BIS

BIS Working Papers
No 1103

Privacy regulation and fintech lending
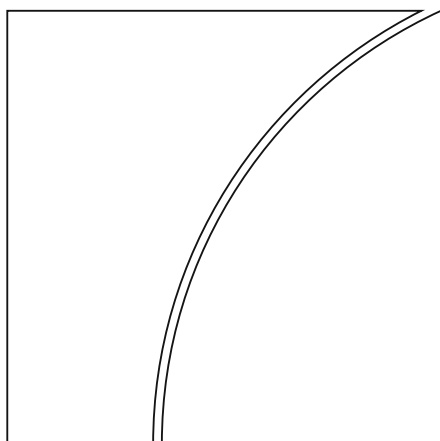
by Sebastian Doerr, Leonardo Gambacorta, Luigi Guiso and Marina Sanchez del Villar

Monetary and Economic Department

June 2023

BIS Working Papers are written by members of the Monetary and Economic Department of the Bank for International Settlements, and from time to time by other economists, and are published by the Bank. The papers are on subjects of topical interest and are technical in character. The views expressed in them are those of their authors and not necessarily the views of the BIS.

This publication is available on the BIS website (www.bis.org).

Privacy regulation and fintech lending

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

June 2023

Abstract

Individuals have concerns about sharing data, but without access to personal information data-intensive fintechs cannot prosper, nor consumers benefit from their services. Well-designed privacy protection regulation needs to address this conflict. This paper studies how the California Consumer Privacy Act (CCPA), a 2020 privacy law that grants users control over data and mitigates concerns over sharing them, affects bank and fintech lending in the mortgage market. Using a difference-in-differences strategy comparing counties along the California border, we establish that the CCPA increases loan applications to fintechs by significantly more than those to banks, raising fintechs' market share by 19%. Consistent with an improved screening process due to applicants' greater willingness to share data with fintechs, fintechs engage in more individualized pricing, deny a greater share of applications, and increase their use of non-traditional data. In turn, fintechs can offer significantly lower loan rates compared with banks after the introduction of the CCPA.

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 on technology to screen and price borrowers, oftentimes by using non-traditional data (Berg et al., 2022; Di Maggio et al., 2022). Better access to data can hence foster the growth of fintechs and promote competition in the financial sector.

There is mounting evidence, however, that consumers dislike sharing their personal data (Goldfarb and Tucker, 2012; Tang, 2022). Concerns range from price discrimination to data abuse, unethical advertising, and financial fraud (Chen et al., 2021; Lin, 2022; Prince and Wallsten, 2022; Bian et al., 2023). Moreover, consumers value with whom they share their data. For a representative sample of U.S. households, Figure 1 documents that consumers have significantly more 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 designing privacy protection regulation hence face a challenging trade-off. Limiting or even prohibiting the collection of personal data protects consumers' privacy, but may come at the cost of reducing the availability and quality of data-intensive services (Acquisti et al., 2016). At the same time, the absence of a regulatory framework that addresses privacy concerns can also impede fintechs' development. Consumers might be reluctant to share data or engage with fintechs altogether, as they are concerned over privacy violations (Chen et al., 2023). Regulators hence need to strike the right balance between protecting users' privacy and promoting the growth of fintechs to foster competition.

This paper focuses on the effects of the California Consumer Privacy Act (CCPA) on bank and fintech lending in the U.S. mortgage market. Introduced in 2020, the CCPA constitutes a landmark change in the design of privacy regulation. By giving California residents control over their data and thus mitigating concerns over sharing them, it attempts to protect consumers without generally restricting information collection. For example, individuals can prevent firms from selling their personal information or request that firms delete the data after their use (Camhi and Lyon, 2018). Recent survey evidence shows that the majority of Californians have already exercised their CCPA rights.[1]

The CCPA differs from other types of privacy regulation that typically limit firms' ability to collect information.[2] It also differs from open banking, which aims to facilitate access to data, rather than protect privacy per se.[3] By now, most other states consider introducing legislation in the spirit of the CCPA, and it serves as U.S. Congress' model to create a federal privacy protection regulation (Mulligan et al., 2019). Understanding the CCPA's impact can yield important insights on the benefits of similar legislation outside of California.

Our analysis uses data on residential mortgages from the Home Mortgage Disclosure Act (HMDA) database from 2018–2021. HMDA provides a wealth of information on lenders, applicants, and loan terms. We classify lenders into banks and fintechs following Fuster et al. (2019). For identification, we estimate difference-in-differences regressions

---

[1]See Consumer Action and Consumer Federation of America (2021): California Consumer Privacy Act (CCPA) Survey.

[2]For example, a key principle of the the European Union's General Data Protection Regulation (GDPR) is that firms must minimize their data processing activities, eg use data only for the purpose for which they were initially collected. Similar to policies such as bankruptcy flag removal, the GDPR hence effectively prevents the collection and use of data. Since data are valuable, such regulation can reduce access to and increase the cost of credit (Liberman et al., 2019; Jansen et al., 2022; Johnson, 2022).

[3]Open banking mandates financial institutions to share their proprietary data with third parties, including fintech lenders, if users give their consent. The need to share data, however, can discourage incumbents from acquiring and storing data in the first place (Babina et al., 2022). It can also increase market power to the detriment of consumers (He et al., 2023), or distort banks' credit allocation (Goldstein et al., 2022).

with individual borrower data collapsed to the lender–tract–year level. The CCPA, introduced in 2020, acts as treatment for tracts in California, while tracts in the neighboring states constitute the control group.

We first investigate the effects of the introduction of the CCPA on loan applications. Leveraging on the fact that privacy concerns are more salient towards fintechs (see Figure 1, as well as Armantier et al. (2021) and Chen et al. (2023)), we hypothesize that the marginal benefit of privacy regulation that assuages concerns about sharing data is higher for newer intermediaries than for traditional financial institutions.[4] Consistent with this argument, we find that after the introduction of the CCPA loan applications to fintechs compared to banks increase significantly in California, relative to neighboring states. In terms of magnitude, loan applications to fintechs increase by 14.6% after the introduction of the CCPA. This implies an increase in fintechs' market share of 3 percentage points (pp), or as much as 19% of their pre-CCPA market share.

Second, we study the impact of the CCPA on interest rates on approved mortgages. In principle, an increase in demand for fintech loans due to better privacy regulation could lead to an increase in the rates they charge. On the other hand, if users become more willing to share data and thus apply to fintechs, the accrual of additional information could allow fintechs to better screen applicants (Di Maggio et al., 2022). This would lower the risk of their borrower pool and result in lower rates. We find that the CCPA reduced loan rates by around 8 basis points (or 13% of the standard deviation) on mortgages originated by fintechs relative to banks in California. This pattern suggests that the privacy protection offered by the CCPA made applicants more willing to share data, thereby improving fintechs' screening process.

---

[4]Banks 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. Moreover, as established players, banks oftentimes have long-standing relationships with clients, bolstering trust.

We provide empirical evidence in support of the argument that the CCPA has improved fintechs' screening abilities. We first show that, after the introduction of the CCPA, the dispersion in interest rates increases by relatively more for fintechs. Greater rate dispersion is consistent with more individualized pricing due to a more precise signal about applicants' quality (Babina et al., 2022; Jansen et al., 2022). Second, we show that the share of denied loan applications increases for fintechs compared to banks in California after the introduction of the CCPA. The relative increase in denial rates by 1 pp (or 19% of the pre-reform mean) suggests an improvement in fintechs' ability to screen out low-quality applicants. Third, we find that fintechs in California process significantly more mortgage applications with data beyond standardized credit scores after the introduction of the CCPA. While direct information on lenders' use of different types of data is difficult to obtain, the use of non-standardized credit scores in processing applications proxies a greater reliance on alternative personal data (Babina et al., 2022). The relative increase in fintechs' use of non-standardized credit scores is in line with our argument that the CCPA increases applicants' willingness to share data, thereby improving fintechs' screening process.

In further support of fintechs' improved screening capabilities due to better access to data we find the negative effect of the CCPA on fintechs' loan rates to be stronger in tracts with a higher share of 'thin credit file' applicants. This finding is consistent with Blattner and Nelson (2021) and Di Maggio et al. (2022), who show that thin credit file borrowers (oftentimes traditionally under-served groups, including minority and low-income applicants) stand to benefit the most from screening with better data, as standard credit scores are noisier indicators of their default risk. These results raise the question of whether the CCPA has improved access to credit. We find that although denial rates increase, banks and especially fintechs originate more mortgages in California 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 change in financial conditions that affects fintechs and banks differently could be reflected in the rates they charge.

We address this challenge in different ways. For one, our analysis focuses only on 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 similar along many observable characteristics, mitigating concerns about selection effects and omitted variable bias (Allegretto et al., 2017). Indeed, in the border sample we find no discernible difference in the evolution of loan applications to or rates on loans by fintechs compared to banks prior to the introduction of the CCPA in 2020 (ie, there is no evidence of differential pre-trends). In addition, we show that the average fintech and bank applicant have comparable observable characteristics in border tracts before the CCPA was introduced.

We further include granular time-varying fixed effects. Tract*time fixed effects absorb any observable and unobservable differences in tract characteristics over time, including applicants' 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 only within-tract variation and compare applications to different lenders from individuals in the same tract and the same year. In addition, we include lender type*time fixed effects to control for time-varying observed and unobserved heterogeneity in lender characteristics, for example funding conditions or regulation.

We find that our results are robust to the inclusion of granular time-varying fixed effects. Together with the absence of any differential pre-trends, this finding mitigates concerns that the CCPA was introduced because of the rise of fintech lenders in California

border counties, or that our findings reflect differences in applicant characteristics across or within tracts.

We investigate a number of alternative explanations for our findings. First, we address the concern that the decline in the interest rate offered by fintechs could be due to an improvement in the quality of fintech applicants within tracts. As applications increase, it could be that higher-quality applicants decide to apply more to fintechs, allowing them to offer lower rates irrespective of any change in their screening ability. Yet we find no significant change in various measures of the quality of applicants to fintechs vs. banks with the introduction of the CCPA. Further, directly controlling for the quality of the applicant pool leads only to a modest decline in the magnitude of the estimated effect of the CCPA on interest rates by fintechs.[5] Second, we control for the local severity of the Covid-19 pandemic. If movement restrictions limited applicants' ability to visit a bank branch, applications to fintechs could have increased. However, an increase in demand should have put upward pressure on interest rates charged by fintechs. We find that our main results remain qualitatively unaltered when we control for the severity of Covid-19. Moreover, controlling for the potential increase in demand and attendant upward pressure on rates during Covid-19 leads to a larger negative effect of the CCPA on interest rates charged by fintechs relative to banks.

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 applicants of age 62 and above from the sample, ie applicants that could have been more affected by Covid-19 related restrictions; and remain robust to the inclusion of a large set of tract-level applicant controls or when we estimate applicant-level regressions and directly control for applicants' characteristics. Finally, we obtain similar results when we restrict

---

[5]The absence of selection effects is consistent with Armantier et al. (2023), who in a representative survey of U.S. consumers find no systematic correlation between individuals' relative distrust in fintechs and characteristics reflecting default risk, such as income, education, or their self-assessed credit score.

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 observable and unobservable characteristics from fintech loans.

Our findings 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 for example reduce the need for collateral (Gambacorta et al., 2022) or promote financial inclusion and entrepreneurship (Philippon, 2020; Ahnert et al., 2022). At the same time, consumers value their privacy and are concerned about the ab- and misuse of data (Armantier et al., 2021; Tang, 2022; Lin, 2022). These considerations pose a trade-off for policy makers, which need to balance improving efficiency through greater use of data with protecting users' right to privacy (Acquisti et al., 2016). Our results suggest that privacy protection legislation that enhances users' control over data and increases transparency and accountability in their use can mitigate this trade-off, at least in the mortgage market. As the CCPA makes users more willing to share data, it enables technology-savvy lenders to better screen with data and offer lower rates, enhancing the scope of financial services.

Literature and contribution. Our paper provides the first empirical evidence on how the CCPA, one of the world's most comprehensive privacy laws, affects fintech lending. It speaks to the growing literature studying the consequences of data sharing policies for loan markets and firms.

One strand of the literature focuses 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' access to relevant information, bankruptcy flag removal can have large distributional effects across

7

borrowers and create both winners and losers, with ambiguous welfare effects (Liberman et al., 2019; Jansen et al., 2022). Similarly, the U.S. Card Act, which limited credit card lenders' discretion to adjust interest rates in response to new information, reduced prices for high-risk consumers but increased them for others (Nelson, 2018). Europe's General Data Protection Regulation (GDPR), which follows the key principle that firms need to minimize their data collection and processing activities, has decreased venture investments in data-related firms (Jia et al., 2018).[6] Moreover, 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).

Other papers study the consequences of open banking, which allows customers to share their bank account history with third parties such as fintech lenders, but does not target data privacy per se. Improving fintechs' access to previously unavailable data can result in better screening and loan market outcomes, especially for riskier borrowers (Nam, 2022). Yet open banking can 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, resulting in lower borrower welfare (He et al., 2023). It 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). Finally, Babina et al. (2022) 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 examines the effects of the CCPA, ie privacy regulation that grants users control over their data and mitigates privacy concerns by explicitly considering differences in privacy preferences. The setting thus differs from studies on the effects of policies

---

[6]A growing literature investigates the effects of the GDPR on firms and finds that it hurts firm performance (see Johnson (2022) for a survey).

that limit lenders' information set or have the goal to facilitate data sharing. Our results suggest that privacy legislation in the spirit of the CCPA can make applicants more willing to share their data and, in the presence of lenders with superior screening technology, lead to improved loan market outcomes.[7]

Finally, our paper relates to work studying the rise of fintechs and the attendant effects on banks. Important drivers behind fintechs' rising footprint in the U.S. mortgage market are an increased regulatory burden on traditional banks and fintechs' superior technology with faster processing times (Buchak et al., 2018; Fuster et al., 2019). In addition, better access to data, notably from payments, can help fintechs compete with banks. 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.[8] Parlour et al. (2022) show that fintechs competing for payments can disrupt information spillovers from banks' payments to their lending services. Most of the literature studying the rise of fintechs and their competition with banks has not explicitly considered the role of privacy preferences or data protection legislation (see Berg et al. (2022) for a survey). We find that in the presence of privacy-sensitive consumers, privacy legislation can spur the growth of fintechs, which could increase the competitive pressure for banks.

The remainder of the paper is organized as follows. Section 2 discusses the evidence on borrowers' preferences for privacy and provides the institutional background on the CCPA. Section 3 lays out a conceptual framework to derive a set of testable hypotheses on the effects of privacy regulation on banks and fintechs. Section 4 tests the hypotheses with U.S. mortgage data, exploiting the introduction of the CCPA. Section 5 concludes.

---

[7]These results are consistent with Ali et al. (2022), who theoretically show that consumer control over data can improve consumer welfare.

[8]Other papers look at data availability and firm dynamics (Begenau et al., 2018; Farboodi et al., 2019; Canayaz et al., 2022) or the growth of fintechs more broadly (Cornelli et al., 2021).

2 Evidence on privacy preferences and the CCPA

This section first discusses evidence on consumers' privacy preferences and their willingness to share data. It then provides institutional background on the California Consumer Privacy Act, a comprehensive data privacy law that set the standard for privacy legislation across the U.S.

Privacy preferences. As more and more economic activity is moving online, personal data is turning into an increasingly important asset for firms (Acquisti et al., 2016; Jones and Tonetti, 2020). When screening applicants, fintech lenders in particular increasingly rely on personal data to complement traditional credit scores (Jagtiani and Lemieux, 2019; Berg et al., 2022). Such personal data can be directly collected, eg through loan applications, but also purchased from data aggregators and vendors. They can include information ranging from education and employment history (Di Maggio et al., 2022) to social network activity or digital footprints (Berg et al., 2020).

However, there is mounting evidence that consumers value their privacy (Goldfarb and Tucker, 2012). For example, Tang (2022) finds that consumers derive utility from withholding information when applying for a loan. Lin (2022) further shows that a preference for privacy plays an important role for users' decision to share data. Indeed, in a representative survey of U.S. households, around three-quarters of respondents were very concerned about negative consequences when sharing their personal data (see Figure 2). Reasons mentioned by survey respondents include identity theft, personal safety, and reputational concerns (Armantier et al., 2021), which are also commonly reflected in financial fraud complaints (Bian et al., 2023).

Since financial intermediaries make extensive use of consumers' data, an important question is in which counterparties consumers have greater confidence to safely handle

their data and protect them from data ab- or misuse. Figure 1 shows the level of trust expressed by a sample of U.S. households for traditional financial intermediaries (FI), fintechs, and bigtechs. A clear pattern emerges: almost two-thirds of respondents state that they have high trust in traditional FIs, as opposed to only 35% placing high trust in fintechs and 10% in bigtechs. Chen et al. (2023) report a similar pattern for a large sample of countries: 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 that they have higher confidence in traditional banks, relative to fintechs, to safeguard their personal data.

The California Consumer Privacy Act. The CCPA is a data privacy law covering the state of California that went into effect in January 2020. It endows Californians with several rights regarding the personal information that a firm may collect about them. In particular, they 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 its sale (Camhi and Lyon, 2018). The CCPA applies to all data of California residents, irrespective of whether the firm they transact with is based in California or outside of the state. By now, many other states consider introducing legislation in the spirit of the CCPA, and U.S. Congress considers it as the blueprint for a federal privacy protection regulation (Mulligan et al., 2019).[9]

The Office of the Attorney General (OAG) in California monitors compliance with the CCPA via enforcement sweeps. If the sweeps uncover practices that go against the CCPA, the OAG sends the company a notice of alleged noncompliance. The company

---

[9]In 2021 alone, lawmakers in twenty-seven U.S. states proposed CCPA-like privacy legislation (see 'Which States Will Consider CCPA-Like Consumer Privacy Bills in 2022?' and 'CCPA: Congress' Model for Data Privacy - or Oblivion?'.

then has 30 days to respond. Violations of the Act entail hefty fines, and there are active cases against Amazon, Zoom, or TikTok, among others.[10]

By granting consumers control over their data, the CCPA directly addresses several of the concerns that individuals list when it comes to sharing their data. As shown in Figure 2, 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 decreases the uncertainty around the use of personal information. As highlighted in an impact assessment of the CCPA by the California Department of Justice, if the CCPA increases consumers' trust of data protections it could increase the amount of data that consumers are willing to share with firms.[11] Recent survey evidence confirms that the CCPA makes consumers more willing to share their data (Armantier et al., 2023).

The CCPA is expected to have a stronger impact on applicants' attitudes towards sharing data with fintechs compared to traditional financial intermediaries. The reason is that, absent regulation, users have significantly lower confidence in fintechs than in banks to safely store their data and prevent data abuse (see Figure 1). 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. Moreover, banks oftentimes have long-lasting relationships with clients, which mitigate concerns that data are misused. The marginal benefit of privacy regulation is hence likely higher among newer intermediaries such as fintechs, compared to traditional financial institutions.

In granting users control over their data, the CCPA differs from other data initiatives and regulations in important aspects. For example, a key principle of the GDPR is

---

[10]See: Data Grail: The Biggest GDPR & CCPA Fines, Analyzed. Appendix Section B.1.3 provides more information on the CCPA and its enforcement.

[11]See Standardized Regulatory Impact Assessment: California Consumer Privacy Act of 2018 Regulations.

that firms need to minimize their data processing activities. Similar to policies such as bankruptcy flag removal, it hence effectively limits firms' ability to exploit information (Liberman et al., 2019; Jansen et al., 2022; Johnson, 2022). Open banking, on the other hand, mandates financial institutions to share their proprietary data with third parties, including fintech lenders, if users give their consent. As argued, the requirement to share data can benefit some borrowers, but also discourage information acquisition (Babina et al., 2022), increase market concentration (He et al., 2023), or distort banks' credit allocation (Goldstein et al., 2022).

Were California residents aware of the introduction of the CCPA? According to a survey by the Consumer Action and Consumer Federation of America (2021), around 70% of those interviewed had seen the notice of their rights required by the CCPA on websites they visited.[12] Moreover, the majority of respondents had exercised their rights granted by the CCPA. For example, over 50% have asked firms to not share or delete their data. 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.[13] 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 confidence that their data will not be used for unintended purposes. Accordingly, it makes consumers more willing to share data, and in particular with fintech lenders (Armantier et al., 2023).

---

[12]See the Consumer Action and Consumer Federation of America (2021): California Consumer Privacy Act (CCPA) Survey. The survey queried 1,500 adults in California about their awareness of and experience with the CCPA.

[13]See the screenshots of mortgage lenders' websites in Figure OA2 and Figure OA3.

3 Conceptual framework

To guide our empirical analysis, we present a simple conceptual framework on how privacy protection regulation affects loan markets in the presence of privacy concerns among applicants.[14]

Suppose two types of lenders – banks and fintechs – compete for privacy-conscious consumers. Before offering a loan, lenders must analyze personal data to screen applicants. These data can be standard, such as an individual's credit score or proof of income. But in some cases, lenders might request or applicants provide additional information, such as their education history, phone or electricity bills. 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 (Di Maggio et al., 2022; He et al., 2023).

Consumers are concerned about the ab- or misuse of data when applying for a loan. Importantly, they are more concerned about sharing their data with fintechs than with banks. These assumptions follow directly from Figure 1 and Figure 2. Some privacy-conscious individuals might hence be unwilling to contract with a fintech even if it offers a lower interest rate than banks.

We then consider the introduction of a CCPA-like privacy legislation, which provides consumers with greater control over their data. Such legislation reduces concerns about the abuse and misuse of data and hence decreases consumers' dislike to sharing them. This decline is more pronounced for sharing data with fintechs compared to banks, an assumption consistent with the evidence that individuals generally have significantly lower confidence in fintechs to safely handle personal data (Armantier et al., 2021; Chen et al., 2023).

---

[14]The framework is based on a parsimonious model of the loan market with banks and a fintech that use data to screen applicants. The fintech has a superior screening technology, but applicants dislike sharing their data with it. For details, see the Online Appendix B.2.

By reducing applicants' disutility from sharing data with fintechs, privacy regulation is expected to increase applications to fintechs compared to banks:

Hypothesis 1: The introduction of privacy protection legislation leads to an increase in loan applications to fintechs compared to banks.

The impact of privacy regulation on interest rates is a priori unclear. On the one hand, an increase in applications could increase the price of fintech credit and lead to higher interest rates (demand effect). On the other hand, applicants' greater willingness to share data could enable fintechs to use additional data to improve their screening process (Berg et al., 2020; Di Maggio et al., 2022), filtering out more low-quality applicants and offering lower interest rates (screening effect). If the screening effect dominates, the following is true:

Hypothesis 2: The introduction of privacy protection legislation decreases loan rates on loans originated by fintechs compared to banks if the screening effect dominates the demand effect.

A relative decline in fintechs' interest rates hence implies that fintechs see an improvement in their screening process. A more precise signal about borrowers' quality is commonly associated with more individualized pricing across borrowers (Babina et al., 2022; Jansen et al., 2022), implying greater rate dispersion. In addition, it would allow fintechs to better screen out low-quality applicants, which should lead to an increase in the share of denied applications:

Hypothesis 3.1: By enabling better screening through data, privacy regulation increases the dispersion in interest rates across borrowers among fintechs relative to banks.

Hypothesis 3.2: By enabling better screening through data, privacy regulation increases the share of denied loan applications by fintechs relative to banks.

Finally, applicants' greater willingness to share data with fintechs implies that fintechs will increase their use of data beyond traditional credit scoring models:

Hypothesis 3.3: After the introduction of privacy regulation, fintechs' loan applications processed with non-traditional data increase by more than banks.

In our context, data can be understood in a number of ways. First, applicants could provide more information in their application. For example, lenders typically leave applicants the choice to provide one or more documents as proof of income, education history, or employment history. After the introduction of privacy regulation, applicants might be more willing to provide such additional information beyond what is strictly required, which could benefit fintechs' screening process (Di Maggio et al., 2022).[15] Second, lenders might become more willing to use alternative data from other sources. A number of high-profile cases, such as the Facebook–Cambridge Analytica data scandal, have led to high reputational costs for data-intensive firms. In consequence, some fintechs have become reluctant to use third-party data providers without users' consent.[16] Under the CCPA's regulatory umbrella, fintechs might become more willing to use such outside data. And third, as more individuals apply to fintechs, each fintech benefits from an increase in the aggregate amount of data it has on applicants. It hence could have a more comprehensive picture of the state of the local economy and more accurately infer other applicants' creditworthiness (Ichihashi, 2021; Bergemann et al., 2022).

---

[15]In studying a major fintech platform, Di Maggio et al. (2022) show that the main variables accounting for improved default predictability include non-traditional information, such as education and employment history, which supplement traditional credit report variables.

[16]See, for example, Fintech's dirty little secret? Lenddo, Facebook and the challenge of identity and What Specific Challenges Are Fintech CTOs Worried About in 2022?

4 The CCPA and fintech lenders in the U.S. mortgage market

This section exploits the introduction of the California Consumer Privacy Act in 2020 to test the hypotheses developed in Section 3. 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' individualized pricing, application denial rates, and use of alternative credit scores.

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, and gender, among other items. To classify lenders in HMDA as banks or fintechs we follow Fuster et al. (2019), who classify an originator as a fintech lender if they enable a mortgage applicant to obtain a preapproval online.[17]

We collapse the individual applicant data at the lender–applicant tract–year level. In our analysis, we use mortgage applications in counties that lie on the border of California (CA) and its neighboring states Arizona (AZ), Nevada (NV), and Oregon (OR). The sample period covers the years from 2018, the first year for which data on interest rates are available, to including 2021. To eliminate noise stemming from tracts with insignificant

---

[17]See Fuster et al. (2019) for a detailed discussion. We also 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.

amounts of loan applications by a given lender, we restrict our sample to markets where a given lender made at least two loans in each year.

The main outcome variables are the log of the number of applications and the average interest rate on approved mortgages. In addition, we compute the share of denied applications, the dispersion in interest rates across approved mortgages, as well as the share of mortgages that do not use standardized underwriting models.[18]

Descriptive statistics. Our final sample contains 900,270 mortgage applications between 2018 and 2021 in 9,723 census tracts in border counties. Collapsed to the lender–applicant tract–year level, we end up with 75,354 observations. Table 1 provides summary statistics for our main outcome variables in 2018 and 2019, ie prior to the introduction of the CCPA. The average lender–tract cell had 5.6 applications. The interest rate charged in the average tract was 4.4%, with a standard deviation of 0.63; the share of denied applications equaled 5.2%, while the share of applications processed with data beyond traditional credit scores was 24%. In the average tract, out of all applications, a share of 16.1% was to fintechs, with a standard deviation of 11.3%.

Table 2 provides summary statistics for applicants to banks and fintechs prior to the introduction of the CCPA. In 2018–19, applicants are statistically similar in terms of the observable characteristics gender, race, income, the value of the property, loan-to-income ratios, and loan-to-value ratios. An exception is applicant age, with applicants of age 62 and above being more common among fintech lenders (consistent with Fuster et al. (2019)). Beyond statistical significance, most values are similar in terms of economic magnitude. Overall, these patterns suggest that bank and fintech applicants are economically and statistically comparable in border tracts. As we will further show below, the

---

[18]We compute the fraction of mortgages originated using a credit scoring model besides the standard ones Equifax, Experian, FICO, or Vantage Score.

quality of applicants to fintechs relative to banks within tracts does not change with the introduction of the CCPA.

4.2 Empirical strategy and results

In this section, we first test whether the introduction of the CCPA has lead to an increase in loan applications to fintechs compared to banks (Hypothesis 1). We then analyze the CCPA's effect on interest rates (Hypothesis 2). Finally, we investigate the underlying channel (Hypotheses 3.1–3.3).

We estimate variants of the following regression at the lender–tract–year level:

$$y_{l,c,t} = \delta_1 \ CA_c \times post_t + \delta_2 \ fintech_l \times post_t$$
$$+ \delta_3 \ CA_c \times fintech_l \times post_t + \theta_{l,c} + \tau_{c,t} + \phi_{l,t} + \varepsilon_{l,c,t}. \tag{1}$$

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–applicant tract level. We hence only exploit variation within each lender-tract cell. Standard errors are clustered at the tract level.[19]

Based on Hypothesis 1, we expect a coefficient of $\delta_3 > 0$, that is, the CCPA should increase consumers' willingness to apply to fintechs relative to banks. Hypothesis 2 instead suggests $\delta_3 < 0$ when the screening effect dominates the demand effect: The

---

[19]We show the robustness of our findings to different levels of clustering in the Online Appendix.

introduction of the CCPA is expected to lower rates on loans originated by fintechs relative to banks in California; if instead the demand effect dominates we expect the opposite sign.

Identification. Equation (1) 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. To address this challenge, we include granular time-varying fixed effects at the tract level ($\tau_{c,t}$). These fixed effects absorb any observable and unobservable differences in tract characteristics over time. They hence account for changes common to applicants and borrowers within a tract, such as changes in average income, borrower risk, internet access or demographic structure, and credit demand. They also control for potential differences in the severity of the Covid-19 pandemic and associated movement restrictions across tracts, an issue we revisit below. With tract*time fixed effects, we essentially compare applications to different lenders from individuals in the same tract and year. In addition, we include lender type*time fixed effects ($\phi_{l,t}$) to control for changes in observed and unobserved characteristics of each type of lender (ie banks and fintechs). These control for eg the effects of changes in the Fed funds rate on fintechs or banks, or changes in the regulatory environment.

As explained in Section 4.1, to further tighten identification we restrict the analysis to tracts within counties along the border of California with its neighboring states. Border counties are generally similar along many observable characteristics, mitigating concerns about selection effects and omitted variable bias (Allegretto et al., 2017).[20] Moreover, as we show in Table 2, the characteristics of applicants to fintechs and banks did not differ in an economically or statistically significant way prior to the introduction of the CCPA.

---

[20]We provide evidence consistent with these arguments in Table OA7.

4.2.1 The CCPA, loan applications, and rates

Columns (1)–(3) in Table 3 show that loan applications to fintechs increase in California border counties after the introduction of the CCPA. The Online Appendix reports the results for the full sample of tracts. Column (1) includes lender-tract as well as year fixed effects and shows that applications increase in California after the introduction of the CCPA ($\delta_1 > 0$ in Equation (1)). The CCPA hence has a positive effect on loan applications with banks, possibly by also increasing some consumers' willingness to share data with banks. Yet, applications increase by significantly more among fintechs compared to banks ($\delta_3 > 0$), in line with Hypothesis 1.

Column (2) controls for unobservable time-varying applicant tract characteristics by introducing tract*time fixed effects. Comparing lending by fintechs and banks to the same tracts leads to almost identical coefficient estimates. These results are consistent with the argument that applicants in tracts in border counties are comparable in terms of observable and unobservable characteristics. Finally, column (3) introduces lender type*time fixed effects to absorb any time-varying unobservable characteristics for each lender type. The coefficient on the triple interaction effect remains positive and significant at the 1% level. In terms of economic magnitude, applications to fintech lenders increase by 14.6% more than to banks in California after the introduction of the CCPA.

How does the increase in applications to fintechs translate into changes in their market share? In the Online Appendix we show that the share of applications to fintechs increases by 3 percentage points after the introduction of the CCPA (see Table OA1, column (4)), implying a sizeable increase of 19% of the pre-CCPA mean. These results suggest that the CCPA has spurred the growth of fintechs.

In columns (4)–(6) we analyze the effects of the CCPA on loan rates, ie Hypothesis 2. Column (4) shows that on average, loan rates in California increase by significantly

more after 2020 than in neighboring states. Fintechs in California, however, decrease their interest rate compared to banks. To ensure that the effects are not driven by unobservable time-varying tract or lender type characteristics, columns (5) and (6) tighten identification by adding tract*time and lender type*time fixed effects. Results show that, even after holding all observable and unobservable variation across time at the tract and lender type level constant, rates on fintech-approved mortgages in California decrease by an 7.9 bp (0.10 standard deviations) compared to banks.

In sum, Table 3 provides empirical support for Hypotheses 1 and 2. Applications to fintechs, relative to banks, increase by more in California after the introduction of the CCPA, compared to neighboring states. Loan rates on mortgages approved by fintechs decrease by relatively more. The decline in interest rates suggest that the screening effect dominates the demand effect among fintech lenders, which we will investigate in the next section.

Pre-trends. Were there any pre-trends across fintechs and banks in applications and interest rates? To this end, we estimate how applications and rates by fintechs in California change compared to banks in each year. Figure 3 plots the coefficient estimates with 90% confidence intervals from Equation (1), in which we replace the *post* dummy with dummies for each year in the sample. The omitted year is 2019, ie the year before the CCPA came into effect. Panel (a) shows that there is no discernible difference in the share of applications to fintechs prior to the introduction of the CCPA in 2020. While applications to fintechs and banks evolve similarly between 2018 and 2019, applications to fintechs increase by relatively more in 2020, ie when the CCPA came into effect. The gap persists 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 is also no significant difference in the evolution of interest rates prior to the introduction

22

of the CCPA. Yet rates decline by significantly more on mortgages approved by fintechs compared to banks after the introduction of the CCPA in 2020.

4.2.2 Testing the mechanism: rate dispersion, denial rates, and the use of non-traditional data

Privacy legislation in the spirit of the CCPA provides consumers with greater confidence that their data will not be used for unintended purposes, making them more willing to share information with lenders. As a consequence, fintechs could be able to engage in more individualized pricing (Hypothesis 3.1), better screen out low-quality applicants (Hypothesis 3.2), and increase the use of alternative data in their screening process (Hypothesis 3.3).

We test these hypotheses in Table 4, estimating variations of Equation (1). In columns (1)–(2) we test Hypothesis 3.1. The dependent variable is the standard deviation in interest rates across loans within each lender–tract–year cell. Consistent with obtaining a more precise signal from additional data, column (1) shows a significant and positive coefficient on the triple interaction term: the dispersion in interest rates increases by significantly more for fintechs relative to banks in California. Adding tract*time and lender type*time fixed effects in column (2) does not materially affect this conclusion.

To test Hypothesis 3.2, columns (3)–(4) analyze the effects of the CCPA on the share of denied loan applications within each lender–tract–year cell. Column (3) uses lender*tract and year fixed effects, while column (4) uses tract*time and lender type*time fixed effects. Across specifications, rejection rates by fintechs significantly increase after the introduction of the CCPA, relative to banks. Greater rate dispersion and an increase in application denial rates are consistent with improved screening through the use of more data.

23

Direct information on lenders' use of data in the screening process is difficult to obtain. To test Hypothesis 3.3, we thus focus on the share of mortgage applications processed with non-standardized credit scores. As argued in Babina et al. (2022), the use of non-standardized credit scores proxies lenders' reliance on alternative data. Column (5) shows that, conditional on lender*tract and year fixed effects, fintechs in California increase their use of non-traditional data beyond standardized credit scores after the introduction of the CCPA. Column (6) with tract*time and lender type*time fixed effects confirms this finding.

To further provide evidence on our proposed mechanism, we focus on 'thin credit file' applicants. Blattner and Nelson (2021) show that widely-used credit scores are noisier indicators of default risk for traditionally under-served groups (including minority and low-income applicants), which results mostly from thin credit files. Di Maggio et al. (2022), analyzing lending by a U.S. fintech that uses alternative data such as education and job history, confirm that current credit scores do not paint an accurate picture of the future creditworthiness of applicants, in particular for those with short or thin credit histories. We thus expect that the CCPA, by increasing the amount of data available to fintechs, lowers the interest rate on loans by fintechs especially among thin credit file applicants.

Table 5 estimates Equation (1), but focuses on tracts with a plausibly higher share of thin file applicants. Specifically, we focus on tracts with a high share of minority applicants (columns 1–2), a low tract-to-MSA income ratio (columns 3–4), or lower average applicant income (columns 5–6). For each measure, we focus on tracts below the median and in the bottom quartile of the distribution. Consistent with Blattner and Nelson (2021) and Di Maggio et al. (2022), the results show that the negative effect of the CCPA on rates by fintechs is significantly stronger in tracts with a higher share of thin file applicants compared to its baseline estimate of 7.9 bp.

These findings could suggest that the introduction of the CCPA improved access to credit. In Table OA1 in the Online Appendix, we show that banks, and even more so fintechs, originate more mortgages after the introduction of the CCPA – despite the increase in denial rates. This pattern suggests that the CCPA did not just lower the cost of credit, but has led to an expansion of credit on aggregate.

All in all, results in Table 4 and Table 5 are consistent with the channel in our conceptual framework: with the introduction of the CCPA, fintechs' screening ability improves thanks to individuals' greater willingness to share data. This results in more individualized pricing and a greater share of denied applications. In consequence, fintechs can offer lower loan rates compared to banks, especially to thin credit file applicants.

4.3 Alternative explanations and robustness tests

This section provides a series of tests to examine alternative channels and the robustness of our findings.

Quality of the applicant pool. A possibly confounding factor is that the quality of the pool of *applicants* changes as the number of applications increases, thereby affecting the interest rate. For example, fintechs could have a larger footprint in tracts with higher-quality applicants on average. An increase in applications from these tracts would then improve the quality of fintechs' applicant (and hence borrower) pool, so that they can offer lower rates. In our regressions, any such differences in the average quality of applicants or borrowers across tracts are absorbed by tract*time fixed effects. A second possibility, however, is that even within a census tract, as applications increase, higher-quality applicants decide to apply more to fintechs. Tract*time fixed effects would not absorb this variation.

The public HMDA data does not provide applicants' credit score, the most common measure of applicant quality. Instead, we must rely on proxy variables, a caveat that should be kept in mind when interpreting our results. Specifically, we use the log of applicant income, the log of the loan-to-income (LTI) ratio, the log of the loan-to-value (LTV) ratio, and a dummy that takes on a value of one if the debt-to-income (DTI) ratio is below 36%.[21] Finally, we compute a composite measure of applicant quality ('risk PCA') by taking the first principle component of log income, the log of the LTI ratio, and the log of the LTV ratio. It explains a sizeable 52% of the variances of the variables.

Table 6 estimates variants of Equation (1) and shows that there was no statistically or economically significant change in applicant quality among fintechs vs. banks when the CCPA was introduced. In particular, for each measure of quality, which we use as dependent variables in columns (1)–(5), the coefficient on the triple interaction term is statistically and economically insignificant. In columns (6)–(8) we directly control for the quality of the applicant pool in our regressions by including the mean and standard deviation of each risk measure in each lender-tract-year cell as controls. Results show only a modest decline in the magnitude of the estimated effect of the CCPA on interest rates by around 9% (or 1.1 bp from column (6) to column (8)). The (insignificant) change in the average quality of applicants across lenders within tracts hence explains only a small fraction of the magnitude of our estimated effects of the CCPA on interest rates.

All in all, results in Table 6 suggest that there was no systematic selection of higher-quality applicants with fintechs. The absence of selection effects is consistent with Ar-

---

[21]Albanesi et al. (2022) show a positive correlation between income and credit scores. Fuster et al. (2021) show that the loan-to-income ratio is highly correlated with ex-post default, making loan-to-income ratios a standard measure of risk (Mayer et al., 2009). The cut-off value of 36% for the debt-to-income ratio, ie that household spend less than 36% of their gross monthly income on total debt service, is generally considered as a positive factor in assessing household risk in mortgage applications. To eliminate extreme values, we winsorize each continuous variable at the 5th and 95th percentile in each year. Since information on LTI, DTI, and LTV ratios is not available for all applications, the sample size of our lender-tract-year sample declines from around 75,000 to around 68,000.

mantier et al. (2023), who show that in a representative survey of U.S. consumers there is no correlation between individuals' relative distrust in fintechs and characteristics such as income, education, or their self-assessed credit score.

Covid-19 and demand effects. The outbreak of Covid-19 in 2020 led to severe restrictions on movements. Voluntary restrictions or lockdown measures could have increased applications to fintechs, as applicants were unable to visit bank branches. Such an increase in demand could put upward pressure on interest rates charged by fintechs. To account for the severity of Covid-19 at the local level and its impact on loan applications and loan rates, we use four measures that vary at the county level: workplace mobility and transit mobility (both from Google Mobility Trends), as well as Covid cases and deaths per capita. We create time-varying measures by setting the values for these variables to zero in 2018 and 2019, and to their actual realization (averaged across days) in each county in 2020 and 2021. We then interact these measures, which vary at the county-year level, with the *fintech* dummy and include them as controls in Equation (1).

Table 7 shows that controlling for the severity of the Covid-19 pandemic at the local level does not materially affect our estimates. For applications (columns 1–3) we still obtain a highly significant positive coefficient on the triple interaction term $CA \times fintech \times post$. For rates (columns 4–6), the coefficient remains negative and significant, and increases in absolute value. In other words, controlling for the potential demand-induced upward pressure on rates due to Covid-19 leads to a larger negative effect of the CCPA on interest rates charged by fintechs relative to banks. These results suggest that the positive (negative) impact of the CCPA on applications (rates) among fintechs is not explained by differences in the local severity of the Covid-19 pandemic and associated restrictions on movements.

Further robustness tests. In Table 8, we show that applications to fintechs increase and rates decline both among purchase as well as refinance mortgages, which addresses concerns about second homes or out-of-state applicants. Results remain unaffected when we exclude applicants of age 62 and above from the sample, ie applicants 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 average 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 (Buchak et al., 2018). 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, including in terms of compliance costs (Gupta et al., 2023). However, we find that also among sold mortgages applications increase and loan rates fall for fintechs relative to banks in California after the introduction of the CCPA.

The Online Appendix provides additional tests. We confirm our results when we directly compare applications and loan rates for fintechs vs banks in California only (Table OA2). We also obtain similar results when we run the baseline specification leaving one border state out at a time (Table OA3). These exercises ensure that our results are not spuriously driven by movements in the control group. We also ensure that our result are robust to different levels of clustering (Table OA4); and when we include all tracts beyond just those in border counties (Table OA5). Finally, we show that our main results also hold in applicant-level regressions, in which we control for applicants' LTI ratio, LTV ratio, log income, age, sex, race, ethnicity, and debt to income ratio, as well

as the log of the loan amount, log of the property value, a dummy for whether the loan was refinanced or not, and a dummy for whether the loan was sold or not (Table OA6).

5   Conclusion

Individuals are constantly generating a wealth of personal data. They leave a digital footprint every time they purchase a good or service online, use a cell phone, or surf the internet. These data contain valuable information about users' creditworthiness and their ability to repay. Financial intermediaries, and in particular fintechs with superior data-processing technologies, can hence benefit from access to data. Consumers, however, are increasingly concerned that they have lost control over who knows what about them, and worry about the misuse of the information that intermediaries collect about them. These concerns can deter individuals from sharing personal data, hence slowing down the growth of innovative financial companies, and eventually limiting the benefits these innovations could bring to consumers. This poses a trade-off to regulators. Effective privacy regulation must strike a balance between protecting users' privacy without negating the benefits obtained from intermediaries' access to data.

We study whether the 2020 California Consumer Privacy Act is successful in achieving this goal in the mortgage market. The CCPA's philosophy is to grant consumers control over their data, thereby mitigating their privacy concerns. In turn, it makes users more willing to share their information. Comparing fintechs to banks in counties along both sides of the California border in a difference-in-differences setting, we find that granting consumers more control over their data encourages them to apply to fintechs: The introduction of the CCPA increases fintechs' market share by 19%. We present evidence that is consistent with fintechs' greater use of alternative data in processing applications, which improves their screening ability and leads to lower rates for consumers.

Our results have implications for the policy debate on how to regulate the use of personal data. While personal data can improve lenders' screening abilities, users value their privacy, generating potentially conflicting regulatory goals. Our results suggest that a privacy protection legislation that enhances users' control over data can achieve both objectives. The CCPA can hence be seen as a successful regulatory initiative that holds important lessons for other U.S. states and countries designing and implementing privacy legislation.

References

Acquisti, A., C. Taylor, and L. Wagman (2016) "The economics of privacy", *Journal of Economic Literature*, 54 (2), pp. 442–92.

Ahnert, T., S. Doerr, N. Pierri, and Y. Timmer (2022) "Does IT help? Information technology in banking and entrepreneurship", *Working Paper*.

Albanesi, S., G. DeGiorgi, and J. Nosal (2022) "Credit growth and the financial crisis: A new narrative", *Journal of Monetary Economics*, 132, pp. 118–139.

Ali, S. N., G. Lewis, and S. Vasserman (2022) "Voluntary Disclosure and Personalized Pricing", *The Review of Economic Studies*.

Allegretto, S., A. Dube, M. Reich, and B. Zipperer (2017) "Credible research designs for minimum wage studies: A response to Neumark, Salas, and Wascher", *ILR Review*, 70 (3), pp. 559–592.

Aridor, G., Y.-K. Che, and T. Salz (2022) "The effect of privacy regulation on the data industry: Empirical evidence from GDPR", *Working Paper*.

Armantier, O., S. Doerr, J. Frost, A. Fuster, and K. Shue (2021) "Whom do consumers trust with their data? US survey evidence", *BIS Bulletin* (42).

Armantier, O., S. Doerr, J. Frost, A. Fuster, and K. Shue (2023) "Nothing to hide? Survey evidence on who shares their day", *mimeo*.

Babina, T., G. Buchak, and W. Gornall (2022) "Customer data access and fintech entry: Early evidence from Open Banking", *Working Paper*.

Begenau, J., M. Farboodi, and L. Veldkamp (2018) "Big data in finance and the growth of large firms", *Journal of Monetary Economics*, 97, pp. 71–87.

Berg, T., V. Burg, A. Gombović, and M. Puri (2020) "On the rise of fintechs: Credit scoring using digital footprints", *The Review of Financial Studies*, 33 (7), pp. 2845–2897.

Berg, T., A. Fuster, and M. Puri (2022) "Fintech lending", *Annual Review of Financial Economics*, 14, pp. 187–207.

Bergemann, D., A. Bonatti, and T. Gan (2022) "The economics of social data", *The RAND Journal of Economics*, 53 (2), pp. 263–296.

Bhutta, N., A. Hizmo, and D. Ringo (2021) "How much does racial bias affect mortgage lending? evidence from human and algorithmic credit decisions".

Bian, B., M. Pagel, and H. Tang (2023) "Consumer surveillance and financial fraud", *mimeo.*

Blattner, L. and S. Nelson (2021) "How costly is noise? data and disparities in consumer credit", *Working Paper.*

Buchak, G., G. Matvos, T. Piskorski, and A. Seru (2018) "Fintech, regulatory arbitrage, and the rise of shadow banks", *Journal of financial economics*, 130 (3), pp. 453–483.

Burke, J., C. Taylor, and L. Wagman (2012) "Information acquisition in competitive markets: An application to the US mortgage market", *American Economic Journal: Microeconomics*, 4 (4), pp. 65–106.

Camhi, R. and S. Lyon (2018) "What is the California Consumer Privacy Act?", *Risk Management*, 65 (9).

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

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

Chen, S., S. Doerr, J. Frost, L. Gambacorta, and H. S. Shin (2023) "The fintech gender gap", *Journal of Financial Intermediation*, 54 (101026).

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

Cornelli, G., S. Doerr, L. Franco, and J. Frost (2021) "Funding for fintechs: patterns and drivers", *BIS Quarterly Review*, September.

Cornelli, G., J. Frost, L. Gambacorta, R. Rau, R. Wardrop, and T. Ziegler (2020) "Fintech and big tech credit: a new database", *BIS Working Papers* (887).

Di Maggio, M., D. Ratnadiwakara, and D. Carmichael (2022) "Invisible primes: Fintech lending with alternative data", *NBER Working Paper*.

Dobbie, W., P. Goldsmith-Pinkham, N. Mahoney, and J. Song (2020) "Bad credit, no problem? credit and labor market consequences of bad credit reports", *The Journal of Finance*, 75 (5), pp. 2377–2419.

Farboodi, M., R. Mihet, T. Philippon, and L. Veldkamp (2019) "Big data and firm dynamics", *AEA Papers and Proceedings*, 109, pp. 38–42.

Fuster, A., P. Goldsmith-Pinkham, T. Ramadorai, and A. Walther (2022) "Predictably unequal? the effects of machine learning on credit markets", *The Journal of Finance*, 77 (1), pp. 5–47.

Fuster, A., M. C. Plosser, and J. I. Vickery (2021) "Does CFPB oversight crimp credit?", *CEPR Discussion Paper*.

Fuster, A., M. Plosser, P. Schnabl, and J. Vickery (2019) "The role of technology in mortgage lending", *The Review of Financial Studies*, 32 (5), pp. 1854–1899.

Gambacorta, L., Y. Huang, Z. Li, H. Qiu, and S. Chen (2022) "Data vs collateral", *Review of Finance*.

Ghosh, P., B. Vallee, and Y. Zeng (2021) "Fintech lending and cashless payments".

Goldfarb, A. and C. Tucker (2012) "Shifts in privacy concerns", *American Economic Review*, 102 (3), pp. 349–53.

Goldstein, I., C. Huang, and L. Yang (2022) "Open Banking under maturity transformation", *Working Paper*.

Gupta, M., D. McGowan, and S. Ongena (2023) "The cost of privacy: The impact of the California Consumer Protection Act on mortgage markets.", *Working Paper*.

He, Z., J. Huang, and J. Zhou (2023) "Open Banking: Credit market competition when borrowers own the data", *Journal of Financial Economics*, 147 (2), pp. 449–474.

Ichihashi, S. (2021) "The economics of data externalities", *Journal of Economic Theory*, 196, p. 105316.

Jagtiani, J. and C. Lemieux (2019) "The roles of alternative data and machine learning in fintech lending: evidence from the LendingClub consumer platform", *Financial Management*, 48 (4), pp. 1009–1029.

Jansen, M., F. Nagel, C. Yannelis, and A. L. Zhang (2022) "Data and Welfare in Credit Markets", Working Paper 30235, National Bureau of Economic Research.

Jia, J., G. Z. Jin, and L. Wagman (2018) "The short-run effects of GDPR on technology venture investment", Working Paper 25248, National Bureau of Economic Research.

Johnson, G. (2022) "Economic research on privacy regulation: lessons from the GDPR and beyond", Working Paper 30705, National Bureau of Economic Research.

Jones, C. and C. Tonetti (2020) "Nonrivalry and the economics of data", *American Economic Review*, 110 (9), pp. 2819–58.

Kim, J.-H. and L. Wagman (2015) "Screening incentives and privacy protection in financial markets: a theoretical and empirical analysis", *The RAND Journal of Economics*, 46 (1), pp. 1–22.

Liberman, A., C. Neilson, L. Opazo, and S. Zimmerman (2019) "The equilibrium effects of information deletion: Evidence from consumer credit markets", *NBER Working Paper*.

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

Mayer, C., K. Pence, and S. M. Sherlund (2009) "The rise in mortgage defaults", *Journal of Economic Perspectives*, 23 (1), pp. 27–50.

Mulligan, S., W. Freeman, and C. Linebaugh (2019) "Data protection law: An overview", CRS Report R45631, Congressional Research Service.

Nam, R. J. (2022) "Open Banking and customer data sharing: Implications for fintech borrowers", *Working Paper*.

Navaretti, G. B., G. Calzolari, J. M. Mansilla-Fernandez, and A. F. Pozzolo (2018) "Fintech and banking. friends or foes?".

Nelson, S. (2018) "Private information and price regulation in the US credit card market", *Working Paper*.

Parlour, C., U. Rajan, and H. Zhu (2022) "When fintech competes for payment flows", *Review of Financial Studies*, 35 (11), pp. 4985–5024.

Philippon, T. (2020) "On fintech and financial inclusion", *BIS Working Paper*, 841.

Prince, J. and S. Wallsten (2022) "How much is privacy worth around the world and across platforms?", *Journal of Economics & Management Strategy*, 31 (4), pp. 841–861.
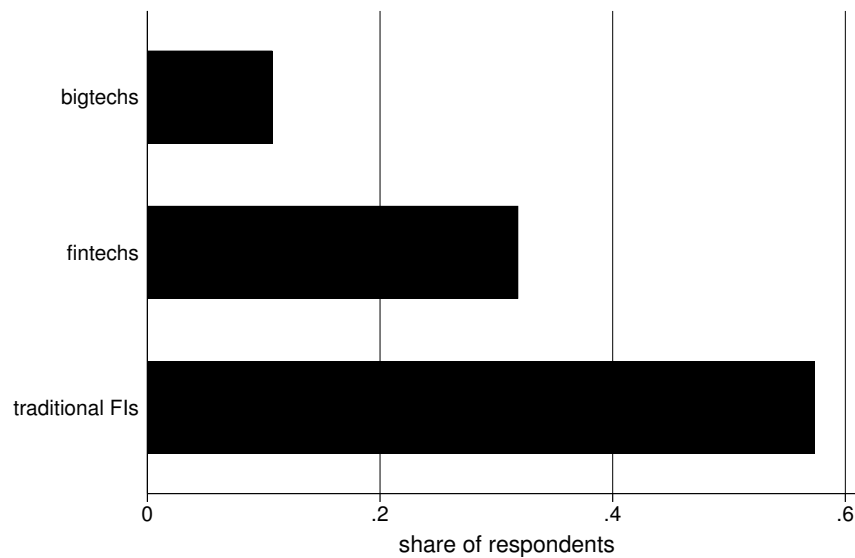
Tang, H. (2022) "The value of privacy: Evidence from online borrowers", *Working Paper*.

Thisse, J.-F. and X. Vives (1988) "On the strategic choice of spatial price policy", *American Economic Review*, pp. 122–137.

Vives, X. and Z. Ye (2022) "Fintech entry, lending market competition and welfare".

A Figures and tables

Figure 1: **Consumers' trust in counterparties to safeguard their data**

This figure shows the share of respondents reporting "high trust" in counterparties, based on a representative sample of 1,361 U.S. households in September 2020 that were part of the Survey of Consumer Expectations (SCE) of the Federal Reserve Bank of New York. Respondents place 'high trust' with a counterparty if they assigned a score of 5 or higher to the question "How much do you trust the following entities to safely store your personal data (that is, your bank transaction history, geolocation or social media data)?", on a scale from 1 (no trust at all in ability to safely store personal data) to 7 (complete trust). Source: SCE and Armantier et al. (2021).

Figure 2: **Consumers are concerned about sharing their data**
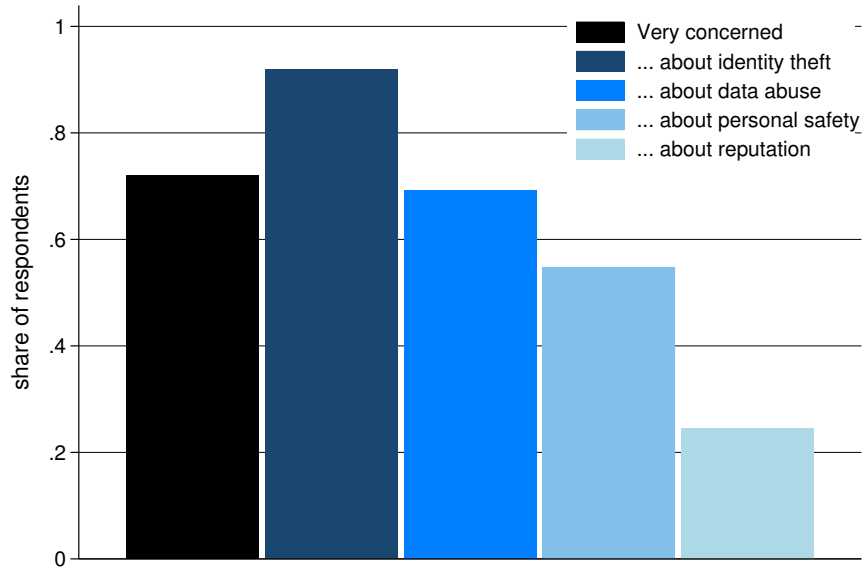
This figure shows concerns about sharing data online, based on a representative sample of 1,361 U.S. households in September 2020 that were part of the Survey of Consumer Expectations (SCE) of the Federal Reserve Bank of New York. Respondents are 'very concerned' about sharing their data online when 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. Respondents could answer yes to more than one option. Source: SCE and Armantier et al. (2021).

Figure 3: **Pre-trends – coefficient estimates**

(a) Loan applications

(b) Interest rates
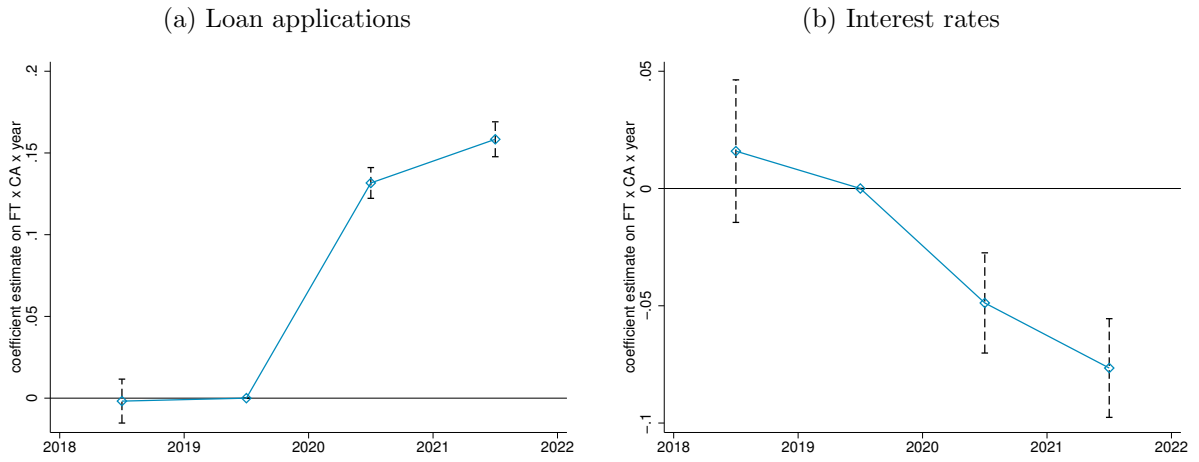
This figure plots the coefficient estimates (blue line) and 95% confidence intervals (gray bars) of the estimated coefficients on the triple interaction term obtained from Equation (1), but with dummies for each year in the sample instead of the *post* dummy. Coefficient β_k indicates the evolution of applications to or interest rates of loans by fintechs in year k before/after the introduction of the CCPA in California. The year prior to the introduction of the CCPA ($k = 2019$) is the omitted category. Panel (a) shows the coefficient estimates for loan applications, panel (b) for interest rates on approved mortgages. All regressions include tract*year fixed effects. Standard errors are clustered at the tract level.

Table 1: **Summary statistics**

Variable	Obs	Mean	Std. Dev.	Min	Max
applications	29215	5.648	6.999	2	181
log(applications)	29215	1.436	.675	.693	5.198
interest rate	29215	4.395	.63	1.875	6.16
other CS model	29215	.24	.346	0	1
sd(int rate)	29215	.515	.437	0	2.21
share denied	29215	.052	.115	0	.818

This table reports summary statistics for the main variables at the lender–tract–year level for the years 2018 and 2019, ie before the introduction of the CCPA. The interest rate is multiplied by 100.

Table 2: **Balancedness in applicant characteristics**

	banks		fintechs		mean diff.
	mean	sd	mean	sd	t
share female	0.25	(0.26)	0.25	(0.26)	-1.19
share Black or African Am.	0.03	(0.11)	0.03	(0.11)	-0.29
income (in USD th)	123.11	(1528.74)	104.14	(45.37)	0.83
property value (in USD th)	383.67	(133.77)	386.14	(140.66)	-1.12
loan-to-income ratio	0.03	(0.01)	0.03	(0.01)	0.19
loan-to-value ratio	0.77	(0.14)	0.77	(0.11)	-1.43
share age 62+	0.17	(0.25)	0.20	(0.26)	-7.00
N	24,789		4,426		29,215

This table reports summary statistics for applicant characteristics in 2018 and 2019 at the lender–tract–year level. The sample is split into banks and fintechs. The column *mean* denotes the mean and *sd* the standard deviation of each variable in each subgroup; *mean diff.* reports the t-value of a test for the statistical significance of the difference in means.

Table 3: **The CCPA, loan applications, and loan rates**

VARIABLES	(1) applications	(2) applications	(3) applications	(4) rate	(5) rate	(6) rate
CA x post	0.120***			0.141***		
	(0.011)			(0.008)		
fintech x post	0.273***	0.284***		0.067***	0.057***	
	(0.014)	(0.014)		(0.011)	(0.011)	
CA x fintech x post	0.133***	0.134***	0.146***	-0.083***	-0.080***	-0.079***
	(0.020)	(0.021)	(0.021)	(0.015)	(0.015)	(0.015)
Observations	75,354	75,354	75,354	75,354	75,354	75,354
R-squared	0.763	0.790	0.791	0.889	0.904	0.904
Lender*Tract FE	✓	✓	✓	✓	✓	✓
Time FE	✓	-	-	✓	-	-
Tract*Time FE	-	✓	✓	-	✓	✓
Lender*Time FE	-	-	✓	-	-	✓

This table reports results for Equation (1) at the lender-tract-year level. The dependent variable is the log of the total number of applications in columns (1)–(3) and the average interest rate on approved mortgages in columns (4)–(6). The dummy *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. Lender*time FE denote time-varying fixed effects at the lender type level. Standard errors are clustered at the tract level. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: **Interest rate dispersion, denial rates, and credit scoring models**

VARIABLES	(1) sd(int rate)	(2) sd(int rate)	(3) denied	(4) denied	(5) alt CS	(6) alt CS
CA x post	-0.069***		0.003*		-0.034***	
	(0.007)		(0.002)		(0.004)	
fintech x post	-0.028***		-0.008**		0.022***	
	(0.010)		(0.003)		(0.003)	
CA x fintech x post	0.111***	0.093***	0.010**	0.011**	0.028***	0.029***
	(0.013)	(0.014)	(0.005)	(0.005)	(0.005)	(0.005)
Observations	75,354	75,354	75,354	75,354	75,354	75,354
R-squared	0.535	0.592	0.550	0.599	0.770	0.796
Lender*Tract FE	✓	✓	✓	✓	✓	✓
Time FE	✓	-	✓	-	✓	-
Tract*Time FE	-	✓	-	✓	-	✓
Lender*Time FE	-	✓	-	✓	-	✓

This table reports results for Equation (1) at the lender-tract-year level. The dependent variable is the dispersion in interest rates in columns (1)–(2); the share of denied loan applications in columns (3)–(4); and the share of mortgages that do not use standardized credit scoring (CS) models in columns (5)–(6). The dummy *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. Lender*time FE denote time-varying fixed effects at the lender type level. Standard errors are clustered at the tract level. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: **Thin credit file borrowers**

VARIABLES	(1) p50 white rate	(2) p25 white rate	(3) p50 tr income rate	(4) p25 tr income rate	(5) p50 app income rate	(6) p25 app income rate
CA x fintech x post	-0.105***	-0.119***	-0.085***	-0.088**	-0.086***	-0.090**
	(0.023)	(0.036)	(0.024)	(0.037)	(0.024)	(0.044)
Observations	33,065	14,213	36,942	18,748	32,947	15,675
R-squared	0.903	0.903	0.898	0.902	0.906	0.900
Lender*Tract FE	✓	✓	✓	✓	✓	✓
Tract*Time FE	✓	✓	✓	✓	✓	✓
Lender*Time FE	✓	✓	✓	✓	✓	✓

This table shows results for Equation (1) for tracts in CA border counties. The dependent variable is the average interest rate on approved mortgages. The dummy *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. Lender*time FE denote time-varying fixed effects at the lender type level. Standard errors are clustered at the tract level. Each column uses a different measure of the share of thin credit file applicants. Columns (1) and (2) focus on tracts with a low share of White applicants, columns (3)–(4) on tracts with a low tract-to-MSA income ratio, and columns (5)–(6) on tracts with low applicant income on average. For each measure, we focus on tracts below the median and in the bottom quartile of the distribution of minority/tract-to-MSA income/applicant income. The negative effect of the CCPA on loan rates offred by fintechs is stronger for 'thin credit file' applicants. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 6: **CCPA, loan rates, and applicant quality**

VARIABLES	(1) DTI < 36pct	(2) LTV ratio	(3) LTI ratio	(4) log(inc)	(5) risk PCA	(6) baseline rate	(7) PCA rate	(8) PCA+other rate
CA x fintech x post	0.004	-0.008	-0.026	0.005	-0.050	-0.112***	-0.107***	-0.101***
	(0.012)	(0.006)	(0.018)	(0.010)	(0.029)	(0.014)	(0.014)	(0.014)
Observations	68,118	68,155	68,155	68,155	68,155	68,118	68,118	68,118
R-squared	0.549	0.693	0.631	0.777	0.651	0.901	0.903	0.908
Lender*Tract FE	✓	✓	✓	✓	✓	✓	✓	✓
Tract*Time FE	✓	✓	✓	✓	✓	✓	✓	✓
Lender*Time FE	✓	✓	✓	✓	✓	✓	✓	✓
Risk controls	-	-	-	-	-	-	✓	✓

This table reports results for Equation (1) at the lender-tract-year level. The dependent variables in columns (1)–(5) are different measures of applicant quality (dummy for low DTI ratio, the log of the LTV ratio, the log of the LTI ratio, the log of income, and the risk PCA). The dependent variable in columns (6)–(8) is the average interest rate on approved mortgages. The dummy *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 (7) and (8) include the mean and standard deviation of the various risk measures in each lender-tract-year cell as control variables. Lender*time FE denote time-varying fixed effects at the lender type level. Standard errors are clustered at the tract level. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: **Covid, applications, and rates**

VARIABLES	(1) baseline applications	(2) applications	(3) applications	(4) baseline rate	(5) rate	(6) rate
CA x fintech x post	0.145***	0.146***	0.140***	-0.080***	-0.100***	-0.100***
	(0.021)	(0.024)	(0.025)	(0.015)	(0.017)	(0.019)
fintech × workplace mob		-0.002	-0.004		0.008***	0.007***
		(0.003)	(0.003)		(0.002)	(0.002)
fintech × transit mob		0.001	0.005***		-0.002*	-0.001
		(0.002)	(0.002)		(0.001)	(0.001)
fintech × cases pc			0.058***			0.014
			(0.013)			(0.010)
fintech × deaths pc			-1.258**			-0.247
			(0.516)			(0.434)
Observations	73,554	73,554	73,554	73,554	73,554	73,554
R-squared	0.791	0.791	0.791	0.903	0.903	0.903
Lender*Tract FE	✓	✓	✓	✓	✓	✓
Tract*Time FE	✓	✓	✓	✓	✓	✓
Lender*Time FE	✓	✓	✓	✓	✓	✓

This table reports results for Equation (1) at the lender-tract-year level. The dependent variable is the log of the total number of applications in columns (1)–(3) and the average interest rate on approved mortgages in columns (4)–(6). The dummy *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. *Workplace mob* and *transit mob* stand for an applicant county's workplace mobility and transit mobility indices, taken from Google Mobility Trends. *Cases pc* and *deaths pc* stand for Covid cases per capita and deaths per capita in the county. Values for these variables are set to zero in 2018 and 2019, and to their actual realization (averaged across days) in 2020 and 2021. Lender*time FE denote time-varying fixed effects at the lender type level. Columns (1) and (4) replicate the baseline finding for the smaller sample of tracts in counties with information on the Covid variables. Standard errors are clustered at the tract level. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: **Robustness tests**

VARIABLES	(1) purchase applications	(2) purchase rate	(3) refinance applications	(4) refinance rate	(5) young applications	(6) young rate	(7) controls applications	(8) controls rate	(9) off-BS applications	(10) off-BS rate
CA x fintech x post	0.092***	-0.106***	0.114***	-0.071***	0.137***	-0.124***	0.140***	-0.117***	0.103***	-0.120***
	(0.035)	(0.023)	(0.042)	(0.020)	(0.022)	(0.017)	(0.022)	(0.013)	(0.035)	(0.015)
Observations	53,972	53,972	37,418	37,418	72,441	72,441	73,767	73,767	52,601	69,446
R-squared	0.768	0.889	0.789	0.906	0.842	0.905	0.841	0.930	0.764	0.911

This table reports results for Equation (1) at the lender-tract-year level. The dependent variable alternates between the log of the total number of applications and the average interest rate on approved mortgages. The dummy *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. Lender*time FE denote time-varying fixed effects at the lender type level. Standard errors are clustered at the tract level. 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.

B Online Appendix

B.1 Additional information on the CCPA

B.1.1 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.

During our period of analysis, the companies that are subject to the 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.[22]

The scope of what constitutes personal information under the CCPA is broad. For example, an IP address of an individual browsing a website is considered personal infor-

---

[22]https://cppa.ca.gov/faq.html. Accessed November 25th, 2022. Starting in 2023 to meet the second threshold, businesses must annually buy, sell, or share the personal information of 100,000 or more consumers or households.

mation. Therefore, the CCPA is likely to cover a large proportion of companies, including numerous small- to medium-sized enterprises.[23] Virtually all banks and fintech lenders fall under the CCPA.

B.1.2 Salience

Google searches for "CCPA" on Google Trends in Figure OA1 suggest that the residents of California were aware of the introduction of the privacy legislation. Google searches for the CCPA in California increased steeply in late 2019 and remained elevated for most of 2020, while trends in Arizona, Nevada and Oregon did not replicate this interest.

Figure OA1: **Internet searches for CCPA**
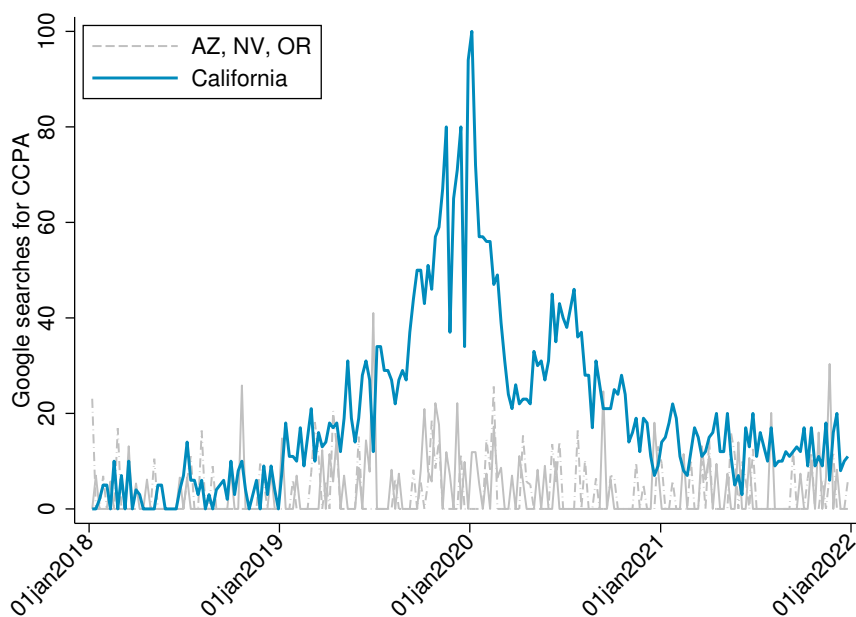
This figure shows internet searches from Google Trends for 'CCPA' in California and neighboring states over time. The CCPA came into effect in January 2020. Note that Google Trends normalizes the search interest so that the day with the most searches over the period of interest takes the value of 100, and all other dates are with respect to this value. To compare searches across states, series are standardized based on average search frequency in California relative to the other states over the period 01Jan2018 to 31Dec2021, with Washington D.C. serving as benchmark.

---

[23]https://iapp.org/news/a/new-california-privacy-law-to-affect-more-than-half-a-million-us-companies/. Accessed November 25th, 2022.

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

B.1.3 Enforcement

Under the CCPA, individuals can file a consumer complaint with the Office of the Attorney General (OAG). Starting in July 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.[24] Individual suing of a business is limited to data breaches where it is clear the company did not take the necessary measures to protect consumers' data.

The OAG regularly sends companies notices of alleged noncompliance.[25] Once a company is notified of alleged noncompliance, it has 30 days to take the neccesary steps to resolve the noncompliance. If it does not solve the issue that prompted the notice, the Attorney General can engage in civil law procedures. The first enforcement settlement related to the CCPA took place in August 2022 and it concerned the French cosmetics brand Sephora. The investigation found that Sephora was selling consumers' personal information without disclosing it, as well as not complying with opt-out requests. Sephora failed to correct these issues within the 30 days of notification by the OAG. Sephora agreed to settle for $1.2 million in fines and agreed to follow its compliance obligations.[26]

---

[24]https://cppa.ca.gov/faq.html. Accessed November 25th, 2022.

[25]https://oag.ca.gov/privacy/ccpa/enforcement. Accessed March 24th, 2023.

[26]https://iapp.org/news/a/the-sephora-case-do-not-sell-but-are-you-selling/. Accessed November 25th, 2022.

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.

B.2 A parsimonious theoretical framework

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

B.2.1 Setup

We consider a competitive loan market where potential borrowers are continuously and uniformly distributed along an arc 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 them.

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, called p_H, are successful in their investment and get a return Y that allows them to repay. The other half (p_L) is unsuccessful so they do not repay. We assume that Y is large enough so that everyone applies for a loan. Applicants are protected by limited liability, and their outside option is normalized to zero.

Applicants are endowed with one unit of personal data and dislike sharing their data with lenders. In line with the evidence in Figure 1, their disutilty 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' dislike 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 dislike of providing data to the fintech vs the bank. Importantly, applicants do not possess the technology to infer their type from their data.

Lenders. Two symmetric banks are located at the extremes of the arc, with bank 1 ($B1$) at the beginning and bank 2 ($B2$) at the end. Each applicant regards the lenders as providing services with different convenience levels, which is captured by the distance x between a consumer on the arc and the banks (Thisse and Vives, 1988).

The fintech (F) is located at the same distance x_F to all consumers, irrespective of their location on the arc. 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).[27] We normalize the fintech's distance to all applicants to zero ($x_F = 0$) for simplicity. The analysis remains qualitatively similar for any positive and sufficiently small distance $x_F > 0$.

Lenders have access to a screening technology that returns either a good signal (η_g) or a bad signal (η_b). The accuracy of the signal depends on two elements: the technology that each lender type $j = \{B, F\}$ 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 technology to one ($\gamma_F = 1$). Moreover, the screening technology's accuracy is increasing and concave in the amount of data: When there is little data available, more data considerably improves the signal accuracy.[28] When lenders already have a considerable amount of data, an additional unit does not increase accuracy by much.[29]

---

[27]Chu and Wei (2021) and Vives and Ye (2022) make a similar assumption in a Salop circle.

[28]As our focus is what happens to data and hence interest rates when there is a change in the dislike to sharing data, we employ a screening technologies that yield interior optimal data requested.

[29]Berg et al. (2020) show that even simple, easily accessible data that proxy for income, character, and reputation are highly valuable for default prediction.

More data hence allow the lenders to filter out, among all applicants, a higher proportion of those that would not repay.

The signal is uncorrelated across lenders and has a bad-news flavor: Only applicants that do not repay send bad signals. Good signals can come from both types of applicants. The signal structure is:

$$\Pr(\eta_{j,g}|p_H) = 1, \qquad \Pr(\eta_{j,b}|p_L) = \gamma_j \sqrt{d_j}.$$

With this signal structure, the probability that lender j observes a good signal is:

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

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

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

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. We follow Vives and Ye (2022) and assume that the fintech can price discriminate their offers to applicants based on their location in the line thanks to its predictive models (Fuster et al., 2022). Allowing the fintech to price discriminate takes a short-term view on the market dynamics: it assumes that the bank is slow to catch up to fintechs' more advanced technology (Navaretti et al., 2018) and it also assumes that regulation is slow in addressing fintechs' discriminatory practices (Bhutta et al., 2021).

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 choose their interest rates: banks first post uniform loan rates, while the fintech moves second and can price discriminate applicants based on their position in the line (Vives and Ye, 2022). The final contract, which consists of a combination of the offered interest rate and requested data $(r(d; x), d)$, is contingent on the applicant qualifying for a loan. 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 arc (ie their relative distance to each lender), they apply to the lender with the offer that maximizes their expected utility.[30] 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 (η_g) and disqualify the applicants that returned a bad signal. We focus on equilibria that are symmetric on the banks' choices.

B.2.2 Equilibrium

We proceed backwards. 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 x_F, normalized to zero for expositional clarity.

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

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

$$\mathbb{E}[u_{s_F}(\ell_F; r_F(d; x), d_F, x)] = \frac{1}{2}(Y - r_F(d; x)) - s_F d_F$$

The expected utility from applying reflects the offered interest rate (which depends on the data requested), the disutility of sharing data (which is greater when applying

---

[30]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.

to 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, for whom the disutility from distance is small, contract with the traditional bank. The fintech receives applications from borrowers 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.

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 whenever the fintech charges the lowest feasible interest rate (that makes it break even), and by $\tilde{x}_2$ the position of the borrower indifferent between applying to lender 2 or the fintech whenever the fintech charges the lowest feasible interest rate.

Proceeding by backwards induction, the fintech sets the interest rate at each location to match the expected utility that applicants get when applying to the bank. Therefore, for each $x \in [\tilde{x}_1, \tilde{x}_2]$, the fintech's interest rate is:

$$
\begin{aligned}
r_F(d, x) &= r_{B1}(d) + 2x + 2d_{B1} - 2d_F s_F && \text{if } x \in [\tilde{x}_1, \tilde{x}] \\
r_F(d, x) &= r_{B2}(d) + 2(1 - x) + 2d_{B2} - 2d_F s_F && \text{if } x \in [\tilde{x}, \tilde{x}_2]
\end{aligned}
$$

Banks anticipate the fintech's pricing strategy, take the data choices as given and internalize the effect of their choices on demand.

$$
\begin{aligned}
\mathbb{E}[\Pi_{Bk}(r_{Bk}(d), d_{Bk}; r_{-j}(d), d_{-j})] &= D_{Bk}(r_{Bk}(d), d_{Bk}; r_{-j}(d), d_{-j}) \, \mathbb{E}[\pi_{Bk}(r_{Bk}(d), d_{Bk})] \\
&= D_{Bk}(r_{Bk}(d), d_{Bk}; r_{-j}(d), d_{-j}) \left[\frac{1}{2} r_{Bk}(d) - 1 + \frac{\gamma_B \sqrt{d_{Bk}}}{2} \right]
\end{aligned}
$$

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

$$d^\star_{B1} = \frac{\gamma_B^2}{16} \tag{2}$$

$$d^\star_{B2} = \frac{\gamma_B^2}{16} \tag{3}$$

$$d^\star_F = \frac{1}{16s_F^2} \tag{4}$$

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: The optimal rates are:

$$r^\star_B = 2 - \frac{3\gamma_B^2}{16} - \frac{1}{16s_F},$$
$$r^\star_F(x) = 2 + 2x - \frac{\gamma_B^2}{16} - \frac{3}{16s_F}.$$

In equilibrium, optimal demand (D_F) is determined by the relative dislike of providing data to the fintech vs the bank (s_F) and the difference in the screening technology (γ_B):[31]

$$D^\star_B = \frac{\gamma_B s_F - 1}{32s_F}$$
$$D^\star_F = 1 - \frac{\gamma_B s_F - 1}{16s_F}$$

Lenders' demand for data is a function of each lender's signal accuracy and the dislike to sharing data 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

---

[31]Given that in the model, market size is constant, demand to the fintech also represents its market share.

costly for applicants. 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 them.

B.2.3 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, ie legislation that provides consumers with greater control over their data and reduces concerns about the abuse and misuse of data. Such legislation decreases borrowers' dislike to sharing their data. Moreover, we assume that this decrease is greater for sharing data with fintechs compared to banks, consistent with the evidence that individuals generally have lower confidence in fintechs to safely handle personal data.

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

$$-\frac{\partial D_F^\star}{\partial s_F} = \frac{1}{16s_F^2} > 0. \tag{5}$$

The increase in demand for the fintech means that the indifferent consumer between banks and fintechs now sits closer to the banks, and hence the range of interest rates offered by the fintech expands, while decreasing the interest rates that the fintech charges relative to banks

$$\left(\frac{\partial r_F^\star}{\partial - s_F}\right) - \left(\frac{\partial r_B^\star}{\partial - s_F}\right) = \left(-\frac{3}{16s_F^2}\right) - \left(-\frac{1}{16s_F^2}\right) < 0. \tag{6}$$

The mechanism behind these results is as follows: as borrowers' dislike sharing data with the fintech decreases, the fintech asks for more data, which makes its screening more accurate. The fintech identifies and rejects more of the credit-unworthy applicants. In turn, this improves the quality of the accepted applicant pool, allowing the fintech to offer lower interest rates. Both because borrowers are less concerned about sharing data and because they are charged a lower interest rate, more borrowers apply to the fintech.

In sum, the model generates the following testable hypotheses:

Hypothesis 1: The introduction of privacy protection legislation in the spirit of the CCPA leads to an increase in loan applications with fintechs, compared to banks.

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

Hypothesis 3.1: The introduction of privacy protection regulation in the spirit of the CCPA leads to the fintech offering more individualized pricing and its dispersion in interest rates increases relative to banks.[32]

Hypothesis 3.2: The introduction of privacy protection regulation in the spirit of the CCPA increases the share of denied loan applications by the fintech by more relative to banks.

Hypothesis 3.3: The introduction of privacy protection legislation in the spirit of the CCPA implies that the fintech will ask for relatively more personal data than banks.

---

[32]Babina et al. (2022) relate the use of non-traditional data beyond standardized credit scores to more individualized pricing. Jansen et al. (2022) suggest that a more precise signal about borrowers' quality leads to more accurate pricing and hence greater dispersion in interest rates.

B.3 Further Figures and Tables

Figure OA4: **Pre-trends**

(a) Loan applications

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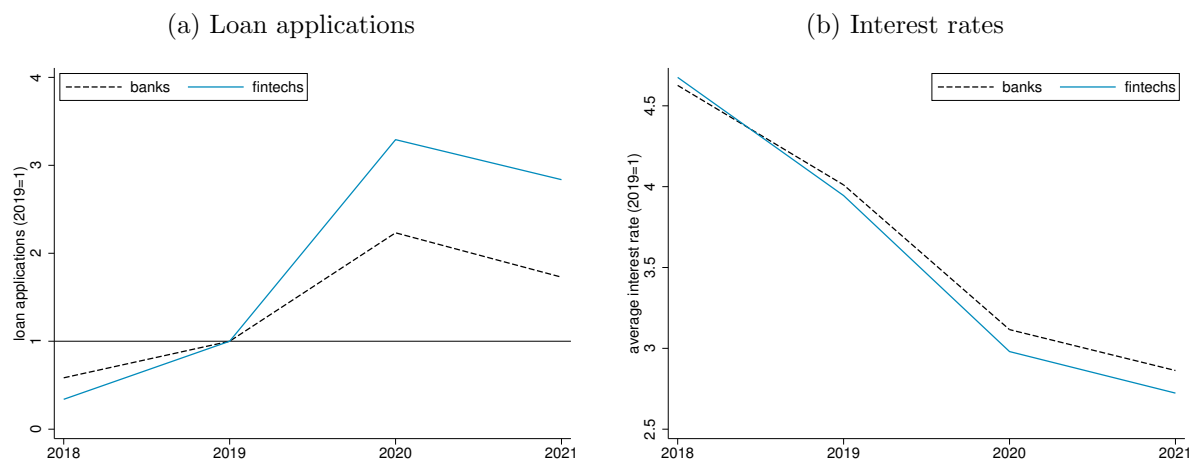
Panel (a) shows the evolution of loan applications with banks (black dashed line) and fintechs (blue solid line) in California during the sample period. Applications with each lender are standardized to 1 in 2019, the year before the CCPA came into effect. Panel (b) plots the respective interest rates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

Table OA1: **The CCPA, fintechs' market share, and mortgage origination**

VARIABLES	(1) FT mkt share	(2) FT mkt share	(3) FT mkt share	(4) FT mkt share	(5) log(orig)	(6) log(orig)
CA x post	0.033***	0.035***	0.031***	0.030***	0.116***	
	(0.005)	(0.005)	(0.006)	(0.010)	(0.011)	
fintech x post					0.284***	
					(0.014)	
CA x fintech x post					0.118***	0.129***
					(0.020)	(0.021)
Observations	6,848	6,848	6,848	6,848	75,354	75,354
R-squared	0.506	0.508	0.509	0.525	0.753	0.783
Tract FE	✓	✓	✓	✓	-	-
Time FE	✓	✓	✓	-	✓	-
Controls	-	C	C+A	C+A	-	-
Border*Time FE	-	-	-	✓	-	-
Lender*Tract FE	-	-	-	-	✓	✓
Tract*Time FE	-	-	-	-	-	✓
Lender*Time FE	-	-	-	-	-	✓

Columns (1)–(4) reports results from the following difference-in-differences specification at the applicant tract–year level: $fintech\ market\ share_{c,t} = \beta\ CA_c \times post_t + \theta_c + \tau_t + controls_{c/a} + \varepsilon_{c,t}$. The dependent variable is the share of applications to fintechs in census tract c and year t (ie fintechs' market share). The dummy 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. Regressions include tract (θ) and year (τ) fixed effects. Census tract-level (C) controls 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. Applicant-level (A) controls, averaged to the tract level, include the pre-period values of 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, all interacted with the post dummy. Standard errors are clustered at the tract level. Column (1), with tract and time fixed effects, shows that after the introduction of the CCPA, the market share of fintechs increased by more in California compared to other states (ie, $\beta > 0$). Columns (2) and (3) include a battery of tract or applicant control variables. Column (4) includes border county*time fixed effects to account for observable and unobservable factors common to neighboring counties, including loan demand. Across specifications, the estimated coefficient remains positive and economically and statistically significant. In terms of magnitude, the share of applications to fintechs increases by 3.0 percentage points in column (4). In 2018–2019, the share of loan applications to fintechs was 16.1% with a standard deviation of 11.3%, implying an increase in their market share of 19% of the pre-treatment mean. Columns (5)–(6) show results for Equation (1). The dependent variable is the log of the total number of originated mortgages. The dummy variable $fintech$ takes on a value of one if the lender is a fintech and a value of zero if it is a bank. *** p<0.01, ** p<0.05, * p<0.1.

Table OA2: **The CCPA, applications, and loan rates — California sample**

VARIABLES	(1) applications	(2) applications	(3) rate	(4) rate
fintech x post	0.513***	0.516***	-0.063***	-0.069***
	(0.006)	(0.006)	(0.004)	(0.004)
Observations	259,156	259,156	259,156	259,156
R-squared	0.764	0.799	0.907	0.921
Lender*Tract FE	✓	✓	✓	✓
Time FE	✓	-	✓	-
Tract*Time FE	-	✓	-	✓

This table shows results for the following equation at the lender-tract-year level: $y_{l,c,t} = \delta\ fintech_l \times post_t + \theta_{l,c} + \tau_t + \varepsilon_{l,c,t}$ for tracts in California only. The dependent variable is the log of the total number of applications in columns (1)–(2) and the average interest rate on approved mortgages in columns (3)–(4). The dummy variable *fintech* takes on a value of one if the lender is a fintech and a value of zero if it is a bank. The dummy *post* takes on a value of one after the CCPA was enacted. Standard errors are clustered at the tract level. Results show that the main findings obtained in Table 3 also hold for California tracts only. *** p<0.01, ** p<0.05, * p<0.1.

Table OA3: **Excluding individual states**

VARIABLES	(1) no AZ applications	(2) no AZ rate	(3) no NV applications	(4) no NV rate	(5) no OR applications	(6) no OR rate
CA x fintech x post	0.142***	-0.072***	0.228***	-0.084***	0.136***	-0.086***
	(0.021)	(0.015)	(0.036)	(0.027)	(0.022)	(0.016)
Observations	72,863	72,863	47,189	47,189	72,197	72,197
R-squared	0.791	0.904	0.788	0.909	0.792	0.902

This table shows results for Equation (1). The dependent variable alternates between the log of the total number of applications and the average interest rate on approved mortgages. The dummy *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. Lender*time FE denote time-varying fixed effects at the lender type level. Standard errors are clustered at the tract level. Columns (1)–(2) [(3)–(4); (5)–(6)] exclude all tracts in AZ [NV, OR]. Results show that the findings obtained in Table 3 are not driven by movements in the control group. *** p<0.01, ** p<0.05, * p<0.1.

Table OA4: **Clustering**

VARIABLES	(1) T applications	(2) T rate	(3) T*Y applications	(4) T*Y rate	(5) C applications	(6) C rate	(7) C*Y applications	(8) C*Y rate	(9) S*Y applications	(10) S*Y' rate
CA x fintech x post	0.146***	-0.079***	0.146***	-0.079***	0.146***	-0.079**	0.146***	-0.079***	0.146**	-0.079***
	(0.021)	(0.015)	(0.021)	(0.015)	(0.041)	(0.029)	(0.048)	(0.025)	(0.064)	(0.019)
Observations	75,354	75,354	75,354	75,354	75,354	75,354	75,354	75,354	75,354	75,354
R-squared	0.791	0.904	0.791	0.904	0.791	0.904	0.791	0.904	0.791	0.904

This table shows results for Equation (1) for tracts in CA border counties. The dependent variable alternates between the log of the total number of applications and the average interest rate on approved mortgages. The dummy *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. Lender*time FE denote time-varying fixed effects at the lender type level. Standard errors are clustered at the level indicated in the column header, where T is tract, C is county, S, is state, and Y is year. Results show that the findings obtained in Table 3 are robust to different levels of clustering. *** p<0.01, ** p<0.05, * p<0.1.

Table OA5: **All tracts in CA and neighboring states**

VARIABLES	(1) applications	(2) rate	(3) sd(int rate)	(4) denied	(5) alt CS
CA x fintech x post	0.227***	-0.132***	0.065***	0.007***	0.012***
	(0.011)	(0.007)	(0.006)	(0.002)	(0.002)
Observations	404,536	404,536	404,536	404,536	404,536
R-squared	0.796	0.917	0.583	0.602	0.815
Lender*Tract FE	✓	✓	✓	✓	✓
Tract*Time FE	✓	✓	✓	✓	✓
Lender*Time FE	✓	✓	✓	✓	✓

This table reports results at the lender-tract-year level for Equation (1) for all tracts in California and neighboring states. The dependent variable is the log of the total number of applications in column (1), the average interest rate on approved mortgages in column (2), the dispersion in interest rates in column (3), the share of denied loan applications in column (4), and the share of mortgages that did not use standardized underwriting models in column (5). The dummy *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. Lender*time FE denote time-varying fixed effects at the lender type level. Standard errors are clustered at the tract level. Results show that the main findings obtained in Table 3 and Table 4 for border tracts also hold for the full sample of tracts. *** p<0.01, ** p<0.05, * p<0.1.

Table OA6: **Applicant-level regressions**

VARIABLES	(1) app to FT	(2) app to FT	(3) interest rate	(4) interest rate	(5) interest rate
CA x post	0.025***	0.021***	0.127***		
	(0.003)	(0.003)	(0.006)		
fintech x post			-0.055***		
			(0.005)		
CA x fintech x post			-0.058***	-0.053***	-0.043***
			(0.006)	(0.006)	(0.004)
Observations	674,720	674,720	595,226	595,188	595,188
R-squared	0.017	0.036	0.513	0.524	0.658
Tract FE	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	-	-
Tract*Time FE	-	-	-	✓	✓
Lender*Time FE	-	-	-	✓	✓
Applicant Controls	-	✓	-	-	✓

This table reports results at the applicant level. Columns (1)–(2) estimate regressions of the following form: $app\ FT_{i(c,t)} = \beta\ CA_c \times post_t + \theta_c + \tau_t + controls_i + \varepsilon_i$. The dependent variable is a dummy with a value of one if the application of applicant i in census tract c and year t is with a fintech and zero otherwise. Columns (3)–(5) estimate regressions of the following form: $rate_{i(c,t)} = \delta_1\ CA_c \times post_t + \delta_2\ fintech_l \times post_t + \delta_3\ CA_c \times fintech_l \times post_t + \theta_c + \tau_t + controls_i + \varepsilon_i$. The dependent variable is the interest rate on approved mortgages of applicant i. The dummy 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. Standard errors are clustered at the tract level. To control for applicant characteristics, $controls_i$ include applicants' LTI ratio, LTV ratio, log income, age, sex, race, ethnicity, and DTI ratio, as well as the log loan amount, log of the property value, a dummy for whether the loan was refinanced or not, and a dummy for whether the loan was sold or not. Results show that the main findings obtained in Table 3 also hold in applicant-level regressions. *** p<0.01, ** p<0.05, * p<0.1.

Table OA7: **Applicant characteristics full vs border sample**

	(1) mean (full sample)	(2) se (full sample)	(3) mean (border sample)	(4) se (border sample)
black (%)	.007	.007	-.001	-.001
female (%)	-.027	-.027	-.042	-.042
age 62+ (%)	-.044	-.044	-.037	-.037
white (%)	-.165	-.165	-.032	-.032
log(income)	.298	.298	.179	.179
log(property value)	.481	.481	.226	.226
LTI ratio (%)	.002	.002	-.001	-.001
LTV ratio (%)	-.057	-.057	-.027	-.027

For average applicant characteristics at the lender–tract–year level, this table shows the difference in means and associated standard errors between tracts in California and neighboring states. Columns (1) and (2) report values for all tracts, columns (3) and (4) for border tracts. Except for *female*, the difference in applicant characteristics is smaller in border tracts than in the full sample of tracts.

Previous volumes in this series

All volumes are available on our website www.bis.org.