Privacy regulation and fintech lending

S. Doerr  L. Gambacorta  L. Guiso  M. Sanchez del Villar

February 2023

Abstract

This paper studies how the California Consumer Privacy Act (CCPA), a comprehensive privacy law that grants users control over their data, affects fintech lending. To develop hypotheses we build a parsimonious screening model. Consumers apply for loans with banks and a fintech that has a superior but data-intensive screening technology. However, consumers dislike sharing their data, and in particular with the fintech. We empirically show that the introduction of the CCPA, by assuaging concerns about data sharing, increases mortgage applications with fintechs relative to banks. Consistent with applicants’ greater willingness to share data, fintechs make greater use of non-traditional data to improve screening. In turn, they deny more applications and can offer lower interest rates.

JEL Codes: G21, G23, G28.

Keywords: Privacy, data sharing, fintech, data regulation, CCPA.

Sebastian Doerr (sebastian.doerr@bis.org) and Leonardo Gambacorta (leonardo.gambacorta@bis.org) are at the Bank for International Settlements, Luigi Guiso (luigi.guiso55@gmail.com) is at the Einaudi Institute for Economics and Finance, and Marina Sanchez del Villar (marina.sanchez@eui.eu) is at the European University Institute. We would like to thank Thorsten Beck, Giacomo Calzolari, Thomas Crossley, Andreas Fuster, Rod Garrat, David Martinez-Miera, Thomas G. Ruchti, Amit Seru, Roberto Steri, and Sergio Vicente, as well as seminar participants at the European University Institute, the HEC Paris conference on Banking in the Age of Challenges Conference, the OCC conference on The Implications of Financial Technology for Banking, and the Bank for International Settlements. We also thank Nikita Divissenko and Maria Magierska. The views in this paper are those of the authors only and do not necessarily reflect those of the Bank for International Settlements.
1 Introduction

Over the last decade, the market share of fintech lenders has increased rapidly in many countries (Buchak et al., 2018; Fuster et al., 2019; Cornelli et al., 2020). A defining feature of these lenders is that they rely heavily on non-traditional personal data to screen and price their borrowers (Berg et al., 2020, 2022).

There is mounting evidence, however, that consumers dislike sharing their personal data (Goldfarb and Tucker, 2012; Tang, 2019). Concerns range from price discrimination to data abuse or unethical advertising (Chen et al., 2021; Lin, 2022; Prince and Wallsten, 2022). Moreover, consumers value with whom they share their data. In a representative sample of US consumers, respondents have greater confidence in traditional banks than in fintechs to safely handle their data and protect them from data abuse or misuse (Armantier et al., 2021).

Regulators see an increasing need to protect consumers and their privacy through privacy protection regulation. In the United States, a key step towards protecting users’ privacy was the introduction of the California Consumer Privacy Act (CCPA) in 2020. The CCPA gives residents the right to control their data, and does so even after they have shared it with a firm. For example, individuals can prevent firms from selling their personal information or request that firms delete the data after use (Camhi and Lyon, 2018). The CCPA hence increases transparency and accountability, and reduces consumers’ concerns about the ab- and misuse of personal data for unintended purposes, a concern that is more salient vis-a-vis fintechs (Armantier et al., 2021; Chen et al., 2021).¹ Numerous other states and countries consider introducing legislation in the spirit of the

¹Banks are already subject to a variety of regulations that in some jurisdictions also include data sharing agreements, which could partly explain consumers’ higher trust in banks to safely store their data. The marginal benefit of privacy regulation is hence expected to be higher among newer intermediaries such as fintechs, compared to traditional financial institutions.
The CCPA differs from other data initiatives and regulations in important respects. For example, a key principle of the European Union’s General Data Protection Regulation (GDPR) is that firms need to minimize their data processing activities. Similar to policies such as bankruptcy flag removal (Liberman et al., 2019; Jansen et al., 2022), it hence effectively limits firms’ ability to collect information (Johnson, 2022). Open banking, on the other hand, mandates incumbent institutions to share their proprietary data with third parties, including fintech lenders, if users give their consent. The need to share data can discourage intermediaries from acquiring and storing data in the first place (Babina et al., 2022) and can even make borrowers worse off (He et al., 2023).

This paper investigates how the interplay between privacy protection legislation and privacy preferences shapes loan markets. To derive hypotheses, we first build a parsimonious screening model in which banks and a fintech request data from potential borrowers before offering a loan. Lenders differ in their data-intensity and borrowers have a preference for privacy. We then exploit the introduction of the CCPA to test the model’s predictions with loan-level data in the U.S. mortgage market. For identification, we compare census tracts at the border of California to those in neighboring states and employ granular time-varying fixed effects.

In our model, two types of lenders – banks and a fintech – compete for privacy-sensitive borrowers. Borrowers are evenly distributed along a line of unit length and differ in their creditworthiness. The two banks are located at the respective end points of the line, while the fintech is located at an equal distance to all consumers. This setup captures that fintechs offer online platforms available to all consumers with greater convenience and speed over banks, two key drivers behind their rapid growth (Buchak et al., 2018;
Before offering a loan, lenders must analyze personal data to screen borrowers. More data improves the screening process for both types of lenders, but fintechs are better at extracting a precise signal from a given amount of data provided by an applicant (as in He et al. (2023)). More data, ie a better signal, in turn improves lenders’ pool of borrowers and allows them to offer a lower interest rate.

Applying for a loan entails a cost for consumers that consists of two parts. The first part is proportional to the distance between the location of the applicant and the location of the lender. The second part reflects applicants’ dislike for sharing their personal data and is a key feature of our model. Specifically, the more data applicants have to provide, the lower their utility from obtaining a loan at a given rate. Motivated by recent evidence, we assume that the disutility from sharing data is stronger when sharing data with the fintech.

In equilibrium, borrowers sort into different lenders depending on the distance and their privacy preferences. Borrowers closer to the banks will prefer to contract with them, while borrowers located near the midpoint of the line prefer the fintech. The position of the consumer indifferent between a bank and the fintech is determined by her relative distance and the applicant’s privacy preference that determines the amount of data shared and thereby the interest rate.

We use the model to derive predictions on the effects of privacy protection legislation in the spirit of the CCPA. Specifically, we analyze regulation that increases consumers’ control over their data and assuages concerns regarding data abuse, which disproportionately benefits the fintech. The first prediction is that loan applications with the fintech increase, as some consumers that initially applied to banks now apply with the fintech. The second prediction is that the fintech offers loans at a lower interest rate. The reason

---

3For banks, these costs capture eg the need for consumers to go to a branch and interact with a loan officer, or the lower speed and convenience of filing applications compared to fintechs.
is that applicants now face a lower disutility from disclosing personal data, so the fintech asks for more data to better screen out low-quality borrowers and faces a better pool of borrowers. Two additional implications are hence that the fintech uses more personal data in its decision making process, and that more precise screening increases the share of denied loan applications for the fintech.

We empirically tests these hypotheses in the U.S. mortgage market by exploiting the introduction of the CCPA in 2020. Our analysis is based on Home Mortgage Disclosure Act (HMDA) data from 2018–2021 that provides a wealth of information on lenders, borrowers, and loan terms. We classify lenders into banks and fintechs following Fuster et al. (2019). For identification, we estimate regressions at the lender–borrower tract–year level and compare loan applications with both lender types in counties on the border of California with its neighboring states.

We first investigate the effects of the introduction of the CCPA on loan applications and loan rates. We find that after the introduction of the CCPA, loan applications with fintechs compared to banks increase significantly in California, relative to border counties in neighboring states. In terms of magnitude, loan applications with fintechs in California increase by 13.9% after the introduction of the CCPA. Fintechs’ relative market share increases by 2.9 percentage points (pp). In 2019 (before CCPA came into effect), the share of loan applications with fintechs averaged 14.1%, implying an increase of 24% of the mean. We also find that the CCPA reduces loan rates by relatively more on mortgages originated by fintechs in California.

Second, we provide empirical support for the two additional implications of the model. After the CCPA was introduced, the share of fintech mortgages that did not use standardized underwriting models increases significantly. As argued in Babina et al. (2022), the use of non-standardized underwriting models reflects the use of non-traditional data
beyond standardized credit scores. Consistent with obtaining a more precise signal to screen borrowers from more data, we further find that the share of denied loan applications increases for fintechs compared to banks after the introduction of the CCPA.

Our analysis faces the common identification challenge that any observed change in loan applications or rates could be due to unobservable factors at the lender or borrower level. For example, fintechs could serve tracts with higher income growth over the sample period, leading to an increase in applications. Likewise, a decision by banks’ management to reduce their lending business in California for reasons unrelated to the CCPA could affect the estimated coefficients.

We address this challenge in different ways. For one, we restrict our analysis to the set of tracts within counties that lie on the border of California with its neighboring states. As has been shown in a large literature, border counties are generally similar along many observable characteristics, mitigating concerns about selection effects and omitted variable bias (Allegretto et al., 2017). Indeed, in the border tract sample we find no discernible difference in the evolution of loan applications or loan rates with fintechs compared to banks prior to the introduction of the CCPA in 2020. In addition, we show that the average fintech and bank borrower have comparable observable characteristics in border tracts.

We further include granular time-varying fixed effects. Borrower tract*time fixed effects absorb any observable and unobservable differences in tract characteristics over time. These fixed effects account for eg differences in borrower income, demographic structure, or credit demand. They also control for potential differences in the severity of the Covid-19 pandemic and associated movement restrictions across tracts. In essence, we exploit

---

4Babina et al. (2022) link the use of non-standardized underwriting models to more individualized loan pricing – in line with models that show that a more precise signal leads to greater dispersion in interest rates (Jansen et al., 2022). Consistent with these findings, we also show that the dispersion in interest rates increases by relatively more for fintechs.

5We also show that there were no significant differences in the number of cases per capita and death
only within-borrower tract variation and compare applications to different lenders from borrowers in the same tract and the same year. In addition, we include lender*time fixed effects to control for time-varying observed and unobserved heterogeneity in lender characteristics, for example size, risk-taking, or funding models. With the full set of fixed effects, we are comparing how mortgage applications to the same lender change with the introduction of the CCPA in border tracts in California compared to comparable tracts in neighboring states.

Across specifications, we find that the introduction of the CCPA has led to an economically and statistically significant increase (decrease) in loan applications (loan rates) with fintechs compared to banks. The fact that our results are robust to the inclusion of granular time-varying fixed effects, together with the absence of any differential pre-trends in loan applications and rates, mitigates concerns that the CCPA was introduced because of the rise of fintech lenders in Californian border counties. To further ensure that our findings are not driven by movements in the control group (ie states other than California), we also show that our findings hold when we compare fintechs to banks in California only.

Our results are robust to a wide range of alternative specifications. They are present among purchase as well as refinance mortgages; are unaffected when we exclude borrowers of age 62 and above from the sample, ie borrowers that could have been more affected by Covid-19 related restrictions; and remain robust to the inclusion of a large set of lender-tract level controls. Finally, we obtain similar results when we restrict the sample to mortgages that were sold in the respective calendar year. The latter exercise addresses the concern that banks are more likely to make on-balance sheet loans that are funded by deposits and could differ in (unobservable) characteristics from fintech loans.
Our results have implications for the policy debate on how to regulate the use of personal data. Personal data lie at the heart of the digital economy. By allowing lenders to better assess the riskiness of borrowers (Berg et al., 2022), the use of data can reduce the need for collateral (Gambacorta et al., 2022) and promote financial inclusion (Philippon, 2020). At the same time, consumers value their privacy and are concerned about the abandonment and misuse of data (Armantier et al., 2021; Tang, 2019; Lin, 2022). These considerations pose a trade-off for privacy policy that needs to balance improving efficiency and protecting users’ right to privacy (Acquisti et al., 2016). Our results suggest that privacy protection legislation that enhances users’ control over data can mitigate this trade-off: by increasing transparency and accountability in the use of data, the CCPA makes users more willing to share data, enabling technology-savvy lenders to better screen with data and offer lower rates.\footnote{These results are also highlighted by Ali et al. (2022), who focus on the effects of voluntary disclosure (users’ control) on consumer welfare. They show that when consumers can choose whether and to whom disclose information, competition between firms can increase, hence improving consumer welfare.}

**Literature and contribution.** Our paper speaks to the literature that studies the consequence of data sharing policies for loan markets.

Several papers focus on policies that restrict information sharing. Dobbie et al. (2020) show that bankruptcy flag removal leads to economically large increases in affected borrowers’ credit limits and borrowing. However, by limiting lenders’ ability to process information, bankruptcy flag removal can have large distributional effects across borrowers, and create both winners and losers with ambiguous welfare effects (Liberman et al., 2019; Jansen et al., 2022). Similarly, the US Card Act, which limited credit card lenders’ to adjust interest rates on the basis of new information, reduced prices for high-risk consumers but increased them for others (Nelson, 2018; Jansen et al., 2022) Europe’s GDPR, which has the key principle that firms need to minimize their data processing...
activities, has decreased venture investments in data-related firms (Jia et al., 2018). Privacy-conscious consumers have used the GDPR to opt for reporting less data, thereby creating externalities for the remaining consumers and losses for intermediaries (Aridor et al., 2022).

Recent work studies the consequences of open banking, an initiative launched by several governments including the European Union and the United Kingdom. It allows customers to share their bank account history with third parties, e.g., fintech lenders. Improving fintechs’ access to previously unavailable data can result in better screening and loan market outcomes. For example, Nam (2022) shows that riskier borrowers share data through open banking more readily with a fintech. Subsequently the probability of loan approval increases and interest rates fall. Yet other work highlights that open banking can also have unintended negative consequences. If fintechs have a sufficiently superior screening technology, open banking could allow them to achieve market power beyond that of banks before open banking (He et al., 2023). Borrower welfare would then be lower than without open banking (He et al., 2023). Open banking can further hamper the efficient allocation of credit, as banks may endogenously adjust their liabilities once data become open to challengers (Goldstein et al., 2022). Babina et al. (2022) empirically show that while open banking can spur fintech venture capital investments and innovation, it can discourage intermediaries’ data production, as they reap fewer benefits from collecting data.

Our paper studies the effects of the CCPA, i.e., privacy regulation that grants users control over their data and mitigates privacy concerns. Our setting hence differs from studies on policies that limit lenders’ information set or force banks to share information with competitors. We find that privacy protection legislation in the spirit of the CCPA

\footnote{A growing literature investigates the effects of the GDPR on firms and finds that it hurts firm performance (see Johnson (2022) for a survey).}
can make borrowers more willing to share their data and, in the presence of lenders with superior screening technology, lead to lower rates and improved loan market outcomes—especially for privacy-sensitive borrowers.

We also relate to papers studying the rise of fintechs and the attendant effects on banks. An increased regulatory burden on traditional banks and fintechs’ superior technology with faster processing times are important drivers underlying their growth in the US mortgage market (Buchak et al., 2018; Fuster et al., 2019). In addition, better access to payments data can affect the competition between banks and fintechs. Ghosh et al. (2021) find that fintech lenders can use payments data to obtain information about potential borrowers that compensates for the lack of an existing lending relationship. Parlour et al. (2022) theoretically show that fintechs competing for payments can disrupt information spillovers from banks’ payments to their lending services, with ambiguous results for welfare. Most of the literature studying the rise of fintechs and their competition with banks have not explicitly considered the role of privacy preferences or data protection legislation. We find that privacy protection legislation can spur the growth of fintechs, which could increase the competitive pressure for incumbent banks.

The remainder of the paper is organized as follows. Section 2 discusses evidence on borrowers’ preference for privacy and provides institutional background on the CCPA. Section 3 provides a theoretical framework to derive predictions on the effects of privacy regulation on banks and fintechs. Section 4 tests the predictions with US mortgage data, exploiting the introduction of the CCPA. Section 5 concludes.

---

8For unsecured personal loans, Di Maggio et al. (2022) show that fintechs initially target riskier borrowers but eventually compete with banks for higher-quality borrowers. Other papers look at data availability and firm dynamics (Begenau et al., 2018; Farboodi et al., 2019; Canayaz et al., 2022).
2 Evidence on privacy preferences and the CCPA

This section first discusses recent evidence on consumers’ privacy preferences and then provides institutional background on the California Consumer Privacy Act.

Privacy preferences. As more and more economic activity is moving online, personal data is turning into an increasingly important input for firms (Acquisti et al., 2016; Jones and Tonetti, 2020). Data, for example, allow firms to ascertain otherwise hidden characteristics of their potential buyers, which can give rise to personalized pricing (Rhodes and Zhou, 2021), by which lenders charge each consumer the maximum acceptable price, leaving consumers worse off.

The avoidance of personalized pricing (or first-degree price discrimination) is one of the reasons why consumers might want to keep their data private. Other important reasons mentioned by survey respondents when asked about their concerns when sharing data include identity theft, personal safety and reputation (Armantier et al., 2021). Indeed, in a representative survey of US households, around three-quarters of respondents were very concerned about negative consequences when sharing their personal data (see Figure 1).

Beyond survey responses, there is mounting empirical evidence that consumers value keeping some data private (Goldfarb and Tucker, 2012). For example, Tang (2019) finds that, when applying for a loan to a fintech, consumers derive utility from withholding information. Lin (2022) further shows that a preference for privacy plays an important role for users’ decision to share data.

Alternative data, however, are an essential input in fintech lending, and in particular when it comes to screening applicants (Berg et al., 2022). Such data can be directly collected, eg through loan applications, but also purchased from data aggregators and vendors. It can include information from social network activity, online footprints,
Figure 1: Consumers are concerned about sharing their data

This figure shows concerns about sharing data online, based on a representative sample of 1,361 US households in September 2020 that were part of the Survey of Consumer Expectations of the Federal Reserve Bank of New York. Respondents are ‘very concerned’ about sharing their data online when they assigned a score of 5 or higher to the question “Are you concerned that sharing your personal data could have negative consequences for you?”, on a scale from 1 (not at all concerned) to 7 (extremely concerned). Regarding specific concerns, the numbers provided denote the share of respondents that answered yes to the question “What are you specifically concerned about if your personal data were to become publicly available?”, where specific concerns are identity theft, data abuse, personal safety, and personal reputation. Source: Armantier et al. (2021) and Carstens et al. (2021).

shopping habits. For example, Jagtiani and Lemieux (2019) find that Lending Club is increasingly using alternative data.\(^9\) Berg et al. (2020) show that default predictions can be improved by using the digital footprint of credit applicants (e.g., the type of device used to access a website or the time of the connection).

Since firms and financial intermediaries use consumers’ data, an important question is in which counterparties consumers have greater confidence to safely handle their data and prevent from data abuse or misuse. Figure 2 shows the level of trust expressed by a sample of US households for traditional financial intermediaries (FI), fintechs, and bigtechs. A clear pattern emerges: almost two-thirds of respondents state that they have high trust

\(^9\) They conclude that Lending Club is increasingly using non-traditional data because the correlation between their proprietary ratings and the FICO score has decreased from 80% to 35% over time.
in traditional FIs, as opposed to only 35% placing high trust in fintechs and 10% in bigtechs. For a large sample of countries, Chen et al. (2021) report a similar pattern: survey respondents are significantly more willing to share their data with traditional FIs than with fintechs.

Taken together, the empirical evidence suggests that consumers have a preference for keeping their personal data private; and they have higher confidence in traditional banks, relative to fintechs, to safeguard their personal data. We will use these insights to motivate our theoretical framework.

The California Consumer Privacy Act. The CCPA is a data privacy law covering the state of California that went into effect in the beginning of 2020. It endows Californians with several rights regarding the personal information that a firm may collect about
them. In particular, Californians have the right to know what personal information is being collected, and whether it is being sold and to whom. They also have the right to access their personal information, delete it, and to opt-out of the sale of their personal data (Camhi and Lyon, 2018).

By granting consumers control over the use of their data, the CCPA directly addresses several of the concerns that individuals list when it comes to sharing their data. As shown in Figure 1, these concerns include identity theft or data abuse. A consumer concerned with these issues can request under the CCPA that her data not be sold or be deleted after transacting with a firm. Therefore, the CCPA increases the certainty around the use of personal data.

The CCPA is expected to have a stronger impact on borrowers’ attitudes towards sharing data with fintechs compared to traditional financial intermediaries. Users have significantly lower confidence in fintechs than in banks to safely store their data and prevent data abuse (see Figure 2). This pattern could arise from the fact that banks are already subject to a variety of regulations that in some jurisdictions also include data sharing agreements. The marginal benefit of privacy regulation is hence likely higher among newer intermediaries such as fintechs, compared to traditional financial institutions.

The CCPA differs from other data initiatives and regulations in important respects. For example, a key principle of GDPR is that firms need to minimize their data processing activities. Similar to policies such as bankruptcy flag removal, it hence effectively limits firms’ ability to collect information (Jansen et al., 2022; Johnson, 2022). Open banking, on the other hand, mandates incumbent institutions to share their proprietary data with third parties, including fintech lenders, if users give their consent. The need to share data can discourage intermediaries from acquiring and storing data in the first place (Babina et al., 2022) and can even make borrowers worse off (He et al., 2023).
Were Californian residents aware of the introduction of the CCPA? According to a 2021 survey of 1,507 adults in California, around 70% answered they saw the notice of their rights required by the CCPA on websites they visited. Moreover, the majority of Californians have already exercised their rights granted by the CCPA. For example, they have asked firms to not share or delete their data.\footnote{See the Consumer Action and Consumer Federation of America: California Consumer Privacy Act (CCPA) Survey.} Consistent with these survey results, mortgage lenders often provide didactic and simple explanations of what the law entails for California residents when they apply for a mortgage.\footnote{See Figure OA2 and Figure OA3. Appendix Section A.2.3 provides more information on the CCPA and an enforcement example.} In the Online Appendix we further show that Google searches for the CCPA in California increased steeply in late 2019 and remained elevated for most of 2020 (see Figure OA1). Overall, the CCPA gives consumers control over their data and provides them with greater certainty that their data will not be used for unintended purposes. It hence has made consumers more willing to share data with lenders (Armantier et al., 2023).

3 Theoretical framework

This section develops the theoretical framework guiding the empirical analysis. In particular, we generate predictions about loan applications and interest rates in a loan market following the introduction of privacy protection regulation.

3.1 Setup

We consider a competitive loan market where potential borrowers are continuously and uniformly distributed along a line of unit length. Banks and a fintech request data to screen applicants before offering a loan. Fintechs are better at extracting precise
information from personal information but applicants dislike sharing their data with the fintech.

**Applicants.** Applicants are risk-neutral and penniless, but are endowed with a risky investment opportunity that requires one unit of funding. An applicant could be a firm looking for a loan to fund a project, or an individual looking for a mortgage to buy a house. Half of the individuals are successful in their investment and get a return $Y$ that allows them to repay the loan. The other half is unsuccessful and gets a return of zero, so they cannot repay the loan.\(^{12}\) Applicants are endowed with one unit of personal data, but do not possess the technology to infer their type from their data.\(^{13}\) Moreover, applicants are protected by limited liability, and their outside option is normalized to zero.

Applicants dislike sharing their data with lenders. In line with the evidence in Figure 2, their disutility from sharing data is higher when they share their data with a fintech than with a traditional bank. We capture this stylized fact by assuming that an applicants’ sensitivity to sharing data with a bank is $s = 1$, while it is $s_F > s$ when sharing data with a fintech. Therefore, $s_F$ denotes the relative sensitivity of providing data to the fintech vs the bank.

**Lenders.** Two symmetric banks are located at the extremes of the line, with bank 1 ($B1$) at the beginning of the line and bank 2 ($B2$) at the end. The distance $x$ between a consumer on the line and the banks can be interpreted as a convenience difference: Each applicant regards the lenders as providing services with different convenience levels, and not all applicants agree on their preferred lender (Thisse and Vives, 1988).

The fintech ($F$) is located at the same distance $x_F$ to all consumers, irrespective of

---

\(^{12}\)We assume that $Y$ is large enough so that everyone prefers to apply for a loan.  
\(^{13}\)This assumption implies that we abstract from issues of adverse selection in our model. We do so to better isolate the effects of screening and preferences over data sharing in our results.
their location on the line. This assumption captures the idea that fintechs offer online platforms with higher convenience and speed over banks to all consumers, two key drivers behind their rapid growth (Buchak et al., 2018; Berg et al., 2022).\footnote{Chu and Wei (2021) and Vives and Ye (2022) make the same assumption in a Salop model, where banks are located around the circumference and the fintech at the center of the circle.} We normalize the fintech’s distance to zero ($x_F = 0$) for simplicity. The analysis would remain qualitatively similar with a positive, sufficiently small distance $x_F > 0$.

Lenders have access to a screening technology that returns either a good signal ($\eta_g$) or a bad signal ($\eta_b$). The accuracy of the signal depends on two elements: the technology that each lender $j$ has ($\gamma_j \in (0, 1]$) and the data each lender collects on the borrower ($d_j \in [0, 1]$). Similar to He et al. (2023), we assume that the fintech is better at extracting information from a given amount of data, so $\gamma_F > \gamma_B$, and we normalize its accuracy to one ($\gamma_F = 1$). Moreover, the screening technology’s accuracy is concave in the amount of data: When there is little data available, more data considerably improves the signal accuracy. When lenders already have a considerable amount of data, an additional unit does not increase accuracy by much.\footnote{Berg et al. (2020) show that even simple, easily accessible data that proxy for income, character, and reputation are highly valuable for default prediction.}

The signal has a bad-news flavour: Bad signals perfectly identify applicants that will not repay. Good signals can come from both types of applicants. The signal structure is:

$$\Pr(\eta_{j,g}|\text{repay}) = 1, \quad \Pr(\eta_{j,b}|\text{no repay}) = \gamma_j \sqrt{d_j}.$$  

Both types of lenders face a perfectly elastic supply of funds at the risk free rate and derive revenue from the interest rate they charge on their loans. Lenders offer the same rate to all consumers, irrespective of their position in the line. While the convenience cost enters applicants’ disutility, lenders do not explicitly consider it when setting their interest rates. Convenience is not verifiable, so lenders would not have a way to distinguish applicants.
with low convenience match from applicants with high convenience match. Alternatively, convenience-pricing is banned by regulation: interest rates can vary with individual loan characteristics (income, LTV), but varying a rate because a bank is considered more or less convenient is not permitted. Finally, more data allow the lenders to filter out, among all applicants, a higher proportion of those that would not repay.

**Timing.** The timing of the game is as follows: First, lenders simultaneously choose the amount of data they request from applicants. Observing each others’ data choices, the lenders simultaneously choose their interest rates. The final contract, which consists of a combination of \((R, d)\), is contingent on the applicant qualifying for a loan (returning a good signal). If the applicant does not qualify, the lenders withdraw the offer. Applicants then observe the data requested and the interest rate offered by each lender. Considering their position on the line (ie their relative distance to each lender), they apply to the lender with the offer that maximizes their expected utility.\(^{16}\) The game concludes as follows: Lenders receive applicants’ data and process them to extract the signal. They then extend credit to the applicants that returned a good signal \((\eta_g)\) and disqualify the applicants that returned a bad signal.

### 3.2 Equilibrium

We now describe the equilibrium. Appendix A.1 provides detailed derivations. We proceed backwards. Observing lenders’ offers, an applicant located at position \(x \in [0, 1]\) has three choices: either apply to bank 1, which is at a distance \(x\); apply to bank 2 at distance \(1 - x\), or apply to the fintech at distance \(x_F = 0\). The expected utility from applying reflects the offered interest rate (which depends on the data requested), the disutility

\(^{16}\)Price commitments under screening are as in Kim and Wagman (2015); Burke et al. (2012). Contracts are exclusive, so borrowers cannot apply to two lenders at the same time.
of sharing data (which is greater for the fintech), as well as the distance to the lender (which is greater for a bank). The applicant indifferent between a bank and a fintech is located where the expected utility from applying to a bank equals that of applying to the fintech. All else equal, applicants sufficiently close to the endpoints of the line contract with the traditional bank. The reason is that for those applicants the disutility from distance is small. The fintech receives applications from borrowers that are further away from both banks, ie near the midpoint of the line. While applying to the fintech entails greater disutility from sharing a given amount of data, this disutility is offset by the lower convenience cost.

Lenders’ demand for data is a function of each lender’s signal accuracy and the sensitivity of the applicants vis-a-vis that lender. In essence, when asking for data, lenders face a trade-off. More data implies greater signal accuracy to screen out the credit-unworthy applicants, and hence lenders can offer a lower interest rate to qualifying applicants. At the same time, asking for more data lowers demand, as sharing data is costly for applicants.\textsuperscript{17} This trade-off is more pronounced for fintechs, as they have a better screening technology but applicants have a greater dislike for sharing data with fintechs than with banks.

In equilibrium, the demand for the fintech is hence determined by the relative sensitivity of providing data to the fintech vs the bank ($s_F$) and the relative advantage in the screening technology ($\gamma_B$):

$$D_F^* = \frac{1 - \gamma_B^2 s_F + 16 s_F}{24 s_F} > 0.$$  

\textsuperscript{17}As both banks are identical, we will focus on symmetric solutions to the banks’ optimal choices, such that $d_1 = d_2 = d_B$ and $R_1 = R_2 = R_B$.  

18
The equilibrium interest rates for the bank and the fintech are

\[
r_B^* = \frac{s_F (56 - 5\gamma_B^2) - 1}{24s_F}, \quad r_F^* = \frac{s_F (64 - \gamma_B^2) - 5}{24s_F}.
\]

### 3.3 Introducing privacy protection legislation

We can perform comparative statics exercises to understand how the introduction of privacy legislation affects the loan market. We consider privacy legislation in the spirit of the CCPA, i.e., legislation that provides consumers with greater control over their data and reduces concerns about the abuse and misuse of data. Such legislation decreases borrowers’ sensitivity to sharing their data. Moreover, we assume that this decrease is greater for sharing data with fintechs compared to sharing them with banks, consistent with the evidence that individuals generally have lower confidence in fintechs to safely handle personal data to begin with.

In equilibrium, the introduction of a privacy protection regulation decreases borrowers’ sensitivity to sharing data with the fintech \(s_F\). It hence increases the relative demand for the fintech

\[
-\frac{\partial D_F}{\partial s_F} = \frac{1}{24s_F^2},
\]

while decreasing the interest rates that the fintech charges relative to banks

\[
\left(-\frac{\partial r_F^*}{\partial s_F}\right) - \left(-\frac{\partial r_B^*}{\partial s_F}\right) = \left(-\frac{5}{24s_F^2}\right) - \left(-\frac{1}{24s_F^2}\right) < 0.
\]

The mechanism behind these results is as follows: as borrowers become less sensitive to sharing data with the fintech, the fintech asks for more data. It does so because more data allows it to screen with a higher accuracy, and hence improves the quality of the accepted applicants. A higher average quality of the borrower pool allows the fintech to
offer lower interest rates. Although borrowers dislike that they now have to provide more
data, not only are they less sensitive to sharing data, but they are also being offered a
lower interest rate. Therefore, more borrowers apply to the fintech.

In sum, the model generates the following predictions.

**Prediction 1**: The introduction of privacy protection legislation leads to an increase in
loan applications with fintechs, compared to banks.

**Prediction 2**: The introduction of privacy protection regulation leads to a decrease in
loan rates on loans originated by fintechs, compared to loans originated by banks.

Moreover, the model yields the following two implications:¹⁸

**Implication 1**: The introduction of privacy protection legislation implies that the fintech will ask for relatively more personal data than banks.

**Implication 2**: The introduction of privacy protection regulation leads to a greater share of denied loan applications by the fintech, relative to banks.

In what follows, we will empirically investigate the effects of privacy protection legis-
lation in the US mortgage market.

## 4 The CCPA and fintech lenders in the US mortgage market

This section exploits the introduction of the California Consumer Privacy Act in 2020
to test the key predictions of the theoretical framework. We first investigate how the
CCPA affects mortgage applications and loan rates on mortgages offered by fintechs and
traditional banks. We then analyze the effects of data protection legislation on fintechs’

¹⁸Appendix A.1.3 provides further details.
use of alternative data and application denial rates.

4.1 Data and summary statistics

HMDA provides home mortgage application data, covering the vast majority of applications and approved mortgages in the U.S. The yearly data include the application outcome, loan amount, and, for granted loans, the interest rate. Additionally, they contain detailed information on applicant income, race, gender, age, and ethnicity, among other items. To classify lenders in HMDA as either banks or fintechs, we follow Fuster et al. (2019).

We aggregate the data to the lender–borrower tract–year level. In our analysis, we will use two samples. The ‘full sample’ contains all mortgage applications in California (CA) and its neighboring states Arizona (AZ), Nevada (NV), and Oregon (OR). The ‘border sample’ only contains mortgage applications in counties that lie along both sides of the California border. The sample period covers the years from 2018, the first year for which data on interest rates is available, to 2021.

The main outcome variables are the log of the number of applications and the interest rate on approved mortgages. In addition, we compute the share of denied applications, as well as the share of mortgages that did not use standardized underwriting models. The use of non-standardized underwriting models reflects the use of non-traditional data beyond standardized credit scores, and is associated with more individualized pricing (Babina et al., 2022).

\[^{19}\text{We follow what is standard in the literature to select our sample of mortgages. We focus on conventional mortgages for purchase or refinancing as principal residence; we drop reverse mortgages, those with business or commercial purpose, with interest only or balloon payment, more than one unit. Further, we drop applications with missing applicant age or sex, non-conforming loans and open-end line of credits, as well as files that were closed for incompleteness.}\]

\[^{20}\text{We follow Babina et al. (2022) to compute the fraction of mortgages originated using a credit scoring model besides standardized Equifax, Experian, FICO, or Vantage Score.}\]
**Descriptive statistics.** Our final sample contains 4,769,745 mortgage applications between 2018 and 2021 in California, Arizona, Nevada, and Oregon, in a total of 11,350 census tracts. Among all applications, 900,270 were in border counties. In the average tract, out of all applications a share of 14.6% were with fintechs and the remainder with traditional or shadow banks. Across tracts, the share has a standard deviation of 8.5%. The respective figures are 14.1% and 7.8% in the border sample. The interest rate charged in the average tract was 3.7%, with a standard deviation of 0.79 (3.8% and 0.82 for the border sample).

Table 1 provides summary statistics at the lender–borrower tract–year level for banks and fintechs. Panel (a) provides information for all tracts in the four states, while panel (b) focuses on the tracts in border counties only. In the border sample, fintechs have on average more applications, charge a lower rate, and rely significantly less on non-standardized credit scoring models than banks.\footnote{The discrepancy in the use of non-standardized credit scoring models is also shown by Babina et al. (2022).} The composition of fintech and bank borrowers across tracts is statistically similar in terms of observable characteristics gender, race, income, and the loan-to-income ratio. Fintechs, however, tend to lend more to households with a favourable debt-to-income ratio below 36%.

### 4.2 Empirical strategy and results

The model suggests that privacy regulation such as the CCPA increases consumers’ control over their data. By assuaging concerns regarding data abuse, which disproportionately benefits the fintech, regulation leads some consumers that initially applied to banks to apply with the fintech. As applicants face a lower disutility from disclosing personal data, the fintech asks for more data and is better able to screen out low-quality borrowers. A better pool of borrowers allows the fintech to offer loans at a lower interest rate. In
### Table 1: Summary statistics

#### Panel (a): all tracts

<table>
<thead>
<tr>
<th></th>
<th>banks</th>
<th>fintechs</th>
<th>mean</th>
<th>sd</th>
<th>mean</th>
<th>sd</th>
<th>mean diff.</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>nr of applications</td>
<td>6.61</td>
<td>(8.25)</td>
<td>9.53</td>
<td>(16.27)</td>
<td>-67.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>interest rate</td>
<td>3.53</td>
<td>(0.80)</td>
<td>3.37</td>
<td>(0.82)</td>
<td>45.53</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>other CS model</td>
<td>0.23</td>
<td>(0.33)</td>
<td>0.01</td>
<td>(0.06)</td>
<td>172.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>loan amount (in USD th)</td>
<td>346241.50</td>
<td>(132078.06)</td>
<td>373298.98</td>
<td>(134444.59)</td>
<td>-47.42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>share female</td>
<td>0.24</td>
<td>(0.25)</td>
<td>0.24</td>
<td>(0.23)</td>
<td>4.71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>share Black or African Am.</td>
<td>0.02</td>
<td>(0.09)</td>
<td>0.03</td>
<td>(0.09)</td>
<td>-4.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>income (in USD th)</td>
<td>121.55</td>
<td>(56.44)</td>
<td>128.99</td>
<td>(53.74)</td>
<td>-30.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>loan-to-income ratio</td>
<td>3.29</td>
<td>(0.91)</td>
<td>3.31</td>
<td>(0.80)</td>
<td>-4.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>share debt-to-income ratio below 36%</td>
<td>0.43</td>
<td>(0.30)</td>
<td>0.49</td>
<td>(0.28)</td>
<td>-43.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>340,507</td>
<td>64,029</td>
<td>404,536</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Panel (b): border tracts

<table>
<thead>
<tr>
<th></th>
<th>banks</th>
<th>fintechs</th>
<th>mean</th>
<th>sd</th>
<th>mean</th>
<th>sd</th>
<th>mean diff.</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>nr of applications</td>
<td>6.78</td>
<td>(9.54)</td>
<td>8.71</td>
<td>(11.91)</td>
<td>-19.59</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>interest rate</td>
<td>3.58</td>
<td>(0.84)</td>
<td>3.52</td>
<td>(0.87)</td>
<td>7.92</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>other CS model</td>
<td>0.23</td>
<td>(0.33)</td>
<td>0.02</td>
<td>(0.09)</td>
<td>70.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>loan amount (in USD th)</td>
<td>290275.69</td>
<td>(94145.46)</td>
<td>295514.27</td>
<td>(89631.92)</td>
<td>-5.65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>share female</td>
<td>0.25</td>
<td>(0.25)</td>
<td>0.25</td>
<td>(0.24)</td>
<td>-1.41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>share Black or African Am.</td>
<td>0.03</td>
<td>(0.11)</td>
<td>0.03</td>
<td>(0.10)</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>income (in USD th)</td>
<td>106.03</td>
<td>(48.98)</td>
<td>106.69</td>
<td>(42.09)</td>
<td>-1.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>loan-to-income ratio</td>
<td>3.21</td>
<td>(0.92)</td>
<td>3.22</td>
<td>(0.82)</td>
<td>-0.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>share debt-to-income ratio below 36%</td>
<td>0.40</td>
<td>(0.30)</td>
<td>0.45</td>
<td>(0.27)</td>
<td>-16.61</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>63,249</td>
<td>12,105</td>
<td>75,354</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: This table shows summary statistics for our main HMDA variables at the lender–tract–year level, where the sample is split into banks and fintechs. Panel (a) includes all census tracts in CA, NV, AZ, and OR, Panel (b) only those tracts in the border counties. The column *mean* denotes the mean and *sd* the standard deviation; *mean diff.* reports the t-value of a test for the statistical significance of the difference in means.

In this section, we first test the key predictions of the model. We then test implications for the use of data and denial rates.

**Prediction 1:** The introduction of privacy protection regulation leads to an increase in loan applications with fintechs, compared to banks.

**Prediction 2:** The introduction of privacy protection regulation leads to a decrease
in loan rates on loans originated by fintechs, compared to banks.

To investigate the effects of the introduction of the CCPA on loan applications and interest rates, we estimate variants of the following regression at the lender–borrower tract–year level:

\[
y_{l,c,t} = \delta_1 CA_c \times post_t + \delta_2 fintech_l \times post_t \\
+ \delta_3 CA_c \times fintech_l \times post_t + \theta_{l,c} + \tau_{c,t} + \phi_{l,t} + \varepsilon_{l,c,t}
\]  

(3)

The dependent variable \( y \) is the log of the number of applications or the average rate charged on approved mortgages by lender \( l \) in census tract \( c \) in year \( t \). The dummy variable \( CA \) varies at the state level and takes on a value of one if the property is located in a tract in California, and zero otherwise. The dummy \( post \) takes on a value of one after the CCPA was enacted (ie for years 2020 and 2021), and a value of zero in 2018 and 2019. \( FinTech \) is a dummy that takes on a value of one if the lender is a fintech, and a value of zero otherwise. All regressions include lender–tract \( (\theta_{l,c}) \) fixed effects that absorb any time-invariant characteristics at the lender–borrower tract level. We hence only exploit variation within each lender-tract cell. Standard errors are clustered at the borrower tract level.

Based on Prediction 1, we expect a coefficient of \( \delta_3 > 0 \): the introduction of the CCPA should increase consumers’ confidence in sharing their personal data with fintechs. Applications to fintechs are hence expected to increase relative to banks in California after the introduction of the CCPA. Prediction 2 instead suggests that \( \delta_3 < 0 \), ie the introduction of the CCPA leads to lower rates on loans originated by fintechs relative to banks.

Identification. Equation (3) faces the common identification challenge that any observed change in applications or rates could be due to unobservable factors, rather than
due to the introduction of the CCPA. For example, fintechs could serve tracts with higher income growth over the sample period, leading to an increase in applications. Likewise, changes at the lender level, for example the decision by bank management to reduce the lending business in California, would affect the estimated coefficients.

We address this challenge in several ways. For one, we include granular time-varying fixed effects. First, borrower tract*time fixed effects ($\tau_{c,t}$) absorb any observable and unobservable differences in tract characteristics over time. These fixed effects account for eg changes in income, borrower risk, internet access or demographic structure, or credit demand. They also control for potential differences in the severity of the Covid-19 pandemic and associated movement restrictions across tracts.\footnote{For example, California could have seen a stronger rise in cases and/or a stronger decline in mobility. Residents in California might hence have been more inclined to apply for mortgages with online lenders, rather than visit the branch of a traditional bank. However, Table OA7 in the Appendix shows that there were no significant differences in the number of cases per capita and death rates per capita in California’s border counties to those in neighboring states. We obtain similar results for mobility rates (from Google’s COVID-19 Community Mobility Reports; not available for all counties). In robustness tests, we further show that excluding applications of age 62 and above, ie the demographic group most at-risk from Covid-19 and hence most likely to restrict its mobility, does not affect our results.}

With tract*time fixed effects we essentially compare applications to different lenders from borrowers in the same tract and year.

In addition, we include lender*time fixed effects ($\phi_{l,t}$) to control for time-varying observed and unobserved heterogeneity in lender characteristics. These control for changes in eg lenders’ size, risk-taking, or funding models. With the full set of fixed effects, we are comparing how mortgage applications with or rates on mortgages by the same lender change with the introduction of the CCPA in tracts in California compared to comparable tracts in neighboring states.

Finally, to further tighten identification and ensure that our results are not driven by unobservable differences in the pool of borrowers across lenders within tracts, we restrict the analysis to tracts within counties along the border of California with its
neighboring states. As has been shown in a large literature, border counties are generally similar along many observable characteristics, mitigating concerns about selection effects and omitted variable bias (Allegretto et al., 2017). Indeed, Table 1 shows that overall borrower characteristics are similar across lenders in tracts in border counties.

4.2.1 The CCPA, loan applications, and rates

We start with a graphical inspection of the evolution of applications and interest rates. Figure 3 shows that there was no discernible difference in the share of applications with fintechs prior to the introduction of the CCPA in 2020. Panel (a) shows the evolution of loan applications with banks (black dashed line) and fintechs (blue solid line) in California during the sample period. While applications with fintechs and banks evolved similarly between 2018 and 2019, applications with fintechs increased by relatively more in 2020, i.e., when the CCPA came into effect. The gap persisted in 2021, suggesting a lasting effect: individuals know that they are protected by the legislation and they permanently adjust their behavior. Panel (b) shows that there was also no significant difference in interest rates prior to the introduction of the CCPA. Yet rates declined by more on mortgages approved by fintechs compared to banks after the CCPA was introduced in 2020.\footnote{Figure OA4 provides the corresponding coefficient estimates in regressions with a set of applicant-level controls and borrower tract*year fixed effects.}

Moving to the regressions, Table 2 shows that loan applications with fintechs increase in California after the introduction of the CCPA. For all tracts, column (1) shows that applications increased in general in California after the introduction of the CCPA ($\delta_1 > 0$ in Equation (3)). The specification includes lender-tract as well as year fixed effects. The overall increase in applications suggests that the CCPA had an overall positive effect on loan applications, possibly by also increasing some consumers’ willingness to share data with banks. Adding interaction terms in column (2) shows that applications increased
Figure 3: Pre-trends

(a) Loan applications

(b) Interest rates

Panel (a) shows the evolution of loan applications with banks (black dashed line) and fintechs (blue solid line) in California during the sample period. Applications with each lender are standardized to 1 in 2019, the year before CCPA came into effect. Panel (b) plots interest rates on approved mortgages for banks and fintechs.

by significantly more among fintechs compared to banks ($\delta_3 > 0$).

To address the concern that the observed differential change in applications between fintechs and banks is explained by differences in borrower-tract characteristics, column (3) restricts the sample to tracts in border counties. As in the full sample, applications with fintechs increase by significantly more than with banks in California after the introduction of the CCPA. Compared to column (2), the coefficient on the triple interaction term declines in magnitude, which could suggest that some of the observed difference in column (1) is explained by (unobservable) differences in borrower characteristics across non-border tracts.

Column (4) further controls for unobservable time-varying borrower-tract characteristics by introducing tract*time fixed effects. Exploiting within-tract variation, ie comparing lending by fintechs and banks to the same tracts, leads to almost identical coefficient estimates. These results suggest that borrowers in tracts in border counties are comparable in terms of observable and unobservable characteristics. Finally, column (5) introduces
lender*time fixed effects to absorb any time-varying unobservable lender characteristics. The coefficient on the triple interaction effect remains positive and significant at the 1% level. In terms of economic magnitude, applications with fintech lenders increase by 13.9% more than with banks in California after the introduction of the CCPA.

How does the increase in applications with fintech lenders translate into changes in fintechs’ market share? In the Online Appendix we show that the share of applications with fintechs in Californian border counties increases by 2.9 percentage points after the introduction of the CCPA (see Table OA1, column (5)). In 2018–2019, the share of loan applications with fintechs was 11.6%, implying a sizeable increase of 24% of the mean.

Table 2: The introduction of the CCPA and loan applications

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA x post</td>
<td>0.147***</td>
<td>0.097***</td>
<td>0.120***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.011)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fintech x post</td>
<td>0.291***</td>
<td>0.273***</td>
<td>0.284***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.014)</td>
<td>(0.014)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA x fintech x post</td>
<td>0.219***</td>
<td>0.133***</td>
<td>0.134***</td>
<td>0.139***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.020)</td>
<td>(0.021)</td>
<td>(0.022)</td>
<td></td>
</tr>
</tbody>
</table>

Observations               | 404,536 | 404,536 | 75,354  | 75,354  | 75,211  |
R-squared                   | 0.754   | 0.763   | 0.763   | 0.790   | 0.839   |
Lender*Tract FE            | ✓       | ✓       | ✓       | ✓       | ✓       |
Time FE                    | ✓       | ✓       | ✓       | ✓       | ✓       |
Tract*Time FE              | -       | -       | -       | ✓       | ✓       |
Lender*Time FE             | -       | -       | -       | -       | ✓       |

Note: This table reports results at the lender-borrower tract-year level. The dependent variable is the log of the total number of applications. The dummy variable CA takes on a value of one if the property is located in a tract in California. The dummy post takes on a value of one after the CCPA was enacted. The dummy fintech takes on a value of one if the lender is a fintech, and a value of zero if it is a bank. Columns (1)–(2) include all tracts in CA, NV, AZ, and OR; columns (3)–(6) include only tracts in CA border counties. Standard errors are clustered at the tract level. *** p<0.01, ** p<0.05, * p<0.1.

In Table 3 we analyze the effects of the CCPA on loan rates, ie Prediction 2. Column (1) shows that on average, loan rates in California increased by significantly more after
2020 than in neighboring states. Specifically, interest rates increased by around 15 basis points (bp), equaling around 4% of the average (or 0.18 standard deviations). Adding interaction effects with the fintech dummy in column (2) shows that while rates on loans originated increased on average in California, they declined among fintechs.

To ensure effects are not driven by unobservable time-varying borrower tract or lender characteristics, columns (3)–(5) tighten identification. Column (3) focuses on the sample of border counties. Results are similar in terms of sign, size and significance. Columns (4) and (5) add tract*time and lender*time fixed effects. Results show that, after holding all observable and unobservable variation across time at the tract and lender level constant, rates on fintech-approved mortgages in California decrease by 13.9 bp compared to banks.

Table 3: **The introduction of the CCPA and loan rates**

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>all tracts</td>
<td>rate</td>
<td>rate</td>
<td>rate</td>
<td>rate</td>
<td>rate</td>
</tr>
<tr>
<td>CA x post</td>
<td>0.104***</td>
<td>0.123***</td>
<td>0.141***</td>
<td>0.137***</td>
<td>0.141***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.008)</td>
<td>(0.003)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>fintech x post</td>
<td>0.062***</td>
<td>0.067***</td>
<td>0.057***</td>
<td>0.058***</td>
<td>0.057***</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>CA x fintech x post</td>
<td>-0.127***</td>
<td>-0.083***</td>
<td>-0.080***</td>
<td>-0.119***</td>
<td>-0.119***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.015)</td>
<td>(0.015)</td>
<td>(0.014)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Observations</td>
<td>404,536</td>
<td>404,536</td>
<td>75,354</td>
<td>75,354</td>
<td>75,211</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.903</td>
<td>0.903</td>
<td>0.889</td>
<td>0.904</td>
<td>0.922</td>
</tr>
<tr>
<td>Lender*Tract FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Time FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tract*Time FE</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Lender*Time FE</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>✓</td>
</tr>
</tbody>
</table>

Note: This table reports results at the lender-borrower tract-year level. The dependent variable is the average interest rate on approved mortgages. The dummy variable CA takes on a value of one if the property is located in a tract in California. The dummy post takes on a value of one after the CCPA was enacted. The dummy fintech takes on a value of one if the lender is a fintech, and a value of zero if it is a bank. Columns (1)–(2) include all tracts in CA, NV, AZ, and OR; columns (3)–(6) include only tracts in CA border counties. Standard errors are clustered at the tract level. *** p<0.01, ** p<0.05, * p<0.1.
In sum, the results in Figure 3 and Table 3 provide empirical support for Predictions 1 & 2. Applications to fintechs, relative to banks, increased by more in California after the introduction of the CCPA, compared to neighboring states. Loan rates on mortgages approved by fintechs decreased by relatively more.

4.2.2 The use of non-traditional data and denial rates

In the model, privacy legislation provides consumers with greater certainty that their data will not be used for unintended purposes, making them more willing to share data with lenders. In consequence, lenders ask for more data. As fintechs can use the data to obtain a better signal about the quality of their prospective borrowers, their data increases by relatively more than that of banks. As more data and the better signal allow fintechs to better screen out low-quality borrowers, the share of denied applications increases. The model hence yields the following two implications:

Implication 1: Fintech lenders ask for additional data from applicants after the introduction of privacy protection legislation.

Implication 2: The introduction of privacy protection regulation leads to an increase in the share of denied applications.

We test these implications in Table 4, where we estimate variations of regression equation (3) among border tracts. Columns (1)–(3) use the share of mortgages that did not use standardized underwriting models as dependent variable. Consistent with Implication 1, column (1) shows a significant and positive coefficient on the triple interaction term: conditional on lender*tract and year fixed effects, fintechs in California increased their use of non-traditional data beyond standardized credit scores after the introduction of the CCPA. Adding tract*time or lender*time fixed effects in columns (2) and (3) does not materially affect this conclusion. Consistent with obtaining a more precise signal, in the
Online Appendix we show that the dispersion in interest rates increases by relatively more for fintechs (see Table OA2). Babina et al. (2022) argue that the use of non-standardized underwriting models reflects the use of non-traditional data beyond standardized credit scores, and show that it is associated with more individualized pricing. This is in line with theoretical findings (Jansen et al., 2022) that suggest that a more precise signal about borrowers’ quality leads to more accurate pricing and hence greater dispersion in interest rates.

Table 4: Credit scoring models and denial rates

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1) alt CS</th>
<th>(2) alt CS</th>
<th>(3) alt CS</th>
<th>(4) denied</th>
<th>(5) denied</th>
<th>(6) denied</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA x post</td>
<td>-0.034***</td>
<td></td>
<td></td>
<td>0.003*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td></td>
<td>(0.002)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fintech x post</td>
<td>0.022***</td>
<td>0.020***</td>
<td></td>
<td>-0.008**</td>
<td>-0.010***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.003)</td>
<td></td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td>CA x fintech x post</td>
<td>0.028***</td>
<td>0.028***</td>
<td>0.014***</td>
<td>0.010**</td>
<td>0.011**</td>
<td>0.011**</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Observations</td>
<td>75,354</td>
<td>75,354</td>
<td>75,211</td>
<td>75,354</td>
<td>75,354</td>
<td>75,211</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.770</td>
<td>0.796</td>
<td>0.853</td>
<td>0.550</td>
<td>0.598</td>
<td>0.599</td>
</tr>
<tr>
<td>Lender*Tract FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Time FE</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>✓</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tract*Time FE</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Lender*Time FE</td>
<td>-</td>
<td>-</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>✓</td>
</tr>
</tbody>
</table>

Note: This table reports results at the lender-borrower tract-year level. The dependent variable is the share of mortgages that did not use standardized underwriting models in columns (1)–(3); and the share of denied loan applications in columns (4)–(6). The dummy variable CA takes on a value of one if the property is located in a tract in California. The dummy post takes on a value of one after the CCPA was enacted. The dummy fintech takes on a value of one if the lender is a fintech, and a value of zero if it is a bank. The sample includes only tracts in CA border counties. Standard errors are clustered at the tract level. *** p<0.01, ** p<0.05, * p<0.1.

Columns (4)–(6) test Implication 2 and analyze at the effects of the CCPA on the share of denied loan applications within each lender–tract–year cell. Column (4) uses lender*tract and year fixed effects, column (5) adds tract*time fixed effects, and column (6) further adds lender*time fixed effects. Across specifications, rejection rates by fintechs
significantly increase after the introduction of the CCPA, relative to banks.

4.3 Robustness tests

In the Online Appendix, we provide a series of additional tests to examine the robustness of our findings (see Table OA3). We show that applications with fintechs increase and rates decline both among purchase as well as refinance mortgages. Results remain unaffected when we exclude borrowers of age 62 and above from the sample, ie borrowers that could have been more affected by Covid-19 related restrictions. They also remain robust when we include a large set of applicant characteristics as controls at the lender–tract–year level (eg the age, gender or race composition, as well as income levels or loan-to-income ratios).

We also examine our hypotheses in a sample restricted to mortgage loans that were sold in the respective calendar year. Fintechs usually sell mortgages on the secondary market, while banks generally keep a higher share of their mortgages on their balance sheet. By focusing on sold loans, we address the concern that banks are more likely to make on-balance sheet loans that are funded by deposits and could differ in (unobservable) characteristics from fintech loans. For example, they could reflect a riskier borrower pool that requires more careful screening. However, we find that also among sold mortgages applications increase and loan rates fall among fintechs in California after the introduction of the CCPA.

We confirm our results when we directly compare applications and loan rates for fintechs vs banks in California only (see Table OA4). This exercise ensures that our results are not spuriously driven by movements in the control group. In We also ensure that our result are robust to different levels of clustering (see Table OA6).

\footnote{Note that we are also focusing on conventional and conforming loans, which are relatively homogeneous and can be easily sold to government sponsored enterprises.}
5 Conclusion

This paper investigates how preferences for privacy and privacy regulation shape loan market outcomes. A parsimonious screening model of the loan market predicts that the introduction of privacy protection legislation leads to an increase in loan applications with fintechs as well as a lower rate on loans offered by the fintech compared to banks.

Exploiting the introduction of the California Consumer Privacy Act in 2020, we find strong empirical support for these predictions in the US mortgage market. Applications to fintechs, relative to banks, increased by more in California after the introduction of the CCPA, compared to neighboring states. Further, loan rates on mortgages approved by fintechs decreased by relatively more after the CCPA came into effect.

Our results have implications for the policy debate on how to regulate the use of personal data. While personal data can improve lenders’ screening ability, users value their privacy – leading to a trade-off for policy makers. Our results suggest that privacy protection legislation that enhances users’ control over data can mitigate this trade-off.
References


Babina, T., G. Buchak, and W. Gornall (2022) “Customer data access and fintech entry: Early evidence from open banking”, *Available at SSRN*.


Canayaz, M., I. Kantorovitch, and R. Mihet (2022) “Consumer privacy and value of consumer data”, *Available at SSRN 3986562*.


Chu, Y. and J. Wei (2021) “Fintech entry and credit market competition”, *Available at SSRN 3827598*.


A Appendix

A.1 Model derivations

A.1.1 Preliminary analysis

The probability that a good signal is observed is equal to:

\[
\sigma_j = \Pr(\eta_{j,g}) = \Pr(\eta_{j,g}|p_H) \Pr(p_H) + \Pr(\eta_{j,g}|p_L) \Pr(p_L) = \frac{1}{2} + \frac{1}{2}(1 - \gamma_j \sqrt{d_j}) = 1 - \frac{\gamma_j \sqrt{d_j}}{2}
\]

Conditional on observing a good signal, the probability of project success is:

\[
\Pr(p_H|\eta_{j,g}) = \frac{\Pr(\eta_{j,g}|p_H) \Pr(p_H)}{\Pr(\eta_{j,g})} = \frac{1}{2\sigma_j}
\]

A.1.2 Subgames

The game described in the main text has the following subgames with associated payoffs.

- In the smallest sub-game, the lenders know their own and rival’s choices of \((r, d)\) and observe the signal from their applicants. Observing the signal, lenders can either qualify \(q_j = 1\) or disqualify \(q_j = 0\) the applicants. The expected profits per applicant conditional on the signal realization (\(\mathbb{E}[\pi|\eta]\)) are:

  \[
  \mathbb{E}[\pi_j(q_j, r_j, d_j; \ell_j)|\eta_{j,g}] = \begin{cases} 
  \mathbb{E}[p|\eta_{j,g}]r - 1 & \text{if } q_j = 1 \\
  0 & \text{if } q_j = 0
  \end{cases} \quad (4)
  \]

  \[
  \mathbb{E}[\pi_j(q_j, r_j, d_j; \ell_j)|\eta_{j,b}] = \begin{cases} 
  -1 & \text{if } q_j = 1 \\
  0 & \text{if } q_j = 0
  \end{cases} \quad (5)
  \]
• In the second-to-smallest subgame, the applicants choose which lender to apply.

Observing lenders’ offers \((r,d)\), applicants decide to which lender apply. Denote the choice of each borrower by \(\ell\), so that \(\ell_j\) denotes the choice to apply to lender \(j\).

For a borrower with sensitivity \(s\), the expected utility of applying for a loan from lender \(j\) with an interest rate \(r_j\) and requested data \(d_j\) is:

\[
\mathbb{E}[u_s(\ell_j; r_j, d_j)] = \mathbb{E}[u_j | \eta_{j,g}] \Pr(\eta_{j,g}) + \mathbb{E}[u_j | \eta_{j,b}] \Pr(\eta_{j,b})
\]

\[
= \left[ \mathbb{E}\left[p_i(Y - r_j) | p_H; \eta_{j,g}\right] \Pr(p_H | \eta_{j,g}) + \mathbb{E}\left[p_i(Y - r_j) | p_L; \eta_{j,g}\right] \Pr(p_L | \eta_{j,g}) \right] \Pr(\eta_{j,g})
\]

\[
- c(d_j; x_j, s_j)
\]

\[
= \sigma_j \left[ \frac{\theta}{\sigma_j} (Y - r_j) \right] - c(d_j; x_j)
\]

\[
= \theta(Y - r_j) - c(d_j; x_j, s_j)
\]

Where \(c(d_j; x_j)\) denotes the transportation costs and disutility for giving data away, which is lender specific, and \(\theta = 1/2\) is the proportion of high-probability borrowers.

An applicant located at position \(x \in [0,1]\) has three choices: either go to bank 1, which is at a distance \(x\); go to bank 2, at a distance \(1 - x\), or go to the FinTech, which is at a distance \(\bar{x}_F\), normalized to zero for expositional clarity.

\[
\mathbb{E}[u_s(\ell_{B1}; r_{B1}, d_{B1}, x)] = \frac{1}{2}(Y - r_{B1}) - t_B x - d_{B1}
\]

\[
\mathbb{E}[u_s(\ell_{B2}; r_{B2}, d_{B2}, x)] = \frac{1}{2}(Y - r_{B2}) - t_B (1 - x) - d_{B2}
\]

\[
\mathbb{E}[u_s(\ell_{F}; r_{F}, d_{F}, x)] = \frac{1}{2}(Y - r_{F}) - s_F d_{F}
\]

The borrowers’ expected utility of applying to the bank depends on the expected benefit from the project \(\mathbb{E}[p](Y - r_{Bk})\), minus the cost of applying to the bank. Two elements impact the costs in which the borrower incurs: one is the convenience dis-
tance to the bank, $t_B x_{Bk}$, which is the standard Hotelling costs. The parameter $t_B$ measures the importance of the convenience differentiation and provides a proxy for market competitiveness: if $t_B = 0$, the loan market is perfectly competitive; while if $t_B > 0$, there is some level of differentiation between the banks and applicants see each lender as offering them different convenience depending on where they are on the line. In the main text we normalize the parameter $t_B = 1$. The second element impacting costs is the amount of data that the borrower gives to the lender. The sensitivity to sharing data $s_j$ denotes how important these privacy costs are.

The borrower’s expected utility of applying to the fintech also depends on the expected benefit from the project. The costs, however, differ from those of contracting with the bank: first, the convenience to the fintech is the same for all borrowers along the line $x_F$, and does not depend on the intensity of competition in the banking sector $t_B$; second, following the evidence displayed in Figure 2, the sensitivity of consumers with respect to contracting with a traditional bank is higher $s_F > s$.

In what follows, we normalize the sensitivity to contracting with a traditional bank to be $s = 1$ and focus on the cases when $s_F$ is not too large, i.e., cases where the applicant’s dislike for sharing data does not preclude her from applying.

Individuals’ choice of lender gives rise to the demand for each lender. Denote by $\tilde{x}_1$ the position of the borrower indifferent between applying to bank 1 or the fintech, and by $\tilde{x}_2$ the position of the borrower indifferent between applying to lender 2 or the fintech. Therefore the demand that each lender faces is:

$$D_1(\cdot) = \begin{cases} 0 & \text{if } \tilde{x}_1 \leq 0 \\ \tilde{x}_1 & \text{if } \tilde{x}_1 \in (0, 1) \\ 1 & \text{if } \tilde{x}_1 \geq 1 \end{cases} \quad D_2(\cdot) = \begin{cases} 1 & \text{if } \tilde{x}_2 \leq 0 \\ 1 - \tilde{x}_2 & \text{if } \tilde{x}_2 \in (0, 1) \\ 0 & \text{if } \tilde{x}_2 \geq 1 \end{cases}$$ (6)
While the demand for the fintech is:

\[
D_F(\cdot) = \begin{cases} 
0 & \text{if } \tilde{x}_1 = \tilde{x}_2 \\
\tilde{x}_j - \tilde{x}_k & \text{if } 1 > \tilde{x}_j > \tilde{x}_k > 0 \\
1 & \text{if } \tilde{x}_1 \leq 0 \land \tilde{x}_2 \geq 1 
\end{cases} \tag{7}
\]

The resulting demands are as follows:

\[
D_{B1}(\cdot) = \int_{0}^{\tilde{x}_1} 1 df(x) = \tilde{x}_1 = \frac{-2d_{B1} + 2d_{FSF} - r_{B1} + r_F}{2t_B} \\
D_{B2}(\cdot) = \int_{\tilde{x}_2}^{1} 1 df(x) = 1 - \tilde{x}_2 = \frac{-2d_{B2} + 2d_{FSF} - r_{B2} + r_F}{2t_B} \\
D_F(\cdot) = \int_{\tilde{x}_1}^{\tilde{x}_2} 1 df(x) = \tilde{x}_2 - \tilde{x}_1 = \frac{2t_B + 2d_{B1} + 2d_{B2} - 4d_{FSF} + r_{B1} + r_{B2} - 2r_F}{2t_B}
\]

As standard in Hotelling models with firms on opposite sides of the line, demand is continuous: a small change in the interest rate of any lender implies a small change in their number of applicants.

- At the beginning of the game, the expected profits of the lenders per applicant are:

\[
\mathbb{E}[\pi_j(r_j, d_j)] = \left[ \mathbb{E}[\pi_j(\eta_{j,g}) \Pr(\eta_{j,g})] + \mathbb{E}[\pi_j(\eta_{j,b}) \Pr(\eta_{j,b})] \right] \\
= \left[ \left( \mathbb{E}[\pi_j(p_H; \eta_{j,g}) \Pr(p_H|\eta_{j,g})] + \mathbb{E}[\pi_j(p_L; \eta_{j,g}) \Pr(p_L|\eta_{j,g})] \right) \Pr(\eta_{j,g}) \right] \\
= \left[ \left( (r_j - 1) \Pr(p_H|\eta_{j,g}) + (-1) \Pr(p_L|\eta_{j,g}) \right) \Pr(\eta_{j,g}) \right] \\
= \left[ \left( \frac{\theta}{\sigma_j} r_j - 1 \right) \frac{\gamma_j}{\sigma_j} \right] \\
= \left[ \left( \frac{\theta}{\sigma_j} r_j - 1 \right) \sqrt{d_j} \right] \\
\mathbb{E}[\pi_j(r_j, d_j)] = \frac{1}{2} r_j - 1 + \frac{1}{2} \gamma_j \sqrt{d_j} \tag{8}
\]
The impact of the signal on lenders’ profits is evident in Equation (8). The signal allows lenders to only use their funds on those borrowers likely to repay. Effectively, data reduces lenders’ cost of making a wrong decision, as it filters some of the creditunworthy applicants. This allows the lender to offer more attractive interest rates.

Total expected profits are hence:

$$E[\Pi_j(r_j, d_j; r_{-j}, d_{-j})] = D_j(r_j, d_j; r_{-j}, d_{-j}) \mathbb{E}[\pi_j(r_j, d_j)]$$

$$= D_j(r_j, d_j; r_{-j}, d_{-j}) \left[ \frac{1}{2} r_j - 1 + \frac{\gamma_j \sqrt{d_j}}{2} \right]$$

(9)  

Proceeding by backwards induction, lenders take the data choices as given and internalize the effect of their choices on demand. The optimal interest rates as a function of data are given by the solution to the first order conditions:

$$\frac{\partial E[\Pi_j]}{\partial r_j} \left[ \frac{1}{2} r_j - 1 + \frac{1}{2} \gamma_j \sqrt{d_j} \right] \frac{\partial D_j(\cdot)}{\partial r_j} + D_j(\cdot) \frac{1}{2} = 0$$

(10)

Call the solutions to the FOCs $r_j^*(d) = r_j^*(d_j; d_{-j})$:

$$r_{B1}^*(d) = \frac{1}{12} \left( -7 \gamma_B \sqrt{d_{B1}} - \gamma_B \sqrt{d_{B2}} + 4 \left( t_B + 2 d_F s_F - \sqrt{d_F} + 6 \right) - 10 d_{B1} + 2 d_{B2} \right)$$

$$r_{B2}^*(d) = \frac{1}{12} \left( -\gamma_B \sqrt{d_{B1}} - 7 \gamma_B \sqrt{d_{B2}} + 4 \left( t_B + 2 d_F s_F - \sqrt{d_F} + 6 \right) + 2 d_{B1} - 10 d_{B2} \right)$$

$$r_F^*(d) = \frac{1}{6} \left( -\gamma_B \sqrt{d_{B1}} - \gamma_B \sqrt{d_{B2}} - 4 \left( -t_B + d_F s_F + \sqrt{d_F} - 3 \right) + 2 d_{B1} + 2 d_{B2} \right)$$

At the beginning of the game, the lenders choose the accuracy of their screening technology taking into account the effect that the data requested have on the optimal
rates, as well as on the demand:

\[
E[\Pi_j(d)] = D_j\left(r_j^*(d), r_{-j}^*(d), d\right) \left[\frac{1}{2}r_j^*(d) - 1 + \frac{1}{2}\gamma_j\sqrt{d_j}\right]
\]  \hspace{1cm} (11)

Optimal choices for data maximize the profits in Equation (11) subject to the constraints that \(d_j \in [0, 1] \; \forall j = 1, 2\). Since both banks are identical, we will focus on symmetric solutions, such that \(d_{B1}^* = d_{B2}^* = d_F^*\). The profit function in Equation (11) is highly non-linear, so we need to check the SOCs.

\[
\frac{\partial E[\Pi_B(d)]}{\partial d_B} = -\frac{5(4\sqrt{d_B} - \gamma_B)(\gamma_B\sqrt{d_B} - 2d_B + t_B + 2d_Fs_F - \sqrt{d_F})}{144\sqrt{d_B}t_B} = 0
\]

\[
\frac{\partial E[\Pi_F(d)]}{\partial d_F} = -\frac{(4s_F - \frac{1}{\sqrt{d_F}})(-\gamma_B\sqrt{d_B} + 2d_B + 2t_B - 2d_Fs_F + \sqrt{d_F})}{18t_B} = 0
\]

The solution to the FOCs that satisfies the SOCs are given by:

\[
d_{B1}^* = \frac{\gamma_B^2}{16} \hspace{1cm} (12)
\]

\[
d_{B2}^* = \frac{\gamma_B^2}{16} \hspace{1cm} (13)
\]

\[
d_F^* = \frac{1}{16s_F^2} \hspace{1cm} (14)
\]

The optimal rates are:

\[
r_{B1}^* = \frac{s_F(48 + 8t_B - 5\gamma_B^2) - 1}{24s_F} \hspace{1cm} (15)
\]

\[
r_{B2}^* = \frac{s_F(48 + 8t_B - 5\gamma_B^2) - 1}{24s_F} \hspace{1cm} (16)
\]

\[
r_F^* = \frac{s_F(48 + 16t_B - \gamma_B^2) - 5}{24s_F} \hspace{1cm} (17)
\]
While the equilibrium demand for each lender is:

\[
D_{B1}^* = \frac{s_F (\gamma_B^2 + 8t_B) - 1}{48t_B s_F},
\]
\[
D_{B2}^* = \frac{s_F (\gamma_B^2 + 8t_B) - 1}{48t_B s_F},
\]
\[
D_F^* = \frac{1 - s_F (\gamma_B^2 - 16t_B)}{24t_B s_F}.
\]

We limit ourselves to the cases when the demands for all lenders are strictly positive, which imply that:

\[
\begin{align*}
t_B &> \frac{1 - \gamma_B^2 s_F}{8s_F} & \text{if } s_F \in \left(1, \frac{1}{\gamma_B}\right) \\
t_B &> \frac{\gamma_B^2 s_F - 1}{16s_F} & \text{if } s_F > \frac{1}{\gamma_B}
\end{align*}
\]

This condition also implies that the solution in Equations (12) to (14) are a maximum.

A.1.3 Comparative Statics

**Implication 1:** Fintechs will employ more data.

A decrease in the sensitivity to sharing data mechanically implies that borrowers will share more data with the fintech. This comes from the optimal choice for data in (14). As the cost of asking for more data decreases, the fintech has an incentive to ask for more data.

**Implication 2:** Fintechs will deny more applications.

In the model, the denial rate of the lenders is represented by the amount of creditworthy applicants that return a bad signal: \((1 - \theta)\gamma_F \sqrt{d_F}\).

The additional amount of data improves the fintech’s screening accuracy, which allows it to filter out more creditworthy applicants, increasing its denial rates.
Prediction 2: Further to the decrease in the interest rate of fintechs, note that traditional banking sector also lowers their interest rate because interest rates are strategic complements. Yet the effect is more pronounced for the fintech.

A.1.4 Consumer surplus

We can calculate ex-ante total consumer surplus by integrating over the whole line. Recall that the expected utility of a consumer of contracting with a lender \( j \) at distance \( x_j \) is:

\[
E[u_s(\ell_j; r_j^*, d_j^*)] = \theta(Y - r_j^*) - c(d_j^*; x_j, s_j)
\]

Abusing notation, let us call the position of the indifferent consumers between bank \( b \) and the fintech as \( \tilde{x}_b^* = \tilde{x}_b(r_b^*, r_F^*, d_b^*, d_F^*; s_F) \)

\[
CS = \int_0^{\tilde{x}_1^*} \frac{1}{2} (Y - r_1^*) - t_B^* dx - d_1^* dx + \int_{\tilde{x}_1^*}^{\tilde{x}_2^*} \frac{1}{2} (Y - R_F^*) - d_F^* dx + \int_{\tilde{x}_2^*}^{Y} \frac{1}{2} (Y - R_2^*) - t_B^* dx - d_2^* dx \\
= s_F (56t_B^* - 2\gamma_B^2) + s_F^2 (\gamma_F^2 + 16t_B^* (\gamma_B^2 + 36Y - 72) - 512t_B^2) + 1 \\
1152t_B^*s_F^2
\]

A.2 Additional information on the CCPA

A.2.1 General information

The California Consumer Privacy Act is a law passed in June 2018 that applies to companies handling personal information of California residents. It went into effect in January 2020. It endows Californians with several rights:

- The right to delete personal information collected from them;
- The right to know what personal information a business has collected about them
and how it is used and shared;

• The right to opt-out of the sale of their personal information; and

• The right to non-discrimination for exercising their CCPA rights.

Starting in 2023, the law will also include the right to correct inaccurate information and the right to limit the use and disclosure of sensitive personal information.

The companies that are subject to CCPA are those that: Have a gross annual revenue of over $25 million; Buy, receive, or sell the personal information of 50,000 or more California residents, households, or devices in one year; Or derive 50% or more of their annual revenue from selling California residents personal information.25

The scope of what constitutes personal information under CCPA is very broad. For example, an IP address of an individual browsing a website is considered personal information. Therefore, CCPA is likely to cover a large proportion of companies, including numerous small- to medium-sized enterprises.26

A.2.2 Salience

To show that the residents of California were aware of the introduction of the privacy legislation, we can look at Google searches for CCPA using Google trends.

Additionally, Figure OA2 and Figure OA3 provide two screenshots of the information on CCPA consumers can find on mortgage application websites.

25https://cppa.ca.gov/faq.html
26https://iapp.org/news/a/new-california-privacy-law-to-affect-more-than-half-a-million-us-companies/
A.2.3 Enforcement

Under CCPA, individuals can file a consumer complaint with the Office of the Attorney General. Starting in January 2023, claims can also be filed with the recently founded California Privacy Protection Agency. The Attorney General and the Agency investigate violations, either following a consumer complaint or from their own initiative and take enforcement actions. Individual suing of a business is limited to the event of particular data breaches, where it is clear the company did not take the necessary measures to protect consumers’ data.

The first enforcement settlement related to CCPA took place in August 2022 and it concerned the French cosmetics brand Sephora. The investigation, led by the California

https://cppa.ca.gov/faq.html
California Consumer Privacy Act Privacy Notice:

This privacy notice supplements the general privacy policy notice of View Mortgage, LLC NMLS# 2185181 which can be found at https://viewmortgage.com/privacy-policy. This supplemental privacy policy applies solely to consumers who reside in the State of California ("you"). View Mortgage, LLC adopts this notice to comply with the California Consumer Privacy Act of 2018, as amended ("CCPA") and other applicable California privacy laws. Any terms defined in the CCPA have the same meaning when used in this notice.

The CCPA was passed by the State of California in 2018 and provides California residents with the following rights over their personal information: (a) the right to access, transfer, edit and delete their personal data with a verifiable consumer request; and (b) the ability to opt out of certain data-processing practices. In addition, California residents have the right to: (a) know what information is being collected about them; (b) know if their personal information is sold or disclosed, and to whom; (c) say “no” to the sale of personal information; and (d) equal service and price, even if they exercise their privacy rights under the CCPA.

View Mortgage, LLC does not and will not discriminate in any way against any consumers who choose to exercise their rights under the CCPA.

What Information We Collect About You:

We may obtain certain personal information (such as name and other contact details) through our Sites. Here are the most common types of information: Contact information (such as name, postal address, e-mail address, telephone number and fax number); Login and access credentials (such as username and password); Information about your property or mortgage loan; Age and gender; Real estate license number.

Figure OA2: This figure displays the CCPA information available in the website from View Mortgage, available at https://www.viewmortgage.com/ccpa.
California Consumer Privacy Act

Open Mortgage is committed to compliance with the new California Consumer Privacy Act (CCPA) and to providing options to Opt-Out. Additionally, Open Mortgage never sells any of your data. That said, it is important to note that CCPA provides exemptions to companies that have consumer data that is necessary to carry out their business. A majority of the data that Open Mortgage collects is exempt from CCPA as it falls under federal privacy laws set out by the Gramm Leach Bliley Act (GLBA) and cannot, as a result, be part of the Opt-Out request. We are GLBA compliant and protect your data to our fullest capability. So, Open Mortgage will receive your Opt-Out request via the option you select, and will work to remove any data that is not exempt.

What does that mean?

That means that if you have applied for a loan from Open Mortgage, we are exempt from removing personal data that was used in the loan process. Personal data that is exempt will include data like your name, phone number, social security number, date of birth, address, employment data, and income data. So while we have to keep that data, we want you to rest assured that Open Mortgage does not engage in the sale of data. Lastly, please keep in mind that data outside the loan process may not be exempt.

What kind of data is outside the loan process?

Information that was not collected in your loan application and was collected from marketing activities, promotional activities, information from a non-borrowing spouse, IP addresses, geolocation data through the use of a third-party application, web page tracking, and information collected through a third party like a marketing list.

To proceed with your request, please be prepared to provide your first and last name, your phone number, your email address, your loan number (if applicable), and the Loan Officer you have been or were working with. Providing this will help us locate your information and validate your request. Open Mortgage may request additional information if your request cannot be validated. You must be a resident of California to be eligible for this request.

Click here for the full CCPA privacy policy

You can fill out and submit this form to request to opt-out.

Figure OA3: This figure displays the CCPA information available in the website from Open Mortgage, available at https://openmortgage.com/ccpa.
Attorney General, found that Sephora was selling consumers’ personal information without disclosing it, as well as did not comply with opt-out requests. Sephora agreed to pay $1.2 million in fines and agreed to follow its compliance obligations.\(^{28}\)

\(^{28}\)https://iapp.org/news/a/the-sephora-case-do-not-sell-but-are-you-selling/
A.3 Further Figures and Tables

To further investigate the effects of the introduction of the CCPA on loan applications with banks and fintechs, we estimate difference-in-differences (DiD) specifications at the borrower tract–year level:

\[ \text{fintech application share}_{c,t} = \beta \text{CA}_c \times \text{post}_t + \theta_c + \tau_t + \text{controls}_{c/a} + \varepsilon_c. \]  

(21)

The dependent variable is the share of applications filed with fintechs in census tract \( c \) and year \( t \). The dummy variable \( \text{CA} \) varies at the state \( (s) \) level and takes on a value of one if the property is located in California, and zero otherwise. The dummy \( \text{post} \) takes on a value of one after the CCPA was enacted (ie for 2020 and 2021), and a value of zero in the years prior. Regressions include tract \( (\theta) \) and year \( (\tau) \) fixed effects. As discussed in the main text, in some specifications we restrict the sample to tracts in border counties and include granular border-pair fixed effects.

To account for differences in economic and applicant characteristics across tracts, we include a battery of control variables. As tract-level \( (c) \) controls, we include the pre-period values of the minority share, tract-to-MSA income ratio, and the log of the total tract population, all interacted with the post dummy. Further, we include the following pre-treatment applicant-level \( (a) \) controls, averaged to the tract level, interacted with the post dummy: the share of female applicants, the share of black applicants, the share of Hispanic applicants, as well as the log of the average application amount and average applicant income. Standard errors are clustered at the tract level.

Based on Prediction 1, we expect a coefficient of \( \beta > 0 \): the introduction of the CCPA should increase consumers’ confidence in sharing their personal data with fintechs. Applications to fintechs are hence expected to increase relative to banks in California after the introduction of the CCPA.
Table OA1, columns (1)–(5), show that the share of loan applications to fintechs increases, relative to banks, after the introduction of the CCPA. Column (1), with tract and time fixed effects, shows that after the introduction of the CCPA, applications with fintechs, relative to banks, increased by more in California compared to other states (ie, $\beta > 0$). The coefficient is significant at the 1% level. Further including a battery of tract or applicant control variables in columns (2) and (3) only leads to a modest change in the magnitude of the estimated coefficient. Across specifications, it remains highly significant at the 1% level. In terms of economic magnitude, the relative increase in the share of applications with fintechs in California equals 3.6 percentage points (pp) in column (3). In 2018–2019, the share of loan applications with fintechs was 11.6% (with a standard deviation of 8.2%), implying an increase of 31% of the mean (or 0.44 sd).

To tighten identification, column (4) focuses on the set of census tracts that lie in counties along the border of California with its neighboring states. The coefficient of interest remains highly significant and positive, but declines in magnitude to 2.6 pp. To account for time-varying trends within each border-pair, column (5) includes border county*time fixed effects. Accounting for observable and unobservable factors, including loan demand, common to neighboring counties yields a positive and highly significant coefficient on the interaction term. In terms of magnitude, the share of applications with fintechs increases by 2.9 pp (24% of the mean and 0.37 standard deviations).
Table OA1: **The introduction of the CCPA and loan applications (share)**

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>all tracts fintech</td>
<td>all tracts fintech</td>
<td>all tracts fintech</td>
<td>border fintech</td>
<td>border fintech</td>
</tr>
<tr>
<td>CA x post</td>
<td>0.040***</td>
<td>0.040***</td>
<td>0.036***</td>
<td>0.026***</td>
<td>0.029***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.003)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Observations</td>
<td>42,644</td>
<td>42,644</td>
<td>42,644</td>
<td>7,341</td>
<td>7,341</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.594</td>
<td>0.598</td>
<td>0.601</td>
<td>0.569</td>
<td>0.583</td>
</tr>
<tr>
<td>Tract FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Time FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>Controls</td>
<td>-</td>
<td>T</td>
<td>T+A</td>
<td>T+A</td>
<td>T+A</td>
</tr>
<tr>
<td>Border FE</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Border*Time FE</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>✓</td>
</tr>
</tbody>
</table>

Note: This table shows results for Equation (21). T and A refer to controls at the tract level or average borrower level. Columns (1)–(3) include all tracts in CA, NV, AZ, and OR; columns (4)–(6) only those tracts in CA border counties. *** p<0.01, ** p<0.05, * p<0.1.
Figure OA4: Pre-trends – coefficient estimates

(a) Loan applications

(b) Interest rates

This figure plots the coefficient estimates (blue line) and 95% confidence intervals (grey bars) of the fintech dummy on loan applications and rates on approved mortgages. The x-axis covers the time window spanning from 2018 to 2021. Specifically, the estimated coefficients from the following regression for the sample of border counties are reported: \[ y_{l,i,t} = \sum_{K=2018}^{K=2021} \beta_k \text{fintech}_k \times \text{year}_k + \theta_i + \text{controls}_{l,i,t} + \tau_{i,t} + \epsilon_{i,t}. \] Coefficient \( \beta_k \) indicates the evolution of applications or interest rates with fintechs in year \( k \) before/after the introduction of the CCPA in California. The year of prior to the introduction of the CCPA (\( k = 2019 \)) is excluded. Standard errors are clustered at the borrower county level. All regressions include borrower-tract*year fixed effects. Panel (a) shows the coefficient estimates for loan applications, panel (b) for interest rates on approved mortgages.
### Table OA2: CCPA and rate dispersion

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>sd(int rate)</td>
<td>sd(int rate)</td>
<td>sd(int rate)</td>
<td>sd(int rate)</td>
<td>sd(int rate)</td>
</tr>
<tr>
<td>CA x post</td>
<td>-0.037***</td>
<td>-0.069***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.007)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fintech x post</td>
<td>-0.012**</td>
<td>-0.028***</td>
<td>-0.021**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.010)</td>
<td>(0.010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA x fintech x post</td>
<td>0.077***</td>
<td>0.111***</td>
<td>0.098***</td>
<td>0.083***</td>
<td>0.072***</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.013)</td>
<td>(0.014)</td>
<td>(0.014)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>applications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.079***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.004)</td>
</tr>
</tbody>
</table>

Observations: 404,536 75,354 75,354 75,211 75,211
R-squared: 0.527 0.535 0.591 0.622 0.626
Lender*Tract FE: ✓ ✓ ✓ ✓ ✓
Time FE: ✓ ✓ - - -
Tract*Time FE: ✓ ✓ ✓ ✓ ✓
Lender*Time FE: ✓ ✓ ✓ ✓ ✓

Note: This table shows results for Equation (3) for tracts in CA border counties. The dependent variable is the standard deviation of loan rates within tracts. Column (5) holds loan applications constant. Standard errors are clustered at the tract level. *** p<0.01, ** p<0.05, * p<0.1.

### Table OA3: CCPA, applications, and loan rates – robustness

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>purchase</td>
<td>purchase</td>
<td>refinance</td>
<td>refinance</td>
<td>young</td>
<td>young</td>
<td>controls</td>
<td>controls</td>
<td>off-BS</td>
<td>off-BS</td>
</tr>
<tr>
<td></td>
<td>applications</td>
<td>rate</td>
<td>applications</td>
<td>rate</td>
<td>rate</td>
<td>applications</td>
<td>rate</td>
<td>applications</td>
<td>rate</td>
<td>rate</td>
</tr>
<tr>
<td>CA x fintech x post</td>
<td>0.092***</td>
<td>-0.106***</td>
<td>0.114***</td>
<td>-0.071***</td>
<td>0.137***</td>
<td>-0.124***</td>
<td>0.140***</td>
<td>-0.117***</td>
<td>0.103***</td>
<td>-0.129***</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.023)</td>
<td>(0.042)</td>
<td>(0.020)</td>
<td>(0.022)</td>
<td>(0.017)</td>
<td>(0.022)</td>
<td>(0.013)</td>
<td>(0.035)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>Observations</td>
<td>53,972</td>
<td>53,972</td>
<td>37,418</td>
<td>37,418</td>
<td>72,441</td>
<td>72,441</td>
<td>73,767</td>
<td>73,767</td>
<td>52,601</td>
<td>69,446</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.768</td>
<td>0.880</td>
<td>0.789</td>
<td>0.906</td>
<td>0.842</td>
<td>0.905</td>
<td>0.841</td>
<td>0.930</td>
<td>0.764</td>
<td>0.911</td>
</tr>
<tr>
<td>Lender*Tract FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Tract*Time FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Lender*Time FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Note: This table shows results for Equation (3) for those tracts in CA border counties. Columns (1)–(2) restrict the sample to purchase mortgages, columns (3)–(4) to refinance mortgages. Columns (5)–(6) exclude all applicants of age 62 and above from the sample. Columns (7)–(8) control for a rich set of controls for the average applicant at the lender–tract–year level: the share of female applicants, the share of applicants age 62 and above, the share of black applicants, the share of Hispanic applicants, the share of white applicants, the log of income, the log of the loan amount, the loan-to-income ratio, the share of applications with a debt-to-income ratio below 36%, as well as the fractions of applications that were denied, used an alternative credit scoring model, or are purchase mortgages. Columns (9)–(10) restrict the sample to mortgage loans that were sold in the respective calendar year. *** p<0.01, ** p<0.05, * p<0.1.
Table OA4: CCPA, applications, and loan rates — California sample

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1) applications</th>
<th>(2) applications</th>
<th>(3) rate</th>
<th>(4) rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>fintech x post</td>
<td>0.513***</td>
<td>0.516***</td>
<td>-0.063***</td>
<td>-0.069***</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Observations</td>
<td>259,156</td>
<td>259,156</td>
<td>259,156</td>
<td>259,156</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.764</td>
<td>0.799</td>
<td>0.907</td>
<td>0.921</td>
</tr>
<tr>
<td>Lender*Tract FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Time FE</td>
<td>✓</td>
<td>-</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>Tract*Time FE</td>
<td>-</td>
<td>✓</td>
<td>-</td>
<td>✓</td>
</tr>
</tbody>
</table>

Note: This table shows results for the following equation: $y_{i,t} = \delta \text{fintech}_{i} \times \text{post}_{t} + \theta_{i} + \tau_{t} + \epsilon_{i,t}$. Columns (1)–(2) use the log of loan applications as dependent variable, columns (3)–(4) loan rates. Standard errors are clustered at the tract level. *** p<0.01, ** p<0.05, * p<0.1.

Table OA5: Main results for individual states

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1) no AZ applications</th>
<th>(2) no AZ rate</th>
<th>(3) no NV applications</th>
<th>(4) no NV rate</th>
<th>(5) no OR applications</th>
<th>(6) no OR rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA x fintech x post</td>
<td>0.127***</td>
<td>-0.116***</td>
<td>0.304***</td>
<td>-0.090***</td>
<td>0.119***</td>
<td>-0.130***</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.015)</td>
<td>(0.040)</td>
<td>(0.026)</td>
<td>(0.022)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>Observations</td>
<td>72,718</td>
<td>72,718</td>
<td>47,011</td>
<td>47,011</td>
<td>72,058</td>
<td>72,058</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.840</td>
<td>0.919</td>
<td>0.848</td>
<td>0.927</td>
<td>0.841</td>
<td>0.919</td>
</tr>
<tr>
<td>Lender*Tract FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Tract*Time FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Lender*Time FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Note: This table shows results for Equation (3) for tracts in CA border counties. Standard errors are clustered at the tract level. *** p<0.01, ** p<0.05, * p<0.1.
### Table OA6: Clustering

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA x fintech x post</td>
<td>0.139***</td>
<td>-0.119***</td>
<td>0.139***</td>
<td>-0.119***</td>
<td>0.139***</td>
<td>-0.119***</td>
<td>0.139***</td>
<td>-0.119***</td>
<td>0.139**</td>
<td>-0.119***</td>
<td>0.139*</td>
<td>-0.119***</td>
</tr>
<tr>
<td>(0.022)</td>
<td>(0.015)</td>
<td>(0.022)</td>
<td>(0.015)</td>
<td>(0.060)</td>
<td>(0.036)</td>
<td>(0.060)</td>
<td>(0.036)</td>
<td>(0.079)</td>
<td>(0.016)</td>
<td>(0.067)</td>
<td>(0.016)</td>
<td></td>
</tr>
</tbody>
</table>

Observations: 75,211

R-squared: 0.839

Lender*Tract FE: ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓

Tract*Time FE: ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓

Lender*Time FE: ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓

Note: This table shows results for Equation (3) for tracts in CA border counties. Standard errors are clustered at the level indicated in the column header, where T is tract, C is county, S is state, Y is year, and L is lender.

*** p<0.01, ** p<0.05, * p<0.1.

### Table OA7: Covid cases and mobility

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA</td>
<td>0.046</td>
<td>-0.217</td>
<td>-0.468</td>
<td>-0.350</td>
<td>0.080</td>
<td>-0.157</td>
<td>0.025</td>
<td>-0.571</td>
</tr>
<tr>
<td>(0.375)</td>
<td>(0.370)</td>
<td>(0.520)</td>
<td>(0.404)</td>
<td>(0.242)</td>
<td>(0.313)</td>
<td>(0.933)</td>
<td>(0.386)</td>
<td></td>
</tr>
</tbody>
</table>

Observations: 30

R-squared: 0.001

Border FE: - - - - ✓ ✓ ✓ ✓

Note: This table shows differences in Covid cases and mobility in 2020 in the border county sample. All dependent variables are standardized to a mean of zero and a standard deviation of one. *** p<0.01, ** p<0.05, * p<0.1.

58