Advertising versus Screening? Online-collected data in lending markets*

Federico Lenzi†

Job Market Paper

18 October 2026

Latest version available **here**

Abstract

As mortgage shopping moves online, lenders gain data on borrowers' search and browsing activity. These data improve screening, lower rates, and expand credit. But they also support targeted advertising that reduces comparison shopping and raises markups. Which force dominates determines who benefits from online data. I exploit the staggered adoption of state privacy laws limiting these uses, comparing borrowers at the same lender across states. Restricting data lowers mortgage rates by 14 basis points at advertising-reliant lenders: lower markups outweigh weaker screening. Consistent with weaker screening, approval rates fall one percentage point for credit scores below 750 and rise half a point above it. Consistent with weaker targeting, paid-search clicks fall 7% at lenders using advertising-platform data. A search model shows that privacy shifts surplus from lenders to borrowers but reduces total surplus because search becomes costlier and screening weaker. In digital lending, data-use rules shape search and credit access.

Keywords: Mortgage markets; Consumer privacy; Alternative data; Targeted advertising; Consumer search; Screening and adverse selection; Price discrimination
JEL Classification: G21, G28, G51, D18, D82, D83, L86, M37.

*I thank my advisor, Efraim Benmelech, for his guidance, and my dissertation committee—Mike Fishman, Sean Higgins, Filippo Mezzanotti, Jacopo Ponticelli, and Regina Wittenberg-Moerman—for their advice. I also thank Gregor Matvos, Nicola Limodio, Alexander Ober, and Lulu Wang for helpful comments and discussions; Irving Birkner for assistance with data procurement; and Scott Coughlin for computing and data support. I gratefully acknowledge financial support from the Guthrie Center for Real Estate Research and the Kellogg School of Management Dean's Office PhD Research Fund. This paper uses HMDA data made publicly available by the Consumer Financial Protection Bureau and the Federal Financial Institutions Examination Council, data obtained from FHFA.gov, and NMLS Consumer Access℠ data provided by the Conference of State Bank Supervisors. The views, analyses, and conclusions are solely my own and do not necessarily reflect those of the CFPB, FFIEC, FHFA, CSBS, or any other federal agency. These institutions do not attest to the quality, accuracy, or timeliness of the data or of analyses derived from them. All errors are my own.

1 Introduction

Seven in ten recent homebuyers prefer to select their mortgage lender online.[1] As borrowers browse the web, search for rates, and start applications, they generate online-collected data that reveal two things to lenders: who is likely to repay, and who is about to shop for a mortgage. I study how these data shape mortgage competition by exploiting the staggered enforcement of state privacy laws that restrict their use, comparing similar borrowers at the same lender across states. Used for screening, the data sharpen risk assessment: lenders approve more creditworthy applicants and reject more risky ones, lowering expected lending costs and rates. Used for search advertising, the largest share of mortgage lenders' ad spending, the data tell a lender when a borrower starts shopping. Advertising can either strengthen or weaken price competition. Broad-reach advertising can strengthen it by adding lenders to the set a borrower compares (Honka et al. 2017; Mendes 2024). Search advertising can weaken it by reaching the borrower as she starts shopping, before she obtains competing quotes: the lender then need only beat her home bank's rate by enough that further search is not worth its cost. The competitive effect of advertising thus depends not only on whom it reaches, but on when. Screening lowers mortgage rates; search advertising raises them. Which force dominates determines whether restricting online-collected data helps or hurts borrowers. For likely shoppers at advertising lenders, search advertising dominates: privacy-law enforcement lowers their mortgage rates by 14 basis points, even as screening weakens. Privacy regulation therefore shapes not only what lenders know about borrowers but also which lenders borrowers consider: it is competition policy as well as information policy.

The two uses affect different decisions and move rates in opposite directions, which makes it possible to separate them in the data. A key distinction is between borrowers who shop, comparing offers across lenders, and those who do not, borrowing instead from a lender they already know, such as their home bank. Screening governs which applicants receive credit and at what price. If privacy laws weaken a screening signal, lenders distinguish risk less precisely among borrowers with similar credit scores: the originated pool should become riskier, defaults should rise, and rates should rise as they incorporate larger expected losses, whether or not borrowers shop. The effect on rejections is ambiguous because weaker screening can lead lenders both to approve more risky applicants and to reject more creditworthy ones. Search advertising governs which lenders a borrower considers, and when. A query for "mortgage rates" or a rival lender's name can place a sponsored lender ahead of the organic results. Under limited attention and costly search, borrowers place greater weight on lenders encountered early (Huang et al. 2026; Günsberg and Kuhnen 2026), and some accept an early offer even when further search would yield a lower rate. A lender that appears first can

[1] Fannie Mae, *Survey of Recent Homebuyers on Digital Verification Technology*, August 2024. The survey covers 1,201 recent homebuyers whose mortgages were acquired by Fannie Mae. (https://www.fanniemae.com/media/52916/display)

therefore price against what the borrower expects to gain from further search, net of its cost, rather than against a competing quote already in hand. If privacy laws weaken this signal, the effects should be concentrated where the advertising rents were: among likely shoppers at advertising lenders. For these lenders, rates on such borrowers should fall, their share of such borrowers should shrink, and paid search clicks should decline, while borrowers unlikely to shop should be affected primarily through screening. No single outcome separates the two channels. Their joint response across rejections, defaults, paid search clicks, borrower composition, and pricing does.

The privacy laws provide the variation in what lenders may know. Beginning with the California Consumer Privacy Act, state privacy regulation became salient for lenders: other states introduced or enacted similar laws, and a federal statute was under active debate (Dubé et al. 2025). Leveraging novel data on lenders' website architecture, I classify lenders by their use of online-data technologies before the laws took effect and focus on those that had already adopted them. Pre-enforcement adoption provides a stable classification: lenders operate a common website infrastructure across states, while increasing regulatory scrutiny at both the state and federal levels sharply curtailed new installations of data-collection technologies. The key feature of the setting is that lenders operate the same technology across states, while enforcement limits its access to borrower data only in the regulating state. I therefore compare observationally similar borrowers applying to the same lender, at the same time, for similar properties, who differ only in whether a privacy law limits the use of their online data, and I estimate dynamic effects with the Sun and Abraham (2021) event study, dating treatment to state enforcement. Holding the lender's business model and technology stack fixed, the design isolates the effect of the data themselves.

Search advertising reaches borrowers who shop. Testing this prediction requires a measure of borrower search, which administrative mortgage data do not record. I address this using the National Survey of Mortgage Originations, in which 55% of borrowers report considering more than one lender. I estimate how reported shopping varies with predetermined demographics and income, and use the resulting estimates to predict search propensity for borrowers in the loan-level sample.

The pricing response splits along exposure to search advertising, as the model implies. Among advertising lenders, and after conditioning on borrower, lender, and property characteristics, privacy-law enforcement lowers mortgage rates for likely shoppers, borrowers with high predicted search propensity, by 14 basis points—a decline of 3.43% relative to the mean rate—while the corresponding estimate for low-propensity borrowers is small and statistically indistinguishable from zero. Screening alone cannot generate this pattern: weaker screening worsens the risk composition of approved loans and should push rates up, and it operates on approved borrowers whether or not they shop. By contrast, borrowers unlikely to shop are less likely to enter the search channel, leaving little scope for search advertising to affect their rates. A decline concentrated among likely shoppers is instead the signature of the search-advertising channel—privacy laws weakening lenders' ability

to reach these borrowers before they obtain competing quotes.

The allocation of borrowers moves with the price. Following enforcement, the share of likely shoppers in advertising lenders' originations falls by 0.85 percentage points (3.4% of the pre-treatment mean). Advertising lenders thus lose likely shoppers, and those they still originate receive lower rates. This joint decline in rates and borrower share is difficult to explain through screening alone but follows naturally when online-collected data also support search advertising.

The empirical patterns provide evidence for both the screening and search-advertising uses of online data. First, after enforcement, rejection rates at advertising lenders fall by 0.22 percentage points (2.9% of the pre-treatment mean), while one-year bankruptcy among their originated borrowers rises by 2.09 percentage points (80.1% of the pre-treatment mean). These patterns are consistent with weaker screening: once the signal is removed, lenders distinguish risk less precisely within credit-score buckets and approve applicants they would previously have rejected. Second, privacy-law exposure reduces paid search clicks among lenders that rely on third-party advertising platforms: a 10-percentage-point increase in exposure lowers clicks by about 7.5%. Clicks show no significant change at lenders that also collect data on their own websites, consistent with state privacy laws mainly restricting cross-site targeting, and they rise at lenders without advertising technology, consistent with these lenders winning more valuable ad auctions once data-reliant rivals lose information on which borrowers to bid for. The rate effect then shows which use dominates: among likely shoppers, rates fall by 14 basis points, implying that the loss of pricing power from weaker search advertising more than offsets the higher expected losses from weaker screening.

The empirical design faces two principal identification challenges. First, privacy-law enforcement could coincide with changes in local credit markets that independently affect mortgage outcomes. Several pieces of evidence mitigate this concern. Privacy-law adoption is more strongly predicted by political preferences and exposure to the technology sector than by borrower credit conditions. Outcomes exhibit no differential pre-trends around either signature or enforcement dates, limiting concerns that the estimates capture pre-existing trends or anticipation. I also compare borrowers served by the same lender within metropolitan areas that cross state borders. These borrowers face similar local housing and credit-market conditions but differ in whether their state enforces a privacy law. Second, privacy laws may proxy for contemporaneous state policies or pandemic-era shocks to mortgage markets. Staggered enforcement and the fixed-effect structure address shocks common to states or to time. California is the only state enforcing during the pandemic window, and its estimated effect is already present before March 2020 and remains present through 2024, so the result is not an artifact of that period. A separate measurement concern is that the predicted search propensity classifies some borrowers incorrectly. Such noise should make likely and unlikely shoppers look more similar, not more different: measurement error works against finding a sharp difference between the two groups unless prediction error is systematically related to privacy-law

enforcement. Taken together, these patterns support interpreting privacy-law enforcement as a decline in the quality of lender information rather than as a proxy for contemporaneous policy or macroeconomic shocks.

I develop a search model to quantify the two channels identified in the reduced-form evidence. In the model, screening lowers expected lending costs, while search advertising lets lenders charge for the search that borrowers skip. Privacy weakens both channels, raising rates through worse screening but lowering them by shifting borrowers from advertised offers toward comparison shopping. I discipline the model with approval, default, advertising, and mortgage-rate responses and use it to quantify incidence and surplus.

The estimated model delivers three results. First, privacy reduces credit access overall, with substantial differences across borrowers. Model-implied approval falls in every credit-score bin below 750, is mixed above 750, and declines by 0.48 percentage points in aggregate. This differs from the reduced-form decline in rejections at advertising lenders because the two estimates measure different objects: the regression isolates changes in approval decisions within the same lender, whereas the model aggregates across lender types and credit-score groups and lets borrowers move across lenders. Second, privacy transfers surplus from lenders to borrowers, but at an efficiency cost. Without privacy laws, search advertising reaches borrowers before they compare offers, so lenders can charge for part of the search friction through the mortgage rate even though borrowers never incur the search effort. Privacy weakens this channel: some borrowers return to comparison shopping, and part of the friction once priced into mortgage rates is instead spent as real time and effort. Using an external estimate of the share of mortgage-search frictions that reflects real resources (Agarwal et al. 2024), borrower surplus rises by 5.8 basis points of mortgage balance, while total surplus on originated loans falls by 18.5 basis points, reflecting weaker screening and additional search. This loss persists for any assumption about how much of the search friction reflects real resources. Third, a narrower rule would do better. If lenders could still use online data to assess borrower risk but not to target advertising, borrowers would gain 16.0 basis points rather than 5.8, and the surplus loss would shrink from 18.5 to 8.9 basis points.

This paper contributes to three literatures, reviewed in Section 2. First, work on alternative data shows that nontraditional information improves screening and credit access (Berg et al. 2020; Di Maggio et al. 2022). I show that the same data also support search advertising, which raises prices, and that for likely shoppers this second use dominates. This dual use offers a common explanation for the opposite effects of the California Consumer Privacy Act on mortgage rates documented by Gupta et al. (2024) and Doerr et al. (2026). Second, work on mortgage search models search frictions primarily on the borrower side (Allen et al. 2019; Agarwal et al. 2024; Bhutta et al. 2026). I add a lender-side margin: lenders use online-collected data to reach borrowers early and offer just enough that further search no longer pays. This margin also helps explain why

data-rich lenders, such as many fintechs, can screen better yet charge higher rates (Buchak et al. 2018): the same data that lower their lending costs also let them earn acquisition markups. Third, work on advertising and privacy regulation emphasizes the informative and persuasive roles of advertising (Honka et al. 2017; Gurun et al. 2016) and how restricting data affects online advertising (Goldfarb and Tucker 2011; Dubé et al. 2025). I show that the timing of advertising also matters: search advertising can soften price competition without persuasion, by cutting search short. I also quantify the trade-off privacy rules create between borrower gains and total surplus, and show that restricting only the advertising use of data would reduce its efficiency cost. As mortgage search moves online, data-use rules become competition rules.

The rest of the paper is organized as follows. Section 2 reviews the related literature. Section 3 describes the institutional setting, and Section 4 introduces the data. Section 5 defines borrower search, classifies lenders by their pre-period adoption of online-data technologies, and presents the empirical design. Section 6 reports the main reduced-form results on pricing and borrower allocation. Section 7 examines the underlying mechanisms, and Section 8 presents robustness analyses. Section 9 develops a simple model that rationalizes the empirical findings, and Section 10 describes its quantitative implementation. Section 11 evaluates the surplus effects of privacy regulation. Section 12 concludes.

2 Related Literature

2.1 Mortgage Search and Lender Market Power

One line of work uses price dispersion and equilibrium pricing to quantify search frictions and lender market power (Allen et al. 2019; Bhutta et al. 2026), while Agarwal et al. (2024) show that search in credit markets is inseparable from screening, because rejection risk shapes both borrowers' search decisions and which transactions are observed. In both approaches, heterogeneity in search frictions is modeled primarily on the borrower side. I add a lender-side margin: lenders use online-collected data and search advertising to influence which lender a borrower encounters and whether she continues searching. Whereas Argyle et al. (2023) study search when access to lenders depends on their geographic distribution, I study online acquisition, where a lender's reach depends less on its branch network. Search frictions persist nonetheless, because lenders can shape attention and consideration before competing on price. By entering a borrower's consideration set before she obtains a competing quote, a lender can price against her continuation value: the rate she expects from further search plus the cost of obtaining another approved offer. The logic is that of Diamond (1971): when search is costly, the presence of many potential lenders does not eliminate markups, because the current lender is disciplined only by offers the borrower actually obtains. What is new

is that lenders themselves partly shape effective search: the order in which borrowers see offers and when they stop looking. This mechanism offers a common explanation for limited shopping through mortgage brokers (Woodward and Hall 2012), the pricing power of realtor–loan-officer referral networks (Zhang et al. 2026), the premium fintech lenders earn among convenience-sensitive borrowers (Buchak et al. 2018), and pricing disparities that persist across fintech and traditional lenders (Bartlett et al. 2022). It also provides a customer-acquisition mechanism for the incomplete pass-through of improvements in processing and verification technology documented by Jiang et al. (2026) and D'Andrea et al. (2025): in the model, adopters' lower marginal costs reduce their posted search rates, yet acquisition markups can cause them to charge higher average book rates than non-adopters.

2.2 Alternative Data, Fintech Lending, and Privacy Regulation

Existing work shows that nontraditional information improves risk prediction and credit access (Berg et al. 2020), allows some borrowers rejected by traditional models to obtain credit at lower rates (Di Maggio et al. 2022), and extends lending to applicants without conventional credit histories (Chioda et al. 2024); restricting such data can disproportionately exclude marginal applicants (Agarwal et al. 2024). I identify a strategic use of the same information: online-collected data can also generate lender rents through customer acquisition. The closest empirical work studies the California Consumer Privacy Act (CCPA), the first of these state laws. Gupta et al. (2024) document that compliance costs raise mortgage costs, with stronger effects among banks, whereas Doerr et al. (2026) find that the CCPA lowers fintech mortgage rates by encouraging data sharing and improving screening; the two estimates rest on different comparison groups. I also find that rates fall, but on a distinct margin: the decline is concentrated among likely shoppers at advertising lenders and is accompanied by fewer risk-based rejections and higher subsequent default. The dual-use mechanism helps reconcile these contrasting findings. Restricting data weakens screening, raising rates through adverse selection, but also weakens search advertising, lowering rates by reducing customer-acquisition rents. Rates alone therefore do not reveal which channel dominates, whereas the joint response of rejections, subsequent default, paid search clicks, and borrower allocation does. Separating the two also requires holding technology fixed. Related work studies how financial firms adopt digital technologies to collect customer information, attract customers, and produce signals (Zanotti 2025; Frésard and Zanotti 2025; Chen et al. 2026, 2023); I instead hold the lender's technology stack fixed and exploit privacy-law enforcement to vary the data available to already-adopting lenders across otherwise similar borrowers. This isolates the effects of online-collected information on screening, pricing, customer acquisition, and lender market power.

More broadly, whereas work on financial technology emphasizes screening, processing efficiency,

and customer service (Berg et al. 2022; Fuster et al. 2019), I highlight customer acquisition as an additional use: lenders use data not only to evaluate and serve borrowers, but also to shape their search and capture part of the resulting surplus. This connects the paper to work on customer capital, in which marketing builds a persistent customer base (Gourio and Rudanko 2014) and is an important intangible investment in industries with severe information frictions (He et al. 2024, 2026). It also complements Eeckhout and Veldkamp (2023) by identifying a channel through which data generate markups: influencing consumer search. The same lender-side view changes the market-structure implications of data adoption. Models of open banking and information span study how broader data access changes lenders' information sets and competitive incentives (He et al. 2023, 2024). I add asymmetric adoption and costly borrower search. Data-rich lenders then need not displace data-poor lenders and can even benefit from their presence: a targeted offer is priced against what the borrower expects to obtain by continuing to shop, and less-informed non-adopters post higher rates in the search market, which raise this benchmark. Non-adopters survive on the other side of the same friction: incumbent relationships and high search costs let them retain borrowers who do not shop. Coexistence is therefore an equilibrium outcome rather than a transitional stage of technology diffusion, and better information can shift competition toward advertising and customer acquisition rather than passing fully into lower rates. This prediction is consistent with Babina et al. (2025), who find that the new-lending effects of UK open banking are concentrated among firms already served by the credit market.

2.3 Advertising and Non-Price Competition

Existing work emphasizes two roles for advertising. It can inform consumers and expand the set of firms they consider, a role studied in banking (Mendes 2024; Honka et al. 2017), mutual funds (Hortaçsu and Syverson 2004; Roussanov et al. 2021), and consumer-product markets (Dubé et al. 2026; Goeree 2008); it can also persuade households to buy costly financial products (Gurun et al. 2016). I emphasize that the timing of advertising shapes which role dominates. Broad-reach advertising can introduce lenders before a borrower begins shopping, whereas search advertising reaches her during search, before she obtains a competing quote. Search advertising can therefore reduce competition without persuasion, by cutting search short. The search fatigue documented by Ursu et al. (2023) and the sequential-search framework of Fershtman and Pavan (2025) explain why the timing and position of lender exposure matter: consumers expand their consideration sets gradually and become less likely to inspect additional results. In mortgages, switching lenders becomes costly once a borrower has been pre-approved or is under contract with a closing deadline, so an early contact can sustain lender markups, as in the models of search advertising (De Corniere 2016; De Corniere and De Nijs 2016; Haan et al. 2024; Burguet and Petrikaitė 2023) on which my

model builds. My evidence shows that this channel operates in practice: when privacy laws curtail search advertising, mortgage rates fall for likely shoppers at advertising lenders, with no detectable effect for borrowers unlikely to shop.

I depart from this literature by embedding search advertising in a credit market in which the same online-collected data also provide a signal of repayment risk. This dual use creates opposing price effects: better screening reduces expected credit losses, whereas search advertising allows lenders to monetize borrowers' search costs. The paper also adds two features to research on privacy regulation and digital marketing (Goldfarb and Tucker 2011; Aridor et al. 2023; Deisenroth et al. 2024; Dubé et al. 2025). First, technology adoption is asymmetric, so restricting data reallocates borrowers across lender types rather than uniformly reducing advertising. Second, costly search under imperfect information about market rates allows lenders to capture part of the value of targeting as acquisition rents, so restricting targeting transfers surplus to borrowers rather than merely reducing advertising effectiveness. The results provide microeconomic evidence for the price-increasing force emphasized by Bergemann et al. (2025), while showing that it operates through borrowers' lender consideration sets and is concentrated among likely shoppers. More broadly, the paper contributes to emerging work on marketing as non-price competition among financial intermediaries (Benetton et al. 2025; Chen et al. 2025): data-driven advertising is not merely a demand shifter, but a technology that changes which lenders borrowers consider, how borrowers are allocated across lenders, and the equilibrium price of credit.

3 Institutional Setting

3.1 Online-Collected Data

Mortgage lenders increasingly use data generated when consumers search online, visit websites, and interact with digital platforms. These data differ from traditional mortgage variables such as FICO scores, loan-to-value ratios, debt-to-income ratios, income, and employment. They describe how a potential borrower behaves before or during the application process: what mortgage terms she searches for, which lender websites she visits, whether she returns to the same product page, whether she starts an application, and how she arrived at the lender's website.

By January 2020, lenders in my sample differed substantially in their use of these technologies. About 26.4% of lenders did not use online-data technologies. Another 39.4% had adopted advertising technologies. These technologies allow lenders to reach potential borrowers through search ads, display ads, and other digital platforms. They also provide analytics about who saw or clicked an ad, where the user was located, what search terms led to the ad, and whether the user later visited the lender's website. The remaining 34.2% of lenders had adopted both advertising technologies and

customer data platforms. Customer data platforms represent a deeper form of data use: they allow lenders to store and combine data from their own websites, advertising platforms, and other digital sources.

The distinction between these two technologies is central to the paper. Advertising technologies serve two functions at once. First, they place a lender in front of potential borrowers on outside platforms such as search engines and advertising networks. They reveal traffic sources, campaigns, keywords, location, audience characteristics, and campaign responses. This is the customer acquisition function. See Dai et al. (2023) for a technical explanation of search advertisement and Wernerfelt et al. (2025) for measuring the value of online data in targeted advertising. Second, the same platforms return analytics on the audiences that saw or clicked an ad, including attributes such as income, education, and spending patterns. These attributes provide a signal of borrower quality within a credit score bucket. This is the screening function. Alternative online data of this kind has been shown to improve lender screening beyond traditional credit bureau variables (Di Maggio et al. 2022). Advertising technologies therefore carry both uses of online collected data. No technology lenders have neither. The comparison between the two groups thus identifies the acquisition and the screening channels jointly. Customer data platforms add richer first party data from the lender's own website, such as page visits, calculator use, and application starts. They represent a more intensive version of the same two uses. Lenders that operate them form the advertising plus data group, which I analyze separately in the appendix. This distinction is consistent with evidence that financial firms use website and analytics technologies to learn about potential customers and target them more effectively (Zanotti 2025). Appendix A provides additional detail.

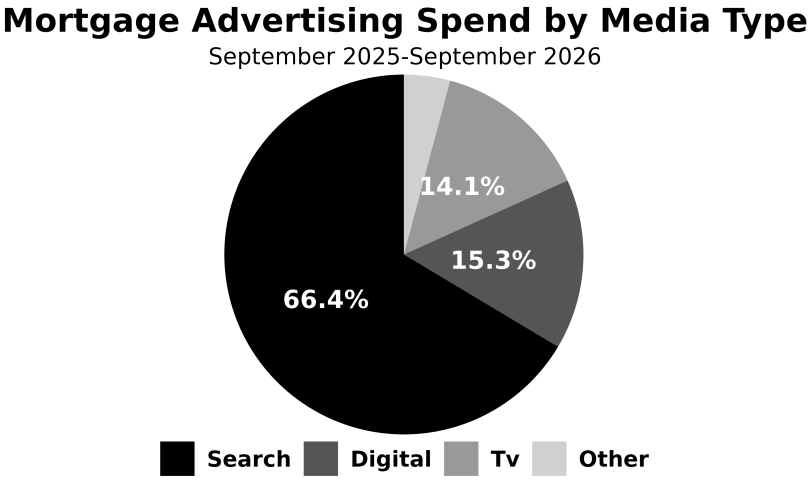
These technologies matter for mortgage lending because the same data can serve two economic functions. First, they help lenders acquire customers. A lender can appear in front of borrowers when they search for mortgage rates, compare lenders, or visit mortgage-related websites. Second, the data can improve information production. Online behavior can reveal demand, attention, product interest, location, and the intensity of a borrower's search. This does not require advertising and underwriting to be the same system. The point is that data generated through advertising and website technologies can be informative both for customer acquisition and for screening, consistent with prior evidence that nontraditional data can improve screening and expand credit access (Berg et al. 2020; Chioda et al. 2024; Di Maggio et al. 2022).

Search advertising is especially relevant because it reaches borrowers at the moment they reveal demand for mortgage credit. A borrower who searches for 'mortgage rates," 'refinance mortgage," or the name of a competing lender can be shown sponsored results before organic search results. These ads can place a lender in the borrower's consideration set before the borrower has finished shopping. Appendix Figure A.1 provides an illustrative example of this search-ad placement and its connection to lender website data collection.

The economic importance of this channel is visible in mortgage advertising expenditures. Figure 3.1 shows that search advertising accounted for 66.4% of mortgage advertising spending between September 2025 and September 2026, more than four times the share of either other digital advertising (15.3%) or television (14.1%). This allocation is consistent with lenders competing not only on mortgage rates but also for a place in borrowers' consideration sets at the outset of search.

This visibility margin is economically important because mortgage consideration sets are small. In the National Survey of Mortgage Originations, 44.8% of borrowers report considering only one lender, 38.1% consider two, and 17.1% consider more than two. If many borrowers consider only one or two lenders, then entering the consideration set is a central margin of competition. Online-data technologies therefore link three objects that are central to the paper: customer acquisition, information production, and credit-market competition.

Figure 3.1. Mortgage Advertising Spending by Media Type



Notes: This figure reports the share of mortgage advertising expenditures by media type from September 2025 to September 2026. "Digital (excl. search)" covers display, video, and other online advertising other than search; "Other" includes all remaining media.
Source: Author's calculations using Vivvix advertising expenditure data.

The adoption of online-data technologies took off around 2013, coinciding with the diffusion of smartphones and the broader growth of online consumer data. Advertising technologies diffused quickly because they are typically provided by third-party platforms and can be adopted without rebuilding a lender's internal data systems. Customer data platforms diffused more slowly because they require larger investments in data infrastructure and technical expertise. Appendix A.4 reports adoption patterns and Cox proportional hazards models based on HMDA originations and NMLS establishment and loan-officer data.

These adoption results are descriptive. The empirical analysis does not use adoption timing as the source of identification. Instead, I use pre-period adoption to classify lenders, condition on lenders that had already adopted these technologies, and study how privacy-law enforcement

changes the data they can use for otherwise similar borrowers within the same lender.

3.2 Privacy Laws

U.S. state privacy laws restrict how firms collect, share, sell, and use personal data. These laws were not written as mortgage-credit policy. They emerged from broader concerns about the growth of online consumer data, the role of large digital platforms, and the limited ability of consumers to observe or control how firms use their personal information. The first major statute was the California Consumer Privacy Act (CCPA), enacted in June 2018 after a privacy ballot initiative qualified and was withdrawn. California later adopted the California Privacy Rights Act (CPRA), approved by voters as Proposition 24 in November 2020. Other states subsequently adopted similar comprehensive privacy statutes.

The laws share a common structure. They give consumers rights to access, delete, correct, and port personal data, and they restrict some uses of personal information for sale, sharing, targeted advertising, profiling, and sensitive-data processing. They also differ in strength. Some laws restrict data sharing as well as data sales; some require opt-in consent for sensitive personal information; some require firms to honor browser-level opt-out signals; and some give firms a cure period before enforcement. Appendix B reports the full state-law timeline, including enactment dates, enforcement dates, and the legal-intensity measures used in the empirical analysis. Appendix Figure B.1 illustrates how formal privacy laws differ from the self-regulatory cookie-banner regime that preceded them.

The empirical analysis uses privacy-law enforcement dates, rather than enactment dates, as the treatment dates. This distinction is central. Enactment dates reveal when a law enters the legal environment, while enforcement dates determine when firms face legal exposure for noncompliance. The gap between enactment and enforcement also allows me to test for anticipation. The treatment is therefore not the political debate over privacy regulation, but the point at which firms face enforceable restrictions on data collection, sharing, and use.

Although these laws were aimed at the broader market for consumer data, they can bind mortgage lenders. The reason is simple: not all data collected by a financial institution are the same kind of data. Traditional customer financial information is governed by federal financial-privacy rules, including the Gramm-Leach-Bliley Act (GLBA). But mortgage lenders also collect online data before a borrower becomes a customer or outside a logged-in customer relationship. Examples include website visits, advertising identifiers, referral sources, campaign responses, device and location signals, and browsing behavior generated when a potential borrower searches for mortgage products or visits a lender's website. State privacy laws can therefore affect lenders through the online-data infrastructure used for advertising, attribution, customer acquisition, and demand assessment.

This boundary is central to the paper. A potential borrower may search for mortgage rates, click on a paid advertisement, visit a lender's website, use a calculator, or begin an application without completing it. These interactions can generate information about demand, attention, search intensity, and product interest. Privacy-law enforcement can therefore change the data inputs available to lenders even though the laws were not designed as mortgage regulation.

Privacy-law enforcement can affect lenders through two related channels. First, it can raise compliance costs by requiring firms to build systems for disclosure, opt-out, deletion, consent, and data-governance obligations. Prior work shows that the California privacy laws increased compliance costs and affected mortgage-market outcomes (Gupta et al. 2024). Second, enforcement can restrict the data that lenders use for targeting, customer acquisition, and information production. Tang (2019) documents that households are willing to pay a premium to preserve their privacy. My empirical design focuses on this second channel by comparing borrowers and lenders with different exposure to online-collected data before privacy-law enforcement.

The adoption patterns in Appendix A.4 support this design. Advertising technologies and customer data platforms diffused rapidly before the first California privacy law and then plateaued after the emergence of state privacy regulation. At the same time, lenders that had already adopted these tools could continue using them where legal restrictions did not bind or for consumers who did not exercise privacy rights. This makes pre-period adoption a useful measure of exposure: it classifies lenders by their reliance on online-collected data before the enforcement shocks. Appendix B.1 also discusses the state-specific applicability thresholds. In my sample, the lenders adopting these technologies before enforcement are large enough that statutory thresholds are unlikely to drive the exposure classification.

A final identification concern is whether privacy-law timing reflects local credit-market conditions. I address this concern directly in Appendix B.3. The California CPRA referendum provides a setting in which privacy preferences can be measured at a granular level. Using census-tract voting records, I show that support for stronger privacy protection is primarily related to political preferences and employment in software-related sectors, which proxy for awareness of online data practices. It is not systematically related to mortgage-market conditions or to borrowers' propensity to search for mortgage prices. This evidence supports the interpretation that privacy-law enforcement changes lenders' data environment rather than responding to mortgage-market trends.

4 Data

4.1 Sources

The analysis combines four main datasets. The first is a novel panel of mortgage lenders' website technologies from 2010 to 2024, constructed from BuiltWith. The provider collects repeated snapshots of website source code and identifies the third-party technologies installed on each website, including analytics tools, advertising pixels, tags, customer data platforms, and other software embedded in the page source. I use these data to measure lenders' adoption of online-data technologies. Prior research uses the same source to measure firms' digital infrastructure (Stroube and Dushnitsky 2023). I manually classify each software product by its primary function. The two categories used in the analysis are advertising technologies and customer data platforms. Advertising technologies link lender websites to advertising platforms and allow lenders to target potential borrowers, measure campaign performance, and receive information on traffic sources and campaign responses. Customer data platforms allow lenders to collect, store, and combine first-party data generated on their own websites, such as page visits, referral sources, form starts, calculator use, repeated visits, and application behavior.

The second dataset combines the public Home Mortgage Disclosure Act (HMDA) data with mortgage-originations data from Equifax. HMDA provides loan-level information on mortgage applications, originations, lender identities, loan characteristics, borrower demographics, and census-tract location. Equifax provides credit-bureau information for a sample covering approximately 14% of the U.S. population, including variables that are not available in the public HMDA data, such as credit scores and origination month. I use the matched HMDA–Equifax data as the main loan-level panel for the analysis of mortgage rates and borrower allocation.

The third dataset is Vivvix advertising data, which reports advertising activity by advertiser, media channel, and time. I use advertising expenditures from 2022 to 2025 to document the importance of search advertising in mortgage customer acquisition, and monthly paid search clicks, observed at the national lender level, as the outcome in the advertising-channel analysis. I match Vivvix advertisers to HMDA lender identifiers by name, which links each lender's clicks to its technology classification and to its 2018–2019 HMDA origination shares across states.

The fourth dataset is the public National Survey of Mortgage Originations (NSMO), covering 2015Q3 to 2024Q2. The NSMO asks borrowers how many lenders they considered, how many lenders they applied to, and, for borrowers with multiple applications, why they submitted more than one application. I use these questions to measure mortgage search behavior. The survey also reports borrower income, credit-score bins, demographics, loan characteristics, and application quarter.

I supplement these four datasets with lender-level information from the Avery HMDA files and the Nationwide Multistate Licensing System (NMLS). These data provide lender characteristics

and organizational structure used in the adoption analysis. I also manually collect information on state privacy laws, including enactment dates, enforcement dates, and legal provisions, and validate the classification against data from the Social Catalyst Lab (Lab 2026). Appendix B reports the privacy-law timeline and provision-level classification.

4.2 Sample Construction and Matching

The mortgage sample focuses on home-purchase mortgages. I exclude refinancing, home-improvement loans, and home-equity lines of credit. I also exclude loans with balloon payments, interest-only features, second liens, and HOEPA status. These restrictions produce a sample of standard home-purchase loans for which mortgage rates and borrower allocation are most directly comparable.

The public HMDA data do not report origination month or borrower credit score. To recover these variables, I probabilistically match HMDA originations to the Equifax mortgage sample. The match is conducted separately by year and geography. I first map HMDA census tracts to ZIP codes using the HUD–USPS ZIP Code Crosswalk. Within year and ZIP code, I restrict candidate matches to loans in the same borrower age bin, government-insurance category, and credit-union-origination status. Because Equifax does not report the lender name, I then select the nearest match using cosine distance based on loan maturity, borrower income, and loan amount. I drop candidate matches with maturity differences greater than 12 months or with income or loan-amount differences greater than $25,000$.

The final HMDA–Equifax match links 18.25% of eligible HMDA originations to Equifax records and 51.6% of Equifax first-mortgage records to HMDA originations. The HMDA-to-Equifax match rate is consistent with the 14% population coverage of the credit-bureau sample. The lower Equifax-to-HMDA match rate reflects the coarser mortgage-purpose information in Equifax. In the credit-bureau data, first mortgages are identified by the account type *FM*, which pools home-purchase and refinance mortgages. By contrast, the HMDA analysis sample keeps only home-purchase mortgages and excludes refinances, home-improvement loans, home-equity lines, balloon-payment loans, interest-only loans, second liens, and HOEPA loans. Some Equifax first-mortgage records therefore fall outside the eligible HMDA analysis sample even before matching. Appendix C reports additional match diagnostics.

I also match Vivvix advertisers to mortgage lenders. Vivvix does not provide a lender identifier, so I match advertisers to lenders using website URLs and advertiser names. I first standardize website domains by removing protocols, prefixes such as `www`, URL paths, and capitalization differences. Remaining unmatched records are matched using string similarity after standardizing names, removing punctuation, removing organizational and legal suffixes, converting numbers to

words, and replacing ampersands with "AND." High-confidence matches are retained automatically, intermediate-score matches are AI reviewed, and the remaining cases are manually validated. The final Vivvix match contains 919 unique advertiser–lender matches, corresponding to 30% of the HMDA–Equifax lender sample and 41% of the Vivvix mortgage-advertiser sample. Appendix C reports match rates and validation statistics.

4.3 Main Variables

The first main variable is lender exposure to online-collected data. I classify lenders into three mutually exclusive groups using their pre-period website technologies. Lenders that use neither advertising technologies nor customer data platforms are classified as *no technology*. Lenders that use advertising technologies but not customer data platforms are classified as *advertising only*. Lenders that use both advertising technologies and customer data platforms are classified as *advertising plus data*. By January 2020, 26.4% of lenders are in the no-technology group, 39.4% are in the advertising-only group, and 34.2% are in the advertising-plus-data group. The empirical analysis uses pre-period adoption status to classify lenders before privacy-law enforcement.

The second main variable is borrower search propensity. I observe search behavior directly in the NSMO, where borrowers report how many lenders they considered and whether they submitted multiple applications. I use these questions to classify borrowers by search behavior and to estimate search propensity as a function of borrower and loan characteristics. I then map the estimated search propensity into the HMDA–Equifax sample using the overlapping variables available in both datasets, such as borrower demographics, income, loan amount, credit-score bins, loan type, and application timing. Section 5.1 describes the imputation procedure and inference.

The third main set of variables comes from the HMDA–Equifax loan panel. The primary price outcome is the mortgage interest rate. Because originations concentrate in the low-rate vintages of 2020–2021, the pooled mean note rate (4.17%) lies below the simple time-average of prevailing rates over 2018–2024; annual means track the Freddie Mac 30-year fixed-rate series closely (Appendix Table C.2). The primary quantity outcome is borrower allocation across lenders, measured by lender market shares within borrower-search groups and local markets. I also use HMDA application data to study screening margins, including rejection outcomes and denial reasons, and Equifax information to measure borrower credit characteristics and post-origination outcomes where available.

The fourth main variable is lenders' paid search clicks, measured by Vivvix at the national lender-month level. A paid click occurs when a consumer clicks on a lender's sponsored search result; it is the unit in which lenders pay search platforms and thus a direct measure of the traffic their search advertising generates. The sample covers 527 lenders from January 2022 to December 2025 (Panel C of Table 4.1). Clicks are highly concentrated: only 34% of lender-months record any

paid clicks, and the mean (80,180) far exceeds the median (zero). To measure exposure to privacy regulation, I compute the share of each lender's 2018–2019 originations located in states with a privacy law in force, dating each law by its first full month of enforcement. Fixing shares before the laws ensures that exposure changes only when new laws take effect, not when lenders adjust where they lend. Because California's law took effect in 2020, before the sample begins, exposure is positive at sample entry and rises as additional states enforce their laws between 2023 and 2025. It averages 0.198, about one-fifth of a lender's pre-period customer base, and its within-sample change, the variation used in estimation, averages 0.122.

The final mortgage sample contains 3,722,265 originated mortgages from 3,466 lenders between 2018 and 2024. Table 4.1 reports summary statistics for the principal variables in the mortgage-origination, loan-application, and advertising samples. Appendix Table D.1 reports the full distributions of the categorical controls and summary statistics for the remaining borrower, loan, and census-tract characteristics.

5 Empirical Strategy

Online collected data affect mortgage markets through two channels. They help lenders screen borrowers within credit score buckets, and they help lenders reach borrowers through targeted advertising. The empirical strategy separates these forces along two dimensions. The first dimension is lender technology. The primary comparison is between no technology lenders and advertising technology lenders. The latter hold both uses of online collected data. The former hold neither. Lenders that also operate customer data platforms hold a more intensive version of the same two uses, and I examine them separately in the appendix. The second dimension is borrower search. The advertising channel is not equally active for all households. Many borrowers never compare lenders, so targeted advertising matters mainly for borrowers willing to consider an outside lender. I therefore measure each borrower's propensity to search for better mortgage prices in the National Survey of Mortgage Originations and map this measure into the HMDA Equifax loan panel. I then use the staggered enforcement of state privacy laws as a shock to lenders' ability to use online collected data. The identifying comparison is within lender and quarter, across borrower states. For the same lender at the same time, I compare borrowers in states with enforceable privacy restrictions to borrowers in states not yet treated. This design holds fixed lender level pricing, funding, underwriting, and advertising shocks. It allows the effect of enforcement to vary with borrower search propensity. The resulting estimates identify whether privacy laws reduce screening, disrupt targeted acquisition, or reallocate borrowers across search channels.

Table 4.1. Summary Statistics: Main Variables

	N	Mean	SD	P25	Median	P75
Panel A. Mortgage-originations sample, 2018–2024						
Lender classification						
No technology	3,722,265	0.131				
Advertising only	3,722,265	0.292				
Advertising and data	3,722,265	0.578				
Outcomes						
Search-for-price propensity ($\widehat{\pi}^P$)	3,228,191	0.249	0.065	0.210	0.256	0.293
Mortgage interest rate	3,716,010	4.172	1.299	3.125	3.875	4.875
Bankruptcy indicator (t+1)	3,228,686	0.027				
Selected borrower and location characteristics						
Vantage score	3,712,681	747.009	68.794	704.000	761.000	798.000
Consumer income score	3,712,681	65.307	42.752	37.000	54.000	79.000
asinh(property value)	3,708,497	13.299	0.510	12.972	13.321	13.629
Tract-to-MSA income (%)	3,722,265	112.824	37.321	88.000	109.000	133.000
Tract minority population (%)	3,722,265	32.548	24.326	13.230	25.860	46.430
Panel B. HMDA loan-application sample						
Rejection indicator	335,687,587	0.067				
Panel C. Lender-month paid search sample, 2022–2025						
Paid search clicks	25,296	80,180.067	747,082.529	0.000	0.000	3,194.253
Any paid search clicks	25,296	0.338				
Privacy-law exposure	25,296	0.198	0.337	0.000	0.000	0.250
Change in exposure since sample entry	25,296	0.122	0.276	0.000	0.000	0.055

Notes: Panel A uses the merged HMDA–Equifax sample of originated mortgages. Panel B uses the HMDA loan-application sample employed in the rejection analysis. Panel C uses the Vivvix lender-month paid search sample employed in Table 7.2, covering 527 lenders; privacy-law exposure is the share of the lender's 2018–2019 originations located in states with a privacy law in force. Because PPML drops lenders whose clicks are zero throughout a fixed-effect cell, estimation samples in Table 7.2 are smaller than Panel C. For indicator variables and mutually exclusive lender classifications, the mean reports the sample share; standard deviations and percentiles are omitted. Variable-specific sample sizes reflect missing observations. Appendix Table D.1 reports the full distributions of categorical controls and summary statistics for additional borrower, loan, and census-tract characteristics.

5.1 Measuring Mortgage Search

Search behavior is not observed in HMDA or in the Equifax mortgage panel. I therefore measure it using the National Survey of Mortgage Originations (NSMO), which contains the main self-reported measures of mortgage shopping in the United States. The survey asks borrowers how many lenders they considered, how many lenders they applied to, and, among borrowers who submitted multiple applications, why they did so. I use these questions to construct a borrower-level propensity to search for better mortgage prices and then map that propensity into the HMDA–Equifax loan panel.

I focus on the application margin because it is the point at which borrowers can obtain comparable price offers. Advertising and online search affect the borrower's consideration set, but applications are the price-revealing margin: borrowers can observe complete offers, including rates, points, and fees, only after engaging lenders sufficiently to obtain an offer. The application measure is therefore downstream of consideration, but it is the margin most directly connected to realized mortgage pricing.

Let X_i denote the set of borrower and loan characteristics observed in both the NSMO and HMDA–Equifax samples. In the NSMO, define

$$A_i = \mathbf{1}\{\text{borrower applied to more than one lender}\}$$

and, among borrowers who applied to more than one lender,

$$P_i = \mathbf{1}\{\text{borrower reports applying to multiple lenders to obtain better terms}\}.$$

I estimate two survey-weighted logit models on the NSMO home-purchase sample:

$$\Pr(A_i = 1 \mid X_i) = \Lambda(X_i'\beta_A), \tag{5.1}$$

$$\Pr(P_i = 1 \mid A_i = 1, X_i) = \Lambda(X_i'\beta_P), \tag{5.2}$$

where $\Lambda(z) = e^z/(1 + e^z)$. The predicted probability that borrower i searches for better prices is

$$\widehat{\pi}_i^P = \Lambda(X_i'\widehat{\beta}_A) \times \Lambda(X_i'\widehat{\beta}_P). \tag{5.3}$$

This measure captures the probability that a borrower submits multiple applications for price-related reasons. It is narrower than simply considering more than one lender, but it corresponds more closely to the point at which borrowers obtain comparable mortgage offers.

In the NSMO home-purchase sample, 55% of borrowers report considering more than one lender, while 28% apply to more than one lender. Among borrowers who submit multiple applications, 83.7% report doing so to obtain a better price or better loan terms. The index $\widehat{\pi}_i^P$ therefore captures

variation in the ex-ante propensity to engage in price-motivated mortgage shopping, rather than ex-post responses to realized offers.

The prediction model uses only variables observed in both the NSMO and HMDA–Equifax data. These include year, quarter, loan amount, loan-to-value ratio, debt-to-income ratio, loan-level price adjustment, rural applicant status, low-income urban status, credit-union origination, government-insured loan status, credit score, income, marital status, sex, age, and race.[2] The retained predictors are standard determinants of mortgage-shopping behavior: age, race, income, and loan size (Agarwal et al. 2007, 2024; Coen et al. 2023).

I then apply the estimated coefficients to the HMDA–Equifax sample after harmonizing the common variables into the same categorical format. Predictions are made only on the support of the NSMO covariates. Observations with covariate cells not observed in the NSMO are left missing. Appendix E reports the share of HMDA–Equifax loans with complete search predictions and the out-of-support rate.

The search measure is a generated index, not a directly observed borrower type. I therefore interpret the estimates as heterogeneity by predicted price-search propensity. The prediction rule is fixed: I estimate the mapping from observables to search behavior in the NSMO and apply the same mapping to HMDA–Equifax borrowers using characteristics observed at origination. Privacy-law enforcement may change the distribution of borrower characteristics, but the mapping from characteristics to predicted search is held fixed. Differential misclassification would require the mapping from observables to latent search behavior to change systematically with enforcement, conditional on the variables used to construct the index. I assess this concern by reporting stability checks for the search prediction and by showing that the main results are robust to alternative thresholds and to excluding race from the prediction model.

The search measure enters the empirical analysis in two ways. In borrower-allocation regressions, I use the continuous predicted propensity $\widehat{\pi}_i^P$. In mortgage-rate regressions, I compare borrowers with high and low predicted price-search propensity and report robustness to alternative thresholds. In the continuous specifications, I treat $\widehat{\pi}_i^P$ as a generated regressor. In the threshold specifications, imperfect classification mixes true searchers and non-searchers across groups. Under nondifferential classification error with respect to privacy-law enforcement, this mixing attenuates the high-versus-low contrast. I use this attenuation logic only for the threshold specifications.

I account for first-stage estimation error by bootstrapping the two-stage procedure. In each bootstrap draw, I re-estimate the NSMO search model, recompute $\widehat{\pi}_i^P$ in the HMDA–Equifax sample, reconstruct the high- and low-search groups when relevant, and re-estimate the second-stage

[2]Education is observed in the NSMO but not directly in HMDA–Equifax. Detailed geography is observed in HMDA–Equifax but not in the public NSMO. I exclude variables that are not measured consistently across the two datasets.

specification. The reported confidence intervals therefore incorporate both uncertainty in the search-prediction model and sampling variation in the second stage.

Table 5.1 reports out-of-sample validation in the NSMO. The prediction exercise is intentionally low-dimensional because the same covariates must be available in both datasets. The models are not designed to classify individual borrowers perfectly; they are designed to rank borrowers by ex-ante search propensity using common observables. In five-fold out-of-sample validation, the AUC is 0.581 for applying to more than one lender and 0.639 for reporting a price motive among multiple applicants. Because the empirical analysis uses the index primarily to rank borrowers or form high- and low-search groups, ranking performance rather than exact calibration is the relevant validation target. Pregibon link tests do not reject the logit specification, and probit and complementary-log-log links produce nearly identical rankings. Appendix E reports the full set of out-of-sample diagnostics, propensity-stability checks, and imputation cross-checks.

Table 5.1. Validation of Price-Search Propensity

	Apply > 1 lender	Price motive | apply > 1
Estimation N	23,341	6,538
Base rate	0.280	0.837
OOS weighted AUC	0.581	0.639
Calibration slope	0.907	0.831
Pregibon link-test p	0.267	0.501
Race robustness		
OOS AUC, without race	0.569	0.639
$\widehat{\pi}^P$: corr(with, no race)	Pearson 0.934 / Spearman 0.939	
$\widehat{\pi}^P$: weighted mean	0.247 with race ; 0.247 without race	

Notes: This table reports five-fold out-of-sample validation in the NSMO home-purchase sample. AUC and calibration statistics are survey-weighted. The price-search propensity is defined in equation (5.3). Alternative links, including probit and complementary log-log, produce similar out-of-sample rankings. The race-robustness rows re-estimate the prediction model after excluding race from the predictor set.

Because race is both a predictor of mortgage-shopping behavior and a central variable in credit-market analysis, I report a robustness exercise excluding race from the prediction model. Removing race leaves the predicted price-search propensity highly correlated with the baseline measure and does not change its mean. This robustness check shows that the empirical split is not driven mechanically by race, while preserving the baseline specification that uses the strongest common predictors of search behavior.

5.2 Regulatory Shock and Empirical Design

I use the staggered enforcement of state privacy laws as a shock to the value of online-collected data in mortgage lending. The laws are not the object of the paper. They provide variation in lenders' ability to collect, combine, and use online data on borrowers located in treated states. Treatment is defined by the enforcement date of a state privacy law, rather than by the enactment date, because enforcement is the point at which firms face legal exposure for noncompliance. I use enactment dates in robustness tests to assess anticipation.

The main analysis focuses on lenders that had adopted online-data technologies before privacy-law enforcement. This sample restriction is central to the design. Lenders operate a common website and advertising infrastructure across states, but privacy-law enforcement restricts the data they may use for borrowers located in protected states. The identifying comparison is therefore within lender and time: for the same lender in the same quarter, I compare borrowers in states where privacy-law enforcement restricts the use of online-collected data to borrowers in states that are not yet treated. This design holds fixed the lender's technology stack, pricing model, funding conditions, national advertising strategy, and other time-varying lender-level shocks.

No-technology lenders do not generate the paper's main treatment effect, but they serve as an informative comparison group. Because these lenders do not use online-collected data, privacy-law enforcement has no direct channel through which to affect their screening or advertising decisions. Their borrowers' outcomes may still move indirectly—through changes in market competition, borrower reallocation, or the pricing responses of rival lenders. Estimates for no-technology lenders therefore separate the direct effect of restricting data use from the broader equilibrium effects of privacy regulation, and, combined with the within-lender comparison in the main design, help isolate the direct data channel from market-level reallocation. A large differential response among these lenders that spillovers cannot explain would instead point to a confound—a correlated state policy or a general shift in mortgage-market conditions. I obtain these estimates by re-estimating equation (5.4) on the sample of no-technology lenders and report them separately.

The identifying assumption is that, absent privacy-law enforcement, borrowers of the same lender in treated and not-yet-treated states would have followed parallel trends, conditional on the fixed effects and controls described below. The timing of enforcement is plausibly unrelated to contemporaneous mortgage-market shocks. These laws were driven by state-level privacy regulation and consumer concerns about personal-data use rather than by mortgage-credit policy. Appendix B.3 shows that the propensity to adopt these laws tracks state political orientation and technology-sector exposure rather than borrower credit conditions. The design further absorbs lender-quarter shocks and local economic shocks, so identification comes from within-lender differences across states at the same time. Event-study estimates test for differential pre-trends before enforcement. The appendix reports additional robustness tests using enactment dates, alternative staggered-adoption

estimators, and tests for cross-state spillovers.

The shock can affect the use of online-collected data even if lenders do not remove installed website technologies. Installed advertising pixels, tags, and customer-data tools remain useful for borrowers in states without enforceable privacy laws and for consumers who do not exercise opt-out rights. Enforcement therefore changes the expected payoff from using online-collected data in treated states, rather than requiring lenders to remove the underlying technology. This distinction is important: the treatment is not technology adoption or technology removal. The treatment is the enforceable restriction on the use of online-collected data for borrowers in treated states.

Because privacy laws are adopted at different times, conventional two-way fixed-effects event studies can be contaminated by comparisons between newly treated and already treated states when treatment effects are heterogeneous (Goodman-Bacon 2021). I therefore use a staggered-adoption-robust event-study design. The baseline estimates follow the interaction-weighted logic of Sun and Abraham (2021), comparing treated cohorts to not-yet-treated states and aggregating cohort-specific effects by event time. Because only nine states are treated during the sample period, the never-treated and not-yet-treated control group is large and clean. I report stacked difference-in-differences estimates (Wooldridge 2025) and group-time average treatment effects following Callaway and Sant'Anna (2021) as robustness checks. The stacked specifications are also useful for estimating interactions with borrower search propensity and lender characteristics in a transparent way.

Let m index mortgages, ℓ lenders, s borrower states, c metropolitan areas, and q calendar quarters. Let E_s denote the enforcement quarter of the privacy law in state s. In the baseline sample of lenders that had adopted online-data technologies before enforcement, I estimate

$$y_{m\ell scq} = \sum_{\substack{k=\underline{k} \\ k\neq -1}}^{\overline{k}} \delta_k \, \mathbf{1}\{q - E_s = k\} + X_m'\gamma + \alpha_{\ell q} + \mu_{c,y(q)} + \eta_{s,\mathrm{mo}(q)} + \varepsilon_{m\ell scq}. \tag{5.4}$$

The omitted event time is $k = -1$. The coefficients δ_k measure the change in outcomes for borrowers in treated states relative to borrowers of the same lenders in not-yet-treated states, around privacy-law enforcement. In the Sun–Abraham implementation, these coefficients are interaction-weighted averages of cohort-specific event-time effects. I report average post-enforcement effects as the main treatment effects and plot the dynamic coefficients to assess pre-trends.

The vector X_m includes borrower, loan, property, and census-tract controls. Borrower controls include age, race, ethnicity, sex, credit score, and income. Loan and property controls include loan amount, loan-to-value ratio, debt-to-income ratio, loan type, government-insurance status, occupancy status, property value, and number of units. Census-tract controls include tract-to-MSA income, minority population share, share of one-to-four-family homes, owner-occupied share, and total population.

The fixed effects absorb the main confounding shocks. Lender-by-quarter fixed effects, $\alpha_{\ell q}$, absorb time-varying lender-level changes in pricing, approval standards, funding costs, advertising strategy, and overall demand. MSA-by-year fixed effects, $\mu_{c,y(q)}$, absorb local economic and housing-market shocks. State-by-month-of-year fixed effects, $\eta_{s,\mathrm{mo}(q)}$, absorb state-specific seasonality in mortgage demand. I do not include state-by-calendar-quarter fixed effects because these would absorb the state-level treatment variation.

To study heterogeneity by borrower search propensity, I estimate equation (5.4) separately for borrowers with high and low predicted search propensity, as defined in Section 5.1. This preserves the staggered-adoption-robust design within each group, comparing treated cohorts to not-yet-treated states and aggregating cohort-specific effects by event time. I plot the two dynamic paths in a single figure and report each group's average post-enforcement effect. Because the two subsamples can place different weights on adoption cohorts, I do not interpret the raw difference between the separately aggregated effects as a formal high-minus-low contrast. The appendix estimates that contrast directly using a stacked difference-in-differences specification that interacts treatment with search-propensity group and yields a single standard error for the difference. Because search propensity is predicted and measured with error, I interpret the results as heterogeneity by predicted search propensity.

Standard errors are clustered at the state level because treatment varies at the state level. The sample contains 51 state clusters, of which 9 are treated by enforcement during the sample period. Because inference with few treated clusters can be unreliable under conventional cluster-robust standard errors, I report wild-cluster bootstrap p-values, using Webb six-point weights, for the average treatment effects and for joint tests of pre-treatment coefficients.[3] The appendix reports robustness to alternative clustering choices, enactment-date timing, stacked difference-in-differences, and cross-state spillover tests within lenders.

6 Results

Online collected data from advertising technologies serve a dual purpose. They extend a lender's outreach, and they allow the lender to pre-screen borrowers within a credit-score bucket. The tables that follow exploit this duality by comparing two lender arms. No-technology lenders have neither use. Advertising-technology lenders have both. When an outcome moves for advertising lenders but not for no-technology lenders, the difference reflects the loss of online-collected data rather than a market-wide effect of the law. When an outcome moves for both, the no-technology response

[3]With nine treated clusters, Rademacher weights admit only 2^9 distinct bootstrap draws; Webb six-point weights provide a finer bootstrap distribution and more reliable inference with few treated clusters (Cameron et al. 2008; MacKinnon and Webb 2018).

identifies the equilibrium reallocation that the data restriction sets in motion. Borrowers fall into two broad groups. Non-searchers prefer the rate offered by their home bank, either to economize on search costs or out of concern about rejection elsewhere. Searchers open their consideration set to competing lenders through organic search and targeted advertising. Within this framework, privacy laws restrict a lender's ability to reach borrowers through targeted advertising and thereby reduce its share of searchers. As the theoretical model shows, targeted advertising lowers a prospective borrower's search cost, and the lender captures part of this saving in the rate it charges. The mechanism can be summarized simply: a lender may post a slightly higher rate and still appear as the first search result, but to be noticed through organic search it must compete on price.

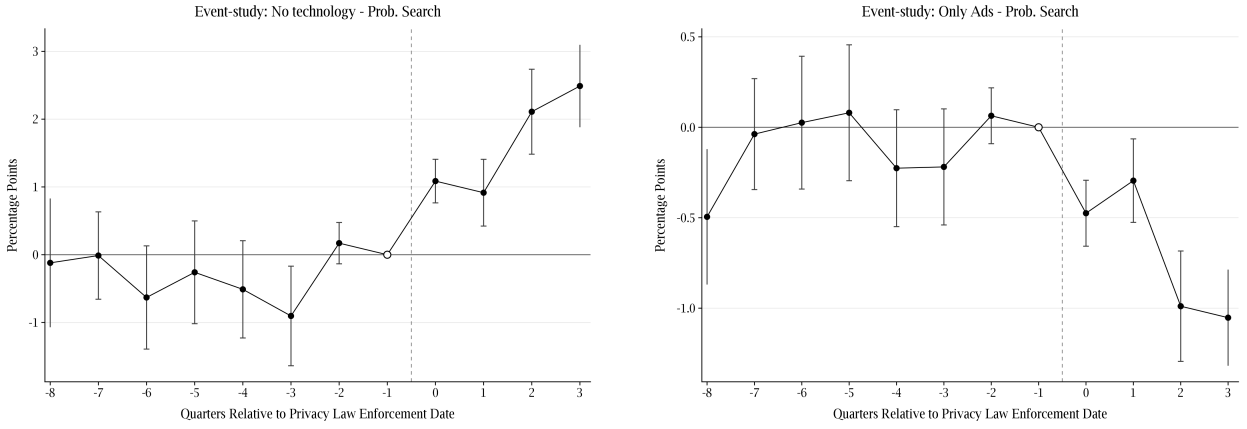
Table 6.1 reports the average treatment effect on the treated for the probability that a mortgage is originated to a price-shopping borrower. Throughout, I report effects in regression units, with the effect relative to the baseline mean in parentheses at first mention. Enforcement reduces the probability that an advertising lender's borrower is a searcher by 0.85 percentage points (3.4% of the baseline searcher share) and raises it for no-technology lenders by 2.07 percentage points (8.1%). The opposite-signed estimates are consistent with borrowers reallocating toward no-technology lenders, although the regressions do not permit a direct accounting of the number of borrowers moving between the two groups. Once access to targeted advertising is curtailed, adopting lenders can be reached only through organic search and lose their visibility advantage, which enlarges the pool of borrowers who discover non-adopting banks. The specification compares borrowers with identical credit scores, demographics, and property characteristics applying to the same lender at the same time, one located in a state with an enforced privacy law and one in a state without. Metropolitan-statistical-area-by-year fixed effects, which span state borders, absorb differential local trends, and state-by-origination-month effects absorb seasonality. Because only nine of fifty states are treated during the sample period, I obtain inference from a wild-cluster bootstrap, for both the treatment effects and the joint nullity of the pre-trends (Figure 6.1).

Table 6.2 reports rate effects by borrower type and lender arm. The estimated rate effects for non-searchers are not statistically distinguishable from zero. Searchers at advertising lenders pay 13.95 basis points less (3.43% of the pre-treatment rate). For a $360,000 thirty-year mortgage, this is roughly $350 per year and about $10,500 over the life of the loan. The discount arises because adopting lenders now acquire these borrowers through the organic channel and must forgo the markup they previously extracted from saved search costs. The placebo cells behave as the mechanism requires. Searchers at no-technology lenders show no effect, because their originations already arrived through organic search and were never exposed to the advertising restriction. Non-searchers show no effect in either arm, because they do not compare lenders and are unresponsive to the advertising channel. The rate effect therefore appears in exactly one of the four cells, the one where the mechanism operates. Figure 6.2 reports wild-cluster bootstrap inference and flat pre-trends.

Table 6.1. Enforcement Effects on Search-for-Price Probability

	Dependent variable: Search-for-price indicator	
	No-Tech	Ads
	(1)	(2)
Post × Treated	0.0207***	-0.0085***
	(0.0026)	(0.0011)
Fixed effects		
Year-Month × Lender	Yes	Yes
MSA × Year	Yes	Yes
State × Month	Yes	Yes
Loan type	Yes	Yes
Observations	420,158	942,716
R^2	0.2297	0.1997
Within R^2	0.0002	0.0001

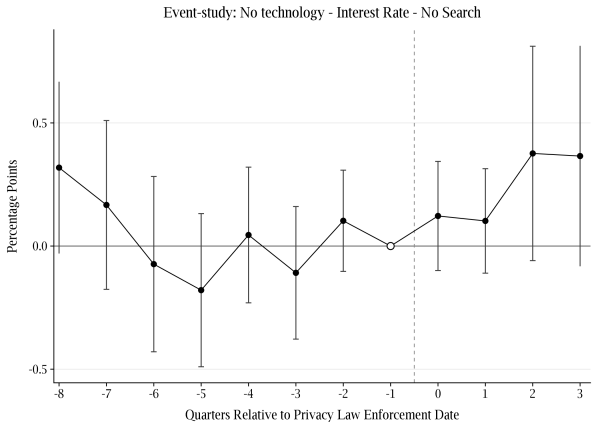
Notes: This table reports difference-in-differences estimates of the effect of privacy-law enforcement on the probability that a borrower searches for price. The coefficient on Post × Treated is the average treatment effect on the treated. Standard errors, clustered at the state level, are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.



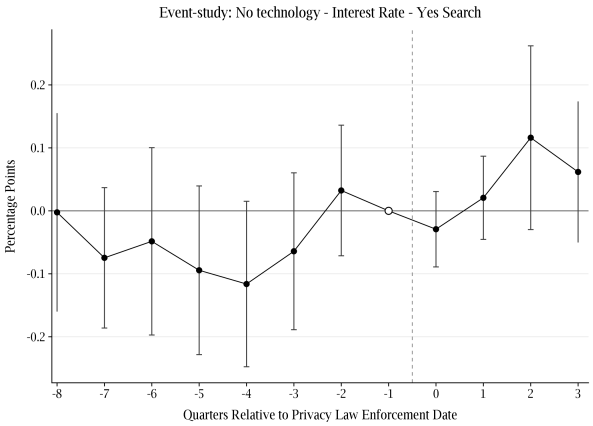
(a): No-Tech **(b):** Ads

Figure 6.1. Event-Study Estimates: Enforcement and Search-for-Price Probability

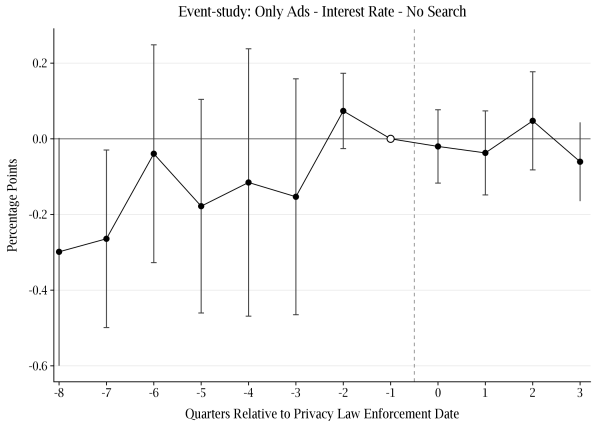
Notes: This figure reports event-study estimates of the effect of privacy-law enforcement on the probability that borrowers search for price. Panel (a) reports estimates for the No-Tech sample and panel (b) for the Ads sample. The omitted period is the last pre-enforcement period. Confidence intervals are based on standard errors clustered at the state level.



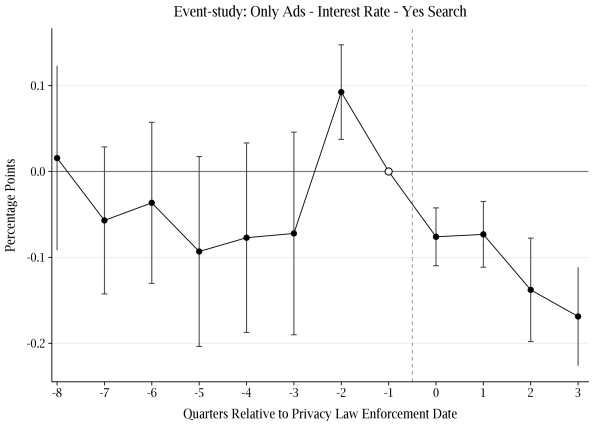
(a): No-Tech, No Search



(b): No-Tech, Search



(c): Ads, No Search



(d): Ads, Search

Figure 6.2. Event-Study Estimates: Enforcement and Mortgage Interest Rates by Search Status

Notes: This figure reports event-study estimates of the effect of privacy-law enforcement on mortgage interest rates, separately by technology exposure and search-for-price status. Panels (a)–(b) report the No-Tech sample for non-searchers and searchers, and panels (c)–(d) report the Ads sample for non-searchers and searchers, respectively. The omitted period is the last pre-enforcement period. Confidence intervals are based on standard errors clustered at the state level.

Table 6.2. Enforcement Effects on Mortgage Interest Rates by Search Status

	Dependent variable: Mortgage interest rate			
	No-Tech		Ads	
	No Search	Search	No Search	Search
	(1)	(2)	(3)	(4)
Post × Treated	0.2944	0.0481	-0.0413	-0.1395***
	(0.1832)	(0.0481)	(0.0472)	(0.0250)
Fixed effects				
Year-Month × Lender	Yes	Yes	Yes	Yes
MSA × Year	Yes	Yes	Yes	Yes
State × Month	Yes	Yes	Yes	Yes
Loan type	Yes	Yes	Yes	Yes
Observations	72,415	334,557	176,410	750,520
R^2	0.6658	0.6700	0.6525	0.6616
Within R^2	0.0253	0.0128	0.0153	0.0149

Notes: This table reports difference-in-differences estimates of the effect of privacy-law enforcement on mortgage interest rates, separately by technology exposure and search-for-price status. The coefficient on Post × Treated is the average treatment effect on the treated. Standard errors, clustered at the state level, are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

In sum, privacy laws contract the targeted-advertising channel and the market data acquired from these platforms. Targeted advertising had lowered borrowers' search costs, raised adopting lenders' shares among searchers, and allowed those lenders to price at the expected search-market rate plus the search cost, generating a rent. After enforcement, more originations flow through organic search, the searcher share of adopting lenders falls, and the rates paid by the searchers they retain decline. Competition among adopters shifts from entering consideration sets back toward price. The model section decomposes this dynamic.

7 Mechanism

The decline in rates traces to the two roles of advertising technology: absorbing search costs and providing information for screening. The theoretical model disentangles these forces. This section provides empirical evidence for each, first on the screening margin through rejections and realized defaults, then on the acquisition margin through paid search clicks.

Rejection rates are available only in the public HMDA release, which lacks origination month and credit score. I therefore adapt the design to compare similar households borrowing from the same lender in the same year across geographic units, with metropolitan-statistical-area-by-year and lender-by-state fixed effects absorbing local variation. Table 7.1 reports the results. Throughout, I

report effects in regression units, with the effect relative to the baseline mean in parentheses at first mention. Rejection rates at advertising lenders fall by 0.22 percentage points after enforcement, or 2.9 percent of the 7.59 percent baseline rejection rate. This is consistent with online-collected data having sharpened screening beforehand. As the signal is withdrawn, lenders separate risk less finely within a credit-score bucket and approve applicants they would previously have declined. Rejections at no-technology lenders fall by 0.53 percentage points, or 6.63 percent. This larger drop reflects composition rather than screening. Price-shopping borrowers reallocate toward non-adopters after enforcement, and these borrowers tend to have higher credit scores, a variable the public HMDA controls cannot capture, so the no-technology applicant pool improves. The same reallocation runs in the opposite direction for advertising lenders, whose applicant pool deteriorates as high-quality searchers leave. The 0.22 percentage point decline is therefore a lower bound on the direct screening effect, because it is net of an adverse composition shift. The cohort analysis in the robustness section supports this reading, since milder privacy laws generate lower rejections only among adopters.

The realized-default evidence completes the screening picture. One-year bankruptcy among advertising lenders' borrowers rises by 2.09 percentage points, a 80.1 percent increase over the 2.61 percent baseline, while the no-technology estimate is small and statistically insignificant. Rejections falling while defaults rise, at the treated arm only, is the joint signature of degraded screening rather than of a market-wide credit shock. The size of the default response also indicates that the withdrawn signal carried substantial information about repayment risk within credit-score buckets. In Figure 7.1, both the treatment effects and the joint nullity of the pre-trends are tested with a wild-cluster bootstrap.

As illustrated in Figure 3.1, search advertising accounts for roughly two-thirds of mortgage lenders' marketing expenditures. I examine whether restrictions on online-collected data reduce lenders' ability to generate paid search clicks. Vivvix reports clicks at the lender-month level but not the location of the households generating them, preventing me from using the borrower-level staggered difference-in-differences design above. I therefore construct lender exposure to privacy regulation as

$$\text{Exposure}_{it} = \sum_{s} \omega_{is}^{pre} \mathbf{1}\{t \geq E_s\}, \tag{7.1}$$

where ω_{is}^{pre} is the share of lender i's 2018–2019 mortgage originations located in state s, and E_s is that state's privacy-law enforcement date. In column (1) of Table 7.2, I estimate

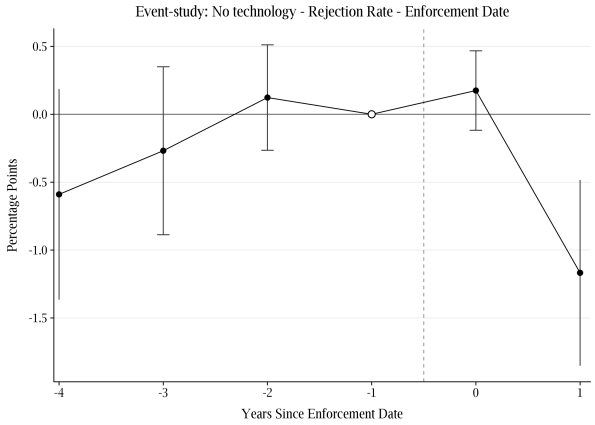
$$\mathbb{E}[\text{Clicks}_{it} \mid \cdot] = \exp\left[\sum_{g} \beta_g \left(10\Delta\text{Exposure}_{it}\right) \mathbf{1}\{G_i = g\} + \alpha_{G_i,t} + \mu_{i,m(t)} + \lambda_{i,y(t)}\right], \tag{7.2}$$

where $\Delta\text{Exposure}_{it} = \text{Exposure}_{it} - \text{Exposure}_{i0}$ and Exposure_{i0} is exposure in lender i's first month in the sample. Given the skewed, nonnegative distribution of paid clicks, I estimate the model using

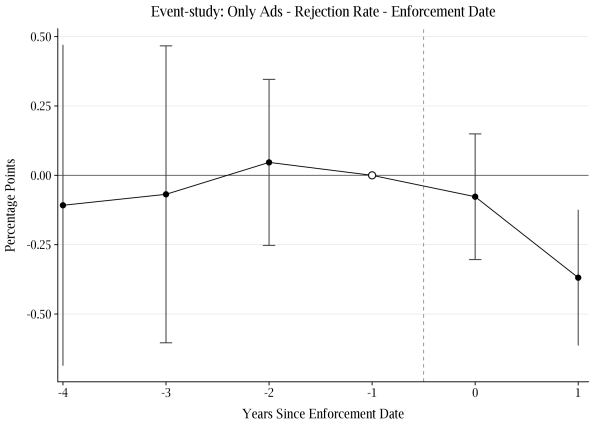
Table 7.1. Enforcement Effects on Rejection Rates and Borrower Bankruptcy

	Rejection indicator		Bankruptcy indicator	
	No-Tech	Ads	No-Tech	Ads
	(1)	(2)	(3)	(4)
Post × Treated	-0.0053**	-0.0022**	-0.0064	0.0209***
	(0.0022)	(0.0010)	(0.0084)	(0.0028)
Fixed effects				
Year × Lender	Yes	Yes	–	–
State × Lender	Yes	Yes	–	–
Year-Month × Lender	–	–	Yes	Yes
State × Month	–	–	Yes	Yes
MSA × Year	Yes	Yes	Yes	Yes
Loan type	Yes	Yes	Yes	Yes
Observations	28,104,664	100,603,792	405,930	925,779
R^2	0.0885	0.0589	0.1079	0.0850
Within R^2	0.0052	0.0063	0.0215	0.0196

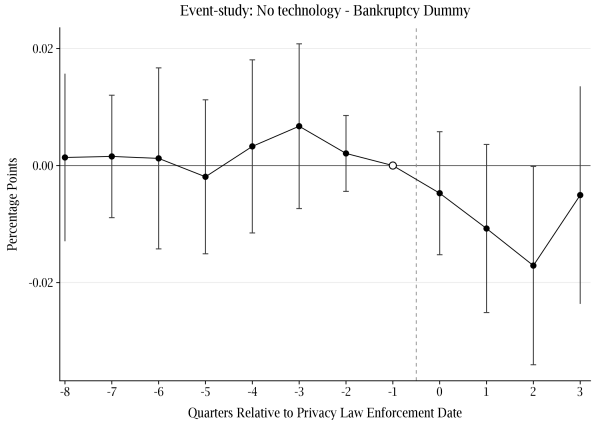
Notes: This table reports difference-in-differences estimates of the effect of privacy-law enforcement on mortgage rejection rates and borrower bankruptcy outcomes. Columns (1)–(2) report effects on the probability that a mortgage application is rejected; columns (3)–(4) report effects on an indicator for borrower bankruptcy. The rejection and bankruptcy specifications use different fixed-effect structures, as indicated: the rejection regressions are estimated on the public HMDA release, which lacks the origination month and credit score, and therefore use Year × Lender and State × Lender fixed effects, whereas the bankruptcy regressions use the merged HMDA–Equifax panel with Year-Month × Lender and State × Month fixed effects. The coefficient on Post × Treated is the average treatment effect on the treated. Standard errors, clustered at the state level, are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.



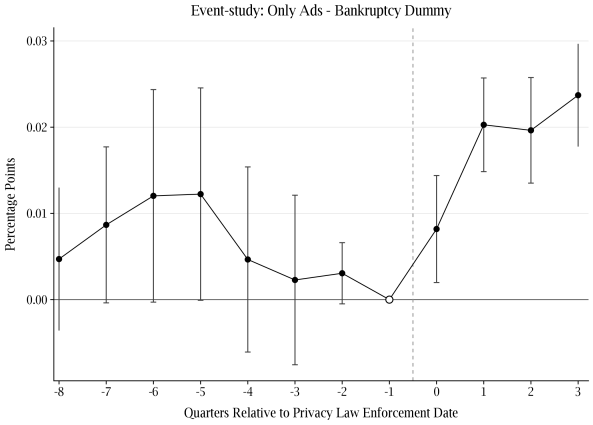
(a): Rejection Rate: No-Tech



(b): Rejection Rate: Ads



(c): Bankruptcy Indicator: No-Tech



(d): Bankruptcy Indicator: Ads

Figure 7.1. Event-Study Estimates: Enforcement, Rejection Rates, and Borrower Bankruptcy

Notes: This figure reports event-study estimates of the effect of privacy-law enforcement on mortgage rejection rates and borrower bankruptcy outcomes. Panels (a)–(b) report estimates for rejection rates in the No-Tech and Ads samples, respectively, and panels (c)–(d) report estimates for borrower bankruptcy in the No-Tech and Ads samples, respectively. The omitted period is the last pre-enforcement period. Confidence intervals are based on standard errors clustered at the state level.

Poisson pseudo-maximum likelihood. The coefficients β_g capture the effect of a 10-percentage-point increase in exposure separately by lender technology group. If restrictions on online-collected data impair search advertising, the decline should be strongest for lenders that use advertising technology without customer data, while lenders that also use customer data may be less affected. Column (2) provides a complementary stacked design, constructing six-month windows around each privacy-law enforcement wave and controlling separately for exposure to other privacy laws. Both specifications include lender-by-year fixed effects, which absorb lender-specific annual shocks, together with lender-specific seasonality and technology-group-specific monthly shocks; the stacked design additionally includes lender-by-wave and technology-group-by-year-month-by-wave fixed effects. Identification therefore comes from within-lender, within-year changes in exposure generated by laws enforced during the calendar year. Standard errors are clustered at the lender level.

Column (1) of Table 7.2 shows that a 10-percentage-point increase in privacy-law exposure reduces paid search clicks by approximately 7.2% among lenders that use advertising technology without customer data (Only Ads). Lenders that combine advertising technology with customer data experience no decline; if anything, their point estimates are positive. Column (2), which uses the stacked design around privacy-law enforcement waves, yields a similar pattern: paid clicks of Only Ads lenders fall by approximately 6.8%. Both estimates are statistically significant at the 5% level with lender-clustered standard errors. Because exposure is driven by a small number of enforcement dates, I also conduct randomization inference that reassigns enforcement dates across states; under this more conservative procedure, the p-values are 0.14 and 0.13, respectively. The corresponding event study in Figure F.2 shows no differential pre-trends for Only Ads lenders ($p = 0.68$) and a gradual decline in paid clicks following enforcement, and a lead test in the panel likewise finds that future exposure changes do not predict current clicks ($p = 0.35$). Finally, Appendix Table F.1 shows that when lender-by-year fixed effects are removed, the decline is concentrated in mid-year enforcement waves (−6.6%), whereas January waves, which coincide with lenders' annual budget cycles, do not generate a comparable decline. Excluding the July 2023 Colorado–Connecticut wave leaves the estimate negative (−5.3%) but less precise. Taken together, these results suggest that restrictions on online-collected data impair search advertising for lenders that rely on third-party tracking, while lenders with their own customer data are insulated. The No Technology group contains only 33 lenders, and I do not interpret its estimates.

This section establishes two results. Online-collected data are used in screening, as the joint movement of rejections and realized defaults shows, and support customer acquisition through search advertising, as the paid-click response shows. Privacy laws weaken the advertising channel sharply by curtailing the flow of personal data across third-party platforms. The screening effect is present but more muted, because lenders use online-collected data both to complement (Di Maggio et al. 2022) and to substitute for (Frésard and Zanotti 2025) other data sources and can therefore

Table 7.2. Privacy-Law Exposure and Paid Search Clicks

	Paid search clicks	
	Shift-share panel	Stacked DiD
	(1)	(2)
Exposure × Only ads	-0.0752**	-0.0701**
	(0.0303)	(0.0295)
Exposure × Ads+Data	0.0605	0.1064
	(0.0701)	(0.0769)
Exposure × No technology	0.0210**	-0.0037
	(0.0082)	(0.0146)
Only ads: effect (%)	-7.2	-6.8
RI p-value, Only ads	0.137	0.126
Pre-trend p-value, Only ads	0.350	0.683
Fixed effects		
Group × Year-Month	Yes	–
Group × Year-Month × Wave	–	Yes
Lender × Wave	–	Yes
Lender × Month-of-year	Yes	Yes
Lender × Year	Yes	Yes
Observations	11,643	29,106
Lenders	368	402
Pseudo R^2	0.992	0.994

Notes: This table reports Poisson pseudo-maximum-likelihood estimates of the effect of state privacy-law exposure on lenders' monthly paid search clicks. Exposure is the share of the lender's 2018–2019 mortgage originations located in states with a privacy law in force, interacted with the lender's technology group; coefficients measure the effect of a 10 percentage point increase, and the implied percentage effect is $100(e^{\beta} - 1)$. Column (1) uses the lender-month panel with exposure measured relative to sample entry. Column (2) stacks six-month windows around each enforcement wave and controls for exposure to other laws. Because lender-by-year fixed effects absorb annual marketing-budget decisions, identification comes from laws enforced within the calendar year. The RI p-value is the share of 1,000 random reassignments of enforcement dates across states yielding an absolute coefficient at least as large as the estimate. The pre-trend p-value tests that the exposure change over the following six months does not predict clicks in column (1), and that event-study coefficients for quarters -4 to -2 are jointly zero in column (2). The No technology group contains 33 lenders. Standard errors, clustered at the lender level, are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

partly offset the loss. The reallocation of borrowers induced by reduced advertising also transmits composition changes to non-adopting lenders, which is what the no-technology rejection estimate captures.

8 Robustness

The empirical analysis and the model treat lenders that collect online data directly through their own websites as an extension of lenders that acquire such data from third-party ad-tech platforms. These advertising-and-data lenders have greater screening and profiling capacity, but the underlying data source is the same, so the paid-search-click regression already includes them. Appendix G.3 re-estimates the main analysis on the Equifax–HMDA data for this group alone. It confirms parallel trends (Figure G.2) and validates the estimates through a wild-cluster bootstrap. Almost all the baseline results carry over, though with smaller magnitudes. Enforcement reduces the share of searcher borrowers by 0.49 percentage points, or 1.96% of the baseline mean, versus 0.85 percentage points (3.4%) in the sample of advertisement-only lenders . It lowers rejection rates by 0.14 percentage points, or 2% of the mean, against 0.22 percentage points (2.9%) in the previous sample (Table G.2). The attenuation reflects partial exposure to privacy laws, since some of these lenders' borrowers still consent to online data collection, allowing imputation on the remaining sample.

The interest-rate results are the one point of divergence. The reduction for searchers is 11.7 basis points, or 2.87% of the mean rate, rather than 14.0 basis points (3.43%) for the full sample, and non-searchers also show a statistically significant 9.5 basis point (2.3%) reduction (Table G.2). This pattern is consistent with imperfect classification at the search cutoff: if some true searchers are assigned to the non-searcher group, part of the searcher rate reduction is carried over into the non-searcher estimate. The following paragraphs address this concern through a sensitivity analysis that varies the classification threshold. Only the one-year-ahead bankruptcy rate does not carry over, and it is small and statistically insignificant for this group, with a point estimate of 0.39 percentage points versus 2.09 percentage points in the full sample. A plausible explanation is that these lenders can classify website-collected data as essential to the origination process, thereby bypassing the privacy-law restrictions that bind data acquired from third-party platforms.

The searcher and non-searcher groups are defined by a cutoff on the imputed search propensity $\widehat{\pi}_i^P$, and this cutoff is not uniquely determined. I therefore re-estimate the interest-rate effects at the 25th, 50th, and 75th percentiles of the full incoming-loan distribution ($c = 0.19$, 0.24, and 0.28), fixed before splitting by lender technology. Table G.7 reports the results. The pattern does not depend on the cutoff. Among ad-tech lenders, the rate reduction is present at every threshold and is larger for high-search borrowers than for low-search borrowers, for instance -0.14 against -0.04 at $c = 0.19$ and -0.21 against -0.06 at $c = 0.24$. Among no-tech lenders, the effect for high-search

borrowers is statistically indistinguishable from zero at every threshold. The point estimates move with the cutoff, as expected. Raising c reassigns borrowers from the high-search group to the low-search group, so the two columns become more similar as the split becomes less selective. This is the attenuation described in the introduction: imperfect classification mixes true searchers and non-searchers and compresses the contrast between them. The presence of this attenuation, rather than any single estimate, is what makes the heterogeneity credible. The ordering does not reverse at any cutoff, which would be difficult to generate from classification error unless that error were systematically related to enforcement.

The results and mechanism sections established parallel trends for the main specification. To validate them formally, I apply the HonestDiD procedure of Rambachan and Roth (2023), reporting for each estimate the breakdown value beyond which the robust confidence set no longer excludes zero, under both the relative-magnitudes ($\bar{M}$) and smoothness (M) restrictions (Table G.1). Under the smoothness restriction $\Delta^{SD}(M)$, which fits the flat pre-trends documented above, the core results hold across the full grid. This covers the searcher interest-rate effect for Ads+Data lenders, all three search-for-price effects, the bankruptcy effect for Ads lenders, and both significant rejection-rate effects. Under the stricter relative-magnitudes restriction $\Delta^{RM}(\bar{M})$ the effects are more fragile, breaking down below $\bar{M} = 0.5$ (the bankruptcy effect for Ads lenders, at $\bar{M} = 0.66$, is the exception), as expected when one or two pre-periods are noisy. The few-treated-clusters concern is handled separately through the wild-cluster bootstrap, so I do not pursue it here.

Appendix G.1 examines anticipation using signature dates rather than enforcement dates (Figure G.1). Because enforcement follows signature by four to eight quarters (median: six), the post-signature estimates through the third quarter are free of enforcement effects, whereas the fourth-quarter estimate may partly include them. The event studies show some movement before enforcement, most notably a decline in the searcher share of Only-ads lenders beginning two quarters after signature. The interest rate paid by non-searchers at no-technology lenders also spikes three quarters after signature and then partly reverts; the remaining post-signature estimates are small. These pre-enforcement movements may reflect early adoption of compliance practices. The period between signature and enforcement allows firms to build consent and opt-out systems, and lenders and advertising platforms that complete this work early curtail targeted advertising before the law binds. If so, the omitted pre-enforcement quarter is already partly treated, which biases the enforcement-date estimates toward zero and makes the baseline effects conservative.

Appendix G.6 examines whether the results reflect transitory shocks associated with the COVID-19 pandemic, whose onset coincided with the enforcement of the California law. Aggregating the post-enforcement period by year, I find that the declines in the searcher share and in searchers' interest rates at adopting lenders persist, or deepen, through 2024 rather than reverting after 2020 (Table G.5; Figures G.8 and G.9). Although some pre-enforcement estimates are significant, this

persistence is inconsistent with a purely transitory pandemic effect. Gupta et al. (2024) and Doerr et al. (2026) provide separate evidence on the pandemic's role in the effects of the California law.

A related concern is that California, which has the most stringent law in the sample, drives the results. Table G.4 therefore reports the main outcomes for each of the six enforcement cohorts, except paid-search clicks, which are observed only at the national lender-month level. The decline in the searcher share is significant in three cohorts for Ads+Data lenders (California, Colorado–Connecticut, and Montana) but only in California for Only-ads lenders. The pricing evidence also extends beyond California: searchers' rates decline significantly at Only-ads lenders in California and Colorado–Connecticut, and at Ads+Data lenders in California and Montana, with a marginally significant decline in Florida–Oregon–Texas. Bankruptcy rises significantly among Only-ads borrowers in California, Virginia, and Colorado–Connecticut. The estimate for Florida–Oregon–Texas is imprecise, and Montana cannot yet be estimated, because bankruptcy is measured one year after origination. Utah stands out: Ads+Data lenders attract a larger share of searchers and charge them higher rates, and bankruptcy declines in both technology groups. Virginia shows a smaller, marginally significant rate increase among Ads+Data searchers. Utah's divergence may reflect its comparatively permissive framework. It is the only law in the sample that imposes no restrictions on sensitive data such as precise geolocation; it also lacks a global opt-out requirement and grants firms a permanent 30-day cure period. This pattern is consistent with the importance of the legal environment emphasized by Doerr et al. (2026). Finally, rejection rates decline significantly only in California, which may reflect its larger mortgage market, its stricter regulation, and the annual frequency of the rejection data. Overall, stringency shapes the magnitude and, in Utah, the direction of the estimates, but the main pricing and borrower-allocation effects are not confined to California.

The staggered design also requires that enforcement in treated states does not spill over to control states. To test this assumption, Table G.6 examines whether a lender's exposure to privacy laws elsewhere affects its outcomes in states that have neither signed nor enforced a privacy law. I interact the lender-level exposure measure from the advertising-channel analysis with the lender's technology group, retain fixed effects comparable to the baseline, and cluster standard errors at the lender level. The main evidence of spillovers comes from borrower allocation. Greater exposure is associated with a higher searcher share at no-technology lenders, whereas the interactions for Only-ads and Ads+Data lenders are negative and the implied net effects for these lenders are small. The interest-rate interactions for Only-ads lenders are also negative but only marginally significant, and rejection and bankruptcy show no significant response. These patterns are consistent with economies of scale in online data. Targeting models are trained on data aggregated across all the markets a lender serves, so privacy laws that shrink the data from treated states also degrade targeting in untreated ones. More exposed lenders therefore reach fewer searchers through targeted advertising

even where no law applies, which shifts searchers toward non-adopters, while Ads+Data lenders, whose own website data partly substitute for third-party data, show no rate response. Nationwide compliance, whereby firms extend the strictest state rules to all customers, would produce a similar pattern. In either case, spillovers are largely limited to borrower allocation and, because they move control-state outcomes in the same direction as the treatment effects, would if anything attenuate the baseline estimates.

Finally, because the Sun–Abraham estimator does not accommodate interactions with lender type, I obtain baseline estimates separately for each technology group, and separate regressions on correlated subsamples may distort cross-group comparisons. Appendix G.4 therefore re-estimates the main results in a single pooled stacked ETWFE specification with cohort-specific controls and fixed effects (Table G.3). The estimates closely reproduce the baseline. Enforcement raises the searcher share at no-technology lenders by 2.09 percentage points and lowers it at Only-ads and Ads+Data lenders by 0.87 and 0.50 percentage points, against 2.07, 0.85, and 0.49 in the baseline. Searchers at Only-ads lenders pay 13.97 basis points less (baseline: 13.95), and those at Ads+Data lenders 11.65 basis points less (baseline: 11.65), while the effect for searchers at no-technology lenders remains insignificant (4.84 versus 4.81 basis points). Bankruptcy remains statistically unchanged at no-technology lenders (−0.66 versus −0.64 percentage points) and rises at Only-ads lenders (2.01 versus 2.09). Rejection rates decline by 0.53 and 0.22 percentage points at no-technology and Only-ads lenders, respectively, matching the baseline. The pooled specification thus leaves the sign and magnitude of the main results essentially unchanged.

I report the robustness checks specific to the advertising mechanism and the estimation of the search-for-price probability in the respective Mechanism (Section 7) and Empirical Strategy (Section 5) sections. I omit them here for simplicity.

9 A Model of Screening and Targeted Advertising

Privacy regulation affects two uses of online-collected data with opposite implications for mortgage rates. Better information improves screening, lowers expected credit losses, and therefore lowers rates within a lending channel. It also helps technology lenders reach borrowers through targeted advertising before they compare outside offers. Because a borrower must incur the search friction K_i to access the competitive comparison market, early contact creates room for the targeted lender to charge above the competitive search rate while still deterring search. Without a targeted offer, low-K borrowers compare, while high-K borrowers remain with the incumbent. Privacy therefore raises rates within channels through weaker screening, but can lower the observed average rate by shifting originations away from the higher-rate advertising channel. Section 10 maps this two-lender economy into the lender groups observed in the data.

9.1 Environment, Information, and Timing

This subsection defines the agents and the sequence of decisions. The model is solved within a credit-score bucket b. There is a unit mass of prospective mortgage borrowers, each requiring a mortgage with principal normalized to one. Lenders are risk neutral and come in two analytical types: $\tau \in \{N, T\}$. Conventional lenders N do not use online-collected data. Technology lenders T use these data for screening and early targeted acquisition. Each borrower begins with a home lender. Let $h_N + h_T = 1$ denote the pre-underwriting home-lender shares. The richer distinction between Only Ads and Ads+Data lenders is introduced only in Section 10.

Borrowers differ in repayment type and in the friction governing active comparison. Repayment type is $\nu_i \in \{S, R\}$, with $\Pr(\nu_i = R) = \mu_R$ and $\mu_S = 1 - \mu_R$. Safe borrowers S repay. Risky borrowers R default. This binary support is a tractability device: the mechanism requires only that risky borrowers generate larger expected losses. Comparison type is $K_i \in \{K_\ell, K_h\}$, where $0 < K_\ell < K_h$ and $\Pr(K_i = K_h) = \lambda \in (0, 1)$. In the analytical benchmark, repayment type, comparison type, and home-lender type are independent. Conditional on a live approved mortgage at rate r, borrower i has rate-equivalent decision utility

$$u_i(r, z) = v_i - r - K_i z, \qquad z \in \{0, 1\}, \tag{9.1}$$

where $z = 1$ denotes active comparison. The participation value v_i is high enough that participation is nonbinding once a live approved offer exists. Thus K_i is an effective stopping wedge: it may reflect time and effort, but also attention, inertia, primacy, or other non-resource frictions. Low-K and high-K refer throughout this section to latent theoretical types, not to the empirical Searcher classification.

Lenders offer mortgage rates through three channels: active comparison, where borrowers seek outside offers; incumbent retention, where borrowers remain with their home lender; and paid acquisition, where a lender reaches a borrower through targeted advertising. Rates are strategic choices. Screening precision, usable information, technology adoption, and capacity are primitives of the analytical model. A current loan originated at rate r yields lender payoff $r - \mathrm{MC}_{\tau b}$. A paid conversion may additionally generate a private continuation value $B \geq 0$.[4]

Let $I \geq 0$ denote usable online information. It affects exactly two primitives. First, it improves the technology lender's ability to classify repayment risk. Second, it raises the probability that a technology lender contacts and prequalifies a borrower before active comparison. Information I does not enter borrower utility, does not mechanically change K_i, and does not directly set mortgage

[4]The sole role of B is to select the pooled paid-acquisition equilibrium in the maintained pricing region. Appendix H.3 derives a sufficient supporting value. Within that region, B does not affect the model-implied rates, quantities, current-loan margins, or reported surplus measures.

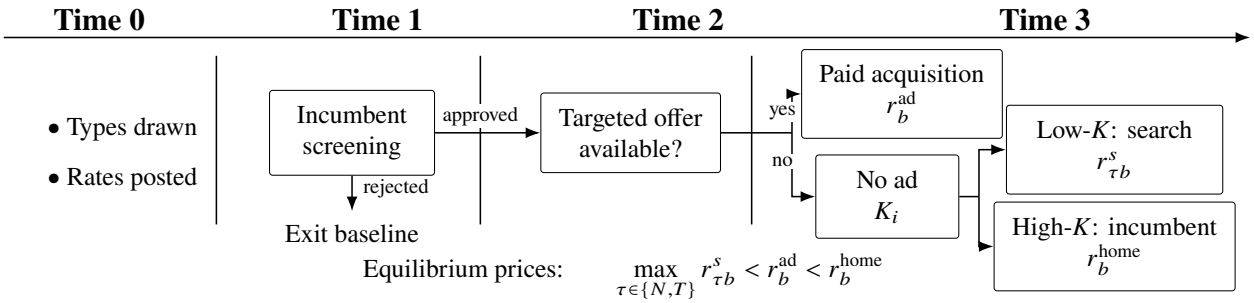


Figure 9.1. Timing, borrower routing, and equilibrium channel prices. The home lender screens first, and rejected borrowers exit the baseline. Conditional on approval, a technology lender may reach and screen the borrower. A targeted offer is available only if that screening results in approval. If no targeted offer is available, low-K borrowers search and high-K borrowers remain with the incumbent. Search lenders also screen the borrower, and the model conditions on two approved outside quotes. In equilibrium, search rates are lowest, targeted-offer rates are intermediate, and incumbent rates are highest.

rates. Its effects operate through screening and the timing of contact.

Borrowers observe their comparison type, approval outcomes, and all live offers available when they choose. Lenders observe the credit-score bucket, aggregate type distributions, lender shares, and the information environment, but they do not observe an individual borrower's repayment or comparison type when rates are posted. Each lender observes its own screening outcome before issuing an approved offer. Every lender therefore screens the borrower before making an executable mortgage offer. Hence rates cannot be conditioned directly on latent comparison type.

Assumption 1 (Timing and information)**.** Lenders first post score-bucket comparison, targeted-offer, and incumbent rate schedules. The borrower then applies to the home lender and is screened. Conditional on home-lender approval, a technology lender may contact, screen, and successfully prequalify the borrower through targeted advertising. If the borrower instead searches, she pays K_i and obtains two approved, executable outside quotes after those lenders' screening. Lenders know aggregate type distributions but not the borrower's individual repayment or comparison type when rates are posted.

Rejected borrowers exit the baseline, keeping the model focused on price comparison after approval rather than repeated search for credit. Approval still depends on lender screening, while the borrower's search friction K_i determines the next step: low-K borrowers are more willing to compare outside offers, whereas high-K borrowers are more likely to remain with the incumbent. Appendix H.6 allows search after rejection and relates the extension to Agarwal et al. (2024).

9.2 Screening and the Competitive Comparison Market

This subsection shows how online information changes technology-lender screening and, through screening, the cost of lending and comparison rates. Conventional lenders, indexed by N, do not use online-collected data. They approve safe and risky applicants with probabilities q_{NS} and q_{NR}, respectively, where $0 < q_{NR} < q_{NS} < 1$.

Technology lenders, indexed by T, use online information I to screen applicants more precisely. Their approval probabilities are

$$\begin{aligned} q_{TS}(I) &= q_{NS} + \alpha_S(I), \\ q_{TR}(I) &= q_{NR} - \alpha_R(I), \end{aligned} \tag{9.2}$$

where $\alpha_S(0) = \alpha_R(0) = 0$ and $\alpha'_S(I), \alpha'_R(I) > 0$, subject to bounds that keep approval probabilities in $(0, 1)$. Thus, more information allows technology lenders to approve more safe borrowers, through α_S, and reject more risky borrowers, through α_R.

For lender type τ, approval, funded-pool risk, and resource marginal cost are

$$A_\tau = \mu_S q_{\tau S} + \mu_R q_{\tau R}, \qquad \rho_\tau = \frac{\mu_R q_{\tau R}}{A_\tau}, \qquad \mathrm{MC}_{\tau b} = f_b + d\rho_{\tau b}, \tag{9.3}$$

where $d \in (0, 1]$ is loss given default and f_b collects funding, servicing, and operating costs. Better information unambiguously improves the composition of the technology lender's funded pool. Its effect on the overall approval rate, however, is ambiguous. Approval rises if the increase in approvals of safe borrowers dominates the decline in approvals of risky borrowers, and falls otherwise: $A'_T(I) = \mu_S \alpha'_S(I) - \mu_R \alpha'_R(I)$.

Assumption 2 (Competitive comparison market)**.** If a borrower shops for another mortgage offer, she receives two live outside offers. Each offer comes from either a conventional lender N or a technology lender T. Within each lender type, at least two lenders compete to make these offers. Lenders of the same type have the same marginal cost within a credit-score bucket, are risk neutral, and face no binding capacity constraints.

Bertrand competition within each lender-type quote pool implies

$$r^s_{\tau b} = \mathrm{MC}_{\tau b} = f_b + d\rho_{\tau b}. \tag{9.4}$$

Let $s_T \in (0, 1)$ denote the probability that any one of the borrower's two outside offers comes from a technology lender. A searching borrower draws two outside lender types independently, so a technology lender appears at least once with probability $\beta(s_T) = 1 - (1 - s_T)^2$. Since $r^s_{Tb} < r^s_{Nb}$,

the expected funded comparison rate is

$$\widetilde{r}_b^s = \beta(s_T) r_{Tb}^s + [1 - \beta(s_T)] r_{Nb}^s. \tag{9.5}$$

Lemma 1 (Screening and comparison rates)**.** *For $I > 0$, technology lenders originate a safer pool than conventional lenders, so funded-pool risk is lower: $\rho_T < \rho_N$. Holding s_T fixed, more usable information lowers the technology lender's funded-pool risk ρ_T, marginal cost MC_{Tb}, shopping rate r_{Tb}^s, and the borrower's expected shopping rate $\widetilde{r}_b^s$. Its effect on the technology lender's overall approval rate A_T is ambiguous.*

Among borrowers who shop for outside offers, Bertrand competition pins the technology-lender quote to marginal cost. Better screening therefore lowers both the technology-lender search rate r_{Tb}^s and the expected rate a borrower obtains from search, $\widetilde{r}_b^s$. Proof is in Appendix H.2.

9.3 Incumbent and Targeted-Offer Pricing

This subsection explains why mortgage rates differ across three acquisition channels: search, incumbent retention, and targeted advertising. Consider a borrower who has already been screened and approved by a conventional home lender N. Before she searches, a technology lender may reach her through targeted advertising and screen her using online information. Let $\gamma(I) \in (0, 1)$ denote the probability that this process results in a successful prequalification, with $\gamma'(I) > 0$.

Screening also occurs in the search channel. If the borrower chooses to search, each outside lender screens her before making an offer; the model conditions on the two approved, executable quotes that result from this process. Thus every lending channel involves screening. The channels differ in timing: the targeted technology lender reaches and screens the borrower before she searches, whereas outside lenders screen her only after she incurs K_i and enters the comparison market. The targeted offer is no longer available once search begins.

Conditional on an approved incumbent offer, generalized borrowing cost equals the mortgage rate plus any cost of searching for outside offers. Without a targeted advertising offer, borrower i chooses between staying with the incumbent and searching:

$$C_i^0 = \min\left\{r_b^{\text{home}}, \widetilde{r}_b^s + K_i\right\}. \tag{9.6}$$

If a targeted advertising offer is available, the borrower can instead choose among accepting that offer, staying with the incumbent, or searching:

$$C_i^{\text{ad}} = \min\left\{r_b^{\text{ad}}, r_b^{\text{home}}, \widetilde{r}_b^s + K_i\right\}. \tag{9.7}$$

Define the two reservation rates $r_{\ell b} = \widetilde{r}_b^s + K_\ell$ and $r_{hb} = \widetilde{r}_b^s + K_h$. At $r_{\ell b}$, a low-K borrower is indifferent between accepting the current offer and searching. At r_{hb}, the same is true for a high-K borrower. Because $K_h > K_\ell$, we have $r_{hb} > r_{\ell b}$.

Let $Q_{\tau b}^{\text{home}}(r)$ and $Q_{Tb}^{\text{ad}}(r)$ denote the demand generated by an incumbent rate and a targeted advertising rate, respectively. A paid acquisition may also generate a continuation value $B \geq 0$ for the technology lender, such as future business from the newly acquired relationship.[5] Incumbents and the targeted technology lender solve

$$\begin{aligned} r_{\tau b}^{\text{home}} &\in \arg\max_{r \in \mathcal{R}_b \cup \{\emptyset\}} Q_{\tau b}^{\text{home}}(r)\big[r - \text{MC}_{\tau b}\big], \\ r_{Tb}^{\text{ad}} &\in \arg\max_{r \in \mathcal{R}_b \cup \{\emptyset\}} Q_{Tb}^{\text{ad}}(r)\big[r - \text{MC}_{Tb} + B\big]. \end{aligned} \tag{9.8}$$

Because K_i has two values, demand is stepwise. Smooth first-order conditions therefore add no content: each lender compares a finite set of reservation-rate candidates and exit.

Assumption 3 (Pricing region and tie-breaking)**.** The rate grid is sufficiently fine and contains the two search thresholds

$$r_{\ell b} = \widetilde{r}_b^s + K_\ell, \qquad r_{hb} = \widetilde{r}_b^s + K_h.$$

Each threshold equals the expected mortgage rate from searching, $\widetilde{r}_b^s$, plus the borrower's search cost. Thus, a low-K borrower is indifferent between accepting a rate $r_{\ell b}$ and searching, while a high-K borrower is indifferent at r_{hb}.

We assume that $r_{\ell b}$ exceeds every realized search quote. The incumbent prefers charging r_{hb} and retaining high-K borrowers rather than charging $r_{\ell b}$ and retaining both types. The targeted technology lender instead prefers charging $r_{\ell b}$ and serving all borrowers it reaches rather than charging more and serving only high-K borrowers. Ties between the targeted offer and search go to the targeted offer. Ties involving the incumbent go to the incumbent.

Appendix H.3 states the exact candidate payoffs, deviation inequalities, rate-grid details, and the sufficient supporting value of B. Conditional on the comparison-rate schedule, the incumbent therefore prices against the stopping threshold of a high-friction borrower, whereas the targeted lender prices against the comparison threshold of a low-friction borrower.

9.4 Equilibrium Prices, Quantities, and Privacy

This subsection characterizes equilibrium mortgage rates and borrower allocation across acquisition channels. It then shows that privacy can raise rates within a given channel while lowering the

[5] The role of B is only to support the maintained pricing region. Within this region, it does not affect equilibrium rates or quantities.

average rate charged by technology lenders, because fewer borrowers are acquired through targeted advertising and more obtain technology-lender loans through the lower-rate search channel.

Definition 1 (Equilibrium)**.** Given usable information I, an equilibrium specifies mortgage rates in the three acquisition channels—search, incumbent retention, and targeted advertising—together with borrower choices and the resulting approval and origination quantities. It satisfies four conditions: (i) search-market quotes are Bertrand best responses within each lender type; (ii) incumbent and targeted-advertising rates maximize lender payoffs given borrower demand; (iii) at each decision point, borrowers choose the option with the lowest generalized borrowing cost; and (iv) the resulting approvals, targeted acquisitions, search, incumbent retention, and originations are consistent with borrower types, lender shares, screening probabilities, and the process that determines which lenders appear among the borrower's outside offers. I focus on pure-strategy Bayesian equilibria in the pricing region of Assumption 3.

The benchmark with no search frictions makes the source of the advertising rent transparent. If $K_\ell = K_h = 0$, then both search thresholds equal $\widetilde{r}_b^s$. A technology lender that reaches the borrower before she shops therefore cannot charge more than the rate she expects to obtain from searching. If $\gamma = 0$, no borrower is reached through targeted advertising, so the paid-acquisition channel disappears. The mechanism thus requires both early contact through advertising and a positive cost of searching for outside offers.

Proposition 1 (Early-acquisition pricing and allocation)**.** *Suppose Assumptions 1–3 hold. Then the maintained pricing region supports a pure-strategy equilibrium in which*

$$r_b^{\text{home}} = \widetilde{r}_b^s + K_h, \qquad r_b^{\text{ad}} = \widetilde{r}_b^s + K_\ell, \qquad \max_{\tau \in \{N,T\}} r_{\tau b}^s < r_b^{\text{ad}} < r_b^{\text{home}}. \tag{9.9}$$

Every successfully targeted borrower accepts the targeted offer. Without a targeted offer, low-K borrowers enter active comparison and high-K borrowers remain with the incumbent.

The proposition delivers a simple ranking across acquisition channels: search rates are lowest, targeted-advertising rates are intermediate, and incumbent rates are highest. The premium of the targeted-advertising rate over the search rate is the acquisition rent from reaching borrowers before they shop. Proof is in Appendix H.3.

The equilibrium allocates borrowers across three channels: paid acquisition, organic comparison, and incumbent retention. Appendix H.4 reports the complete quantity system and verifies adding-up.

Let w_T^s, w_T^{ad}, and w_T^{home} denote the shares of technology-lender originations arising through comparison, paid acquisition, and incumbent retention, respectively, with $w_T^s + w_T^{\text{ad}} + w_T^{\text{home}} = 1$.

The average technology-lender mortgage rate is

$$\bar{r}_T = w_T^s r_T^s + w_T^{\text{ad}} r^{\text{ad}} + w_T^{\text{home}} r^{\text{home}}. \tag{9.10}$$

Thus observed mortgage rates depend both on prices within acquisition channels and on how borrowers are allocated across those channels.

A privacy restriction lowers usable information from I to $I' < I$ through two channels. First, weaker screening raises technology lenders' expected lending costs and therefore their search rate. This raises the expected rate from search and, because advertising and incumbent rates are priced relative to that outside option, also raises r^{ad} and r^{home}. Second, privacy reduces targeted advertising, $\gamma(I') < \gamma(I)$, shifting borrowers away from paid acquisition toward organic search or the incumbent channel.

For a finite policy change, let $\Delta x = x' - x$, where primes denote the post-privacy equilibrium. Using $w_T^s + w_T^{\text{ad}} + w_T^{\text{home}} = 1$, the exact change in the average technology-lender mortgage rate is

$$\begin{aligned} \Delta \bar{r}_T = \; & \underbrace{w_T^{s\prime} \Delta r_T^s + w_T^{\text{ad}\prime} \Delta r^{\text{ad}} + w_T^{\text{home}\prime} \Delta r^{\text{home}}}_{\text{within-channel screening effects}} \\ & + \underbrace{\left(r^{\text{ad}} - r_T^s\right) \Delta w_T^{\text{ad}} + \left(r^{\text{home}} - r_T^s\right) \Delta w_T^{\text{home}}}_{\text{channel reallocation}}. \end{aligned} \tag{9.11}$$

The first component is positive: weaker screening raises rates within each channel. The second captures the change in the composition of originations across channels. Because $r_T^s < r^{\text{ad}} < r^{\text{home}}$, shifting borrowers toward organic comparison lowers the average rate, whereas shifting them toward incumbent retention raises it.

Proposition 2 (Privacy and the measured-rate reversal)**.** *Let privacy reduce usable online-collected information from I to $I' < I$. In the maintained pricing region, all technology-lender channel rates rise, yet the average technology-lender mortgage rate falls if and only if*

$$-\left[\left(r^{\text{ad}} - r_T^s\right) \Delta w_T^{\text{ad}} + \left(r^{\text{home}} - r_T^s\right) \Delta w_T^{\text{home}}\right] > w_T^{s\prime} \Delta r_T^s + w_T^{\text{ad}\prime} \Delta r^{\text{ad}} + w_T^{\text{home}\prime} \Delta r^{\text{home}}. \tag{9.12}$$

The average rate therefore falls when the rate-reducing effect of borrower reallocation across channels dominates the increase in rates within channels from weaker screening. The aggregate approval response remains unsigned. Proof is in Appendix H.4.

10 Quantitative Implementation, Identification, and Estimation

The analytical model groups lenders into conventional (N) and technology (T) types to isolate the mechanism. In the quantitative implementation, I split technology lenders into Only Ads (A) and Ads+Data (D) to map the model to the lender groups observed in the data. The implementation uses six credit-score buckets and replaces the scalar information stock in the analytical model with the observed pre- and post-enforcement policy states. Appendix Table I.3 reports the score-bin primitives and the construction of the main data-to-model inputs.

Let $u \in \{0, 1\}$ index the privacy-enforcement state, with $u = 0$ denoting pre-enforcement and $u = 1$ post-enforcement. For lender type $g \in \{N, A, D\}$ and credit-score bucket b, marginal resource cost and the borrower-facing comparison quote are

$$\mathrm{MC}_{gbu} = f_b + d\rho_{gbu}, \qquad r^s_{gbu} = f_b + \kappa_\rho d\rho_{gbu}. \tag{10.1}$$

Here ρ_{gbu} is the observed one-year-ahead bankruptcy rate of the funded pool and serves as the empirical proxy for funded-pool risk. The analytical Bertrand benchmark has $\kappa_\rho = 1$. In the quantitative implementation, however, the second relation in equation (10.1) is an empirical measurement equation for borrower-facing comparison quotes rather than the analytical zero-profit restriction $r^s = \mathrm{MC}$. The parameter κ_ρ measures the estimated pass-through of the risk proxy into comparison quotes, while resource-cost accounting throughout uses MC_{gbu}. Appendix I.2 describes its estimation.

Let $\widetilde{r}^s_{bu}$ denote the expected minimum of two approved outside quotes. The targeted-offer and incumbent rates retain the threshold-pricing structure of the analytical model:

$$r^{\mathrm{ad}}_{bu} = \widetilde{r}^s_{bu} + K_\ell, \qquad r^{\mathrm{home}}_{bu} = \widetilde{r}^s_{bu} + K_h. \tag{10.2}$$

Appendix I.1 gives the exact expected-minimum quote in equation (I.1) and the complete allocation and channel-quantity construction.

The quantitative model distinguishes the share of comparison originations funded by group g, π^s_{gbu}, from its quote-draw share, s^q_{gbu}. Because the data provide only one local lender-share measure, the baseline sets $\pi^s_{gbu} = s^q_{gbu}$. This is a measurement normalization, not an equilibrium restriction. The corresponding allocation and adding-up restrictions are given in Appendix I.1.

The three-lender implementation adds a paid-contact matrix $P_{ht,b}$, where target lender type t may contact a borrower initially attached to incumbent type h. I impose the acquisition-only convention $P_{hh,b} = 0$. The parameter τ_g allows paid contact to vary by latent comparison type, while ζ_g captures post-enforcement changes in the composition of organic comparison. Post-enforcement, paid contact directed toward target type t is scaled by a retention parameter δ_t. I calibrate these

retention rates from the independent search-advertising response in Table 7.2, rather than from mortgage-rate moments. Appendix equations (I.3) and (I.4) give the exact pre- and post-enforcement contact matrices.

Let

$$\lambda_b = \Pr(K_i = K_h \mid b)$$

denote the latent high-friction share in score bin b. The empirical Searcher indicator Z_i is a noisy measure of latent comparison type, summarized by

$$\eta_{L,b} = \Pr(Z_i = 1 \mid K_i = K_\ell, b), \qquad \eta_{H,b} = \Pr(Z_i = 1 \mid K_i = K_h, b).$$

Appendix equation (I.6) gives the latent-to-observed Searcher mapping, while Appendix Table I.3 reports the corresponding score-bin inputs.

A final quantitative ingredient describes where borrowers go when privacy eliminates a paid contact. For each technology group $g \in \{A, D\}$, define

$$\omega_g = \Pr(\text{organic comparison} \mid \text{paid contact disappears},\ g,\ Z_i = 1)\,. \tag{10.3}$$

Within the empirical Searcher sample, a fraction ω_g of borrowers displaced from paid acquisition enters organic comparison, while the remaining fraction $1 - \omega_g$ returns to the incumbent channel. The parameter ω_g therefore governs only the destination of borrowers who lose a paid contact. By contrast, ζ_g captures the broader post-enforcement change in the composition of all borrowers reaching the organic-comparison market. Appendix equations (I.8) and (I.16) formalize this distinction.

The quantitative implementation combines direct empirical inputs, external calibrations, pre-enforcement estimates, and parameters inferred from privacy-law responses. Appendix Table I.1 summarizes the mapping between model objects, empirical moments, and estimation status; Appendix Table I.2 reports the underlying empirical moments; and Appendix Table I.3 reports the score-bin primitives and data construction.

Estimation proceeds sequentially. I construct the screening states from approval and bankruptcy moments, estimate κ_ρ from pre-enforcement rate–risk covariation, calibrate paid reach Γ, discipline τ_g and ζ_g from shopping composition, profile K_ℓ, recover f_b and K_h from pre-enforcement rate levels, and finally infer ω_A and ω_D from the technology-group Searcher-rate responses. The Only Ads and Ads+Data Searcher-rate effects are therefore targeted moments; the No-Tech Searcher-rate effect is held out for validation.

Appendix Table I.4 reports the resulting parameter vector, Appendix Table J.3 reports the cell-level pre-enforcement fit, and Appendix Table I.5 reports the numerical and identification

diagnostics. Appendix I.2 provides the identifying equations, the estimation algorithm, and the conditional ATT-implied ranges reported for ω_A and ω_D. Section 11 reports model fit, the mechanism decomposition, and rerouting estimates. The counterfactuals are partial equilibrium and hold lender types, entry, long-run technology adoption, and aggregate funding conditions fixed.

11 Quantitative Results

11.1 Model Fit and Mechanism Decomposition

Table 11.1 reports the model fit and mechanism decomposition. The Only Ads and Ads + Data Searcher-rate treatment effects identify ω_A and ω_D and are therefore matched by construction, while the No-Tech effect is not targeted and provides a held-out validation moment. The model separates two opposing forces: weaker screening raises expected lending costs and puts upward pressure on rates, while the loss of targeted acquisition reallocates borrowers across organic search and incumbent lending. For technology lenders, this reallocation shifts displaced borrowers toward lower-rate organic search and more than offsets the screening effect, accounting for the decline in average mortgage rates. No-Tech lenders' own screening technology remains unchanged, but their rates nevertheless respond to changes in the composition of borrowers reaching the search market. Panel B reports where borrowers go when a paid contact is lost: ω_g is the fraction of displaced borrowers who continue into active comparison rather than return to the home-bank channel. For Only Ads lenders, about 63% continue searching, with a conditional range of 55–72%; for Ads + Data lenders, the point estimate implies nearly complete fallback into search, with a conditional range of 77–100%. These routing responses determine how the loss of targeted acquisition translates into the average rate effects reported in Panel A.

The model fits pre-enforcement mortgage-rate levels with a weighted RMSE of 11.02 basis points, and the largest absolute cell-level residual is 24.46 basis points. As a held-out validation moment, the model predicts a 4.25 basis-point increase in mortgage rates for No-Tech lenders, close to the empirical point estimate of 4.81 basis points (s.e. 4.81). The model also preserves the strict ordering $r_{\text{search}} < r_{\text{ad}} < r_{\text{home}}$ in every applicable lender-group-by-score-bin-by-policy cell. Appendix Table J.2 profiles K_ℓ and shows that the baseline value of 125 basis points delivers the best pre-enforcement rate-level fit within the region that jointly rationalizes the two technology-group mortgage-rate treatment effect.

Appendix Table J.1 reports model-implied mortgage rates by lender type and acquisition channel before and after privacy-law enforcement. Consistent with the model's pricing structure, organic search offers the lowest rates, followed by paid advertising and then home-bank incumbent offers. Enforcement raises rates within each acquisition channel for technology lenders, reflecting the

Table 11.1. Fit, Mechanism Decomposition, and Rerouting

	Mortgage-rate effect (basis points)					
Group	Data ATT	S.E.	Model	Screening	Reallocation	Calibrated?
Panel A. Mortgage-rate effects						
No Tech	4.81	4.81	4.25	1.52	2.72	No
Only Ads	−13.95	2.50	−13.95	3.62	−17.58	Yes
Ads + Data	−11.65	1.77	−11.65	2.81	−14.47	Yes

Panel B. Destination of displaced paid contacts			
Group	ω_g	Conditional range	Interpretation per 100 lost contacts
Only Ads	0.63	[0.55, 0.72]	About 63 compare; 37 return to the incumbent.
Ads + Data	1.00	[0.77, 1.00]	Nearly all compare; as many as 23 may return.

Notes: Panel A compares empirical and model-implied effects of privacy-law enforcement on the mortgage rates paid by Searchers, defined as borrowers with predicted search propensity $\hat{\pi} \geq 0.19$, in basis points. Data ATTs and standard errors are from Tables 6.2 and G.2. The model effect is decomposed into a screening component, the rate change induced by weaker screening with acquisition channels held fixed, and a reallocation component, the remaining change from shifts across paid advertising, organic search, and the incumbent channel. No-Tech lenders' own screening technology is unchanged; their screening component captures indirect effects of weaker screening at technology lenders. The search-fallback parameters ω_g for Only Ads and Ads + Data are calibrated to match their groups' Searcher rate effects exactly (Appendix I.2); the No-Tech effect is not targeted. Panel B reports ω_g, the fraction of displaced paid contacts among Searchers that enter active comparison rather than return to the incumbent. Conditional ranges invert the 95% confidence interval of the corresponding empirical rate ATT, holding the other structural parameters at their calibrated values. They therefore reflect uncertainty in the targeted moment only and are not joint confidence intervals for the structural parameters. The Ads + Data range is truncated at the upper bound $\omega_g = 1$.

increase in expected lending costs associated with weaker screening. Thus, the decline in average mortgage rates for these lenders does not arise from lower prices conditional on acquisition channel. Instead, it is generated by borrower reallocation across channels, as displaced borrowers move toward lower-rate search opportunities.

11.2 Reallocation and Surplus Effects of Privacy Laws

Privacy affects both credit access and borrower allocation across acquisition channels. Appendix Figure J.1 shows heterogeneous approval responses across score bins; aggregating across bins and lender groups, approval falls by approximately 0.48 percentage points. Appendix Figure J.2 shows that paid acquisition falls by roughly 24–29 percentage points, with borrowers reallocating primarily toward organic comparison and secondarily toward the incumbent. I report credit-access effects separately because the surplus analysis below is conditional on origination.

The comparison friction K_i can reflect both real search costs—the time and effort required to obtain and compare offers—and non-resource stopping frictions such as inattention or inertia. Borrower choices identify the total wedge, but not its composition. Let $\theta \in [0, 1]$ denote its resource share:

$$K_j^{\text{resource}} = \theta K_j, \qquad K_j^{\text{nonresource}} = (1-\theta)K_j, \qquad j \in \{\ell, h\}. \tag{11.1}$$

Varying θ affects only surplus accounting; prices, quantities, lender margins, and estimated rerouting remain fixed.

The distinction is important for surplus. Before privacy, a targeted lender reaches the borrower before she searches and monetizes her stopping wedge through the mortgage rate. No search resources are then spent. Privacy removes some of these targeted offers. The non-resource part of the lost markup is simply transferred from lenders to borrowers through lower rates. The resource part is different: borrowers who are rerouted to organic comparison must now actually spend time and effort searching, so this part of the wedge is no longer capitalized into the mortgage rate and is instead dissipated as a real resource cost.

Moreover, the targeted offer is pooled at the low-K reservation rate, including for targeted high-K borrowers. In the quantitative implementation, privacy can reroute some of these high-K borrowers into organic comparison, where they incur the larger search cost K_h. Thus the real search cost generated after privacy can exceed the resource component of the acquisition markup that was previously captured by lenders.

Figure J.4 reports borrower surplus and identified intensive-margin surplus over $\theta \in [0, 1]$. The model-implied borrower break-even point is $\theta = 0.50$: borrowers gain whenever at least half of the comparison wedge is non-resource. Identified intensive-margin surplus is negative throughout the full range of θ. For orientation, Agarwal et al. (2024)'s mortgage search-cost estimate maps into

$\theta = 0.238$, or a non-resource share of 76.2%. This is an external benchmark for surplus accounting, not an identifying restriction of the model.

At this benchmark, the gross acquisition-related rate gain is 14.25 basis points. Of this amount, 10.86 basis points reflect non-resource rents returned from lenders to borrowers, while 3.39 basis points reflect the resource-search component that had previously been monetized in the targeted rate. Privacy, however, induces 5.27 basis points of actual search costs as displaced borrowers enter comparison, including some high-K borrowers. The net resource-search effect on borrower surplus is therefore $3.39 - 5.27 = -1.88$ basis points. Together with the -3.16 basis-point effect of weaker screening,

$$\Delta CS = \underbrace{-3.16}_{\text{weaker screening}} + \underbrace{10.86}_{\substack{\text{non-resource rents} \\ \text{returned to borrowers}}} + \underbrace{(3.39 - 5.27)}_{\substack{\text{avoided search cost returned} \\ \text{minus actual search cost}}} = 5.82 \text{ bp.} \tag{11.2}$$

Appendix Figure J.3 reports the corresponding decomposition across credit-score bins.

Current-loan operating margins change by +1.57 basis points for No Tech, −34.89 for Only Ads, and +8.97 for Ads+Data, for an aggregate change of −24.35 basis points. Rate transfers between borrowers and lenders cancel in the efficiency calculation. What remains are changes in real lending and search costs:

$$\Delta W^{\text{int}} = \underbrace{-18.60}_{\substack{\text{weaker screening} \\ \text{higher lending costs}}} + \underbrace{5.33}_{\substack{\text{lender reallocation} \\ \text{real-cost effect}}} - \underbrace{5.27}_{\substack{\text{search costs} \\ \text{actually dissipated}}} = -18.54 \text{ bp.} \tag{11.3}$$

Thus privacy returns non-resource acquisition rents to borrowers, but converts part of the search wedge that lenders previously monetized into real search effort. The former is a transfer; the latter is a deadweight loss.[6]

Finally, Table J.4 considers a targeting-only restriction that preserves pre-privacy underwriting information but prohibits its use for paid cross-lender acquisition. At $\theta = 0.238$, this policy raises borrower surplus by 16.02 basis points, compared with 5.82 under broad privacy, and reduces the identified intensive-margin efficiency loss from 18.54 to 8.90 basis points. It still causes some previously monetized search costs to be dissipated through active comparison, but avoids the additional loss from weaker screening. Separating data uses therefore preserves more of the borrower gain while substantially reducing the efficiency cost of broad privacy regulation.

[6]The surplus calculation is conditional on origination and does not monetize the value of changes in credit access, continuation or franchise value, advertising or platform resource costs, entry, or long-run technology adoption.

12 Conclusion

This paper shows that online-collected consumer data shape mortgage competition through two uses with opposing effects on prices, and that for likely shoppers the advertising use dominates. The same information that helps lenders assess repayment risk also helps them reach borrowers before they compare offers. Exploiting the staggered enforcement of state privacy laws, I find evidence for both channels. Enforcement weakens screening: rejections decline and one-year bankruptcy among originated borrowers rises. It also weakens search advertising: advertising lenders lose likely shoppers, and paid search clicks fall among lenders that rely on third-party advertising platforms. On net, mortgage rates fall by about 14 basis points for likely shoppers, with no detectable effect for borrowers unlikely to shop. Privacy can therefore lower mortgage rates even as the originated pool becomes riskier, consistent with a decline in acquisition rents rather than in lending costs.

The model explains this result by making borrower acquisition part of mortgage competition. Search advertising reaches borrowers before they compare offers, letting lenders charge for part of the cost of comparison through the mortgage rate; the competitive effect of advertising thus depends on when it reaches borrowers. Privacy weakens this channel and reroutes borrowers toward active comparison and their existing lenders. In the estimated model, the resulting rate decline is several times larger than the rate increase from weaker screening. Aggregating across all lenders and borrowers, the model also implies that privacy reduces credit access, with approval falling by about 0.48 percentage points, even as rejections decline within advertising lenders.

This mechanism shapes incidence and surplus. Part of the cost of comparison reflects real time and effort, and part reflects frictions such as inattention or inertia. Privacy returns the second part to borrowers through lower rates, but borrowers now spend the first on active comparison, and weaker screening raises lending costs. Using an external estimate of the share of search frictions that reflects real resources (Agarwal et al. 2024), borrowers gain about 5.8 basis points of mortgage balance, while total surplus on originated loans falls by about 18.5 basis points, a loss that persists across the full range of assumptions about that share. The calculation is partial equilibrium and conditional on origination, so it does not value changes in credit access or borrowers' direct preferences for privacy.

More broadly, consumer data are both a productive input into financial intermediation and a strategic input into customer acquisition, and broad privacy rules restrict both at once. A use-based counterfactual shows the value of separating them: preserving underwriting information while prohibiting its use for cross-lender advertising raises the borrower gain from 5.8 to about 16.0 basis points and reduces the surplus loss from 18.5 to 8.9 basis points. The relevant policy question is therefore not how much information financial intermediaries possess, but how they are permitted to use it. As mortgage search moves online, data-use rules become competition rules.

References

Agarwal, S., J. C. Driscoll, X. Gabaix, and D. Laibson (2007). The age of reason: Financial decisions over the lifecycle.

Agarwal, S., P. Ghosh, P. Jin, S. Kundu, N. Vats, X. Wang, and Y. Xu (2024). When privacy protects but excludes: The costs and benefits of privacy regulation in credit markets. *Olin Business School Center for Finance & Accounting Research Paper* (2026/01).

Agarwal, S., J. Grigsby, A. Hortaçsu, G. Matvos, A. Seru, and V. Yao (2024). Searching for approval. *Econometrica 92*(4), 1195–1231.

Allen, J., R. Clark, and J.-F. Houde (2019). Search frictions and market power in negotiated-price markets. *Journal of Political Economy 127*(4), 1550–1598.

Argyle, B., T. Nadauld, and C. Palmer (2023). Real effects of search frictions in consumer credit markets. *The Review of Financial Studies 36*(7), 2685–2720.

Aridor, G., Y.-K. Che, and T. Salz (2023). The effect of privacy regulation on the data industry: Empirical evidence from gdpr. *The RAND Journal of Economics 54*(4), 695–730.

Babina, T., S. Bahaj, G. Buchak, F. De Marco, A. Foulis, W. Gornall, F. Mazzola, and T. Yu (2025). Customer data access and fintech entry: Early evidence from open banking. *Journal of Financial Economics 169*, 103950.

Bartlett, R., A. Morse, R. Stanton, and N. Wallace (2022). Consumer-lending discrimination in the fintech era. *Journal of Financial Economics 143*(1), 30–56.

Benetton, M., B. Hébert, and T. McQuade (2025). Deposit competition beyond rates.

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), 2845–2897.

Berg, T., A. Fuster, and M. Puri (2022). Fintech lending. *Annual Review of Financial Economics 14*(1), 187–207.

Bergemann, D., A. Bonatti, and N. Wu (2025). How do digital advertising auctions impact product prices? *Review of Economic Studies 92*(4), 2330–2358.

Bhutta, N., A. Fuster, and A. Hizmo (2026). Paying too much? borrower sophistication and overpayment in the us mortgage market. *The Journal of Finance 81*(1), 49–90.

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), 453–483.

Burguet, R. and V. Petrikaitė (2023). Targeted advertising and costly consumer search. *Economic Inquiry 61*(2), 430–450.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-differences with multiple time periods. *Journal of econometrics 225*(2), 200–230.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *The review of economics and statistics 90*(3), 414–427.

Chen, W., J. K. Kang, S. Kim, and L. Lin (2026). Customer-facing technologies and banks' macroeconomic information production. *Harvard Business School Accounting & Management Unit Working Paper* (25-057).

Chen, W., J. K. Kang, and A. Mohan (2023). Data-driven technologies and local information advantages in small business lending.

Chen, X., A. Hu, and S. Ma (2025). Banks' images: Evidence from advertising videos. Technical report, National Bureau of Economic Research.

Chioda, L., P. J. Gertler, S. Higgins, and P. Medina (2024). Fintech lending to borrowers with no credit history. *NBER Working Paper* (w33208).

Coen, J., A. K. Kashyap, and M. Rostom (2023). Price discrimination and mortgage choice. Technical report, National Bureau of Economic Research.

Dai, W., H. Kim, and M. Luca (2023). Frontiers: which firms gain from digital advertising? evidence from a field experiment. *Marketing science 42*(3), 429–439.

D'Andrea, A., P. Hitayezu, R. K. Kpodar, N. Limodio, and A. Presbitero (2025). Developing the mortgage market: Technology, property rights, and banking. Technical report, Centre for Economic Policy Research.

De Corniere, A. (2016). Search advertising. *American Economic Journal: Microeconomics 8*(3), 156–188.

De Corniere, A. and R. De Nijs (2016). Online advertising and privacy. *The RAND Journal of Economics 47*(1), 48–72.

Deisenroth, D., U. Manjeer, Z. Sohail, S. Tadelis, and N. Wernerfelt (2024). Digital advertising and market structure: Implications for privacy regulation. Technical report, National Bureau of Economic Research.

Di Maggio, M., D. Ratnadiwakara, and D. Carmichael (2022). Invisible primes: Fintech lending with alternative data. Technical report, National Bureau of Economic Research.

Diamond, P. A. (1971). A model of price adjustment. *Journal of economic theory 3*(2), 156–168.

Doerr, S., L. Gambacorta, L. Guiso, and M. Sanchez del Villar (2026). Privacy regulation and fintech lending. *Management Science*.

Dubé, J.-P., J. G. Lynch, D. Bergemann, M. Demirer, A. Goldfarb, G. Johnson, A. Lambrecht, T. Lin, A. Tuchman, and C. Tucker (2025). Frontiers: The intended and unintended consequences of privacy regulation for consumer marketing. *Marketing Science 44*(5), 975–984.

Dubé, J.-P. H., I. Morozov, F. She, and A. Tuchman (2026). The effects of ads on beliefs and implications for consumer search. Technical report, National Bureau of Economic Research.

Eeckhout, J. and L. Veldkamp (2023). Data and markups: A macro-finance perspective. *NBER Working Paper* (w30022).

Fershtman, D. and A. Pavan (2025). Non-sequential-non-cascading clicking in online consumer search.

Frésard, L. and M. Zanotti (2025). Are new technologies replacing the information produced by financial markets?

Fuster, A., M. Plosser, P. Schnabl, and J. Vickery (2019). The role of technology in mortgage lending. *The review of financial studies 32*(5), 1854–1899.

Goeree, M. S. (2008). Limited information and advertising in the us personal computer industry. *Econometrica 76*(5), 1017–1074.

Goldfarb, A. and C. E. Tucker (2011). Privacy regulation and online advertising. *Management science 57*(1), 57–71.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of econometrics 225*(2), 254–277.

Gourio, F. and L. Rudanko (2014). Customer capital. *Review of Economic Studies 81*(3), 1102–1136.

Günsberg, A. and C. M. Kuhnen (2026). Learning how to borrow in a fintech world: Consumer behavior when search costs are (near) zero. Technical report, National Bureau of Economic Research.

Gupta, M., D. McGowan, and S. Ongena (2024). The cost of data privacy law: Evidence from the california mortgage market. *Swiss Finance Institute Research Paper* (23-25).

Gurun, U. G., G. Matvos, and A. Seru (2016). Advertising expensive mortgages. *The Journal of Finance 71*(5), 2371–2416.

Haan, M., P. Heijnen, and J. Knol (2024). Search-inducing informative advertising.

He, B., L. Mostrom, and A. Sufi (2026). Fintech and customer capital. *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2026-14).

He, B., L. I. Mostrom, and A. Sufi (2024). Investing in customer capital. Technical report, National Bureau of Economic Research.

He, Z., J. Huang, and C. Parlatore (2024). Information span in credit market competition. Technical report, National Bureau of Economic Research.

He, Z., J. Huang, and J. Zhou (2023). Open banking: Credit market competition when borrowers own the data. *Journal of financial economics 147*(2), 449–474.

Honka, E., A. Hortaçsu, and M. A. Vitorino (2017). Advertising, consumer awareness, and choice: Evidence from the us banking industry. *The RAND Journal of Economics 48*(3), 611–646.

Hortaçsu, A. and C. Syverson (2004). Product differentiation, search costs, and competition in the mutual fund industry: A case study of s&p 500 index funds. *The Quarterly journal of economics 119*(2), 403–456.

Huang, K., M. Tang, R. Wittenberg-Moerman, J. S. Yoon, and T. Zhang (2026). Scroll to the top: Visual cues and primacy bias in consumer lending. Working paper.

Jiang, E. X., Y. Lee, and Q. Maingi (2026). Credit without proximity: Informational frictions and unequal gains from technology. *Available at SSRN 6713098*.

Lab, S. C. (2026). The innovation cost of privacy: how state data privacy laws reshape the technology sector? *Working Paper*.

MacKinnon, J. G. and M. D. Webb (2018). The wild bootstrap for few (treated) clusters. *The Econometrics Journal 21*(2), 114–135.

Mendes, B. (2024). Bank advertising and deposit demand. *Kilts Center at Chicago Booth Marketing Data Center Paper (forthcoming)*.

Rambachan, A. and J. Roth (2023). A more credible approach to parallel trends. *Review of Economic Studies 90*(5), 2555–2591.

Roussanov, N., H. Ruan, and Y. Wei (2021). Marketing mutual funds. *The Review of Financial Studies 34*(6), 3045–3094.

Stroube, B. and G. Dushnitsky (2023). Built with builtwith: A technical note. *Available at SSRN 4475309*.

Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of econometrics 225*(2), 175–199.

Tang, H. (2019). The value of privacy: Evidence from online borrowers. *Available at SSRN 3880119*.

Ursu, R. M., Q. Zhang, and E. Honka (2023). Search gaps and consumer fatigue. *Marketing Science 42*(1), 110–136.

Wernerfelt, N., A. Tuchman, B. T. Shapiro, and R. Moakler (2025). Estimating the value of offsite tracking data to advertisers: Evidence from meta. *Marketing Science 44*(2), 268–286.

Woodward, S. E. and R. E. Hall (2012). Diagnosing consumer confusion and sub-optimal shopping effort: Theory and mortgage-market evidence. *American Economic Review 102*(7), 3249–3276.

Wooldridge, J. M. (2025). Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators. *Empirical Economics 69*(5), 2545–2587.

Zanotti, M. (2025). Asset (and data) managers. *Available at SSRN 5734507*.

Zhang, D., P. J. Barwick, L. Han, and J. Kroah (2026). Market power in mortgage pricing: the role of referral lending. Technical report, National Bureau of Economic Research.

A Appendix – Online-Data Technologies and Adoption

This appendix provides additional detail on the online-data technologies discussed in Section 3, explaining in practical terms what they do and why they matter for mortgage lending, and documenting the lender characteristics associated with adoption.

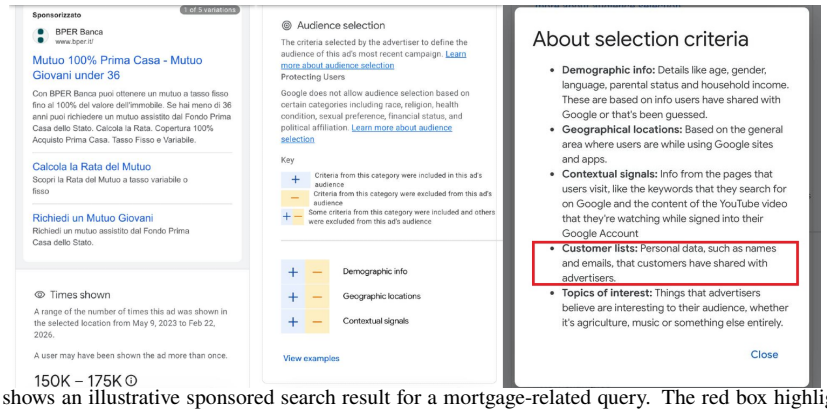
A.1 Search Advertising and Borrower Acquisition

Borrowers often begin the mortgage process before contacting a lender—searching for rates, refinancing options, calculators, or a specific lender's name. Search advertising lets lenders appear at this point: when a user enters a mortgage-related query, the search engine can display sponsored results above or alongside the organic ones, shaping which lenders the borrower considers.

This matters because mortgage shopping is limited. When a borrower considers only one or two lenders, appearing early is valuable, so search advertising can affect competition before any application is submitted. It also lets a lender enter the consideration set, measure which search terms generate visits, and learn which users respond to its ads.

Figure A.1 illustrates search-ad placement for mortgage queries. It is not used as an outcome in the empirical analysis. It clarifies the institutional setting, in which online advertising is not merely a way to display ads but part of a data infrastructure connecting borrower search, website visits, and information production.

Figure A.1. Illustrative Search Advertisement and Lender Website Link



Notes: The figure shows an illustrative sponsored search result for a mortgage-related query. The red box highlights the platform's disclosure that data collected through the lender's website were used in targeting the advertisement. This information is available under Article 39 of the EU Digital Services Act, which requires designated online platforms and search engines to disclose key advertising and targeting information. The example features an Italian lender because the disclosure requirement applies in the European Union. It is included solely to illustrate the advertising infrastructure studied in the U.S. analysis and is not used in the empirical results. *Source:* Google advertising-transparency disclosure pursuant to Article 39 of Regulation (EU) 2022/2065.

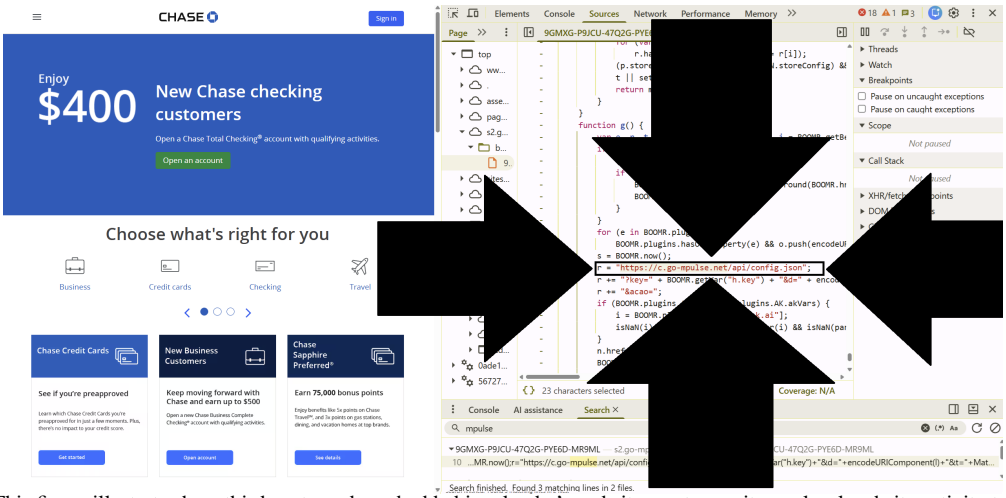
A.2 Website Data Collection and Transmission

After a borrower clicks on a search result, advertisement, or referral link, the lender can collect information about the visit. This collection usually occurs through code embedded in the lender's website. The code can include cookies, tracking pixels, tags, software development kits, and related scripts. These tools are common across the modern internet. They allow websites and third-party platforms to record how users arrive at a page, what they view, and how they interact with the site. Deisenroth et al. (2024) provides a more detailed technical illustration on the use of cookies and pixel by Meta.

For example, a lender may observe that a user arrived from a paid search advertisement, visited a mortgage-rate page, interacted with a calculator, began an application, and returned to the same page several days later. The lender may also observe technical information such as device type, browser, referral source, approximate location, and time of visit. These data can be used to measure advertising performance, retarget users, improve website design, and produce information about borrower demand.

Figure A.2 illustrates this process. The important point is not the specific technology used, but the flow of information. A user visits a lender's website. Code embedded in the website records the visit and transmits information to the lender or to third-party platforms. These platforms can then combine the visit with information from advertising campaigns, other websites, or prior user behavior.

Figure A.2. Website Tracking and Data Transmission



Notes: This figure illustrates how third-party code embedded in a lender's website can transmit user-level website activity to outside platforms. Cookies, tracking pixels, tags, software development kits, and related scripts record information on page visits, referral sources, device characteristics, browser information, and user interactions. These data can then be combined with information from advertising platforms, third-party websites, and customer data platforms. The figure is illustrative and is used to describe the data infrastructure underlying online advertising and website analytics.

A.3 Advertising Technologies and Customer Data Platforms

The paper distinguishes between two levels of online-data use. The first is advertising technologies. These technologies help lenders target potential borrowers and measure the performance of advertising campaigns. They can tell a lender which ads were shown, which users clicked, which search terms generated traffic, and which users later visited the lender's website. They are usually provided by third-party platforms and can be adopted without building a large internal data system.

The second is customer data platforms. These platforms allow lenders to store and combine first-party data from their own websites. First-party data are data generated inside the lender's own digital environment. Examples include visits to mortgage pages, calculator use, form starts, repeated visits, referral sources, and application behavior. A customer data platform can combine these first-party data with information from advertising platforms and third-party websites.

Both types of data can be informative, but they differ in depth. Advertising-platform data can reveal how users arrive at a lender's website, including traffic sources, campaigns, keywords, search terms associated with advertising campaigns, location, audience characteristics, and campaign responses. These signals can help lenders target potential borrowers and learn about borrower demand. They can also be relevant for screening because search terms, acquisition channels, location, repeated visits, and product interest help distinguish borrowers by demand, attention, and likely fit for a product. First-party website data provide a richer version of the same logic because they are generated inside the lender's own product environment. A borrower who repeatedly visits a mortgage page, starts an application, uses a calculator, or returns through a specific referral source reveals lender-specific information about product interest and search intensity. This is the sense in which customer data platforms represent a deeper form of information production than advertising technologies alone.

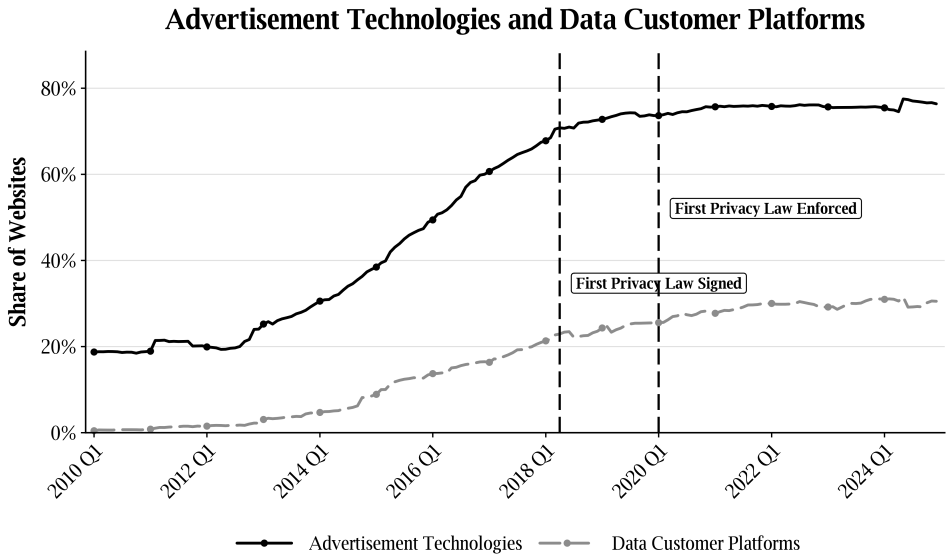
This distinction is consistent with recent evidence from asset management. Zanotti (2025) shows that financial firms use website technologies such as Google Analytics and A/B testing tools to collect information on potential customers, including browsing behavior, demographics, product interest, acquisition channels, search terms, and audience overlap with competitors. The setting is asset management rather than mortgage lending, but the institutional lesson is directly relevant: digital customer data can help financial firms learn about customer demand and target customers more effectively. In my setting, the same logic applies to mortgage lenders, with first-party website data providing especially lender-specific information about borrower demand and search behavior.

The empirical analysis does not require observing the exact internal use of each signal. The relevant institutional fact is that these technologies change the information available to lenders before and during the application process. Privacy-law enforcement can therefore affect credit markets by changing the data lenders can collect, combine, or use.

A.4 Adoption of Online-Data Technologies

The adoption of online-data technologies rises sharply after 2013, coinciding with the diffusion of smartphones and the broader growth of online consumer data. Advertising technologies diffuse faster because they are commonly provided as third-party services. Customer data platforms diffuse more slowly because full deployment requires greater investment in data infrastructure and technical expertise. Figure A.3 shows these adoption patterns over time.

Figure A.3. Adoption of Online-Data Technologies Over Time



Notes: This figure plots the evolution of lenders' adoption of online-data technologies. Advertising technologies are third-party tools used for digital advertising, targeting, and campaign measurement. Customer data platforms allow lenders to store and combine first-party data generated on their own websites with information from advertising platforms and other digital sources.

Source: Author's calculations based on website source-code inspections of mortgage lenders' websites.

To study adoption correlates, I estimate Cox proportional hazards models using HMDA originations data and NMLS establishment and loan-officer data. I do not include balance sheet variables because they are unavailable for many fintech lenders. Table A.1 reports adoption correlates for advertising technologies, while Table A.2 reports adoption correlates for customer data platforms.

The results show that larger lenders and credit unions have higher adoption hazards. For advertising technologies, credit-union supervision is associated with a hazard ratio of 1.303, meaning a 30.3% higher instantaneous adoption hazard, conditional on the other covariates. A one-unit increase in log applications is associated with a hazard ratio of 1.298, or a 29.8% higher instantaneous adoption hazard. Website characteristics also predict adoption. Google Analytics is associated with a hazard ratio of 1.523, or a 52.3% higher adoption hazard, while dynamic websites, online support, social-media presence, and mobile optimization are also associated with higher adoption hazards.

Customer data platforms display a similar but stronger size gradient. A one-unit increase in log applications is associated with a hazard ratio of 1.469, or a 46.9% higher instantaneous adoption hazard. Credit-union supervision is associated with a hazard ratio of 1.299, or a 29.9% higher adoption hazard. Website characteristics are also important. Google Analytics is associated with a hazard ratio of 1.570, or a 57.0% higher adoption hazard, while dynamic websites are associated with a hazard ratio of 1.505, or a 50.5% higher adoption hazard. These patterns are consistent with customer data platforms requiring a broader digital infrastructure before adoption.

These results serve two purposes. First, they document that online-data technologies were economically meaningful before privacy-law enforcement. Second, they help classify lenders by their pre-period use of online-collected data. They are not the source of identification in the paper. The empirical design conditions on lenders that had already adopted these technologies and uses privacy-law enforcement to vary the data these lenders can use across otherwise similar borrowers within the same lender.

Table A.1. Adoption of Advertising Technologies

Variable	Hazard ratio	(Std. error)	% change in hazard
Panel A: Lender characteristics			
Supervised by National Credit Union Administration	1.303***	(0.025)	+30.3%
Applications (log)	1.298***	(0.029)	+29.8%
Employees per Remote Work Locations	1.072***	(0.022)	+7.2%
Origination rate	1.043	(0.029)	+4.3%
Mean tenure (Headquarter)	0.929**	(0.028)	−7.1%
Headquarter/(Branch+Remote Work+Branches)	0.926**	(0.029)	−7.4%
Panel B: Website characteristics			
Google Analytics	1.523***	(0.029)	+52.3%
Dynamic website	1.182***	(0.016)	+18.2%
Online support	1.140***	(0.020)	+14.0%
Social media presence	1.091***	(0.020)	+9.1%
Mobile optimized	1.066***	(0.016)	+6.6%

Notes: This table reports Cox proportional hazards estimates for adoption of advertising technologies. Hazard ratios above one indicate a higher instantaneous adoption hazard, conditional on the other covariates; hazard ratios below one indicate a lower adoption hazard. The percentage-change column reports (hazard ratio − 1) × 100. Standard errors are in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Table A.2. Adoption of Customer Data Platforms

Variable	Hazard ratio	(Std. error)	% change in hazard
Panel A: Lender characteristics			
Applications (log)	1.469***	(0.037)	+46.9%
Supervised by National Credit Union Administration	1.299***	(0.029)	+29.9%
Number of Remote Work Locations (log)	1.146***	(0.026)	+14.6%
Employees per Remote Work Locations	1.141***	(0.023)	+14.1%
Origination rate	1.073*	(0.038)	+7.3%
Panel B: Website characteristics			
Google Analytics	1.570***	(0.049)	+57.0%
Dynamic website	1.505***	(0.027)	+50.5%
Social media presence	1.268***	(0.030)	+26.8%
Twitter widgets	1.040**	(0.018)	+4.0%
Multiple languages	0.981	(0.023)	−1.9%

Notes: This table reports Cox proportional hazards estimates for adoption of customer data platforms. Hazard ratios above one indicate a higher instantaneous adoption hazard, conditional on the other covariates; hazard ratios below one indicate a lower adoption hazard. The percentage-change column reports (hazard ratio − 1) × 100. Standard errors are in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

B Appendix – State Privacy Laws

This appendix reports the state privacy laws used in the empirical analysis. Table B.1 records each law's enactment date, enforcement date, and key provisions used to measure legal intensity. These provisions capture whether the law restricts data sales, data sharing, sensitive-data collection, browser-level opt-out signals, opt-in consent for sensitive data, cure periods, and private rights of action.

The table shows that state privacy laws vary along two dimensions. The first is timing: enactment dates and enforcement dates differ both within and across states. I use enforcement dates in the empirical analysis because enforcement determines when firms face legal exposure for noncompliance. The second is intensity: laws differ in whether they restrict sales, sharing, sensitive-data collection, opt-in consent, global opt-out signals, cure periods, and private rights of action. These differences allow me to distinguish the existence of a privacy law from the strength of the restrictions it imposes.

B.1 Applicability to Financial Intermediaries

State privacy laws can apply to financial intermediaries when the data at issue fall outside traditional customer financial records. In my setting, the relevant data are generated through online advertising, website visits, attribution tools, and pre-application interactions. These data can include advertising identifiers, referral sources, device and location signals, campaign responses, page views, calculator

Table B.1. Timeline and Provisions of U.S. State Privacy Laws

		Date		Consumer protections						Cure period		
State	Law	Enacted	Effective	Sale opt-out	Sharing opt-out	Sensitive limit	Global opt-out	Sensitive opt-in	Promoted by	Days	Expiration	Private action
California	CCPA/CPRA	28-Jun-18 / 3-Nov-20	1-Jan-20 / 1-Jan-23	✓	✓	✓	✓	✓	Consumers	0	1-Jan-23	✓
Virginia	VCDPA	2-Mar-21	1-Jan-23	✓	✓	✓		✓	Firms	30	—	
Colorado	CPA	8-Jul-21	1-Jul-23	✓	✓	✓	✓	✓	Consumers	0	1-Jan-25	
Connecticut	CTDPA	10-May-22	1-Jul-23	✓	✓	✓	✓	✓	Consumers	0	31-Dec-24	
Utah	UCPA	24-Mar-22	31-Dec-23	✓	✓				Firms	30	—	
Florida	FDBR	6-Jun-23	1-Jul-24	✓	✓	✓		✓	Firms	45	—	
Oregon	OCPA	18-Jul-23	1-Jul-24	✓	✓	✓	✓	✓	Consumers	0	1-Jan-26	
Texas	TDPSA	18-Jun-23	1-Jul-24	✓	✓	✓	✓	✓	Firms	30	—	
Montana	MCDPA	19-May-23	1-Oct-24	✓	✓	✓	✓	✓	Consumers	0	1-Oct-25	
Delaware	DPDPA	11-Sep-23	1-Jan-25	✓	✓	✓	✓	✓	Consumers	0	31-Dec-25	
Iowa	ICDPA	28-Mar-23	1-Jan-25	✓					Firms	90	—	
Nebraska	NEDPA	17-Apr-24	1-Jan-25	✓	✓	✓	✓	✓	Firms	30	—	
New Hampshire	NHPA	6-Mar-24	1-Jan-25	✓	✓	✓	✓	✓	Consumers	0	31-Dec-25	
New Jersey	NJDPA	16-Jan-24	15-Jan-25	✓	✓	✓	✓	✓	Consumers	30	1-Jul-26	
Tennessee	TIPA	11-May-23	1-Jul-25	✓	✓	✓		✓	Firms	60	—	
Minnesota	MNCDPA	24-May-24	31-Jul-25	✓	✓	✓	✓	✓	Consumers	0	31-Jan-26	
Maryland	MODPA	9-May-24	1-Oct-25	✓	✓	✓	✓	✓	Consumers	60	1-Apr-27	
Indiana	INCDPA	1-May-23	1-Jan-26	✓	✓	✓		✓	Firms	30	—	
Rhode Island	RIDTPPA	28-Jun-24	1-Jan-26	✓	✓	✓		✓	Firms	0	N/A	
Kentucky	KCDPA	4-Apr-24	1-Jan-26	✓	✓	✓		✓	Firms	30	—	

Notes: This table reports the state privacy laws used in the empirical analysis. *Enacted* is the date the law was signed or approved; *Effective* is the date it became enforceable. *Sale opt-out* and *Sharing opt-out* indicate whether consumers can restrict the sale or sharing of personal data. *Sensitive limit* and *Sensitive opt-in* indicate whether the law imposes additional restrictions on sensitive personal information (e.g., geolocation, biometric identifiers, government IDs, health information, gender, race, religion, or union membership). *Global opt-out* indicates whether the law requires firms to honor browser- or device-level opt-out signals. *Promoted by* records whether the law was primarily advanced by consumer advocates or by firms. *Cure period* reports the statutory number of days firms have to remediate violations before enforcement and, where applicable, the date the cure provision expires ("—" denotes a permanent cure period; "N/A" denotes no cure period). *Private action* indicates whether private consumers may sue for covered violations. A check mark (✓) denotes that the provision is present; a blank cell denotes that it is absent.

use, application starts, and repeated visits. They are distinct from the standard mortgage-file variables regulated under federal financial-privacy rules.

This distinction matters for the exposure measure. The lenders classified as adopters are those that had already installed advertising technologies or customer data platforms before privacy-law enforcement. These tools are precisely the infrastructure through which lenders collect or combine online data outside the traditional loan file. The exposure classification therefore captures pre-period reliance on online-data technologies, not whether a lender is generally subject to financial regulation.

State privacy laws also contain applicability thresholds. These thresholds vary across states and often depend on revenue, the number of consumers whose data a firm processes, or the share of revenue derived from selling or sharing personal data. In the mortgage-lender sample, lenders that had adopted advertising technologies or customer data platforms before enforcement are generally large enough that statutory thresholds are unlikely to determine treatment status. The identifying variation comes from pre-period reliance on online-data technologies, not from marginal crossing of a statutory size threshold.

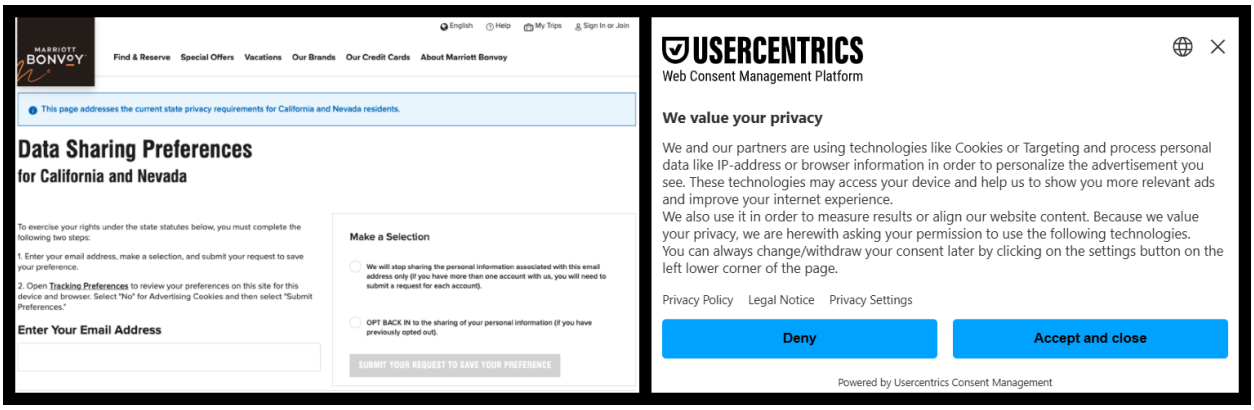
Recent enforcement materials and private litigation make this boundary visible by focusing on website tracking, advertising-platform transfers, opt-out mechanisms, and global opt-out signals. I do not use these cases as evidence of liability. They are useful only to illustrate that website and advertising data are legally distinct from the traditional customer financial records covered by federal financial-privacy rules.

B.2 Cookie Banners and Formal Privacy Laws

Before comprehensive state privacy laws, many websites relied on cookie banners and platform-specific consent tools. These mechanisms are different from formal privacy-law enforcement. Cookie banners often require users to navigate website-specific forms, make repeated choices across websites, and understand how data are collected or shared. Formal privacy laws can instead create enforceable rights to access, delete, correct, or port data, and in some states require firms to honor browser-level opt-out signals.

Figure B.1 illustrates the difference between website-level consent interfaces and the broader legal regime studied in the paper. The figure is not used as an outcome in the empirical analysis. It is included to clarify the institutional setting: the empirical shock comes from enforceable state privacy laws, not from voluntary self-regulation by websites.

Figure B.1. Cookie Banner and Website-Level Consent



Notes: This figure provides an illustrative example of a website-level cookie banner. Cookie banners allow users to accept, reject, or customize some forms of website tracking, but they typically operate at the website level and depend on user action within each website. State privacy laws can impose broader legal obligations, including access, deletion, correction, portability, opt-out rights, sensitive-data restrictions, and, in some states, browser-level global opt-out signals. The figure is illustrative and is not used as an outcome in the empirical analysis.

B.3 Privacy-Law Voting Patterns and Mortgage-Market Dynamics

I construct a tract-level dataset linking support for California Proposition 24 to local mortgage-market conditions, borrower search propensity, political composition, and resident industry exposure. I begin with precinct-level Proposition 24 returns and voter-registration data from the California Statewide Database. Because voting precincts do not map one-to-one into Census tracts, I use the Statewide Database precinct-to-block geographic crosswalks to allocate Proposition 24 yes votes, no votes, total votes, turnout, and party registration from precincts to Census blocks, and then aggregate the weighted block-level values to Census tracts. I then merge these tract-level election outcomes with HMDA mortgage-market variables measured before the vote, including mortgage originations, denial rates, preapproval requests, loan amounts, rate spreads, loan-to-value ratios, lender composition, and the tract-level mean of the borrower search-for-price propensity measure. I further add ACS demographic, housing, income, education, broadband, and resident-industry controls; LODES workplace employment measures; and FHFA tract-level house-price growth from 2017 to 2019. The final regressions are weighted by Proposition 24 votes, include county fixed effects, and cluster standard errors by county.

Table B.2 uses this dataset to test whether support for Proposition 24 reflects local mortgage-market conditions or borrowers' propensity to search for mortgage prices. This test is important because a potential concern is that privacy-law adoption may be correlated with pre-existing mortgage-market behavior: places where borrowers search more, face different credit conditions, or interact with different types of lenders might also be more likely to support privacy regulation. The evidence does not support this interpretation. The coefficient on mortgage search propensity is close

to zero, economically small, and statistically insignificant both in the full-controls specification and in the specification that excludes the income, loan amount, age, and race variables used to construct the search measure. Thus, the absence of a relationship is not mechanically driven by over-controlling for the inputs of the search measure. More broadly, mortgage-market dynamics do not provide a coherent explanation for the vote. Origination volume, denial rates, rate spreads, loan-to-value ratios, lender concentration, and ad-tech lender shares are either weakly related to support or fail to explain away the main pattern.

Instead, support for Proposition 24 is explained primarily by political composition and local sectoral exposure. Democratic, unaffiliated, and other-party registration shares strongly predict higher support relative to Republican registration, indicating that the vote has a first-order political component. Conditional on these political controls, the most robust nonpolitical predictor is the resident real-estate-sector share. A 10 percentage point increase in the share of residents employed in real estate is associated with a 1.4–1.7 percentage point higher yes-vote share. This pattern is consistent with the institutional origin of the initiative: Proposition 24 was promoted by Alastair Mactaggart, a real-estate developer. By contrast, support is lower in tracts with larger professional, scientific, and technical resident employment shares, and in the full-controls specification it is also lower in more educated, more homeowner-intensive, and more prime-age areas. These patterns suggest that support for privacy regulation is not driven by mortgage borrowers' search behavior or by local mortgage-market conditions. Rather, the vote reflects broader political and sectoral interests.

Table B.2. Prop. 24 Support And Mortgage Market Dynamics

	Dependent variable: Prop. 24 yes-vote share	
	Full controls (1)	Excl. `pi_price` inputs (2)
Mortgage search propensity	-0.0001	0.0023
	(0.0024)	(0.0029)
Mortgage-market controls		
Log mortgage originations	0.0053*	0.0040†
	(0.0021)	(0.0023)
Denial rate	0.0277*	0.0409*
	(0.0113)	(0.0171)
Preapproval request rate	0.0410	0.0319
	(0.0758)	(0.0813)
Median loan amount	-0.0000	–
	(0.0000)	
Median rate spread	0.0081†	0.0079
	(0.0044)	(0.0052)
Median LTV	0.0000	0.0000
	(0.0003)	(0.0004)
Share ad-tech lenders	-0.0080	-0.0129
	(0.0075)	(0.0093)
Share ad-tech/data lenders	-0.0046	-0.0012
	(0.0071)	(0.0104)
Number of lenders	0.0003	0.0005*
	(0.0002)	(0.0002)
FHFA log HPI growth, 2017–2019	0.0329*	0.0466**
	(0.0164)	(0.0148)
Political registration controls		
Democratic registration share	0.4603***	0.4949***
	(0.0350)	(0.0418)
Unaffiliated registration share	0.9355***	1.0170***
	(0.1237)	(0.1312)
Other-party registration share	0.9610***	0.4269**
	(0.1730)	(0.1466)
Socioeconomic and demographic controls		
Log owner-occupied units	-0.0047†	-0.0022
	(0.0028)	(0.0036)
Log median household income	0.0101	–
	(0.0084)	
Bachelor's degree share	-0.0734***	-0.0692**
	(0.0197)	(0.0214)
Homeownership rate	-0.0268*	0.0024
	(0.0104)	(0.0072)
Black non-Hispanic share	0.1727***	–
	(0.0213)	
Asian non-Hispanic share	0.0810***	–
	(0.0159)	
Hispanic share	0.0268**	–
	(0.0099)	
Other race share	0.0255	–
	(0.0332)	
Age 25–44 share	-0.0958**	–
	(0.0330)	
Age 65+ share	0.0461†	–
	(0.0232)	
Poverty share	-0.0239*	-0.0047
	(0.0094)	(0.0143)
Unemployment rate	0.0174	0.0799**
	(0.0199)	(0.0285)
Broadband share	0.0411**	0.0105
	(0.0139)	(0.0275)
Resident-sector controls		
Resident information-sector share	0.0637	-0.0556*
	(0.0456)	(0.0258)
Resident professional/scientific/tech share	-0.0592**	-0.0996***
	(0.0192)	(0.0230)
Resident finance/insurance share	0.0544	0.0727*
	(0.0343)	(0.0310)
Resident real-estate share	0.1366***	0.1664**
	(0.0387)	(0.0531)
Resident management-company share	0.1251	0.0508
	(0.2980)	(0.2470)
Resident admin-support share	0.0863**	0.0702*
	(0.0252)	(0.0268)
County fixed effects	Yes	Yes
Observations	3,855	3,855
R^2	0.884	0.856
Within R^2	0.837	0.798

Notes: This table reports tract-level regressions of Proposition 24 support on mortgage search propensity, mortgage-market controls, political registration, socioeconomic characteristics, resident-sector employment shares, and house-price growth. The outcome is the tract-level yes-vote share on Proposition 24. Mortgage search propensity is the standardized tract-level mean of the search-for-price measure estimated from HMDA. Column 1 includes the demographic and loan variables used in constructing the search measure: income, loan amount, race, and age. Column 2 excludes those inputs. Regressions are weighted by Proposition 24 votes and include county fixed effects. Standard errors are clustered by county and reported in parentheses. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, and $^{\dagger}p < 0.10$.

C Appendix – Data

Table C.1. Vivvix Advertiser–Lender Matches by Match Type

Match type	Matches	Share (%)
Website (URL)		
Exact	50	5.4
Fuzzy	4	0.4
Name		
String similarity, high score (automatic)	613	66.7
String similarity, intermediate (reviewed)	131	14.3
Tokenization (reviewed)	46	5.0
Manual		
Manually validated	75	8.2
Total	919	100.0

Notes: This table decomposes the 919 unique Vivvix advertiser-to-lender matches by the step that produced each match, following the tiered procedure described in the sample-construction section. Advertisers are matched first on standardized website domains—exactly, then by fuzzy comparison; the remaining advertisers are matched on standardized names, where high-scoring string-similarity pairs are accepted automatically, intermediate-scoring pairs and tokenization matches are reviewed, and residual cases are validated manually. Shares are percentages of the 919 matches and may not sum to 100 exactly because of rounding. The matched set covers 30% of the HMDA–Equifax lender sample and 41% of the Vivvix mortgage-advertiser sample.

Table C.2. Vintage Composition and the Pooled Mean Note Rate

Origination year	Share of originations (%)	30-Yr FRM PMMS (%)	Contribution to mean (pp)
2018	10.98	4.54	0.498
2019	15.41	3.94	0.607
2020	19.20	3.11	0.597
2021	21.69	2.96	0.642
2022	14.47	5.34	0.773
2023	9.19	6.81	0.626
2024	9.06	6.72	0.609
Origination-weighted mean	100.00		4.352
Memo: observed sample mean			4.172

Notes: Decomposition of the pooled mean note rate into vintage components. Column 1 is each year's share of the 3,722,265 home-purchase originations; Column 2 is the annual average Freddie Mac 30-year fixed rate from the Primary Mortgage Market Survey (PMMS); Column 3 is their product, in percentage points. The origination-weighted mixture of prevailing rates (4.35%) lies just above the observed sample mean note rate (4.17%); the gap is consistent with the sample restrictions (first-lien, conventional-heavy, fully amortizing home-purchase loans, excluding balloon-payment, interest-only, second-lien, and HOEPA loans), which remove higher-cost originations. Shares and the weighted mean are computed from exact origination counts, so displayed products may differ from ratios of the rounded entries.

Source: Freddie Mac PMMS, 30-year fixed-rate series (annual averages of weekly readings); FRED series `MORTGAGE30US` (https://fred.stlouisfed.org/series/MORTGAGE30US).

D Appendix – Additional Summary Statistics

Table D.1. Summary Statistics: Additional Covariates

	N	Mean	SD	P25	Median	P75
Panel A. Categorical borrower and loan controls						
Consumer age						
< 25	129,771	0.035				
25–34	1,135,597	0.305				
35–44	1,018,758	0.274				
45–54	656,291	0.176				
55–64	466,716	0.125				
65–74	257,963	0.069				
> 74	57,169	0.015				
Race						
Two or more minority races	7,858	0.002				
American Indian or Alaska Native	19,708	0.005				
Asian	252,941	0.068				
Black or African American	290,401	0.078				
Free-form text only	489	0.000				
Joint	81,298	0.022				
Native Hawaiian or Other Pacific Islander	8,458	0.002				
Race not available	471,939	0.127				
White	2,589,173	0.696				
Ethnicity						
Ethnicity not available	444,482	0.119				
Free-form text only	1,107	0.000				
Hispanic or Latino	497,890	0.134				
Joint	90,429	0.024				
Not Hispanic or Latino	2,688,357	0.722				
Sex						
Female	983,117	0.264				
Joint	1,272,557	0.342				
Male	1,279,722	0.344				
Sex not available	186,869	0.050				
Loan type						
Conventional	2,618,917	0.704				
FHA insured	650,815	0.175				
VA guaranteed	407,170	0.109				
RHS/FSA guaranteed	45,363	0.012				
Panel B. Additional continuous controls						
Total units	3,722,265	1.015	0.144	1.000	1.000	1.000
Tract median age of housing	3,722,265	33.125	17.537	19.000	31.000	44.000
One-to-four-family homes	3,722,265	2,030.507	1,144.536	1,366.000	1,836.000	2,423.000
Owner-occupied units	3,722,265	1,524.961	933.713	969.000	1,375.000	1,864.000
Tract population	3,722,265	5,923.616	3,323.858	3,992.000	5,307.000	6,935.000

Notes: This table complements Table 4.1. All variables are measured in the merged HMDA–Equifax mortgage-origination sample. For categorical variables, *N* reports the number of observations in each category and the mean reports the corresponding sample share. For continuous variables, the table reports the mean, standard deviation, and the 25th, 50th, and 75th percentiles. Variable-specific sample sizes reflect missing observations.

E Appendix – Measuring Search

Table E.1. Out-of-sample diagnostics, all five stages, with vs. without race.

Stage	Race	OOS AUC	Calib. slope	Calib. int.	Pregibon $I(\hat{\ell}^2)$	p
consider	with	0.576	0.925	−0.000	0.309	0.004
consider	without	0.573	0.920	−0.000	0.281	0.033
apply	with	0.581	0.907	−0.001	−0.128	0.267
apply	without	0.569	0.895	−0.000	−0.050	0.761
union	with	0.579	0.633	0.002	0.085	0.668
union	without	0.585	0.696	−0.001	0.064	0.766
fear	with	0.647	0.912	0.002	0.057	0.502
fear	without	0.636	0.900	0.001	0.151	0.145
price	with	0.639	0.831	0.001	0.071	0.501
price	without	0.639	0.854	−0.000	0.066	0.541

Notes: Five-fold out-of-sample validation in the NSMO home-purchase sample; all statistics are survey-weighted. Each of the five stage models is estimated with and without race in the predictor set. OOS AUC measures rank discrimination on held-out folds; the calibration slope and intercept come from regressing the realized outcome on the out-of-sample linear predictor (equal to 1 and 0 under perfect calibration); $I(\hat{\ell}^2)$ is the Pregibon link-test statistic and p its p-value, with a small p rejecting the logit link. The *consider* stage fails the link test ($p < 0.05$) and therefore enters only the structural weights, not the reduced form; the *price* stage used in the empirical analysis passes. Because the index is used to rank borrowers, discrimination rather than exact calibration is the relevant validation target.

Table E.2. Propensity stability: with vs. without race (full sample, weighted).

Quantity	mean (with)	mean (no race)	Pearson	Spearman	MAD
pip (empirical)	0.247	0.247	0.934	0.939	0.018
π_{12} (apply> 1)	0.297	0.297	0.890	0.898	0.025
p^P (price|app> 1)	0.828	0.827	0.987	0.985	0.011
π^M (consider> 1)	0.547	0.547	0.934	0.939	0.018
π_{fear}	0.090	0.090	0.858	0.902	0.019
w_0	0.453	0.453	0.934	0.939	0.018
w_{se}	0.269	0.269	0.912	0.916	0.022
w_{ad}	0.278	0.278	0.930	0.930	0.011

Notes: Stability of the imputed search quantities to dropping race from the prediction model, on the full weighted sample. For each object the table reports its weighted mean with and without race, the Pearson and Spearman correlations between the two imputations, and MAD, the weighted mean absolute difference. pip is the price-search propensity of equation (5.3); π_{12}, p^P, π^M, and π_{fear} are its apply-> 1, price-motive, consider-> 1, and fear-motive components, and w_0, w_{se}, and w_{ad} are the structural weights. Means are essentially unchanged and the ranking is preserved: pip has Spearman 0.939, so the searcher ordering used by the Sun–Abraham estimator is largely retained when race is dropped.

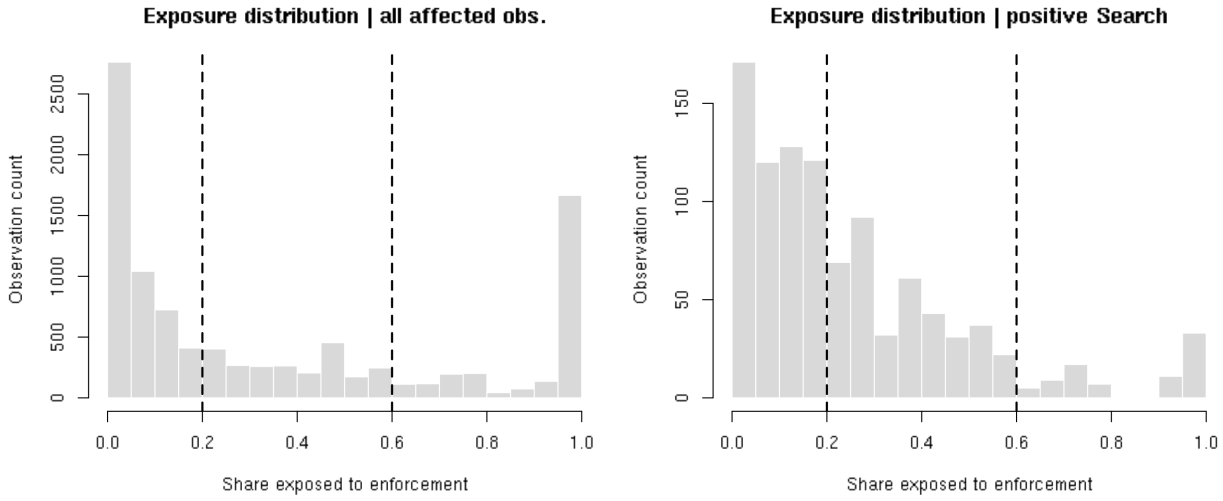
Table E.3. Imputation onto the Equifax–HMDA sample: coverage and the marginal cross-check.

Stage	NSMO fitted mean	Equifax mean	Target coverage	Dropped (no support)
consider > 1 (π^M)	0.539	0.518	0.851	870,146 (14.9%)
apply > 1 (π_{12})	0.282	0.303	0.851	870,146 (14.9%)
union (p^S)	0.904	0.887	0.851	870,146 (14.9%)
fear (p^F)	0.296	0.378	0.851	870,146 (14.9%)
price (p^P)	0.829	0.785	0.851	870,146 (14.9%)
Cross-check (marginal):				
pip marginal rate	0.247	0.240	should match NSMO	

Notes: Imputation of the five NSMO stage models onto the Equifax–HMDA loan panel. The NSMO column reports the full-sample in-sample fitted mean; the Equifax column reports the mean imputed propensity on the loan-level Equifax–HMDA sample. Both are unweighted and therefore directly comparable. A loan is scored only when all four harmonized predictors are present and on the NSMO support; coverage and drops are common to all five stages (0.851; 870,146 of ≈5.83M loans) and are driven by loans missing a harmonized covariate rather than by stage-specific out-of-support (essentially nil here). Loans with any missing or out-of-support predictor are set to missing, not zero. The marginal cross-check validates the two-sample bridge: the marginal pip on Equifax (0.240) lands near the NSMO marginal (0.247) even though the conditional price rate p^P averages ≈0.79–0.83.

F Appendix – Advertisement

Figure F.1. Advertisement Channel: Exposure Distribution



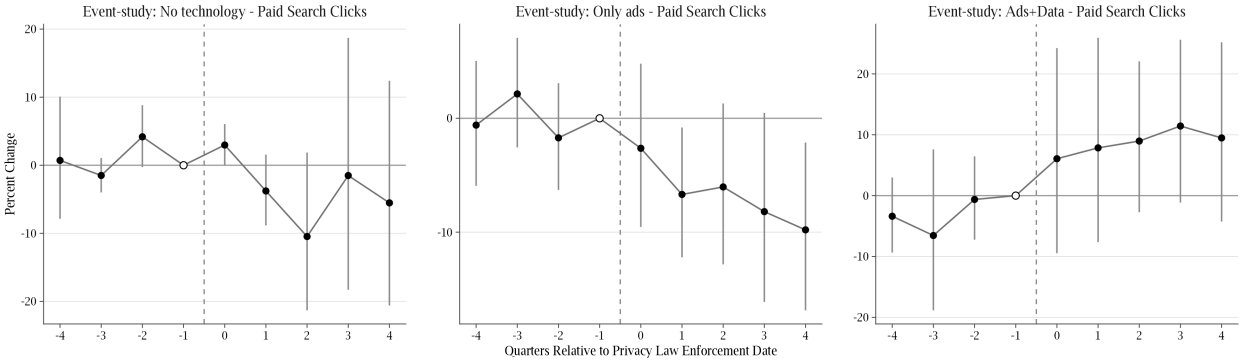
Notes: This figure plots the distribution of lenders' privacy-law *exposure share*, the shift-share regressor used to group lenders in the advertising-spend regression (Table 7.2). For a given lender and month, exposure is the fraction of the lender's pre-period (2018–2019) customer base located in states that have enforced a comprehensive privacy law as of that month; because enforcement rolls out gradually across states, a lender's exposure rises over time and differs across lenders. The histogram pools observations over the lender-month advertising panel and is restricted to the larger lenders that engage in measurable advertising. The distribution is markedly non-linear—mass is concentrated near zero, with additional clusters at higher exposure—so imposing a linear exposure–spend relationship would be misleading and sensitive to outliers. The cutoffs at 0.1, 0.2, and 0.6 define the fixed bins used in the regression: $[0, 0.1)$ (the omitted reference category), $[0.1, 0.2)$, $[0.2, 0.6)$, and $[0.6, 1]$. These breaks are chosen to track the natural mass points of the distribution while retaining enough observations in each bin for inference; the main text describes the construction of the exposure measure and the advertising sample.

Table F.1. Privacy-Law Exposure and Paid Search Clicks: Robustness

	Paid search clicks			
	All waves	Mid-year waves	January waves	Excl. CO/CT
	(1)	(2)	(3)	(4)
Exposure × Only ads	-0.0252	-0.0684**	0.0457	-0.0547
	(0.0431)	(0.0282)	(0.0705)	(0.0726)
Exposure × Ads+Data	0.0624**	0.0923*	-0.0549	0.0949
	(0.0250)	(0.0490)	(0.1295)	(0.0664)
Exposure × No technology	0.0175	0.0153	0.7403**	-0.0246**
	(0.0186)	(0.0180)	(0.3530)	(0.0112)
Only ads: effect (%)	-2.5	-6.6	4.7	-5.3
Enforcement waves	10	7	3	9
Fixed effects				
Group × Year-Month × Wave	Yes	Yes	Yes	Yes
Lender × Wave	Yes	Yes	Yes	Yes
Lender × Month-of-year	Yes	Yes	Yes	Yes
Lender × Year	–	–	–	Yes
Observations	30,350	30,350	30,350	25,672
Lenders	433	433	433	365
Pseudo R^2	0.991	0.991	0.991	0.995

Notes: This table reports Poisson pseudo-maximum-likelihood estimates of the stacked design. Columns (1)–(3) replace lender-by-year with lender-by-wave fixed effects, so that laws enforced in January also contribute to identification. Column (1) pools all enforcement waves. Columns (2) and (3) report coefficients from a single regression that allows separate effects for waves enforced mid-year and in January. Column (4) repeats the specification of Table 7.2, column (2), excluding the July 2023 Colorado–Connecticut wave. The No technology group contains 33 lenders, so its coefficients, particularly in column (3), are imprecisely estimated. Standard errors, clustered at the lender level, are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Figure F.2. Privacy-Law Enforcement and Paid Search Clicks: Event Study



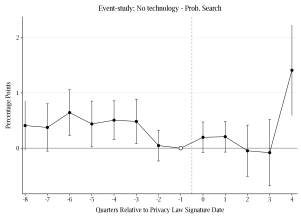
Notes: This figure reports Poisson pseudo-maximum-likelihood event-study estimates corresponding to column (2) of Table 7.2. The specification stacks windows around each privacy-law enforcement wave and includes lender-by-wave, technology-group-by-year-month-by-wave, lender-by-month-of-year, and lender-by-year fixed effects, controlling for exposure to other privacy laws. Each point is the percentage change in paid search clicks, $100(e^{\beta} - 1)$, associated with a 10 percentage point increase in the share of the lender's 2018–2019 originations located in newly regulated states, by quarter relative to enforcement. Quarter -1 (hollow circle) is the reference period. Bars show 95% confidence intervals based on standard errors clustered at the lender level; vertical axes differ across panels. Joint tests that the coefficients for quarters -4 to -2 are zero yield p-values of 0.68 (Only Ads), 0.44 (Ads+Data), and 0.00 (No Technology). The No Technology group contains only 33 lenders, and its estimates are not interpreted.

G Appendix – Robustness

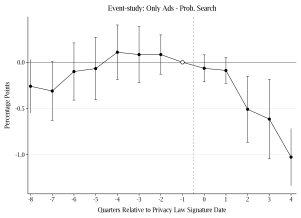
G.1 Signature Date Parallel Trends

Figure G.1. Parallel Trends, Signature Date

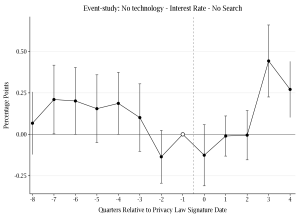
(a): Search-for-Price – No Technology



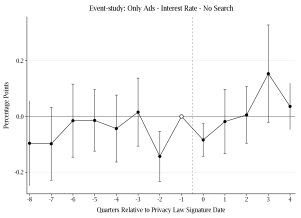
(b): Search-for-Price – Only Ads



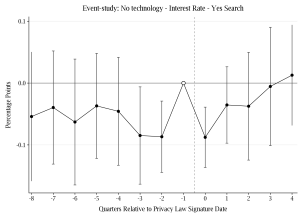
(c): Interest Rate – No Technology, No Search



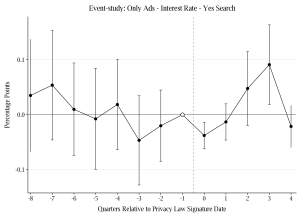
(d): Interest Rate – Only Ads, No Search



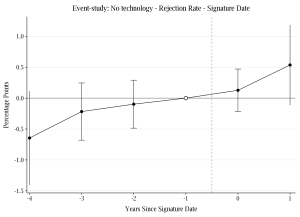
(e): Interest Rate – No Technology, Search



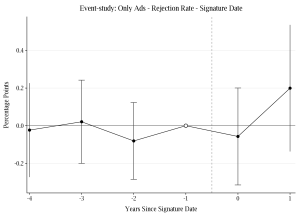
(f): Interest Rate – Only Ads, Search



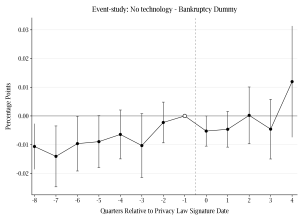
(g): Rejection Rate – No Technology



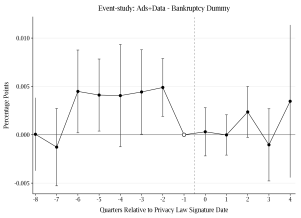
(h): Rejection Rate – Only Ads



(i): Bankruptcy – No Technology



(j): Bankruptcy – Ads + Data



Notes: Event-study coefficients under the alternative "signature" definition of enforcement, plotted in the canonical outcome order: search-for-price probability, mortgage interest rate (no search then search), rejection rate, and borrower bankruptcy. The omitted period is the last pre-enforcement period. Confidence intervals are based on standard errors clustered at the state level.

G.2 Honest DiD

Table G.1. Sensitivity to Violations of Parallel Trends

	Baseline		Breakdown value		
Specification	ATT	95% CI	$\bar{M}$ (RM)	M (SD)	Robust?
Panel A. Search-for-price probability (percentage points)					
No technology	2.07	[1.57, 2.61]	0.32	>0.030	Yes (SD)
Only ads	−0.85	[−1.09, −0.64]	0.43	>0.030	Yes (SD)
Ads + Data	−0.49	[−0.88, −0.12]	0.19	0.018	Yes (SD)
Panel B. Interest rate, non-searchers (basis points)					
No technology	29.44	[−6.63, 65.45]	—	—	(n.s.)
Only ads	−4.13	[−13.36, 5.19]	—	—	(n.s.)
Ads + Data	−9.47	[−15.19, −3.48]	0.09	0.000	No
Panel C. Interest rate, searchers (basis points)					
No technology	4.81	[−4.66, 14.34]	—	—	(n.s.)
Only ads	−13.95	[−18.87, −9.08]	0.12	0.000	No
Ads + Data	−11.65	[−15.12, −8.17]	0.38	>0.030	Yes (SD)
Panel D. Rejection rate (percentage points, annual)					
No technology	−0.53	[−0.95, −0.10]	0.14	>0.030	Yes (SD)
Only ads	−0.22	[−0.42, −0.01]	0.07	0.012	Yes (SD)
Ads + Data	−0.14	[−0.29, 0.01]	—	0.023	(n.s.)
Panel E. Borrower bankruptcy (percentage points)					
No technology	−0.64	[−2.27, 0.89]	—	—	(n.s.)
Only ads	2.09	[1.48, 2.53]	0.66	>0.030	Yes (SD)
Ads + Data	0.39	[−0.48, 1.22]	—	—	(n.s.)

Notes: Sensitivity of the baseline event-study estimates to violations of parallel trends, following Rambachan and Roth (2023). Each ATT averages the effect over the post-enforcement window with Sun–Abraham cohort weights: event quarters 0–3 for the quarterly outcomes (Panels A–C, E) and event years 0–1 for the annual rejection rate (Panel D); the local pre-period window is −4 to −2 and period −1 is omitted. "95% CI" is the baseline interval, clustered by state. The breakdown value is the largest restriction parameter for which the robust 95% set still excludes zero: $\bar{M}$ under the relative-magnitudes restriction $\Delta^{RM}(\bar{M})$ (each post-treatment violation bounded by $\bar{M}$ times the largest pre-treatment violation) and M under the smoothness restriction $\Delta^{SD}(M)$ (bound on the change in slope of the differential trend). Robust sets use the conditional–least-favorable method for Δ^{RM} and the FLCI for Δ^{SD} (RM grid step 0.5; SD grid step 0.005 up to 0.030). ">0.030" means the robust set excludes zero across the whole smoothness grid; "—" means it includes zero at every node; "(n.s.)" denotes specifications not significant at baseline. Under Δ^{RM} every effect breaks down below $\bar{M} = 0.5$ except borrower bankruptcy for Only-ads lenders ($\bar{M} = 0.66$).

G.3 Ads + Data

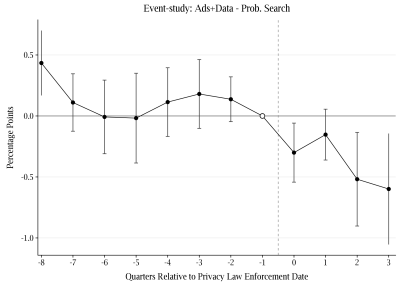
Table G.2. Enforcement: Summary Results (Ads + Data lenders)

Dependent variable:	Search-for-price	Mortgage interest rate		Rejection rate	Bankruptcy
	probability	No search	Search		
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
ATT	-0.0049**	-0.0947***	-0.1165***	-0.0014*	0.0039
	(0.0019)	(0.0298)	(0.0177)	(0.0008)	(0.0046)
Fixed effects					
Year-month × Lender	Yes	Yes	Yes	–	Yes
Year × Lender	–	–	–	Yes	–
State × Lender	–	–	–	Yes	–
MSA × Year	Yes	Yes	Yes	Yes	Yes
State × Month	Yes	Yes	Yes	–	Yes
Loan type	Yes	Yes	Yes	Yes	Yes
Loan purpose	Yes	Yes	Yes	–	Yes
Fit statistics					
Observations	1,834,266	353,922	1,459,400	206,978,641	1,849,622
R^2	0.18137	0.63003	0.65168	0.05500	0.06733
Within R^2	6.66×10^{-5}	0.01368	0.01664	0.00524	0.01999

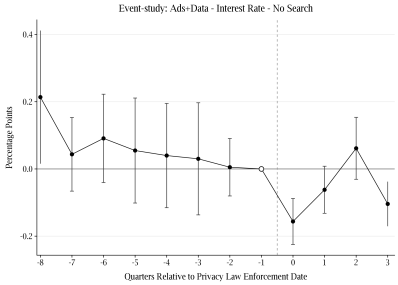
Notes. This table reports average post-enforcement treatment effects (ATT) of state consumer privacy-law enforcement on mortgage-market outcomes for the subsample of lenders that had adopted *both* advertising technologies and customer data platforms ("Ads + Data" lenders) prior to enforcement, estimated on the merged Equifax–HMDA sample. Treatment is defined by the state-level enforcement quarter E_s of the privacy law; the identifying comparison is within lender and calendar period, between borrowers located in treated states and borrowers of the *same* lender in not-yet-treated states. Reported coefficients are the interaction-weighted, staggered-adoption-robust estimates of Sun and Abraham (2021), aggregated across treatment cohorts into a single average post-enforcement effect. The dependent variable is the predicted search-for-price probability in column (1); the mortgage interest rate, in percentage points, in columns (2)–(3), estimated separately for non-searcher and searcher borrowers; the application rejection rate in column (4); and an indicator for subsequent borrower bankruptcy in column (5). Coefficients on the indicator and probability outcomes in columns (1), (4), and (5) are in native units (multiply by 100 for percentage points). Columns (1)–(3) and (5) are estimated on the Equifax-matched origination sample, whereas column (4) is estimated on the HMDA application universe, which accounts for the difference in sample sizes and in the fixed-effect structure. All specifications include the borrower, loan, property, and census-tract controls described in Section 5 (borrower age, race, ethnicity, sex, credit score, and income; loan amount, loan-to-value and debt-to-income ratios, loan type, government-insurance status, occupancy, property value, and number of units; and tract-level demographic and housing controls); the fixed effects included in each column are indicated in the lower panel. Standard errors, in parentheses, are clustered at the borrower-state level (51 clusters, of which 9 are treated during the sample period). Because cluster-robust inference is unreliable with few treated clusters, statistical significance is assessed using wild-cluster bootstrap p-values computed with Webb six-point weights. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Figure G.2. Parallel Trends, Enforcement (Ads + Data lenders)

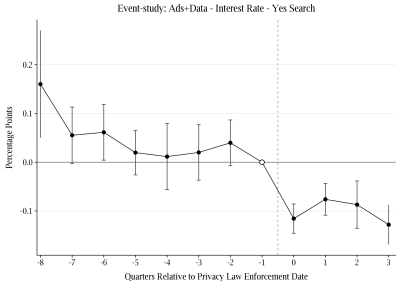
(a): Search-for-Price Probability



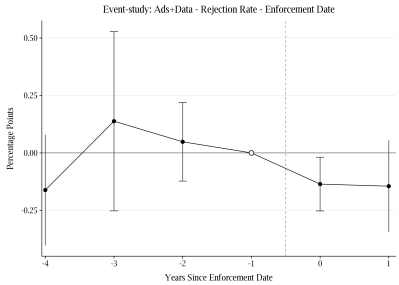
(b): Interest Rate – No Search



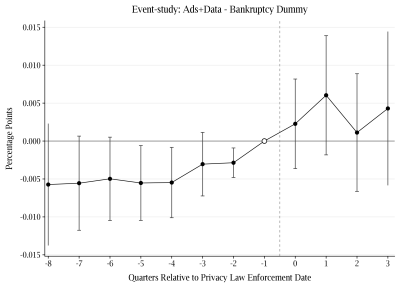
(c): Interest Rate – Search



(d): Rejection Rate



(e): Bankruptcy



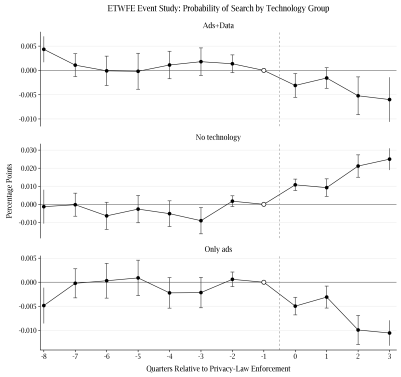
Notes: Event-study estimates for the Ads + Data group, in the canonical outcome order. The omitted period is the last pre-enforcement period. Confidence intervals are based on standard errors clustered at the state level.

G.4 Alternative Estimator: Stacked ETWFE

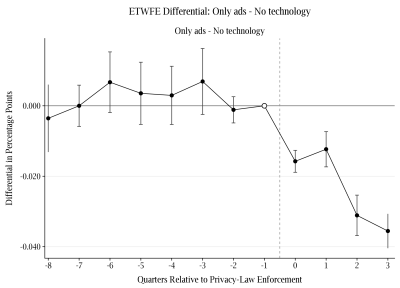
Table G.3. Stacked ETWFE Estimates: Enforcement Effects Across Outcomes (All Cohorts)

	Search-for-price	Mortgage interest rate		Rejection rate	Borrower bankruptcy
	probability	No search	Search		$(t+1)$
	(pp)	(bp)	(bp)	(pp)	(pp)
	(1)	(2)	(3)	(4)	(5)
Average treatment effect on the treated					
No technology	2.088***	29.410*	4.842	-0.525**	-0.664
	(0.262)	(17.106)	(4.793)	(0.213)	(0.801)
Only ads	-0.866***	-4.085	-13.969***	-0.215**	2.006***
	(0.115)	(4.540)	(2.503)	(0.103)	(0.270)
Ads + Data	-0.498**	-9.335***	-11.646***	-0.140*	0.358
	(0.197)	(2.951)	(1.799)	(0.076)	(0.441)
Differences relative to No technology					
Only ads − No technology	-2.955***	-33.495**	-18.812***	0.309	2.670***
	(0.222)	(15.144)	(6.734)	(0.212)	(0.700)
p-value	<0.0001	0.0270	0.0052	0.1438	0.0001
Ads + Data − No technology	-2.587***	-38.745**	-16.489***	0.384*	1.022*
	(0.188)	(15.758)	(5.234)	(0.228)	(0.564)
p-value	<0.0001	0.0139	0.0016	0.0912	0.0699
Fixed effects and controls					
Year-month × Lender	Yes	Yes	Yes	–	Yes
Year × Lender	–	–	–	Yes	–
State × Lender	–	–	–	Yes	–
MSA × Year	Yes	Yes	Yes	Yes	Yes
State × Month	Yes	Yes	Yes	–	Yes
Loan type	Yes	Yes	Yes	Yes	Yes
Loan purpose	Yes	Yes	Yes	Yes	Yes
Borrower controls†	No	Yes	Yes	Yes	Yes
Observations	3,197,140	3,147,224		6,375,539	3,181,331
Clusters	51	51		51	51
R^2	0.1932	0.6539		0.1866	0.0779
Within R^2	0.0001	0.0160		0.0200	0.0200

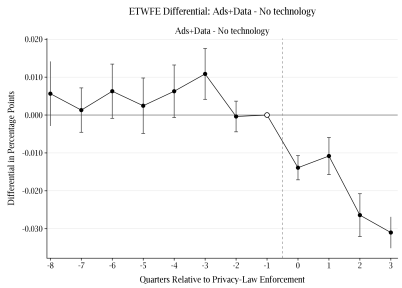
Notes: Stacked event-time ETWFE estimates of the effect of enforcement on each outcome, pooling all enforcement cohorts (including California), with one column per outcome. Columns (2)–(3) come from a single interacted regression, so their observation counts and fit statistics are shared. The top block reports the ATT for each technology group; the middle block reports between-group differences relative to the No-technology group with their p-values. All fixed effects and controls are interacted with the stacking cohort (group). The ATT window pools event-time quarters 0–3 (final bin pooling quarter 3 and later) for columns (1)–(3) and (5); for the annual rejection rate (column 4) it pools event years 0–1 (bin −4 pools −4 and earlier, period 1 pools 1 and later). Estimates are in percentage points (pp), except the interest rate in basis points (bp). Column (4) is estimated on 6,375,539 collapsed cells aggregating 335,669,940 loan-level records. †Borrower controls are borrower and census-tract controls in all columns except (4), which includes borrower controls only. Standard errors clustered at the state level in parentheses; ***, **, and * denote significance at the 1%, 5%, and 10% levels.



(a): ATT by technology group



(b): Difference: Only ads − No technology



(c): Difference: Ads + Data − No technology

Figure G.3. Stacked ETWFE Event Study: Search-for-Price Probability (All Cohorts)

Notes: This figure reports stacked ETWFE event-study estimates for the search-for-price probability, pooling all enforcement cohorts (including California). Panel (a) plots the ATT path by technology group; Panels (b)–(c) plot the between-group differences relative to the No-technology group. The omitted period is the last pre-enforcement period. Confidence intervals are based on standard errors clustered at the state level.

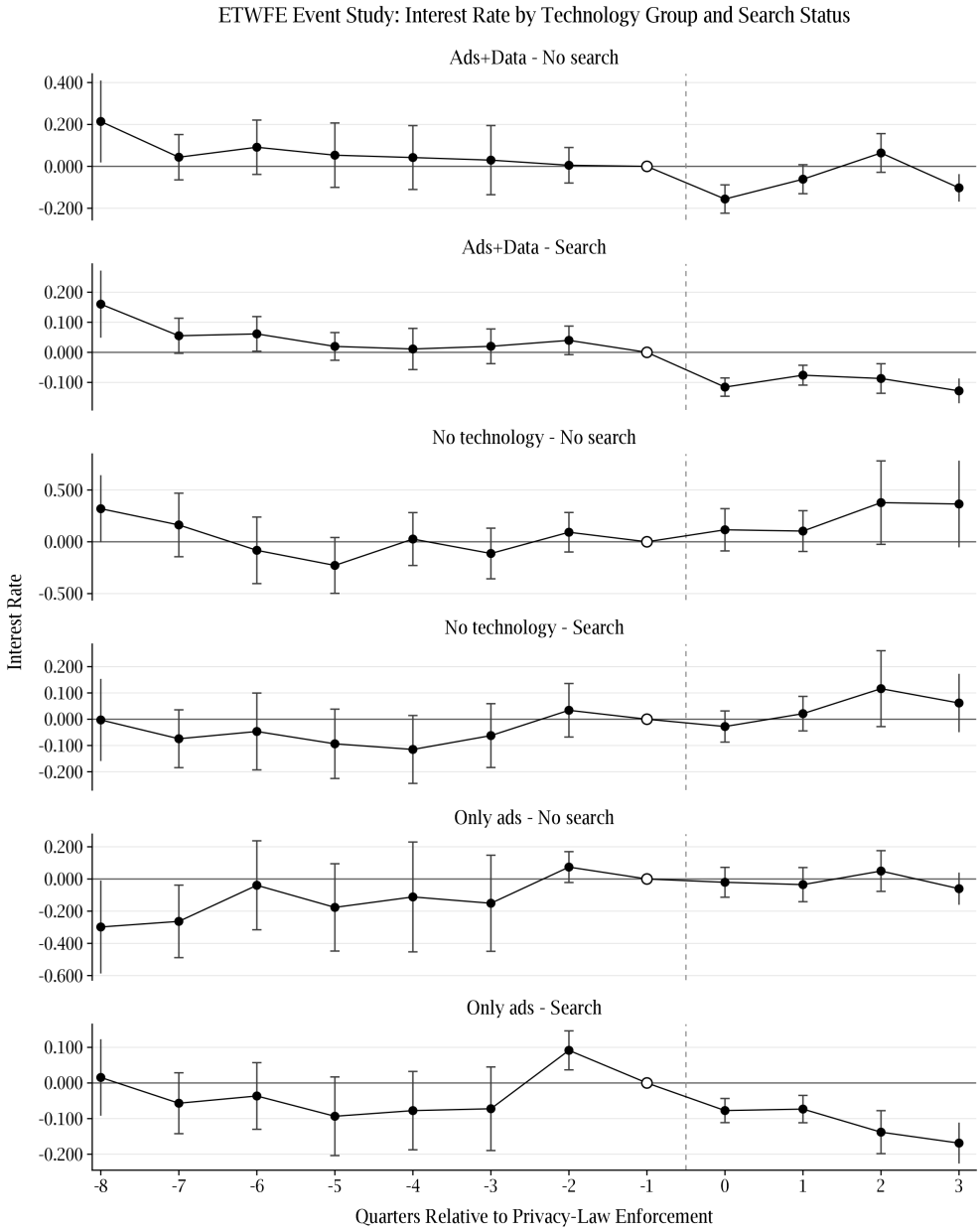
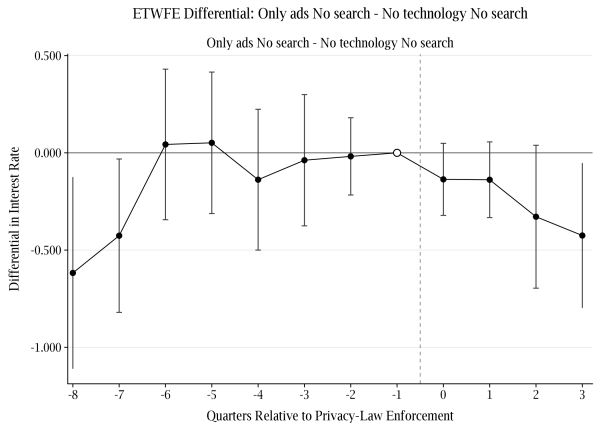
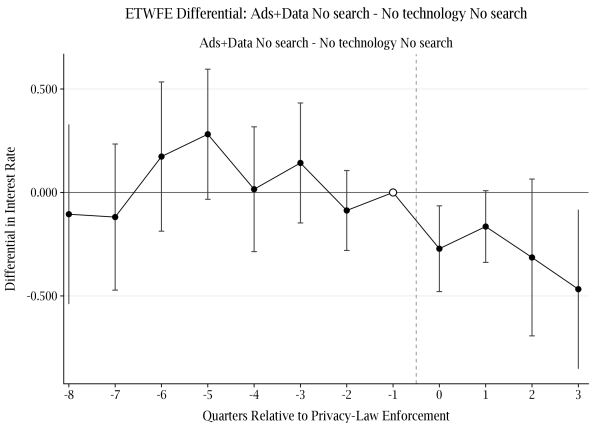


Figure G.4. Stacked ETWFE Event Study: Mortgage Interest Rate by Technology Group and Search Status

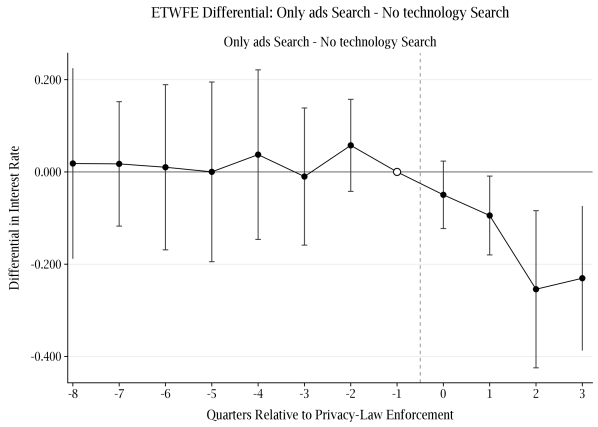
Notes: This figure plots the stacked ETWFE ATT path for mortgage interest rates by technology group and search status. The omitted period is the last pre-enforcement period. Confidence intervals are based on standard errors clustered at the state level.



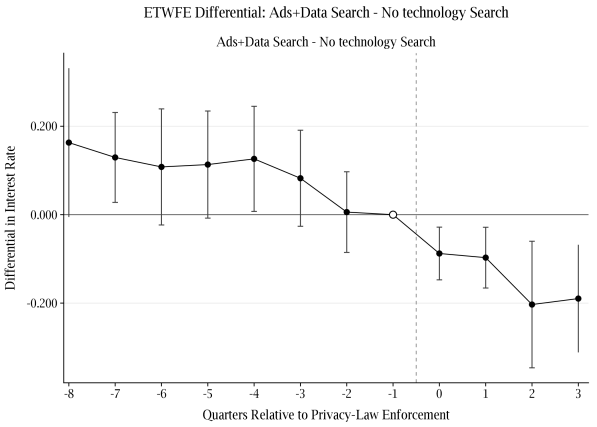
(a): Only ads − No technology (No search)



(b): Ads + Data − No technology (No search)



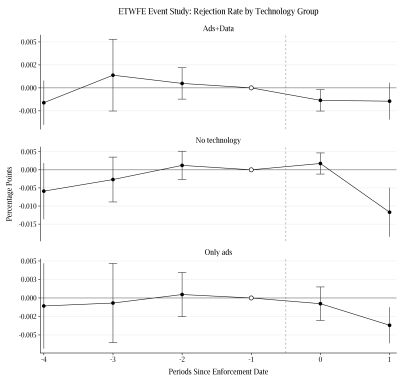
(c): Only ads − No technology (Search)



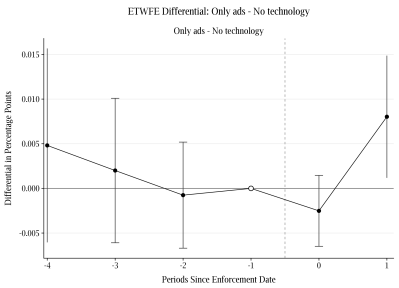
(d): Ads + Data − No technology (Search)

Figure G.5. Stacked ETWFE Event Study: Between-Group Differences in Mortgage Interest Rate

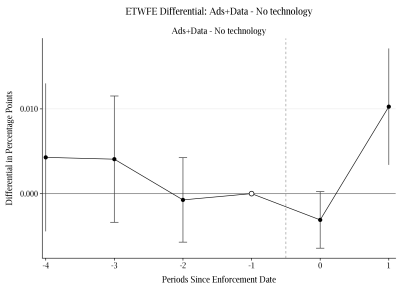
Notes: This figure plots the between-group differences in the stacked ETWFE ATT path for mortgage interest rates, relative to the No-technology group, separately by search-for-price status (no search first, then search). The omitted period is the last pre-enforcement period. Confidence intervals are based on standard errors clustered at the state level.



(a): ATT by technology group



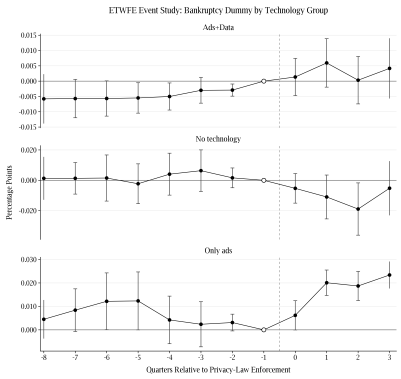
(b): Difference: Only ads − No technology



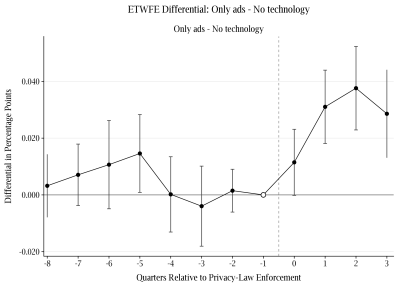
(c): Difference: Ads + Data − No technology

Figure G.6. Stacked ETWFE Event Study: Rejection Rate

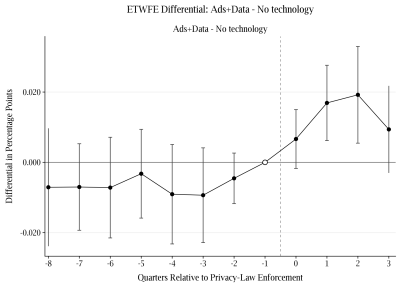
Notes: This figure reports stacked ETWFE event-study estimates for the loan rejection rate. Panel (a) plots the ATT path by technology group; Panels (b)–(c) plot the between-group differences relative to the No-technology group. The omitted period is the last pre-enforcement period. Confidence intervals are based on standard errors clustered at the state level.



(a): ATT by technology group



(b): Difference: Only ads − No technology



(c): Difference: Ads + Data − No technology

Figure G.7. Stacked ETWFE Event Study: Bankruptcy Indicator

Notes: This figure reports stacked ETWFE event-study estimates for the one-period-ahead bankruptcy indicator. Panel (a) plots the ATT path by technology group; Panels (b)–(c) plot the between-group differences relative to the No-technology group. The omitted period is the last pre-enforcement period. Confidence intervals are based on standard errors clustered at the state level.

G.5 Cohort Composition

Table G.4. Cohort Composition of the Enforcement Effects

Dependent variable:	Search-for-price		Interest rate: No search		Interest rate: Search		Rejection rate		Bankruptcy	
Model:	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
	Only ads	Ads+Data	Only ads	Ads+Data	Only ads	Ads+Data	Only ads	Ads+Data	Only ads	Ads+Data
Variables										
cohort = CA	-0.0123***	-0.0064**	-0.0959	-0.1632***	-0.1804***	-0.1595***			0.0239***	0.0048
	(0.0015)	(0.0026)	(0.0580)	(0.0355)	(0.0338)	(0.0235)			(0.0032)	(0.0055)
cohort = VA	-0.0009	-0.0005	-0.0289	0.1599	0.0003	0.0886*			0.0164**	0.0100***
	(0.0019)	(0.0016)	(0.1539)	(0.1683)	(0.0642)	(0.0459)			(0.0070)	(0.0025)
cohort = CO, CT	0.0019	-0.0068***	0.3162	0.3005*	-0.2717***	-0.0283			0.0162**	-0.0023
	(0.0040)	(0.0014)	(0.2212)	(0.1735)	(0.0874)	(0.0683)			(0.0064)	(0.0043)
cohort = UT	0.0030	0.0128***	0.7618	0.2231	0.1526	0.1551***			-0.0233**	-0.0316***
	(0.0053)	(0.0028)	(0.5309)	(0.1966)	(0.1536)	(0.0537)			(0.0095)	(0.0070)
cohort = FL, OR, TX	0.0016	-0.0005	0.0126	-0.1051**	0.0014	-0.0457*			-0.0036	-0.0033
	(0.0014)	(0.0012)	(0.0627)	(0.0401)	(0.0154)	(0.0241)			(0.0038)	(0.0055)
cohort = MT	-0.0013	-0.0032**	-0.6969*	-0.4571***	-0.0980	-0.1915***				
	(0.0044)	(0.0013)	(0.3565)	(0.0865)	(0.1546)	(0.0285)				
cohort = CA							-0.0042***	-0.0023**		
							(0.0011)	(0.0011)		
cohort = VA, CO, CT							-0.0001	-0.0006		
							(0.0014)	(0.0012)		
cohort = UT, FL, OR, TX, MT							0.0007	0.0001		
							(0.0020)	(0.0005)		
Fixed effects										
Year-month × Lender	Yes	Yes	Yes	Yes	Yes	Yes	–	–	Yes	Yes
Year × Lender	–	–	–	–	–	–	Yes	Yes	–	–
State × Lender	–	–	–	–	–	–	Yes	Yes	–	–
MSA × Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Month	Yes	Yes	Yes	Yes	Yes	Yes	–	–	Yes	Yes
Loan type	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan purpose	Yes	Yes	Yes	Yes	Yes	Yes	–	–	Yes	Yes
Fit statistics										
Observations	942,716	1,834,266	176,410	353,922	750,520	1,459,400	100,603,792	206,978,641	925,779	1,849,622
R^2	0.19974	0.18137	0.65251	0.63003	0.66162	0.65168	0.05887	0.05500	0.08498	0.06733
Within R^2	0.00010	6.66×10^{-5}	0.01525	0.01368	0.01486	0.01664	0.00632	0.00524	0.01956	0.01999

Notes: This table decomposes the pooled enforcement effect into cohort-specific estimates, where a cohort is a group of states that began enforcing comprehensive privacy laws in the same period. Each entry is the cohort-specific difference-in-differences estimate for the labeled state group; the aggregate ATT reported in the main text is a cohort-share-weighted average of these interaction-weighted building blocks of the Sun–Abraham estimator. Columns follow the canonical outcome order—search-for-price probability, mortgage interest rate (no search, then search), rejection rate, and borrower bankruptcy—each reported for lenders using only advertising technologies ("Only ads") and for lenders that additionally collect data on their own websites ("Ads + Data"). Search-for-price, rejection, and bankruptcy are linear-probability estimates, so their coefficients are in probability units; the interest rate is in percentage points. Because the rejection rate is observed only in the annual public HMDA data, its cohorts are pooled more coarsely and reported in the lower block. Quarterly outcomes include year-month×lender, MSA×year, and state×month fixed effects, together with loan-type and loan-purpose fixed effects; the rejection specification uses year×lender, state×lender, and MSA×year fixed effects with loan-type controls. Standard errors clustered at the state level are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

G.6 COVID Effect

Table G.5. COVID Robustness: California Event Study Across Outcomes

	Search-for-price probability		Mortgage interest rate			
			No search		Search	
	Only ads	Ads + Data	Only ads	Ads + Data	Only ads	Ads + Data
	(1)	(2)	(3)	(4)	(5)	(6)
California × time-window interactions						
CA × pre [−8, −5]	0.0055	-0.0051***	0.1164	-0.1188***	-0.0493	0.0270**
	(0.0033)	(0.0007)	(0.0815)	(0.0332)	(0.0302)	(0.0119)
CA × pre [−4, −2]	0.0004	0.0005	-0.0023	-0.0083	-0.0384***	-0.0209***
	(0.0009)	(0.0004)	(0.0231)	(0.0149)	(0.0111)	(0.0053)
CA × 2020	-0.0139***	-0.0024	0.1418***	-0.0160	0.1099**	-0.0192*
	(0.0027)	(0.0016)	(0.0441)	(0.0289)	(0.0538)	(0.0103)
CA × 2021	-0.0111***	-0.0094***	0.0002	-0.0840***	0.0425	-0.0582***
	(0.0013)	(0.0014)	(0.0651)	(0.0258)	(0.0587)	(0.0177)
CA × 2022	-0.0117***	-0.0133***	-0.1648	-0.1737***	-0.3074***	-0.1685***
	(0.0037)	(0.0015)	(0.2201)	(0.0451)	(0.0759)	(0.0336)
CA × 2023	-0.0072**	-0.0168***	0.2832	0.2651***	-1.104***	-0.3559***
	(0.0031)	(0.0038)	(0.2241)	(0.0584)	(0.1092)	(0.1148)
CA × 2024	-0.0184***	-0.0146***	-1.449***	-0.8893***	-0.5313***	-0.4474***
	(0.0034)	(0.0014)	(0.2882)	(0.0613)	(0.1174)	(0.0615)
Fixed effects						
Year-month × Lender	Yes	Yes	Yes	Yes	Yes	Yes
MSA × Year	Yes	Yes	Yes	Yes	Yes	Yes
State × Month	Yes	Yes	Yes	Yes	Yes	Yes
Loan type	Yes	Yes	Yes	Yes	Yes	Yes
Loan purpose	Yes	Yes	Yes	Yes	Yes	Yes
Observations	648,047	1,343,837	121,062	258,220	513,310	1,067,979
R^2	0.21451	0.19117	0.67106	0.65640	0.67494	0.66119
Within R^2	1.85×10^{-5}	1.61×10^{-5}	0.01629	0.01383	0.01524	0.01656

Notes: California-specific dynamic difference-in-differences testing whether the enforcement effects are confounded by COVID-19. California—the first jurisdiction to enforce a comprehensive privacy law—is the treated unit; each coefficient is a California indicator interacted with a time-window dummy and measures California's differential change relative to the rest of the sample in that window. The pre-period windows ([−8, −5] and [−4, −2] event quarters) are pre-trend placebo checks, with the last pre-enforcement period omitted; the 2020 window isolates the pandemic onset, so the persistence of the effects in 2021–2024 separates enforcement from a purely COVID-driven shift. Columns (1)–(2) are linear-probability estimates for the search-for-price indicator (probability units); columns (3)–(6) are the interest rate in percentage points, by borrower search status. "Only ads" lenders use third-party advertising technologies; "Ads + Data" lenders also collect data on their own websites. All columns include year-month×lender, MSA×year, and state×month fixed effects plus loan-type and loan-purpose fixed effects, and are estimated separately by column. Standard errors clustered at the state level in parentheses; ***, **, and * denote significance at the 1%, 5%, and 10% levels.

Figure G.8. COVID Robustness: Search-for-Price – Ads + Data, Search

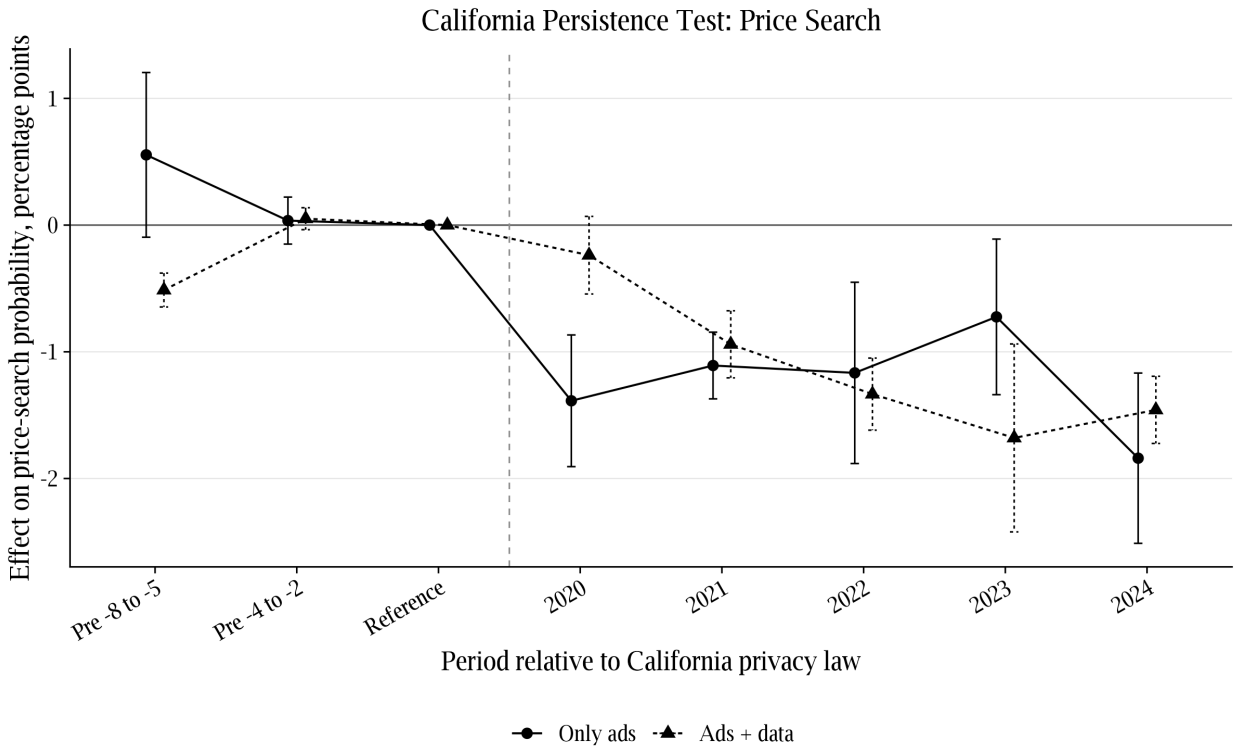
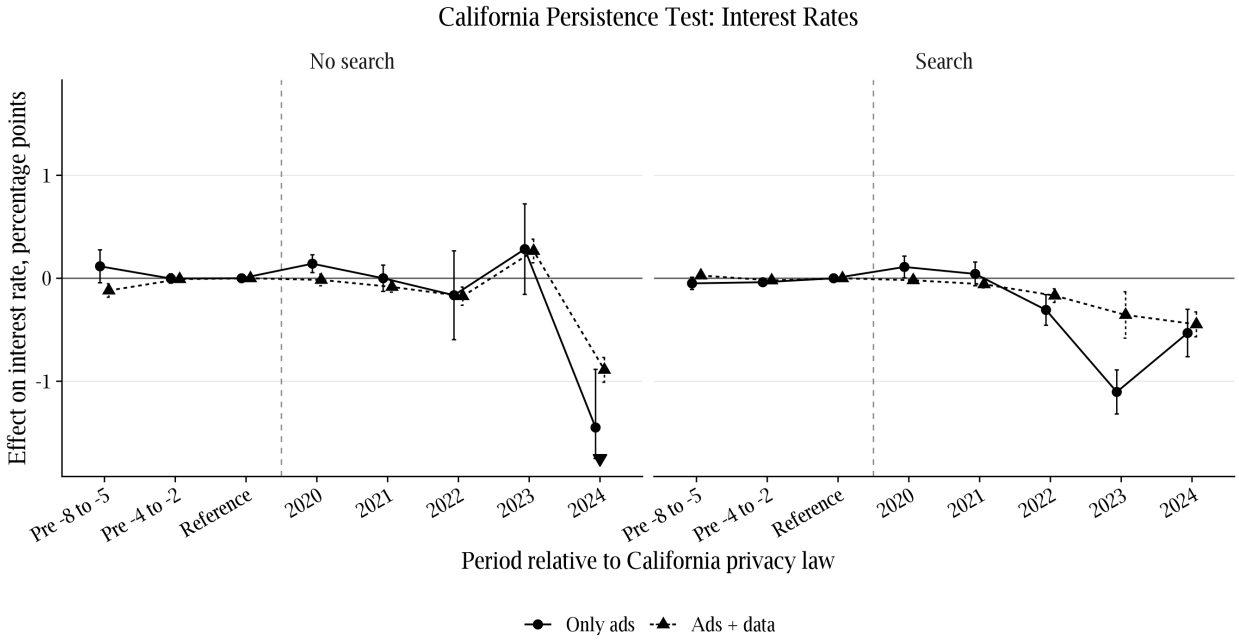


Figure G.9. COVID Robustness: Interest Rate – Ads + Data, Search



95% confidence intervals shown. Triangle markers indicate intervals capped for readability.

G.7 SUTVA

Table G.6. SUTVA: Exposure-Share (Spillover) Test

Dependent variable:	Search-for-price	Mortgage interest rate		Rejection rate	Bankruptcy
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Exposure share	0.0081***	0.0476	0.1892	-0.0293	-0.0093
	(0.0028)	(0.1164)	(0.2382)	(0.0435)	(0.0059)
Exposure share × Only ads	-0.0109*	-0.3416*	-0.4980*	0.0375	0.0087
	(0.0059)	(0.1745)	(0.2723)	(0.0454)	(0.0064)
Exposure share × Ads + Data	-0.0082**	0.1248	-0.0771	0.0358	0.0011
	(0.0034)	(0.1397)	(0.2699)	(0.0443)	(0.0063)
Fixed effects					
Year-month × Lender	Yes	Yes	Yes	–	Yes
Lender	–	–	–	Yes	–
State × Lender	–	–	–	Yes	–
MSA × Year-month	Yes	Yes	Yes	–	Yes
MSA × Year	–	–	–	Yes	–
Loan type	Yes	Yes	Yes	Yes	Yes
Loan purpose	Yes	Yes	Yes	–	Yes
Fit statistics					
Observations	2,313,732	1,683,336	607,745	163,292,370	2,359,406
R^2	0.19925	0.64233	0.61919	0.05060	0.06522
Within R^2	1.62×10^{-5}	0.01387	0.01744	0.00605	0.01945

Notes: This table tests for cross-state spillovers that would violate the stable unit treatment value assumption (SUTVA). In place of a discrete treatment, the regressor is a continuous *exposure share*—the fraction of a lender's pre-enforcement customer base located in states that subsequently enforce a privacy law. If enforcement in treated states spilled over to lending elsewhere, this exposure would predict outcomes beyond the direct treatment effect; the small and largely insignificant coefficients indicate little such contamination. Exposure share is interacted with lender type ("Only ads" and "Ads + Data," with No-technology lenders as the omitted baseline). Column (1) is the search-for-price probability, columns (2)–(3) the mortgage interest rate under two specifications, column (4) the rejection rate, and column (5) borrower bankruptcy. Search-for-price, rejection, and bankruptcy are linear-probability estimates, so their coefficients are in probability units; the interest rate is in percentage points. Fixed effects are as marked in the table: the rejection specification (column 4) uses lender, state×lender, and MSA×year fixed effects, while the remaining columns use year-month×lender and MSA×year-month fixed effects; loan-type and loan-purpose controls are included throughout. Standard errors clustered at the lender level are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

G.8 Sensitivity to Search-Propensity Classification

Table G.7. Interest-Rate Effects by Borrower Search Propensity and Lender Technology

	Low search propensity ($\widehat{\pi} < c$)	High search propensity ($\widehat{\pi} \geq c$)
Panel A: No-technology lenders		
$c = 0.19$	0.294	0.048
	(0.183)	(0.048)
Observations	72,415	334,557
$c = 0.24$	0.291**	-0.043
	(0.117)	(0.054)
Observations	165,619	241,246
$c = 0.28$	0.188**	-0.097
	(0.074)	(0.068)
Observations	271,375	136,238
Panel B: Only-ads lenders		
$c = 0.19$	-0.041	-0.140***
	(0.047)	(0.025)
Observations	176,410	750,520
$c = 0.24$	-0.060*	-0.213***
	(0.033)	(0.024)
Observations	379,796	547,169
$c = 0.28$	-0.098***	-0.223***
	(0.018)	(0.060)
Observations	613,325	315,112
Panel C: Ads + Data lenders		
$c = 0.19$	-0.095***	-0.117***
	(0.030)	(0.018)
Observations	353,922	1,459,400
$c = 0.24$	-0.134***	-0.115***
	(0.027)	(0.019)
Observations	751,448	1,062,049
$c = 0.28$	-0.120***	-0.093***
	(0.021)	(0.031)
Observations	1,209,325	605,884

Notes: This table examines whether the interest-rate effects of privacy-law enforcement are sensitive to the cutoff used to classify borrowers by search propensity. The outcome is the mortgage interest rate, measured in percentage points. Each entry reports the Sun and Abraham aggregated ATT, estimated separately by lender technology regime and borrower search-propensity group. Borrower search propensity, $\widehat{\pi}$, is the predicted probability that a borrower searches for a better mortgage price, imputed from the NSMO-based first-stage model. The thresholds $c = 0.19$, 0.24, and 0.28 are fixed using the full incoming-loan dataset before splitting by lender technology: I rank the incoming mortgage flow by predicted $\widehat{\pi}$ and use the first quartile, median, and third quartile of that full-sample distribution. Low-search-propensity borrowers satisfy $\widehat{\pi} < c$; high-search-propensity borrowers satisfy $\widehat{\pi} \geq c$. All specifications include the baseline controls and fixed effects. Standard errors are reported in parentheses and account for first-stage uncertainty in the imputed search-propensity measure. The sample is restricted to loans with non-missing $\widehat{\pi}$. $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$.

H Appendix – A Model of Screening and Targeted Advertising

This appendix collects the formal environment, proofs, pricing-program details, and quantity accounting underlying Section 9. The order mirrors the main text so that each result can be verified without interrupting the economic argument.

H.1 Complete Environment, Strategies, and Information

Fix a credit-score bucket b. Nature draws repayment type $\nu_i \in \{S, R\}$, comparison type $K_i \in \{K_\ell, K_h\}$, and home-lender type $\tau_i^0 \in \{N, T\}$. In the analytical benchmark,

$$\Pr(\nu_i = R) = \mu_R, \qquad \Pr(K_i = K_h) = \lambda, \qquad \Pr(\tau_i^0 = T) = h_T,$$

and the three draws are independent. The mortgage principal is one. A safe borrower repays. A risky borrower generates loss given default d. Let s_T denote the share of technology lenders in the live comparison market, equivalently the probability that any one outside quote comes from a technology lender.

The borrower observes K_i, all realized approval decisions, and all live offers before choosing among them. Whether the borrower observes ν_i is behaviorally irrelevant in the baseline because repayment type affects lender losses but not borrower utility conditional on approval. A convenient borrower information set at the comparison stage is

$$\mathcal{I}_i^B = \left\{K_i,\, b,\, a_i^{\text{home}},\, r_b^{\text{home}},\, r_b^{\text{ad}}\mathbf{1}\{\text{targeted offer}\},\, \widetilde{r}_b^{s},\, s_T\right\}. \tag{H.1}$$

Before screening an individual applicant, lenders observe the score bucket, aggregate type distributions, home-lender shares, the comparison-market lender share, and the information environment, but not ν_i or K_i. Conditional on the incumbent approval decision, a common pre-screening lender information set can be written

$$\mathcal{I}_{ib}^{L,0} = \left\{b,\, \Pr(\nu_i \mid b),\, \Pr(K_i \mid b),\, h_N, h_T,\, s_T,\, I,\, a_i^{\text{home}}\right\}. \tag{H.2}$$

Lender ℓ additionally observes its private screening outcome $\sigma_{i\ell}$ before deciding whether to issue an approved offer:

$$\mathcal{I}_{i\ell b}^{L,1} = \mathcal{I}_{ib}^{L,0} \cup \{\sigma_{i\ell}\}. \tag{H.3}$$

A lender strategy consists of score-bucket rate schedules for active comparison, incumbent retention, and—for technology lenders—paid acquisition, together with an exit option. Borrower strategies specify sequential choices after each history: incumbent rejection, incumbent approval

without targeted contact, and incumbent approval with successful targeted contact. Assumption 1 restricts the baseline after incumbent rejection to immediate exit. The extension in Appendix H.6 relaxes that restriction.

H.2 Screening and the Comparison Market

This subsection proves Lemma 1. At $I = 0$, both lender types use the same screening probabilities. For $I > 0$, $q_{TS} > q_{NS}$ and $q_{TR} < q_{NR}$, so

$$\frac{q_{TR}}{q_{TS}} < \frac{q_{NR}}{q_{NS}}.$$

Because

$$\frac{\rho_\tau}{1-\rho_\tau} = \frac{\mu_R}{\mu_S}\frac{q_{\tau R}}{q_{\tau S}},$$

it follows that $\rho_T < \rho_N$.

For the derivative, write $q_S = q_{TS}$ and $q_R = q_{TR}$. From equation (9.3),

$$\rho_T'(I) = \frac{\mu_S\mu_R\left(q_R'(I)q_S(I) - q_R(I)q_S'(I)\right)}{A_T(I)^2} < 0, \tag{H.4}$$

because $q_S'(I) = \alpha_S'(I) > 0$ and $q_R'(I) = -\alpha_R'(I) < 0$. With f_b and d fixed, equation (9.4) then gives $\mathrm{MC}_{Tb}'(I) < 0$ and $\partial r_{Tb}^s/\partial I < 0$. Aggregate approval instead satisfies

$$A_T'(I) = \mu_S\alpha_S'(I) - \mu_R\alpha_R'(I), \tag{H.5}$$

which is not signed.

Under Assumption 2, a symmetric live quote above $\mathrm{MC}_{\tau b}$ can be undercut by another lender of the same type, whereas a quote below marginal cost yields a loss. Hence $r_{\tau b}^s = \mathrm{MC}_{\tau b}$.

A searching borrower draws two outside lender types independently. A technology lender appears at least once with probability

$$\beta(s_T) = 1 - (1-s_T)^2 = 2s_T - s_T^2. \tag{H.6}$$

Since $r_{Tb}^s < r_{Nb}^s$, the borrower funds at the technology quote whenever at least one technology lender is drawn. Therefore

$$\widetilde{r}_b^s = \beta(s_T)r_{Tb}^s + [1-\beta(s_T)]r_{Nb}^s = r_{Nb}^s - \beta(s_T)d(\rho_{Nb} - \rho_{Tb}). \tag{H.7}$$

Holding s_T fixed,

$$\frac{\partial \widetilde{r}_b^s}{\partial I} = \beta(s_T) d \rho'_{Tb}(I) < 0. \tag{H.8}$$

This completes the proof of Lemma 1.

H.3 Pricing Programs, the Maintained Region, and Proposition 1

Fix I, score bucket b, and the comparison-rate schedule. Because demand is constant between the two reservation thresholds, only a finite set of rates can maximize profit. Let r_{hb}^- be the highest grid rate below r_{hb}. Normalizing each relevant reached pool to one, the candidate payoffs are:

Table H.1. Finite Pricing Candidates

Pricing problem	Full-pool candidate	High-K candidate	Exit
Technology incumbent	$r_{\ell b} - \mathrm{MC}_{Tb}$	$\lambda(r_{hb} - \mathrm{MC}_{Tb})$	0
Conventional incumbent	$r_{\ell b} - \mathrm{MC}_{Nb}$	$\lambda(1-\gamma)(r_{hb} - \mathrm{MC}_{Nb})$	0
Targeted technology lender	$(r_{\ell b} - \mathrm{MC}_{Tb}) + B$	$\lambda[(r_{hb}^- - \mathrm{MC}_{Tb}) + B]$	0

Notes: The high-price conventional-incumbent mass is reduced by successful targeting. The continuation value B is earned only after a paid conversion and therefore appears only in the paid-channel problem.

The rate grid has tick $\varepsilon < K_h - K_\ell$ and contains both reservation thresholds. Ties between the targeted offer and active comparison go to the targeted offer; ties involving incumbent retention go to the incumbent. The condition that the low-friction reservation rate exceed every realized comparison quote is

$$K_\ell > \beta(s_T) d (\rho_{Nb} - \rho_{Tb}). \tag{H.9}$$

The exact candidate comparisons supporting Assumption 3 are

$$\lambda(r_{hb} - \mathrm{MC}_{Tb}) \geq \max\{r_{\ell b} - \mathrm{MC}_{Tb}, 0\}, \tag{H.10}$$

$$\lambda(1-\gamma)(r_{hb} - \mathrm{MC}_{Nb}) \geq \max\{r_{\ell b} - \mathrm{MC}_{Nb}, 0\}, \tag{H.11}$$

$$(r_{\ell b} - \mathrm{MC}_{Tb}) + B > \max\left\{\lambda[(r_{hb}^- - \mathrm{MC}_{Tb}) + B], 0\right\}. \tag{H.12}$$

The first two inequalities select the high-threshold incumbent candidate. The third selects the pooled targeted candidate. These are global discrete comparisons, not first-order conditions.

Rearranging equation (H.12), a sufficient supporting value of the paid-conversion continuation value is

$$B > \max\left\{0, \frac{\lambda(r_{hb}^- - \mathrm{MC}_{Tb}) - (r_{\ell b} - \mathrm{MC}_{Tb})}{1-\lambda}\right\}. \tag{H.13}$$

A finite value exists because $\lambda < 1$. The model does not identify B and the quantitative implementation does not use equation (H.13) as a moment. Within the maintained pricing region, B selects the

pooled paid equilibrium but does not affect the equilibrium rates or quantities used in estimation.

Proof of Proposition 1. For an incumbent, any rate below $r_{\ell b}$ serves the same pool as $r_{\ell b}$ but earns a lower margin. Any rate in $(r_{\ell b}, r_{hb}]$ serves only the relevant high-K segment, and profit is increasing in the rate within that segment, so the best such rate is r_{hb}. Equations (H.10)–(H.11) select this candidate over pooling and exit.

For a targeted lender, charging $r_{\ell b}$ makes a low-K borrower indifferent between accepting and comparing and makes a high-K borrower strictly prefer the targeted offer to the incumbent. The targeted–comparison tie is assigned to the targeted offer. A lower rate serves the same pool at a lower margin. Any rate between the two thresholds serves only high-K borrowers, so r_{hb}^{-} weakly dominates all lower rates in that interval. Equation (H.12) selects $r_{\ell b}$ over that candidate and exit.

Finally, equation (H.9) implies $r_{\ell b} > \max_\tau r_{\tau b}^s$, while $K_h > K_\ell$ implies $r_{hb} > r_{\ell b}$. Borrower choices are therefore sequentially optimal and the candidate comparison exhausts all grid deviations. This proves Proposition 1.

H.4 Quantity Accounting and Proof of Proposition 2

The mass of borrowers reaching the pricing game is

$$M_0 = h_N A_N + h_T A_T. \tag{H.14}$$

A paid technology origination requires approval by an N home lender and successful early targeting:

$$Q_T^{\text{ad}} = h_N A_N \gamma. \tag{H.15}$$

Low-K borrowers who do not accept a targeted offer enter active comparison, so the comparison mass is

$$\Sigma = (1-\lambda)\left[h_T A_T + h_N A_N (1-\gamma)\right]. \tag{H.16}$$

Since a technology lender appears in at least one of the two comparison draws with probability $\beta(s_T)$,

$$Q_T^s = \beta(s_T)\Sigma, \qquad Q_N^s = (1-s_T)^2\Sigma. \tag{H.17}$$

High-K borrowers remain with the home lender unless an N-home borrower is successfully acquired through the paid channel:

$$Q_T^{\text{home}} = \lambda h_T A_T, \qquad Q_N^{\text{home}} = \lambda h_N A_N (1-\gamma). \tag{H.18}$$

Adding up. Substituting equations (H.15)–(H.18),

$$\begin{aligned} &Q_T^{\text{ad}} + Q_T^s + Q_N^s + Q_T^{\text{home}} + Q_N^{\text{home}} \\ &\quad = h_N A_N \gamma + (1-\lambda)\left[h_T A_T + h_N A_N (1-\gamma)\right] + \lambda h_T A_T + \lambda h_N A_N (1-\gamma) \\ &\quad = h_N A_N + h_T A_T = M_0. \end{aligned} \tag{H.19}$$

Among new technology-lender acquisitions, the paid share and average rate are

$$w_T^{\text{ad}} = \frac{Q_T^{\text{ad}}}{Q_T^{\text{ad}} + Q_T^s}, \qquad r_T^{\text{new}} = (1 - w_T^{\text{ad}}) r_T^s + w_T^{\text{ad}} r^{\text{ad}}. \tag{H.20}$$

Proof of Proposition 2. Let privacy reduce usable information from I to $I' < I$. By Lemma 1, ρ_T rises. Equations (9.3), (9.4), and (9.5) then imply

$$\Delta \text{MC}_{Tb} > 0, \qquad \Delta r_T^s > 0, \qquad \Delta \widetilde{r}_b^s > 0.$$

Proposition 1 therefore gives $\Delta r^{\text{ad}} > 0$ and $\Delta r^{\text{home}} > 0$. The conventional comparison quote is unchanged.

Because $Q_T^{\text{ad}} = h_N A_N \gamma(I)$ and the conventional screening state is independent of I, $\gamma(I') < \gamma(I)$ implies $\Delta Q_T^{\text{ad}} < 0$. The paid share need not fall because Q_T^s also changes.

Starting from equation (H.20), add and subtract $(1 - w_T^{\text{ad}\prime}) r_T^s + w_T^{\text{ad}\prime} r^{\text{ad}}$ to obtain

$$\Delta r_T^{\text{new}} = (1 - w_T^{\text{ad}\prime}) \Delta r_T^s + w_T^{\text{ad}\prime} \Delta r^{\text{ad}} + (r^{\text{ad}} - r_T^s) \Delta w_T^{\text{ad}}. \tag{H.21}$$

The first two terms are positive. If $\Delta w_T^{\text{ad}} < 0$, the final term is negative, and the total rate falls exactly when its absolute value exceeds the first two terms, which is condition (9.12). If $\Delta w_T^{\text{ad}} \geq 0$, all three terms are nonnegative. Finally, equation (H.5) leaves the aggregate approval response unsigned. This proves Proposition 2.

H.5 Continuous Comparison Frictions: Optional Robustness

The two-point distribution of K_i makes the pricing logic transparent but is not essential for defining the stopping problem. Suppose instead that $K_i \sim F_K$ on support $[\underline{K}, \overline{K}]$, with F_K continuous. Conditional on a live incumbent offer and no targeted offer, the borrower enters comparison whenever

$$K_i < r_b^{\text{home}} - \widetilde{r}_b^s. \tag{H.22}$$

If a targeted offer satisfies $r_b^{\mathrm{ad}} \le r_b^{\mathrm{home}}$, a reached borrower accepts it rather than entering comparison whenever

$$K_i \ge r_b^{\mathrm{ad}} - \widetilde{r}_b^{s}. \tag{H.23}$$

Thus the incumbent and targeted lender continue to price against different quantiles of the comparison-friction distribution. In this extension, however, the optimal thresholds need not coincide with two exogenously given mass points; they are determined by the trade-off between margin and the mass served. A formal continuous-K proposition should therefore be stated only after the relevant regularity conditions on F_K and the global pricing problem are verified.

H.6 Search After Rejection and Other Robustness Extensions

Search after rejection

The baseline conditions active price comparison on a live incumbent approval. This isolates the paper's shopping mechanism from repeated search for approval. A natural extension allows a borrower rejected by the incumbent to submit another application. Such an extension requires an additional continuation value from another application that incorporates both the probability of approval and any additional application cost. Approval risk then enters the borrower's stopping problem directly, unlike in the baseline, where all compared quotes are already live and approved. The key object to verify is whether an early targeted prequalification still arrives before outside comparison and therefore continues to alter the borrower's stopping threshold. A formal robustness proposition should be added only after the continuation problem and lender information after rejection are fully specified.

Correlation between repayment and comparison types

The analytical benchmark assumes $\Pr(K_i \mid \nu_i) = \Pr(K_i)$. This assumption can be relaxed by allowing the distribution of comparison frictions to differ across safe and risky borrowers. The screening formulas then use the type-specific joint masses rather than the product of marginal probabilities. The central screening result continues to require that better information lower the risky share of technology-lender originations. The measured-rate reversal additionally requires that the resulting change in channel weights remain large enough to satisfy condition (9.12). Because these restrictions depend on the joint distribution, no general sign claim is imposed here.

More than two outside quotes

If a borrower who activates comparison draws $n \geq 2$ lender types independently, a technology lender appears at least once with probability

$$\beta_n(s_T) = 1 - (1 - s_T)^n. \tag{H.24}$$

When $r^s_{Tb} < r^s_{Nb}$, the expected funded comparison rate becomes

$$\widetilde{r}^s_{b,n} = \beta_n(s_T) r^s_{Tb} + [1 - \beta_n(s_T)] r^s_{Nb}. \tag{H.25}$$

More outside quotes increase the probability of encountering the lower-cost technology lender and therefore weakly reduce the expected comparison rate. The reservation thresholds become $\widetilde{r}^s_{b,n} + K_\ell$ and $\widetilde{r}^s_{b,n} + K_h$; the early-acquisition logic is otherwise unchanged as long as the comparison wedge remains positive.

Selection without continuation value

Equation (H.13) makes clear when the continuation value is not needed. Setting $B = 0$ supports the pooled paid candidate whenever

$$r_{\ell b} - \mathrm{MC}_{Tb} > \max\left\{\lambda(r^-_{hb} - \mathrm{MC}_{Tb}),\, 0\right\}. \tag{H.26}$$

Thus B is not intrinsically tied to the economic mechanism; it is a convenient device for selecting the pricing region when current-loan margins alone do not satisfy this inequality.

I Appendix – Quantitative Implementation, Identification, and Estimation

I.1 Three-Lender Quantitative Economy

This appendix provides the complete three-lender quantitative economy underlying Section 10. It defines the quote-draw process, paid-contact technology, latent-to-observed Searcher mapping, channel quantities, and sequential policy experiment used in the final calibration. Appendix I.2 then describes how these objects are disciplined by the data and distinguishes targeted moments from held-out validation.

Rate Block and Quote Draws

Let $g, h, t \in \{N, A, D\}$ denote lender types, b a credit-score bucket, and $u \in \{0, 1\}$ the privacy-enforcement state. Marginal resource cost and the borrower-facing comparison quote are those defined in Section 10. If two approved outside quotes are drawn independently from lender-type distribution $s^q_{\cdot bu}$, the expected comparison rate is

$$\widetilde{r}^s_{bu} = \sum_{j\in\{N,A,D\}} \sum_{k\in\{N,A,D\}} s^q_{jbu} s^q_{kbu} \min\left\{r^s_{jbu}, r^s_{kbu}\right\}. \tag{I.1}$$

In the two-type analytical benchmark, the corresponding expression reduces to the probability-weighted comparison-rate formula under the maintained type ordering. The quantitative implementation retains the expected-minimum form in equation (I.1) because the relative ordering of N, A, and D comparison quotes may vary across score bins and policy states.

The targeted-offer and incumbent rates then satisfy the threshold-pricing relations in equation (40). The quantitative implementation imposes the corresponding probability bounds and channel-ordering restrictions described in Appendix I.2.

Paid Contact and Target Selectivity

Let $P_{ht,bu}$ denote the probability that a paid campaign by target lender type t reaches a borrower initially attached to home-lender type h in score bin b and policy state u. Rows of P_{bu} index the home lender and columns index the target lender. Define the Only-Ads share of the local technology search market by

$$\phi_{A,b} = \frac{s^q_{A,b0}}{s^q_{A,b0} + s^q_{D,b0}}. \tag{I.2}$$

The pre-enforcement reach matrix used in the quantitative implementation is

$$P_{b0} = \begin{pmatrix} 0 & \Gamma\phi_{A,b} & \Gamma(1-\phi_{A,b}) \\ 0 & 0 & \Gamma\xi \\ 0 & \Gamma\xi & 0 \end{pmatrix}, \qquad \xi = 1 \text{ in the frozen V5-MC3.1 specification.} \tag{I.3}$$

Thus No-Tech lenders never target another borrower, and own-customer marketing is absorbed into the home channel. In the frozen V5-MC3.1 specification, $\xi = 1$, so Only Ads and Ads+Data may target one another. The first row allocates paid contacts originating from No-Tech home borrowers between the two technology groups using their observed local technology shares.

Privacy scales contacts by the technology of the target lender:

$$P_{ht,b1} = P_{ht,b0}\delta_t, \qquad t \in \{A, D\}, \qquad \delta_A = 0.240, \quad \delta_D = 0.645, \tag{I.4}$$

where the retention parameters are calibrated independently from the Vivvix advertising-spending evidence. Conditional on approval by home lender h, the probability that target lender t also approves the borrower is

$$a_{ht,bu} = \frac{(1 - \mu_{R,b})q_{hS,bu}q_{tS,bu} + \mu_{R,b}q_{hR,bu}q_{tR,bu}}{A_{hbu}}. \tag{I.5}$$

The parameter τ_t then allows paid acquisition to be selective with respect to latent comparison type by scaling successful high-K contacts relative to low-K contacts. This selection margin is distinct from ζ_t, which governs the composition of the post-privacy organic-search market rather than the survival of a paid contact.

Latent Comparison Types and Empirical Searcher Status

Let the latent comparison type be $K_i \in \{K_\ell, K_h\}$, with score-bin high-friction share $\lambda_b = \Pr(K_i = K_h \mid b)$ and low-friction share $1 - \lambda_b$. The empirical Searcher indicator Z_i is linked to latent type through

$$\eta_{L,b} = \Pr(Z_i = 1 \mid K_i = K_\ell, b), \qquad \eta_{H,b} = \Pr(Z_i = 1 \mid K_i = K_h, b), \qquad \eta_{L,b} > \eta_{H,b}. \tag{I.6}$$

The mapping is deterministic given the imported search-propensity first stage; no additional measurement-error parameter is estimated. Once the latent channel quantities are constructed below, their unnormalized masses in the empirical Searcher sample are

$$\begin{aligned} M^{\text{paid}}_{gbu}(Z = 1) &= \eta_{L,b}Q^{\text{ad},L}_{gbu} + \eta_{H,b}Q^{\text{ad},H}_{gbu}, \\ M^{\text{search}}_{gbu}(Z = 1) &= \eta_{L,b}Q^{s}_{gbu}, \\ M^{\text{home}}_{gbu}(Z = 1) &= \eta_{H,b}Q^{0}_{gbu}. \end{aligned} \tag{I.7}$$

The corresponding non-Searcher masses replace $\eta_{L,b}$ and $\eta_{H,b}$ with $1-\eta_{L,b}$ and $1-\eta_{H,b}$. Normalizing the three masses within (g, b, u, Z) gives the paid, search, and home weights used in the model counterpart of the empirical mortgage-rate regression. Paid originations can therefore contain both latent comparison types even though low-K borrowers are more likely to be classified as Searchers.

Comparison Allocation and Measurement Normalization

The model distinguishes the probability that lender type g appears in an outside quote draw, s^{q}_{gbu}, from the share of funded comparison originations allocated to that type, π^{s}_{gbu}. Pre-enforcement quote-draw weights are borrower-weighted shares of active lender types in county-by-month markets, as documented in Table I.3. The post-privacy organic-search environment reweights those observed

pre-law shares according to

$$
\begin{aligned}
s^q_{N,b1} &= \frac{s^q_{N,b0}}{\mathcal{D}_b}, \\
s^q_{A,b1} &= \frac{\zeta_A s^q_{A,b0}}{\mathcal{D}_b}, \qquad \mathcal{D}_b = s^q_{N,b0} + \zeta_A s^q_{A,b0} + \zeta_D s^q_{D,b0}. \\
s^q_{D,b1} &= \frac{\zeta_D s^q_{D,b0}}{\mathcal{D}_b},
\end{aligned}
\tag{I.8}
$$

Thus ζ_g is a composition multiplier rather than a probability; values above one are admissible. The normalization in (I.8) ensures that the post-privacy quote-draw shares sum to one.

The baseline allocation normalization is

$$
\pi^s_{gbu} = s^q_{gbu}, \qquad g \in \{N, A, D\},\ u \in \{0, 1\}. \tag{I.9}
$$

This restriction closes the mapping from observed local lender availability to funded comparison quantities; it is an empirical allocation normalization, not an equilibrium condition. A winner-probability allocation based directly on the two-quote price game is a natural robustness specification.

Channel Quantities and Lost-Contact Rerouting

For readability, suppress (b, u) in this subsection. Let h_h denote the home-lender weight and let

$$
H_h = h_h A_h, \qquad H^L_h = (1 - \lambda) H_h, \qquad H^H_h = \lambda H_h \tag{I.10}
$$

be the approved home mass and its latent low- and high-friction components. Successful paid-acquisition masses from home type h to target type t are

$$
S^L_{ht} = H^L_h P_{ht} a_{ht}, \qquad S^H_{ht} = H^H_h \tau_t P_{ht} a_{ht}, \tag{I.11}
$$

so the residual masses after paid targeting are

$$
R^L_h = H^L_h - \sum_t S^L_{ht}, \qquad R^H_h = H^H_h - \sum_t S^H_{ht}. \tag{I.12}
$$

All residual low-K borrowers enter organic comparison. Using the allocation normalization in (I.9), funded comparison, incumbent, and paid quantities are

$$\begin{aligned} Q_g^s &= \pi_g^s \sum_h R_h^L, \\ Q_g^0 &= R_g^H, \\ Q_g^{\mathrm{ad},L} &= \sum_h S_{hg}^L, \qquad Q_g^{\mathrm{ad},H} = \sum_h S_{hg}^H. \end{aligned} \tag{I.13}$$

Hence

$$Q_g^L = Q_g^s + Q_g^{\mathrm{ad},L}, \qquad Q_g^H = Q_g^0 + Q_g^{\mathrm{ad},H}, \qquad Q_g = Q_g^L + Q_g^H. \tag{I.14}$$

The construction is required to satisfy nonnegativity and the accounting identity $\sum_g Q_g = \sum_h H_h$ up to numerical tolerance.

Privacy is evaluated sequentially. First screening changes while acquisition shares are held at their pre-law configuration. Second the paid-contact matrix is scaled by the retention parameters δ_A and δ_D. Let $w_{gb}^{p,\mathrm{scr}}$ and $w_{gb}^{p,\mathrm{tar}}$ denote the empirical Searcher-sample paid weights after the screening stage and after the paid-retention stage. Lost paid mass is

$$L_{gb} = \max\left\{ w_{gb}^{p,\mathrm{scr}} - w_{gb}^{p,\mathrm{tar}}, 0 \right\}. \tag{I.15}$$

The sufficient-statistic rerouting step is

$$\begin{aligned} w_{gb}^{s,\mathrm{tar,new}} &= w_{gb}^{s,\mathrm{scr}} + \omega_g L_{gb}, \\ w_{gb}^{h,\mathrm{tar,new}} &= 1 - w_{gb}^{p,\mathrm{tar}} - w_{gb}^{s,\mathrm{tar,new}}. \end{aligned} \tag{I.16}$$

The separately estimated organic-search composition shift generated by ζ_A and ζ_D is then restored to obtain the full post-privacy state. Thus ω_g governs the destination of lost paid contacts and does not replace the broader organic-comparison fit. No Tech receives no direct paid-contact shock; only the common organic-search environment affects its channel mix.

For observed Searchers, the model-implied mortgage rate is the resulting channel-weighted average

$$R_{gbu}(Z=1) = w_{gbu}^s r_{gbu}^s + w_{gbu}^p r_{bu}^{\mathrm{ad}} + w_{gbu}^h r_{bu}^{\mathrm{home}}, \qquad w_{Nbu}^p = 0. \tag{I.17}$$

The implementation therefore evaluates four states in sequence: pre-privacy → post-screening → post-paid-retention → full post-privacy with organic reweighting and rerouting.

I.2 Identification, Estimation, Inference, and Diagnostics

This appendix documents the empirical mapping for the quantitative economy in Appendix I.1. It follows the identification sequence summarized in Section 10 and distinguishes direct inputs, external calibrations, pre-enforcement estimates, profiled parameters, targeted causal moments, and held-out validation moments. Table I.1 summarizes the full sequence before the details below.

Table I.1. Identification and Calibration Map

Model object	Empirical discipline	Status
Screening states A_{gbu}, ρ_{gbu}	Pre-law approval and bankruptcy risk; post-law approval and bankruptcy ATTs	Empirical input
κ_ρ	Pre-law Searcher mortgage rates against funded-pool risk, with score-bin fixed effects	Estimated
Γ, P_{b0}	Nested No-Tech + Only-Ads pre-law composition; No-Tech-home contacts allocated across A and D using $\phi_{A,b}$	Calibrated / constructed
δ_A, δ_D	Vivvix advertising-retention estimates	External calibration
$\tau_A, \tau_D, \zeta_A, \zeta_D$	Pre-law low-K composition and post-law search-propensity changes; τ_D fixed at 0.30	Estimated / preset
K_ℓ	Profile grid; admissible rows must fit both technology rate ATTs and satisfy channel ordering	Profiled
$f_b, K_h \mid K_\ell$	Pre-law Searcher and non-Searcher mortgage-rate levels	Estimated
ω_A, ω_D	Only-Ads and Ads+Data Searcher-rate ATTs	Calibrated / inverted
No-Tech Searcher-rate ATT	Never targeted in calibration	Held-out validation

Notes: The estimation is sequential. "Estimated" refers to pre-law or composition moments; "calibrated / inverted" refers to parameters chosen to match the stated causal policy moment. The No-Tech rate effect is reserved for validation.

Empirical Moment Ledger

Table I.2 records the reduced-form moments that enter the quantitative exercise and separates inputs, targeted moments, and held-out validation. Table I.3 then reports the score-bin primitives and the data construction used to map empirical Searcher status and lender availability into the model. Together, these tables make the data–model mapping explicit before the sequential estimation steps below.

Screening States

The screening block starts from pre-enforcement approval rates and funded-pool risk by lender group and score bucket. Approval and bankruptcy treatment effects map these baseline objects into post-enforcement screening states for Only Ads and Ads+Data, while the structural No-Tech

Table I.2. Empirical Moments Used in the Quantitative Model

	No Tech	Only Ads	Ads + Data	Source	Role
Approval ATT (pp)	–	0.38 (0.13)	0.21 (0.08)	HMDA	Screening input
One-year bankruptcy ATT (pp)	–	2.09 (0.28)	0.39 (0.46)	Equifax–HMDA	Screening input
Mean search propensity ATT (pp)	–	-0.85 (0.11)	-0.49 (0.19)	Equifax–HMDA	Composition target
Searcher mortgage-rate ATT (bp)	4.81 (4.81)	-13.95 (2.50)	-11.65 (1.77)	Equifax–HMDA	Validation / target
Paid-ad retention after privacy	–	0.240	0.645	Vivvix	External calibration

Notes: Standard errors are in parentheses. Approval and bankruptcy effects shift the screening state. Mean search-propensity effects discipline post-law composition. Only Ads and Ads+Data Searcher-rate effects identify ω_A and ω_D; the No-Tech effect is held out. Paid-ad retention is calibrated independently from advertising expenditures. The approval moments are estimated on the score-imputed HMDA application sample used to form lender-group × score-bin structural cells, and therefore need not coincide numerically with the pooled rejection estimates reported in the reduced-form section. For consistency with the headline reduced-form table, the held-out No-Tech validation moment is reported as 4.81 bp (s.e. 4.81). No-Tech is never targeted in calibration, so this reporting harmonization does not alter any calibrated parameter or counterfactual.

Table I.3. Quantitative Inputs and Data Construction

VantageScore bin	Risky share $\mu_{R,b}$	High-K share λ_b	$\eta_{L,b}$	$\eta_{H,b}$
Panel A. Score-bin primitives				
< 670	0.164	0.761	0.876	0.788
670-709	0.116	0.759	0.878	0.786
710-749	0.091	0.756	0.882	0.788
750-789	0.049	0.753	0.875	0.768
790-829	0.029	0.763	0.824	0.687
> 829	0.023	0.781	0.734	0.577
Panel B. Empirical construction				
Pre-law window	2018–2019			
Searcher indicator	$Z_i = 1\{\widehat{\pi}_i^P \geq 0.19\}$			
Quote-draw weights s^q_{gb0}	Borrower-weighted county × month active-lender-type shares			
Home-bank weights h_{g0}	Lender-group shares of 2018 branch deposits			
Risky-population share $\mu_{R,b}$	Midpoint of the admissible interval implied by approval and funded-pool risk			

Notes: $\lambda_b = \Pr(K = K_h \mid b)$, so $1 - \lambda_b$ is the latent low-K share. $\eta_{L,b} = \Pr(Z = 1 \mid K_\ell, b)$ and $\eta_{H,b} = \Pr(Z = 1 \mid K_h, b)$ map latent shopping types into the empirical Searcher indicator Z. The six bins are < 670, 670–709, 710–749, 750–789, 790–829, and > 829. The empirical Searcher split uses $\widehat{\pi}_i^P \geq 0.19$. The K-type distribution and measurement map are pooled across lender groups within each score bin. Quote-draw weights are borrower-weighted shares of active lender types in county-by-month markets; home-bank weights are lender-group shares of 2018 branch deposits.

screening state is held fixed. Given the maintained within-bin risky-population share, the implied safe- and risky-type approval probabilities are recovered subject to probability bounds. Specifically, for each lender group, score bin, and policy state,

$$q_{gR,bu} = \frac{A_{gbu}\rho_{gbu}}{\mu_{R,b}}, \qquad q_{gS,bu} = \frac{A_{gbu}(1-\rho_{gbu})}{1-\mu_{R,b}}. \tag{I.18}$$

The maintained $\mu_{R,b}$ values are reported in Table I.3. These objects feed directly into the marginal-cost and comparison-rate block in Section 10.

Risk Pass-Through, Paid Reach, and Composition

The quantitative rate equation treats κ_ρ as an empirical loading of funded-pool risk into comparison quotes. I estimate it from pre-enforcement covariation between Searcher mortgage rates and funded-pool risk across lender-group-by-score-bin cells, with score-bin fixed effects and the cell weights used in the structural moment construction. No privacy-law mortgage-rate effect enters this step. The risk measure ρ_{gbu} is the observed one-year-ahead bankruptcy rate used as a funded-pool risk proxy. Accordingly, $r^s_{gbu} = f_b + \kappa_\rho d\rho_{gbu}$ is an empirical measurement equation for comparison quotes rather than the analytical zero-profit restriction $r^s = \text{MC}$. Resource-cost accounting separately uses $\text{MC}_{gbu} = f_b + d\rho_{gbu}$; the difference between these two measurement equations is therefore not interpreted as a current-loan search-channel margin.

I calibrate total pre-law paid reach Γ in the nested No-Tech + Only-Ads economy using pre-law latent low-K composition. In the three-lender economy, No-Tech-home contacts are allocated between Only Ads and Ads+Data according to $\phi_{A,b}$ in equation (I.2). The frozen V5-MC3.1 specification sets the cross-technology scaling parameter ξ to one. Advertising-retention rates $\delta_A = 0.240$ and $\delta_D = 0.645$ come from the independent Vivvix advertising-spending evidence.

Conditional on $\tau_D = 0.30$, the implementation estimates τ_A, ζ_A, and ζ_D from pre-law low-K composition and privacy-induced changes in shopping propensity. The τ parameters govern selection into successful paid contact, whereas the ζ parameters reweight the post-privacy organic search environment according to equation (I.8).

Comparison Wedges and Operating Costs

I profile the low-friction comparison wedge K_ℓ. Conditional on each profile value, I recover the score-bin operating-cost terms f_b and the common high-friction wedge K_h from pre-enforcement Searcher and non-Searcher mortgage-rate levels. For every profile value, the remaining blocks are recomputed and the quantitative economy is required to satisfy probability, adding-up, feasibility, and channel-ordering restrictions. The selected specification minimizes pre-enforcement rate-level

RMSE within the admissible region. The corresponding profile and supporting fit diagnostics are reported in the existing quantitative-results appendix, Appendix J.1.

Rerouting Parameters and Targeted Moments

Conditional on the preceding blocks, ω_A and ω_D solve

$$\Delta r^{\text{model}}_{g,Z=1}(\omega_g) = \widehat{\text{ATT}}^{\text{rate}}_{g,Z=1}, \qquad g \in \{A, D\}. \tag{I.19}$$

These two effects are therefore targeted moments rather than goodness-of-fit tests. For each technology group, I verify whether the mapping from ω_g to the model-implied rate effect is monotone on $[0, 1]$, whether the solution is unique, and whether the optimum lies on a boundary. The No-Tech Searcher-rate effect is not used in this inversion and is reserved for the held-out validation exercise reported in Section 11.

Estimation Algorithm

The sequential routine is:

a. construct six VantageScore-bin-by-lender-group empirical cells from pre-law approval, bankruptcy, shopping propensity, home-lender shares, local search-market shares, and Searcher/non-Searcher mortgage rates;

b. apply the approval and bankruptcy treatment effects to Only Ads and Ads+Data and recover pre- and post-enforcement screening states, while holding the structural No-Tech screening state fixed;

c. estimate κ_ρ using pre-law rate–risk covariation only;

d. calibrate total pre-law paid reach Γ in the nested No-Tech + Only-Ads economy and construct the three-lender contact matrix in equation (I.3);

e. fix $\tau_D = 0.30$ and estimate τ_A, ζ_A, and ζ_D from pre-law low-K composition and post-law shopping-propensity changes;

f. for each K_ℓ on the profile grid, recover the six f_b terms and the common K_h by weighted least squares from pre-law Searcher and non-Searcher mortgage-rate levels;

g. for each admissible profile row, solve ω_A and ω_D by one-dimensional inversion to reproduce the two technology-group Searcher-rate ATTs;

h. retain rows satisfying probability, nonnegativity, adding-up, and strict channel-rate ordering restrictions and select the exact-fit row with the lowest pre-law rate-level RMSE;

i. invert the empirical rate-ATT endpoints through the frozen structural block to obtain the conditional ω_g intervals reported below;

j. hold all equilibrium objects fixed and vary only the resource share used in surplus accounting $\theta \in [0, 1]$; and

k. for the targeting-only policy, restore pre-privacy screening and organic-search shares, set $P_{ht,b} = 0$, and reroute all displaced technology paid mass using the calibrated ω_g.

The two technology Searcher-rate ATTs are targeted moments. The No-Tech rate ATT, the detailed pre-law rate residuals, and the channel-ordering restrictions are validation and diagnostic objects. The resulting estimates are summarized in Table I.4 and feed the quantitative results in Section 11.

Inference and Numerical Diagnostics

The current implementation reports conditional ATT-implied ranges for ω_A and ω_D. Holding the remaining structural block fixed, I invert the empirical Searcher-rate ATT interval through equation (I.19). These ranges are therefore not joint structural confidence regions. For Ads+Data, the upper endpoint reaches the economic support restriction $\omega_D \leq 1$, so the reported upper endpoint is constraint-limited rather than an interior data-determined bound.

Table I.5 reports the corresponding numerical and identification diagnostics, including probability bounds, adding-up errors, targeted-moment residuals, channel ordering, and proximity to parameter bounds. Supporting channel-rate, profile, and pre-enforcement fit results are reported in Appendix J.1.

J Appendix – Quantitative Results

J.1 Model Fit and Mechanism Decomposition

This appendix reports model-implied channel prices, the profile over the comparison friction, and pre-enforcement rate-level fit. The first two tables show the equilibrium objects behind the mechanism decomposition; the final fit table makes transparent which lender-group-by-score-by-search-status cells the parsimonious pricing block matches less closely.

Pre-Enforcement Rate-Level Fit

Table J.3 compares the pre-enforcement mortgage-rate levels used to recover f_b and K_h with their model counterparts. The weighted RMSE is 11.02 basis points and the maximum absolute residual is 24.46 basis points. The misses are systematic rather than hidden: the model understates No-Tech Searcher rates and overstates Ads+Data non-Searcher rates, reflecting the discipline imposed by a common K_h and a common score-bin operating-cost component across lender types.

Table I.4. Calibrated and Estimated Parameters

Parameter	Estimate	Status	Discipline / interpretation
Panel A. Common parameters			
d	0.150	External	Loss given default
κ_ρ	0.380	Estimated	Pre-law rate–risk WLS
Γ	0.557	Calibrated	Total pre-law paid-contact reach
ξ	1.000	Fixed	Cross-technology paid-contact scaling in frozen V5-MC3.1 specification
τ_A	0.838	Estimated	High-K contact fraction, Only Ads
τ_D	0.300	Preset	High-K contact fraction, Ads+Data
ζ_A	0.870	Estimated	Post-law organic-search reweighting, Only Ads
ζ_D	1.195	Estimated	Post-law organic-search reweighting, Ads+Data
δ_A	0.240	External	Post-law paid-ad retention, Only Ads
δ_D	0.645	External	Post-law paid-ad retention, Ads+Data
K_ℓ	125.0 bp	Profiled	Low-friction comparison wedge
K_h	164.8 bp	Estimated	High-friction comparison wedge
ω_A	0.634 [0.550,0.718]	ATT inversion	Fallback to organic search after lost paid contact
ω_D	0.996 [0.766,1.000]	ATT inversion	Fallback to organic search; upper CI hits support bound
Panel B. Score-specific operating-cost component			
f_b: < 670	2.792%	Estimated	Pre-law Searcher and non-Searcher rate levels
f_b: 670-709	2.922%	Estimated	Pre-law Searcher and non-Searcher rate levels
f_b: 710-749	2.969%	Estimated	Pre-law Searcher and non-Searcher rate levels
f_b: 750-789	3.110%	Estimated	Pre-law Searcher and non-Searcher rate levels
f_b: 790-829	3.161%	Estimated	Pre-law Searcher and non-Searcher rate levels
f_b: > 829	3.153%	Estimated	Pre-law Searcher and non-Searcher rate levels

Notes: Brackets for ω_g are conditional 95% ATT-implied intervals obtained by inverting the empirical rate-ATT interval while holding the remaining structural block fixed; they are not joint structural confidence regions. The Ads+Data upper endpoint is support-limited at one. The ζ_g objects are composition reweighting multipliers, not probabilities, so values above one are admissible.

Table I.5. Estimation and Numerical Diagnostics

Diagnostic	Value	Interpretation
Pre-law weighted rate RMSE	11.02 bp	Level fit
Maximum absolute pre-law rate residual	24.46 bp	Level fit
Held-out No-Tech Searcher-rate ATT	4.25 vs. 4.81 (0.12 SE gap)	Not targeted
Only-Ads ATT residual	$< 10^{-12}$ bp	Exact targeted fit
Ads+Data ATT residual	$< 10^{-12}$ bp	Exact targeted fit
Market-share adding-up error	1.11e-16	Pass
Acquisition-channel adding-up error	1.39e-14 pp	Pass
Minimum home-minus-paid rate gap	39.77 bp	Strict ordering
Minimum paid-minus-search rate gap	115.18 bp	Strict ordering
ω_A conditional 95% interval	[0.550, 0.718]	Interior
ω_D conditional 95% interval	[0.766, 1.000]	Upper endpoint hits $\omega_D = 1$ support bound
Γ first-stage weighted RMSE	0.0206	Low-K share

Notes: The ω_g intervals invert the empirical Searcher-rate ATT intervals with the structural block frozen. For Ads+Data, the empirical lower ATT endpoint would imply $\omega_D > 1$, so the reported upper endpoint is censored by the economic support restriction. These intervals are conditional ATT-implied ranges, not joint structural confidence regions. The held-out No-Tech comparison uses the 4.81 bp headline reduced-form estimate (s.e. 4.81); the model prediction of 4.25 bp is not used to calibrate any parameter.

Table J.1. Mortgage Rates by Lender Type and Acquisition Channel

	Organic Search		Paid Advertising		Home / Incumbent	
Lender type	Pre	Post	Pre	Post	Pre	Post
No Tech	3.244	3.244	—	—	4.876	4.904
Only Ads	3.227	3.344	4.477	4.504	4.877	4.904
Ads + Data	3.231	3.251	4.474	4.502	4.875	4.903

Notes: This table reports model-implied mortgage rates before and after privacy-law enforcement by lender technology and borrower acquisition channel. Rates are in percentage points and are aggregated across VantageScore bins using the model's rate weights. Organic Search denotes borrowers who actively compare offers, Paid Advertising denotes borrowers reached through paid acquisition, and Home / Incumbent denotes borrowers remaining with their incumbent lender. No-Tech lenders do not operate a paid-advertising acquisition channel in the model. For every technology-lender-by-credit-score cell, both before and after enforcement, the model satisfies $r_{\text{search}} < r_{\text{ad}} < r_{\text{home}}$; for No-Tech lenders, $r_{\text{search}} < r_{\text{home}}$.

Table J.2. Profile over the Comparison Friction

K_ℓ	K_h (bp)	Pre-law RMSE (bp)	ω_A	ω_D	Min. search rate (%)	Exact fit	Rate ordering
100	132.6	9.06	0.720	—	—	No	—
125	164.8	11.02	0.634	0.996	3.16	Yes	Yes
150	197.0	13.02	0.577	0.839	2.90	Yes	Yes
175	229.2	15.04	0.537	0.725	2.64	Yes	Yes
200	261.4	17.07	0.508	0.640	2.39	Yes	Yes

Notes: This table profiles the model over the common low-search-cost comparison friction K_ℓ. Conditional on each value of K_ℓ, K_h and the credit-score-specific cost components are recovered from pre-enforcement Searcher and non-Searcher mortgage-rate levels. The columns ω_A and ω_D report the search-fallback shares required to reproduce the headline Searcher mortgage-rate treatment effects for Only Ads and Ads + Data, respectively. "Exact fit" indicates that both technology-group treatment effects can be reproduced with admissible $\omega_g \in [0, 1]$. "Rate ordering" indicates whether $r_{\text{search}} < r_{\text{ad}} < r_{\text{home}}$ holds in every applicable technology-lender-by-credit-score cell before and after enforcement. The baseline, $K_\ell = 125$ basis points, is the lowest value for which both technology-group treatment effects can be jointly rationalized and therefore minimizes the pre-enforcement rate-level RMSE within the exact-fit region. Goodness-of-fit statistics for this baseline are discussed in Section J.1, and the corresponding pooled channel-specific mortgage rates are reported in Table J.1.

Table J.3. Pre-Law Mortgage-Rate Fit by Score and Search Status

	No Tech			Only Ads			Ads + Data		
VantageScore	Data	Model	Gap	Data	Model	Gap	Data	Model	Gap
	percent		bp	percent		bp	percent		bp
Panel A. Searchers									
< 670	4.570	4.389	-18.2	4.555	4.576	2.1	4.535	4.602	6.7
670-709	4.557	4.359	-19.7	4.530	4.545	1.5	4.517	4.572	5.5
710-749	4.547	4.334	-21.3	4.508	4.520	1.2	4.489	4.544	5.4
750-789	4.523	4.294	-23.0	4.479	4.492	1.4	4.453	4.510	5.7
790-829	4.506	4.261	-24.5	4.452	4.460	0.9	4.423	4.479	5.6
> 829	4.483	4.251	-23.2	4.430	4.440	1.0	4.406	4.462	5.6
Panel B. Non-searchers									
< 670	4.650	4.596	-5.4	4.630	4.635	0.4	4.578	4.726	14.8
670-709	4.617	4.573	-4.4	4.576	4.601	2.5	4.553	4.698	14.5
710-749	4.578	4.555	-2.3	4.556	4.576	2.1	4.530	4.673	14.4
750-789	4.552	4.533	-1.9	4.525	4.551	2.5	4.502	4.649	14.6
790-829	4.511	4.503	-0.7	4.492	4.518	2.6	4.465	4.617	15.2
> 829	4.504	4.475	-2.9	4.456	4.492	3.6	4.437	4.590	15.3

Notes: Gap is model minus data in basis points. The weighted RMSE across pre-law rate-level cells is 11.02 bp and the maximum absolute residual is 24.46 bp. These level moments identify f_b and K_h conditional on the profiled K_ℓ. The two systematic misses are transparent in the table: the model understates No-Tech Searcher rates by roughly 18–24 bp and overstates Ads+Data non-Searcher rates by roughly 14–15 bp.

J.2 Reallocation and Surplus Effects of Privacy Laws

The quantitative exercise uses the estimated model to separate extensive-margin credit access from intensive-margin incidence among originated loans. The figures below report credit access, acquisition-channel reallocation, and the borrower-surplus decomposition by score. Table J.4 then consolidates the aggregate incidence, surplus accounting, and the targeting-only policy counterfactual.

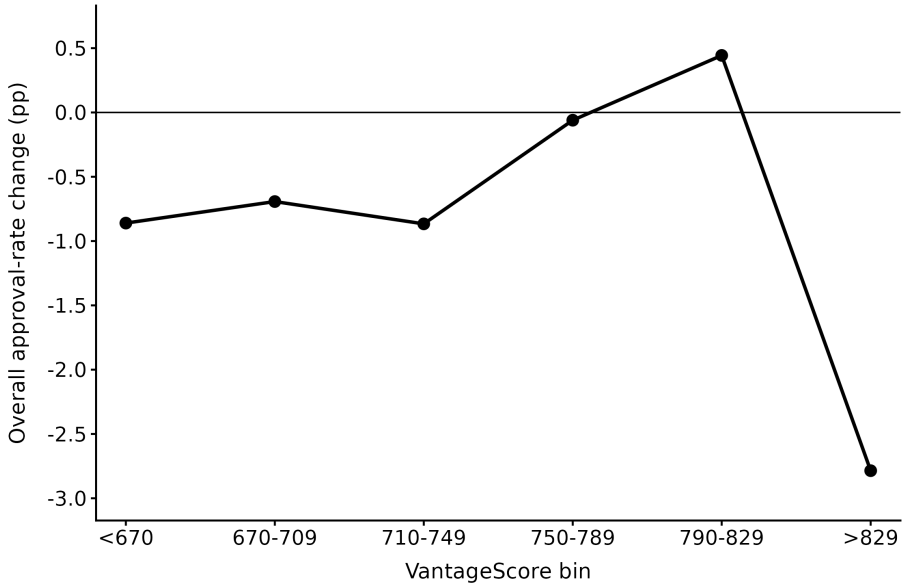


Figure J.1. Credit Access by Credit Score

Notes: This figure reports the overall change in mortgage approval probabilities following privacy-law enforcement across VantageScore bins. Changes are measured in percentage points. These approval effects capture the extensive margin of credit access and are reported separately from the surplus analysis conditional on origination.

Surplus Decomposition and Use-Based Counterfactual

Let ΔS_ℓ and ΔS_h denote changes in active organic comparison by latent low- and high-friction borrowers. Borrower choice responds to the full comparison wedge K_j, whereas surplus accounting treats only the resource component, θK_j, as a real cost. The change in resource-intensive search cost is therefore

$$\Delta C^{\text{search}}(\theta) = \theta \left(K_\ell \Delta S_\ell + K_h \Delta S_h\right). \tag{J.1}$$

Let G^{acq} denote the borrower rate gain associated with the acquisition-channel change and let $\Delta CS^{\text{screen}}$ denote the borrower effect of weaker screening at fixed channel allocation. Borrower surplus can then be written as

$$\Delta CS(\theta) = \Delta CS^{\text{screen}} + (1-\theta) G^{\text{acq}} + \left[\theta G^{\text{acq}} - \Delta C^{\text{search}}(\theta)\right]. \tag{J.2}$$

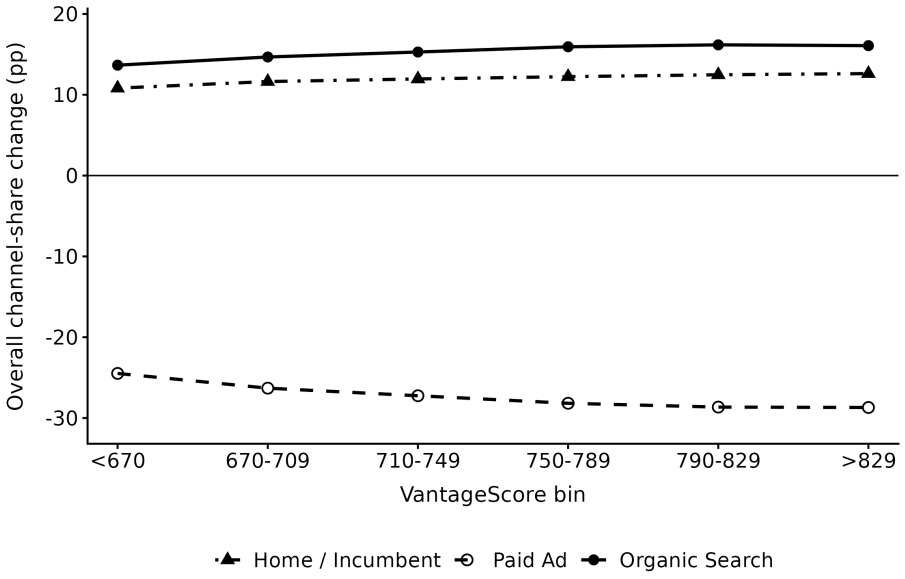


Figure J.2. Acquisition-Channel Reallocation by Credit Score

Notes: This figure reports the model-implied change in acquisition-channel shares following privacy-law enforcement across VantageScore bins. Changes are measured in percentage points. Privacy reduces originations through paid advertising and reallocates borrowers toward organic search and the home/incumbent channel.

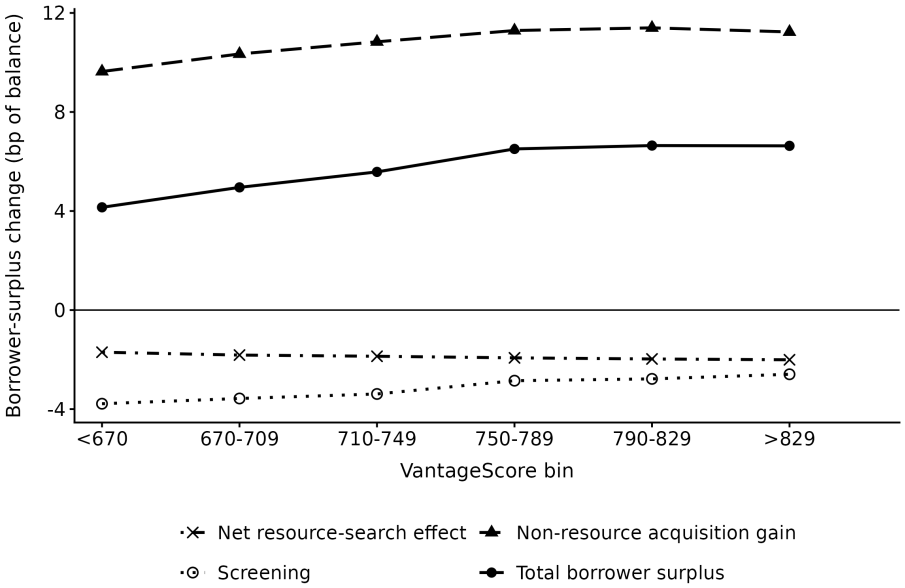


Figure J.3. Borrower-Surplus Decomposition by Credit Score

Notes: This figure decomposes the model-implied change in borrower surplus following privacy-law enforcement across VantageScore bins. Values are expressed in basis points of mortgage balance. The decomposition separates the effect of weaker screening, the non-resource gain from acquisition-channel reallocation, and the net resource cost of additional search. Total borrower surplus is the sum of these components.

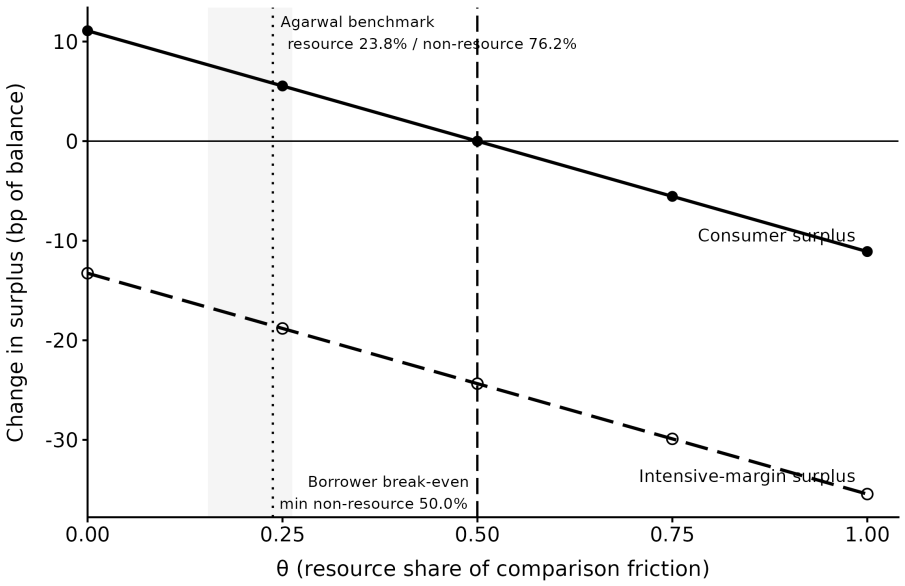


Figure J.4. Surplus Effects by the Resource Share of Comparison Frictions

Notes: This figure reports borrower surplus and model-implied intensive-margin surplus for alternative values of θ, the share of the estimated comparison friction that represents a real resource cost rather than a non-resource stopping friction. The shaded region denotes the benchmark range implied by Agarwal et al. (2024), and the dotted vertical line marks the baseline value, $\theta = 0.238$. The dashed vertical line at $\theta = 0.50$ marks the borrower-surplus break-even point. Prices, quantities, and rerouting parameters are held fixed as θ varies. Borrower surplus is positive when the non-resource component of comparison frictions is sufficiently large, whereas model-implied intensive-margin surplus remains negative throughout.

The second term is the transfer back to borrowers associated with the non-resource portion of the stopping wedge; the bracketed term is the net effect of the resource component once borrowers resume active comparison.

For efficiency, let E^{screen} denote the real-lending-cost effect of weaker screening at fixed channel allocation and let E^{realloc} denote the real-cost effect of reallocating originations across lender types. Model-implied intensive-margin surplus is

$$\Delta W^{\text{int}}(\theta) = E^{\text{screen}} + E^{\text{realloc}} - \Delta C^{\text{search}}(\theta). \tag{J.3}$$

Mortgage-rate changes themselves are transfers between borrowers and lenders. The calculation is conditional on origination and does not monetize the value of marginal credit access, continuation or franchise value, platform payments, advertising expenditures, entry, or long-run technology adoption.

At the literature benchmark $\theta^{\text{lit}} = 29.7/125 = 0.2376$, broad privacy raises borrower surplus by 5.82 basis points of mortgage balance but lowers model-implied intensive-margin surplus by 18.54 basis points. The targeting-only counterfactual preserves pre-privacy underwriting information and the pre-privacy organic-search environment while setting the paid-contact matrix to zero. It therefore isolates the acquisition channel without imposing the screening loss generated by broad privacy. All remaining structural and price parameters, including the fallback shares ω_g, are held fixed; the exercise is partial equilibrium.

Let $T0$ denote the targeting-only state and 0 the pre-privacy baseline. With mortgage rates expressed in percentage points, the corresponding borrower and intensive-margin surplus changes, in basis points of mortgage balance, are

$$\Delta CS^{T0}(\theta) = -100\left(\overline{R}_{T0} - \overline{R}_0\right) - \theta\left[K_\ell(S_\ell^{T0} - S_\ell^0) + K_h(S_h^{T0} - S_h^0)\right], \tag{J.4}$$

$$\Delta W^{\text{int},T0}(\theta) = -100\left(\overline{MC}_{T0} - \overline{MC}_0\right) - \theta\left[K_\ell(S_\ell^{T0} - S_\ell^0) + K_h(S_h^{T0} - S_h^0)\right]. \tag{J.5}$$

At θ^{lit}, these equations deliver the 16.02 bp borrower gain and the −8.90 bp intensive-margin surplus change reported in Table J.4.

Table J.4. Surplus Accounting and Use-Based Counterfactual

	Consumers	No Tech	Only Ads	Ads + Data	Intensive-margin surplus
		basis points of mortgage balance			
Panel A. Broad privacy: incidence at $\theta^{\text{lit}} = 0.2376$					
Total change	5.82	1.57	-34.89	8.97	-18.54
Panel B. Broad privacy: decomposition					
Consumer: weaker screening	-3.16	–	–	–	–
Consumer: non-resource acquisition gain	10.86	–	–	–	–
Consumer: net resource-search effect	-1.88	–	–	–	–
Efficiency: screening / real lending cost	–	–	–	–	-18.60
Efficiency: lender-type reallocation	–	–	–	–	5.33
Efficiency: active-search resource cost	–	–	–	–	-5.27
Panel C. Policy comparison (aggregate)					
Policy	Consumers	Aggregate lender margins			Intensive-margin surplus
Broad privacy	5.82	-24.35			-18.54
Targeting-only prohibition	16.02	-24.91			-8.90

Notes: $\theta^{\text{lit}} = 29.7/125 = 0.2376$ is the literature benchmark for the resource share of the effective comparison wedge; the implied non-resource share is 76.2%. Broad privacy uses post-law screening and paid-ad retention of 0.24 for Only Ads and 0.645 for Ads+Data. The targeting-only counterfactual preserves pre-law screening and organic-search shares, sets the full paid cross-lender contact matrix to zero, and reroutes displaced paid borrowers using the calibrated ω_g. Intensive-margin surplus is borrower surplus plus current-loan lender margins and is conditional on origination. The counterfactual is partial equilibrium: lender types, entry, long-run technology adoption, and market-wide funding conditions are held fixed; credit-access value and continuation value are not monetized.