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

Amit Kumar†

Abstract

Bad credit history is the dominant reason for mortgage rejection. Exploiting the 2004 nationwide extension of local laws existing in seven U.S. states that allowed free credit reports, I document that cost reduction of reports for consumers resulted in increased mortgage demand, approval ratio and first-time homebuyers; and decreased defaults, persisting through the 2008 financial crisis. Consistent with a demand-pull (rather than a supply-push) effect, mortgages became costlier. Moreover, higher approvals, reflecting an improved borrower pool, occurred mainly among prime, more educated, and bottom income quartile consumers. Results emphasize that cheaper credit history information to consumers improves credit market outcomes.

JEL Classification: G21, G28, D83

Keywords: Credit Reports, Information Provision, Household Finance, Retail Mortgage Market.

*I am profoundly grateful to Utpal Bhattacharya, Emilio Bisetti and John 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 the doctoral sessions. I also want to thank the seminar participants at HKUST brownbag seminar, Southern Denmark University 5th Finance Workshop, PhD Mentoring Day at 28th Finance Forum, and the PhD Symposium at McCombs Business School, UT Austin for helpful suggestions.

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

1 Introduction

Housing is a dominant asset class for middle-class homeowners (Campbell, 2006), and consumers primarily use mortgage debt to finance housing. In the US, consumers made about 150 million mortgage applications between 2000 and 2008, and interestingly, one in five applications was rejected. A deeper look at these rejections reveals a surprising pattern. The most frequent cause for mortgage rejection is consumers' own credit history. 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). Barring errors in the application assessment, why do numerous consumers apply for mortgages only to be subsequently rejected for their own credit history despite incurring real costs (application fees, and increased interest rate and rejection probability on future applications)? The data seem to suggest that consumers have imperfect information of their creditworthiness. Under this informational imperfection, the above phenomenon can rationally arise.[1]

Another widely observed phenomenon further sheds light on the issue of lack of information about creditworthiness among the US households. About 15% of the US households *choose* not to apply for credit because they anticipate rejection (Survey of Consumer Finances, 1998—2007). Suppressing the credit demand would be rational if households were perfectly aware of their creditworthiness. However, in the presence of informational issues, many households may mistakenly underestimate their creditworthiness, or overestimate the rejection probability, and incorrectly suppress their demand. Overall, these facts suggest that consumers have imperfect information about their creditworthiness.

The purpose of this paper is to show that provision of information to consumers about their credit history at lower cost can affect credit market outcomes. Specifically, I examine the causal link between consumers' economic cost of accessing their credit reports — an authoritative information source of their creditworthiness — and the mortgage demand and approval ratio.

Credit reports record crucial information about the financial situation of consumers e.g., their credit history and available borrowing capacity, and are a credible creditworthiness signal (Figure 1). Yet, tendency among consumers to use credit reports is shockingly low. Out of approximately 1 billion credit reports generated annually in the US, a mere 1.6% end up in

[1] In a credit market with perfectly informed consumers (and lenders), (1) credit-unworthy consumers *would not apply* because credit application and rejection is costly; (2) no creditworthy consumer needing credit would choose *not to apply* for credit.

the hands of consumers (Avery, Calem, & Canner, 2004).[2] The high economic cost of accessing credit reports — monetary cost, search cost, and cost imposed by unawareness that such reports exist — might be the reason for such low levels of credit reports usage among consumers.

I propose a *consumer self-learning* mechanism that links consumers' credit report cost with the aggregate market outcomes (credit demand and approval ratio). Consumers can use credit reports to self-assess their creditworthiness more accurately before making a credit application.[3] If the information learned from a consumer's report matches or exceeds his or her *ex-ante* expectation, consumers can continue with the application. Otherwise, he or she can delay the application to take steps to improve the credit records, or correct any inaccurate information therein (Avery et al., 2004), search for a more suitable lender (say, a sub-prime lender), or altogether abandon the application. As a result of this self-learning-induced sorting among consumers, credit-worthy consumers stay in / enter the market, while marginal creditworthy consumers leave the market or search for a more suitable lender (to save the rejection costs). This improves the average quality of the applicant pool in the market, leading to a higher approval ratio. Such sorting also shapes the total credit demand — it may result in increased or decreased total demand for credit 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 that credit reports exist and matter for their credit application. Hence, under this demand-driven mechanism, the lower economic cost of credit reports leads to their wider usage and increased self-learning among consumer, and ultimately influences the credit market outcomes.

The key challenge in examining the link between the cost of credit reports and credit market outcomes is to establish that the effect is causal. An exogenous change in consumer's credit report cost can address this challenge. A natural experiment in the US provides such an opportunity — the nationwide extension, under the enactment of the federal *Fair and Accurate Transaction Act of 2003* (FACTA), of local laws already existing in seven states offering their residents *free credit reports*. Figure 2 depicts the seven early-adoption states (the pre-FACTA states) — Colorado, Georgia, Maine, Maryland, Massachusetts, New Jersey, and Vermont — and the

[2] Perhaps this is why, experts including the Federal Reserve Board actively encourage consumers to check their credit reports (Federal Reserve Bank of Philadelphia, 2015).

[3] Inaccurate self-assessment of creditworthiness leads to worse financial outcomes (Courchane, Gailey, & Zorn, 2008). Lack of financial knowledge and complexity of debt products contribute to such inaccuracies. Brown, Haughwout, Lee, and Van Der Klaauw (2011) estimates that consumers underestimate their student (credit card) debt by as much as 25% (37%). Using a consumer credit survey collected in 2000, Perry (2008) finds that 32% of consumers overestimate their credit ratings, while only 4% underestimate it.

year in which they adopted the free credit report laws. Owing to the provisions of FACTA, consumers throughout the US could obtain three credit reports annually for free from the website www.annualcreditreport.com, which became operational in January 2005. However, credit reports were already available annually for free in the pre-FACTA states. Therefore, consumers across all states, other than the pre-FACTA states, saw a reduction in the cost of credit reports since 2005, the treatment year. My empirical strategy centers on this change. I use a difference-in-differences (DID) design in which the pre-FACTA states constitute the control group, while the states surrounding them constitute the treatment group (Figure 3). Further, I restrict the focus only to the counties lying along the border of the control and treatment states, to control for prevailing local socioeconomic conditions, and also to cleanly identify the treatment effects. Figure 4 shows the treatment and control counties on the map of the US.

This empirical strategy accords us certain advantages over typical state-law-based DID settings. Here, the assignment of treatment to states is not due to the state's adoption of the law, which arguably is endogenous. Instead, the assignment is by the federal enactment of FACTA resulting in nationwide extension of the *existing* state laws to *all* states, whereas the pre-FACTA states had already enacted their local laws in the past. Notwithstanding the past adoption of local laws, the endogeneity concern remains to the extent that the enactment of the federal law was in response to some prevailing unobserved characteristics (around 2003). The timing and circumstances of FACTA enactment does mitigate this concern to a great extent. As discussed in detail later, FACTA was mostly a consolidation of already existing provisions of another federal law, *Fair Credit Reporting Act of 1970* (*FCRA*). The main difference is addition of the provision of free annual disclosure of credit reports. Further, the FCRA was set to expire in 2003 due to the sunset clause embedded in its 1996 amendment. Thus, the timing of the enactment of FACTA in 2003 was independent of any contemporary socio-economic characteristics, and was necessitated by the sunset clause of the FCRA set in the past. Moreover, the concern that other provisions of FACTA might affect the outcomes is mitigated as well, because most of its provisions were already in place under the FCRA, while the free disclosure of credit reports provision was newly added.

A critical assumption underpinning the identification strategy in this setting is that, prior to the nationwide extension of the free credit report laws, the usage of credit reports among consumers was higher in the pre-FACTA states than in rest of the US. The credit report usage data (take-up rates) from the US senate hearing confirm this. The usage of credit reports, relative to the national average, 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). 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).

This paper contributes three key findings. First, free credit reports led to increased mortgage demand and approval ratio, mainly for occupancy purposes, and a weakly significant increase in house price growth rate. Second, good quality borrowers seem to have been instrumental in driving up the demand: origination was higher in more creditworthy areas, and among prime consumers; and, the mortgages from treated areas were less likely to be defaulted upon. Third, more first-time homebuyers took mortgages in the treated areas after free credit reports became available, suggesting that new homebuyers are more uncertain over their creditworthiness, and benefit from free credit reports.

I now discuss the empirical results. I begin the empirical analysis with quantifying the changes in mortgage demand and approval ratio. Comparing the census tracts in the control and treated counties lying at the borders of the pre-FACTA (control) states and the surrounding (treatment) states, I find that the extension of free credit report laws resulted in an increase of 13.8% to 16.0% in mortgage applications, and an increase of 1 percentage point in approval ratio. In volume terms, these represent an increase of ~$38.1 billion in mortgage demand across the treated areas, and an increase of ~2.6 mortgage approvals per census tract, amounting to a ~$2.8 billion increase in mortgage origination in treated areas. These estimates are robust to the inclusion of *Census Tract* fixed effects and "*Border* × *Year*" fixed effects, and a host of time-varying controls capturing the economic conditions at county and state level.

I next examine the nature of the increased demand, specifically whether the increased demand came from occupancy-seeking or investment-seeking borrowers? When I restrict the sample to the owner-occupied mortgages category, the estimates of changes in demand and approval ratio are similar to the previous estimates for all categories. Also, tests for changes in the non-owner-occupied mortgages fraction yield insignificant coefficients, suggesting that the composition of mortgage demand between occupancy vis-à-vis non-occupancy purposes did not change after the free credit reports laws. Thus, free reports do not appear to fuel an investment-motivated house-buying frenzy.

A natural question follows: did the house prices increase in the treated areas on the back of a robust ~13% increase in the mortgage demand? I find that there was a statistically weak increase in house prices in the treated areas: the growth rate of house prices increased by 1.7–1.8 percentage point. This is in line with a previous estimate by Di Maggio and Kermani (2017), which reports an increase in the growth rate of house prices by 3.3 percentage point following a 10% increase in mortgage origination.

It is crucial to understand whether the increase in mortgage demand was accompanied by an increase in default. To analyze this, I compute the *adjusted default rate* as the difference between the default rate on the mortgages originated in the event year and the year before in the treated areas minus the same for the control areas. Its time trend shows that the treated mortgages were less likely to be defaulted upon, in relation to the control areas, even during and after the years of the 2008 financial crisis. This result is noteworthy. Free credit reports not only resulted in more consumers applying for mortgages and more lenders approving them, but also resulted in lower probability of default, *not higher*, even in and after the financial crisis. Thus, the findings that mortgage demand and origination increased, and default rates decreased *ex-post* are consistent with improvement in the borrower pool.

Having established that the mortgage demand and approval ratios increased in the treated areas, I next examine three outcomes that speak to the proposed *consumer self-learning* channel. First, I examine the likelihood of mortgage rejection due to credit history. If consumers self-learn their creditworthiness from credit reports, then in aggregate, the likelihood of rejections due to credit history should decrease. I find that the fraction of applications denied due to credit history declines in the treated areas by 0.3 percentage point in *ex-ante* high rejection areas, but the fraction of applications denied for debt-to-income ratio did not change significantly. This is consistent with increased self-learning among consumers from credit reports.

Second, I examine the search accuracy of the consumers by exploiting a unique feature of the mortgage market — *in-process application withdrawals*. Owing to uncertainty over the success of their application, consumers tend to apply to multiple lenders at once. In doing so, they incur multiple non-refundable application costs (~US$ 400 per application), while ultimately obtaining the mortgage from only one lender (if approved) and withdrawing from the others. In a limiting case in which an applicant *ex-ante* is certain her mortgage would be approved by a given lender, she would apply to just one lender, and save the costs associated with multiple applications. Since the certainty over application acceptance is higher if applicants are better-learned about their creditworthiness, in aggregate, free credit report laws should lead to a reduction in in-process application withdrawals, through increased self-learning among consumers. I find that the in-process applications fraction decreased by 0.9 percentage point (~2.34 applications per census tract) in treated areas. This represents a saving of ~US $6.6 million in upfront application fees across the treated areas.

Finally, I assess the self-learning effect of free credit reports among the consumers for whom it is the most valuable — the first-time homebuyers. As previously mentioned, about 15% of US households do not apply for credit because they anticipate rejection. Similarly, there may be non-trivial fraction of consumers who do not understand the important role of credit reports in mortgage application process. Among such consumers, if the fraction who learn from credit reports that they have higher (or the same) creditworthiness relative to their expectation is larger than those who learn otherwise, the proportion of first-time homebuyers in the market will increase. Using a subset of mortgages purchased by Fannie Mae and Freddie Mac (Government Sponsored Entities or GSEs), I find that the first-time homebuyers fraction increased in the treated areas by 1 percentage point.

Lower cost of credit reports not only impact the borrowers, but also the lenders as they may respond to the knowledge that free credit reports have become available (to their borrower clients). Thus, approval ratios may increase in two ways: first, lenders may start extending more mortgages by adjusting their lending standards (*supply-push effect*); second, borrowers may start self-learning more, leading to an improved applicant pool (*demand-pull effect*). I argue that the latter explanation is the primary driver of the results in this paper. We saw earlier a preliminary evidence favoring the demand-side mechanism: the increase in number of mortgage applications in the treated areas. Further supporting evidence comes from increase in consumer interest in free credit reports measured from Google searches for the phrase *'Free Credit Reports'*. I see that in 2004 — the year before the treatment i.e. the establishment of the website www.annualcreditreport.com — this search phrase was equally popular in both the treated and control states. However, from 2005 onwards, this phrase became increasingly popular in the treated states relative to the control states. In addition to these preliminary evidences, I employ two strategies to ascertain that demand-side is the primary driver of the results.

First, I appeal to the supply-demand framework. Assuming the market is in equilibrium, the quantity as well as the price of a good increases only when the demand curve shifts outward.

Using the mean contracted interest rates on the mortgages purchased by the GSEs, I find that the interest rates increased by 1.1–1.2 basis points in the treated areas in comparison with the control areas. This increase occurred notwithstanding the highly elastic mortgage supply in the US owing to the GSEs' mandates. This finding supports the demand-pull explanation.

In the second strategy, I utilize the heterogeneity in the effects of free credit reports to argue that demand-side factors are the primary mechanism at work in the current setting. The ideas is that if the findings were driven by credit demand- (supply-) related factors, they would be stronger where the demand (supply) related factors were stronger. I find that the two key outcome variables — mortgage demand and approval ratio — respond to different degrees across areas in a manner consistent (inconsistent) with the demand- (supply-) driven explanation. Firstly, results are similar across areas with high and low lender density, which is a supply-related factor. Then, I find that the results are stronger in areas where consumers have higher education and creditworthiness, which are demand-related factors. Finally, an increase in approval ratio is significant among the applicants in the lowest income quartile. Agarwal, Chomsisengphet, Mahoney, and Stroebel (2018) shows that marginal propensity to lend to low income consumers is low, while marginal propensity to borrow is high among them. Therefore, the increase in approval ratio for low income consumers does suggest a pool improvement among the low income quartile borrowers. Overall, these four heterogeneity-related findings, together with the finding that equilibrium interest rates rise in treated areas, underscores the role of the demand-pull mechanism in driving the results.

I present three supplementary findings in the final set of results. Firstly, I examine weather increased private securitization of mortgages led to the increased mortgage supply in the treated areas (Keys, Mukherjee, Seru, & Vig, 2010). I find no change in the fraction of mortgages that were approved and sold to non-government entities, ruling out this alternative explanation.

Secondly, I test the subprime-supply hypothesis that the increase in the mortgage supply was due to lenders increasing supply in subprime areas (Mian & Sufi, 2009). We already saw that the effect of free credit reports is stronger in creditworthy areas. I now examine whether the increased origination went to prime consumers using the application level credit scores in the GSE data. I indeed find that the increase in number of mortgages to prime borrowers (credit score ≥ 620) in the treated areas was a staggering 30 times higher than to subprime borrowers in the same area.

Finally, I evaluate the effect of free credit reports on commercial banks, notwithstanding the caveats that these banks are not the dominant mortgage lenders, and that the assignment of treatment to banks is noisy due to their multi-state operation. I find that the financial performance of the treated banks improved. Their net interest margin increased by 6 basis points, return on equity rose by 0.74 percentage points, and return on assets improved by 0.07 percentage points relative to the control banks.

Besides the consumer self-learning mechanism proposed in this paper, one may suggest an alternative mechanism based on asymmetric information. In this mechanism, borrowers *privately know* their true creditworthiness type, but do not know what the bank knows about their true type. Using free credit reports, borrowers learn that the bank has information about them that is proportional to their true type. Given that the search/application cost is non-trivial, the bad borrowers would self-select out. Hence, compared to a situation in which borrowers do not know that the bank has information about their true type, and expect that the information about them at the bank is better than would be warranted from learning their true type from their credit reports, the borrower pool would improve due to the self-select-out of the bad borrowers. Thus, under this alternative mechanism based on the asymmetric information, free credit reports would lead to pool improvement 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, recall that under the self-learning mechanism proposed in this paper, borrowers have imperfect information of their true type. So it predicts selecting-in by good borrowers (who underestimate their creditworthiness in absence of true information) as well as selecting-out by bad borrowers (who overestimate their creditworthiness).

The empirical findings of the paper allow me to attribute the results primarily to the proposed self-learning mechanism, rather than to the alternative asymmetric information mechanism described above. First, we see that mortgage applications increased, not decreased, in the treated areas. An increase is plausible under the proposed *self-learning* mechanism through self-select-in effect, but is not plausible under the asymmetric information mechanism. The result suggests that there must be more (good) borrowers selecting-in the market than the bad borrowers selecting-out. Further, we saw that the first-time homebuyers increased in the treated areas, which is again suggestive of a selecting-in effect plausible under the *self-learning* mechanism.

An additional concern is that banks may use private information in addition to the credit scores, such as soft information accumulated through relationship lending, to assess mortgage applications of borrowers. Two features of the mortgage market allay this concern. First, the dominant mortgage lenders are not banks, but mortgage companies, which do not have avenues to accumulate soft information from activities such as credit card lending or deposit accounts. Avery, Brevoort, and Canner (2007) shows that banks accounted for just 37% of the mortgage lending activity, while mortgage companies accounted for the rest, in 2005. Second, while different lenders may evaluate borrowers on additional criteria such as income, credit scores are one key piece of information they necessarily look at.[4] Admittedly, the orthogonality of banks' private information (that goes into the mortgage decision) with that contained in the credit reports makes my estimates noisy. One could estimate the exclusive effects of learning among consumers from credit reports on approval ratio by purging the role of soft information usage by lenders in the approval process, if one could identify which mortgage decisions are exclusively based on the credit reports. However, data limitation does not allow me to examine this.

Overall, the results in this paper highlight that imprecise information among consumers about their creditworthiness brings inefficiency in the market, and provision of information to consumers at reduced economic cost improves aggregate market outcomes, benefiting both lenders and consumers. Even though the setting in this paper builds upon the data from the mortgage market, the findings remain relevant to any credit-related decision-making by consumers under imperfect knowledge of their creditworthiness. Banking on the causal interpretation of these results, we can reasonably expect that any policy or intervention aimed at educating consumers of their creditworthiness would yield similar results.

Related Literature

This paper contributes to multiple strands of the literature. The most relevant is the nascent literature on information provision and its effect on market participants. This is the first paper to document that the borrower pool in the market improves from usage of credit reports among consumers. In a field experiment, Homonoff, O'Brien, and Sussman (2019) finds that borrowers who are randomly provided information about their own FICO® scores are less likely

[4] Expeiran, one of the three credit reporting agencies, 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 (e.g. CCJs and IVAs)." https://www.experian.co.uk/consumer/mortgages/guides/credit-and-mortgages.html
Transunion too explains the importance of credit score on mortgage applications here: https://www.transunion.com/mortgage

to default. Kulkarni, Truffa, and Iberti (2018) shows that standardized financial contracts reduce consumer delinquency by 40%, and sophisticated (unsophisticated) borrowers are helped the most by the increased product disclosure (product standardization). Liberman, Neilson, Opazo, and Zimmerman (2018) documents aggregate welfare loss, and a reduction (increase) in the cost of credit for poorer defaulters (non-defaulters) when consumer reporting agencies (CRAs) are barred from reporting consumer defaults to lenders. Dobbie, Goldsmith-Pinkham, Mahoney, and Song (2016) shows that removal of bankruptcy flags from credit reports results in a significant increase in credit card balance and mortgage borrowing. In the investment market, consumers become more sensitive to expense ratios and short-term performance after regulation made the fee and performance disclosure of 401(k) plans mandatory (Kronlund, Pool, Sialm, & Stefanescu, 2019).

This paper also contributes to the literature on financial literacy and household behavior. Low financial literacy has been shown to have detrimental economic outcomes: high mortgage delinquency and home foreclosure (Gerardi, Goette, & Meier, 2010); poor mortgage choice (Moore, 2003) and large debt accumulation (Lusardi & Tufano, 2009; Stango & Zinman, 2009). Perry (2008) reports that about one-third of the population overestimates its own credit rating, while Courchane et al. (2008) shows that inaccurate assessment leads to higher financing charges and higher probability of denial. Using a field experiment, Balakina, Balasubramaniam, Dimri, and Sane (2020) show that an educational intervention reduces the likelihood of borrowers purchasing a sub-optimal financial product. In another field experiment, Hundtofte (2017) shows that even though distressed borrowers increase repayments in response to loan modification programs, they ultimately fail to realize the financial benefit of the loan modification, due to imperfect financial sophistication and misvaluation of the contract. In this paper, I document that free credit reports result in more consumers applying for mortgages, and a higher probability of mortgage approval.

This paper further relates to the extensive literature on the mortgage expansion in the US before the great recession. The literature has advanced several supply side arguments for excessive mortgage expansion. Mian and Sufi (2009) argues that subprime borrowers had disproportionately large credit growth but lower income growth, and eventually suffered large mortgage delinquencies. Adelino, Schoar, and Severino (2016); Foote, Loewenstein, and Willen (2016) and Conklin, Frame, Gerardi, and Liu (2018) document contradictory results. Keys et al. (2010) shows that the originate-to-distribute model of securitization resulted in lenders relaxing screening of borrowers. In this paper, I document an increase in mortgage applications and approval ratio due to free credit reports, which are suggestive of the demand-side channel of self-learning.

Finally, this research also speaks to the nascent literature on issues related to the information contained in credit reports. A few reports by government agencies — the Federal Reserve Board (FED) and Congressional Research Service (CRS) — and consumer advocacy groups provide a perspective on issues related to credit reports: information recorded in credit reports (Avery, Calem, Canner, & Bostic, 2003); limitation of the credit report data and its consequences on credit (Avery et al., 2004); effect of free credit reports on the CRA industry (Nott & Welborn, 2003); and the extent of errors in credit reports and consumer loss (Cassady & Mierzwinski, 2004; Consumer Federation of America, 2002; Golinger & Mierzwinski, 1998). In this paper, one of the motivations for the increase in demand for mortgage credit is the reduction in demand suppression by consumers incorrectly anticipating rejection.

The rest of the paper is organized as follows. Section 2 discusses the institutional background. Section 3 describes the empirical research design, and Section 4 provides description of the data. In Section 5, I describe key results of the paper, in Section 6 I discuss supplementary results, and in Section 7, I show the robustness tests. Finally, Section 8 concludes the paper.

2 Institutional Background

Laws Governing Consumers' Access to Credit Reports in the US: Before the enactment of FACTA in 2004, the FCRA governed consumer credit information-related laws in the US. Even under the FCRA, consumers had the right to see the contents of their credit report except for the credit score (Avery et al., 2003). 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. Under the FCRA, consumers could also receive free credit reports under specific circumstances, for example, if a consumer 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.[5]

[5] 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.

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).[6] 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, residents of seven states – Colorado (CO), Georgia (GA), Maine (ME), Maryland (MD), Massachusetts (MA), New Jersey (NJ), and Vermont (VT) – could access free credit reports under respective state laws enacted over the years (Figure 2). I call these seven states the pre-FACTA states. Consider, for example, the state of Colorado. It enacted a law for providing free credit reports 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 provisions of the FCRA were to expire in 2003. Thus, in order to make those provisions permanent, the FACTA was enacted on December 4, 2003. One new key provision of FACTA was to provide free annual disclosure of credit reports to consumers by each of the three national credit reporting agencies. As a federal law, FACTA is applicable to US consumers from all states. Thus, residents of the pre-FACTA states can enjoy free credit reports under FACTA as well as under their respective state laws, while residents of other states can enjoy free credit reports only under FACTA.

3 Empirical Research Design

As discussed, residents of the seven pre-FACTA states could access credit reports for free prior to 2005. These seven states and the year in which they adopted free credit report laws are as follows: Colorado (1997), Georgia (1996), Maine (2003), Maryland (1992), Massachusetts (1995), New Jersey (1996), and Vermont (1992). Figure 2 depicts this information on a map of

[6] 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.

the US. Owing to the nationwide extension of these local laws by the enactment of FACTA in December 2003, and the establishment of the website www.annualcreditreport.com in January 2005, residents from all the US states could begin to access free credit reports annually from three national credit reporting agencies.[7]

To evaluate the effect of free credit reports, I designate the seven pre-FACTA states as a control group, and the states bordering these pre-FACTA states as the treatment group in a DID setting. Figure 3 illustrates the control and the treatment states on the US map. To isolate the treatment effect of free credit reports from the effect of local economic conditions, I use a *border county strategy*. Here, I restrict the sample to a narrow geographic area around the borders of the pre-FACTA states and surrounding states. I keep only the counties lying at the border of the pre-FACTA control states and surrounding treatment states in the sample, and remove all the inner counties of the treatment and control states. Figure 4 illustrates this strategy. For example, Garrett County, Maryland (a border county of a control state) and Somerset County, Pennsylvania (a border county of a treatment state) will be included in the sample, while inner counties of Maryland (e.g., Howard County) and Pennsylvania (e.g., Montgomery County) will be excluded.

As discussed previously, by utilizing the federal extension of the local laws as the event, and not the adoption of the local laws by the state, my research design mitigates the endogeneity concern that states enact legislation in response to prevailing local socioeconomic conditions. The case of Vermont illustrates the endogeneity problem if one were to focus on a state's adoption of free credit report laws. In 1991, Experian (then, TRW Inc.) erroneously garbled the credit report of every citizen residing in Norwich, claiming that everyone had failed to pay property taxes. This could have resulted in banks or other lenders rejecting these residents if they applied for new credit. TRW settled the subsequent state lawsuit in 1992 and paid compensation to each affected homeowner. Vermont responded with the 1992 legislation mandating free annual credit reports for every requesting state resident.

[7] The website was rolled-out in four phases from Dec 2004 to Jan 2005. Phase I roll-out was on Dec 1, 2004 across 13 states: *Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming*. Phase II roll-out was on Jan 3, 2005 in 12 states: *Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin*. Phase III roll-out was on Jan 6, 2005 in 11 states: *Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, Oklahoma, South Carolina, Tennessee, and Texas*. Phase IV roll-out was on Jan 9, 2005 in 14 states and one territory: *Connecticut, Delaware, DC, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, North Carolina, Pennsylvania, Rhode Island, Vermont, Virginia, and West Virginia*.

The main regression specification is the following:

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

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. Hence 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 controls for time-invariant and census tract-specific fixed effects. I cluster the standard errors at the county level in all specifications to account for any potential correlation in error terms from census tracts belonging to the same county, and also to account for any serial correlation.

I control for local economic conditions in two ways: $Economic\ controls$ 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 states (surrounding treatment states). 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 important of all, $\gamma_{j,t}$ flexibly controls for any time-varying regional economic shocks that may affect bordering states.

4 Data

The key data used in this paper come from the US mortgage data available under the *Home Mortgage Disclosure Act of 1975* (HMDA). These data provide application-level details on applicants' demographics (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. The period of the study is from 2000 to 2008. I use the data until 2008 to allow for enough post-experiment observations, since the experiment occurs at the beginning of 2005. 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 Servie Agency (FSA)/Rural Housing Administration (RHS), and loans insured by Federal Housing Administration (FHA)).

The coverage of mortgages in the US under the HMDA is the widest. It contains 190.3 million application-level observations over the sample period.[8] I start processing this data, containing 190.4 million applications, by first removing observations for which state, county or census tract information is missing or "NA", or state Federal Information Processing Standard (FIPS) code is "0", "00" or "0 ". This drops 2.5% of the observations, leaving 185.6 million mortgages with an identifiable county. Next, I drop three types of mortgages: first, the covered loan purchased by the financial institutions from other institutions (18.80%), as these mortgages are not borrower-initiated transactions; second, pre-approval requests denied by financial institutions (0.01%), as these data were included in HMDA reporting only from 2004; third, 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 also because their reporting is optional (0.025%). This leaves 150.7 million mortgage applications in the sample. I aggregate these mortgages at the census tract and year level, yielding 77,525 unique census tracts and 603,840 *Census Tract* $\times$ *year* observations. Utilizing the *county adjacency* data, I select the census tracts that lie in the counties at the border of the control and treatment states (see Figure 4) (Census Bureau, n.d.). These census tracts constitute the final sample of 12,047 unique census tracts — 7,057 in the treatment group and 4,990 in the control group — resulting in 90,353 observations at the "*Census Tract* $\times$ *Year*" level.

Though the coverage of HMDA data is the largest, it lacks some key application-level information, such as the credit score of the applicant. Hence, I utilize data from two main GSEs — the Federal National Mortgage Agency (Fannie), and the Federal National Home Loan Mort-

[8] 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).

gage Corporation (Freddie). These data, henceforth called the GSE data, pertain to the 30-year fixed rate single family mortgages, the most popular mortgage type in the US. This consists of 33 million mortgage-level observations on debt-to-income ratio, credit score, first-time home-buyer flag, investment purpose and more. The property location is available as a 3-digit zip code (henceforth, zip3) and state. To map zip3 locations to the counties at the borders of the states, I use the 2010 Q3 version crosswalk files provided by the US Department of Housing.[9] I aggregate the mortgage-level observations to zip3-state level. This creates a panel of 225 unique zip3-states, leading to 7,711 "$Zip3\text{-}State \times Quarter$" observations.

To understand the effect of free credit reports on the commercial banks, I use the quarterly "Call Reports" (FFIEC Forms 031/041) data. The key task here is to match the mortgage lenders in the HMDA data with the commercial banks in the Call Reports. I do this using the HMDA Ultimate Panel Data. HMDA identifies a mortgage lender by a combination of an Agency Code (banking regulator of the mortgage lender) and a Respondent ID (a unique identifier assigned by the lender's regulator), while Call Reports uniquely identify the commercial banks using an RSSD ID. In the HMDA data, Respondent ID represents a Federal Deposit Insurance Corporation (FDIC) Certificate ID if the lender's regulator is FDIC; and it represents the Office of the Comptroller of the Currency (OCC) Charter Number if the lender's regulator is OCC. Further, HMDA started providing the RSSD ID of lenders from 2004 onwards. So, until 2003, I accomplish the matching using the Respondent ID, while from 2004 onwards, I use the RSSD ID. Finally, many HMDA lenders operate as affiliates of commercial banks, and report the bank as "parent" in the HMDA Ultimate Panel data. Whenever a lender has a parent, I first use the lender's Respondent ID to identify the corresponding bank; and if this fails, I use lender's parent's ID for matching. In cases that an HMDA lender and its parent both are a commercial bank (both file Call Reports and are matched successfully), I keep the parent's match as the corresponding commercial bank.

I use a few other data sources in this paper. To gauge the creditworthiness of a county, I use Federal Reserve Bank of New York and Equifax (2000–2008) data on subprime county population. To measure county economic characteristics, such as employment and number of establishments, I use data from the annual survey of County Business Patterns (Census Bureau,

[9] 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.

2000—2008). To map the zip code-level variables from County Business Patterns to census tract-level, I use data from the Missouri Census Data Center (2010). The data on state level economic conditions comes from the Bureau of Economic Analysis, and the data on population characteristics 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 and approval ratio. Other variables that I use 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. I define an application to be successful if the mortgage has either been originated or the application has been approved but not accepted by the applicant. The approval ratio is the ratio of the number of successful applications to the number of total applications in a census tract. Similarly, I calculate all the other ratios as a fraction of total applications in a given census tract.

Table 1 provides the summary statistics of the HMDA sample. Panel A shows the summary statistics for the entire sample, while panel B shows the statistics for the pre-treatment sample. We see that the treated census tracts have fewer mortgage applications per 1000 adults, a lower mortgage approval ratio, and a higher fraction of applications denied due to credit history and debt-to-income ratio. 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.

Panel B of Table 1 shows the comparison of the treatment and control groups in the pre-treatment period. We see that the differences between the treatment and control census tracts in the pre-treatment sample are similar to those observed in the full sample in Panel A. Panel B also shows the p-value for the t-test of the difference in mean between the two groups for each variable. Results from the t-test suggest that the control and treatment census tracts are different in pre-treatment years on these observed characteristics. This also raises the endogeneity concern that these groups may differ on unobservable characteristics.

A DID setting does mitigate the above concern as it can accommodate any pre-existing difference between the treatment and control subjects so long as they satisfy the parallel trends assumption. Though we cannot formally test this assumption, in Figure (5) I examine whether the parallel trend assumption is plausible in the current setting. Panel A shows the mean approval ratio across the treated and control census tracts over the sample period. We see that approval ratio across the two groups trend in parallel.

Furthermore, in Panel B of Figure (5), I plot the coefficients that represent the difference in approval ratio for the two groups over the years relative to the pre-event year (2004). I obtain these coefficients (β_k) from regressing "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 $t = T - j$. $\text{Event}_k = 0$ if $t \neq T - j, j = \{-3, 4\}$. $T =$ Event year 2005.

We see from the plot in Panel B that no significant difference exists in the approval ratio between the two groups before the experiment, but the difference becomes significant afterwards. Overall, the two plots in Figure (5) provides reasonable assurance that the parallel trend assumption is satisfied in the current setting.

Finally, in all regressions I include "*Census Tract*" fixed effects that controls for any time-invariant differences at a granular geographic level. Recall that a census tract covers an area accommodating about 2,500–8,000 population. Further, I also include "*Border* $\times$ *Year*" fixed effects that flexibly controls for any time-varying regional shock impacting neighboring states.

One may argue that the enactment of FACTA in Dec 2003, and the subsequent establishment of www.annualcreditreport.com, is not a salient event for consumers, and hence, this natural experiment is irrelevant to examine the link between cost of credit reports and credit market outcomes. I address this concern in two ways. First, I plot the search interest for the key phrase *Free Credit Report* from 2004 to 2010 in Figure (6).[10] Search interest shows a significant peak in Jan 2005, the time of the establishment of the website, suggesting that there was significant consumer interest in free credit reports in 2005.

Second, I show in Figure (7) that the popularity of the key phrase *Free Credit Report* increased significantly in the treatment states after the event, but there was no difference in popularity before the event. To do this, I begin by sourcing the interest-by-subregion data from Google

[10] 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 starts from January 2004.

Trends for the key phrase *Free Credit Report* for each year from 2004 to 2008. I calculate the mean of the popularity rank for the treatment and control states in each year. Then, I calculate relative popularity as ($\text{Popularity}_{trt} - \text{Popularity}_{ctrl}$) each year. Figure (7) shows the plot of relative popularity over time. We see that key phrase *Free Credit Report* 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 from 2005, the experiment year.[11]

5 Results

In this section, I provide the key empirical results of the study. I first provide the baseline results on how free credit reports affect mortgage demand and approval ratio, house prices, and default rates. Then, I discuss the results that highlight the proposed self-learning mechanism. Finally, I show several findings centered on understanding the role of supply-side and demand-side factors in driving the results.

§A Baseline Results

§A.1 Mortgage demand and approval ratio

The baseline regressions use Equation (1) to examine the effect of free credit report laws on the number of applications per 1000 adults and the approval ratio. Each outcome variable is calculated at the census tract level. All specifications include *Census Tract* fixed effects and "$Border \times Year$" fixed effects. As discussed previously, *Census Tract* fixed effects control for any pre-existing difference across different census tracts; and "$Border \times Year$" fixed effects flexibly control for any regional shocks that may arise across different years, and also ensures that the comparison between the treatment and control census tracts is restricted to geographically close areas.

Table 2 shows the results of regressing the number of applications per 1000 adults in a census tract using Equation (1). Column (1) shows the result for the simplest DID specification without any controls, while (2) shows the result with economic controls (annual growth rate of county income per capita, county aggregate employment, and state GDP) added. The coefficient of interest is "$Treat \times Post$". It captures the change in number of mortgage applications per

[11] It is important to note that the methodology Google uses to calculate the popularity of a keyword in a region implies that increase in popularity of one area would mechanically reduce the popularity in other areas. Also, we cannot interpret how much the popularity increased over the year, as the popularity rank is re-based to 100 and is assigned to the region where the keyword is most popular, every year.

1000 adults in treated areas relative to the control areas after free credit reports became available in the former. In columns (1) and (2) we see that a treatment census tract sees an increase of 13.29–15.45 mortgage applications per 1000 adults. This is 13.8–16.0% increase in applications (over the pre-treatment average of 96.65 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 million per 1000 adults in a typical census tract ($150,597 $\times$ 13.29), by about $5.4 million per treated census tract ($2 million $\times$ 2.7 thousand adults per census tract), or, by about $38.1 billion in the entire treated area after free credit reports became available ($5.4 million $\times$ 7,057 treated tracts).

Coefficients on "$Treat \times Post$" in columns (3) and (4) estimate the effect of free credit reports on the approval ratio at the census tract level. The ratio increased by 1 percentage point in the treatment census tracts after free credit reports became available. This increase represents ~2.6 more successful applications per treated tract (96.65 applications per 1000 audlts in pre-treatment period $\times$ 0.01 $\times$ 2.7 thousand adults per treated census tract), ~18,348 more successful applications in the entire treated area (2.62 applications $\times$ 7,057 treated census tracts), or a ~$2.8 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, could be interpreted to be a result of improvement in the borrower pool if the increased origination is primarily driven by demand-side factors, which I argue in the later section to be the case in the current setting.

Next, I examine whether the mortgages in the treated areas were utilized for occupancy purposes, or investment purposes. For this, I re-estimate the baseline regression for only the mortgages in the owner-occupied category. Panel A of Table (3) shows the results. The estimates are broadly similar to those from the baseline specifications: increase of 9.8–11.5 (10.1–11.9%) in applications per 1000 adults, and increase of 1 percentage point in approval ratio. Further, in Panel B of Table (3), I examine whether the fraction of non-owner-occupied mortgages at the application or the origination stage changed. For this, I regress 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). Insignificant coefficients in the first two columns imply that not-for-occupancy mortgage applications did not increase in the treated areas . Also, coefficients in the last two columns provide weak evidence of a small increase in the fraction of not-for-occupancy mortgages at the origination stage by 1 percentage point. As ~86% of the applications in the HMDA data are for occupancy purposes, this increase is economically small, and we can reasonably conclude that free credit report laws led to an increase in demand mostly for occupancy purposes.

§A.2 Effect on house prices

Since free credit reports led to an increase in demand for mortgages, a natural question then arises, did the house prices increase in the treated areas? I examine the effect on house prices using a highly granular census tract-level house price index developed by Bogin, Doerner, and Larson (2016). I use the 2000 version of the index, which means that its value for all census tracts is 100 in the year 2000.

Table 4 shows the results of regressing the growth in the house price index using Equation (1). The marginally significant coefficients on "*Treat* × *Post*" in columns (1) and (2) suggest that the growth rate of house prices increased in the treated areas by 1.7–1.8 percentage point after the event. This magnitude of increase is in line with, but slightly lower than, that in Di Maggio and Kermani (2017), which reports a 3.3 percentage point increase in the growth rate of house prices, resulting from a 10% increase in mortgage origination in their sample.

§A.3 Effect on 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 on these mortgages. Another possibility is that the consumers in the treated areas self-learned their creditworthiness better after the free credit reports became available, and the increased demand came from good borrowers who could repay the mortgages. In this latter case, mortgages from the treated areas would be equally or less likely to be defaulted upon after the event. So, which of these two predictions holds in the data?

I use the publicly available default data for mortgages purchased by the two GSEs, Fannie Mae and Freddie Mac, for comparison of default rates. I match the GSE data to the bordering counties of control and treatment states using the steps detailed in the Data section. I then define two vintages of the mortgages: those originated in the event year 2005 and those in the pre-event year 2004. The default rate for a given vintage-year is the fraction of total mortgages of a given vintage that subsequently miss a scheduled payment in a given month by 30–59 days for the first time, namely $\text{Def}_{2005,age}$ and $\text{Def}_{2004,age}$, where age is measured in months since origination. I calculate this rate for both treated and control areas separately, and then define the adjusted default rate as follows:

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

I calculate the above measure every month over six years. A positive (negative) adjusted default rate implies that the mortgages from the treated areas are more (less) likely to be defaulted upon than those from the control areas at a given age in the event year than the year before.

Figure 8 shows the plot of adjusted default rate with age. The plot 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 point over a six-years period, and is statistically significant at a 1% level (p-value = 0.0000). This means that, on average, mortgages from the treated areas were 0.012 percentage point less likely to be defaulted upon up to six years after origination, relative to a mortgage from the control area, after the free credit reports became available in the treated areas.

Another important observation from this plot is the adjusted performance of the mortgages during the financial crisis of 2007. In this plot, age greater than 48 months corresponds to the bust years post the financial crisis (48 months after 2005 is 2009). We see that even during the bust years, the mortgages from the treated areas were less likely to be defaulted upon than those from the control areas.

Overall, we learn that free credit reports stimulated the mortgage demand, raised the approval rates, and resulted in lower *ex-post* defaults. This is consistent with an improved borrower pool in the mortgage market owing to better self-learning among consumers.

§B Consumer Self-Learning Channel

Having established that the mortgages and approval ratio increased in the treated areas, I proceed next to examine three outcomes that shed light on the proposed *self-learning* channel through which free credit reports may affect the mortgage demand and approval ratio. The three outcomes are credit history-induced rejection, consumer search accuracy, and first-time homebuyers.

§B.1 Contraction in credit history-related rejections

We can proxy for self-learning among consumers by analyzing the likelihood of mortgage rejections due to credit history. To understand why, recall that free credit report laws do not change the financial condition of consumers. The only real change is that consumers get to know their credit history free of cost, and also understand what lenders see about them. If we were to assume that knowledge about their own creditworthiness among the consumers was perfect before the experiment, and self-learning did not increase after free credit reports became available, the likelihood that an application would be rejected due to credit history would remain the same. Further, likelihood of rejections due to debt-to-income ratio, the second most frequent reason for mortgage rejection, too, would remain unchanged.

Using Equation (1), I test whether the likelihood of mortgage rejection due to credit history and debt-to-income ratio changes 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. I estimate this 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).[12] 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 see improvement by self-learning from free reports.

Table 5 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 point in the treatment census tracts relative to the control census tracts, but it is significant only in the areas that had high rejection rates before the experiment (in columns 3 and 4). Plausible reasons for the treatment effect to be significant only in the high rejection areas is 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 less for consumers who have a higher probability to receive credit. 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.

Note also that not only we see contraction in the likelihood of credit history-related rejec-

[12]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.

tions (columns 3 and 4), but we also see no change in debt-to-income ratio-related rejections (columns 5 through 8). Albeit weak, these results together are suggestive of the self-learning mechanism.

A further caveat limits the interpretation of the results of this test. 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.

§B.2 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 (in)accuracy in their search for a suitable mortgage lender. To discern this link, we need to understand a typical mortgage application process. Owing to the uncertainty over their application, and also 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. 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.[13] In this context, a high rate of withdrawals represents uncertainty among consumers over outcome, or their inaccuracy in searching for a lender. Importantly, this tendency is not small: consumers withdrew about 12% of all mortgage applications over the 2000–2008 period. Furthermore, credit reporting agencies do not penalize multiple applications by consumers if they are made within a short window.[14]

All applications that consumers withdraw before the lender has made the credit decision are reported as *withdrawn* under HMDA regulations. I argue that under the free credit reports regime, consumers self-learn their creditworthiness better before the application, and be-

[13] Anecdotal evidence suggest that consumers tend to withdraw an application when they find a better offer from other lenders. https://www.reddit.com/r/personalfinance/comments/38k1l5/withdrawing_a_mortgage_application/

[14] For example, Experian, one of the three national CRAs, explicitly mentions its policy of treating multiple applications as one. "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." https://www.equifax.com/personal/education/credit/report/understanding-hard-inquiries-on-your-credit-report/

come more certain over their application outcome (approval probability and mortgage terms). Hence, they apply at fewer lenders *ex-ante*, save multiple application costs along the way, and ultimately, withdraw fewer applications in aggregate.

I test the above prediction that the withdrawal tendency among consumers will decrease if they self-learn better after the free credit reports. I regress the fraction of total mortgages that are withdrawn while in-process using Equation (1). Table 6 shows the results. We see that the fraction drops by 0.9–0.11 percentage point 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 entire treatment area. At an average cost of ~US $400 per withdrawn application, consumers saved ~US $6.6 million 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.

§B.3 Increase in first-time homebuyers

The third outcome that directly points to an increase in self-learning among consumers is the enlarged fraction of first-time homebuyers in the mortgage market. 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 we do not know what fraction of these consumers overestimate the rejection, and incorrectly suppress their credit demand, the provision of free credit reports certainly can aid them in better assessing their creditworthiness. If we see that the fraction of first-time homebuyers in the mortgage market increases after the experiment, we can conclude that first-time homebuyers on average overestimate the rejection probability, and hence more of them enter the credit market after self-learning, and that free credit reports reduce the demand-suppression phenomenon.

While the HMDA data are comprehensive, they do not record the information as to whether an applicant is a first-time homebuyer. However, the subset of mortgages purchased by the two GSEs do capture this information. From the location data of these mortgages (available as the 3-digit zip codes), I approximately identify those for the properties residing within the border counties of the treatment and control states, using the steps outlined in the Data section. Another empirical issue with the GSE data is that about 6.71% of the observations over the sample period do not have information as to whether the mortgagor is a first-time homebuyer.

To address this issue of missing information, I calculate the fraction of first-time homebuyers as a fraction of all GSE mortgages, or as a fraction of all GSE mortgages with valid first-time homebuyers information.

The outcome variable for this test is the fraction of first-time homebuyers, and I use the following regression equation:

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

Here, z indexes the areas delineated by a combination of 3-digit zip codes and states that lie along the border counties of treatment and control states. Note that the geographic aggregation for this regression is the zip3-state level (unlike the previous regressions for which aggregation was at the census tract level). The remaining terms in (3) are the same as in Equation (1). Table 7 shows the results of this regression. The percentage of first-time homebuyers increases by 1 percentage point in the treatment areas relative to the control areas. Hence, this finding confirms that self-learning increased after the free credit reports, and it addresses 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: GSEs selected these mortgages to be fit for purchase (conforming mortgages). However, selection by GSEs would only be a concern if their incentives or mandate to purchase first-time homebuyer mortgages relative to their overall purchase differentially changed across the treatment and control counties in the year 2005. Such a change seems implausible, and so the issue of selection by the GSEs does not seem critical in the above regressions.

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

We have seen that the free credit reports resulted in increased demand and origination in the treated areas. We also saw some preliminary facts suggestive of a demand-driven mechanism: increased self-learning among consumers (a demand-pull effect). However, supply-side factors, too, offer an alternative explanation for the observed increase. Since this experiment treated not only the consumers, but also the lenders, with the knowledge that consumers could now get free credit reports, lenders in response could increase the approval rates (supply-push effect). In the limiting case, one may argue that the entire rise in mortgage origination was due to lenders increasing the approval ratio, and not due to a better borrower pool stemming from improvement in self-learning among consumers about their creditworthiness.

In this section, I lay out the evidence suggesting that, while supply-side factors may have had a role in the outcomes after the experiment, it was the demand-side factors that primarily drove the results. I begin with the observation that the number of applications increased in the treated areas, over which lenders had only limited control, if any. Similarly, the drop in in-process withdrawals is a consumer-driven result and is free from influence from the supply-side factors. Notwithstanding these preliminary arguments suggestive of a demand-driven effect, I employ two strategies to disentangle the contribution of supply and demand factors.

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

First, I appeal to the basic supply and demand framework to identify if the results are driven by the supply- or demand-side factors. We know that, if markets clear, an increase in quantity of a good as well as its price is possible only if the demand curve shifts outward, and the supply curve either stays the same, or shifts inadequately so that the equilibrium price needs to rise to match the increased demand. Thus, in the current context, I test if the mortgage interest rates (price of mortgage) increased in the treated areas.

As the HMDA data do not capture interest rates of the mortgages, I again turn to the GSE data, the subset of mortgages purchased by Fannie Mae and Fredie Mac, to examine the effect on interest rates. I regress the mean and median interest rate in a zip3-state area using Equation 3. Table 8 shows the results. We see that the mean (median) interest rate in the treated areas increased by 1.1–1.2 (1–1.1) basis points in the treated areas, while accounting for changes in the control areas. The increase in interest rate is consistent with an outward shift in the demand curve, confirming a demand-pull effect after the free credit reports became available.

We see that the increase in the mean (median) interest rates in the treated areas is statistically significant, but quantitatively small. The increase is small for a few reasons: firstly, the supply of mortgages in the US is highly elastic because of the GSEs' mandates to purchase conforming mortgages from lenders; secondly, the pricing of standard 30-year mortgages fluctuates little across lenders as compared with other credit instruments, such as credit cards or auto loans. Also, note that even if the increase in interest rate were statistically non-significant in the treated areas, it would have indicated that the demand in the treated areas increased, and the supply could match the demand such that the prices did not need to rise. For the supply-side factors to be the primary driver of increase in mortgage origination, the interest rates would need to decrease in a statistically significant manner in the treated areas. Hence, increase in the interest rate in treated areas confirms that a demand-driven effect is at work.

§C.2 Distinguishing supply and demand: Heterogeneity tests

My second strategy to disentangle the effect of supply- and demand-side factors in driving the results is to employ four heterogeneity tests on effects of free credit reports across different groups. I rely upon the idea that the effect of free credit reports will vary with the characteristics of the primary driver of the mechanism— lenders, consumers, or both. In other words, if lenders were the primary driver, the effect would be heterogeneous in the characteristics of the lenders; if consumers are the primary driver, the effect should be heterogeneous in the characteristics of the consumers; and finally, the effect would be heterogeneous in both sets of characteristics to the extent that these two sets of characteristics correlate with one another. Hence, by examining the heterogeneous effects of free credit reports across characteristics of lenders and consumers, we can shed light on the role of supply- and demand- side factors in increasing the mortgage origination.

(i) Heterogeneous effects by density of lenders: Firstly, I investigate if the effects are heterogeneous in the density of lenders. I test whether the areas having a higher density of lenders saw higher origination and approval ratios compared to the areas having a lower density. If the increase in mortgage origination were driven by lenders, we would expect more mortgage origination and a greater increase in approval ratio in areas where the density of lenders is high. I calculate the density of lenders as the number of HMDA lenders per adult in each census tract in the pre-event year 2004. I classify a census tract to have a high density of lenders if this density is more than the *regional mean* (see Footnote 12).

I separately regress the volume of mortgages originated and the approval ratio using the regression in Equation (1) for the areas with a high and low density of lenders. Table 9 shows the results. Columns (1) through (4) show the results for total amount of mortgages originated (in 1000 USD) per adult in a census tract. We see that the coefficient of $Treat \times Post$ is smaller in specification (2) than in specification (1), and smaller in (4) than in (3). These estimates suggest that the increase in mortgage origination in the high-lender-density areas in the treated group is smaller than the increase in the low-lender-density areas in the treated group, controlling for concurrent changes in the control group. At the bottom of the table, I report the t-test for the difference in the coefficient of the interaction term $Treat \times Post$ in high- and low- lender-density areas $(High - Low)$. The results confirm that the increase in mortgage origination in high-lender-density areas is not statistically different from that in low-lender-density areas. In columns (4) through (8), I repeat the analysis for the approval ratio. 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 mortgage approval ratio could have increased solely due to lenders increasing the supply of mortgages.

(ii) Heterogeneous effects by creditworthiness of consumers: Secondly, I look for heterogeneous effect of free credit reports across creditworthiness of consumers. If after the availability of free credit reports, consumers self-learn about their creditworthiness and apply for a mortgage after they become more certain over the outcome, the areas that *ex-ante* had a larger prime population fraction should see a larger increase in approvals. Furthermore, the creditworthiness of an area cannot increase simply because free credit reports become available. It would improve if consumers with better than average creditworthiness entered the credit markets in the treated areas, or, if consumers with worse than average creditworthiness took steps to improve it after learning the true value from their reports. Thus, if we see an increase in origination in high creditworthiness areas after the experiment, it must be due to a consumer-driven self-learning effect, rather than a lender-driven increase in mortgage supply to such areas.

To examine this prediction, I begin with Federal Reserve Bank of New York and Equifax (2000–2008) data on the percentage of the population that is subprime (with credit score <660) in a county. I classify a county into the high (low) creditworthiness category if its prime population fraction is higher (lower) than the *regional mean* (see Footnote 12). For this classification, I use data from the year 1999. This is because the creditworthiness of the population in a county might endogenously change with the onset of a housing boom. Mian and Sufi (2009) suggests that such classifications should be done at a time well before the start of the housing boom.[15]

Table 10 shows the results of regressing the number of applications per 1000 adults and the approval ratio at census tract level using Equation (1), estimated separately for prime and subprime counties. The estimates in columns (1) through (4) show that applications per 1000 adults increased by 16–18 (17.3–18.7%) and approval ratio increased by 2 percentage points

[15] Mian and Sufi (2009) uses year 1996 to classify zip codes into prime and subprime. I use the year 1999, as this is the earliest year for which the data are publicly available.

in treated high creditworthiness counties relative to the control counties with similar creditworthiness. However, coefficients in columns (5) through (8) show that in the low creditworthiness treated counties applications per 1000 adults increased by just 8–10 (8.8%–11.1%) and approval ratio increased by 1 percentage point. These results provide two insights. First, free credit reports seem to result in an increase in the number of applications in prime areas. Second, the higher increase in approval ratio in high creditworthiness counties suggests that increased mortgage supply in this setting is not driven by increased subprime credit.

Notwithstanding the noisy county level creditworthiness measure in the above regression, I turn to a more precise proxy for the creditworthiness of a locality — the number of payday lenders establishments. The number of payday lenders in a locality correlates negatively with local creditworthiness, because they tend to operate in subprime areas (Prager, 2009). Since the location of payday establishments is available at 5-digit zip code level from the Survey of County Business Patterns, their density is a more precise proxy than the county-level measure used in the previous regressions. However, many states restrict payday lending activities (Prager (2009) and Bhutta (2014)). In my sample, only Colorado and the bordering states allow unrestricted payday lending activity. Thus, I conduct this test only for all the census tracts in the counties at the border between CO, the control state, and WY, UT, AZ, NM, OK, KS and NE (the treatment states). I classify these census tracts into two sub-samples using the average number of payday lenders in census tracts within the bordering counties of these states in the pre-treatment year 2004.

Table 11 shows the results of estimating the previous specification with the new proxy. These results show dramatic differences in the effect of free credit reports across high and low creditworthiness areas. In columns (1) through (4) we see that the applications per 1000 adults increased by 68–72, while the approval ratio increased by a staggering 5 percentage point in treated high-creditworthiness areas compared to control high-creditworthiness areas. In columns (5) through (8), we see that an increase in applications is smaller, at 43, while there is no significant increase in approval ratio at all in treated but low-creditworthiness areas compared with control areas with low creditworthiness. These results corroborate the earlier finding using the county-level measure.

Overall, the finding that the effect of free credit reports on approval ratio varies with consumer creditworthiness, a demand-side characteristic, suggests that the increased origination is primarily due to a consumer-driven demand-pull effect.

(iii) Heterogeneous effects by education level of consumers: Thirdly, I examine the heterogeneous effect of free credit reports across education levels of consumers. As with debt products, the information about credit products contained in a consumer credit report may be complex. Thus, if the free credit reports affect approval ratio through consumer self-learning, their effect on approval ratios should be stronger for more educated consumers. I test this prediction next.

HMDA data do not record the applicants' education. Thus I use a location-based proxy for testing this hypothesis. I calculate the fraction of the adult population, aged between 18 and 64 years, having a graduate or equivalent education in each census tract using data from Census 2000. A person is classified as a *Graduate or equivalent* if he or she has an associate degree, a bachelor's degree, or a graduate *or* professional degree. I calculate the *regional mean* of the fraction of graduates in a census tract in each region (see Footnote 12). I then classify a census tract as high- (low-) education if its graduate population fraction is higher (equal or lower) than the *regional mean*.

Table 12 shows the results of regressing the number of applications per 1000 adults and the approval ratio in high- and low- education areas, estimated separately using Equation (1). We see in columns (1) through (4) that the increase in number of applications per 1000 adults (approval ratio) is 6.0–6.8 (1 percentage point), and almost statistically insignificant for treated low-education areas compared with control low-education areas. However, the estimates in columns (4) through (8) show that applications per 1000 adults increased by 12.5–15.0 (12.9%–15.5%) in the high-education treated areas relative to the high-education control areas, and the approval ratio increased by 2 percentage point (more than in low-education areas), both results being statistically significant at the 1% level.

The finding that the effect of free credit reports varies with the education level of consumers, a demand-side characteristics, corroborates the demand-pull effect coming through the self-learning channel: a more educated population can assess the information in credit reports better than a less educated population, leading to higher improvement in the applicant pool and approval ratios.

(iv) Heterogeneous effects by income level of consumers: Finally, I examine the effect of free credit reports across consumers from different income groups. The idea behind this test is twofold. First, the cost of mortgage rejection is higher for low-income consumers. Second, marginal propensity to borrow is high for such 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 a simple rise in lender-driven mortgage supply. However, we cannot infer the same from a similar change observed for high-income consumers.

To test the effect of free credit reports across income groups, I divide the mortgage applications each year into four income groups (income quartiles), and calculate the number of applications per 1000 adults and approval ratio for each quartile. Table 13 shows the results of separately regressing these variables for each income quartile using Equation (1). In Panel (A), we see that the applications increased in each quatrile except the lowest, and the effect is larger as we move from the lower quartile (column 1 and 2) to the higher (column 7 and 8). This suggests that mortgage demand increased more among higher income consumers.

The results on approval ratio in Panel B shows an interesting pattern. From columns (1) and (2), we learn that the increase in approval ratio is 2 percentage point in the lowest income quartile, and is statistically significant at the 1% level. The increase is not significant for quartile 2 and 3 (columns 3 through 6). Columns 7 and 8 show that for the highest income quartile consumers, the increase is marginally significant but smaller, at 1 percentage point, than that for the lowest income quartile, at 2 percentage point. This finding is remarkable. If one were to argue that the increase in mortgage origination was supply-driven, we would have observed 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. My preferred interpretation is that the scope of error in self-evaluation of creditworthiness is higher for low-income consumers, making these consumers more likely to benefit from accurate self-assessment stemming from free credit reports. Heterogeneity in effects across income groups again corroborates the demand-pull effect.

All in all, both the strategies that I employed — equilibrium interest rates and multiple heterogeneity tests — consistently point towards a significant role of a consumer-driven demand-pull effect in driving the increased origination.

6 Supplementary Results

We have seen that the mortgage demand and origination increased in the treated areas, and it was driven primarily by the demand-side factors. I now provide three supplementary findings, two of which address alternative explanations, while the third quantifies the effect of free credit reports on commercial banks.

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

I examine whether the mortgage origination increased in treated areas due to private (non-government) securitization, instead of my preferred explanation of an improved borrower pool stemming from increased self-learning among consumers. Keys et al. (2010) shows that lenders did lax screening for the just-prime borrowers (credit score$\geq$620) and sold these mortgages to private securitization entities in the lead up to the financial crisis of 2008. If increased approval was due to private securitization, then we should observe a higher fraction of originated mortgages being sold to non-government (private securitization) entities in the treated areas.

To examine the above prediction, I regress 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, using the specification in Equation 1. Table 14 shows the results. 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 concern that private securitization, not a better borrower pool, caused the increase in mortgage origination.

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

I now test the subprime lending hypothesis motivated from Mian and Sufi (2009). It could be argued that the increased approval ratio is due to lenders increasing credit supply to subprime borrowers in the treatment areas. We saw previously in Table (10) and (11) that the effect of free credit reports is stronger in the prime counties (census tracts) and not significant in subprime counties (census tracts). Area-based proxies, though widely used (Di Maggio

& Kermani, 2017; Mian & Sufi, 2009), are still noisy proxies of consumers' creditworthiness. Hence, I now examine this hypothesis using the apllication-level credit scores. I turn to the GSE data for this test as HMDA does not provide the credit score data. I regress the number of mortgages in the treatment and control zip3-state areas separately for prime (credit score $\geq$ 620) and subprime consumers using Equation (3).

Table 15 shows the results. The coefficients in columns (1) and (2) show that the number of mortgages extended to the prime consumers increased (by 308–312) in the treatment zip3-state areas relative to the control zip3-state areas, while those in columns (3) and (4) show that it only increased by ~10 for subprime consumers. Hence, we can be reasonably sure that the increased origination in the context of this paper did not go to subprime consumers. It should be noted, however, that the magnitude estimated in this table is not directly comparable with those estimated in the previous tables using the HMDA data, because the observation unit is the zip3-state in the current regression, but the census tract in the previous regressions.

A caution is warranted in interpreting these results. Since this sample is a selected sample of GSE-purchased mortgages only, a change in prime-subprime composition reflects the combined effect of demand-side factors and variation in incentives of GSEs to purchase prime and subprime mortgages. However, two arguments allay this concern. First, as these changes are estimated using a DID estimator, the time-varying GSE incentives would be a valid concern 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 a precise change in incentive, though not impossible, seems unlikely. Second, Elul, Gupta, and Musto (2020) shows that GSEs increased purchasing subprime mortgages, not prime mortgages, before 2007 in a bid to combat the housing bust. Thus, the current result that the origination increased more for prime consumers in treated areas in the GSE sample cannot be explained by their changing incentives. Hence, we can reasonably attribute this increase to demand-side factors.

§C Effect of free credit reports on banks

Since free credit reports appear to have stimulated mortgage demand, it is informative to understand its effect on the second most relevant agent in this natural experiment, the commercial banks. However, before proceeding, it is important to note the limitations of such an exercise. First, contrary to the general perception, commercial banks are not the dominant mortgage originators: despite being 80% of mortgage lenders by number, commercial banks accounted for just 37% of the mortgage lending activity in 2005 (Avery et al., 2007). Thus, any attempt at identifying effects on commercial banks — the lenders that provide financial data publicly — is incomplete as it excludes the dominant lenders. Second, since most commercial banks operate in multiple states, the natural experiment of this paper may assign a bank to the treatment as well as the control group. Specifically, if a bank operates in both treated and control states, their treatment assignment is not discrete, but rather continuous: such banks receive different intensities of treatment depending on their mortgage activity in the treated states vis-à-vis the control states in the pre-event year.

Notwithstanding these limitations, I evaluate the effect of free credit reports on the commercial banks. To classify these banks into treated or control group, I use the cross-sectional average of the ratio of mortgages originated by a bank in control states to that in control and treated states combined in the pre-event year 2004. I examine the effect on a bank's net interest margin (NIM), return on equity (RoE), and return on assets (RoA) using the following regression specification:

$$Y_{lt} = \beta_0 + \beta_1 \text{Treatment}_l \times \text{Post}_t + \delta \times \text{Bank controls}_{lt} + \alpha_l + \gamma_t + \varepsilon_{lt} \tag{4}$$

where Y represents the outcome variable; l indexes the banks; $Treatment$ is 0 if the bank's ratio of mortgages originated in control states to that in control and treated states combined is greater than the benchmark, and 0 otherwise; $Post_t$ 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 log of total assets of the bank in the given quarter, share of liquid assets to total assets, and cost of deposit.[16]

Table 16 shows the results of the above regression. The estimates suggest that the treated banks saw a robust 6 basis point increase in NIM (columns 1 and 2); a 0.74 percentage point increase in RoE (columns 3 and 4); and a 0.07 percentage point increase in RoA (columns 5 and 6) relative to the control banks. Keeping the limitations mentioned above in mind, these estimates suggest that not only consumers, but also the banks, experienced positive outcomes from the increased demand and improved borrower pool resulting from the free credit report laws.

[16] 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.

7 Robustness

The natural experiment utilized in this paper takes place in the year 2005. The years included in the analysis are between 2000 and 2008 to allow for enough post-experiment observation. As subprime mortgage origination has often been argued to be the reason behind the financial crisis of 2008, it is a valid concern that the results in this paper might be due to the later years 2007 and 2008. To mitigate this concern, I re-estimate all the tests in this paper by excluding 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 (corresponding to the years 2005 and 2006). Hence, the findings of this paper are not driven by the distinctive lending environment that existed prior to the financial crisis.

8 Conclusion

Consumer credit markets suffer from many imperfections. Various data from the US market suggest that a non-trivial fraction of consumers makes credit decisions consistent with them having imperfect information of their creditworthiness. In this paper, I focus on the role of credit reports in shaping the aggregate mortgage market outcomes. Specifically, I examine the causal link between the cost of credit reports for consumers and aggregate mortgage market outcomes. Market level outcomes change presumably because credit reports induce better credit-related decision making among consumers.

I utilize a natural experiment in a difference-in-differences (DID) setting to establish the causal link between the cost of credit report and mortgage demand and approval ratio. The enactment of the federal *Fair and Accurate Transactions Act of 2003* (FACTA) allowed all US consumers access to three free credit reports annually, while seven states had already enacted local laws that allowed their residents to obtain free credit reports. I operationalize the DID setting by designating the border counties of the early-adopting states as the control group and the border counties of the surrounding states as the treatment group. An advantage of this setting is that it relies on the federal enactment of the law while control states already had enacted similar law in the past, thus this strategy avoids the usual endogeneity-related concern in the DID strategies that rely on adoption of laws by states.

I find that a reduction in consumers' economic cost of accessing credit reports resulted in increased credit demand, approval ratio, and first-time homebuyers. The increased origination seem to have come from borrowers of good quality — the approval ratio was higher among prime consumers, and in areas having high creditworthiness population; and these mortgages were less likely to be defaulted upon after origination. Various cross-sectional tests yield results that are consistent with these changes being primarily driven by demand-related factors, than by supply-related factors such as increased credit supply or private securitization.

The implications of the findings in this paper are not limited to only the mortgage-related decisions of consumers, but in general apply to any credit-related decision under imperfect information of their creditworthiness. All in all, we learn that information provision to consumers at lower cost can improve credit market outcomes and may benefit both consumers and lenders.

References

Adelino, M., Schoar, A., & Severino, F. (2016). Loan Originations and Defaults in the Mortgage Crisis: The Role of the Middle Class. *The Review of Financial Studies*, *29*(7), 1635–1670.

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.

Cassady, A., & Mierzwinski, E. (2004). *Mistakes Do Happen: A Look at Errors in Consumer Credit Report*. US PIRG Education Fund.

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 Re-*

lationship Files (*CTRF*) [Database]. Inter-university Consortium for Political and Social Research. Retrieved from https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/13287

Conklin, J., Frame, W. S., Gerardi, K., & Liu, H. (2018). Villains or Scapegoats? The Role of Subprime Borrowers in Driving the Us Housing Boom [Working Paper]. *Unpublished*. doi: https://doi.org/10.29338/wp2018-10

Consumer Federation of America. (2002). *Implications for Consumers* (Tech. Rep.). Consumer Federation of America and National Credit Reporting Association.

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

Federal Reserve Bank of New York, & Equifax. (2000–2008). *Equifax Subprime Credit Population* [Database]. FRED, Federal Reserve Bank of St. Louis.

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

Foote, C. L., Loewenstein, L., & Willen, P. S. (2016, December). Cross-sectional Patterns of Mortgage Debt During the Housing Boom: Evidence and Implications [Working Paper]. *Unpublished*. Retrieved from http://www.nber.org/papers/w22985

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

Golinger, J., & Mierzwinski, E. (1998). *Mistakes Do Happen: Credit Report Errors Mean Consumers Lose*. US PIRG Education Fund.

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]* [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.

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

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)

Figure 1:
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 2:
States Providing Free Credit Reports prior to FACTA (pre-FACTA states)

This figure shows the seven US states which had enacted a free credit reports laws prior to 2004 — Colorado, Georgia, Maine, Maryland, Massachusetts, New Jersey, and Vermont — in 1997, 1996, 2003, 1992, 1995, 1996, and 1992, respectively. These seven states are referred to as the pre-FACTA states or the control states. In December 2003, with the enactment of the Fair and Accurate Credit Transactions Act (FACTA), free credit reports became mandatory in all US states. The website www.annualcreditreport.com was subsequently established in Dec 2004 to distribute the free credit reports (see Footnote 7).

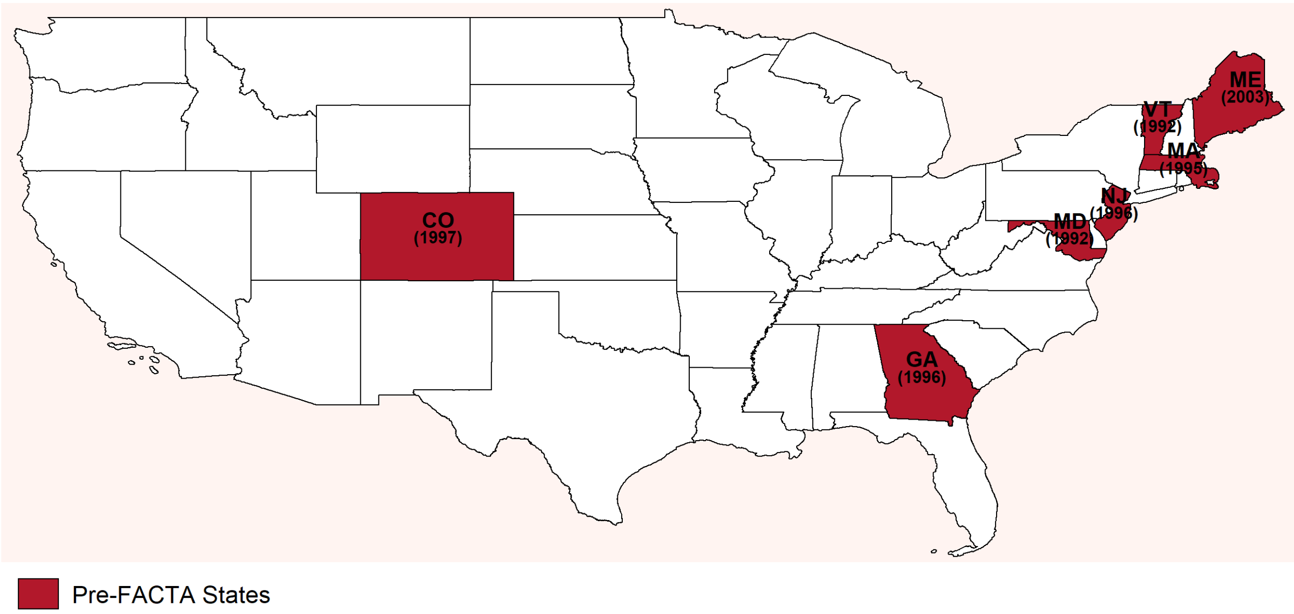


Figure 3:
Empirical Setting: Control and Treatment States

This figure shows the seven pre-FACTA states, serving as the control states, and the 26 states that surround the control states, serving as the treatment states.

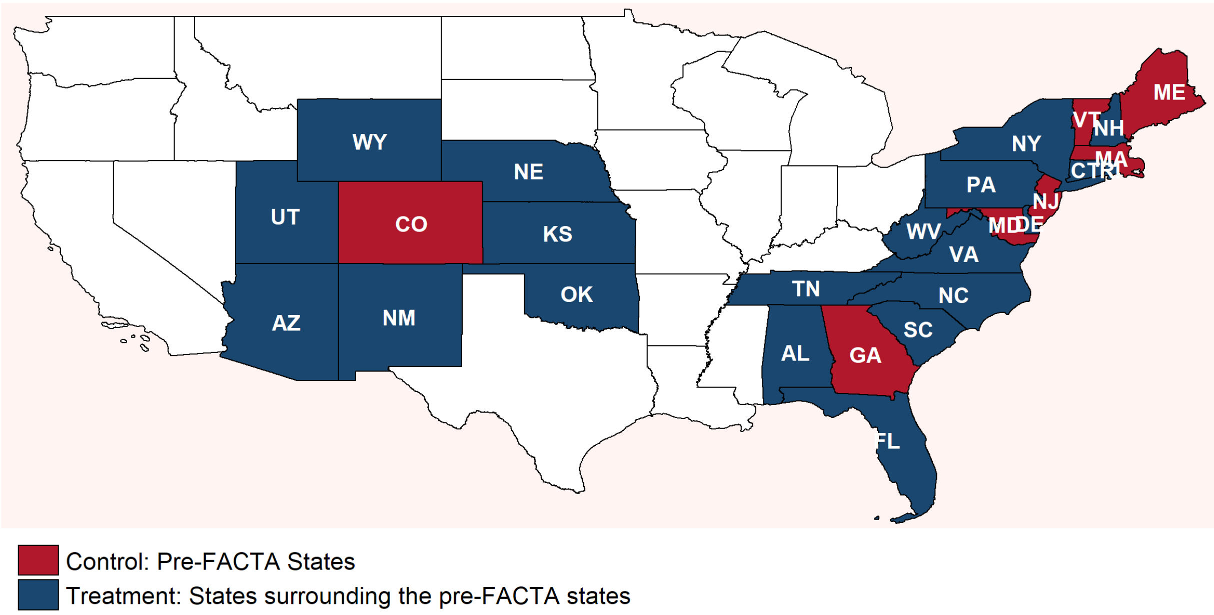


Figure 4:
Empirical Setting: Control and Treatment Counties

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.

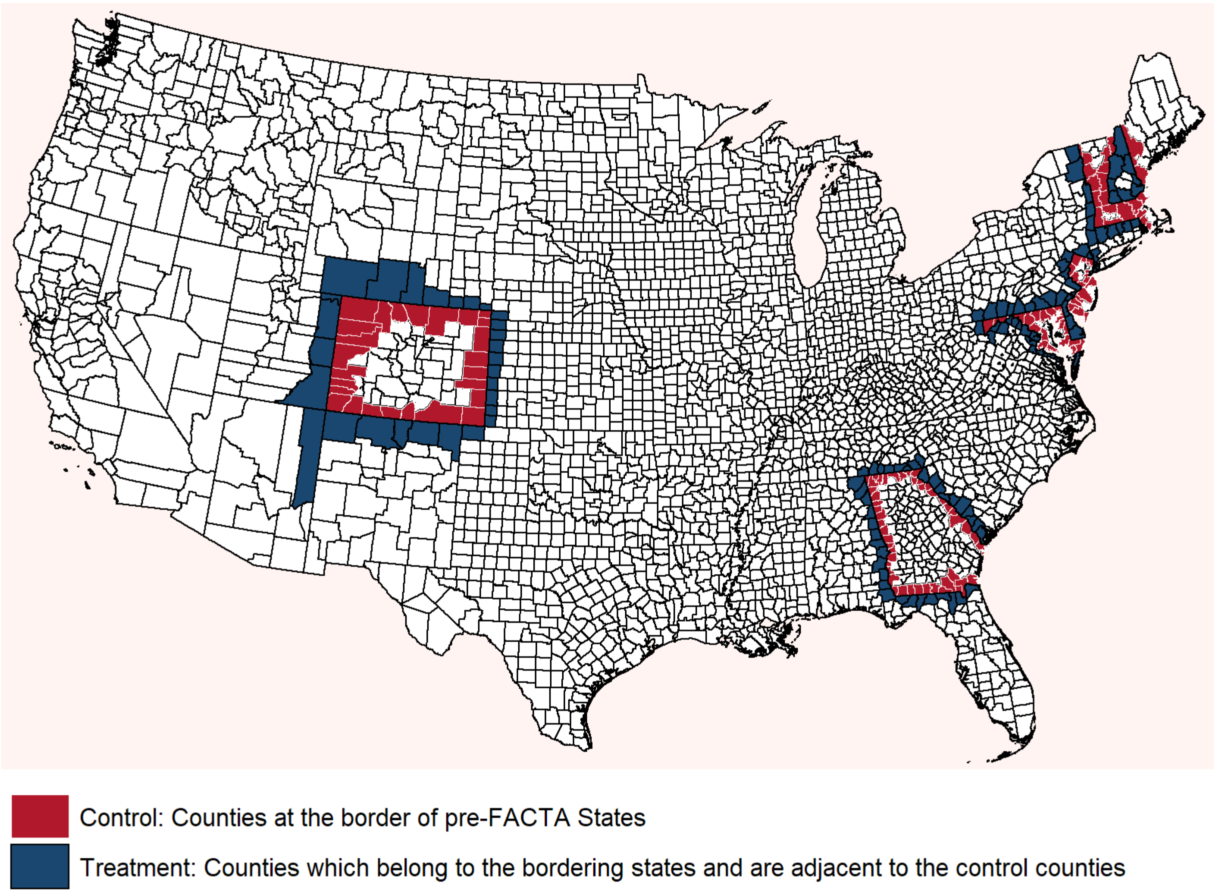


Figure 5: Examining the Parallel Trends

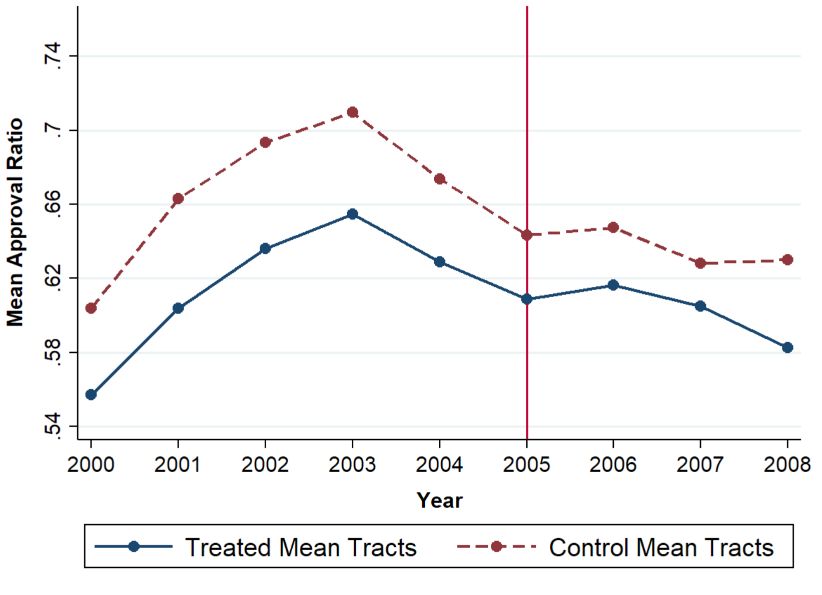
Panel A shows the mean approval ratio in the treated and control census tracts. Panel B shows the coefficients β_k from regressing *Approval Ratio* using the specification:

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

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

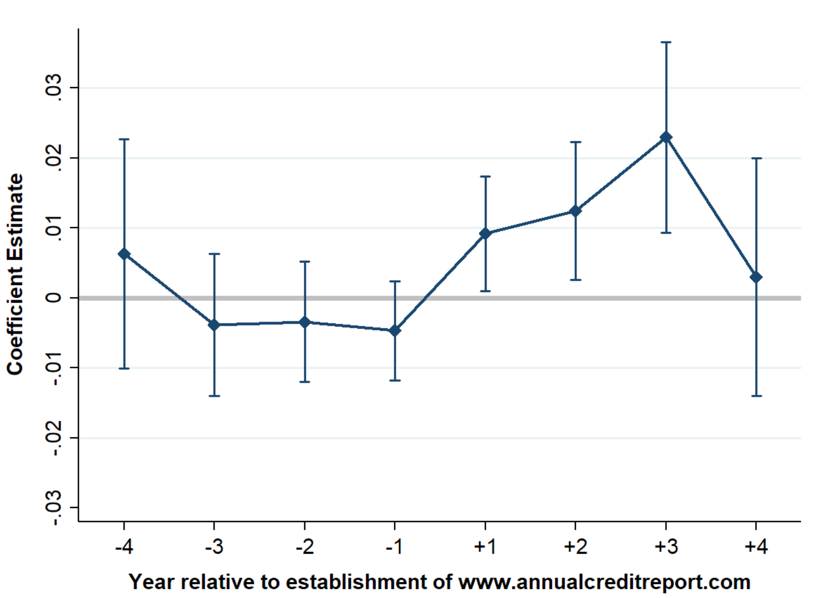


Figure 6:
Google Search Interest for the Term "Free Credit Report" in the U.S.

This figure shows the plot of *Search Interest* for the terms *Free Credit Report* in the US over time, from Jan 1, 2004 till Dec 31, 2011. Numbers on the vertical axis represent search interest relative to the highest point on the chart in the US during the time period of the plot. A value of 100 is the peak popularity for the term. A value of 50 means that the term is half as popular. A score of 0 means there was not enough data for this term.

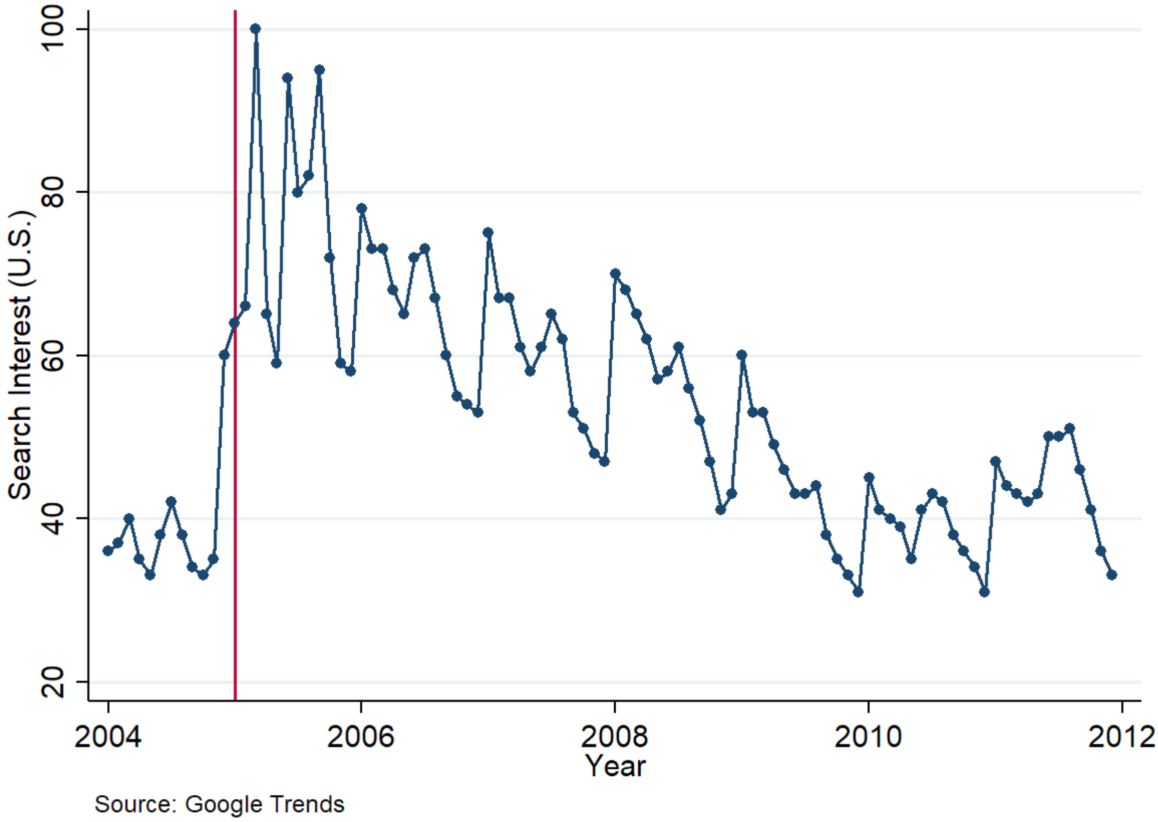


Figure 7:
Popularity of the keyphrase *Free Credit Report*

This figure shows the difference in mean popularity rank of treatment and control states for the keyphrase *Free Credit Report* from 2004 to 2008. The popularity score of each state is ranges from 0 to 100, calculated each 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. Here, I plot the difference in mean annual popularity in treatment and control states.

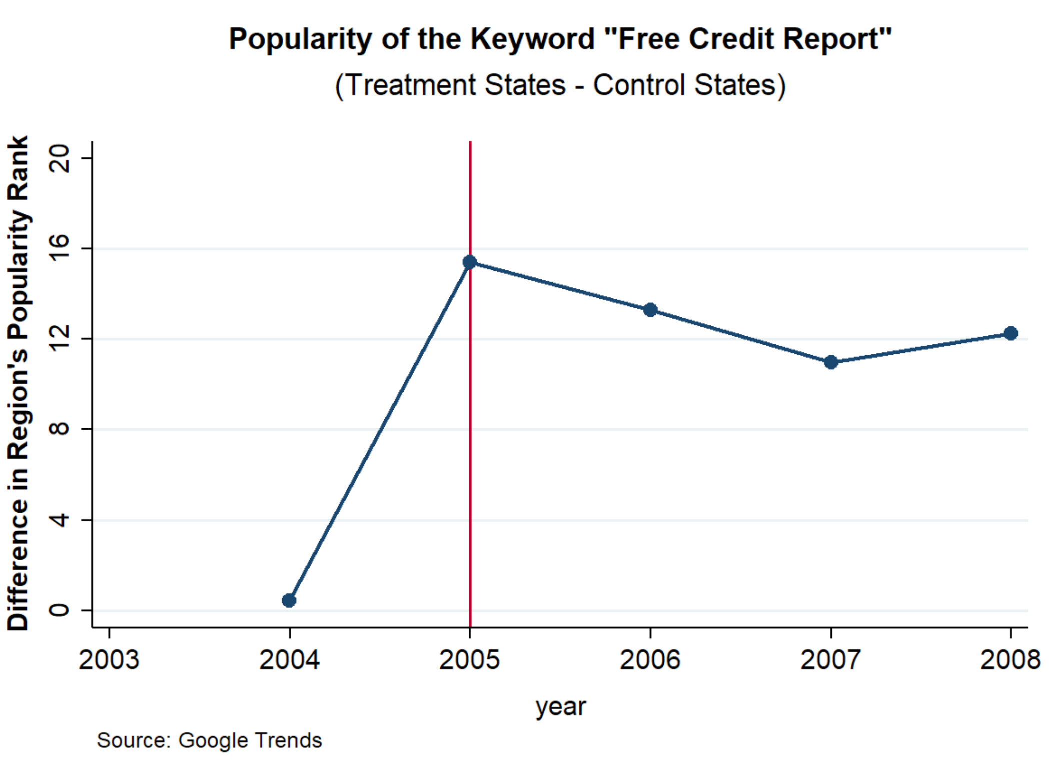


Figure 8:
Effect of Free Credit Report on Mortgage Defaults

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.

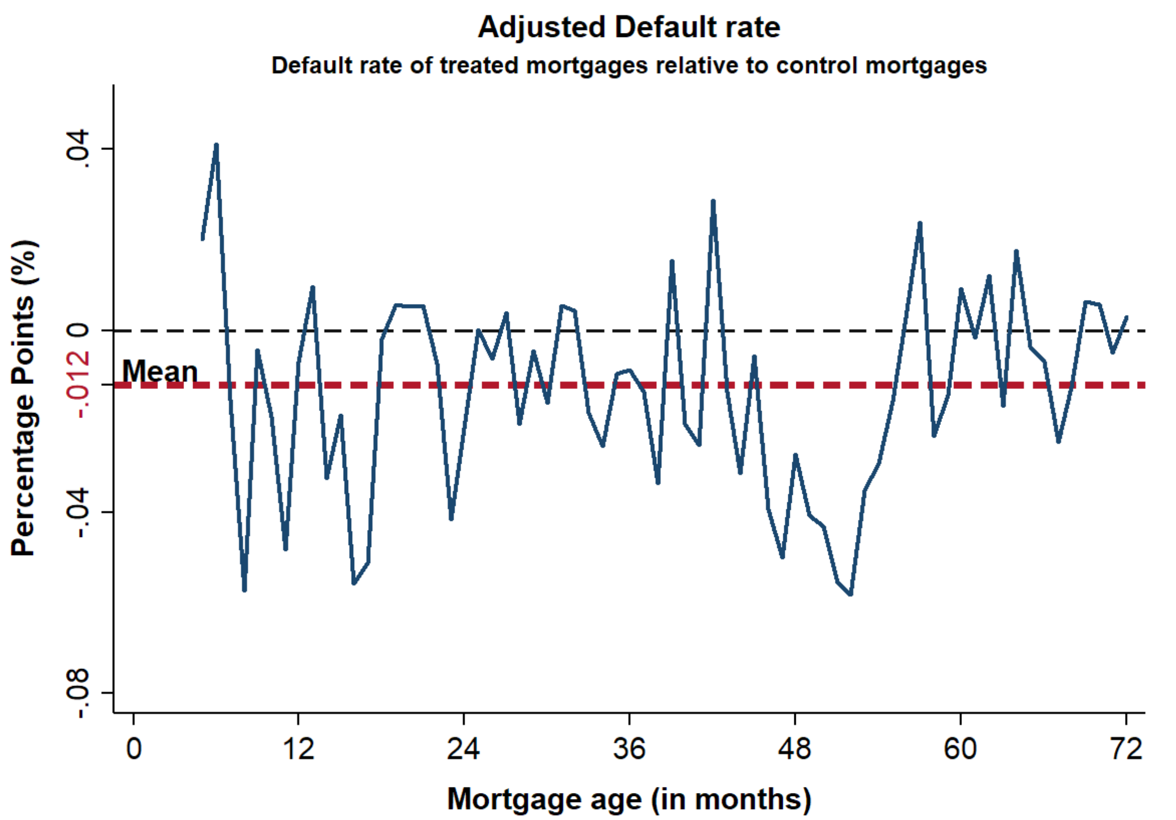


Table 1: Summary Statistics

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 county income per capita. Δ *Emp.* is the annual growth rate of the county employment by all establishments, 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	83.42	74.76	66.41	36494	98.42	77.58	77.74	50323	72.54	70.68	56.48
Approval Ratio	82713	0.63	0.13	0.64	35879	0.65	0.12	0.66	46834	0.61	0.13	0.62
Deny Credit Hist Ratio	82713	0.06	0.04	0.05	35879	0.05	0.04	0.04	46834	0.06	0.05	0.05
Deny Debt-to-inc Ratio	82713	0.03	0.03	0.03	35879	0.03	0.02	0.03	46834	0.03	0.03	0.03
Withdrawn Ratio	82713	0.12	0.05	0.12	35879	0.12	0.04	0.11	46834	0.12	0.06	0.12
Δ Inc per capita	88349	0.04	0.03	0.04	37644	0.04	0.03	0.04	50705	0.04	0.03	0.04
Δ Emp	90231	0.01	0.05	0.01	37635	0.01	0.04	0.01	52596	0.01	0.06	0.01
Δ State GDP	90353	0.04	0.02	0.05	37644	0.05	0.02	0.04	52709	0.04	0.02	0.05

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

	Full Sample				Control Group (C)				Treatment Group (T)				(C-T)
	N	Mean	SD	Med.	N	Mean	SD	Med.	N	Mean	SD	Med.	p-val
Num of App per 1000 adult	48367	110.52	83.57	93.50	20291	129.70	86.15	108.66	28076	96.65	78.79	83.22	0.000
Approval Ratio	47028	0.64	0.13	0.65	20074	0.67	0.12	0.68	26954	0.62	0.13	0.63	0.000
Deny Credit Hist Ratio	47028	0.06	0.04	0.05	20074	0.06	0.04	0.05	26954	0.07	0.05	0.06	0.000
Deny Debt-to-inc Ratio	47028	0.03	0.02	0.03	20074	0.03	0.02	0.03	26954	0.03	0.02	0.03	0.000
Withdrawn Ratio	47028	0.12	0.05	0.11	20074	0.12	0.04	0.11	26954	0.13	0.05	0.12	0.000
Δ Inc per capita	50293	0.04	0.03	0.04	21372	0.04	0.03	0.04	28921	0.04	0.03	0.04	0.000
Δ Emp	51342	0.01	0.06	0.01	21366	0.01	0.04	0.01	29976	0.01	0.07	0.01	0.000
Δ State GDP	51452	0.05	0.02	0.05	21372	0.05	0.02	0.04	30080	0.04	0.02	0.05	0.000

Table 2: Mortgage Applications and the Approval Ratio

This table reports the estimates of the treatment effect of free credit reports on the number of mortgage applications and the approval ratio. 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. *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 the *Border* × *Year* fixed effects (FE) and the *Census Tract* FE. All variables are defined in Table 1. 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)
	# Applications	# Applications	Approval Ratio	Approval Ratio
Treat × Post	13.29***	15.45***	0.01***	0.01***
	(2.94)	(3.65)	(2.73)	(2.97)
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.)	0.807	0.808	0.739	0.735
Observations	86807	84786	82667	80665

Table 3: 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 adultss 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 1. 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)
	# Applications	# Applications	Approval Ratio	Approval Ratio
Treat × Post	9.81***	11.50***	0.01***	0.01***
	(2.83)	(3.48)	(2.89)	(3.04)
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.)	0.640	0.645	0.618	0.613
Observations	86807	84786	84190	82256

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*
	(1.55)	(1.60)	(1.58)	(1.69)
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.)	0.129	0.127	0.105	0.104
Observations	82667	80665	82626	80624

Table 4: Effect of Free Credit Reports on House Prices

This table reports estimates of the treatment effect of free credit reports on the house price and growth in house price. 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 variable is the growth rate of the house price index $((Index_t - Index_{t-1})/Index_{t-1})$. The value of the house price index is 100 in year 2000 for all 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* × *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 1. 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)
	Growth in House Prices	Growth in House Prices
Treat × Post	1.74*	1.82*
	(1.83)	(1.82)
Economic Controls	No	Yes
Census Tract FE	Yes	Yes
Border × Year FE	Yes	Yes
Cluster (County)	Yes	Yes
R^2 (Adj.)	0.683	0.686
Observations	25387	25361

Table 5: 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* × *Post* interaction term captures the change in the fraction of mortgage applications denied due to a given reason in the treated census tracts relative to the control census tracts. All regressions include *Border* × *Year* fixed effects (FE) and *Census Tract* FE. All variables are defined in Table 1. 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
	(-1.50)	(-1.51)	(-2.04)	(-1.88)	(-1.07)	(-0.96)	(-1.48)	(-1.19)
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	82667	80665	39071	38692	82667	80665	39071	38692

Table 6: 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:

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

% 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 1. 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***
	(-2.81)	(-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	82667	80665

Table 7: 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:

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

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* × *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 *Border* × *Quarter* fixed effects (FE) and *Zip3-State* FE. All variables are defined in Table 1. Standard errors are clustered by county. t-statistics are reported below the coefficients in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

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

Table 8: Distinguishing Supply and Demand: Equilibrium Interest Rate

This table reports the estimates of the treatment effect of free credit reports on mean and median interest rate in a zip3-state area. The regression specification is:

$y_{cjt} = \beta_0 + \beta_1 Treatment_c \times Post_T + \delta \times Economic\_controls + \alpha_i + \gamma_{j,t} + \varepsilon_{ct}$. See Eq. (3).

Mean (*Median*) *Int Rate* is the mean (median) interest rate expressed in percentage points, calculated in a 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* × *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 *County* FE. All variables are defined in Table 1. 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 point)			
	(1)	(2)	(3)	(4)
	Mean	Mean	Median	Median
Treat × Post	0.012*	0.011*	0.011*	0.010*
	(1.93)	(1.77)	(1.93)	(1.72)
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.)	0.991	0.991	0.989	0.989
Observations	7711	7711	7711	7711

Table 9: 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 12). *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 1. 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 $\times$ Post	0.002**	0.001	0.003***	0.002*	0.015***	0.009*	0.015***	0.010**
	(2.16)	(1.11)	(2.84)	(1.78)	(2.98)	(1.90)	(3.21)	(2.03)
Difference [High - Low]		-0.001		-0.001		-0.006		-0.006
p-value		(0.577)		(0.581)		(0.501)		(0.475)
Economic Controls	No	No	Yes	Yes	No	No	Yes	Yes
Census Tract FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Border $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster (County)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (Adj.)	0.663	0.595	0.652	0.589	0.750	0.716	0.746	0.712
Observations	60699	25805	59206	25289	57623	24936	56143	24426

Table 10: Heterogeneous Effects by Consumer Creditworthiness (County Measure)

This table reports estimates of the treatment effect of free credit reports on the number of mortgage applications and the approval ratio separately for prime and subprime counties. A county is categorized as subprime if its subprime population fraction is more than the *regional mean* subprime population fraction. 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). *#App.* and *Apv. Ratio* are the number of applications per 1000 adultss 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* $\times$ *Post* interaction term captures the change in the dependent variable in the treated census tracts relative to the control census tracts. All regressions include *Border* $\times$ *Year* fixed effects (FE) and *Census Tract* FE. All variables are defined in Table 1. 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.

	Prime Counties				Subprime Counties			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	# App.	# App.	Apv. Ratio	Apv. Ratio	# App.	# App.	Apv. Ratio	Apv. Ratio
Treat $\times$ Post	16.73**	18.11**	0.02***	0.02***	8.49	10.69**	0.01*	0.01**
	(2.29)	(2.59)	(3.14)	(3.32)	(1.64)	(2.25)	(1.70)	(2.01)
Economic Controls	No	Yes	No	Yes	No	Yes	No	Yes
Census Tract FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Border $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster (County)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (Adj.)	0.802	0.802	0.777	0.773	0.826	0.828	0.681	0.680
Observations	39263	37698	38184	36631	47241	46797	44375	43938

Table 11: Heterogeneous Effects by Consumer Creditworthiness (Census Tract Measure) This table reports estimates of the treatment effect of free credit reports on the number of mortgage applications and the approval ratio separately for census tracts with a low and high density of payday lenders. 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). *#App.* and *Apv. Ratio* are the number of applications per 1000 adultss and the approval ratio in a census tract, respectively. The benchmark for creating the high and low number of payday lenders areas is the average number of payday lenders in census tracts in counties at the border of Colorado (CO) and surrounding states in 2004. *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. All variables are defined in Table 1. 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.

	# Payday Lenders - Low				# Payday Lenders - High			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	# App.	# App.	Apv. Ratio	Apv. Ratio	# App.	# App.	Apv. Ratio	Apv. Ratio
Treat × Post	68.43***	72.78***	0.05***	0.05***	43.72***	43.27***	0.01	0.01
	(5.31)	(5.36)	(3.46)	(3.43)	(3.82)	(4.00)	(0.60)	(0.64)
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	0.751	0.751	0.816	0.818	0.773	0.774
Observations	1452	1452	1395	1395	872	872	865	865

Table 12: Heterogeneous Effects by education 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 separately for the census tracts with a high and low graduate fraction population. The cut-off fraction for defining the high and low graduate fraction population is the *regional mean* graduate population fraction (see Footnote 12). 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). *#App.* and *Apv. Ratio* are the number of applications per 1000 adultss and the approval ratio in a census tract, respectively. Graduate education is defined as *associate degree*, *bachelor's degree*, or *graduate or professional degree*. The graduate fraction population is calculated as the number of adults aged 18 to 64 having graduate degree to the total number of adults aged 18 to 64 in a given census tract, as recorded in Census 2000. *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. All variables are defined in Table 1. 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.

	Low Education Area				High Education Area			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	# App.	# App.	Apv. Ratio	Apv. Ratio	# App.	# App.	Apv. Ratio	Apv. Ratio
Treat × Post	6.03*	6.88**	0.01	0.01*	12.52***	15.04***	0.01***	0.01***
	(1.95)	(2.20)	(1.55)	(1.66)	(2.67)	(3.55)	(2.92)	(3.08)
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.)	0.828	0.831	0.582	0.575	0.824	0.823	0.743	0.739
Observations	35945	35419	33182	32666	50720	49261	49363	47913

Table 13: 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:

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

#App. and *Aprv.* are the number of applications per 1000 adultss 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 1. 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 adultss

	Income Quartile 1		Income quartile 2		Income Quartile 3		Income quartile 4	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	# App.	# App.	# App.	# App.	# App.	# App.	# App.	# App.
Treat × Post	-0.25	-0.41	2.44**	2.67***	3.30***	3.88***	6.10**	7.50***
	(-0.18)	(-0.29)	(2.47)	(2.68)	(2.77)	(3.55)	(2.20)	(3.08)
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.)	0.779	0.780	0.802	0.804	0.795	0.797	0.739	0.739
Observations	88283	86252	88283	86252	88283	86252	88283	86252

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***	0.02***	0.01	0.01	0.01	0.01	0.01*	0.01**
	(2.99)	(3.97)	(1.03)	(1.39)	(1.13)	(1.23)	(1.87)	(2.06)
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.)	0.257	0.259	0.340	0.339	0.399	0.394	0.379	0.373
Observations	79920	77983	81114	79136	81634	79639	81337	79339

Table 14: 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 1. 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	-0.007	-0.006	0.039**	0.042**	-0.004	-0.001
	(-0.54)	(-0.48)	(2.22)	(2.33)	(-0.58)	(-0.12)
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.)	0.003	0.002	-0.000	-0.000	0.032	0.034
Observations	82667	80665	82667	80665	82667	80665

Table 15: 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* × *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 1. 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)
	#Prime App.	#Prime App.	#Subprime App.	#Subprime App.
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 16: Effect of Free Credit Reports on Banks

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

$$Y_{lt} = \beta_0 + \beta_1 Treatment_l \times Post_t + \delta \times Bank\_controls + \alpha_l + \gamma_t + \varepsilon_{lt}. \quad \text{See Eq. (4).}$$

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. All variables are defined in Table 1. 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