information about their creditworthiness seems to be a reason consumers make erroneous credit-related decisions.

Multiple evidence highlight that consumers' information of their creditworthiness is shockingly inadequate. Survey of Consumer Expectations (SCE) Credit Access Survey reveals that about 12% of the US consumers don't know their credit score, and 20% have either never checked their credit reports, or checked it more than two years ago (FRBNY, 2013–2020). Similarly, ~37% of the respondents in a US mail survey study wrongly assessed their credit rating (Perry, 2008). When faced with uncertainty over creditworthiness, one may not even apply for credit anticipating a rejection despite needing credit. Such consumers are commonly referred to as the *discouraged borrowers*. SCE data show that among those consumers unlikely to apply for mortgage or refinance, 13% are discouraged. Similarly, across all credit types, a staggering ~15% of the US households are discouraged (Survey of Consumer Finances, 1998–2007). Under uncertainty and poor understanding of creditworthiness, many potential borrowers might be discouraged from applying for credit because they overestimate the rejection probability.

The purpose of this paper is to examine the effect of lowering consumers' economic cost of accessing their credit-history information on credit market outcomes. Specifically, this paper evaluates the causal effect of providing consumers free credit reports—an authoritative information source of their creditworthiness—on mortgage demand and approval ratio. This question is important because on one hand consumers seem to have poor understanding of their creditworthiness, on the other hand they rarely seek to access their credit reports. Out of approximately 1 billion credit reports generated annually in the US in early 2000's, a mere 1.6% were used by consumers (Avery, Calem, & Canner, 2004). High economic costs of accessing the reports may contribute to its low usage among consumers, as accessing it involves, in addition to monetary and search cost, the costs imposed by limited financial literacy and awareness.

Making the credit reports free had three main effects on the mortgage market outcomes. First, mortgage demand and approval ratio increased, and subprime population fraction decreased. Second, good quality borrowers seem to be behind the increased demand: origination was higher in more creditworthy areas and among prime consumers, and new mortgages were less likely to be defaulted upon. Finally, more first-time homebuyers took out mortgages, and lenders experienced improvements in their financial performance.

What mechanism might be at play here? This paper proposes that consumers self-assess their creditworthiness better when economic cost of accessing credit reports is lowered, ultimately shaping the credit market outcomes (*consumer self-learning*). Lenders rely on credit reports to assess creditworthiness of applicants as these reports convey crucial financial information on consumers e.g., their credit history and borrowing capacity (Figure I). Thus, before making a credit application, by using the information in their credit reports, consumers can acquire accurate knowledge of their creditworthiness.[1] With this, they can sort themselves better in the credit market: good borrowers may stay-in/enter the market, while those with bad credit history may search for a suitable lender, say a subprime lender, or do not apply/exit the market. This sorting leads to changes in the credit demand and approval ratio.

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 US—the enactment of the federal *Fair and Accurate Transaction Act of 2003* (FACTA)—that led to a close-to-exogenous reduction in the cost of 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 been allowing their residents free credit reports before the FACTA. So, consumers from all states except these seven

[1] Inaccurate self-assessment of creditworthiness, contributed by lack of financial knowledge and complexity of debt products, leads to worse financial outcomes (Courchane, Gailey, & Zorn, 2008). Further, consumers underestimate their student (credit card) debt by as much as 25% (37%) (Brown, Haughwout, Lee, & Van Der Klaauw, 2011). Also, credit reports itself may contain mistakes large enough to affect assessment by lenders (Avery et al., 2004).

(the pre-FACTA states) saw a reduction in cost of the reports in 2005. The law also raised general awareness about the reports, thereby reducing the overall economic costs of the reports.

This paper deploys the aforementioned reduction in cost of credit reports in a difference-in-differences (DID) design to establish the causal link. In this design, 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 analysis filters out the effect of local economic conditions on the outcomes of interest.

This empirical strategy 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 empirical investigation focuses on mortgage-related outcomes aggregated at the census tracts level—a sub-county geographic area roughly encompassing a population of about 4,000. Comparing the census tracts in the control and treated counties, the DID estimates suggest that free credit reports resulted in an increase of 13.8%–16.0% in mortgage applications, and an increase of 1–2 percentage points in the approval ratio. In dollar terms, the demand increase was about \$38.1 billions, and increase in origination due to higher approval was about \$5.5 billions, aggregated only across the treated bordering counties (not all counties from all states). Finally, the growth rate of house prices increased, albeit statistically weakly, by ~1.7 percentage points. These estimates account for *Census Tract* and "$Border \times Year$" fixed effects, and are robust to the addition of time-varying controls capturing county- and state-level economic conditions.

Since the event took place in 2005, close to the 2008 financial crisis, a natural question arises: whether the increase in demand came from occupancy- or investment-seeking borrowers? Within the owner-occupied mortgages, the estimates for increase in demand and approval ratio are similar to those observed in the whole sample. Then, tests for changes in the composition of demand show a minor increase of 1 percentage point in the non-owner-occupied mortgages as the fraction of total applications (or, successful applications). In essence, surge in demand appears to have mostly come from the occupancy-motivated borrowers.

A spike in mortgage origination could be detrimental to the economy if the defaults too increase subsequently. Over a six-year period since inception, mortgages from the treated areas originated in the event year were less likely to be defaulted upon than those originated in the pre-event year, after accounting for the trends in the control areas. Besides, this superior performance persisted through the 2008 financial crisis. The overall result—more origination but fewer defaults—suggest that the borrower pool improved in response to the free credit reports.

The next set of tests examines the *consumer self-learning* channel. To start with, the subprime population fraction in the treated counties relative to the control began to decline a year after the event. The reason could be market entry by new prime borrowers and/or subprime-to-prime transition by some. Any which way, the treated areas saw slight improvements in creditworthiness, plausibly from improved creditworthiness knowledge among consumers.

Three further mortgage-market measures shed light on the channel. *First*, if consumers self-learn their creditworthiness from credit reports, then in aggregate, the likelihood of rejections due to credit history should decrease. Indeed, the credit-history denials decreased in the treated areas by 0.3 percentage points, though significant only in the *ex-ante* high rejection areas, while debt-to-income denials did not decrease. *Second*, the certainty over application acceptance should be higher among borrowers that are better-learned of their creditworthiness, allowing them to apply to fewer lenders *ex-ante* (confident searching) and saving the costs of multiple applications. Consistent with it, withdrawal of in-process applications dropped by 0.9 percentage points in the treated areas. *Finally*, among the originated mortgages, the fraction of first-time homebuyers increased by 1 percentage point in the treated areas, clearly showing entry by *new* borrowers.

It is important to examine if the results are supply-driven, instead of demand-driven as proposed. Lenders may react to the knowledge that their potential clients could now access free credit reports under the FACTA. Multiple evidence, however, favor the demand-explanation.

To start with, the mortgage applications increased, and consumer interest in free credit reports, measured using Google Search Interest for the keyphrase *"Free Credit Reports"*, heightened in the treated areas after the event. Then, interest rates on the mortgages purchased by Government-sponsored Enterprise (GSE) in the treated areas rose by ~1.1 basis points, controlling for credit score and property attributes, and despite a highly elastic mortgage supply. Had the increased origination been supply-driven, the interest rates would have decreased.

Furthermore, three heterogeneity tests show that the demand and approval ratio responded to different degrees across areas in a manner consistent (inconsistent) with the demand- (supply-) related factors. *Firstly,* there was no difference in areas with high and low lender density, a supply-related factor. *Secondly,* effects were stronger in areas with high consumer creditworthiness, a demand-related factor. *Finally,* the increase in approval ratio was more for the lowest income-quartile applicants. As the marginal propensity to lend to low-income applicants is low, (Agarwal, Chomsisengphet, Mahoney, & Stroebel, 2018), approvals to them may not have increased without lender-perceived improvement in the pool. Overall, tests examining these heterogeneity and interest rates underscore that demand-related factors were instrumental in the increased origination.

Finally, this paper presents three supplementary results that aid in interpreting the findings. *First,* there was no increase in the fraction of mortgages sold to non-GSEs, thus private securitization cannot explain the increased origination (Keys, Mukherjee, Seru, & Vig, 2010). *Second,* analysis of the GSE-purchased mortgages show that increase in origination to prime borrowers (credit score ≥ 620) was staggering 30 times larger than that for the subprime borrowers, ruling out subprime origination as the reason for the increase (Mian & Sufi, 2009). *Finally,* with a few caveats that are detailed later, data show that banks with *ex-ante* higher mortgage activity in the treated areas enjoyed increase in net interest margin, return on equity, and return on assets.

This paper primarily relates to the literature on effects of information provision on credit market participants. This is the first paper to argue that the borrower pool in the market improves when economic cost of credit reports goes down for consumers. In a field experiment, Homonoff, O'Brien, and Sussman (2019) find that borrowers are less likely to default when they are provided information on their FICO® scores. Similarly, Mikhed (2015) finds that customer enrollment in their bank's free FICO scores program is associated with decreased credit utilization and default probability and increased credit card spending. Kulkarni, Truffa, and Iberti (2018) show that delinquency decreases with product disclosure for sophisticated borrowers and with product standardization for unsophisticated borrowers. Dobbie, Goldsmith-Pinkham, Mahoney, and Song (2016) document increased credit card and mortgage borrowing after removal of bankruptcy flags from credit reports. Kronlund, Pool, Sialm, and Stefanescu (2019) find that consumers become more sensitive to expense ratios and short-term performance after disclosure of fee and performance for 401(k) plans becomes mandatory.

This paper also relates to the extensive literature that shows that 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 free credit reports, that aid consumers in learning their creditworthiness, result in increased credit demand and lower defaults.

All in all, we learn that imprecise creditworthiness information among consumers plays a role in erroneous credit-related decisions: frequent mortgage rejections due to bad credit history and being *discouraged* from applying for credit. This paper shows that provision of credit reports to consumers at reduced economic cost improves credit market outcomes. Specifically, free credit reports led to increased mortgage demand and approval ratio, more origination to creditworthy borrowers, and reduced defaults and subprime population fraction, and also resulted in more entry of first-time homebuyers and better financial performance of lenders.

These findings pertain to the mortgage market, but they are equally relevant to any general consumers credit-related decision-making under imperfect knowledge of their creditworthiness. Moreover, as these findings are causal in nature, a policy intervention aimed at educating consumers of their creditworthiness may yield similar results.

The rest of the paper is organized as follows. Section (1) describes the US laws related to consumers' access to credit reports, Section (2) presents the research design, and Section (3) details the data used in this paper. Section (4) presents the results for mortgage market outcomes, and Section (5) contains supplementary results that aid interpretation of the findings. Finally, Section (6) concludes the paper.

1 US Laws Governing Consumers' Access to Credit Reports

FACTA granted all US consumers a right to receive a free credit report annually from the three national credit reporting agencies. Before FACTA, the FCRA governed consumer credit information-related laws. FCRA too had specific, but restrictive, provisions that allowed consumers the right to see the contents of their credit reports, except for the credit score. 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).[2] 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.

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).[3] Thus, only 0.084% of all credit reports generated were disclosed to consumers as a result of their own request.

In addition to the federal provisions under the FCRA, pre-FACTA states had enacted local laws to allow state residents free credit reports (Panel A of Figure II). For example, Colorado enacted the law on April 21, 1997 through The State of Colorado SENATE BILL 133. 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

[2] 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.

[3] 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.

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, the paper uses a DID setting in which the seven pre-FACTA states constitute the control group, and the states bordering these states constitute 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. The counties at the border of these states, the focus of the empirical analysis, are shown in Figure (III). The event is year 2005 when www.annualcreditreport.com was established to distribute the free credit reports.[4]

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

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

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 $t \geq 2005$. This is because, even though FACTA was passed in Dec 2003, the centralized website for distribution of free credit reports was rolled out from Dec 1, 2004 to Jan 9, 2005. $Treatment_{icsj}$ is 0 for all the census tracts i in counties c from pre-FACTA (control) states j, and is 1 for those from treatment states s. α_i represents "*Census Tract*" fixed effects and controls for any time-invariant differences at a highly granular geographic level, as a census tract covers an area encompassing about 4,000 population. Further, clustering the standard errors at the county level accounts for correlation in error terms for the observations from census tracts belonging to the same county.

The key identifying assumption for this DID estimate to be causal is that the treatment states

---

[4] 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.

would have had similar trends to the control states in the absence of the treatment (parallel-trends assumption). Though it is unverifiable, Figure (IV) examines the trend of mean approval ratio across the two groups before the event. It seems to be parallel (Panel A).

Furthermore, Panel B of Figure (IV) plots the coefficients (β_k) from regression of *Approval Ratio* according to the following 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. 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 no significant difference exists in the approval ratio between the treated and control census tracts before the event, but the difference 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.

It is also necessary to account for the local economic conditions when analyzing mortgage market outcomes. The regression specification in Equation (1) does this through *Economic controls* and "*Border* $\times$ *Year*" fixed effects. *Economic controls* are county- and state-level variables capturing local economic conditions—annual growth rate of county income per capita, county aggregate employment, and state gross domestic product (GDP). "*Border* $\times$ *Year*" fixed effects manifests in Equation (1) as $\gamma_{j,t}$. Here, j refers to the border of a pre-FACTA or control state. Consider a control state, Colorado. All census tracts from counties at the border between Colorado and all the adjacent states—WY, UT, AZ, NM, OK, KS, and NE—are grouped as one unit and take the same value (j). This grouping ensures that all census tracts at the border counties of Colorado (a control state) are compared with all census tracts lying at the border counties from the surrounding seven treatment states *only*. In other words, this also ensures that a control census tract from Colorado is not compared with a treatment census tract from Alabama, which should act as a treatment group for Georgia. Most importantly, $\gamma_{j,t}$ flexibly controls for any time-varying regional economic shocks that may affect bordering states across different years.

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. Recall that 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 these play in a mortgage process.

Further, usage data of credit reports do suggest that its usage was low in general (is low even now, more than a decade after it became free), and was higher in the control states before FACTA enactment. *First,* as described previously, 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, only 0.84 million consumers accessed their credit report, compared to about 16.8 million applications annually around 2004. So, even if all these reports were requested by mortgage-seeking borrowers before applying, the fraction of mortgage borrowers who checked their credit reports proactively is less than 5% .[5] *Second,* 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 Results section). *Third,* the data from testimony in the US 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).[6]

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 plot of search interest for the key phrase *Free Credit Report* from 2004 to 2010 in Panel A of Figure (V) reveals that the search interest heightened in Jan 2005, coinciding perfectly with the establishment of the website.[7] *Second,* the plot of differential search interest across the treated and control states using the Interest-by-subregion Google Trends data also suggests that consumer interest in free credit

[5] 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).

[6] 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).

[7] 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(US 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.

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.[8] Also, some anecdotal evidence suggest that the website issued about 52 million credit reports to consumers in the first two years(Wikipedia, n.d.).

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 US 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 US. HMDA data provide application-level details on applicants' 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.[9] 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.[10]

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;

[8] 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.

[9] 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).

[10] 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)).

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), leading to the final sample of 12,044 census tracts—7,054 treated and 4,990 control—and 90,353 "*Census Tract* $\times$ *Year*" observations.

Though the coverage of mortgages in HMDA data is the largest, it lacks some key application-level information, such as the credit score of the applicant. 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 US. 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 US Department of Housing allows to map the zip3-state level information to the counties in the sample.[11] Aggregation of these observations (properties from the zip3-state areas from bordering counties of the treated and control states) to zip3-state level yields 225 unique zip3-states, or 7,711 "*Zip3-State* $\times$ *Quarter*" observations.

The third important 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 commercial banks' RSSD ID, which is used in the Call Reports, requires several steps. HMDA data identify lenders by a combination of an agency code (lender's regulator) and a respondent ID. A respondent ID equals the Federal Deposit Insurance Corporation (FDIC) Certificate ID if the lender's regulator is the FDIC, and it equals the Office of the Comptroller of the Currency (OCC) charter number if the regulator is the OCC. Many mortgage companies operate as affiliates of commercial banks. Such lenders can be matched to the parent bank

[11] 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.

using the HMDA Ultimate Panel data. From 2004 onward, this panel also started providing RSSD ID for each respondent ID. Thus, matching of lenders across HMDA and Call Reports is feasible using respondent ID and agency code until 2003 and using RSSD ID from 2004 onward. Further, if a lender reports a parent in HMDA Ultimate Panel, first lender's own respondent ID is used for matching with Call Reports. If no match is found, its parent's respondent ID is used. If both the HMDA lender and its parent yield a successful match with a commercial bank, the match found using parent's ID is kept.

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 (FRBNY, 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 that are denied for credit history or debt-to-income ratio; and the fraction of total applications that are withdrawn by applicants while still under processing. All these ratios are calculated as the fraction of total applications in a given census tract.

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.[12]

Panel B of Table (I) shows the comparison of the treatment and control groups in the pre-

[12] 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.

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 facilitate comparison, while the differences in mortgage-related outcomes raise the concern that these groups may also differ on some unobservable characteristics that may lead to endogeneity issue. However, recall that a DID setting can accommodate pre-existing differences between the treatment and control subjects so long as they satisfy the *parallel-trends* assumption.

4 Results

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

§A Survey evidence on usage of credit reports/scores and discouraged borrowers

The SCE Credit Access Survey data provide valuable insights into the credit reports/scores usage and the *discouraged borrowers* phenomenon. Given that it is a panel survey, fixed effects at "Year$\times$month" level, together with clustering of standard errors at the same level, filter out the time trend. Also, use of the *weight* variable included in the micro data as sampling weight in regressions make the estimates derived from these representative of the US population.

Table (II) presents the results. Column (1) shows that about 8% of the respondents have never checked their credit reports/scores. Column (2) suggests that a staggering 20% respondents have either never checked, or checked it more than two years ago (infrequent checkers). Finally, column (3) shows that ~12% of respondents don't know their credit score.

Discouraged borrowers population also remain high. A respondent is *discouraged* if he/she responds "I don't think I would get approved". Column (4) shows that the proportion of the discouraged borrowers among those who were, over the next 12 months, very unlikely or somewhat unlikely to apply for mortgage/home-based loans or refinance or assigned less than 10% probability to these actions, is 13%. Further, column (5) and (6) suggest that the respondents who check their credit reports/scores infrequently and those who are unaware of their credit score are more likely to be discouraged from applying for mortgage.

Together these findings suggest that even after more than a decade since credit reports have been free in the US: (i) a significant proportion of consumers do not check credit reports; (ii) are unaware of their credit score; (iii) suffer from the discouraged borrower phenomenon.

§B Baseline Results

§B.1 Mortgage demand, approval ratio and house prices

The outcome variables of interest—mortgage demand (scaled applications), approval ratio, and house prices—are aggregated at the census tract level. Equation (1) shows the regression specification. "$Treat \times Post$", the coefficient of interest, captures the change in the outcome variable in treated areas relative to the control areas after free credit reports became available in the former. The specification includes $Census\ Tract$ fixed effects and "$Border \times Year$" fixed effects. As discussed previously, $Census\ Tract$ fixed effects control for any existing difference across census tracts, and "$Border \times Year$" fixed effects flexibly control for regional shocks that may arise across different years.

Table (III) shows the regression results. Column (1) shows the result of regressing mortgage demand using the DID specification without any controls, while column (2) shows the results after adding the controls for economic characteristics (annual growth rate of county income per capita, county aggregate employment, and state GDP). We see that scaled application increased in treated census tracts by 13.3–15.4, a 13.8–16.0% increase in applications (over the pre-treatment average of 96.6 in the treated census tracts). Average mortgage size in the treatment counties in the pre-treatment period is ~\$150,597. Thus, the consumer demand for mortgages increased by about \$2.0 millions per 1000 adults per census tract (\$150,597 $\times$ 13.29), by about \$5.4 millions per treated census tract (\$2 millions $\times$ 2.7 thousand adults per census tract), or, by about \$38.1 billions in the treated border counties after free credit reports became available (\$5.4 millions $\times$ 7,054 treated tracts).

Coefficients on "$Treat \times Post$" in columns (3) and (4) estimate the effect of free credit reports on the mortgage approval ratio. The ratio increased by about 1–2 percentage points in the treatment census tracts after free credit reports became available. This increase represents ~5.22 more successful applications per treated tract (96.6 applications per 1000 audlts in pre-treatment period $\times$ 0.02 $\times$ 2.7 thousand adults per treated census tract), ~36,827 more successful applications in the entire treated area (2.62 applications $\times$ 7,054 treated census tracts), or a

~$5.5 billion increase in successful mortgage origination across all treated census tracts (18,348 $\times$ $150,597 average mortgage amount per application). The increase in approval ratio, together with an increase in mortgage demand, perhaps suggest an improvement in the borrower pool. Recall that the availability of free credit reports does not materially alter borrowers' financial condition, such as wealth and collateralizability of existing assets. The only changes it brings is that borrowers are more likely to be aware of their credit history and the information lenders would see when evaluating them.

Coefficients on "$Treat \times Post$" in columns (5) and (6) quantify the effect of free credit reports on the growth rate of house prices using the house price index from Bogin, Doerner, and Larson (2016). This index is available at the census tracts level and starts in the year 2000 with the value of 100. The coefficients, though statistically significant only at the 10% level, suggest that the growth rate of house prices increased in the treated areas by 1.7–1.8 percentage points after the event. This is similar to the finding of Di Maggio and Kermani (2017): an increase of 3.3 percentage points in the growth rate of house prices resulting from a 10% increase in mortgage origination in their sample.

Since this natural experiment occurred in the year 2005, close to the 2008 financial crisis, it is important to investigate whether the increased demand came from occupancy-seeking borrowers or investment-seeking borrowers. Panel A of Table (IV) shows the results from re-estimating the baseline regressions for only the mortgages in the owner-occupied category. The new estimates are broadly similar to those from the baseline specifications: an increase of 12.8–14.9 (13.2–15.4%) in scaled application, and an increase of 1 percentage point in approval ratio.

Regression specifications in Panel B of Table (IV) examine whether there was a change in the fraction of mortgages in the non-owner and owner-occupied category at the application/origination stage. The outcome variables are the fraction of *total* mortgages that are non-owner-occupied (columns 1 and 2), and the fraction of *originated* mortgages that are non-owner-occupied (columns 3 and 4). The coefficients in all columns suggest a small increase of 1 percentage point in the fraction of not-for-occupancy mortgages in the application and origination stage. As ~86% of the applications in the sample are for occupancy purposes, this minor increase in non-occupancy mortgages at the origination stage is economically small, allowing us to conclude that free credit reports led to an increase in demand mostly for occupancy purposes.

§B.2 Mortgage defaults

The increased mortgage demand and approval ratio suggest that more consumers took mortgage debt. Were these consumers good borrowers: were they more, or less, likely to default on these mortgages? It is plausible that free credit reports led to an increased influx of *just-marginal* borrowers who over-borrowed and later defaulted, which may be detrimental to the economy. Another possibility is that free credit reports-induced learning among borrowers led to better sorting and improved pool, so good quality borrowers took those mortgages and were less likely to default.

The GSE data allow us to compare the default rates on the mortgages originated in the treated and control areas in the pre-event year 2004 and the event year 2005 (years vintages). Among all the mortgages originated in treated and control areas in a given vintage year, I calculate every month the fraction that had missed a scheduled payment by 30–59 days for the first time ($\text{Def}_{2004,age}$ and $\text{Def}_{2004,age}$, where age is measured in months since origination). An adjusted default rate, defined below, then measures the differential default probability:

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

A negative (positive) adjusted default rate implies that the mortgages from the treated areas are less (more) 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 adjusted default rate is -0.012 percentage points (*p-value = 0.0000*), suggesting that, mortgages from the treated areas were 0.012 percentage points less likely to be defaulted upon within six years period since origination relative to those from the control area after free credit reports became available. Moreover, the differential default probability around the financial crisis of 2008 shows a (positively) noteworthy pattern. Recall that the age greater than 48 months in this plot corresponds to the bust years post the financial crisis (48 months after 2005 is 2009). We see in the plot that even during these bust years, the mortgages from the treated areas were less likely to be defaulted upon.[13]

[13] 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.

Overall, we learn from the baseline results that free credit reports stimulated the mortgage demand, raised the approval rates, and resulted in lower *ex-post* defaults, consistent with an improved borrower pool presumably due to well-informed borrowers.

§C Mechanism: Consumer Self-learning Channel

What follow next are the results related to the consumer *self-learning* mechanism. Under this mechanism, credit reports aid consumers in better self-assessment of their creditworthiness. Before making a credit application, by relying on the information in the credit report, one can self-assess creditworthiness more accurately. If one learns from the report that the creditworthiness signal in the report is good enough for the application to be approved, he or she would continue with the application. Otherwise, one may delay the application to take steps to improve the credit records, correct any inaccurate information therein, search for a more suitable lender (say, a sub-prime lender), or altogether abandon the application. Thus, owing to the credit reports, learned borrowers can sort themselves in the credit market better: creditworthy borrowers stay-in/enter the market, while those with bad creditworthiness may search for a sub-prime lender, or exit from it. Consequently, the quality of applicant pool improves, resulting in an increase in approval ratio. Also, total credit demand may increase or decrease depending on (1) whether, on average, consumers overestimate or underestimate their creditworthiness in the absence of learning their true type; (2) the proportion of consumers who are unaware of the existence and role of credit reports in the result of their credit application. Hence, under this demand-driven mechanism, the lower economic cost of credit reports leads to its wider usage and increased self-learning among consumers of their creditworthiness, and shapes the credit demand and approval ratio.

§C.1 Reduction in subprime population fraction

A direct implication of better self-assessment of creditworthiness among consumers is that new prime consumers may enter the market and/or some (just-) subprime consumers could transition into the prime category. If this channel is at work, the subprime population fraction in the treated counties would decrease relative to the control counties. The Equifax data on subprime population fraction at the county level facilitate this analysis. Equifax classifies consumers with credit score less than 660 as subprime.

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. Similar pattern emerges when mean, instead of median, is used, 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 improvement in the borrower pool coming from consumer self-learning.

§C.2 Contraction in credit history-related rejections

The probability of mortgage rejection due to credit history speaks to the self-learning among consumers. To understand why, recall that free credit reports do not change financial condition of consumers. The only it brings is that consumers get to know their credit history and what lenders see about them, free of cost. If we were to assume that knowledge about their credit-worthiness among the consumers was perfect before the experiment, and self-learning did not increase after free credit reports became available, the probability of a rejection due to credit history would remain the same. Further, the probability of rejections due to debt-to-income ratio, the second most frequent reason for mortgage rejection, too, would remain unchanged.

Using the generalized DID specification in Equation (1), I test whether the probability of mortgage rejection due to credit history and debt-to-income ratio change after the experiment. The two outcome variables are the fraction of total applications denied due to credit history and due to debt-to-income ratio. The regressions are estimated over the 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 (High rejection areas).[14] The rationale for testing over this sub-sample is that areas where consumers were more often denied mortgages prior to the experiment are the ones more likely to reap benefits by self-learning their creditworthiness.

Table (V) shows the results. In columns (1) through (4) we see that the fraction of mortgage applications denied due to credit history decreases by 0.3 percentage points in the treatment

[14]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.

census tracts relative to the control census tracts. However, the coefficients are statistically significant only in the *ex ante* high rejection rate areas (in columns 3 and 4). Plausible reasons for significant reduction only in *ex-ante* high rejection areas are economic as well as econometric. The economic reason is that self-learning is more likely to help consumers in high denial areas, or in other words, the value of self-learning is smaller for consumers who already enjoy higher acceptance rates. The econometric reason is that we cannot estimate a reduction in rejections due to a given reason if mortgages are not likely to be rejected in the first place.

Not only the credit history-related rejections decreased (columns 3 and 4) in the treated tracts, but the debt-to-income ratio-related rejections did not decrease (columns 5 through 8). Though statistically weak, these results are suggestive of increase in consumer self-learning of their credit history.

However, a caveat applies to these results. Under HMDA regulations, reporting the rejection reason is not mandatory, so if lenders were to adjust their reporting after the experiment, we would incorrectly attribute these changes to consumer self-learning. However, this limitation may not be critical for two reasons. First, lenders reported the rejection reason for over 70.81% of the total rejections over the 2000–2008 period. Second, these results are unbiased to the extent that the incentive of the lenders to report rejection reasons remained unchanged across the treatment and control census tracts in 2005, the experiment year.

§C.3 Increase in consumer search accuracy (Drop in in-process application withdrawals)

A second market level outcome, in-process application withdrawals, can shed light on self-learning among consumers as this represents their uncertainty in their search for a mortgage lender. To discern this link, we need to understand a typical mortgage application process. Owing to the uncertainty over their application, and in a bid to secure better terms, consumers apply to multiple lenders for mortgages, and incur multiple non-refundable application fees (~US $400) along the way.[15] At the end, they will finalize their mortgage with one lender, the one with higher certainty or better terms, and withdraw the remaining in-process applications at other lenders.[16] Importantly, consumers withdrew about 12% of all mortgage applications

[15] 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."

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

over the 2000–2008 period, representing that uncertainty over application remain high.

An application is *withdrawn* if consumer withdraws it before the lender has decided the outcome. If consumers learn their creditworthiness better under the free credit reports regime before making a credit application, they should be more certain over the approval probability and mortgage terms. Hence, these consumers would apply at fewer lenders *ex-ante*, save multiple application costs, leading to fewer withdrawn applications in aggregate.

Table (VI) shows the results of regressing the fraction of total mortgages that are withdrawn in-process, using the specification in Equation (1). We see that the fraction drops by 0.9–0.11 percentage points in the treatment areas relative to the control areas. This decrease represents ~2.34 fewer in-process withdrawals per treated census tract, or ~16,513 fewer withdrawn applications over the treated border counties. At an average cost of ~$400 per withdrawn application, consumers saved ~US $6.6 millions in upfront charges, presumably resulting from better self-learning from free credit reports. Furthermore, this measure of consumer self-learning is free from any influence from lenders' responses to the experiment: application withdrawal is a consumer decision, and lenders have no control over it in any way.

§C.4 Increase in first-time homebuyers

As alluded to previously, about 15% of households in the US do not apply for credit because they anticipate rejection (Survey of Consumer Finances, 1998–2007). While the fraction of these consumers who overestimate the rejection, and thus, overly suppress their credit demand is unknown, the provision of free credit reports may aid them in better self-assessment of creditworthiness. If the fraction of first-time homebuyers in the mortgage market increases after the event, we can conclude that first-time homebuyers on average overestimate the rejection probability and that free credit reports reduce the *discouraged-borrowers* phenomenon.

While the HMDA data do not contain the information on whether an applicant is a first-time homebuyer, the GSE data do. From the location data of these mortgages (available as the 3-digit zip codes), I approximately identify the mortgages for the properties residing within the border counties of the treatment and control states, using the steps outlined in the Data section. An issue with the data is that about 6.71% of the observations in the sample have information missing on the first-time homebuyer status. Hence, the outcome variable of interest, proportion of mortgages by the first-time homebuyer, is defined as the fraction of all GSE mortgages, or as a fraction of all GSE mortgages with the valid first-time homebuyers information. The regression equation for this test is specified at the zip3-state level, different from the census tract-level aggregation in earlier regressions:

$$Y_{zjt} = \beta_0 + \beta_1 \text{Treatment}_{zj} \times \text{Post}_t + \delta \times \text{Economic controls} + \alpha_z + \gamma_{jt} + \varepsilon_{zjt} \qquad (3)$$

Here, z indexes the areas delineated by a 3-digit zip code and state, sample is limited to zip3-state areas that come under the border counties of treatment and control states. The rest of the terms are the same as in Equation (1). Table (VII) show the coefficients obtained from this regression. From the coefficients in columns (1) through (4), we see that the percentage of first-time homebuyers increased by 1 percentage point in the treatment areas relative to the control areas. This confirms that free credit reports addresses to some extent the issue of demand suppression among first-time homebuyers.

One may have a concern that the sample of mortgages used in this regression suffers from a selection issue because GSEs only purchase conforming mortgages. However, such selection would only be a concern if GSEs' incentives or mandate to purchase first-time homebuyer mortgages relative to their overall purchase from the treated counties increased relative to the control counties in the year 2005. Such a time- and location-specific change seems improbable.

§D 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 free credit reports became available. A demand-driven mechanism seems to be at work here, since an increase in demand and a decrease in in-process application withdrawals are consumer-determined, and are mostly independent of lenders action. However, supply-driven increase in mortgage origination is also plausible. Lenders were treated with the knowledge that consumers could now get free credit reports. So, in response, lenders could increase the mortgage supply and approval rates. In a limiting case, one may argue that the entire increase was due to lenders action, and not due to increasing demand from consumers. This section employs two set of tests—equilibrium (contracted) interest rates and heterogeneity-based tests—to argue that while the role of supply-side factors may not be ruled out, the origination increased primarily fue to the demand-side factors.

§D.1 Distinguishing supply and demand: Equilibrium interest rate

In a supply-demand framework, an increase in quantity and price of a good occurs when demand curve shifts outward and the supply curve stays the same or shifts outward inadequately causing the equilibrium price to rise for the market to clear. Since previous results already established that demand for mortgages increased in the treated areas, an increase in mortgage interest rates would confirm that the primary mechanism is demand-driven. Application-level interest rates from the GSE data allows us to examine this. The specification is similar to Equation (3), but specified at the loan level, as follows:

$$\mathrm{Y}_{izjt} = \beta_0 + \beta_1 \mathrm{Treatment}_{izj} \times \mathrm{Post}_t + \delta \times \mathrm{Loan\ controls}_i + \alpha_z + \gamma_{jt} + \varepsilon_{izjt} \tag{4}$$

The loan level controls include credit score, debt-to-income ratio, combined loan-to-value (loan-to-value ratio inclusive of all loans secured by a mortgaged property), number of units comprising the related mortgaged property, and percentage of mortgage insurance coverage. Table (VIII) shows the results. Column (1) is without any control, while Column (2) includes the above loan-level controls. We see that interest rate in the treated areas increased by 0.9–1.1 basis points for a loan with the same observable characteristics, including credit score, debt-to-income-ratio and so on. This finding contradicts a supply-driven increase, but is consistent with a demand-driven increase in origination.

One may wonder if the increase in interest rate is due to increase in risky borrowers in the treated areas. Note that these specifications control for credit scores. Also, in Section (5) we would see that prime mortgages increased 30 times more than subprime mortgages increased in the treated areas relative to the control in the GSE data, so interest rate increase is not due to increase in risky borrowers. Note also that the interest rate increase is small. It may be for two reasons: firstly, the supply of mortgages in the US is highly elastic because of the large-scale GSE purchase of conforming mortgages in the secondary market; secondly, the interest rates for 30-year mortgages fluctuate little across lenders. Even if the increase in interest rate were statistically non-significant, it would be valid to conclude that results are demand-driven, as interest rates would need to decrease in a statistically significant manner for results to be supply-driven.

§D.2 Distinguishing supply and demand effect: Heterogeneity tests

The heterogeneity in the effects of free credit reports may further highlight the role of supply and demand factors. If the results were driven primarily by lenders, the heterogeneity would exist in lenders characteristics, else across consumer characteristics. The effects may vary with either set of characteristics if these characteristics are correlated. Following tests examine heterogeneity across density of lenders, and consumer creditworthiness and income.

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, I classify a census tract 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* (see Footnote 14).

Table (IX) shows the results obtained from separately estimating regression Equation (1) for low and high lender density areas. Columns (1) through (4) show the results for amount of mortgages originated (in 1000 USD) per adult in a census tract. $Treat \times Post$ coefficients are smaller and have weaker statistical significance for high density areas (in specification 2 and 4) than for low density areas (in 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. Columns (4) through (8), repeat 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. 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.

Heterogeneous effects by creditworthiness of consumers: If credit reports aid consumers in assessing their creditworthiness and lenders approve creditworthy borrowers, effects of free credit reports should be higher in areas which had *ex-ante* high fraction of creditworthy consumers. To examine this, I classify a county as having high creditworthiness if its subprime population fraction in year 1999 is less than the *regional mean* subprime population fraction (see Footnote 14).[17]

Panel A of Table (X) shows the results obtained from separately estimating regression Equation (1) for counties with high and low creditworthiness. Coefficients in columns (1) and (2) show that scaled application increased by 16.7–18.1 (17.3–18.7%), while those in columns (3) and (4) show that approval ratio increased by 2 percentage points in treated high-creditworthiness counties relative to control counties with similar creditworthiness. Interestingly, coefficients in columns (5) through (8) show that the increase in low-creditworthiness areas was far smaller: scaled application increased by 8.5–10.6 (8.8%–11%) and approval ratio increased by 1 percentage point, and had weaker statistical significance. These results are consistent with the self-learning channel in which borrower pool improves from better sorting by borrowers in the credit market through increased credit reports usage.

Admittedly, county-level creditworthiness measure is noisy. The number of payday-lenders establishments in a locality may serve as a more precise proxy for creditworthiness as payday lenders tend to operate in subprime areas (Prager, 2009) and data on their location are available at 5-digit zip code level from Census Bureau (2000–2008). However, only Colorado and the bordering states in the sample allow unrestricted payday lending activity, while other sample states restrict it to varying degrees (Prager (2009) and Bhutta (2014)). Thus, I repeat the above regression by taking number of payday lenders in a county as proxy for creditworthiness, and keep only counties at the border between CO, and surrounding treated states (WY, UT, AZ, NM, OK, KS and NE). A tract has high creditworthiness if the number of payday lenders in pre-event year 2004 in it is *less* than the mean number of payday lenders across all census tracts within these counties.

Panel B of Table (X) shows the results. Dramatic differences exist in the effect of free credit reports across high and low creditworthiness areas. Coefficients in columns (1) through (4) suggest that the scaled application increased by 68–72 and approval ratio increased by a staggering 6 percentage points in the treated high-creditworthiness areas compared to control high-creditworthiness areas. However, coefficients in columns (5) through (8) show that the increase was small for scaled application, at 43, and no significant increase at all in approval ratio in treated-low-creditworthiness areas compared with similar control areas.

The finding that the effect of free credit reports varies with creditworthiness of consumers,

[17] The choice of year 1999 follows from 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 may endogenously change with the housing boom. FRBNY and Equifax (n.d.) data on county subprime fraction start from year 1999.

a demand-side characteristics, corroborates a demand-driven self-learning mechanism. These results also rule out that subprime credit is the reason behind the increased origination.

Heterogeneous effects by income level of consumers: Free credit reports may affect borrowers of different income differently for two reasons. First, the cost of mortgage rejection is higher for low-income consumers. Second, marginal propensity to borrow is high for low-income consumers, while marginal propensity (among lenders) to lend to these consumers is low (Agarwal et al., 2018). Thus, any increase in approval ratio for the low-income consumer would strongly point to an improvement in the borrower pool rather than to an increase in lender willingness to originate disproportionately more mortgages to these consumers. However, we cannot infer the same for high-income consumers.

Table (XI) shows the results obtained from separately estimating regression Equation (1) for each income-quartile (here, quartile cut-offs are calculated yearly among the applications from areas in the sample). Coefficients in Panel (A) show that the scaled application increased in each quatrile except the lowest, and the effect is larger as we move from the lowest quartile (column 1 and 2) to the highest (column 7 and 8). Coefficients in Panel B shows that approval ratio increased by 1 percentage point in the lowest income-quartile, and is statistically significant at the 1% level (columns 1 and 2), while it did not increase statistically significantly for higher income-quartiles.

These findings are remarkable. If one were to argue that the increase in mortgage origination was supply-driven, we would observe a larger increase in the approval ratio for higher income consumers, because (i) the higher income consumers demanded more credit, as evidenced from the number of applications; and (ii) propensity to lend to high-income consumers is high. Instead, we see that the approval ratio increased more for low-income consumers. The preferred interpretation is that the scope of error in self-evaluation of creditworthiness and cost of rejection is higher for low-income consumers, making these consumers more likely to benefit from accurate self-assessment. This finding that the effect of free credit reports varies with the income level of consumers, a demand-side characteristics, is also consistent with a demand-driven mechanism coming through the self-learning channel.

All in all, the equilibrium interest rates and three heterogeneity tests consistently point towards a consumer-driven mechanism driving the increased origination.

5 Supplementary Discussion

This section discusses a few alternative mechanisms and also investigates the effect of free credit reports on commercial banks.

§A 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 et al., 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) for the outcome variables 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. The coefficients in column (1) and (2) show that there is no change (no increase) in the fraction of mortgages sold to the private entities involved in securitization (non-GSEs), while those in column (3) and (4) show that the fraction of mortgages sold to the GSEs increased, and those in column 5 and 6 show that the fraction of unsold mortgages did not change. Thus, there is no evidence of an increase in private securitization in the treated areas. Hence, we can rule out the private securitization as the reason behind increased mortgage origination.

§B 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). Relying on the location-based proxies of creditworthiness, findings in Table (X) suggest that effect of free credit reports was stronger in the prime counties/census tracts than in subprime counties/census tracts. Such location-based proxies are informative and widely used (Di Maggio & Kermani, 2017; Mian & Sufi, 2009), but are noisy. So, the next test uses application-level credit scores from the GSE data to revisit this question.

Table (XIII) shows the results of regressing separately the number of prime (credit score $\geq$ 620) and subprime *originated* mortgages in zip3-state areas using Equation (3). The coefficients in 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. However, coefficients in columns (3) and (4) imply that subprime mortgages increased by ~10 applications, which is 30 times smaller. These results imply that the increased origination did not disproportionately go to subprime consumers. Also note that the observation unit here is zip3-state, but in the previous regressions it is census tract, hence these coefficients are not comparable.

A caution applies to these results. Since the data in the above regressions are GSE-purchased mortgages, a change in prime-subprime composition reflects the combined effect of demand factors and changing GSEs' incentives to purchase these. However, two arguments allay this concern. First, as these findings are using a DID estimator, the changing GSE incentives would be an issue if their incentive to purchase prime mortgages in 2005 from the treated counties increased, but stayed the same or decreased from the control counties. Such precise change in incentives seems unlikely. Second, prior to 2007, GSEs sought to buy more subprime, not prime, mortgages to combat the housing bust (Elul, Gupta, & Musto, 2020).

§C Effect of free credit reports on banks

So far we focused on effect of free credit reports on borrowers. It is only natural to ask what was the effect on the lenders. The empirical design of the paper, however, faces important limitations in this regard. 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 effect on banks may be limited (Avery, Brevoort, & Canner, 2007). Second, most commercial banks often operate across states, hence their treatment and control status in this natural experiment is not discreet, but continuous. The treatment intensity is proportional to bank's mortgage activity in the treated states vis-à-vis control states.

Nonetheless, to examine this question, I classify a bank as treated if in the pre-event year 2004 the ratio of its mortgage originated in control states to that in control *and* treated 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.[18]

Coefficients 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) relative to the control banks. Subject to the caveats described above, these results imply that the outcomes from free credit reports were positive not only for the borrowers, but also for the lenders.

§D An alternative mechanism based on information asymmetry

An alternative mechanism based on asymmetric information may seem 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 lender possess information proportional to their true type. Hence, under the non-trivial search/application cost, bad borrowers self-select out. The borrower pool now improves compared to the situation in which borrowers do not know that lender has information about their true type, and expect this information to be better than warranted by their credit reports. Hence, the pool improves only 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 have imperfect information of their true type, thus both selecting-in by good borrowers selecting-out by bad borrowers are plausible.

The empirical findings are consistent primarily with the self-learning mechanism, rather than to the alternative. We saw that in the treated areas mortgage applications increased, not decreased, and the first-time homebuyers increased. Both of these suggest selecting-in by borrowers, plausible only under the *self-learning* mechanism, but not under the alternative.

One may also concern that in assessing mortgage applications, together with the credit reports, banks use private information such as soft information accumulated through relationship lending. In that case, free credit reports to consumers would have muted effects. Two features of the mortgage market mitigate this concern. First, mortgage companies, the dominant mortgage players, are not depository institutions. So unlike banks, they do not have avenues to accumulate soft information from credit cards or deposit accounts. Second, while different

[18] 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.

lenders may evaluate borrowers on additional criteria, credit scores is the key piece of information they necessarily peruse.[19] Admittedly, such soft information that goes into making mortgage decision but is not covered in credit reports make my estimates noisy.

§E 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 a valid concern that the results in this paper might be spuriously caused by the unique lending environment that existed in 2007–2008. To examine this, I re-estimate all the regressions by excluding the observations for years 2007 and 2008. These results, unreported for brevity, are qualitatively and quantitatively similar to those estimated earlier for almost all specifications, despite having only two post-experiment observations.

6 Conclusion

A non-trivial fraction of consumers do not check their credit reports regularly and do not know their credit scores. Various data suggest that borrowers makes erroneous credit decisions consistent with them having imperfect information of their creditworthiness. Credit reports can induce better credit-related decision-making among consumers as these can aid them self-assess their creditworthiness more correctly.

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 US consumers to access three free credit reports annually from 2005, while seven states already had local laws permitting their residents to obtain credit reports for free. This paper estimates the causal effects utilizing this reduction in cost of credit 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 constitute the treatment group.

The key finding is that reduced consumers' economic cost of accessing credit reports results in improved mortgage market outcomes, and benefits both consumers and lenders. Specifically, free credit reports resulted in an increase in mortgage demand and approval ratio, higher credit

[19] 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."

origination to creditworthy borrowers, a reduction in *ex-post* defaults and subprime population fraction, more entry of first-time borrowers, and better financial performance of lenders.

Though these findings concern mortgage-related decisions, their implications apply broadly to any general credit-related consumer decision-making when they have imperfect information of their creditworthiness. Furthermore, the causal nature of these findings imply 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).

Campbell, J. Y. (2006). Household Finance. *The Journal of Finance*, *61*(4), 1553–1604.

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

FRBNY. (2013–2020). *Survey of Consumer Expectations* [Database]. Retrieved from https://www.newyorkfed.org/microeconomics/sce

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

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.

Kronlund, M., Pool, V. K., Sialm, C., & Stefanescu, I. (2019). Out of Sight No More? The Effect of Fee Disclosures on 401 (k) Investment Allocations [Working Paper]. *Unpublished*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3288808

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/

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

Survey of Consumer Finances. (1998–2007). *Survey of Consumer Finances (SCF)* [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

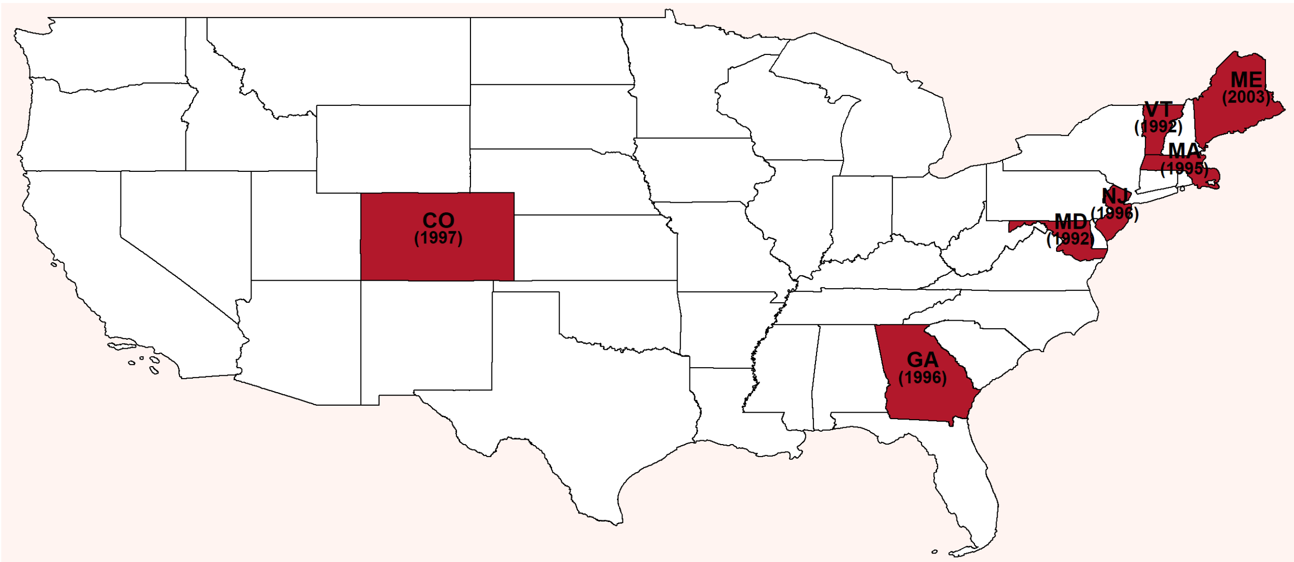
Your credit report includes your Personal Information and, if applicable, Consumer Statements, and could include other items that may affect your credit score and rating.

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

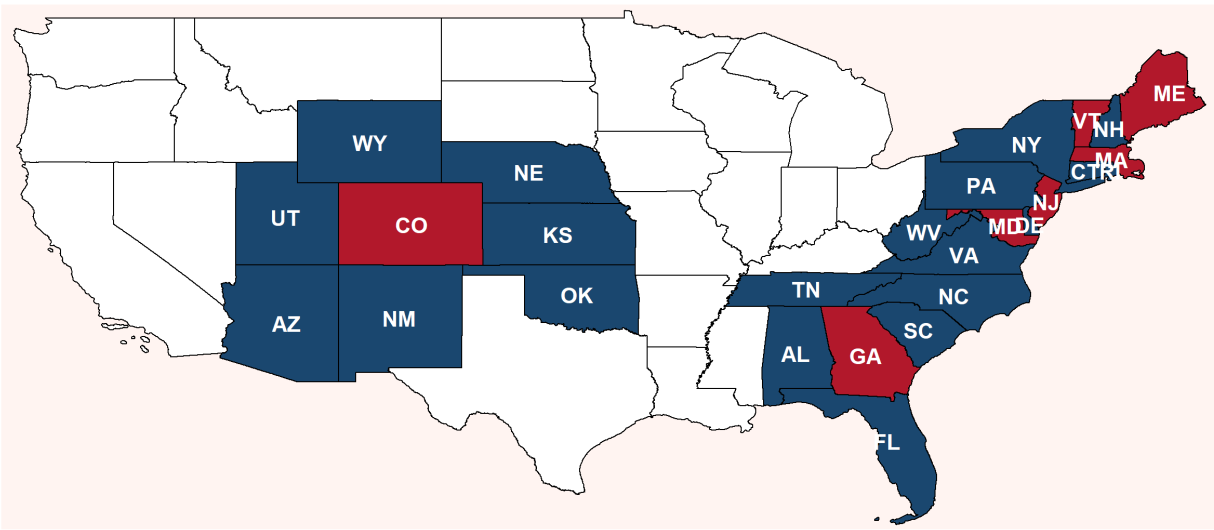
This figure illustrates the empirical setting of the paper. **Panel A** shows the pre-FACTA states — the states that had local laws allowing free credit reports before the enactment of FACTA in 2004 — and the respective years in which those states enacted free credit report laws. **Panel B** illustrates 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

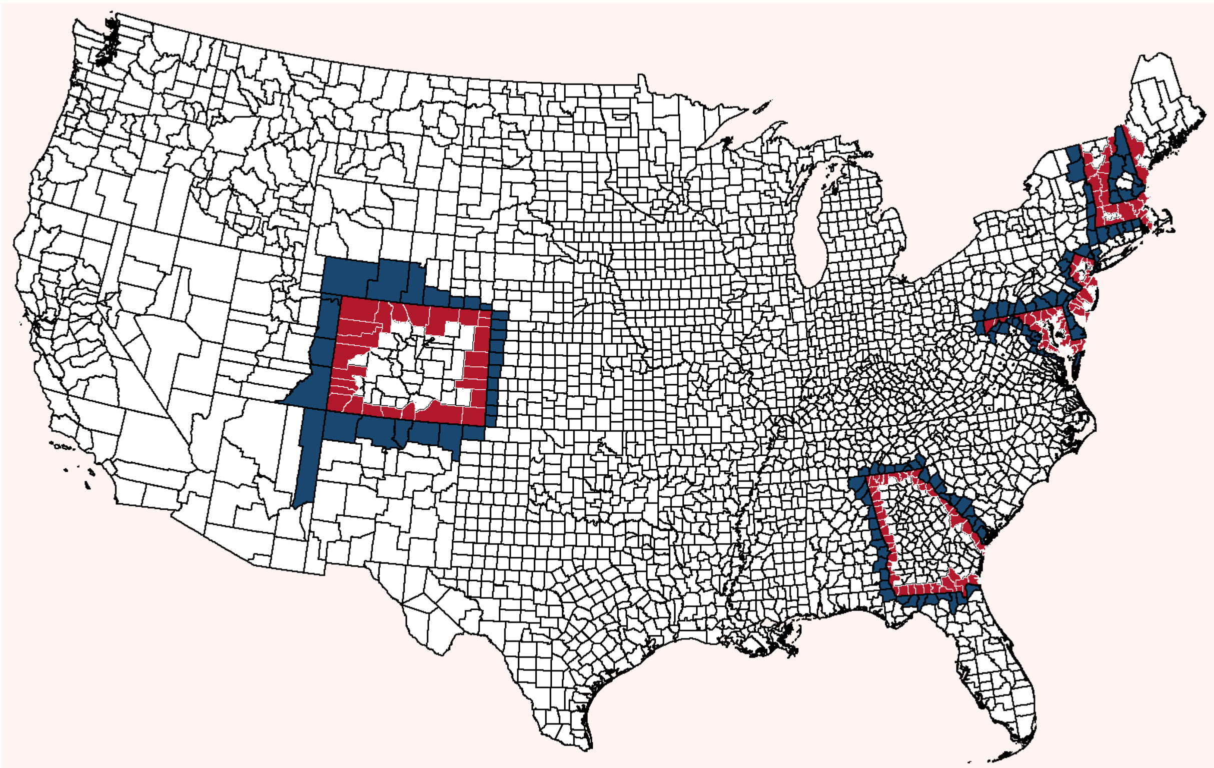
Panel B: Empirical Setting: Control and Treatment States



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

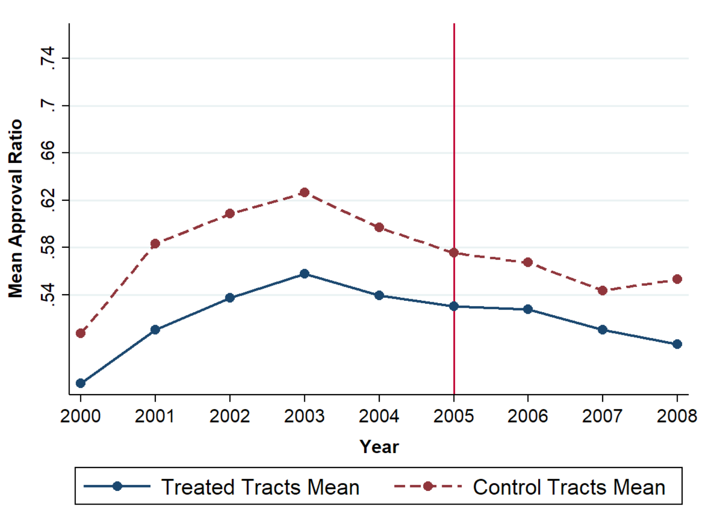
Panel A of this figure shows the mean approval ratio in the treated and control census tracts. **Panel B** of this figure shows the coefficients β_k from regressing *Approval Ratio* using the specification:

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

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

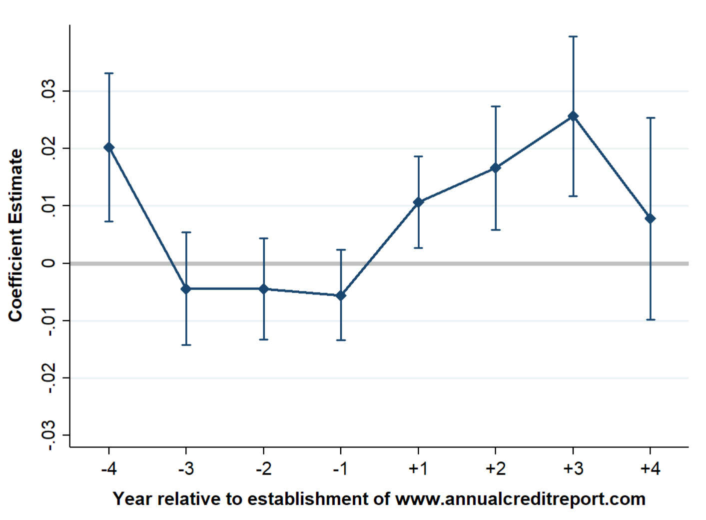
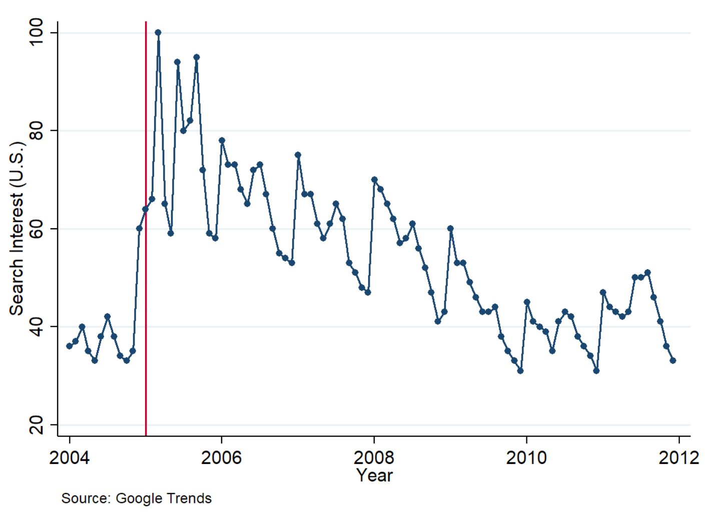


Figure V: Consumer Interest in Free Credit Reports: Google Trends

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*

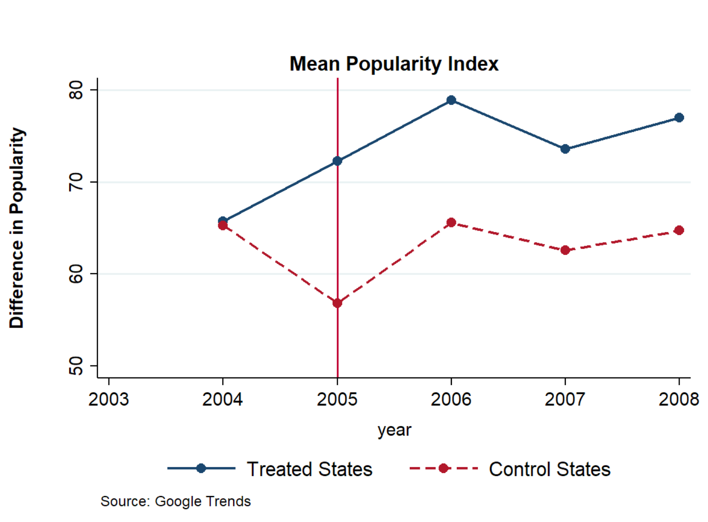
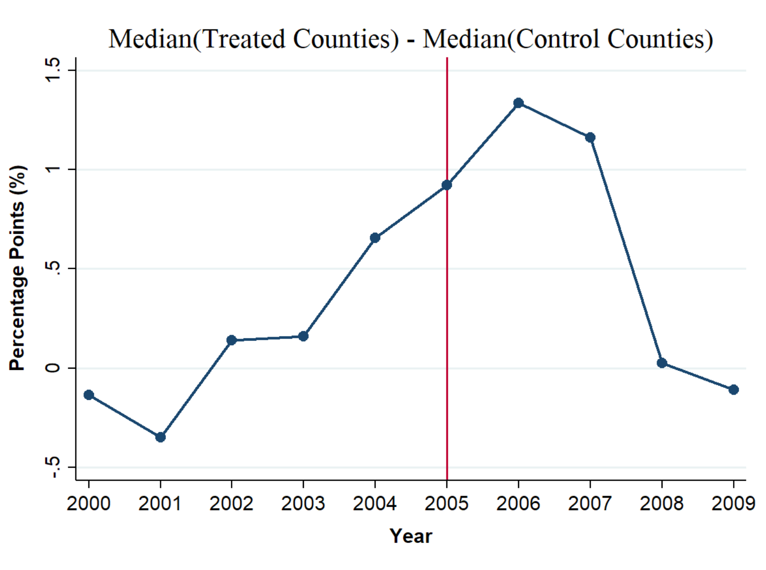


Figure VI: Subprime Population over Time

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)

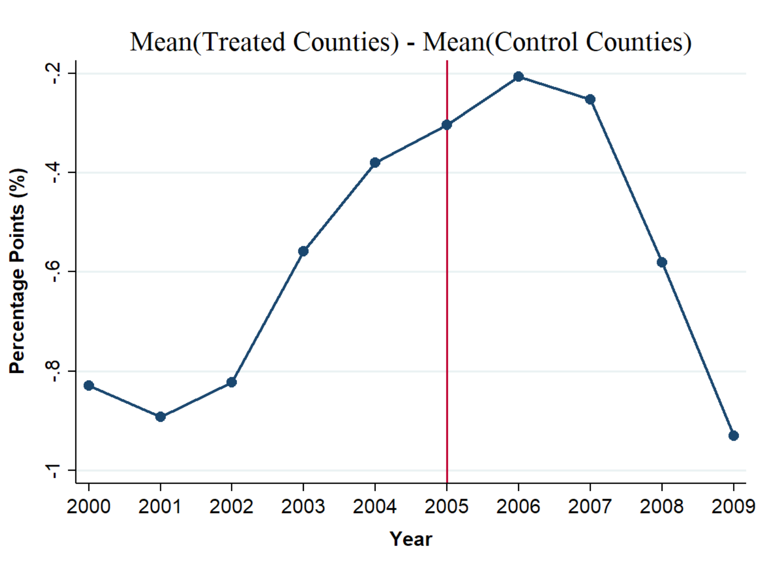


Figure VII: Effect of Free Credit Report on Mortgage Defaults

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.

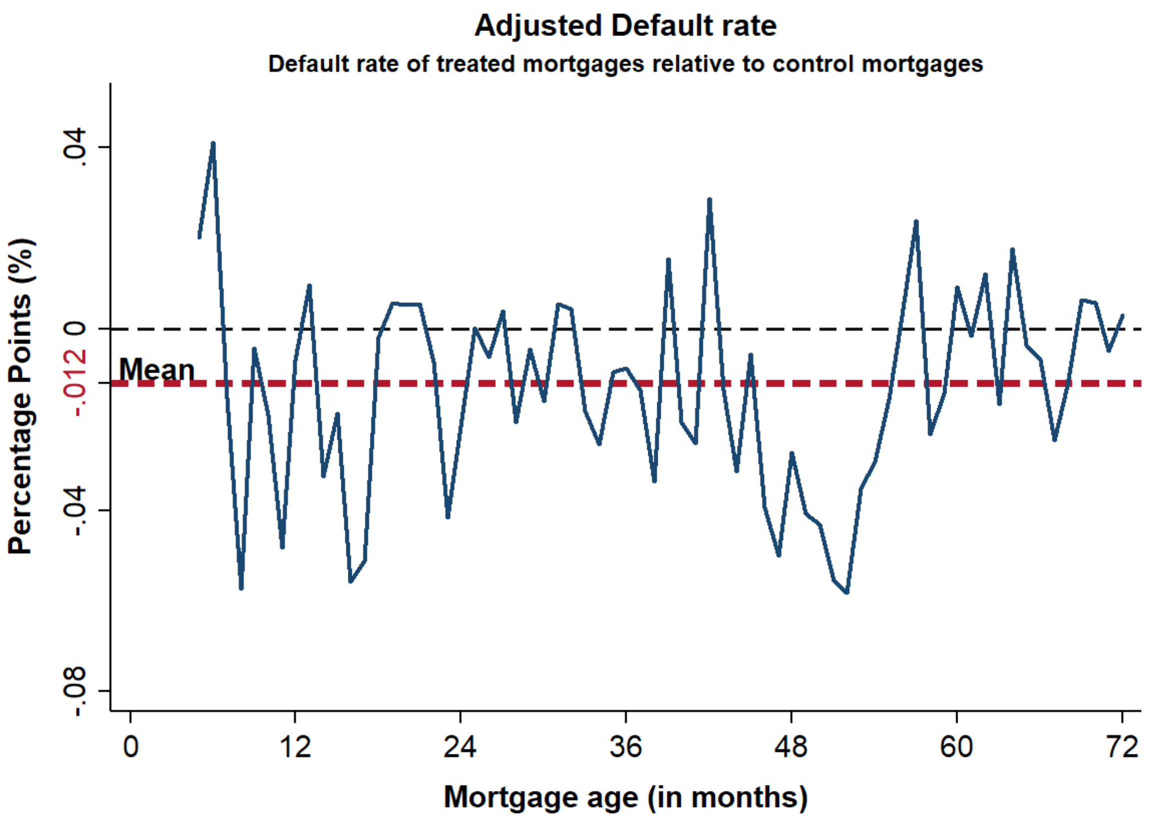


Table I: Summary Statistics

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. *Num of App per 1000 adults* is the number of mortgage applications scaled by the population aged 18 to 64 years in a census tract. *Approval ratio* is the ratio of mortgages originated, or mortgages approved but not accepted, to 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 three variables at the bottom capture the economic conditions (*Economic Controls*). 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.
Num of App per 1000 adult	86817	83.39	74.73	66.39	36501	98.39	77.57	77.73	50316	72.51	70.64	56.46
Approval Ratio	82712	0.54	0.13	0.55	35885	0.57	0.12	0.58	46827	0.52	0.14	0.53
Deny Credit Hist Ratio	82712	0.06	0.04	0.05	35885	0.05	0.04	0.04	46827	0.06	0.05	0.05
Deny Debt-to-inc Ratio	82712	0.03	0.03	0.03	35885	0.03	0.02	0.03	46827	0.03	0.03	0.03
Withdrawn Ratio	82712	0.12	0.05	0.12	35885	0.12	0.04	0.11	46827	0.12	0.06	0.12
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	73	0.05	0.03	0.04	29	0.04	0.02	0.04	44	0.05	0.04	0.04

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
Num of App per 1000 adult	48363	110.48	83.53	93.48	20294	129.68	86.14	108.62	28069	96.59	78.74	83.20	0.000
Approval Ratio	47024	0.55	0.14	0.56	20077	0.58	0.13	0.60	26947	0.52	0.14	0.53	0.000
Deny Credit Hist Ratio	47024	0.06	0.04	0.05	20077	0.06	0.04	0.05	26947	0.07	0.05	0.06	0.000
Deny Debt-to-inc Ratio	47024	0.03	0.02	0.03	20077	0.03	0.02	0.03	26947	0.03	0.02	0.03	0.000
Withdrawn Ratio	47024	0.12	0.05	0.11	20077	0.12	0.04	0.11	26947	0.13	0.05	0.12	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	41	0.05	0.02	0.05	18	0.05	0.02	0.05	23	0.05	0.02	0.06	0.795

Table II: Survey Evidence on usage of Credit Reports/Scores 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*. Since *Q. N19* is a conditional question in the survey, and this paper concerns with only mortgage-related decisions, specifications (4–6) include only those responses in which respondent selected *very unlikely* or *somewhat unlikely* to apply for mortgage/home-based loan, or refinance in *Q. N17A*, or mentioned probability <10% in *Q. N17B* to apply for mortgage, or to refinance. 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/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 ΔHPI 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 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)
	N	N	Aprv.	Aprv.	ΔHPI	ΔHPI
Treat × Post	13.28***	15.39***	0.01**	0.02***	1.74*	1.82*
	(2.94)	(3.63)	(2.42)	(2.82)	(1.83)	(1.82)
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.)	0.807	0.808	0.748	0.744	0.683	0.686
Observations	86806	84789	82665	80667	25390	25365

Table IV: Owner-occupied Mortgages

Panel A of this table reports the estimates of the treatment effect of free credit reports on the number of mortgage applications and the approval ratio only for the sub-sample of owner-occupied mortgage applications. Panel B reports the result of regressing the ratio of not owner-occupied mortgages to total applications (to successful applications) in column 1 and 2 (in column 3 and 4). The regression specification is:

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

#Applications and *Approval Ratio* are the number of applications per 1000 adults and the approval ratio in a census tract, respectively. *Non-ocp.* represents the fraction of total (successful) applications which are not owner occupied in column 1 and 2 (3 and 4). *Economic Controls* are 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: Applications and Approval Ratio for Owner-occupied Mortgages

	(1)	(2)	(3)	(4)
	N	N	Aprv.	Aprv.
Treat × Post	12.78***	14.95***	0.01**	0.01***
	(2.89)	(3.61)	(2.26)	(2.60)
Economic Controls	No	Yes	No	Yes
Census Tract FE	Yes	Yes	Yes	Yes
Border × Year FE	Yes	Yes	Yes	Yes
Cluster (County)	Yes	Yes	Yes	Yes
R^2 (Adj.)	0.808	0.810	0.661	0.654
Observations	86806	84789	86619	84602

Panel B: Non-owner-occupied Mortgages Fraction

	Fraction of total app.		Fraction of successful app.	
	(1)	(2)	(3)	(4)
	Non-ocp.	Non-ocp.	Non-ocp.	Non-ocp.
Treat × Post	0.01**	0.01**	0.01**	0.01**
	(1.99)	(2.03)	(2.02)	(2.13)
Economic Controls	No	Yes	No	Yes
Census Tract FE	Yes	Yes	Yes	Yes
Border × Year FE	Yes	Yes	Yes	Yes
Cluster (County)	Yes	Yes	Yes	Yes
R^2 (Adj.)	0.088	0.087	0.087	0.086
Observations	82665	80667	82579	80581

Table V: Contraction in Credit History-related Rejections

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:

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

	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	-0.001	-0.002	-0.002
	(-1.49)	(-1.51)	(-2.04)	(-1.89)	(-1.08)	(-0.96)	(-1.49)	(-1.20)
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.)	0.542	0.538	0.575	0.575	0.267	0.266	0.319	0.322
Observations	82665	80667	39069	38692	82665	80667	39069	38692

Table VI: Increase in Consumer Search Accuracy

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). *% Application Withdrawn* is the ratio of applications expressly withdrawn by consumers to the number of applications in a census tract. *Economic Controls* are 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)
	% Application Withdrawn	% Application Withdrawn
Treat × Post	-0.009***	-0.011***
	(-2.82)	(-3.51)
Economic Controls	No	Yes
Census Tract FE	Yes	Yes
Border × Year FE	Yes	Yes
Cluster (County)	Yes	Yes
R^2 (Adj.)	0.340	0.341
Observations	82665	80667

Table VII: 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:

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

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 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 $\times$ 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 $\times$ 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 VIII: Distinguishing Supply and Demand: Equilibrium Interest Rate

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:

$$Y_{izjt} = \beta_0 + \beta_1 \text{Treatment}_{izj} \times \text{Post}_t + \delta \times \text{Loan controls}_i + \alpha_z + \gamma_{jt} + \varepsilon_{izjt}. \quad \text{See Eq. (4).}$$

Loan Controls are *credit score, debt-to-income ratio, combined loan-to-value (CLTV), number of units, 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 (in percentage points)	
	(1)	(2)
	%	%
Treat × Post	0.009***	0.011***
	(13.68)	(12.60)
Loan Controls	No	Yes
Zip3-State FE	Yes	Yes
Border × Qtr FE	Yes	Yes
Cluster Zip3-State	Yes	Yes
R^2 (Adj.)	0.728	0.758
Observations	7739882	3548884

Table IX: Distinguishing Supply and Demand: Heterogeneous Effects by Density of Lenders

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 14). *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 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	0.002***	0.002*	0.016**	0.011*	0.017***	0.012**
	(2.20)	(1.16)	(2.87)	(1.83)	(2.52)	(1.94)	(2.87)	(2.28)
Difference [High - Low]		-0.001		-0.001		-0.005		-0.005
p-value		(0.595)		(0.600)		(0.610)		(0.592)
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	0.633	0.563	0.758	0.728	0.754	0.723
Observations	60704	25808	59207	25293	57628	24938	56144	24429

Table X: Heterogeneous Effects by Consumer Creditworthiness

This table reports estimates of the treatment effect of free credit reports on the number of mortgage applications and the approval ratio in *ex ante* low and high creditworthiness areas. To separate areas by creditworthiness, Panel A uses county-level subprime population fraction, while Panel B uses census tracts-level number of payday lenders. In Panel A, a county is categorized as 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 sample size in panel B only includes census tracts from CO and surrounding states. 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. *Economic Controls* are 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: County-level Creditworthiness Measure

	High Creditworthiness (Prime Counties)				Low Creditworthiness (Subprime Counties)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	N	N	Aprv.	Aprv.	N	N	Aprv.	Aprv.
Treat × Post	16.74**	18.09**	0.02**	0.02**	8.49	10.63**	0.01	0.01*
	(2.30)	(2.58)	(2.31)	(2.55)	(1.65)	(2.24)	(1.58)	(1.90)
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	0.803	0.774	0.771	0.826	0.828	0.697	0.697
Observations	39254	37692	38175	36625	47258	46808	44391	43948

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***	72.78***	0.06***	0.06***	43.72***	43.27***	0.02	0.02
	(5.31)	(5.36)	(3.52)	(3.57)	(3.82)	(4.00)	(0.94)	(0.98)
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	0.731	0.816	0.818	0.794	0.795
Observations	1452	1452	1395	1395	872	872	865	865

Table XI: Heterogeneous Effects 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 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	0.05	0.06	1.82**	2.10***	2.62**	3.06***	4.33**	5.29***
	(0.04)	(0.04)	(2.32)	(2.84)	(2.45)	(3.24)	(2.00)	(2.80)
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.)	0.763	0.765	0.776	0.778	0.745	0.747	0.659	0.660
Observations	88282	86255	88282	86255	88282	86255	88282	86255

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
	(2.06)	(2.59)	(0.93)	(1.03)	(0.87)	(0.89)	(1.32)	(1.27)
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	0.346	0.391	0.388	0.404	0.399	0.363	0.357
Observations	71879	70132	72428	70661	72548	70771	71995	70219

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); and approved and not sold by lending institutions (columns 5 and 6). The dependent variables are calculated at the census tract level. *Economic Controls* are 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.

	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**	0.050***	0.000	0.004
	(-0.35)	(-0.10)	(2.54)	(2.76)	(0.05)	(0.94)
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	0.002	0.055	0.035
Observations	82665	80667	82665	80667	82665	80667

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:

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

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 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 zip3-state areas relative to the control zip3-state 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.

	(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 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