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

Federico Lenzi†

9 August 2026

Latest version available **here**

Abstract

Search frictions induce lenders to compete for borrower attention as well as on price. Online consumer data supports both margins: it improves screening and enables targeted advertising before borrowers obtain competing quotes. The two uses push rates in opposite directions—better screening mitigates adverse selection and lowers rates, whereas better targeting lets lenders price against borrowers' reservation rates. I separate the channels using the staggered enforcement of U.S. state privacy laws, combining linked Equifax–HMDA originations from 2018–2024 with lender advertising expenditures and search propensities predicted from a national mortgage-shopping survey. Comparing borrowers within the same lender, effects appear only where likely searchers face data-reliant lenders. For these borrowers, rates fall by 14 basis points, and data-reliant lenders become 0.85 percentage points less likely to attract them, a 3.4 percent decline relative to the mean. Rejection rates fall by 0.22 percentage points, and search-advertising expenditures decline by roughly 70 percent among the most exposed lenders. Effects are near zero for other borrowers and lenders. I develop a search model in which consumer data both sharpens screening and supports targeted advertising; because the channels move rates in opposite directions, the observed decline implies that targeting dominates screening. The model further characterizes how privacy restrictions redistribute surplus between borrowers and lenders. As mortgage search moves online, data-use rules become competition rules.

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

Mortgage lenders cannot compete on price unless borrowers consider them. Competing for that consideration has become more important as mortgage shopping moves online: 71% of recent homebuyers report preferring to select a lender online.[1] Online-collected consumer data tells a lender two things at once: who is likely to repay, and who is about to shop. The first can improve screening by raising rejection rates and reducing interest rates for approved borrowers. The second can raise rates by helping lenders reach borrowers before they obtain competing quotes and preserve a markup because continued search is costly. Restricting online-collected data can therefore move mortgage rates in either direction. I separate these forces by exploiting the staggered enforcement of state privacy laws. Privacy regulation governs not only what lenders know about borrowers, but also which lenders borrowers consider. It therefore shapes competition as well as information.

The two channels operate on different decisions: screening determines which applicants receive credit and on what terms, whereas advertising shapes which lenders borrowers consider during search. If online consumer data is used only for screening, privacy-law enforcement weakens lenders' ability to distinguish risk among borrowers with similar credit scores. Rejection rates should fall, the originated pool should become riskier, and both subsequent default and rates on originated loans should rise. If the data also supports borrower acquisition, enforcement generates an additional set of predictions. Under limited attention and costly search, borrowers place greater weight on lenders encountered early (Huang et al. 2026; Günsberg and Kuhnen 2026). Targeted advertising can therefore place a lender in the borrower's consideration set before competing quotes are obtained. Because obtaining additional quotes is costly, some borrowers accept an early offer even when further search could yield a lower rate. Targeted advertising thus allows lenders to price closer to borrowers' willingness to pay than to the rate they could obtain by continuing to search. Restricting access to the underlying data should reduce search advertising, lower data-reliant lenders' share of high-search-propensity borrowers, and reduce the rates those borrowers pay. Search propensity therefore provides the key cross-sectional margin for separating the channels. Identification comes from their joint empirical signature across rejection, default, advertising expenditure, borrower composition, and pricing, rather than from any single outcome.

To separate these two forces, I exploit the staggered enforcement of U.S. state privacy laws. Beginning with the California Consumer Privacy Act, privacy regulation became salient for lenders: other states introduced or enacted similar laws, and a federal privacy regulation was under active debate (Dubé et al. 2025). I classify lenders by their use of online-data technologies before these laws took effect and focus on lenders that had already adopted them. This pre-enforcement classification is

[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)

important because lenders operate a common website infrastructure across states. Figure A.3 shows that, once privacy regulation became salient, new installations of relevant ad-tech and data-collection tools slowed sharply. I measure lender type using pre-period data. Enforcement does not remove these technologies from lenders' websites. It restricts the data they may use for borrowers in protected states. I compare observationally similar borrowers applying to the same lender at the same time and for similar properties, who differ only in whether privacy law limits the use of their online consumer data. 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 borrower-level variation in data availability, identifying the screening and advertising responses to restricting it.

Search behavior is not observed in the administrative mortgage data. In the National Survey of Mortgage Originations, 55 percent of borrowers report considering more than one lender; I estimate the relationship between reported shopping and predetermined borrower demographics and income in the survey, then project it onto the loan-level sample to obtain each borrower's predicted search propensity. Among lenders using advertising technologies, privacy-law enforcement lowers mortgage rates for borrowers with high predicted search propensity by 14 basis points, or 3.3 percent of the mean rate, while the corresponding estimate for low-propensity borrowers is small and statistically indistinguishable from zero. This heterogeneity is difficult to reconcile with screening alone: weaker screening worsens the risk composition of approved loans and should push rates up, and it operates on approved borrowers regardless of whether they shop. The decline concentrated among likely shoppers is instead consistent with privacy laws weakening lenders' ability to reach these borrowers before they obtain competing quotes.

A change in borrower allocation accompanies the pricing effect. Following enforcement, the share of high-search-propensity borrowers in advertising-technology lenders' originations falls by 0.85 percentage points, or 3.4% relative to its baseline mean. Targeted advertising lets lenders enter shoppers' consideration sets before search is complete and price against the value of continued search rather than the best available quote. Privacy laws weaken this advantage: data-using lenders attract fewer likely shoppers, and the likely shoppers they continue to originate receive lower rates. The joint decline in rates and borrower share is difficult to explain through screening alone but follows naturally when online-collected data also supports targeted borrower acquisition.

The empirical design faces two principal identification challenges. First, enforcement timing could track local credit-market trends. I address this concern in three ways. Privacy-law adoption is better predicted by political preferences and exposure to the technology sector than by borrower credit conditions. Pre-trends are flat around both signature and enforcement dates, limiting concerns about anticipation. I also compare borrowers within lender and within state-border-crossing metropolitan areas, where borrowers face similar local housing and credit-market conditions but differ in whether

their state enforces a privacy law. Second, the laws could affect mortgage outcomes through channels other than lender information, such as broader policy bundles or COVID-era disruptions to mortgage markets. The timing makes this concern less likely: California is the only enforcing state during the pandemic period, and its effect begins before the onset of COVID and persists when the sample is extended through 2024. 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.

Two pieces of direct evidence show that online consumer data supported both screening and borrower acquisition before enforcement. First, rejection rates at data-using lenders fall by 0.22 percentage points after enforcement, or 3.3 percent of the mean, while the one-year incidence of financial distress among their originated borrowers rises by 2.09 percentage points from a 2.7 percent baseline. These patterns are consistent with weaker screening: after losing the signal, lenders distinguish risk less precisely within credit-score buckets and approve applicants they would previously have rejected. Second, lenders that rely on data collected by third-party websites sharply reduce search-engine advertising as privacy-law coverage expands across their customer base, with expenditures falling by approximately 70 percent among the most exposed lenders. The rate effect identifies the dominant pricing force: among borrowers who search, rates fall by 14 basis points, implying that the loss of pricing power from weaker targeting more than offsets the higher expected losses from weaker screening.

The reduced-form estimates identify the dominant pricing channel but cannot recover the value of data in screening and borrower acquisition or its incidence between borrowers and lenders. I therefore develop a model that combines mortgage search and approval, as in Agarwal et al. (2024), with advertising-induced awareness and endogenous consideration sets, as in Honka et al. (2017). Borrowers differ in repayment risk and search costs, while lenders differ in their use of online-data technologies. Lenders reach borrowers through incumbent relationships, targeted advertising, or organic search. Incumbent relationships and high-search-cost borrowers sustain demand for lenders without targeting technologies, allowing both lender types to coexist in equilibrium, while competition between them is concentrated among borrowers who search. Online consumer data enters through two primitives: screening precision and targeting effectiveness. The model maps privacy-induced changes in these primitives into loan pricing, borrower allocation, and surplus, allowing me to decompose the value of data into screening and acquisition components and its incidence between borrowers and lenders.

Estimating the model allows me to quantify the economic value of online-collected data in screening and in targeted borrower acquisition. It also shows how privacy laws, although not designed as credit-market policy, redistribute surplus between borrowers and lenders. The welfare effect is ambiguous ex ante. Restricting online-collected data can reduce surplus by weakening

screening precision, but it can increase borrower surplus by reducing lenders' ability to preserve margins through targeted advertising. Weaker screening extends credit to borrowers who would otherwise be denied, while borrowers approved either way gain only if the rate decline from weaker targeting exceeds the rate increase from weaker screening.

One measurement issue deserves note. To measure borrower search, I use self-reported search behavior in the NSMO and predict the probability that a borrower searches for a better price using demographics, income, and loan amount. I then impute this search propensity to the loan-level sample and carry first-stage uncertainty into the second-stage inference. The main concern is that the predicted measure classifies some borrowers incorrectly. Such noise should make high- and low-search borrowers look more similar, not more different. In other words, measurement error in search propensity works against finding a sharp difference between searchers and non-searchers, unless prediction error is systematically related to privacy-law enforcement. The estimated heterogeneity is therefore unlikely to be mechanically generated by the imputation procedure.

First, the paper contributes to the literature on mortgage search. One strand 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 affects both borrowers' search decisions and selection into observed transactions. 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 targeted 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 examine online acquisition, where lender reach is less tied to local branch networks. Search frictions nevertheless persist 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 expected rate from continued 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 competing offers the borrower actually obtains. What is new is that effective search—its sequence and stopping point—is partly shaped by lenders themselves. This mechanism provides a common interpretation of 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 (Section 8.4). Relative to work emphasizing screening, processing efficiency, and customer service (Berg et al. 2022; Fuster et al. 2019), I therefore identify customer acquisition as an additional use of financial technology: 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 constitutes an important intangible investment in industries with severe information frictions (He et al. 2024, 2026). Online data raise the return to that investment: they are both productive, by improving screening, and strategic, by shaping customer acquisition, market shares, and prices. This complements Eeckhout and Veldkamp (2023) by identifying a channel through which data generate markups by influencing consumer search. My setting isolates this margin statically; how repeated targeting accumulates into a durable customer base is left to future work.

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 benefit from their continued presence. A targeted offer is priced against what the borrower expects to obtain by continuing to shop; because less-informed non-adopters post higher search-market rates, their presence raises this benchmark and increases the acquisition rent earned by an adopter that reaches the borrower early. That rent is the adopter's markup over its own posted rate, $\chi = \kappa + e_N\Delta$ (equation (8.65)), rising in the borrower's search cost κ and in the share of search draws e_N that reach less-informed rivals. Non-adopters survive on the other side of the same friction: incumbent relationships and high search costs allow them to retain borrowers who do not shop. Coexistence is therefore an equilibrium outcome rather than a transitional stage of technological diffusion. Data-poor lenders retain captive demand, while data-rich lenders earn rents by shaping search rather than eliminating their competitors. Better information can consequently redirect competition toward advertising and customer acquisition rather than being passed through 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. My contribution is to show that, under asymmetric adoption and search frictions, an expansion in available credit-market information can increase non-price competition while preserving the coexistence of data-rich and data-poor lenders.

Second, the paper contributes to work on 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 instead disproportionately exclude marginal applicants (Agarwal et al. 2024). I identify a strategic use of the same information: online-collected data can generate lender rents through customer acquisition. The closest empirical work examines the first California privacy law (CCPA). 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 searchers borrowing from 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 targeting, lowering rates by reducing customer-acquisition rents. Rates alone therefore do not reveal which channel dominates; the joint response of rejections, subsequent default, search-advertising expenditures, 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.

Finally, the paper contributes to the literature on advertising, consideration sets, and non-price competition in consumer finance. Existing work emphasizes two roles for advertising: it can persuade households to purchase costly financial products (Gurun et al. 2016), but it can also inform consumers and expand the set of firms they consider. This consideration-set role has been studied in banking (Mendes 2024; Honka et al. 2017), mutual funds (Hortaçsu and Syverson 2004; Roussanov et al. 2021), and consumer-product markets (Goeree 2008). I build on this second view by emphasizing that targeted advertising affects not only whether a lender enters a borrower's consideration set, but also whether it does so before the borrower obtains a competing quote. The search fatigue documented by Ursu et al. (2023) and the sequential-search framework of Fershtman and Pavan (2025), in which consumers expand their consideration sets gradually and become less likely to inspect additional results, explain why the timing and position of lender exposure matter. The model combines this consideration-set logic with the search-advertising mechanisms of De Corniere (2016) and De Corniere and De Nijs (2016): a lender that reaches a borrower early can price against her continuation value rather than against a competing offer already in hand. I depart from this literature by embedding targeted 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 better targeting 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 it 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 those with high predicted search propensity. 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.

The rest of the paper proceeds as follows. Section 2 describes the institutional setting. Section 3 introduces the data. Section 4 defines borrower search, classifies lenders by their pre-period adoption of online-data technologies, and presents the empirical design. Section 5 reports the main reduced-form results on pricing and borrower allocation. Section 6 examines mechanisms. Sections 8 and 9 develop and estimate the structural model. Section 10 concludes.

2 Institutional Setting

2.1 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.

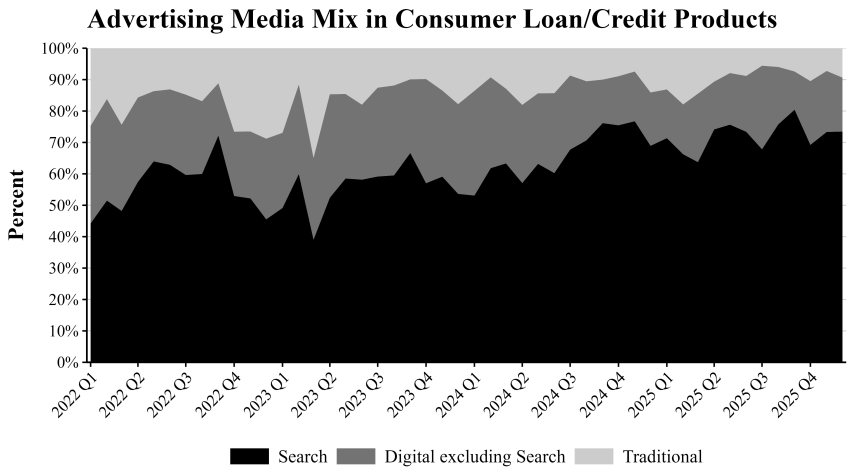
The economic importance of this channel is visible in mortgage advertising expenditures.

Figure 2.1 shows that search advertising accounts for the largest share of mortgage advertising budgets in recent years. Panel B provides examples of competitor-keyword targeting, in which a lender bids on search terms associated with major national lenders. These examples show that lenders compete not only on mortgage rates but also for inclusion 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 2.1. Search Advertising and Competitor Keyword Targeting

Panel A. Search versus other digital advertising



Panel B. Examples of competitor keyword targeting

Date	Advertiser	Display URL	Search keyword	Clicks
2022-01	Aspire Lending	`aspirelending.com`	chase mortgage rates	10
2022-01	Aspire Lending	`aspirelending.com`	wells fargo mortgage rates	11

Notes: Panel A plots digital advertising expenditures by channel. Panel B reports examples of search advertisements in which a lender bid on keywords associated with major national lenders. The sample covers advertising expenditures from 2022 to 2025.
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.

2.2 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.

3 Data

3.1 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. Vivvix reports advertising expenditures by advertiser, media channel, and time. For traditional media and most digital advertising, expenditures are available at the monthly metropolitan-statistical-area level. Search advertising is available at the national level, reflecting the fact that search ads are targeted to individual queries and are harder to assign to fixed geographic markets. I use Vivvix to measure lenders' advertising spending from 2022 to 2025 and to document the importance of search advertising in mortgage customer acquisition.

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.

3.2 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.

3.3 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 4.1 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 set of variables comes from Vivvix. I use these data to measure advertising expenditures by lender, time, media channel, and geography. The analysis distinguishes search advertising, other digital advertising, and traditional advertising. Search advertising is observed at the national level; other media are observed at the metropolitan-statistical-area-by-month level. I restrict attention to mortgage-related advertising and exclude advertisements for other financial

products or general brand campaigns. Lenders' exposure to privacy-law enforcement is measured using their origination shares across states in the 2018–2019 pre-period.

Enforcement exposure is a lender's pre-period (2018–2019) origination share in states that subsequently enforce privacy laws. Because California's law took effect in 2020 and additional states enforced later, exposure is positive from 2020 onward and rises over time as the set of enforcing states expands. This is why the advertising-panel exposure (mean 0.171, 2022–2025) is non-zero and increasing.

The final mortgage sample contains 3,722,265 originated mortgages from 3,466 lenders between 2018 and 2024. Table 3.1 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.

4 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 3.1. 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 advertising sample, 2022–2025						
Enforcement exposure	14,079	0.171	0.303	0.000	0.002	0.177
Search advertising spend	14,079	7,962.857	86,887.228	0.000	0.000	0.000
Digital advertising spend	14,079	8,414.301	85,587.041	0.000	0.000	457.322
Traditional advertising spend	14,079	16,496.565	150,342.800	0.000	0.000	1,073.500
Total advertising spend	14,079	32,873.723	215,021.848	0.000	541.000	4,715.610

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 lender-month advertising sample. For indicator variables and mutually exclusive lender classifications, the mean reports the sample share; standard deviations and percentiles are omitted. Advertising expenditures are reported in the units of the source data. 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.

4.1 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:

$$\Pr(A_i = 1 \mid X_i) = \Lambda(X_i'\beta_A), \tag{4.1}$$

$$\Pr(P_i = 1 \mid A_i = 1, X_i) = \Lambda(X_i'\beta_P), \tag{4.2}$$

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

$$\widehat{\pi}_i^P = \Lambda(X_i'\widehat{\beta}_A) \times \Lambda(X_i'\widehat{\beta}_P). \tag{4.3}$$

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 4.1 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 4.1. 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 (4.3). 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.

4.2 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 (4.4) 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

$$y_{m\ell scq} = \sum_{\substack{k=\underline{k} \\ k\neq -1}}^{\overline{k}} \delta_k \, \mathbf{1}\{q - E_s = k\} + X_m'\gamma + \alpha_{\ell q} + \mu_{c,y(q)} + \eta_{s,\mathrm{mo}(q)} + \varepsilon_{m\ell scq}. \tag{4.4}$$

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 (4.4) separately for borrowers with high and low predicted search propensity, as defined in Section 4.1. 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.

5 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 identifies the equilibrium reallocation that the data restriction sets in motion. Borrowers fall into

[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 (**??**).

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 5.1 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 24.9% baseline searcher share) and raises it for no-technology lenders by 2.07 percentage points (8.3%). 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 5.1).

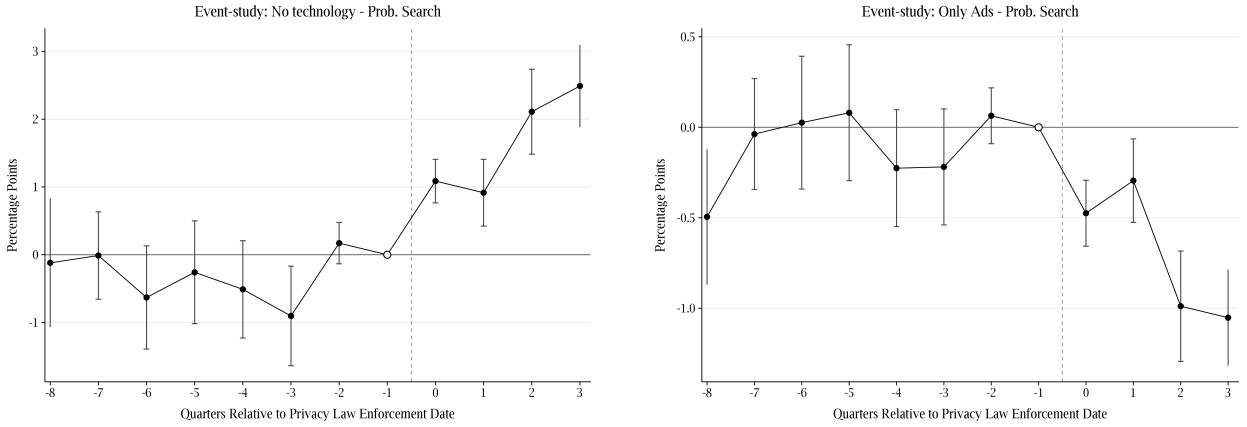
Table 5.2 reports rate effects by borrower type and lender arm. The estimated rate effects for non-searchers are not statistically distinguishable from zero. Searchers at advertising lenders pay 13.95 basis points less (3.3% of the 4.17% mean rate). For a $360,000 thirty-year mortgage, this is roughly $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 5.2 reports wild-cluster bootstrap inference and flat pre-trends.

In sum, privacy laws contract the targeted-advertising channel and the market data acquired from

Table 5.1. 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

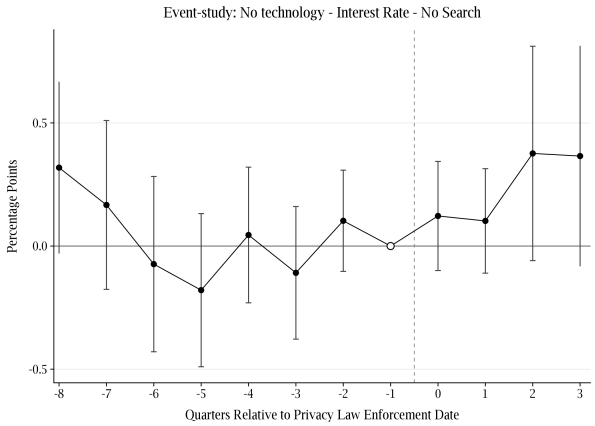
Notes: This table reports difference-in-differences estimates of the effect of privacy-law enforcement on the probability that a borrower searches for price. The coefficient on Post × Treated is the average treatment effect on the treated. Standard errors, clustered at the state level, are reported in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



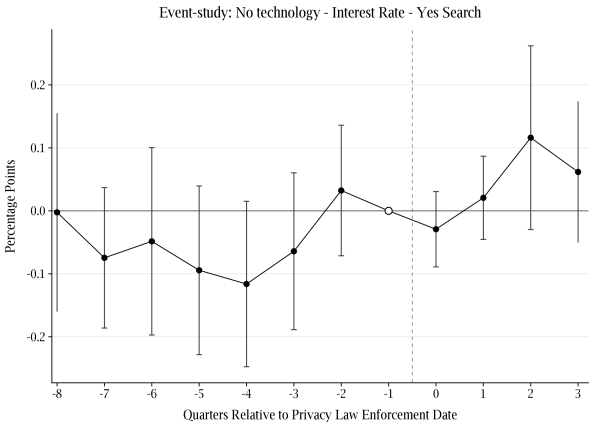
(a): No-Tech **(b):** Ads

Figure 5.1. Event-Study Estimates: Enforcement and Search-for-Price Probability

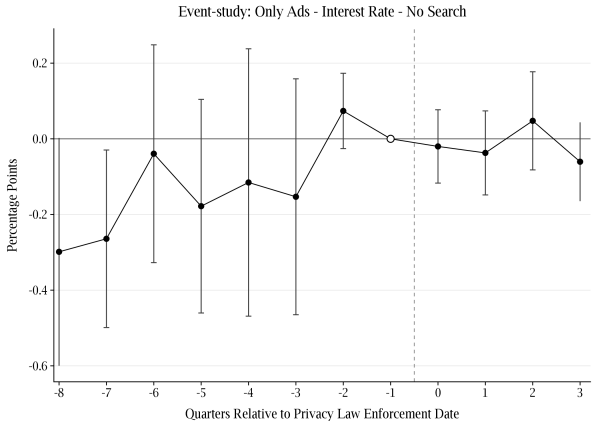
Notes: This figure reports event-study estimates of the effect of privacy-law enforcement on the probability that borrowers search for price. Panel (a) reports estimates for the No-Tech sample and panel (b) for the Ads sample. The omitted period is the last pre-enforcement period. Confidence intervals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



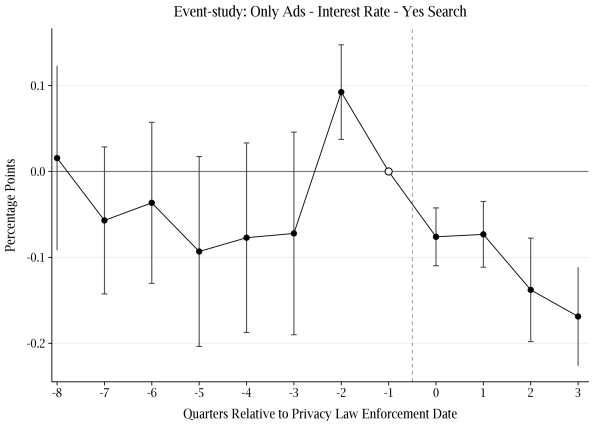
(a): No-Tech, No Search



(b): No-Tech, Search



(c): Ads, No Search



(d): Ads, Search

Figure 5.2. 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 5.2. 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.

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.

6 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 advertising budgets.

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 6.1 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 3.3 percent of the 6.7 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 7.9 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.

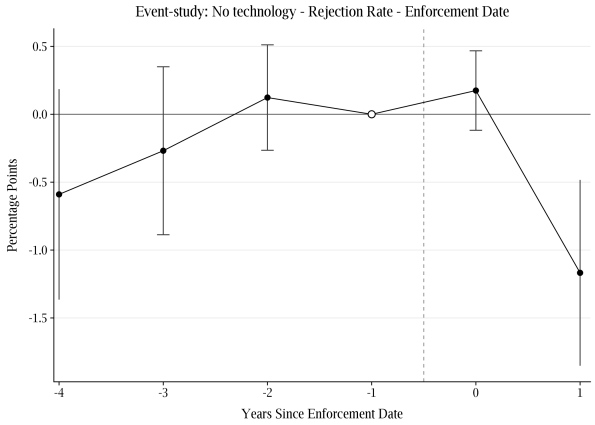
The realized-default evidence completes the screening picture. One-year bankruptcy among advertising lenders' borrowers rises by 2.09 percentage points, a 77 percent increase over the 2.7 percent baseline, while the no-technology estimate is small and statistically insignificant. Rejections falling while 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 6.1, both the treatment effects and the joint nullity of the pre-trends are tested with a wild-cluster bootstrap.

Advertising falls into three types: traditional media such as television and magazines, digital media such as online banners and video, and search advertising such as appearing among the top search results. The first two can be mapped to a geographic unit, since the decision to appear on a given television channel or local website is inherently local. The targeted nature of search advertising, by contrast, makes it orthogonal to the geographic location of the targeted households. Accordingly, the Vivvix data report search advertising only as a national monthly total at the lender level, while the other two media are available at the MSA-lender level. This structure precludes the staggered difference-in-differences design used above. I instead compute lenders' state-level market shares in the pre-privacy-law period (2018–2019). Combining these with a lender-month panel of advertising budgets, the gradual rollout of privacy laws increases the share of each lender's customer base exposed to the new restrictions. This yields a shift-share design in which I measure the effect of privacy-law exposure on advertising budgets. Fixed effects absorb time and lender trends, and standard errors are clustered at the lender level. Because only the largest U.S. lenders advertise heavily, especially in search, I address the sparsity of the dependent variable with a two-part estimator of the extensive and intensive margins, following Chen and Roth (2024). Table 6.2 reports the extensive margin, an indicator for positive spend, and the intensive margin, log spend conditional

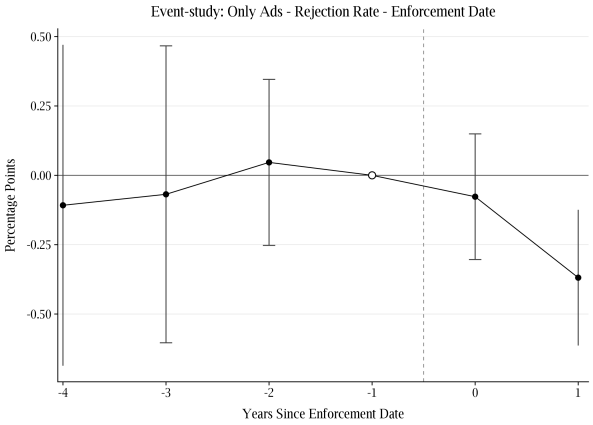
Table 6.1. Enforcement Effects on Rejection Rates and Borrower Bankruptcy

	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

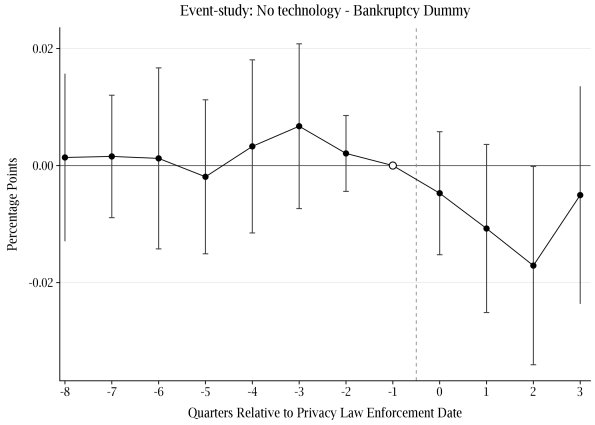
Notes: This table reports difference-in-differences estimates of the effect of privacy-law enforcement on mortgage rejection rates and borrower bankruptcy outcomes. Columns (1)–(2) report effects on the probability that a mortgage application is rejected; columns (3)–(4) report effects on an indicator 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.



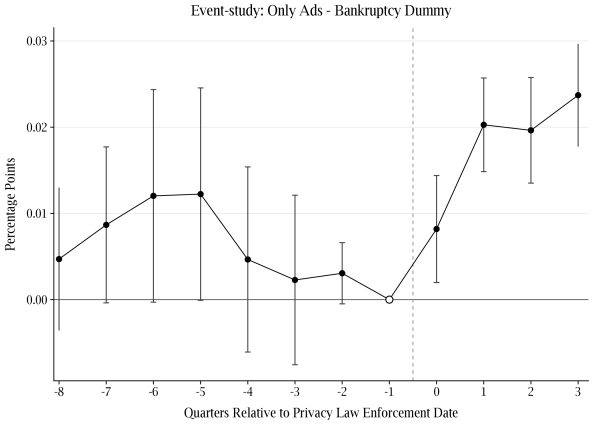
(a): Rejection Rate: No-Tech



(b): Rejection Rate: Ads



(c): Bankruptcy Indicator: No-Tech



(d): Bankruptcy Indicator: Ads

Figure 6.1. 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.

on positive spend, for the three media. Because the specification interacts exposure bins with lender type, I report the combined effect for each lender type, obtained from a linear combination of the exposure baseline and the interaction, and I state magnitudes as percent changes relative to lightly exposed lenders of the same type. Enforcement exposure primarily affects search advertising. For lenders that use only advertising technologies, search spend falls by 81 percent in the 10–20 percent exposure bin and by 71 percent in the 20–60 percent bin. The reduction attenuates above 60 percent exposure, to 3 percent, tracking a declining baseline of search advertising at high exposure: the Vivvix panel begins in 2022 and therefore excludes the first large shock from the California law. The reduction is smaller for lenders that also collect data on their own websites, at 42 percent and 29 percent in the same two bins, because these lenders can retarget their own customers and partially offset the restriction. Panel B reports inference for the underlying coefficients. The analytic tests are strongly significant. The randomization-inference p-values, which are the appropriate benchmark given only 54 lenders on the intensive margin, remain below 0.05 for the only-advertising lenders in the 20–60 percent bin and below 0.10 in the 10–20 percent bin. The joint permutation tests are borderline, an outcome consistent with the small sample induced by restricting attention to search advertisements that contain the word "mortgage."

This section establishes two results. Online-collected data are used in screening, as the joint movement of rejections and realized defaults shows, and are deployed in search-advertising campaigns, as the spending 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 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.

7 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 advertising-budget 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 2% of the baseline mean, against 0.85 percentage points (3.4%) in the full advertising sample. It lowers rejection rates by 0.14 percentage points,

Table 6.2. Search Advertising and Enforcement: Main Estimates and Mechanism Inference

Panel A. Effect of enforcement exposure on advertising spend

(1)	Pr(Search>0) (2)	Pr(Digital>0) (3)	Pr(Trad.>0) (4)	log Search (5)	log Digital (6)	log Trad.
Exposure = 0.1–0.2	0.0774*	-0.0425	-0.0646	1.704***	-1.189*	0.3624
	(0.0405)	(0.0682)	(0.0644)	(0.4027)	(0.6288)	(0.2396)
Exposure = 0.2–0.6	0.0798	-0.1515	0.2108**	3.702***	-1.127**	0.2784
	(0.0845)	(0.0947)	(0.1045)	(0.8126)	(0.5418)	(0.2812)
Exposure > 0.6	-0.0162	-0.2187***	0.0491	0.8940***	-0.3059	0.0532
	(0.0357)	(0.0812)	(0.0801)	(0.2950)	(0.2849)	(0.3171)
Exposure = 0.1–0.2 × Only Ads	-0.0848**	0.0095	0.0707	-3.382***	1.058	-0.6474*
	(0.0431)	(0.0805)	(0.0779)	(0.7351)	(0.7327)	(0.3308)
Exposure = 0.1–0.2 × Ads + Data	-0.0569	-0.0134	0.0683	-2.252***	1.034	-0.4503
	(0.0445)	(0.0746)	(0.0758)	(0.7027)	(0.6764)	(0.3723)
Exposure = 0.2–0.6 × Only Ads	-0.1468	0.1009	-0.2504**	-4.936***	1.155	-0.6942
	(0.0931)	(0.1240)	(0.1159)	(1.018)	(0.7338)	(0.4716)
Exposure = 0.2–0.6 × Ads + Data	-0.0654	0.0648	-0.2161*	-4.045***	1.391**	-0.4922
	(0.0885)	(0.1071)	(0.1149)	(1.194)	(0.6069)	(0.3928)
Exposure > 0.6 × Only Ads	-0.0196	0.2544***	-0.0499	-0.9219**	1.125*	-0.4908
	(0.0440)	(0.0926)	(0.0876)	(0.4127)	(0.5897)	(0.3733)
Exposure > 0.6 × Ads + Data	0.0303	0.1788*	-0.0767	-0.9775**	0.7994*	0.1374
	(0.0387)	(0.0952)	(0.0862)	(0.4750)	(0.4769)	(0.3991)
Lender FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month × Type FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	19,466	19,466	19,466	1,344	5,112	7,327
Effective sample	767	767	767	54	486	485

Panel B. Inference for the intensive-margin Search coefficients (column 4)

	Estimate and analytic inference				Permutation p-value	
Coefficient / Test	Estimate	S.E.	t	p	p_β	p_t
Individual coefficients						
Exposure = 0.1–0.2 × Only Ads	-3.382	0.735	-4.600	<0.001	0.080	0.092
Exposure = 0.1–0.2 × Ads + Data	-2.252	0.703	-3.205	0.002	0.139	0.163
Exposure = 0.2–0.6 × Only Ads	-4.936	1.018	-4.850	<0.001	0.042	0.049
Exposure = 0.2–0.6 × Ads + Data	-4.045	1.194	-3.388	0.001	0.045	0.150
Joint tests (Wald statistic)						
Only Ads: both bins = 0	11.799	—	—	<0.001	0.161	—
Ads + Data: both bins = 0	5.746	—	—	0.003	0.302	—
All four coefficients = 0	5.920	—	—	<0.001	0.303	—

Notes. The table reports the effect of state privacy-law enforcement on lender advertising spend. *Panel A* reports the full specification: columns (1)–(3) are linear probability models for the extensive margin (an indicator for positive spend in the channel), and columns (4)–(6) are OLS regressions of log spend on the intensive margin, conditional on positive spend. Enforcement exposure is binned as 0 (omitted), 0.1–0.2, 0.2–0.6, and > 0.6, each interacted with lender type (Only Ads, Ads + Data; No Ads omitted). All specifications include lender fixed effects and year-month × type fixed effects. Standard errors clustered by lender are in parentheses; *, **, and *** denote significance at the 10%, 5%, and 1% levels. *Panel B* reports inference for the four intensive-margin Search coefficients from column (4). For each coefficient it gives the point estimate, clustered standard error, t-statistic, and analytic p-value, alongside randomization-inference p-values from 9,999 permutations of enforcement dates across states, computed on the coefficient scale (p_β) and using the studentized statistic (p_t). In the joint-test rows the Estimate column reports the Wald statistic, the p column its asymptotic p-value, and the p_β column the permutation p-value of the Wald statistic; dashes denote columns that do not apply.

or 2.1% of the mean, against 0.22 percentage points (3.3%) in the full sample (Table G.2). The attenuation reflects partial exposure to privacy laws, since part of these lenders' borrower base still consents to online data collection, which permits 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 2.8% of the mean rate, rather than the 14.0 basis points (3.3%) of the full sample, and there is now a statistically significant reduction of 9.5 basis points, or 2.3%, for non-searchers as well (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. The only result that does not carry over is the one-year-ahead bankruptcy rate, which is small and statistically insignificant for this group, with a point estimate of 0.39 percentage points against 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.

Moeover, Appendix G.1 also reports the event study for the signature dates (Figure G.1). The goal is to identify eventual anticipation in the first year following the signature. Indeed, the nine privacy laws enforced in our dataset have a minimum gap of four quarters and a median of five before enforcement. The latest estimate may contains part of the enforcement results. The only significance are the interest rate for searchers provided by adopters starts to decline two quarters after signature, as well as the interest rate for non-searchers in non-adopters spikeand decline after three quarters. The main specificaton manage to control for these early dynamics, maybe related to temporal trends not fully captured in the signature specification. In any case, Appendix G.6 verify wether the main results of the paper are only short term uncaptured temporal local trends or permanent transiction under the framework of the California privacy laws implemented during the Covid 19 pandemics. Aggregating at the yearly level, the main results for searchers allocation and paid interest rate from adopters converge over a lower steady state in the long run (Table G.5; Figures G.8 and G.9). The work of Gupta et al. (2024) and Doerr et al. (2026) already provided separate tests for ruling out the role of the pandemic from the California privacy law.

The stringency of the California privacy law direclty map into another source of concern. All the results of the study may be driven by this single privacy law. To address this point, I propose all the main results of the paper by cohort with the exception of the advertisement budget where only national aggregated resulted available (Table G.4). Starting from the loss in market share among searchers borrowers, I observe that the result present and statistically significant across three cohorts out of five for lenders also directly collecting online data, while only in California for lenders only exploiting advertisement platform data. This partially relates to the more stringent requirement of the law, but it does not exclude it on lenders with a more intensive use of online collected data in other states. Moving to the reduction in interest rates for searchers, the result for lenders exploiting ad-tech data only is not confined in California being present also for Colorado and Connectictut. Both the laws are slightly less stringent than the California one. Indeed, once again, I document a larger drop on interest rate for the lenders also directly collecting online data under the weaker laws of Montana, Florida, Oregon and Texas. Utah represents the only state where privacy laws translate in higher interest rate for lenders directly collecting online data and no effect on the ones using only advertisemtn platform data. At the same time, Utah is also the only cohort witness lower bankruptcy rates controlling for observable. All the other cohorts result in a bankrutpcy increase for lenders acquiring data from advertisement platforms, with the last batch of implementers where few

data points are available being our measure collected one year post origination. The divergence of Utah may be explained by one of the most business friendly approach among the state privacy laws. Consumers have to opt-out from data collection from each single website, while the state grants a cure period for the sanctioned firms and mantain a very low enforcement level. Similar results matches with a more certain trust and legal environement leading to a more pronounced use of these technologies, as predicted by Doerr et al. (2026). Finally, the results on rejection rates are statistically significant only for California and non-signficant with a confidence interval around zero for the other privacy laws. Such a result may originate from the high aggregation of these yearly data and the California mortgage market size, as well as from the higherst stringency of this law. Coming to a conclusion, privacy laws stringency plays a role in driving the magnitude and the direction of the results. However, under different intensity, the main findings of the present study applies to all the privacy laws.

The staggered difference in difference relies on the assumption of stable units tretment value assumption (SUTVA). A privacy law in state A should not have affect on state B used for the comparison analysis. To validate this concern, I borrow the exposure to privacy laws from the advertising budget analysis and I interact it with a factor variables indicating the use of online data of the lender. The used sample contains only states not enforcing, or signing a privacy law. The fixed effects remains similar to the main specification used in the paper, while the standard errors are now clustered at the lender level (Table G.6). The only source of concerns relates to the market share of searchers. The more lender are exposed to privacy laws and the more no-tech lenders gain market share at the expense of tech one. Such a result matches what observed on the advertising result. Lenders exploiting online data for targetted advertising find it less profittable when a large part of their customer basis result exposed to privacy law, while no-tech lenders find it profittable at the resulting lower prices of the new equilibrium.

The SUNAB estimator does not permits interaction in the staggered difference in difference estimator. As a result, I split the population into several samples to test different hypothesis. Such a procedure can generate a measurement problems whenever different samples result correlated among them. To solve this issue, Appendix G.4 repropose the main results of the study through a stacked regression that hold all the lenders' types into a single regressions (Table G.3). I interact controls and fixed effects with the lender's type to keep it comparable with the SUNAB estimator. For the probability of searching for price, enforcement raises search among no-technology lenders by 2.09 percentage points and lowers it among advertising adopters by 0.87 percentage points, against 2.07 and 0.85 percentage points in the baseline. The pass-through to interest rates is likewise unchanged: searchers of advertising adopters pay 13.97 basis points less, essentially identical to the 13.95 basis points of the main specification, while no-technology searchers remain unaffected (4.84 versus 4.81 basis points, both insignificant). The two remaining outcomes agree as well. Enforcement leaves

the subsequent bankruptcy probability of no-technology borrowers statistically unchanged (0.66 percentage points, against 0.64 in the baseline) while raising it by 2.01 percentage points among advertising adopters, close to the baseline estimate of 2.09 percentage points. Rejection rates decline modestly for both families—by 0.53 and 0.22 percentage points for no-technology and advertising lenders, respectively—reproducing the baseline estimates of 0.53 and 0.22 percentage points almost exactly. Summing up, the sample splitting required by the SUNAB estimator does not distort the sign and nor the magnitude of the estimated effects.

The robustness relative to the advertising budget mechanism and to estimation of the searching for price probability are reported into the respective Mechanism (Section 6) and Empircal Strategy (Section 4) of the paper. I do not report them here for the sake of simplicity.

8 Model

To discipline the mechanism, I develop a search model of mortgage pricing with lender screening and targeted advertising. The model is designed to capture two features of the mortgage market. First, borrowers enter the market with imperfect information about the relevant price of credit. Publicly available mortgage-rate benchmarks are typically reported within credit-score buckets, but they do not fully account for fees, discount points, and other contract terms. Obtaining a complete and comparable loan estimate requires time, and borrowers face a limited window during the house-purchase process. Second, borrowers do not approach lenders symmetrically. Each borrower begins with an incumbent financial institution, which provides an initial quote at low cost and may benefit from both trust and informational advantages. The borrower then decides whether to accept an available quote or incur the cost of searching for additional lenders.

Borrowers differ along two dimensions. They differ in repayment risk and in the cost of searching for additional mortgage quotes. To keep the screening channel transparent, I use a binary repayment technology: low-risk borrowers repay, whereas high-risk borrowers generate a default loss for the lender. This stark assumption is a tractable normalization rather than a literal description of repayment behavior. The model only requires that high-risk borrowers have higher expected default costs than low-risk borrowers. High-risk borrowers also face a greater risk of rejection when applying for credit. Rejection risk makes search more costly for risky borrowers because a quoted rate is valuable only if the application is ultimately approved. As a result, some borrowers may accept an early quote rather than continue searching, while others may search until they obtain an approved offer. Search frictions therefore affect both the borrower's outside option and the composition of borrowers who reach different lenders.

Lenders differ in whether they use online-collected data technologies. Non-adopting lenders acquire borrowers through incumbent relationships and through borrowers' organic search. Adopting

lenders have access to an additional technology that serves two functions. First, online-collected data helps identify borrowers who are actively considering a mortgage, allowing adopting lenders to enter the borrower's consideration set through targeted advertising. Second, the same data provides a signal about repayment risk within a credit-score bucket, improving pre-screening before the loan is originated. The technology therefore affects both borrower acquisition and lender screening.

These two uses of online-collected data have opposing implications for mortgage rates. Better screening reduces adverse selection among approved borrowers and lowers the expected default cost of originated loans. This force pushes rates down for borrowers who are approved by data-using lenders. Targeted advertising works in the opposite direction. By reaching borrowers early in the search process, an adopting lender can price relative to the borrower's reservation value: the expected rate from continued search plus the rate-equivalent cost of searching and applying elsewhere. This force allows the lender to preserve a search-cost wedge over marginal cost. The equilibrium price effect of online-collected data is therefore ambiguous ex ante and depends on whether the screening channel or the borrower-acquisition channel dominates.

The model maps these forces into three borrower-acquisition channels: incumbency, targeted advertising, and borrower-initiated search. Incumbent borrowers receive an initial quote from their home bank. Borrowers reached through targeted advertising may receive an additional quote from a data-using lender before conducting costly search. Borrowers who continue searching draw lenders from the broader market and face posted search rates. Thus, each lender type may serve borrowers through distinct channels, and rates differ across channels because borrowers' outside options and lenders' information sets differ across them.

Conceptually, the model combines the mortgage-search-and-approval framework of Agarwal et al. (2024) with endogenous consideration sets under advertising, as in Honka et al. (2017), and the sequential-search logic of paid search advertising in De Corniere (2016). Its departure from these frameworks is that the same online-collected data governs both targeting and screening. The model is solved within credit-score buckets, so adverse selection arises from residual heterogeneity in repayment risk among observationally similar borrowers. The equilibrium is recursive: posted rates determine the expected search-market rate, which determines borrowers' reservation values, channel-specific rates, search decisions, approval outcomes, and lender volumes. Estimating the model allows me to decompose the value of online-collected data into screening and targeted-acquisition components and to evaluate how privacy restrictions redistribute surplus between borrowers and lenders.

8.1 Environment

Agents, loans, and information

Online-collected data used in targeted advertising is the technology that separates lenders into two types: technology adopters T and non-adopters N. The technology has two components. First, it allows adopters to identify borrowers who are actively interested in obtaining a mortgage, summarized by the advertising-reach parameter γ. Second, it provides a signal about repayment risk, summarized by the screening index z_τ. Non-adopters do not use this technology, while adopters receive both the advertising and screening components:

$$\begin{aligned} N &= (\gamma_N, z_N) = (0, 0), \\ T &= (\gamma_T, z_T) = (\gamma, \alpha), \qquad \gamma \in (0, 1), \quad \alpha > 0. \end{aligned} \tag{8.1}$$

Thus, α indexes the additional screening precision available to technology adopters.

There is a unit mass of borrowers, indexed by i. Each borrower demands one unit of credit, interpreted as a normalized mortgage, obtains utility v_i from the loan, and promises to repay $1 + r$ after one period. Borrowers differ in repayment risk. There are two types, $\theta_i \in \{L, H\}$: low-risk borrowers L repay, whereas high-risk borrowers H generate a default loss for the lender. This binary repayment technology is a tractable normalization. The model only requires that high-risk borrowers have higher expected default costs than low-risk borrowers. Let default probabilities satisfy

$$0 \leq d_L < d_H \leq 1. \tag{8.2}$$

The population shares are

$$\mu_H = \Pr(\theta_i = H), \qquad \mu_L = 1 - \mu_H. \tag{8.3}$$

Borrower type is not observed by the lender. Therefore, the cost of originating one unit of credit equals the funding rate f plus the expected default cost of the borrower pool being served. A lender charging rate r obtains payoff $r - f - d_\theta$ from a borrower with expected default cost d_θ.

Each borrower has an incumbent home bank that provides the first quote at zero search cost. Incumbency shares by lender type satisfy $h_T + h_N = 1$. The model is solved within credit-score buckets s, so the relevant adverse-selection problem comes from residual heterogeneity in repayment risk among borrowers with the same credit score. I omit s from the notation to simplify the analysis.

Timing

The model proceeds as follows.

a. Each lender type posts a public interest rate r_τ^{se} for searching borrowers and an incumbency rate $\bar{r}$ for borrowers for whom it is the home bank.

b. Borrowers search for houses and obtain an initial quote $\bar{r}$ from their home bank. This search activity generates a signal on advertising platforms. With probability γ, a technology lender T wins the top slot in the borrower's search results and captures the borrower's attention. The lender can then deliver an advertising rate r^{ad}. Borrowers flagged as high-risk by the signal and borrowers for whom T is already the home bank do not receive r^{ad}.

c. The borrower decides whether to accept one of the received rates, $\bar{r}$ or, when available, r^{ad}, or instead enter costly search.

d. A searching borrower pays a cash application cost κ_i. In rate-equivalent terms, the cost of obtaining an approved outside option is

$$\tilde{\kappa}_i = \kappa_i / q^e_{\theta_i}.$$

The borrower draws a lender of type τ with probability e_τ, where $e_N + e_T = 1$. The searched rate is below the borrower's reservation rate, so the relevant friction in this stage is the cost of generating an approved outside option.

e. The borrower accepts the first approved searched rate. The lender then decides whether to approve or reject the application. If the application is rejected, the borrower falls back on any previously approved home-bank rate $\bar{r}$ or continues searching.

f. The loan matures. Low-risk borrowers repay, while high-risk borrowers misclassified as low-risk default.

The screening technology

Assumption 1. *Every lender type τ can flag a high-risk applicant with baseline probability $p_0 \in (0, 1)$ and misclassifies the applicant as low-risk with probability $q \equiv 1 - p_0$. Using online-collected data, a technology adopter T can flag a high-risk applicant with probability*

$$P(z) = p_0 + (1 - p_0)\, p(z) < 1, \tag{8.4}$$

where $p(0) = 0$, $p'(z) > 0$, and $p''(z) < 0$. The signal is observed before underwriting because targeted-advertising platforms allow lenders to pre-screen potential borrowers using demographics, location, and search behavior. Low-risk borrowers are never flagged as high-risk and are always approved. The signal can also be used for borrowers reaching the lender through the search channel.

Its informative value has been documented by Chen et al. (2023) and Chen et al. (2026).

Assumption 1 implies that some high-risk borrowers survive screening. This creates a role for rejection risk in borrower search. High-risk borrowers may accept an early approved offer to avoid the risk of rejection from additional applications, while rejected high-risk borrowers may continue searching until approval materializes.

As shown in Appendix I.1, in the paid and posted-search channels it is optimal to reject borrowers flagged as high-risk. The approval probability of a high-risk borrower at lender type τ is

$$q_H(z_\tau) = q\left(1 - p(z_\tau)\right), \qquad q_H(0) = q, \qquad q_T \equiv q_H(\alpha) < q_H(0). \tag{8.5}$$

The first equality defines the high-risk approval probability as a function of the screening signal. The second equality follows from the normalization $p(0) = 0$ and gives the approval probability at non-adopters, where $z_N = 0$. The third defines q_T as the approval probability at technology adopters, where $z_T = \alpha$.

The approval probability A_τ, default probability D_τ, marginal mortgage cost c_τ, and rejection rate R_τ for lender type τ are

$$A_\tau = \mu_L + \mu_H q_H(z_\tau), \qquad D_\tau = \frac{\mu_L d_L + \mu_H q_H(z_\tau)\, d_H}{A_\tau}, \qquad c_\tau = f + D_\tau, \qquad R_\tau = \mu_H\left(1 - q_H(z_\tau)\right). \tag{8.6}$$

Online-collected data z_τ improves pre-screening and reduces the default probability among approved borrowers. Hence,

$$\bar{D} \equiv D_N > D_T, \qquad c_T < c_N, \qquad \Delta \equiv \bar{D} - D_T > 0. \tag{8.7}$$

Thus, technology adopters have a marginal-cost advantage for approved borrowers.

Given cross-selling incentives and relationship banking, the home bank provides the first rate $\bar{r}$ based on the unscreened borrower pool. The corresponding default probability is

$$D_0 \equiv \mu_L d_L + \mu_H d_H. \tag{8.8}$$

This pooling can cross-subsidize high-risk borrowers with inattentive low-risk borrowers, while attentive low-risk borrowers are more likely to search.

Search costs

Assumption 2. *Independently of borrower type θ_i, borrowers differ in the cost of finding and applying to a new lender:*

$$\kappa_i \in \{\kappa, \kappa^c\}, \qquad \kappa < \kappa^c, \tag{8.9}$$

with

$$\Pr(\kappa_i = \kappa^c) = \lambda \in (0, 1). \tag{8.10}$$

Borrowers observe interest rates but do not observe lender type or the lender's use of online-collected data, (γ_τ, z_τ). Therefore, when entering search, they form expectations over lender types. Since a searching borrower draws a lender of type τ with probability e_τ, the expected approval probability of a high-risk borrower in the search market is

$$q_H^e \equiv e_N q_H(0) + e_T q_H(\alpha). \tag{8.11}$$

Low-risk borrowers are always approved, so $q_L^e = 1$. A quote is valuable only conditional on approval. Rejection risk therefore raises the effective cost of search. The per-quote search cost is

$$\tilde{\kappa}_i = \kappa_i / q_{\theta_i}^e, \tag{8.12}$$

where $q_L^e = 1$ and $q_H^e = e_N q_H(0) + e_T q_H(\alpha)$. I impose the search-cost separation

$$\frac{\kappa}{q_H^e} < \kappa^c. \tag{8.13}$$

This condition ensures that low-cost borrowers search even when they are high risk, while high-cost borrowers remain captives at the home-bank contract characterized in Lemma 4.

Low-risk borrowers are always approved, so their effective search cost is κ_i. High-risk borrowers are approved only with probability q_H^e, and therefore require $1/q_H^e > 1$ expected attempts to obtain an approved offer. This mechanism discourages further search after approval and makes search for approval costly. It can also generate self-exclusion from mortgage markets. For low-search-cost borrowers, the type-weighted effective search cost is

$$\tilde{\kappa}^L \equiv \kappa \left(\mu_L + \mu_H / q_H^e\right). \tag{8.14}$$

Ad rates, posted rates, and incumbency rates

Assumption 3. *Ad rates. When borrowers search for housing options and home-bank mortgage rates, web-search activity transmits a signal to advertising platforms. These platforms sell prominent search-result slots to lenders. A lender of type N does not use this technology. A lender of type T wins the relevant slot and captures the borrower's attention with probability γ. The advertising probability is also the probability of a real interaction, since targeted advertising is paid only on impressions that generate borrower attention. A lender T does not target borrowers flagged as high-risk and does not target borrowers already attached to its own incumbency pool.*

Posted rates. Each lender competes in a differentiated-product Bertrand posted-rate subgame for searching borrowers. Mortgage products are differentiated because borrowers value lender-specific attributes, including service quality, processing speed, brand, branch relationships, and platform convenience. This differentiation allows lenders to earn a common equilibrium markup m_0 over marginal cost. The equilibrium rate, derived in Appendix H.3, is

$$r_\tau^{\text{se}} = c_\tau + m_0. \tag{8.15}$$

A searching borrower matches with lender type τ with probability $e_\tau > 0$, where $e_N + e_T = 1$. The expected search rate is therefore

$$\tilde{r} \equiv e_N r_N^{\text{se}} + e_T r_T^{\text{se}} = f + m_0 + D^e, \qquad D^e \equiv e_N \bar{D} + e_T D_T = \bar{D} - e_T \Delta. \tag{8.16}$$

I impose the posted-rate cap

$$\kappa > r_N^{\text{se}} - \tilde{r} = e_T \Delta. \tag{8.17}$$

This condition prevents borrowers from searching only to obtain a technology adopter. It also implies that every searched posted rate is strictly below the reservation rate of any borrower who enters search. Since $c_T < c_N$ and $r_\tau^{\text{se}} = c_\tau + m_0$, the highest posted rate is r_N^{se}. Therefore, for any searching borrower,

$$r_N^{\text{se}} - \tilde{r} = e_T \Delta < \kappa \leq \tilde{\kappa}_i, \tag{8.18}$$

so

$$r_N^{\text{se}} < \tilde{r} + \tilde{\kappa}_i = \rho_i. \tag{8.19}$$

The screening advantage arises from ad-platform data acquired during the online application and search process.

Incumbency rate. Each borrower receives a rate $\bar{r}$ from the home lender. This rate can prevent search by borrowers for whom additional search is sufficiently costly, and it is chosen optimally

according to Lemma 4. The home bank excludes no borrower at this stage and prices the mortgage using the default probability of the unscreened pool, D_0. The incumbency channel is exposed to poaching through targeted advertising, but because advertising reach is imperfect, $\gamma < 1$, the incumbency channel does not disappear.

8.2 The Borrower's Problem

This subsection characterizes borrower search and acceptance decisions. Since the utility from obtaining the mortgage, v_i, is common across the relevant alternatives, the borrower chooses the option with the lowest expected rate-equivalent cost. Let r_i^0 denote the best approved rate initially available to borrower i. The borrower always receives the home-bank rate $\bar{r}$. If an advertising quote is delivered, the borrower also observes r^{ad}; if no advertising quote is delivered, set $r^{\text{ad}} = +\infty$. Hence,

$$r_i^0 = \min\{\bar{r}, r^{\text{ad}}\}. \tag{8.20}$$

A searched quote is valuable only if the application is approved. Borrowers know their own type but do not observe lender type or the lender's use of online-collected data. Therefore, when entering search, they form expectations over lender types. Let $q_L^e = 1$ and let q_H^e denote the expected approval probability of a high-risk borrower in the search market, as defined in Assumption 2. The effective per-quote search cost is

$$\tilde{\kappa}_i = \frac{\kappa_i}{q_{\theta_i}^e}. \tag{8.21}$$

The cash cost of one application is κ_i. The factor $1/q_{\theta_i}^e$ is the rejection loading induced by the need to obtain an approved offer.

Before receiving any advertising quote, borrower i evaluates search using the organic information set

$$\mathcal{I}_i^{\text{org}} = \{\tilde{r}, \bar{r}, \kappa_i, q_{\theta_i}^e\}. \tag{8.22}$$

The advertising quote is not part of this information set. If an advertising quote is delivered, the borrower evaluates it as an additional held quote against the same continuation value.

Given a currently held approved rate r, the borrower's continuation value is

$$V_i(r) = \max\left\{v_i - (1+r), -\tilde{\kappa}_i + \sum_{\tau} e_\tau V_i\left(\min\{r, r_\tau^{\text{se}}\}\right)\right\}. \tag{8.23}$$

The first term is the payoff from accepting the current rate. The second term is the value of continuing to search: the borrower pays the approval-adjusted search cost, draws lender type τ with probability e_τ, and keeps the better of the current approved rate and the searched posted rate. Free recall implies

that previously approved rates remain available.

The analysis uses two maintained restrictions from the environment. First, the search-cost separation in Assumption 2 requires

$$\frac{\kappa}{q_H^e} < \kappa^c. \tag{8.24}$$

This condition ensures that low-cost borrowers search even when they are high risk, while high-cost borrowers remain captives at the home-bank contract characterized in Lemma 4. Second, the posted-rate cap in Assumption 3 requires

$$\kappa > r_N^{\text{se}} - \tilde{r} = e_T \Delta. \tag{8.25}$$

This condition prevents borrowers from searching only to draw a technology adopter and implies that every searched posted rate is strictly below the reservation rate of any borrower who enters search.

Lemma 1 (Acceptance, stopping, and participation). *Under Assumptions 1–3, and given the home-bank contract $\bar{r}$, the borrower's problem has the following solution.*

a. *Acceptance rule.* A borrower holding a delivered quote at rate r accepts it and stops searching if and only if

$$r \leq \min\{\rho_i, \bar{r}\}, \qquad \rho_i \equiv \tilde{r} + \tilde{\kappa}_i. \tag{8.26}$$

The reservation rate ρ_i is the expected market rate from continued search plus the borrower's effective search cost. For shoppers, $\rho_i < \bar{r}$, so the relevant outside option is continued search. For captives, $\bar{r} \leq \rho_i$, so the relevant outside option is the home-bank contract. Ties between accepting and continuing to search break toward accepting; ties between two held options at the same rate break toward the home bank.

The optimal advertised quote, $r^{\text{ad}} = \tilde{r} + \kappa$, equals the expected market rate plus the lower bound of search costs. This characterization follows from the lender's paid-quote problem, formalized in Lemma 3. Conditional on that equilibrium quote, every unflagged recipient accepts. Shoppers accept because $\tilde{\kappa}_i \geq \kappa$, so

$$r^{\text{ad}} = \tilde{r} + \kappa \leq \tilde{r} + \tilde{\kappa}_i = \rho_i. \tag{8.27}$$

Captives accept because the advertising quote is strictly below the home-bank rate:

$$r^{\text{ad}} = \tilde{r} + \kappa < \tilde{r} + \kappa^c = \bar{r}. \tag{8.28}$$

b. *First-draw acceptance.* The recursive search problem collapses to first-draw acceptance under

the posted-rate cap. Because online-collected data lowers the default probability among approved borrowers, technology adopters have lower marginal mortgage costs: $c_T = f + D_T < f + D_N = c_N$. Posted search rates equal marginal cost plus the common markup m_0, so

$$r_T^{\text{se}} = c_T + m_0 < c_N + m_0 = r_N^{\text{se}}. \tag{8.29}$$

Thus, r_N^{se} is the highest posted search rate. Under the posted-rate cap,

$$r_N^{\text{se}} - \tilde{r} = e_T \Delta < \kappa \leq \tilde{\kappa}_i. \tag{8.30}$$

It follows that

$$r_N^{\text{se}} < \tilde{r} + \tilde{\kappa}_i = \rho_i. \tag{8.31}$$

All posted rates are weakly below r_N^{se}, so any searched approved quote is acceptable and there is no value to drawing again. A borrower who enters search pays the cash cost κ_i and draws lender type τ with probability e_τ. With probability $q^e_{\theta_i}$, the application is approved and the borrower is funded at the searched rate; with probability $1 - q^e_{\theta_i}$, the application is rejected and the borrower falls back on the previously approved rate.

c. *Participation.* Starting from the home-bank contract, the expected cost of entering search is

$$\kappa_i + q^e_{\theta_i} \tilde{r} + \left(1 - q^e_{\theta_i}\right) \bar{r}. \tag{8.32}$$

The borrower enters search if and only if this expected cost is below the home-bank rate $\bar{r}$. Therefore,

$$\kappa_i + q^e_{\theta_i} \tilde{r} + \left(1 - q^e_{\theta_i}\right) \bar{r} < \bar{r} \quad \Longleftrightarrow \quad \tilde{\kappa}_i < \bar{\kappa} \equiv \bar{r} - \tilde{r}. \tag{8.33}$$

At the contract characterized in Lemma 4, $\bar{r} = \tilde{r} + \kappa^c$, so $\bar{\kappa} = \kappa^c$. By the search-cost separation in Assumption 2, every κ-borrower searches. Low-risk low-cost borrowers have $\tilde{\kappa}_i = \kappa < \kappa^c$, and high-risk low-cost borrowers have $\tilde{\kappa}_i = \kappa / q^e_H < \kappa^c$. Every κ^c-borrower stays. The low-risk high-cost borrower is the marginal captive, with $\tilde{\kappa}_i = \kappa^c = \bar{\kappa}$, and the tie breaks toward the home bank. Equivalently, Lemma 4 prices the home-bank contract at the search reservation value of this marginal captive. The high-risk high-cost borrower strictly prefers staying because $\tilde{\kappa}_i = \kappa^c / q^e_H > \kappa^c = \bar{\kappa}$.

Corollary 1 (High-risk information rents). *Approved high-risk borrowers are pooled with low-risk borrowers. Because rejection risk raises their effective search cost, approved high-risk borrowers receive an information rent equal to the gap between their reservation rate and the rate they pay. This rent equals $\kappa^c (1 - q^e_H)/q^e_H$ on the home-bank rate and $\kappa(1 - q^e_H)/q^e_H$ on the advertising rate.*

Lemma 1 shows that the home-bank pool contains borrowers who do not search because their effective search costs are high. Searchers continue only until the expected gain from a new approved quote equals the effective cost of search. The advertising rate prevents costly search, but the lender captures part of the saved search cost through the advertising quote.

8.3 The Lender's Problem

Lenders serve borrowers through three channels: organic search, targeted advertising, and incumbency. A lender of type $\tau \in \{N, T\}$ chooses a posted search rate r_τ^{se}, an advertising policy a_τ, a home-bank rate $\bar{r}_\tau$, and an underwriting rule w_τ. The posted search rate applies to borrowers who enter organic search. The advertising policy applies to non-customers reached through the paid channel. The home-bank rate applies to incumbent borrowers who first receive a quote from their existing lender.

The underwriting rule in the paid and posted-search channels is

$$w_\tau = w_\tau(\text{high-risk flag}) \in \{\text{fund}, \text{reject}\}. \tag{8.34}$$

The home-bank channel is priced as an unscreened incumbency pool, as described above. The advertising policy is

$$a_\tau = a_\tau\,(\text{credit score bucket}, \text{shopping intent}, \text{high-risk flag}, o_i) \in \{\tilde{r} + x, \text{withhold}\}, \tag{8.35}$$

where $o_i \in \{0, 1\}$ indicates whether borrower i is already a customer of the lender. Consistent with Assumption 3, lenders do not advertise to their own customers, so

$$a_\tau\big|_{o_i=1} = \text{withhold}. \tag{8.36}$$

Advertising can condition on borrower intent and screening flags, but not on the borrower's search cost:

$$\frac{\partial a_\tau}{\partial \kappa_i} = 0. \tag{8.37}$$

A delivered advertising quote $\tilde{r} + x$ is accepted only if it is no greater than the borrower's relevant outside option:

$$\tilde{r} + x \le \min\{\rho_i, \bar{r}\}, \qquad \rho_i \equiv \tilde{r} + \tilde{\kappa}_i. \tag{8.38}$$

The lender's expected profit is

$$\Pi_\tau = \underbrace{Q_\tau^{\text{se}}\left(r_\tau^{\text{se}} - f - D_\tau\right)}_{\text{posted-search channel}} + \underbrace{\gamma_\tau h_{-\tau}\left[\pi_\tau^{\text{ad}}(a_\tau) - c^{\text{ad}}\right]}_{\text{targeted-advertising channel}} + \underbrace{B_\tau\left(\bar{r}_\tau\right)}_{\text{incumbency channel}}. \tag{8.39}$$

The first term is profit from borrowers who enter organic search and draw lender type τ. The second term is profit from targeted advertising to non-customers of lender type $-\tau$, net of the platform cost c^{ad}. Since $\gamma_N = 0$, this channel is active only for technology adopters. The third term, $B_\tau(\bar{r}_\tau)$, is profit from the lender's home-bank customers after borrower search and advertising poaching.

The posted search rate of lender type τ is

$$r_\tau^{\text{se}} = c_\tau + m_0, \qquad c_\tau = f + D_\tau, \tag{8.40}$$

where m_0 is the differentiation markup from the posted-rate subgame in Appendix H.3. The expected search rate is

$$\tilde{r} \equiv e_N r_N^{\text{se}} + e_T r_T^{\text{se}} = f + m_0 + D^e, \tag{8.41}$$

where

$$D^e \equiv e_N \bar{D} + e_T D_T = \bar{D} - e_T \Delta. \tag{8.42}$$

Thus, D^e is the expected default cost of a searched lender draw. It enters borrower reservation values because a searching borrower draws a non-adopter with probability e_N and an adopter with probability e_T.

The incumbency-profit component is

$$B_\tau(\bar{r}_\tau) = \begin{cases} h_T \lambda\left(\bar{r}_T - f - D_0\right), & \tau = T, \\ h_N \lambda M(\gamma)\left(\bar{r}_N - f - D_0^\gamma\right), & \tau = N. \end{cases} \tag{8.43}$$

where h_τ is the share of borrowers whose home bank is of type τ and $h_T + h_N = 1$. For technology adopters, the home-bank book consists of high-search-cost incumbent borrowers and is not exposed to the lender's own advertising. For non-adopters, the home-bank book is exposed to poaching by technology adopters. The residual mass of the non-adopter home-bank book is

$$M(\gamma) = \mu_L(1-\gamma) + \mu_H\left(1 - \gamma q_H(\alpha)\right), \tag{8.44}$$

and its default probability is

$$D_0^\gamma = \frac{\mu_L(1-\gamma)d_L + \mu_H\left(1 - \gamma q_H(\alpha)\right) d_H}{M(\gamma)}. \tag{8.45}$$

Because targeted advertising removes low-risk borrowers more effectively than high-risk borrowers, the residual non-adopter home-bank book becomes riskier as advertising reach increases:

$$\frac{\partial D_0^\gamma}{\partial \gamma} > 0. \tag{8.46}$$

Lemma 2 (Separability). *Conditional on the underwriting rule and on the market anchor $\tilde{r}$, the lender's problem is additively separable across the posted-search, targeted-advertising, and incumbency channels.*

The underwriting rule determines D_τ, the marginal mortgage cost $c_\tau = f + D_\tau$, and the posted search rate $r_\tau^{\text{se}} = c_\tau + m_0$. Conditional on these objects, the profit representation in equation (8.39) contains three distinct controls: r_τ^{se} in the posted-search channel, a_τ in the targeted-advertising channel, and $\bar{r}_\tau$ in the incumbency channel. The rival's advertising reach γ enters the non-adopter's home-bank book through $M(\gamma)$ and D_0^γ, but it is an exposure rather than a choice variable of the non-adopter. Hence the joint problem decomposes into the three subproblems characterized below.

Assumption 4. *Served markets.* Consider a lender writing a mortgage at rate $\tilde{r} + x$ to a borrower or borrower pool with default risk d. The per-loan profit is

$$\pi(x, d) \equiv (\tilde{r} + x) - f - d = m_0 + D^e + x - d. \tag{8.47}$$

The function $\pi(x, d)$ is increasing in the rate margin x and decreasing in default risk d. The following conditions determine which borrower pools are profitable to serve.

(i) Funding a high-risk borrower is unprofitable at every rate she would accept:

$$\pi\left(\kappa^c/q_H^e, d_H\right) \le 0 \quad \Longleftrightarrow \quad d_H - D^e \ge m_0 + \kappa^c/q_H^e. \tag{8.48}$$

(ii) The residual home-bank book is profitable when priced at the captive margin κ^c:

$$\pi\left(\kappa^c, D_0^\gamma\right) \ge 0 \quad \Longleftrightarrow \quad \kappa^c + m_0 \ge D_0^\gamma - D^e. \tag{8.49}$$

(iii) Defending the whole home-bank book against poaching by pricing at the shopper margin κ is unprofitable:

$$\pi\left(\kappa, D_0\right) \le 0 \quad \Longleftrightarrow \quad \kappa + m_0 \le D_0 - D^e. \tag{8.50}$$

(iv) On the paid channel, pooling shoppers and captives at the shopper margin strictly dominates

skimming only captives at the captive margin:

$$(1-\lambda)\left[\kappa + e_N(m_0+\Delta)\right] > \lambda(\kappa^c - \kappa). \tag{8.51}$$

Conditions (ii) and (iii) place the excess default cost of the unscreened book between the shopper and captive margins:

$$\kappa + m_0 \le D_0 - D^e \le D_0^\gamma - D^e \le \kappa^c + m_0. \tag{8.52}$$

Together with the search-cost separation in Assumption 2 and the posted-rate cap in Assumption 3, these restrictions imply

$$e_T\Delta < \kappa < \kappa/q_H^e < \kappa^c < \kappa^c/q_H^e. \tag{8.53}$$

Thus, Assumption 4(i), stated at the highest margin a high-risk borrower would accept, also rules out funding a high-risk borrower at any lower acceptable rate.

Advertising rate

Lemma 3 (Optimal advertising quote). *Under Assumption 4, a technology adopter T uniquely optimizes on the paid channel by withholding advertising quotes from flagged high-risk borrowers and offering the pooled advertising quote*

$$r^{\mathrm{ad}} = \tilde{r} + \kappa \tag{8.54}$$

to unflagged non-customers.

Flagged high-risk borrowers are excluded because Assumption 4(i) makes them unprofitable at every rate they would accept. For unflagged borrowers, the quote $r^{\mathrm{ad}} = \tilde{r} + \kappa$ is accepted by all recipients by Lemma 1. Assumption 4(iv) rules out deviations to higher quotes that skim only captives. The advertising quote therefore extracts the shopper search-cost margin while preserving acceptance by both shoppers and captives.

It is useful to distinguish two approval objects. Borrower reservation values use the search-market approval probability q_H^e for high-risk borrowers and 1 for low-risk borrowers, because borrowers do not know the lender type when evaluating search. The funded pool of a technology adopter uses $q_H(\alpha)$, because approval on the paid channel is governed by T's own screening technology. Let

$$A_T = \mu_L + \mu_H q_H(\alpha) \tag{8.55}$$

denote the approval probability of the pool served by a technology adopter. Since $r^{\mathrm{ad}} = \tilde{r} + \kappa$ and

$D^e - D_T = e_N\Delta$, expected funded profit per impression, derived in Appendix I.5, is

$$\pi^{\text{ad}} = A_T\left(r^{\text{ad}} - f - D_T\right) = A_T\left(\kappa + m_0 + e_N\Delta\right) > 0. \tag{8.56}$$

For high-risk shoppers, the advertising quote lies below the search reservation rate by

$$\frac{\kappa(1-q_H^e)}{q_H^e}. \tag{8.57}$$

Thus, rejection risk creates an information rent for approved high-risk borrowers, while the technology lender's own screening technology determines which high-risk borrowers are funded.

The advertising platform charges a share of the lender's surplus from the slot:

$$c^{\text{ad}} = \eta\bar{S}, \qquad \eta \in [0,1). \tag{8.58}$$

The gross surplus from a non-customer advertising impression is derived in Appendix H.4 and equals

$$\bar{S} = A_T\left[\kappa + e_N(m_0+\Delta) + \lambda e_T m_0\right]. \tag{8.59}$$

Home-bank incumbency contracts

Lemma 4 (Home-bank incumbency rate). *For every $\gamma \in [0,1)$, the optimal home-bank rate is*

$$\bar{r} = \tilde{r} + \kappa^c. \tag{8.60}$$

Holding the search-market anchor $\tilde{r}$ fixed, the optimal home-bank markup is independent of advertising reach:

$$\frac{\partial \bar{r}}{\partial \gamma} = 0. \tag{8.61}$$

The home-bank contract prices the marginal captive borrower at her search reservation value. At this rate, low-search-cost borrowers search, while high-search-cost borrowers remain with the incumbent lender, as characterized in Lemma 1. Assumption 4(ii) guarantees that the residual home-bank book is profitable at the captive margin κ^c. Assumption 4(iii) makes it unprofitable to defend the entire book by lowering the rate to the shopper margin κ. The incumbent lender therefore concedes low-cost borrowers to search and targeted advertising, and extracts the captive margin from high-cost borrowers.

The result can be interpreted as the limit of home-bank rates approaching the low-risk high-cost borrower's search reservation value from below. The tie-breaking rule assigns this marginal borrower to the home bank. High-risk high-cost borrowers strictly remain captives because rejection risk

raises their effective search cost above κ^c.

For technology adopters, the retained home-bank mass is $h_T\lambda$ and the default rate is D_0. For non-adopters, targeted advertising by technology adopters skims part of the home-bank book. The retained mass is $h_N\lambda M(\gamma)$ and the default rate is D_0^γ. Since advertising reach is imperfect, $\gamma < 1$, the incumbency channel does not disappear.

Posted search rates

Lemma 5 (Posted search rates). *In the posted search market, lender type τ charges marginal cost plus the common differentiation markup:*

$$r_\tau^{\text{se}} = c_\tau + m_0. \tag{8.62}$$

The expected search rate is

$$\tilde{r} = e_N r_N^{\text{se}} + e_T r_T^{\text{se}} = r_T^{\text{se}} + e_N\Delta. \tag{8.63}$$

Because technology adopters have lower marginal mortgage costs, $c_T < c_N$, their posted search rate is lower:

$$r_T^{\text{se}} = c_T + m_0 < c_N + m_0 = r_N^{\text{se}}. \tag{8.64}$$

Thus, r_N^{se} is the highest posted search rate. The proof of the differentiated-product posted-rate subgame is in Appendix H.3.

Comparing the advertising quote with the adopter posted rate gives

$$\chi \equiv r^{\text{ad}} - r_T^{\text{se}} = \kappa + e_N\Delta > 0. \tag{8.65}$$

Targeted advertising therefore allows the technology adopter to capture the borrower's saved search cost, κ, and the rival-visibility component of its screening advantage, $e_N\Delta$.

8.4 Equilibrium

An equilibrium consists of posted search rates $\{r_N^{\text{se}}, r_T^{\text{se}}\}$, the expected search rate $\tilde{r}$, an advertising quote r^{ad}, a home-bank rate $\bar{r}$, underwriting rules $\{w_N, w_T\}$, an advertising policy a_T, borrower search and acceptance policies, and the induced origination volumes. In equilibrium, borrowers optimize given their information and search costs. Lenders optimize given borrower acceptance rules and screening technologies. Beliefs about search draws are consistent with $\{e_N, e_T\}$, and origination volumes are induced by the equilibrium policies.

Proposition 1 (Equilibrium characterization). *Under Assumptions 1–4, the equilibrium exists and is unique. It is characterized by forward substitution:*

$$\underbrace{\left\{r_N^{\text{se}}, r_T^{\text{se}}\right\}}_{\substack{\text{posted rates}\\ \text{marginal cost}+m_0}} \longrightarrow \underbrace{\tilde{r}}_{\substack{\text{search-rate}\\ \text{anchor}}} \longrightarrow \underbrace{\left\{\rho_i,\ r^{\text{ad}} = \tilde{r} + \kappa,\ \bar{r} = \tilde{r} + \kappa^c\right\}}_{\substack{\text{reservation, advertising,}\\ \text{and home-bank rates}}}$$
$$\longrightarrow \underbrace{\bar{\kappa} = \kappa^c}_{\substack{\text{search/stay}\\ \text{threshold}}} \longrightarrow \underbrace{\left\{Q_\tau^{\text{se}}, Q_T^{\text{ad}}, Q_\tau^0\right\}}_{\substack{\text{origination}\\ \text{volumes}}} \longrightarrow \underbrace{\left\{r_T^{\text{book}}, r_N^{\text{book}}\right\}}_{\substack{\text{average}\\ \text{book rates}}}. \tag{8.66}$$

Existence follows by construction. The posted-rate subgame pins down $\{r_N^{\text{se}}, r_T^{\text{se}}\}$ uniquely, which pins down $\tilde{r}$. Given $\tilde{r}$, Lemmas 1, 3, and 4 uniquely determine borrower reservation values, the advertising quote, the home-bank rate, and the search partition. These objects then uniquely determine origination volumes and book shares. Hence the equilibrium is unique.

The posted-rate subgame gives

$$r_\tau^{\text{se}} = c_\tau + m_0, \qquad c_\tau = f + D_\tau. \tag{8.67}$$

The expected search rate is

$$\tilde{r} = e_N r_N^{\text{se}} + e_T r_T^{\text{se}} = f + m_0 + D^e, \qquad D^e \equiv e_N \bar{D} + e_T D_T = \bar{D} - e_T \Delta. \tag{8.68}$$

By Lemma 3, the equilibrium advertising quote is

$$r^{\text{ad}} = \tilde{r} + \kappa. \tag{8.69}$$

By Lemma 4, the equilibrium home-bank rate is

$$\bar{r} = \tilde{r} + \kappa^c. \tag{8.70}$$

The underwriting rule rejects flagged high-risk borrowers in the paid and posted-search channels and funds approved borrowers. Borrowers accept and search according to Lemma 1.

Equilibrium volumes and book shares

Let

$$A_\tau = \mu_L + \mu_H q_H(z_\tau) \tag{8.71}$$

denote the approval probability at lender type τ. In particular,

$$A_T = \mu_L + \mu_H q_H(\alpha), \qquad A_N = \mu_L + \mu_H q_H(0). \tag{8.72}$$

Advertising by technology adopters poaches unflagged non-customers from the non-adopter home-bank pool. The paid-channel origination volume of technology adopters is therefore

$$Q_T^{\mathrm{ad}} = \gamma h_N A_T, \qquad Q_N^{\mathrm{ad}} \equiv 0. \tag{8.73}$$

For organic search, low-search-cost borrowers enter search unless they have already accepted an advertising quote. Define the residual approval mass of non-adopter home-bank borrowers who remain available for organic search and draw lender type τ as

$$A_\tau^R(\gamma) \equiv \mu_L(1-\gamma) + \mu_H\left(1-\gamma q_H(\alpha)\right) q_H(z_\tau). \tag{8.74}$$

This object is the residual post-advertising borrower mass from the non-adopter incumbency pool, multiplied by the approval probability of the lender type drawn in organic search. Low-risk borrowers remain with probability $1-\gamma$ and are always approved. High-risk borrowers remain with probability $1-\gamma q_H(\alpha)$ and are approved by the searched lender with probability $q_H(z_\tau)$. Moreover,

$$\frac{\partial A_\tau^R(\gamma)}{\partial \gamma} = -\mu_L - \mu_H q_H(\alpha) q_H(z_\tau) < 0. \tag{8.75}$$

Advertising therefore reduces the residual organic-search pool.

Organic-search originations are

$$Q_\tau^{\mathrm{se}} = (1-\lambda) e_\tau \left[h_T A_\tau + h_N A_\tau^R(\gamma)\right]. \tag{8.76}$$

The first term inside brackets is the contribution of low-search-cost borrowers whose home bank is a technology adopter. These borrowers are not targeted by their own lender's advertising. The second term is the residual low-search-cost pool from non-adopter home banks after advertising poaching.

Home-bank originations are

$$Q_T^0 = h_T \lambda, \qquad Q_N^0 = h_N \lambda M(\gamma), \tag{8.77}$$

where

$$M(\gamma) = \mu_L(1-\gamma) + \mu_H\left(1-\gamma q_H(\alpha)\right). \tag{8.78}$$

Thus, the non-adopter home-bank book is smaller and riskier after advertising poaching. Its default probability is

$$D_0^\gamma = \frac{\mu_L(1-\gamma) d_L + \mu_H\left(1-\gamma q_H(\alpha)\right) d_H}{M(\gamma)}. \tag{8.79}$$

Let total funded originations by lender type be

$$Q_T = Q_T^{\text{se}} + Q_T^{\text{ad}} + Q_T^0, \qquad Q_N = Q_N^{\text{se}} + Q_N^0. \tag{8.80}$$

Define technology-adopter book shares as

$$\hat{\omega}_T^{\text{se}} = \frac{Q_T^{\text{se}}}{Q_T}, \qquad \hat{\omega}_T^{\text{ad}} = \frac{Q_T^{\text{ad}}}{Q_T}, \qquad \hat{\omega}_T^0 = \frac{Q_T^0}{Q_T}. \tag{8.81}$$

Define non-adopter book shares as

$$\hat{\omega}_N^{\text{se}} = \frac{Q_N^{\text{se}}}{Q_N}, \qquad \hat{\omega}_N^0 = \frac{Q_N^0}{Q_N}. \tag{8.82}$$

These shares sum to one within each lender type:

$$\hat{\omega}_T^{\text{se}} + \hat{\omega}_T^{\text{ad}} + \hat{\omega}_T^0 = 1, \qquad \hat{\omega}_N^{\text{se}} + \hat{\omega}_N^0 = 1. \tag{8.83}$$

The paid-channel share of technology-adopter acquisitions is

$$\omega_T^{\text{ad}} \equiv \frac{Q_T^{\text{ad}}}{Q_T^{\text{ad}} + Q_T^{\text{se}}}, \qquad \omega_N^{\text{ad}} \equiv 0. \tag{8.84}$$

Since Q_T^{ad} is increasing in γ and Q_T^{se} is decreasing in γ through $A_T^R(\gamma)$, ω_T^{ad} is strictly increasing in advertising reach. The formal derivative is given in Appendix I.8.

Average book rates

The average charged rate in the technology-adopter book is

$$r_T^{\text{book}} = \hat{\omega}_T^{\text{se}} r_T^{\text{se}} + \hat{\omega}_T^{\text{ad}} r^{\text{ad}} + \hat{\omega}_T^0 \bar{r}. \tag{8.85}$$

Using $r_T^{\text{se}} = c_T + m_0$, $r^{\text{ad}} = \tilde{r} + \kappa$, $\bar{r} = \tilde{r} + \kappa^c$, and $D^e - D_T = e_N \Delta$, this becomes

$$r_T^{\text{book}} = c_T + m_0 + \hat{\omega}_T^{\text{ad}} \left(\kappa + e_N \Delta\right) + \hat{\omega}_T^0 \left(\kappa^c + e_N \Delta\right). \tag{8.86}$$

Equivalently,

$$r_T^{\text{book}} = \underbrace{c_T}_{\text{marginal cost}} + \underbrace{m_0}_{\text{differentiation margin}} + \underbrace{\hat{\omega}_T^{\text{ad}} \left(\kappa + e_N \Delta\right)}_{\text{paid-channel markup}} + \underbrace{\hat{\omega}_T^0 \left(\kappa^c + e_N \Delta\right)}_{\text{home-bank markup}}. \tag{8.87}$$

The average charged rate in the non-adopter book is

$$r_N^{\text{book}} = \hat{\omega}_N^{\text{se}} r_N^{\text{se}} + \hat{\omega}_N^0 \bar{r}. \tag{8.88}$$

Using $r_N^{\text{se}} = c_N + m_0$, $c_N = f + \bar{D}$, and $D^e - \bar{D} = -e_T\Delta$, this becomes

$$r_N^{\text{book}} = c_N + m_0 + \hat{\omega}_N^0 \left(\kappa^c - e_T\Delta\right). \tag{8.89}$$

Equivalently,

$$r_N^{\text{book}} = \underbrace{c_N}_{\text{marginal cost}} + \underbrace{m_0}_{\text{differentiation margin}} + \underbrace{\hat{\omega}_N^0 \left(\kappa^c - e_T\Delta\right)}_{\text{home-bank markup adjusted by the search-market anchor}}. \tag{8.90}$$

Rate decomposition

The average technology-adopter book rate can be written relative to the no-data benchmark $f + \bar{D} + m_0$ as

$$r_T^{\text{book}} = \underbrace{f + \bar{D} + m_0}_{\text{no-data benchmark}} - \underbrace{\left[1 - e_N\left(\hat{\omega}_T^{\text{ad}} + \hat{\omega}_T^0\right)\right]\Delta}_{\text{screening pass-through}} + \underbrace{\hat{\omega}_T^{\text{ad}}\kappa + \hat{\omega}_T^0\kappa^c}_{\text{search-cost extraction}}. \tag{8.91}$$

The first term is the rate that would prevail absent the technology adopter's screening advantage. The second term captures the pass-through of lower default costs. The third term captures the extraction of borrower search costs through targeted advertising and home-bank pricing.

The average non-adopter book rate is

$$r_N^{\text{book}} = \underbrace{f + \bar{D} + m_0}_{\text{no-data benchmark}} - \underbrace{\hat{\omega}_N^0 e_T\Delta}_{\text{rival screening via the search anchor}} + \underbrace{\hat{\omega}_N^0\kappa^c}_{\text{home-bank extraction}}. \tag{8.92}$$

Although non-adopters do not use online-collected data, their home-bank rate is tied to the expected search rate $\tilde{r}$. Therefore, part of the technology adopter's screening advantage is passed through to non-adopter customers through the market-wide search anchor.

Technology adopters need not offer lower average rates

Technology adopters have lower posted search rates because $c_T < c_N$. However, they need not have lower average book rates. The difference in average book rates is

$$r_T^{\text{book}} - r_N^{\text{book}} = -\left(e_N\hat{\omega}_T^{\text{se}} + e_T\hat{\omega}_N^{\text{se}}\right)\Delta + \hat{\omega}_T^{\text{ad}}\kappa + \left(\hat{\omega}_T^0 - \hat{\omega}_N^0\right)\kappa^c. \tag{8.93}$$

Channel	Information set	Rate	Wedge over channel cost
Posted search, τ	Credit-score bucket	$r_\tau^{\text{se}} = c_\tau + m_0$	m_0
Paid, unflagged T	Bucket, intent, flag	$r^{\text{ad}} = \tilde{r} + \kappa$	$m_0 + \kappa + e_N\Delta$
Paid, flagged T	Bucket, intent, flag	No usable quote	Excluded
Home bank, T	Relationship	$\bar{r} = \tilde{r} + \kappa^c$	$m_0 + \kappa^c + (D^e - D_0)$
Home bank, N	Relationship, residual book	$\bar{r} = \tilde{r} + \kappa^c$	$m_0 + \kappa^c + (D^e - D_0^\gamma)$

Table 8.1. Equilibrium rates by channel. The paid channel is available only to technology adopters. Wedges are measured relative to the default cost of the borrower pool served in each channel.

The first term is the net screening pass-through advantage of technology adopters. The remaining terms are net extraction markups from targeted advertising and home-bank pricing. Hence,

$$r_T^{\text{book}} - r_N^{\text{book}} \gtrless 0. \tag{8.94}$$

Technology adopters can offer lower posted search rates and still charge higher average rates on their overall book if the extraction markups from advertising and incumbency are large enough.

Summary

The equilibrium has two opposing forces. Online-collected data improves screening, lowers default costs among approved borrowers, and reduces posted search rates. At the same time, the same data allows technology adopters to target borrowers before they enter costly search. Advertising and home-bank pricing allow lenders to capture part of borrowers' search costs through higher non-posted rates. Approved high-risk borrowers receive information rents because rejection risk raises their search reservation values, while lenders use screening to exclude high-risk borrowers who are identified before funding.

8.5 Privacy Laws

Privacy laws restrict the flow of online-collected data across websites. In the model, this weakens both components of the technology used by adopters. The advertising component is hit more directly because it relies on cross-site tracking and real-time targeting. The screening component is also affected, but less strongly, because lenders may still use coarser or aggregated information.

Assumption 5. *A privacy law reduces the effectiveness of online-collected data for technology adopters according to*

$$\gamma^+ = (1 - \delta_\gamma)\gamma, \qquad \alpha^+ = (1 - \delta_\alpha)\alpha, \qquad 1 > \delta_\gamma > \delta_\alpha \geq 0. \tag{8.95}$$

The law therefore reduces advertising reach more than screening precision. I study the new-origination equilibrium after the law. Because the home-bank contract is re-solved as $\bar{r}^+ = \tilde{r}^+ + \kappa^c$, the search threshold remains $\bar{r}^+ - \tilde{r}^+ = \kappa^c$, so the search partition characterized in Lemma 1 is unchanged.

I also impose that the reach loss dominates the mechanical increase in approvals induced by weaker screening:

$$\delta_\gamma > \varepsilon_A \delta_\alpha, \qquad \varepsilon_A \equiv \frac{\mu_H q\, p'(\alpha)\alpha}{A_T}. \tag{8.96}$$

Here q is the baseline high-risk misclassification probability. The use of q, rather than $q_H(\alpha)$, follows from the derivative of the screening technology: since

$$q_H(\alpha) = q\,(1 - p(\alpha))\,, \tag{8.97}$$

a small reduction in α raises high-risk approval at technology adopters by

$$dq_H(\alpha) = q\, p'(\alpha)\alpha\, \delta_\alpha. \tag{8.98}$$

Thus,

$$dA_T = \mu_H q\, p'(\alpha)\alpha\, \delta_\alpha = A_T \varepsilon_A \delta_\alpha. \tag{8.99}$$

To first order,

$$d(\gamma A_T) = A_T\, d\gamma + \gamma\, dA_T = \gamma A_T \left(\varepsilon_A \delta_\alpha - \delta_\gamma\right) < 0. \tag{8.100}$$

Let X^+ denote the post-law value of object X, and let $dX \equiv X^+ - X$. Since $\alpha^+ < \alpha$, screening deteriorates:

$$q_H(\alpha^+) > q_H(\alpha). \tag{8.101}$$

The funded pool at technology adopters becomes riskier:

$$dD_T \equiv D_T^+ - D_T > 0, \qquad d\Delta \equiv \Delta^+ - \Delta = -dD_T < 0. \tag{8.102}$$

Conditional on dD_T, the changes in the technology-adopter posted rate and in the search-market anchor are

$$dr_T^{\mathrm{se}} = dD_T, \qquad d\tilde{r} = e_T\, dD_T. \tag{8.103}$$

Proposition 2. *Under Assumptions 1–5, privacy laws weaken screening, reduce targeted advertising, and reallocate borrowers across equilibrium channels.*

First, technology adopters reject fewer high-risk borrowers:

$$R_T^+ < R_T. \tag{8.104}$$

This improves access for high-risk borrowers but raises default costs among approved borrowers. Hence, r_T^{se} rises by dD_T, and the search-market anchor $\tilde{r}$ rises by $e_T\, dD_T$.

Second, paid-channel volume falls. Since

$$Q_T^{\text{ad}} = \gamma h_N A_T, \qquad A_T = \mu_L + \mu_H q_H(\alpha), \tag{8.105}$$

Assumption 5 implies

$$dQ_T^{\text{ad}} = h_N\, d(\gamma A_T) < 0. \tag{8.106}$$

The technology-adopter book therefore shifts away from the high-markup advertising channel and toward organic search. In particular, Appendix H.2 shows that $dQ_T^{\text{se}} > 0$, so

$$d\omega_T^{\text{ad}} < 0. \tag{8.107}$$

Third, the non-adopter home-bank book expands and improves in quality. The non-adopter home-bank volume is

$$Q_N^0 = h_N \lambda M(\gamma), \tag{8.108}$$

where

$$M(\gamma) = \mu_L(1-\gamma) + \mu_H\,(1 - \gamma q_H(\alpha)) = 1 - \gamma A_T. \tag{8.109}$$

Therefore,

$$dQ_N^0 = -h_N \lambda\, d(\gamma A_T) > 0. \tag{8.110}$$

The residual non-adopter book also becomes safer. Since $\partial D_0^\gamma / \partial \gamma > 0$ and $\partial D_0^\gamma / \partial q_H(\alpha) < 0$, the decline in γ and the increase in $q_H(\alpha)$ imply

$$dD_0^\gamma < 0. \tag{8.111}$$

Fourth, the home-bank rate rises in levels but not in markup. Lemma 4 implies

$$\bar{r} - \tilde{r} = \kappa^c.$$

The privacy law raises the search-market anchor by $d\tilde{r} = e_T dD_T$. Re-solving the home-bank

problem after the law gives

$$\bar{r}^{+} = \tilde{r}^{+} + \kappa^{c}, \qquad d\bar{r} = e_T dD_T > 0. \tag{8.112}$$

The search threshold remains unchanged:

$$\bar{r}^{+} - \tilde{r}^{+} = \kappa^{c}. \tag{8.113}$$

Finally, measured rates at technology adopters can fall even when borrower costs rise. Let ω_T^{ad} denote the paid-channel share of technology-adopter acquisitions, and define

$$\chi \equiv r^{\mathrm{ad}} - r_T^{\mathrm{se}} = \kappa + e_N \Delta. \tag{8.114}$$

The average acquisition rate of technology adopters satisfies

$$dr_T^{\mathrm{new}} = \left(1 - e_N \omega_T^{\mathrm{ad}}\right) dD_T + \chi \, d\omega_T^{\mathrm{ad}}. \tag{8.115}$$

The first term is positive because weaker screening raises funding costs. The second term is negative because the high-markup advertising channel shrinks. Thus,

$$dr_T^{\mathrm{new}} < 0 \quad \Longleftrightarrow \quad \chi \left| d\omega_T^{\mathrm{ad}} \right| > \left(1 - e_N \omega_T^{\mathrm{ad}}\right) dD_T. \tag{8.116}$$

This is the composition effect: measured rates can fall because the high-markup advertising channel shrinks, not because borrowers necessarily face lower costs.

The advertising rate illustrates the incidence of the screening advantage. The advertising quote prices off the borrower's search outside option:

$$r^{\mathrm{ad}} = \tilde{r} + \kappa = f + m_0 + \bar{D} - e_T \Delta + \kappa. \tag{8.117}$$

Relative to the no-data benchmark $f + m_0 + \bar{D}$, advertising borrowers receive pass-through equal to $e_T \Delta$. The lender does not pass through the full screening advantage, because relative to its own posted rate,

$$r^{\mathrm{ad}} - r_T^{\mathrm{se}} = \kappa + e_N \Delta. \tag{8.118}$$

Thus, the lender captures the search-cost wedge κ and the non-adopter visibility component $e_N \Delta$.

After the privacy law,

$$r^{\mathrm{ad},+} = f + m_0 + \bar{D} - e_T \Delta^{+} + \kappa, \tag{8.119}$$

and

$$r^{\mathrm{ad},+} - r_T^{\mathrm{se},+} = \kappa + e_N \Delta^+. \tag{8.120}$$

Hence, if $\Delta^+ > 0$, advertising borrowers still receive partial screening pass-through. Privacy laws reduce this pass-through but do not eliminate it unless the screening advantage is fully destroyed. Advertising prices off the borrower's search outside option, so some screening advantage passes through via the anchor, while the lender captures the search-cost wedge and the non-adopter visibility component.

The surplus incidence follows directly from these channel effects. Borrower surplus is ambiguous. High-risk borrowers benefit from higher approval probabilities, while low-risk borrowers face higher posted rates because screening becomes less precise. Borrowers displaced from the advertising channel lose the search-cost savings embedded in the advertising quote: shoppers must search directly, while captives return to the home-bank contract. In the post-law equilibrium,

$$\bar{r}^+ = \tilde{r}^+ + \kappa^c, \tag{8.121}$$

so released captives face a higher home-bank rate through both the loss of the advertising discount and the increase in the search-market anchor.

Non-adopters gain in the home-bank channel because fewer borrowers are poached and the residual incumbency book becomes safer:

$$dQ_N^0 > 0, \qquad dD_0^\gamma < 0. \tag{8.122}$$

Technology adopters lose paid-channel volume, although the total effect on their profits is ambiguous because lower advertising expenditure and higher home-bank pass-through partly offset the loss of advertising originations. Advertising platforms lose reach. Their revenue changes by

$$d\Pi_{\mathrm{plat}} = \eta h_N \left(\bar{S}\, d\gamma + \gamma\, d\bar{S} \right), \tag{8.123}$$

and falls whenever the reach loss dominates any increase in surplus per remaining impression. Thus, privacy laws can lower measured rates at technology adopters while increasing the cost of credit for borrowers displaced from targeted advertising. The aggregate welfare effect depends on the relative strength of access gains, default-cost pass-through, search-cost losses, and platform revenue losses.

9 Model Estimation and Calibration

This section connects the reduced-form evidence to the structural model. The empirical analysis identifies how privacy laws change equilibrium quantities, approval decisions, default risk, and rates. The model uses these moments to separate the value of online-collected data in borrower acquisition from its value in screening. I implement the model by constructing borrower-level channel weights from survey evidence and then inverting the equilibrium restrictions to recover the key primitives.

The procedure has two steps. First, I use the National Survey of Mortgage Originations (NSMO) to construct fractional assignments of borrowers to the model's three acquisition channels: home-bank incumbency, organic search, and passive outside quotes. Second, I combine these channel assignments with HMDA, Equifax, and lender-type information to recover the model objects governing organic visibility, advertising reach, search costs, screening, and default-cost pass-through. The Sun–Abraham estimates identify the policy-induced changes in these objects, not the pre-law levels.

9.1 Calibration

Let π_i^M denote the probability that borrower i considers more than one lender. Let π_i^P denote the probability that borrower i considers more than one lender for price reasons, and let π_i^F denote the probability that borrower i considers more than one lender because of fear of rejection:

$$\begin{aligned}
\pi_i^M &= \Pr_i(\text{considers more than one lender}),\\
\pi_i^P &= \Pr_i(\text{considers more than one lender for price}),\\
\pi_i^F &= \Pr_i(\text{considers more than one lender for fear of rejection}).
\end{aligned}$$

The probabilities are estimated as functions of borrower observables, including income, demographics, credit score, and debt-to-income ratio. If the motive probabilities are reported conditional on considering multiple lenders, I first convert them to unconditional probabilities, as described in Appendix J.

I classify active search as the component of multi-lender consideration explained by price search or rejection-fear search:

$$\pi_i^S = \min\{\pi_i^M, \pi_i^P + \pi_i^F\}. \tag{9.1}$$

The resulting channel weights are

$$w_i^0 = 1 - \pi_i^M, \qquad w_i^{\text{se}} = \pi_i^S, \qquad w_i^{\text{ad}} = \pi_i^M - \pi_i^S. \tag{9.2}$$

The three weights form an exhaustive fractional assignment:

$$w_i^0 + w_i^{\text{se}} + w_i^{\text{ad}} = 1, \qquad w_i^k \geq 0 \quad \text{for } k \in \{0, \text{se}, \text{ad}\}. \tag{9.3}$$

Economically, the construction treats stated price-search and rejection-fear motives as the active-search component of multi-lender consideration, total multi-lender consideration as the upper bound on outside-quote exposure, and the remaining gap as passive exposure to an outside quote.

The identifying assumption is that, conditional on borrower observables, lender type, broker use, and referral channels, multi-lender consideration not explained by active price search or rejection-fear search proxies for passive exposure to an outside quote. In the model, this passive outside quote is generated by the targeted-advertising channel.

I validate this interpretation using independent variation in advertising-technology intensity. Let Z_{mt}^{ad} denote the ad-tech or data-collection intensity used in the reduced-form analysis, measured at the market-by-time level. I test whether markets with greater ad-tech intensity have larger technology-lender passive-quote mass:

$$\bar{w}_{mt}^{\text{ad},T} = \rho Z_{mt}^{\text{ad}} + X_{mt}'\Gamma + \eta_m + \eta_t + u_{mt}, \tag{9.4}$$

where

$$\bar{w}_{mt}^{\text{ad},T} = \frac{\sum_{i \in (m,t)} \varpi_i \mathbf{1}\{\tau_i = T\} w_i^{\text{ad}}}{\sum_{i \in (m,t)} \varpi_i \mathbf{1}\{\tau_i = T\}}. \tag{9.5}$$

The model predicts $\rho > 0$. This test is not used to construct the calibration; it validates the interpretation of the passive residual as advertising exposure. Standard errors are clustered at the market level. I also report the corresponding placebo for non-technology lenders, for whom the model predicts no systematic relationship between passive residual mass and ad-tech intensity.

Let r_i denote the observed note rate, let $\tau_i \in \{T, N\}$ denote lender type, and let ϖ_i denote the sampling weight. Table 9.1 shows how the channel weights map observed originations into the rate and quantity moments used in the calibration.

Table 9.1. Channel Weights, Rates, and Quantities

Cell	Economic content	Probability weight	Rate moment	Quantity moment
Home bank / one lender	Borrower considers one lender. In the model this is the incumbent or fallback loan. This row constructs the home-bank rate and home-bank quantities.	$w_i^0 = 1 - \pi_i^M$	$\bar{r} = \frac{\sum_i \varpi_i w_i^0 r_i}{\sum_i \varpi_i w_i^0}$	$Q_\tau^0 = \sum_i \varpi_i \mathbf{1}\{\tau_i = \tau\} w_i^0$ $Q^0 = \sum_\tau Q_\tau^0$
Organic active search	Borrower considers multiple lenders because of price search or rejection-fear search. This row constructs lender-specific organic-search rates and quantities.	$w_i^{\mathrm{se}} = \pi_i^S$	$r_\tau^{\mathrm{se}} = \frac{\sum_i \varpi_i \mathbf{1}\{\tau_i = \tau\} w_i^{\mathrm{se}} r_i}{\sum_i \varpi_i \mathbf{1}\{\tau_i = \tau\} w_i^{\mathrm{se}}}$	$Q_\tau^{\mathrm{se}} = \sum_i \varpi_i \mathbf{1}\{\tau_i = \tau\} w_i^{\mathrm{se}}$
Passive outside quote / advertising	Borrower considers multiple lenders, but the multi-lender event is not explained by active price search or rejection-fear search. This residual cell constructs the passive-quote rate and paid-channel quantity for technology lenders.	$w_i^{\mathrm{ad}} = \pi_i^M - \pi_i^S$	$r^{\mathrm{ad}} = \frac{\sum_i \varpi_i \mathbf{1}\{\tau_i = T\} w_i^{\mathrm{ad}} r_i}{\sum_i \varpi_i \mathbf{1}\{\tau_i = T\} w_i^{\mathrm{ad}}}$	$Q_T^{\mathrm{ad}} = \sum_i \varpi_i \mathbf{1}\{\tau_i = T\} w_i^{\mathrm{ad}}$

The weights form an exhaustive fractional assignment. The row-specific rates feed $\bar{r}$, r_T^{se}, r_N^{se}, r^{ad}, and $\tilde{r}$. The row-specific quantities feed e_T, ω_T^{ad}, γ, λ, and the return-home tests. Passive residual mass ending at non-technology lenders is not interpreted as non-technology advertising and is reported as a diagnostic. If price and fear motives overlap, the appendix replaces $\pi_i^P + \pi_i^F$ with $\pi_i^P + \pi_i^F - \pi_i^{PF}$ and reports the sensitivity of the calibrated objects.

The remaining parameters are recovered recursively from equilibrium restrictions. Table 9.2 reports the main parameter map. A full recursive map, probability audit, advertising-intensity validation, overidentifying restrictions, and inference procedure are reported in Appendix J.

The Sun–Abraham estimates enter the structural inversion by quantifying how privacy laws

Table 9.2. Main Parameter Map

Parameter	Meaning	Estimator / moment equation	Source
A_T, A_N	Approval probabilities	$A_\tau = 1 - R_\tau$, checked against $A_\tau = \mu_L + \mu_H q_H(z_\tau)$	HMDA
q	Baseline high-risk survival probability	$q = 1 - R_N/\mu_H$	HMDA, Equifax
$p(\alpha)$	Cross-site detection probability	$p(\alpha) = (R_T - R_N)/(\mu_H q)$	HMDA
Δ	Screening cost advantage	$\Delta = D_N - D_T$	Equifax
e_T	Organic visibility of technology lenders	$e_T = Q_T^{\text{se}}/(Q_T^{\text{se}} + Q_N^{\text{se}})$	Quantity
e_N	Organic visibility of non-technology lenders	$e_N = 1 - e_T$	Implied
ω_T^{ad}	Paid/passive-quote share of technology-lender acquisitions	$\omega_T^{\text{ad}} = Q_T^{\text{ad}}/(Q_T^{\text{ad}} + Q_T^{\text{se}})$	Quantity
h_T, h_N	Home-bank shares	Pre-period relationship-lender shares; $h_N = 1 - h_T$	Equifax / servicer
γ	Targeting reach	$\gamma = Q_T^{\text{ad}}/(h_N A_T)$	Quantity, HMDA
λ	Captive share	$\lambda = Q^0/[h_T + h_N M(\gamma)] = Q^0/(1 - h_N \gamma A_T)$	Quantity
$\tilde{r}$	Organic-search anchor	$\tilde{r} = e_N r_N^{\text{se}} + e_T r_T^{\text{se}}$	Rate, quantity
κ	Shopper advertising wedge	$\kappa = r^{\text{ad}} - \tilde{r}$	Rate
κ^c	Captive home-bank wedge	$\kappa^c = \bar{r} - \tilde{r}$	Rate
δ_α	Law-induced screening loss	Solves the Sun–Abraham effects on R_T and D_T	Sun–Abraham
δ_γ	Law-induced reach loss	Solves the Sun–Abraham effects on Q_T^{ad} and Q_N^0, given δ_α	Sun–Abraham

Rows are ordered recursively. Approval and default moments identify A_T, A_N, q, $p(\alpha)$, and Δ. Quantities from Table 9.1 identify e_T, ω_T^{ad}, γ, and λ. Rates from Table 9.1 identify $\tilde{r}$, κ, and κ^c. The screening shock δ_α is recovered before the reach shock δ_γ because the quantity response depends on the approval effect induced by weaker screening. The Sun–Abraham estimates identify policy-induced changes in reach and screening, not the pre-law levels.

move the equilibrium moments:

$\widehat{\theta}^{Q_T^{\text{ad}}} < 0 \quad \Rightarrow$ targeting reach falls, $\delta_\gamma > 0$,

$\widehat{\theta}^{Q_N^0} > 0 \quad \Rightarrow$ borrowers displaced from advertising return to incumbent or home-bank channels,

$\widehat{\theta}^{R_T} < 0 \quad \Rightarrow$ screening weakens and high-risk approvals rise,

$\widehat{\theta}^{D_T} > 0 \quad \Rightarrow$ screening losses raise default risk among funded technology-lender loans,

$\widehat{\theta}^{r}_{\text{stock}} = 0 \quad \Rightarrow$ contract-stock placebo: predetermined incumbent rates do not reprice,

$\widehat{\theta}^{r_T} < 0 \quad \Rightarrow$ measured technology-lender rates fall if the paid-share decline dominates the screening-cost incre

The home-bank placebo is estimated on predetermined incumbent relationships or locked contracts. It is not imposed as a new-origination equilibrium restriction. In the model's new-origination equilibrium, the home-bank rate moves with the search-market anchor.

9.2 Quantitative Predictions

The calibrated model translates the reduced-form effects into welfare and industry-incidence predictions. The model separates three forces: the loss of targeted advertising, the deterioration of screening, and the reallocation of borrowers across home-bank, search, and passive-quote channels. This separation allows the counterfactuals to measure whether privacy laws lower observed rates by reducing lender rents or instead by moving borrowers out of high-markup channels.

The counterfactual analysis considers two classes of policies. First, I evaluate privacy rules that restrict only targeted advertising while preserving lenders' ability to use first-party or compliance-protected screening information under the Gramm–Leach–Bliley Act. Second, I re-estimate the model at the state level using the Sun–Abraham effects and compare how alternative privacy-law designs affect borrower surplus, lender profits, platform revenues, approval probabilities, and default-cost pass-through.

10 Conclusion

This paper shows that online-collected consumer data shapes mortgage competition through two opposing channels. The information helps lenders assess who is likely to repay, but it also helps them reach prospective borrowers before search is complete, enter their consideration sets, and price against the value of continuing to shop. The two uses push rates in opposite directions, making the effect of restricting the data ambiguous ex ante. Exploiting the staggered enforcement of state privacy laws, I find evidence that both channels matter. Enforcement weakens screening, but it also weakens targeted borrower acquisition. Among borrowers with high predicted search propensity,

rates fall even as the originated pool becomes riskier, implying that the loss of acquisition rents dominates the increase in expected credit losses. Privacy rules were not designed as credit-market policy, but by changing which lenders borrowers encounter, they also change how lenders compete.

The estimates move together across the margins implicated by the mechanism. Following enforcement, the share of high predicted-search-propensity borrowers in adopting lenders' originations falls by 0.85 percentage points and rises at non-adopting lenders. The roughly offsetting movements are consistent with shoppers reallocating across lender types rather than leaving the mortgage market. At adopting lenders, rates fall by 14 basis points for borrowers with high predicted search propensity, while the corresponding effect for borrowers with low predicted search propensity is small and statistically indistinguishable from zero. Screening deteriorates at the same time: rejection rates fall by 0.22 percentage points, and one-year financial distress among originated borrowers rises by 2.09 percentage points from a 2.7 percent baseline. Lenders' conduct points in the same direction, as search-advertising expenditures fall sharply among those most exposed to enforcement. Taken together, these responses help distinguish the loss of a screening signal from the loss of a targeting advantage: weaker screening expands approval of riskier borrowers regardless of whether they shop, whereas weaker targeting reduces rents earned specifically from borrowers who would otherwise continue shopping.

The central idea is that the same information relaxes one friction while monetizing another. It mitigates adverse selection within credit-score buckets, but it also converts borrowers' search costs into lender rents by placing a lender in front of the borrower before competing offers arrive. Lenders are therefore not passive price-posters waiting for applicants. They influence the sequence and stopping point of search and price against the borrower's continuation value. This helps explain why consideration sets remain small and comparison shopping remains limited even as mortgage search moves online. Consumer data is not only a productive asset that improves risk assessment; it is also a strategic asset that shapes borrower allocation, prices, and lender market power.

The model makes the tension precise. Adopters' screening advantage lowers their marginal costs and posted search rates, yet borrowers reached through advertising pay an acquisition markup, so data-using lenders can offer lower posted rates while charging higher average rates across their mortgage books. Asymmetric adoption need not lead data-rich lenders to displace data-poor lenders, because non-adopters retain borrowers with high search costs while their continued presence raises the search-market benchmark against which adopters price targeted offers. Restricting the data cuts both ways: it worsens screening, raising expected credit losses, and it erodes the acquisition rents lenders extract from borrowers' search costs. Where the second force dominates, as it does for likely shoppers in my estimates, restricting the data transfers surplus from lenders to those borrowers. The aggregate effect is less clear, because that transfer must be weighed against weaker screening, greater repayment risk, and changes in credit access.

Three limits define the scope of these results, and each points to a direction for future research. Search is a propensity predicted from survey behavior and predetermined borrower characteristics rather than observed directly, so the heterogeneity is by predicted search propensity and, under classification error, conservative; linking household search behavior to administrative mortgage records would measure this margin directly. Advertising expenditure is measured at the lender rather than the borrower level, and rests on the small set of lenders that advertise heavily in search; data linking individual applications to browsing, advertising exposure, and lender outreach would identify which borrowers each campaign reaches. Finally, the welfare decomposition inherits the model's assumptions, and the mechanism invites testing wherever screening, advertising, and search meet. Refinancing, home-equity credit, auto lending, and credit cards differ in the relative weight of underwriting risk, urgency, and comparison shopping, and so offer natural settings for asking when data-driven acquisition improves matching and when it becomes a source of market power. More broadly, when firms use behavioral data both to predict consumer value and to compete for attention, regulating the data is also regulating competition.

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.

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

Chen, J. and J. Roth (2024). Logs with zeros? some problems and solutions. *The Quarterly Journal of Economics 139*(2), 891–936.

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.

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.

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*.

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 2, explaining in practical terms what they do and why they matter for mortgage lending, and documenting the lender characteristics associated with adoption.

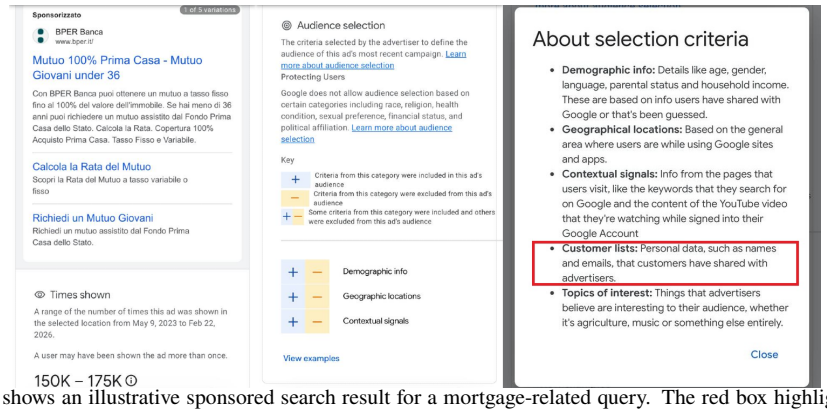
A.1 Search Advertising and Borrower Acquisition

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.

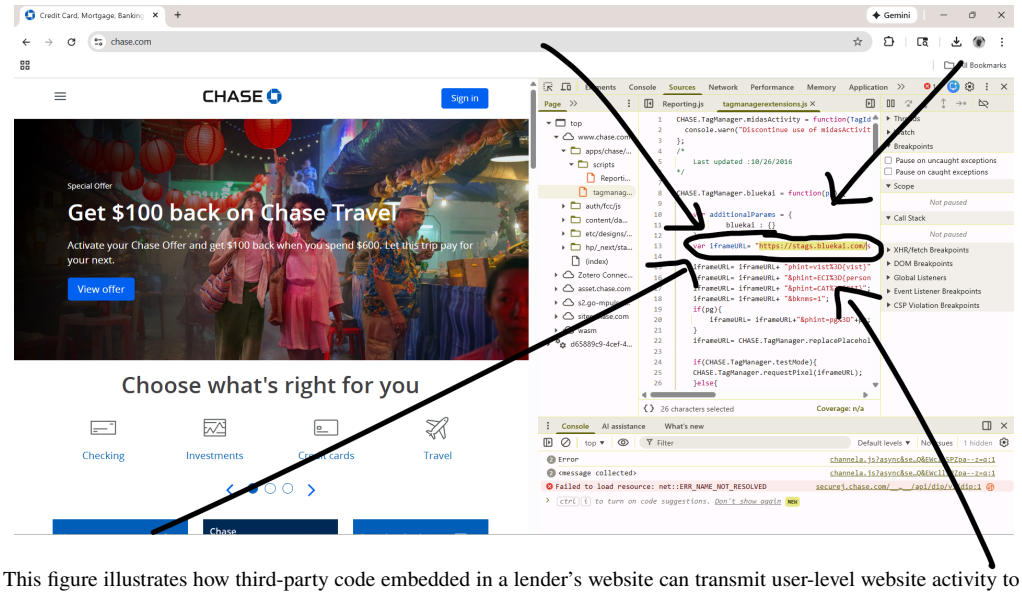
A.2 Website Data Collection and Transmission

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.

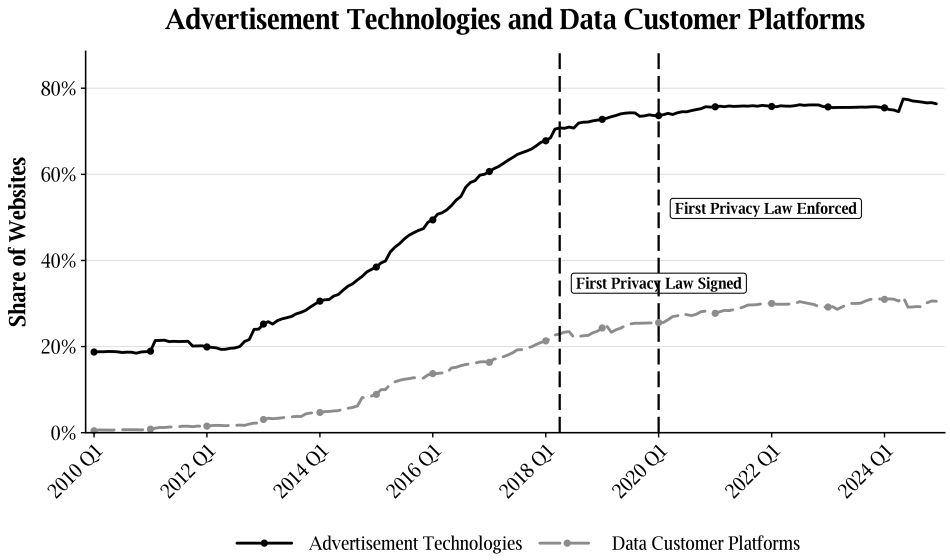
This distinction is consistent with recent evidence from asset management. Zanotti (2025) shows that financial firms use website technologies such as Google Analytics and A/B testing tools to collect information on potential customers, including browsing behavior, demographics, product interest, acquisition 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.

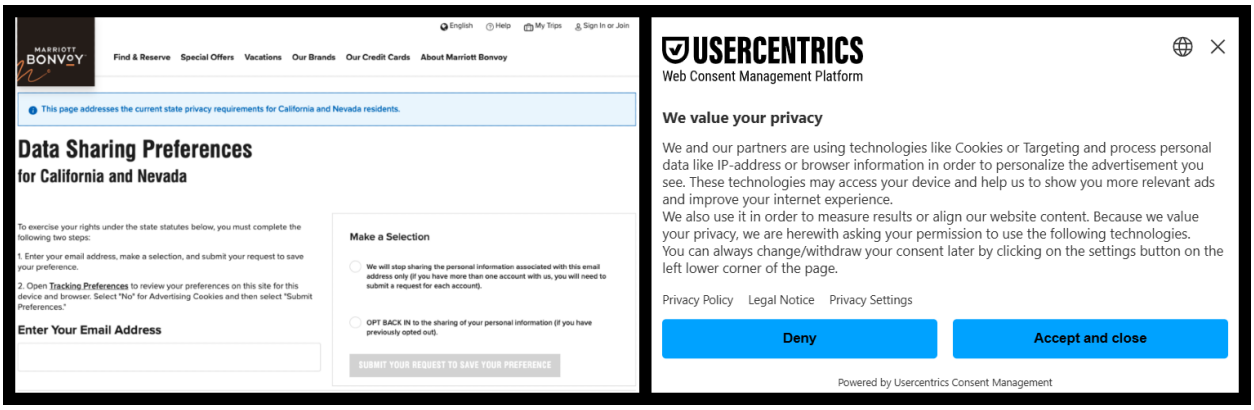
Recent enforcement materials and private litigation make this boundary visible by focusing on website tracking, advertising-platform transfers, opt-out mechanisms, and global opt-out signals. I do not use these cases as evidence of liability. They are useful only to illustrate that website and advertising 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 3.1. 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 (4.3); 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.

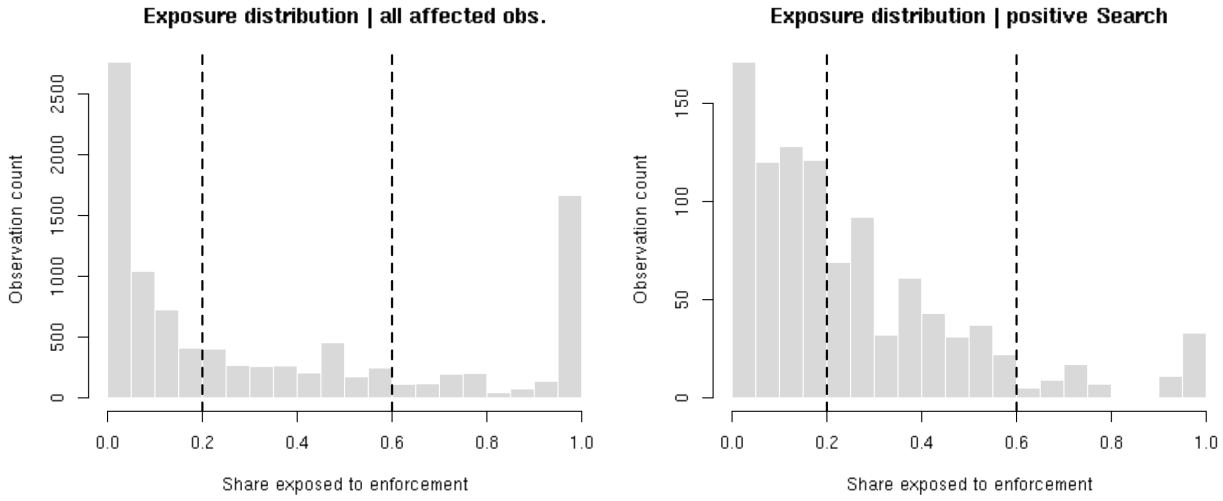
Table E.3. Imputation onto the Equifax–HMDA sample: coverage and the marginal cross-check.

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



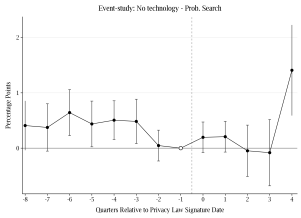
Notes: This figure plots the distribution of lenders' privacy-law *exposure share*, the shift-share regressor used to group lenders in the advertising-spend regression (Table 6.2). For a given lender and month, exposure is the fraction of the lender's pre-period (2018–2019) customer base located in states 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.

G Appendix – Robustness

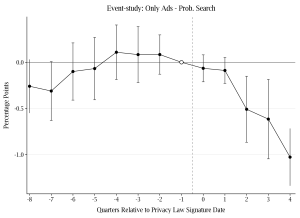
G.1 Signature Parallel Trends

Figure G.1. Parallel Trends, Signature

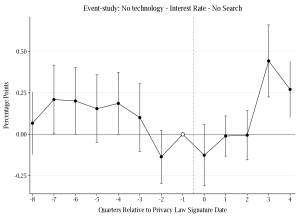
(a): Search-for-Price – No Technology



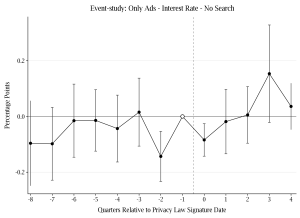
(b): Search-for-Price – Only Ads



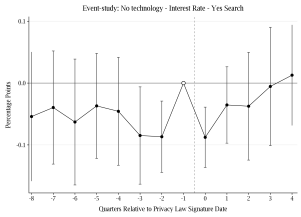
(c): Interest Rate – No Technology, No Search



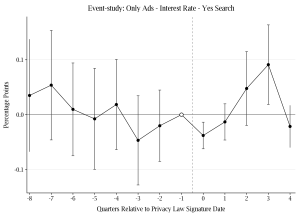
(d): Interest Rate – Only Ads, No Search



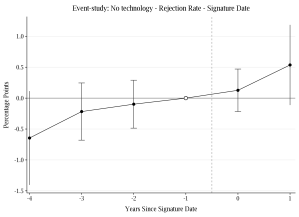
(e): Interest Rate – No Technology, Search



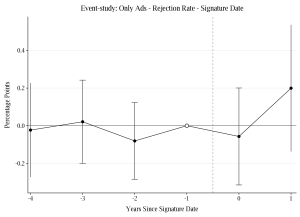
(f): Interest Rate – Only Ads, Search



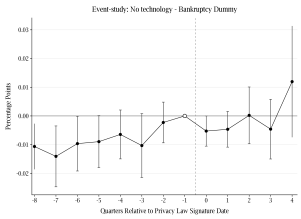
(g): Rejection Rate – No Technology



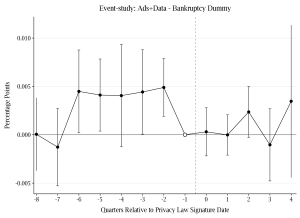
(h): Rejection Rate – Only Ads



(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 **?**. 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

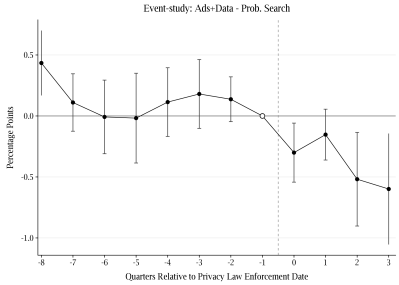
Table G.2. Enforcement: Summary Results (Ads + Data lenders)

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

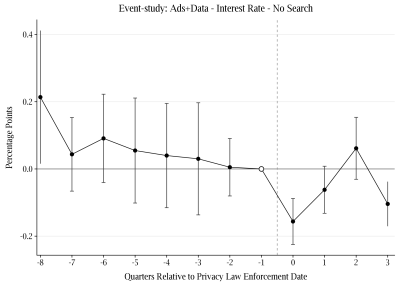
Notes. This table reports average post-enforcement treatment effects (ATT) of state consumer privacy-law enforcement on mortgage-market outcomes for the subsample of lenders that had adopted *both* advertising technologies and customer data platforms ("Ads + Data" lenders) prior to enforcement, estimated 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 4 (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)

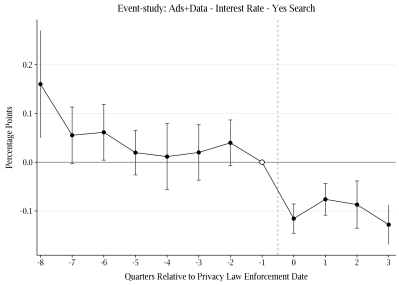
(a): Search-for-Price Probability



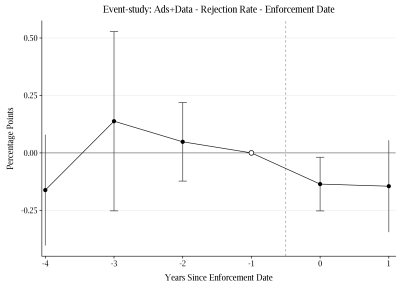
(b): Interest Rate – No Search



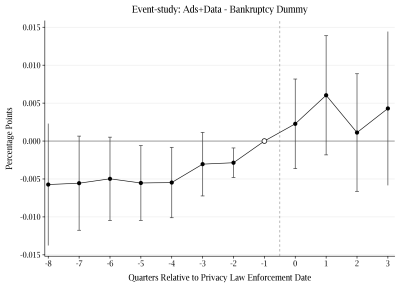
(c): Interest Rate – Search



(d): Rejection Rate



(e): Bankruptcy



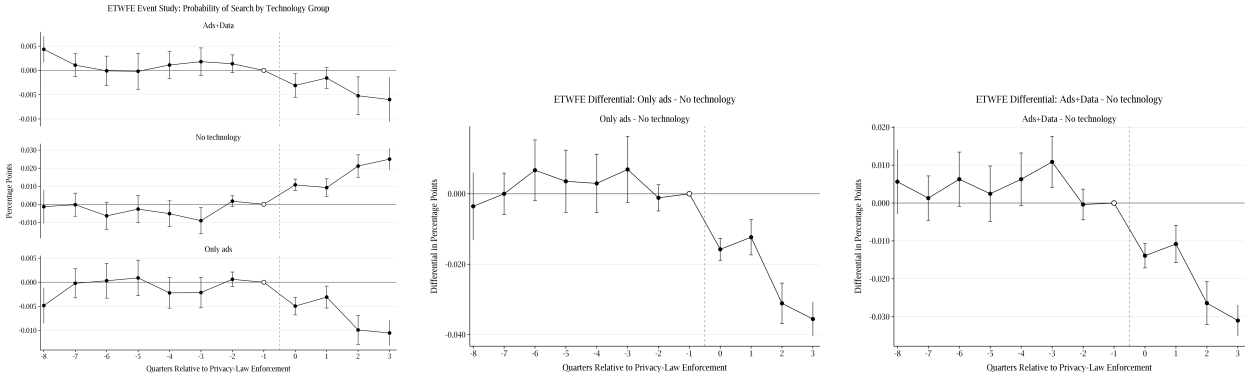
Notes: Event-study estimates for the Ads + Data group, in the canonical outcome order. The omitted period is the last pre-enforcement period. Confidence intervals are based on standard errors clustered at the state level.

G.4 Alternative Estimator: Stacked ETWFE

Table G.3. Stacked ETWFE Estimates: Enforcement Effects Across Outcomes (Excluding California)

	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		52	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, excluding 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 (Excluding California)

Notes: This figure reports stacked ETWFE event-study estimates for the search-for-price probability. 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.

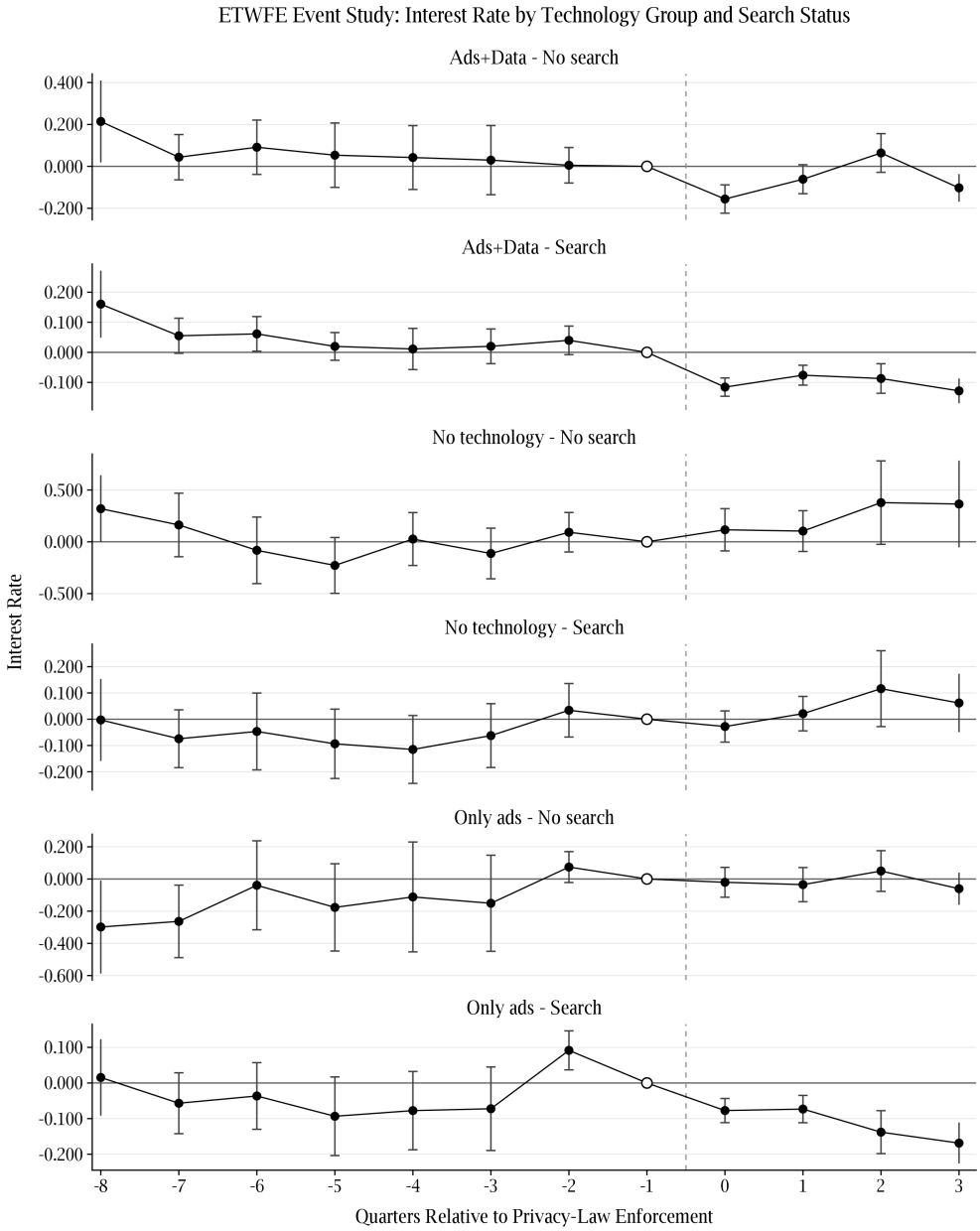
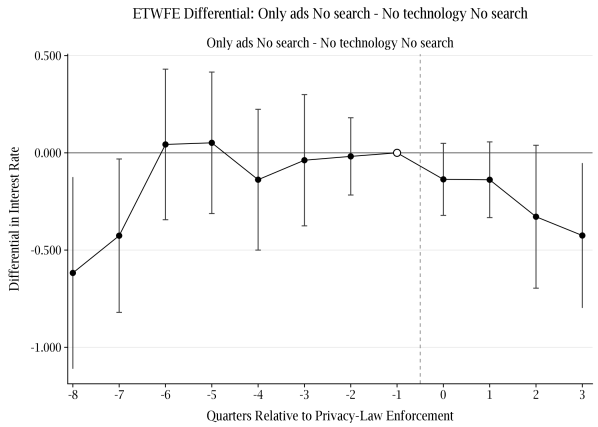
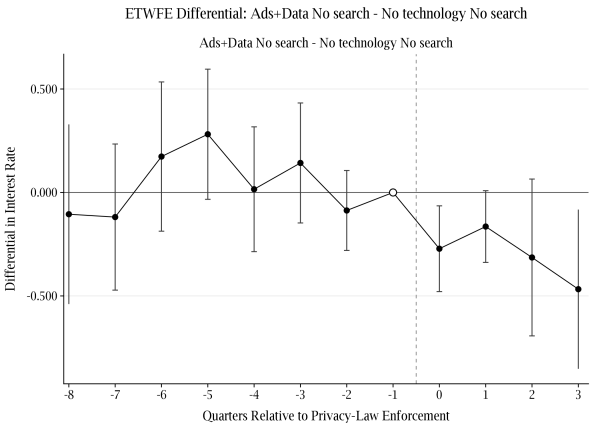


Figure G.4. Stacked ETWFE Event Study: Mortgage Interest Rate by Technology Group and Search Status

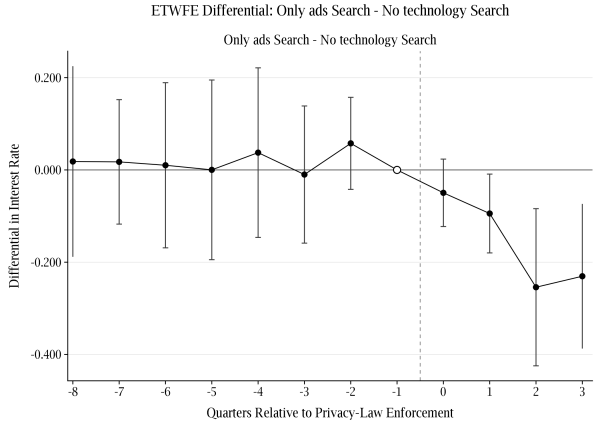
Notes: This figure plots the stacked ETWFE ATT path for mortgage interest rates by technology group and search status. The omitted period is the last pre-enforcement period. Confidence intervals are based on standard errors clustered at the state level.



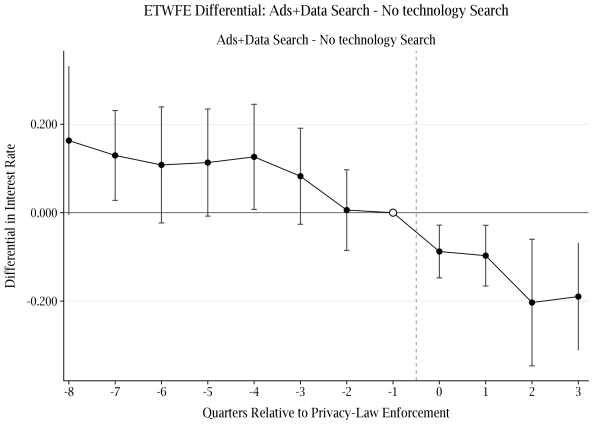
(a): Only ads − No technology (No search)



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



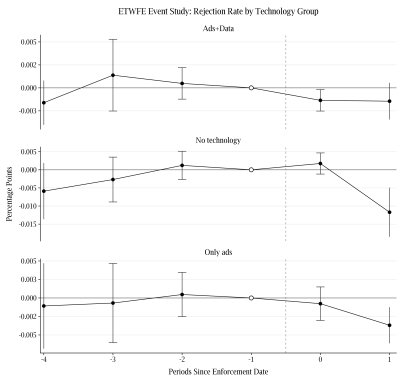
(c): Only ads − No technology (Search)



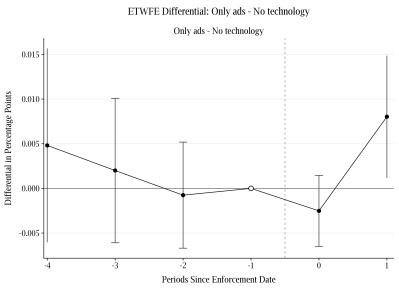
(d): Ads + Data − No technology (Search)

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

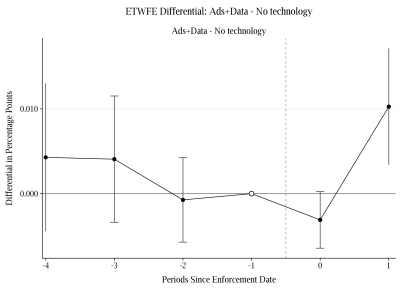
Notes: This figure plots the between-group differences in the stacked ETWFE ATT path for mortgage interest rates, relative to the No-technology group, separately by search-for-price status (no search first, then search). The omitted period is the last pre-enforcement period. Confidence intervals are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



(a): ATT by technology group



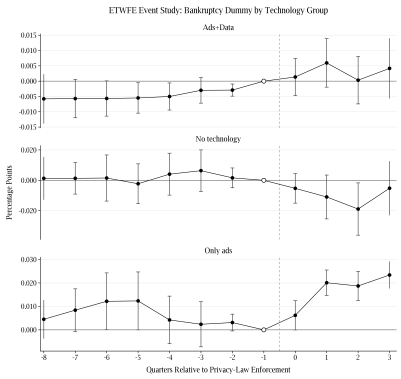
(b): Difference: Only ads − No technology



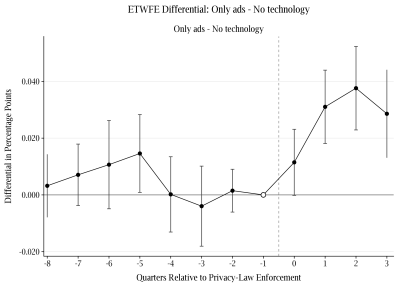
(c): Difference: Ads + Data − No technology

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

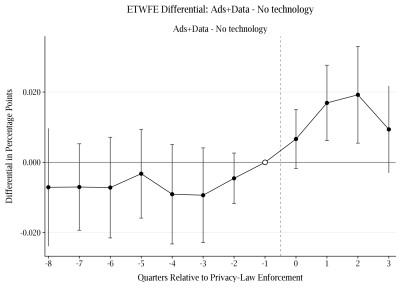
Notes: This figure reports stacked ETWFE event-study estimates for the loan rejection rate. Panel (a) plots the ATT path by technology group; Panels (b)–(c) plot the between-group differences relative to the No-technology group. The omitted period is the last pre-enforcement period. Confidence intervals 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) Only ads	(2) Ads+Data	(1) Only ads	(2) Ads+Data	(1) Only ads	(2) Ads+Data	(1) Only ads	(2) Ads+Data	(1) Only ads	(2) 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

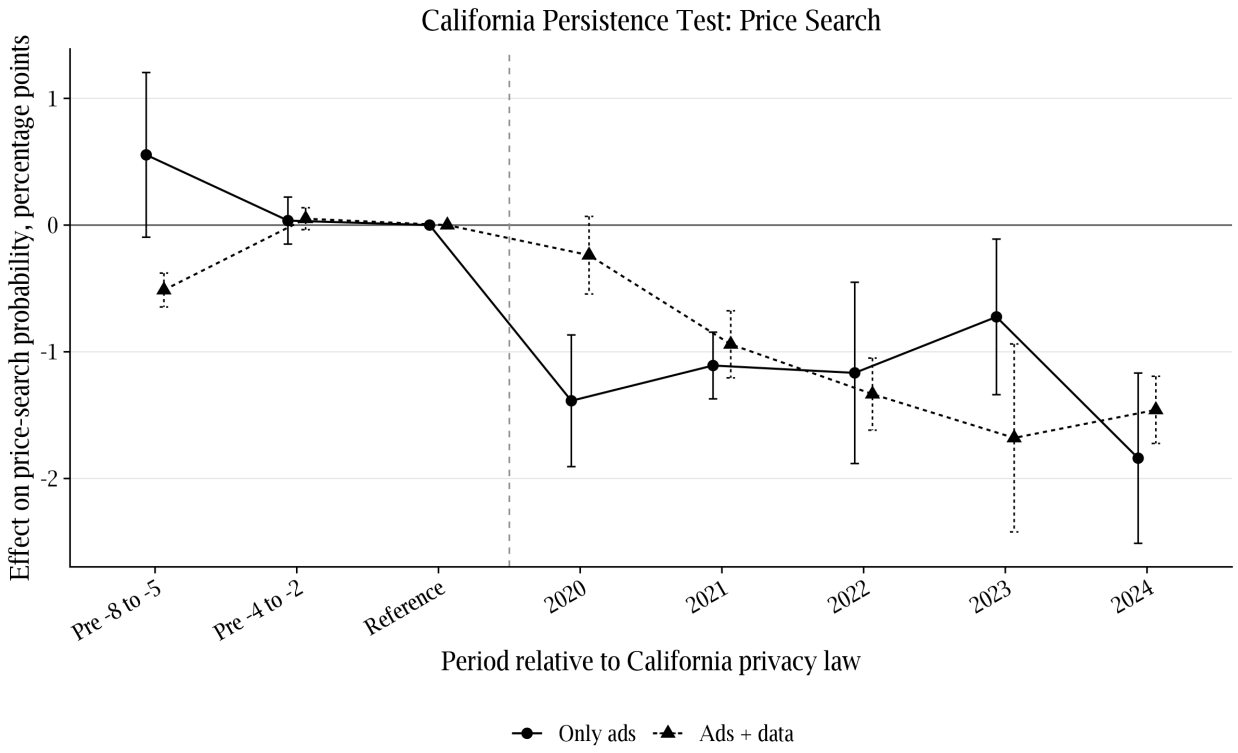
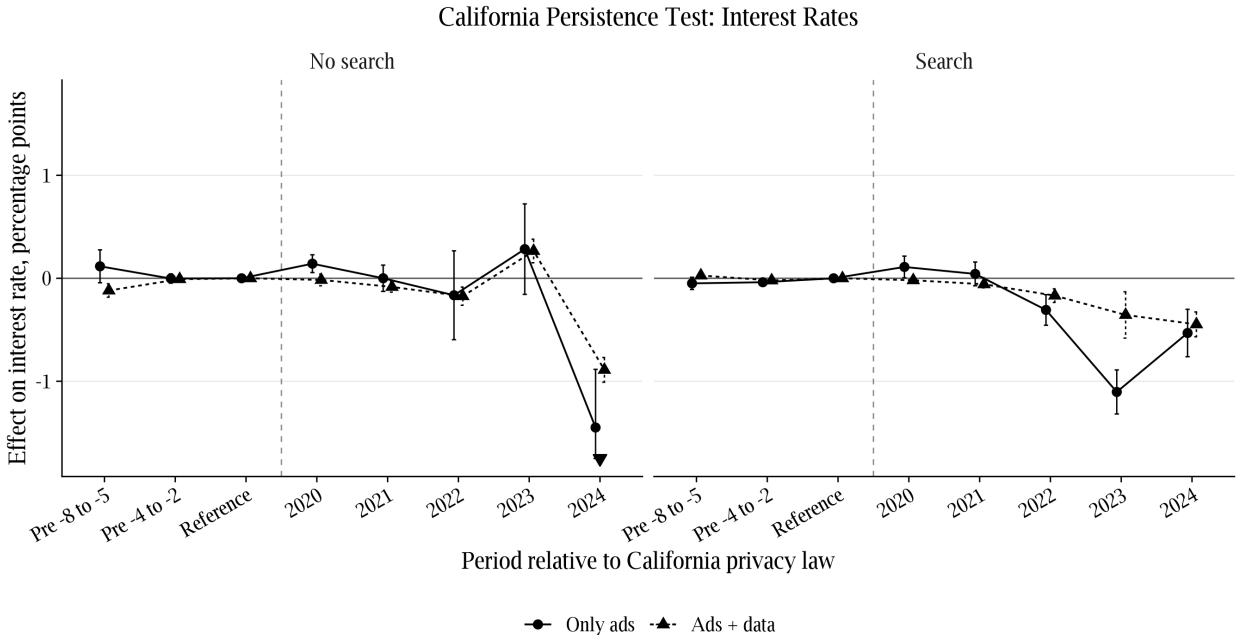


Figure G.9. COVID Robustness: Interest Rate – Ads + Data, Search



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 Technical Preliminaries and Model Microfoundations

This appendix collects notation, auxiliary identities, and the microfoundations used in the proofs. Appendix I then proves the model results in the order in which they appear in the text. To avoid duplicate numbering, each main-text result is restated manually before its proof rather than introduced as a new appendix theorem.

H.1 Notation

The baseline probability that a high-risk borrower is not flagged is

$$q \equiv 1 - p_0.$$

For a lender with screening index z_τ, the high-risk approval probability is

$$q_H(z_\tau) = q\big(1 - p(z_\tau)\big).$$

Thus

$$q_H(0) = q, \qquad q_H(\alpha) = q\big(1 - p(\alpha)\big).$$

The expected high-risk approval probability in the search market is

$$q_H^e = e_N q_H(0) + e_T q_H(\alpha).$$

Low-risk borrowers are always approved, so $q_L^e = 1$. Hence borrower i's effective per-quote search cost is

$$\tilde{\kappa}_i = \frac{\kappa_i}{q_{\theta_i}^e}.$$

For lender type τ, the funded approval mass is

$$A_\tau = \mu_L + \mu_H q_H(z_\tau),$$

and the default probability of the approved pool is

$$D_\tau = \frac{\mu_L d_L + \mu_H q_H(z_\tau) d_H}{\mu_L + \mu_H q_H(z_\tau)}.$$

For non-adopters and technology adopters, respectively, define

$$\bar{D} \equiv D_N, \qquad D_T \equiv D_{\tau=T}, \qquad \Delta \equiv \bar{D} - D_T.$$

The expected default cost of a searched lender draw is

$$D^e = e_N \bar{D} + e_T D_T = \bar{D} - e_T \Delta.$$

The expected search-rate anchor is

$$\tilde{r} = f + m_0 + D^e.$$

The residual mass of the non-adopter home-bank book after advertising poaching is

$$M(\gamma) = \mu_L(1-\gamma) + \mu_H\,(1 - \gamma q_H(\alpha)) = 1 - \gamma A_T,$$

where

$$A_T = \mu_L + \mu_H q_H(\alpha).$$

The default probability in this residual non-adopter book is

$$D_0^\gamma = \frac{\mu_L(1-\gamma)d_L + \mu_H\,(1-\gamma q_H(\alpha))\,d_H}{M(\gamma)}.$$

The default probability in the unscreened home-bank pool is

$$D_0 = \mu_L d_L + \mu_H d_H.$$

The privacy-law shocks are denoted by

$$\gamma^+ = (1-\delta_\gamma)\gamma, \qquad \alpha^+ = (1-\delta_\alpha)\alpha, \qquad 1 > \delta_\gamma > \delta_\alpha \geq 0.$$

The approval semi-elasticity used in Assumption 5 is

$$\varepsilon_A = \frac{\mu_H q\, p'(\alpha)\alpha}{A_T}.$$

H.2 Auxiliary identities

This subsection records identities used throughout the appendix.

First, screening precision lowers default risk among approved borrowers. Write the default probability of an approved pool as a function of the high-risk approval probability q_H:

$$D(q_H) = \frac{\mu_L d_L + \mu_H q_H d_H}{\mu_L + \mu_H q_H}.$$

Then

$$\frac{\partial D}{\partial q_H} = \frac{\mu_L \mu_H (d_H - d_L)}{(\mu_L + \mu_H q_H)^2} > 0.$$

Because

$$q'_H(z) = -qp'(z) < 0,$$

we have

$$\frac{dD(z)}{dz} = \frac{\partial D}{\partial q_H} q'_H(z) < 0.$$

Thus greater screening precision reduces default risk among approved borrowers. Since $z_T = \alpha > 0$ and $z_N = 0$,

$$D_T < D_N = \bar{D}, \qquad \Delta = \bar{D} - D_T > 0.$$

Second, advertising reach raises default risk in the residual non-adopter home-bank book. The high-risk share in that residual book is

$$s_H(\gamma) = \frac{\mu_H \left(1 - \gamma q_H(\alpha)\right)}{\mu_L(1-\gamma) + \mu_H \left(1 - \gamma q_H(\alpha)\right)}.$$

Differentiating with respect to γ gives

$$\frac{\partial s_H(\gamma)}{\partial \gamma} = \frac{\mu_H \mu_L \left(1 - q_H(\alpha)\right)}{M(\gamma)^2} > 0.$$

Since

$$D_0^\gamma = (1 - s_H(\gamma)) d_L + s_H(\gamma) d_H,$$

it follows that

$$\frac{\partial D_0^\gamma}{\partial \gamma} = \frac{\partial s_H(\gamma)}{\partial \gamma} (d_H - d_L) > 0.$$

Advertising reach therefore removes low-risk borrowers from the non-adopter home-bank book more aggressively than high-risk borrowers and makes the residual book riskier.

The residual non-adopter book becomes safer when the technology adopter approves more high-risk borrowers through the paid channel. Indeed,

$$\frac{\partial s_H}{\partial q_H(\alpha)} = -\frac{\mu_H \mu_L \gamma (1-\gamma)}{M(\gamma)^2} < 0.$$

Thus

$$\frac{\partial D_0^\gamma}{\partial q_H(\alpha)} < 0.$$

This identity is used in the proof of Proposition 8.5.

Finally, under Assumption 5, the residual organic-search pool available to technology adopters expands after the privacy law. Let

$$b \equiv q_H(\alpha), \qquad db > 0, \qquad d\gamma < 0.$$

For technology adopters,

$$A_T^R(\gamma) = \mu_L(1-\gamma) + \mu_H(1-\gamma b)b.$$

Therefore

$$dA_T^R = -(\mu_L + \mu_H b^2)d\gamma + \mu_H(1-2\gamma b)db.$$

If $1 - 2\gamma b \geq 0$, then $dA_T^R > 0$ immediately. If $1 - 2\gamma b < 0$, Assumption 5 implies

$$d(\gamma A_T) < 0 \quad \Longleftrightarrow \quad A_T(-d\gamma) > \gamma \mu_H db,$$

where $A_T = \mu_L + \mu_H b$. Hence

$$\begin{aligned} dA_T^R &> \mu_H db \left[\frac{\gamma(\mu_L + \mu_H b^2)}{\mu_L + \mu_H b} + 1 - 2\gamma b \right] \\ &= \frac{\mu_H db}{\mu_L + \mu_H b} \left[\mu_L(1 + \gamma - 2\gamma b) + \mu_H b(1 - \gamma b) \right] > 0, \end{aligned}$$

because $b \leq 1$ and $\gamma < 1$. Thus

$$dA_T^R > 0.$$

Since $dA_T > 0$, this implies

$$dQ_T^{\text{se}} = (1-\lambda)e_T \left[h_T dA_T + h_N dA_T^R \right] > 0.$$

This identity is used to show that the paid-channel share of technology-adopter acquisitions falls after the privacy law.

H.3 Differentiated-product posted-rate subgame

This subsection provides the microfoundation for the posted search rate

$$r_\tau^{\text{se}} = c_\tau + m_0.$$

The role of the subgame is deliberately modest: it delivers a common differentiation markup in the posted search market, allowing the model to focus on screening and targeted borrower acquisition.

A searching borrower who reaches lender type τ faces a differentiated mortgage product.

Differentiation captures service quality, processing speed, brand, branch relationships, platform convenience, and other lender-specific attributes. Let the residual demand facing lender type τ be

$$Q_\tau(r_\tau, r_{-\tau}),$$

with common own-rate semi-elasticity

$$\frac{\partial \log Q_\tau}{\partial r_\tau} = -\beta, \qquad \beta > 0.$$

This is a reduced-form representation of local differentiated-product Bertrand competition.

Given marginal cost

$$c_\tau = f + D_\tau,$$

lender τ's posted-search profit is

$$\Pi_\tau^{\text{se}} = Q_\tau(r_\tau, r_{-\tau})(r_\tau - c_\tau).$$

The first-order condition is

$$Q_\tau(r_\tau, r_{-\tau}) + (r_\tau - c_\tau)\frac{\partial Q_\tau(r_\tau, r_{-\tau})}{\partial r_\tau} = 0.$$

Using

$$\frac{\partial Q_\tau}{\partial r_\tau} = -\beta Q_\tau,$$

the first-order condition becomes

$$Q_\tau - \beta Q_\tau(r_\tau - c_\tau) = 0.$$

Therefore

$$r_\tau - c_\tau = \frac{1}{\beta} \equiv m_0.$$

The second-order condition holds locally. Under the constant semi-elasticity specification, $Q_\tau(r_\tau, \cdot)$ is proportional to $\exp(-\beta r_\tau)$, so at the optimum

$$\frac{\partial^2 \Pi_\tau^{\text{se}}}{\partial r_\tau^2} = -\beta Q_\tau < 0.$$

Thus

$$r_\tau^{\text{se}} = c_\tau + m_0.$$

The markup m_0 is common across lender types. Screening affects posted rates through marginal costs, not through the differentiation markup. Since

$$c_T = f + D_T < f + \bar{D} = c_N,$$

we have

$$r_T^{\text{se}} = c_T + m_0 < c_N + m_0 = r_N^{\text{se}}.$$

The expected search rate is

$$\tilde{r} = e_N r_N^{\text{se}} + e_T r_T^{\text{se}} = f + m_0 + e_N \bar{D} + e_T D_T = f + m_0 + D^e.$$

Moreover,

$$\tilde{r} - r_T^{\text{se}} = e_N (r_N^{\text{se}} - r_T^{\text{se}}) = e_N (c_N - c_T) = e_N \Delta,$$

and

$$r_N^{\text{se}} - \tilde{r} = e_T (r_N^{\text{se}} - r_T^{\text{se}}) = e_T \Delta.$$

These identities are used throughout the proofs.

A symmetric logit or Hotelling demand system would deliver the same qualitative structure: differentiated lenders charge a markup over marginal cost, and screening affects posted rates through marginal costs. The constant semi-elasticity specification is used here only to keep the posted-rate microfoundation transparent.

H.4 Advertising-platform surplus

This subsection derives the gross surplus from an advertising impression. The advertising platform charges

$$c^{\text{ad}} = \eta \bar{S}, \qquad \eta \in [0, 1),$$

where $\bar{S}$ is the technology lender's incremental gross surplus from the slot.

Consider first a prospective shopper. Without the paid impression, the borrower enters organic search. The technology adopter receives the borrower with probability e_T, and conditional on receiving the borrower earns expected funded profit $A_T m_0$. Hence the value of the shopper without the impression is

$$e_T A_T m_0.$$

With the paid impression, the borrower receives

$$r^{\text{ad}} = \tilde{r} + \kappa.$$

The expected funded profit from the impression is

$$A_T(r^{\text{ad}} - f - D_T).$$

Using

$$\tilde{r} = f + m_0 + D^e \qquad \text{and} \qquad D^e - D_T = e_N\Delta,$$

this value equals

$$A_T(\kappa + m_0 + e_N\Delta).$$

Therefore the incremental surplus from a shopper impression is

$$A_T(\kappa + m_0 + e_N\Delta) - e_T A_T m_0 = A_T\left[\kappa + e_N(m_0 + \Delta)\right].$$

For a prospective captive, the technology adopter's value without the impression is zero because the borrower remains with the incumbent non-adopter. With the impression, the value is again

$$A_T(\kappa + m_0 + e_N\Delta).$$

Weighting shoppers by $1 - \lambda$ and captives by λ, the gross advertising surplus is

$$\begin{aligned}\bar{S} &= (1-\lambda)A_T\left[\kappa + e_N(m_0 + \Delta)\right] + \lambda A_T(\kappa + m_0 + e_N\Delta) \\ &= A_T\left[\kappa + e_N(m_0 + \Delta) + \lambda e_T m_0\right],\end{aligned}$$

where the second equality uses $1 - e_N = e_T$. Therefore

$$c^{\text{ad}} = \eta A_T\left[\kappa + e_N(m_0 + \Delta) + \lambda e_T m_0\right].$$

The parameter η affects the division of advertising surplus between the platform and the technology lender, but it does not enter equilibrium rates or origination quantities.

I Proofs of Model Results

I.1 Underwriting rule

Auxiliary Lemma (Underwriting rule). *In the paid and posted-search channels, under Assumption 4(i), the equilibrium underwriting rule rejects borrowers flagged as high-risk and funds borrowers not flagged by the screening technology. The home-bank channel is priced separately as an unscreened incumbency pool, as in the main text.*

Proof. A flagged borrower is high-risk with probability one because the screening technology has no false positives. Consider the highest rate at which a flagged high-risk borrower could be willing to borrow. A high-risk shopper with search cost κ has reservation rate

$$\tilde{r} + \frac{\kappa}{q_H^e}.$$

A high-risk captive with search cost κ^c has search reservation

$$\tilde{r} + \frac{\kappa^c}{q_H^e},$$

but she also holds the approved home-bank contract

$$\bar{r} = \tilde{r} + \kappa^c.$$

Thus the home-bank contract caps her willingness to accept any competing quote at $\bar{r}$, which is weakly below

$$\tilde{r} + \frac{\kappa^c}{q_H^e}.$$

Therefore the highest relevant rate for a flagged high-risk borrower is bounded above by

$$\tilde{r} + \frac{\kappa^c}{q_H^e}.$$

At a rate $\tilde{r} + x$, the lender's per-loan profit from a high-risk borrower is

$$\pi(x, d_H) = m_0 + D^e + x - d_H.$$

Assumption 4(i) states that

$$\pi\left(\frac{\kappa^c}{q_H^e}, d_H\right) \leq 0.$$

Since $\pi(x, d_H)$ is increasing in x, funding a flagged high-risk borrower is unprofitable at every rate she would accept. Rejection yields zero payoff and therefore weakly dominates funding. Hence, in the paid and posted-search channels, the optimal underwriting rule rejects flagged high-risk borrowers and funds borrowers not flagged by the screening technology. Low-risk borrowers are never flagged and are always approved. □

I.2 Proof of Lemma 1

Lemma 1 (Acceptance, stopping, and participation). *Under Assumptions 1–3 and given the home-bank contract $\bar{r}$, borrower i accepts a held quote at rate r if and only if*

$$r \leq \min\{\rho_i, \bar{r}\}, \qquad \rho_i = \tilde{r} + \tilde{\kappa}_i.$$

The posted-rate cap implies first-draw acceptance in the search market. Starting from the home-bank contract, borrower i enters search if and only if

$$\tilde{\kappa}_i < \bar{r} - \tilde{r}.$$

At $\bar{r} = \tilde{r} + \kappa^c$, all low-cost borrowers search and all high-cost borrowers stay.

Proof. Fix $\bar{r}$ and the posted search rates $\{r_N^{\text{se}}, r_T^{\text{se}}\}$. Borrower i's effective search cost is

$$\tilde{\kappa}_i = \frac{\kappa_i}{q^e_{\theta_i}}.$$

The expected posted rate from a searched lender draw is

$$\tilde{r} = e_N r_N^{\text{se}} + e_T r_T^{\text{se}}.$$

By Appendix H.3,

$$r_T^{\text{se}} < r_N^{\text{se}},$$

and

$$r_N^{\text{se}} - \tilde{r} = e_T \Delta.$$

The posted-rate cap gives

$$e_T \Delta < \kappa \leq \tilde{\kappa}_i.$$

Therefore

$$r_N^{\text{se}} < \tilde{r} + \tilde{\kappa}_i.$$

Since r_N^{se} is the highest posted search rate, every searched posted rate is strictly below

$$\rho_i \equiv \tilde{r} + \tilde{\kappa}_i.$$

Thus any approved searched quote is acceptable.

There is also no value to drawing again after an approved searched quote. The largest possible rate improvement from replacing a non-adopter posted quote with a technology-adopter posted

quote is

$$r_N^{\text{se}} - r_T^{\text{se}} = \Delta.$$

The expected gain from an additional draw is bounded by $e_T\Delta$, which is strictly below κ and hence below $\tilde{\kappa}_i$. Therefore the recursive search problem collapses to first-draw acceptance.

Now suppose the borrower holds an approved quote at rate r. Accepting the quote is optimal relative to continued search if and only if

$$r \leq \rho_i = \tilde{r} + \tilde{\kappa}_i.$$

The borrower also holds the home-bank contract $\bar{r}$. Therefore the quote is accepted over the home-bank contract if and only if

$$r \leq \bar{r}.$$

Combining the two comparisons gives

$$r \leq \min\{\rho_i, \bar{r}\}.$$

Ties between accepting and further search are broken toward accepting. Ties between the home-bank contract and another held quote are broken toward the home bank.

Starting from the home-bank contract, if borrower i enters search, she pays the cash application cost κ_i. With probability $q^e_{\theta_i}$, she receives an approved searched quote and pays the expected searched rate $\tilde{r}$. With probability $1 - q^e_{\theta_i}$, she is rejected and falls back on the approved home-bank contract $\bar{r}$. Therefore the expected cost of entering search is

$$\kappa_i + q^e_{\theta_i}\tilde{r} + (1 - q^e_{\theta_i})\bar{r}.$$

Search is optimal if and only if

$$\kappa_i + q^e_{\theta_i}\tilde{r} + (1 - q^e_{\theta_i})\bar{r} < \bar{r}.$$

Rearranging gives

$$\frac{\kappa_i}{q^e_{\theta_i}} < \bar{r} - \tilde{r}.$$

Thus borrower i searches if and only if

$$\tilde{\kappa}_i < \bar{\kappa}, \qquad \bar{\kappa} \equiv \bar{r} - \tilde{r}.$$

At the home-bank contract characterized in Lemma 4,

$$\bar{r} = \tilde{r} + \kappa^c,$$

so

$$\bar{\kappa} = \kappa^c.$$

Low-risk low-cost borrowers have $\tilde{\kappa}_i = \kappa < \kappa^c$ and search. High-risk low-cost borrowers have

$$\tilde{\kappa}_i = \frac{\kappa}{q_H^e} < \kappa^c$$

by the search-cost separation in Assumption 2 and also search. Low-risk high-cost borrowers have

$$\tilde{\kappa}_i = \kappa^c$$

and stay by the tie-breaking rule. High-risk high-cost borrowers have

$$\tilde{\kappa}_i = \frac{\kappa^c}{q_H^e} > \kappa^c$$

and strictly prefer to stay.

Finally, the equilibrium advertising quote is

$$r^{\text{ad}} = \tilde{r} + \kappa.$$

Every unflagged recipient accepts this quote. For shoppers,

$$r^{\text{ad}} = \tilde{r} + \kappa \leq \tilde{r} + \tilde{\kappa}_i = \rho_i.$$

For captives,

$$r^{\text{ad}} = \tilde{r} + \kappa < \tilde{r} + \kappa^c = \bar{r}.$$

Thus all unflagged recipients accept the advertising quote. □

I.3 Proof of Corollary 1

Corollary 1 (High-risk information rents). *Approved high-risk borrowers are pooled with low-risk borrowers. Because rejection risk raises their effective search cost, high-risk borrowers*

receive an information rent equal to

$$\frac{\kappa(1-q_H^e)}{q_H^e}$$

on the advertising rate and

$$\frac{\kappa^c(1-q_H^e)}{q_H^e}$$

on the home-bank rate.

Proof. A high-risk borrower faces rejection risk in the search market. Her effective search cost is

$$\frac{\kappa_i}{q_H^e},$$

so her search reservation rate is

$$\rho_H(\kappa_i) = \tilde{r} + \frac{\kappa_i}{q_H^e}.$$

For a low-search-cost high-risk borrower receiving the advertising quote,

$$r^{\text{ad}} = \tilde{r} + \kappa.$$

Conditional on receiving this quote, the borrower is unflagged by the technology lender's screening technology. Under the equilibrium reject-if-flagged underwriting rule in the paid channel, the borrower is funded at the advertising quote. Her rent relative to her search reservation value is

$$\left(\tilde{r} + \frac{\kappa}{q_H^e}\right) - (\tilde{r} + \kappa) = \frac{\kappa(1-q_H^e)}{q_H^e}.$$

For a high-search-cost high-risk borrower remaining at the home bank, the home-bank rate is

$$\bar{r} = \tilde{r} + \kappa^c.$$

Her search reservation value is

$$\tilde{r} + \frac{\kappa^c}{q_H^e}.$$

Thus her rent is

$$\left(\tilde{r} + \frac{\kappa^c}{q_H^e}\right) - (\tilde{r} + \kappa^c) = \frac{\kappa^c(1-q_H^e)}{q_H^e}.$$

□

I.4 Proof of Lemma 2

Lemma 2 (Separability). *Conditional on the underwriting rule and on the market anchor $\tilde{r}$, the lender's problem is additively separable across the posted-search, targeted-advertising, and incumbency channels.*

Proof. Fix the underwriting rule characterized in Appendix I.1 for paid and posted-search applications. The underwriting rule determines the funded pool at each lender type in those channels and therefore determines

$$D_\tau, \qquad c_\tau = f + D_\tau, \qquad r_\tau^{\text{se}} = c_\tau + m_0.$$

Fix also the market anchor

$$\tilde{r} = e_N r_N^{\text{se}} + e_T r_T^{\text{se}}.$$

Conditional on these objects, lender τ's profit is

$$\Pi_\tau = Q_\tau^{\text{se}} \left(r_\tau^{\text{se}} - f - D_\tau\right) + \gamma_\tau h_{-\tau} \left[\pi_\tau^{\text{ad}}(a_\tau) - c^{\text{ad}}\right] + B_\tau(\bar{r}_\tau).$$

The three terms correspond to distinct controls. The posted-search term depends on r_τ^{se}. The paid-channel term depends on the advertising policy a_τ. The incumbency term depends on the home-bank rate $\bar{r}_\tau$. Conditional on the underwriting rule and $\tilde{r}$, no control appears in more than one term.

The rival's advertising reach γ enters the non-adopter's home-bank book through $M(\gamma)$ and D_0^γ. This is an exposure of the non-adopter's book, not a choice variable of the non-adopter. Hence the optimization problem is additively separable across the three channels. □

I.5 Proof of Lemma 3

Lemma 3 (Optimal advertising quote). *Under Assumption 4, a technology adopter optimally withholds advertising quotes from borrowers flagged as high-risk and offers*

$$r^{\text{ad}} = \tilde{r} + \kappa$$

to unflagged non-customers.

Proof. By Appendix I.1, the technology adopter optimally withholds quotes from borrowers flagged as high-risk. It remains to solve the optimal quote to unflagged non-customers.

Let the technology adopter quote

$$r = \tilde{r} + x.$$

By Lemma 1, a borrower accepts the quote if and only if

$$\tilde{r} + x \leq \min\{\rho_i, \bar{r}\}.$$

For low-cost shoppers, the lowest reservation threshold is the low-risk threshold

$$\rho_L = \tilde{r} + \kappa.$$

Therefore all shoppers accept any quote with $x \leq \kappa$. Since captives have the home-bank contract

$$\bar{r} = \tilde{r} + \kappa^c$$

and $\kappa < \kappa^c$, all captives also accept any quote with $x \leq \kappa$.

For any $x \leq \kappa$, all unflagged recipients accept. The expected funded profit per impression is

$$\begin{aligned}\pi^{\text{ad}}(x) &= \mu_L \left(\tilde{r} + x - f - d_L\right) + \mu_H q_H(\alpha) \left(\tilde{r} + x - f - d_H\right) \\ &= A_T \left(\tilde{r} + x - f - D_T\right).\end{aligned}$$

Using

$$\tilde{r} = f + m_0 + D^e \qquad \text{and} \qquad D^e - D_T = e_N \Delta,$$

this becomes

$$\pi^{\text{ad}}(x) = A_T \left(x + m_0 + e_N \Delta\right).$$

The expression is strictly increasing in x, so the best quote accepted by all recipients is

$$x = \kappa.$$

Consider next the region

$$\kappa < x \leq \frac{\kappa}{q_H^e}.$$

Low-risk shoppers reject the quote and enter organic search. The only shoppers who may accept in this region are unflagged high-risk shoppers, whose search reservation value is $\tilde{r} + \kappa/q_H^e$. Funding such a borrower at any acceptable rate in this region is loss-making under Assumption 4(i), since

$$\frac{\kappa}{q_H^e} < \frac{\kappa^c}{q_H^e}$$

and Assumption 4(i) rules out profitable funding of a high-risk borrower even at the higher margin κ^c/q_H^e. For any such x, a quote $x' \in (\kappa/q_H^e, \kappa^c)$ eliminates this loss-making high-risk shopper

component while preserving captive acceptance and weakly increasing the captive margin. Hence every quote in this intermediate region is dominated by a captive-skimming quote.

For

$$\frac{\kappa}{q_H^e} < x < \kappa^c,$$

all shoppers reject and enter organic search. Only captives accept. The supremum value of such a captive-skimming deviation is obtained as $x \uparrow \kappa^c$. Because ties at $x = \kappa^c$ are assigned to the home bank, the supremum is not attained. The value of the captive-skimming deviation, including the technology lender's organic continuation value from shoppers, is

$$V^{\text{skim}} = (1 - \lambda) e_T A_T m_0 + \lambda A_T (\kappa^c + m_0 + e_N \Delta).$$

The value of the pooled quote $x = \kappa$ is

$$V^{\text{pool}} = A_T (\kappa + m_0 + e_N \Delta).$$

The difference is

$$V^{\text{pool}} - V^{\text{skim}} = A_T \left[(1 - \lambda) \left(\kappa + e_N (m_0 + \Delta) \right) - \lambda (\kappa^c - \kappa) \right].$$

Assumption 4(iv) makes this expression strictly positive. Thus the pooled quote dominates captive skimming.

Finally, for $x \geq \kappa^c$, captives weakly prefer the home-bank contract, with ties assigned to the home bank, and shoppers have already rejected. Hence no profitable deviation exists in this region.

The unique optimal advertising quote is therefore

$$r^{\text{ad}} = \tilde{r} + \kappa.$$

Expected funded profit per impression is

$$\pi^{\text{ad}} = A_T \left(\kappa + m_0 + e_N \Delta \right) > 0.$$

□

I.6 Proof of Lemma 4

Lemma 4 (Home-bank incumbency rate). *For every $\gamma \in [0, 1)$, the optimal home-bank rate is*

$$\bar{r} = \tilde{r} + \kappa^c.$$

Holding the search-market anchor $\tilde{r}$ fixed, the optimal home-bank markup is independent of advertising reach:

$$\frac{\partial \bar{r}}{\partial \gamma} = 0.$$

Proof. Let

$$x \equiv \bar{r} - \tilde{r}$$

denote the home-bank markup over the search-market anchor. The home bank chooses x, taking $\tilde{r}$ as given.

By Lemma 1, a borrower remains with the home bank rather than entering search if and only if

$$\tilde{\kappa}_i \geq x,$$

with ties assigned to the home bank. The relevant effective search costs are

$$\kappa, \qquad \frac{\kappa}{q_H^e}, \qquad \kappa^c, \qquad \frac{\kappa^c}{q_H^e}.$$

By Assumption 2,

$$\kappa < \frac{\kappa}{q_H^e} < \kappa^c < \frac{\kappa^c}{q_H^e}.$$

If $x \leq \kappa$, the home bank attempts to retain the entire borrower pool. The highest payoff in this region occurs at $x = \kappa$. Assumption 4(iii) states that defending the whole book at the shopper margin is unprofitable:

$$\pi(\kappa, D_0) \leq 0.$$

Thus no $x \leq \kappa$ can be optimal.

If

$$\kappa < x \leq \frac{\kappa}{q_H^e},$$

low-risk low-cost borrowers search, while high-risk low-cost borrowers may remain because rejection risk raises their effective search cost. This worsens the risk composition of the retained book relative to the full pool. The additional high-risk retention is not profitable under Assumption 4(i), and the captive component is priced below the margin κ^c available in the next region. Thus this region is dominated.

If

$$\frac{\kappa}{q_H^e} < x \leq \kappa^c,$$

all low-cost borrowers search and all high-cost borrowers remain. In this region, the composition of the retained high-cost book is constant in x, while the rate charged to that book is increasing in x.

Hence the payoff is strictly increasing in x over this region.

For a technology-adopter home bank, the retained mass is $h_T\lambda$, and the default probability is D_0. For a non-adopter home bank, targeted advertising by technology adopters skims part of the high-cost book. The retained mass is

$$h_N\lambda M(\gamma),$$

and the residual default probability is D_0^γ. In either case, changing x within the interval

$$\left(\frac{\kappa}{q_H^e}, \kappa^c\right]$$

changes only the price charged to the same retained set. The optimal choice in this region is therefore

$$x = \kappa^c.$$

Assumption 4(ii) guarantees that the residual home-bank book is profitable at this margin:

$$\pi(\kappa^c, D_0^\gamma) \geq 0.$$

If

$$\kappa^c < x \leq \frac{\kappa^c}{q_H^e},$$

low-risk high-cost borrowers also enter search. The retained book consists only of high-risk high-cost borrowers. By Assumption 4(i), serving high-risk borrowers at any rate they would accept is unprofitable. This region is therefore dominated. For

$$x > \frac{\kappa^c}{q_H^e},$$

all borrowers search and the home-bank book is empty.

Thus the unique optimal home-bank markup is

$$x^* = \kappa^c,$$

and the optimal home-bank rate is

$$\bar{r} = \tilde{r} + \kappa^c.$$

Holding $\tilde{r}$ fixed, the optimal markup is independent of advertising reach γ, so

$$\frac{\partial \bar{r}}{\partial \gamma} = 0.$$

The level of $\bar{r}$, however, moves one-for-one with the search-market anchor. This distinction matters in Proposition 8.5. □

I.7 Proof of Lemma 5

Lemma 5 (Posted search rates). *In the posted search market,*

$$r_\tau^{\text{se}} = c_\tau + m_0.$$

The expected search rate is

$$\tilde{r} = e_N r_N^{\text{se}} + e_T r_T^{\text{se}} = r_T^{\text{se}} + e_N \Delta.$$

Proof. Appendix H.3 derives

$$r_\tau^{\text{se}} = c_\tau + m_0, \qquad c_\tau = f + D_\tau.$$

Since technology adopters have lower default costs among approved borrowers,

$$D_T < \bar{D} = D_N,$$

we have

$$c_T < c_N.$$

Therefore

$$r_T^{\text{se}} = c_T + m_0 < c_N + m_0 = r_N^{\text{se}}.$$

The expected search rate is

$$\begin{aligned} \tilde{r} &= e_N r_N^{\text{se}} + e_T r_T^{\text{se}} \\ &= e_N (f + \bar{D} + m_0) + e_T (f + D_T + m_0) \\ &= f + m_0 + D^e. \end{aligned}$$

Equivalently,

$$\tilde{r} = r_T^{\text{se}} + e_N \Delta.$$

Finally,

$$r^{\text{ad}} - r_T^{\text{se}} = (\tilde{r} + \kappa) - r_T^{\text{se}} = \kappa + e_N \Delta.$$

□

I.8 Proof of Proposition 1

Proposition 1 (Equilibrium characterization). *Under Assumptions 1–4, the equilibrium exists and is unique. It is characterized by the forward-substitution system in Proposition 1. The equilibrium quantities, book shares, and average book rates are those stated in the main text.*

Proof. Appendix H.3 uniquely determines the posted search rates:

$$r_\tau^{\mathrm{se}} = c_\tau + m_0.$$

These rates uniquely determine the search-market anchor

$$\tilde{r} = e_N r_N^{\mathrm{se}} + e_T r_T^{\mathrm{se}}.$$

Given $\tilde{r}$, Lemmas 1, 3, and 4 uniquely determine borrower reservation values, the advertising quote, and the home-bank rate:

$$\rho_i = \tilde{r} + \tilde{\kappa}_i, \qquad r^{\mathrm{ad}} = \tilde{r} + \kappa, \qquad \bar{r} = \tilde{r} + \kappa^c.$$

The search/stay threshold is

$$\bar{\kappa} = \bar{r} - \tilde{r} = \kappa^c.$$

Assumption 2 then uniquely partitions borrowers: all low-cost borrowers search unless they receive and accept an advertising quote, while high-cost borrowers remain with the home bank unless they are poached by advertising. The underwriting rule in the paid and posted-search channels is uniquely determined by Appendix I.1. Hence the equilibrium exists by construction and is unique.

We next derive the quantities. The paid channel is available only to technology adopters and only for non-customers. A technology adopter reaches a non-adopter home-bank borrower with probability γ, and funded approval at the technology lender is

$$A_T = \mu_L + \mu_H q_H(\alpha).$$

Therefore

$$Q_T^{\mathrm{ad}} = \gamma h_N A_T, \qquad Q_N^{\mathrm{ad}} = 0.$$

For organic search, low-cost borrowers search unless they have already accepted an advertising quote. Technology-adopter home-bank borrowers are not targeted by their own lender's advertising. Their contribution to the search pool is

$$h_T A_\tau.$$

For non-adopter home-bank borrowers, targeted advertising removes low-risk borrowers with probability γ, and removes high-risk borrowers when the technology adopter reaches them and they are not flagged, which occurs with probability $\gamma q_H(\alpha)$. The residual approval mass from the non-adopter home-bank pool that remains available for organic search and draws lender type τ is therefore

$$A_\tau^R(\gamma) = \mu_L(1-\gamma) + \mu_H\left(1-\gamma q_H(\alpha)\right) q_H(z_\tau).$$

Thus organic-search originations are

$$Q_\tau^{\text{se}} = (1-\lambda)e_\tau\left[h_T A_\tau + h_N A_\tau^R(\gamma)\right].$$

Home-bank originations consist of high-cost borrowers who remain with the incumbent. For technology adopters,

$$Q_T^0 = h_T\lambda.$$

For non-adopters,

$$Q_N^0 = h_N\lambda M(\gamma).$$

Total originations are

$$Q_T = Q_T^{\text{se}} + Q_T^{\text{ad}} + Q_T^0, \qquad Q_N = Q_N^{\text{se}} + Q_N^0.$$

The within-lender book shares are

$$\hat{\omega}_T^{\text{se}} = \frac{Q_T^{\text{se}}}{Q_T}, \qquad \hat{\omega}_T^{\text{ad}} = \frac{Q_T^{\text{ad}}}{Q_T}, \qquad \hat{\omega}_T^0 = \frac{Q_T^0}{Q_T},$$

and

$$\hat{\omega}_N^{\text{se}} = \frac{Q_N^{\text{se}}}{Q_N}, \qquad \hat{\omega}_N^0 = \frac{Q_N^0}{Q_N}.$$

These shares sum to one within lender type by construction.

The paid-channel share of technology-adopter acquisitions is

$$\omega_T^{\text{ad}} = \frac{Q_T^{\text{ad}}}{Q_T^{\text{ad}} + Q_T^{\text{se}}}.$$

Let

$$N(\gamma) \equiv Q_T^{\text{ad}} = \gamma h_N A_T$$

and

$$S(\gamma) \equiv Q_T^{\text{se}} = (1-\lambda)e_T\left[h_T A_T + h_N A_T^R(\gamma)\right].$$

Holding screening precision fixed, A_T is constant and

$$N'(\gamma) = h_N A_T > 0.$$

Moreover,

$$\frac{\partial A_T^R(\gamma)}{\partial \gamma} = -\mu_L - \mu_H q_H(\alpha) q_H(z_T) < 0.$$

Therefore

$$S'(\gamma) < 0.$$

It follows that

$$\frac{\partial \omega_T^{\text{ad}}}{\partial \gamma} = \frac{S(\gamma)N'(\gamma) - N(\gamma)S'(\gamma)}{[N(\gamma) + S(\gamma)]^2} > 0.$$

Thus the paid-channel share of technology-adopter acquisitions is strictly increasing in advertising reach.

We now derive average book rates. For technology adopters,

$$r_T^{\text{book}} = \hat{\omega}_T^{\text{se}} r_T^{\text{se}} + \hat{\omega}_T^{\text{ad}} r^{\text{ad}} + \hat{\omega}_T^0 \bar{r}.$$

Since

$$r^{\text{ad}} - r_T^{\text{se}} = \kappa + e_N \Delta$$

and

$$\bar{r} - r_T^{\text{se}} = \kappa^c + e_N \Delta,$$

we obtain

$$r_T^{\text{book}} = c_T + m_0 + \hat{\omega}_T^{\text{ad}}(\kappa + e_N \Delta) + \hat{\omega}_T^0(\kappa^c + e_N \Delta).$$

Equivalently,

$$r_T^{\text{book}} = f + \bar{D} + m_0 - \left[1 - e_N \left(\hat{\omega}_T^{\text{ad}} + \hat{\omega}_T^0\right)\right] \Delta + \hat{\omega}_T^{\text{ad}} \kappa + \hat{\omega}_T^0 \kappa^c.$$

For non-adopters,

$$r_N^{\text{book}} = \hat{\omega}_N^{\text{se}} r_N^{\text{se}} + \hat{\omega}_N^0 \bar{r}.$$

Because

$$\bar{r} - r_N^{\text{se}} = \kappa^c - e_T \Delta,$$

we have

$$r_N^{\text{book}} = c_N + m_0 + \hat{\omega}_N^0(\kappa^c - e_T \Delta).$$

Equivalently,

$$r_N^{\text{book}} = f + \bar{D} + m_0 - \hat{\omega}_N^0 e_T \Delta + \hat{\omega}_N^0 \kappa^c.$$

Subtracting gives

$$r_T^{\text{book}} - r_N^{\text{book}} = -\left(e_N \hat{\omega}_T^{\text{se}} + e_T \hat{\omega}_N^{\text{se}}\right) \Delta + \hat{\omega}_T^{\text{ad}} \kappa + \left(\hat{\omega}_T^0 - \hat{\omega}_N^0\right) \kappa^c.$$

The first term is the net screening pass-through advantage of technology adopters. The remaining terms are search-cost extraction through the advertising and home-bank channels. Hence

$$r_T^{\text{book}} - r_N^{\text{book}} \gtrless 0.$$

□

I.9 Proof of Proposition 2

Proposition 2 (Privacy-law comparative statics). *Under Assumptions 1–5, privacy laws weaken screening, reduce targeted advertising, and reallocate borrowers across equilibrium channels as stated in Proposition 8.5.*

Proof. The privacy law changes the technology adopter's reach and screening precision according to

$$\gamma^+ = (1 - \delta_\gamma)\gamma, \qquad \alpha^+ = (1 - \delta_\alpha)\alpha, \qquad 1 > \delta_\gamma > \delta_\alpha \geq 0.$$

Let $dX = X^+ - X$.

Because $\alpha^+ < \alpha$ and $p'(z) > 0$,

$$p(\alpha^+) < p(\alpha).$$

Therefore

$$q_H(\alpha^+) = q\left(1 - p(\alpha^+)\right) > q\left(1 - p(\alpha)\right) = q_H(\alpha).$$

Technology adopters approve more high-risk borrowers and reject fewer of them. Since

$$R_T = \mu_H \left(1 - q_H(\alpha)\right),$$

we obtain

$$R_T^+ < R_T.$$

The approved pool at technology adopters becomes riskier, so

$$dD_T = D_T^+ - D_T > 0.$$

Consequently,

$$dr_T^{\text{se}} = dD_T.$$

Because

$$\tilde{r} = e_N r_N^{\text{se}} + e_T r_T^{\text{se}},$$

and r_N^{se} is unchanged by the technology adopter's screening shock,

$$d\tilde{r} = e_T dD_T.$$

Next consider paid-channel volume. The paid-channel origination volume is

$$Q_T^{\text{ad}} = \gamma h_N A_T, \qquad A_T = \mu_L + \mu_H q_H(\alpha).$$

To first order,

$$dA_T = \mu_H q p'(\alpha)\alpha\delta_\alpha.$$

Thus

$$\frac{dA_T}{A_T} = \varepsilon_A \delta_\alpha, \qquad \varepsilon_A = \frac{\mu_H q p'(\alpha)\alpha}{A_T}.$$

Since

$$d\gamma = -\delta_\gamma \gamma,$$

we have

$$d(\gamma A_T) = A_T d\gamma + \gamma dA_T = \gamma A_T(\varepsilon_A \delta_\alpha - \delta_\gamma).$$

Assumption 5 imposes

$$\delta_\gamma > \varepsilon_A \delta_\alpha,$$

so

$$d(\gamma A_T) < 0.$$

Therefore

$$dQ_T^{\text{ad}} = h_N d(\gamma A_T) < 0.$$

Paid-channel volume falls.

The non-adopter home-bank book moves in the opposite direction. Its volume is

$$Q_N^0 = h_N \lambda M(\gamma),$$

where

$$M(\gamma) = 1 - \gamma A_T.$$

Therefore

$$dQ_N^0 = -h_N \lambda d(\gamma A_T) > 0.$$

The non-adopter home-bank book expands.

The residual non-adopter book also becomes safer. From Appendix H.2,

$$\frac{\partial D_0^\gamma}{\partial \gamma} > 0$$

and

$$\frac{\partial D_0^\gamma}{\partial q_H(\alpha)} < 0.$$

The privacy law reduces γ and increases $q_H(\alpha)$. Both forces reduce D_0^γ. Hence

$$dD_0^\gamma < 0.$$

The paid-channel share of technology-adopter acquisitions also falls. Recall

$$\omega_T^{\text{ad}} = \frac{Q_T^{\text{ad}}}{Q_T^{\text{ad}} + Q_T^{\text{se}}}.$$

We have already shown that

$$dQ_T^{\text{ad}} < 0.$$

Appendix H.2 shows that, under Assumption 5,

$$dQ_T^{\text{se}} > 0.$$

Therefore

$$d\omega_T^{\text{ad}} < 0.$$

We now clarify the home-bank rate. Lemma 4 shows that the optimal home-bank markup over the search-market anchor is

$$\bar{r} - \tilde{r} = \kappa^c.$$

This markup is invariant to advertising reach. However, the privacy law also reduces screening precision, which raises D_T and lifts the search-market anchor by

$$d\tilde{r} = e_T dD_T.$$

Re-solving the home-bank problem after the law gives

$$\bar{r}^{+} = \tilde{r}^{+} + \kappa^{c}.$$

Therefore

$$d\bar{r} = d\tilde{r} = e_T dD_T > 0.$$

The participation threshold is unchanged:

$$\bar{\kappa}^{+} = \bar{r}^{+} - \tilde{r}^{+} = \kappa^{c}.$$

Thus the search partition is preserved.

Finally, consider measured acquisition rates at technology adopters. Define

$$\chi \equiv r^{\text{ad}} - r_T^{\text{se}} = \kappa + e_N \Delta.$$

The average acquisition rate of technology adopters can be written as

$$r_T^{\text{new}} = r_T^{\text{se}} + \omega_T^{\text{ad}} \chi.$$

Because

$$d\Delta = -dD_T,$$

we have

$$d\chi = -e_N dD_T.$$

Therefore

$$\begin{aligned} dr_T^{\text{new}} &= dr_T^{\text{se}} + \omega_T^{\text{ad}} d\chi + \chi d\omega_T^{\text{ad}} \\ &= dD_T - e_N \omega_T^{\text{ad}} dD_T + \chi d\omega_T^{\text{ad}} \\ &= \left(1 - e_N \omega_T^{\text{ad}}\right) dD_T + \chi d\omega_T^{\text{ad}}. \end{aligned}$$

The first term is positive because weaker screening raises funding costs. The second term is negative because $d\omega_T^{\text{ad}} < 0$. Hence measured rates at technology adopters can fall if the composition effect dominates:

$$dr_T^{\text{new}} < 0 \quad \Longleftrightarrow \quad \chi \left| d\omega_T^{\text{ad}} \right| > \left(1 - e_N \omega_T^{\text{ad}}\right) dD_T.$$

The advertising quote illustrates the incidence of the screening advantage. In the pre-law equilibrium,

$$r^{\text{ad}} = \tilde{r} + \kappa = f + m_0 + \bar{D} - e_T \Delta + \kappa.$$

Relative to the no-data benchmark $f + m_0 + \bar{D}$, advertising borrowers receive pass-through equal to

$e_T\Delta$. Relative to the technology adopter's own posted rate,

$$r^{\text{ad}} - r_T^{\text{se}} = \kappa + e_N\Delta.$$

Thus the lender captures the search-cost wedge and the non-adopter visibility component of its screening advantage.

After the privacy law,

$$\Delta^+ = \bar{D} - D_T^+ < \Delta.$$

The advertising quote becomes

$$r^{\text{ad},+} = f + m_0 + \bar{D} - e_T\Delta^+ + \kappa,$$

and the wedge over the technology adopter's posted rate is

$$r^{\text{ad},+} - r_T^{\text{se},+} = \kappa + e_N\Delta^+.$$

Privacy laws reduce the pass-through from screening but do not eliminate it unless the screening advantage is fully destroyed.

The surplus effects follow from these channel movements. High-risk borrowers benefit from higher approval probabilities. Low-risk borrowers face higher posted rates because screening becomes less precise. Borrowers displaced from the advertising channel lose the search-cost savings embedded in the advertising quote. Captives no longer poached by advertising return to the home-bank contract, whose level rises with the search-market anchor:

$$\bar{r}^+ = \tilde{r}^+ + \kappa^c.$$

Non-adopters gain in the home-bank channel because fewer borrowers are poached and the residual book becomes safer:

$$dQ_N^0 > 0, \qquad dD_0^\gamma < 0.$$

Technology adopters lose paid-channel volume, although the total profit effect is ambiguous because reduced advertising expenditure and higher rates on remaining originations partly offset the loss of paid-channel borrowers. Advertising platforms lose reach. Their revenue is

$$\Pi_{\text{plat}} = \eta\gamma h_N \bar{S},$$

so

$$d\Pi_{\text{plat}} = \eta h_N \left(\bar{S}\, d\gamma + \gamma\, d\bar{S}\right).$$

Platform revenue falls whenever the reach loss dominates any increase in surplus per remaining impression. □

J Appendix: Structural Calibration and Moment Inversion

This appendix provides the details behind the calibration in Section **??**. It reports the probability audit, the full moment definitions, the recursive parameter map, the advertising-intensity validation, overidentifying restrictions, and the inference procedure for generated channel weights.

J.1 Probability Audit and Overlap Correction

The baseline construction uses

$$\pi_i^S = \min\{\pi_i^M, \pi_i^P + \pi_i^F\}. \tag{J.1}$$

This rule treats stated price-search and rejection-fear motives as the active-search component of multi-lender consideration, total multi-lender consideration as the ceiling on outside-quote exposure, and the remaining gap as passive outside-quote exposure. The cap prevents active-search mass from exceeding total multi-lender consideration.

If the motive probabilities are conditional on considering multiple lenders, I first form unconditional probabilities:

$$\pi_i^P = \pi_i^M \Pr_i(\text{price reason} \mid n_i > 1),$$
$$\pi_i^F = \pi_i^M \Pr_i(\text{fear reason} \mid n_i > 1).$$

Then I apply equation (9.2). This conversion is necessary because the weights must form an exhaustive fractional assignment.

When the NSMO permits direct measurement of overlapping motives, I use

$$\pi_i^S = \min\{\pi_i^M, \pi_i^P + \pi_i^F - \pi_i^{PF}\}. \tag{J.2}$$

When the overlap is not observed, I report bounds using $\pi_i^{PF} = 0$ and $\pi_i^{PF} = \min\{\pi_i^P, \pi_i^F\}$.

J.2 Full Moment Definitions

Let ϖ_i denote the sampling weight. For any outcome Y_i and any channel $k \in \{0, \text{se}, \text{ad}\}$, define

$$\widehat{Y}^k = \frac{\sum_i \varpi_i w_i^k Y_i}{\sum_i \varpi_i w_i^k}, \qquad \widehat{Y}_\tau^k = \frac{\sum_i \varpi_i \mathbf{1}\{\tau_i = \tau\} w_i^k Y_i}{\sum_i \varpi_i \mathbf{1}\{\tau_i = \tau\} w_i^k}. \tag{J.3}$$

Table J.1. Probability Audit and Diagnostics

Issue	Diagnostic	Treatment
Different probability scales	Check whether $\pi_i^P \leq \pi_i^M$ and $\pi_i^F \leq \pi_i^M$	If violated, motive probabilities are conditional; convert them to unconditional probabilities.
Overlapping price and fear motives	Estimate or bound π_i^{PF}	Preferred correction: $\pi_i^S = \pi_i^P + \pi_i^F - \pi_i^{PF}$. Baseline correction: cap at π_i^M.
Negative residual passive mass before capping	Share with $\pi_i^P + \pi_i^F > \pi_i^M$	Set $w_i^{\text{ad}} = 0$ after capping; report the capped mass as a diagnostic.
Large residual passive mass	Mean of $w_i^{\text{ad}} = \pi_i^M - \pi_i^S$	Interpret as passive outside-quote exposure only after controlling for brokers, referrals, and lender solicitations.
Passive residual ending at non-technology lenders	Share of $\mathbf{1}\{\tau_i = N\} w_i^{\text{ad}}$	Report as diagnostic; baseline paid-channel moments use only technology-lender passive residuals.
Advertising-intensity validation	Estimate whether technology-lender passive mass is increasing in independently measured ad-tech intensity	Validate that w_i^{ad} is advertising-related rather than only a residual survey category.
Price-search risk independence	Relation between π_i^P and ex post default within score bucket	A violation implies search costs are correlated with hidden risk; report richer controls and robustness.
Implied advertising reach	Check whether $\gamma = Q_T^{\text{ad}}/(h_N A_T) \in [0, 1]$	If violated, report the violation and use bounded calibration or alternative passive-quote weights.
Captive-share feasibility	Check whether $\lambda = Q^0/[h_T + h_N M(\gamma)] \in [0, 1]$	If violated, report the violation and revisit channel weights or home-bank share measurement.

The corresponding quantities are

$$Q_\tau^k = \sum_i \varpi_i \mathbf{1}\{\tau_i = \tau\} w_i^k, \qquad Q^k = \sum_\tau Q_\tau^k. \tag{J.4}$$

These formulas apply to rates, default outcomes, approval outcomes, borrower covariates, and event-study outcomes.

The technology-lender passive-quote moments are

$$r^{\text{ad}} = \frac{\sum_i \varpi_i \mathbf{1}\{\tau_i = T\} w_i^{\text{ad}} r_i}{\sum_i \varpi_i \mathbf{1}\{\tau_i = T\} w_i^{\text{ad}}}, \qquad Q_T^{\text{ad}} = \sum_i \varpi_i \mathbf{1}\{\tau_i = T\} w_i^{\text{ad}}. \tag{J.5}$$

The corresponding non-technology passive residual is not interpreted as a non-technology advertising channel in the baseline model. It is reported as a diagnostic and used in robustness checks.

J.3 Advertising-Intensity Validation

The passive-quote residual is an identifying assumption, so I validate it using an independent measure of advertising-technology intensity. Let Z_{mt}^{ad} denote market-by-time ad-tech or data-collection intensity from the reduced-form analysis. Define technology-lender passive residual mass as

$$\bar{w}_{mt}^{\text{ad},T} = \frac{\sum_{i \in (m,t)} \varpi_i \mathbf{1}\{\tau_i = T\} w_i^{\text{ad}}}{\sum_{i \in (m,t)} \varpi_i \mathbf{1}\{\tau_i = T\}}. \tag{J.6}$$

I estimate

$$\bar{w}_{mt}^{\text{ad},T} = \rho Z_{mt}^{\text{ad}} + X_{mt}' \Gamma + \eta_m + \eta_t + u_{mt}. \tag{J.7}$$

The model predicts $\rho > 0$: passive residual mass should be higher in markets and periods with greater independent ad-tech intensity. Standard errors are clustered at the market level. This validation is not used to construct the calibrated moments. It is an overidentifying test of the interpretation of w_i^{ad} as passive advertising exposure.

As a placebo, I estimate the same regression for non-technology lenders:

$$\bar{w}_{mt}^{\text{ad},N} = \rho_N Z_{mt}^{\text{ad}} + X_{mt}' \Gamma_N + \eta_m + \eta_t + u_{mt}^N, \tag{J.8}$$

where

$$\bar{w}_{mt}^{\text{ad},N} = \frac{\sum_{i \in (m,t)} \varpi_i \mathbf{1}\{\tau_i = N\} w_i^{\text{ad}}}{\sum_{i \in (m,t)} \varpi_i \mathbf{1}\{\tau_i = N\}}. \tag{J.9}$$

The model predicts that ρ_N should be small relative to ρ.

J.4 Full Recursive Parameter Map

The rows are ordered recursively. Approval and default moments identify A_T, A_N, q, $p(\alpha)$, and Δ. Quantity moments then identify e_T, ω_T^{ad}, γ, and λ. Rate moments identify $\tilde{r}$, κ, and κ^c. The screening shock δ_α is recovered before the reach shock δ_γ because the quantity response depends on the approval effect induced by weaker screening.

Table J.2. Full Recursive Parameter Map

Parameter	Meaning	Estimator / equation	Source
d_L, d_H	Default probabilities of low- and high-risk types	Two-point mixture from default variation across score buckets; low- and high-risk benchmark buckets discipline the support	Equifax
$\mu_{H,s}$	Hidden high-risk share in score bucket s	$\mu_{H,s} = (D_{0,s} - d_L)/(d_H - d_L)$, where $D_{0,s}$ is the unscreened or home-bank default rate in bucket s	Equifax
A_T, A_N	Approval probabilities	$A_\tau = 1 - R_\tau$, checked against $A_\tau = \mu_L + \mu_H q_H(z_\tau)$	HMDA
q	Baseline high-risk survival at non-tech lenders	$q = 1 - R_{N,s}/\mu_{H,s}$	HMDA, Equifax
$p(\alpha)$	Cross-site detection probability	$p(\alpha) = (R_{T,s} - R_{N,s})/(\mu_{H,s} q)$	HMDA
D_T, D_N	Funded default rates	Weighted default rates among funded loans by lender type	Equifax
Δ	Screening cost advantage	$\Delta = D_N - D_T$	Equifax
e_T	Organic visibility of technology lenders	$e_T = Q_T^{\text{se}}/(Q_T^{\text{se}} + Q_N^{\text{se}})$	Table 9.1
e_N	Organic visibility of non-technology lenders	$e_N = 1 - e_T$	Implied
h_T, h_N	Home-bank shares	Pre-period relationship-lender or servicer shares; $h_N = 1 - h_T$	Equifax / servicer
ω_T^{ad}	Paid/passive-quote share of tech-lender acquisitions	$\omega_T^{\text{ad}} = Q_T^{\text{ad}}/(Q_T^{\text{ad}} + Q_T^{\text{se}})$	Table 9.1
γ	Targeting reach	$\gamma = Q_T^{\text{ad}}/(h_N A_T)$; bounded in $[0, 1]$, reported with feasibility diagnostics	Table 9.1, HMDA

Continued on next page

Table J.2 continued

Parameter	Meaning	Estimator / equation	Source
$M(\gamma)$	Residual non-technology home-bank mass	$M(\gamma) = \mu_L(1-\gamma)+\mu_H(1-\gamma q_H(\alpha)) = 1-\gamma A_T$	Model-implied
λ	Captive share	$\lambda = Q^0/[h_T + h_N M(\gamma)] = Q^0/(1 - h_N \gamma A_T)$	Table 9.1
$\tilde{r}$	Search-market anchor	$\tilde{r} = e_N r_N^{\text{se}} + e_T r_T^{\text{se}}$	Table 9.1
κ	Shopper advertising wedge	$\kappa = r^{\text{ad}} - \tilde{r}$	Table 9.1
κ^c	Captive home-bank wedge	$\kappa^c = \bar{r} - \tilde{r}$	Table 9.1
$f + m_0$	Funding-cost plus organic margin	Common intercept of $r^{\text{se}}_{\tau,s} - D_{\tau,s}$ with time and market fixed effects	Rates, defaults
$D'_T(\alpha)\alpha$	Local screening slope	From response of D_T to privacy-law dose; locally $\widehat{\theta}^{D_T} \approx -D'_T(\alpha)\alpha\delta_\alpha$	Sun–Abraham dose
δ_α	Law-induced screening loss	Solves observed changes in R_T and D_T using $q_H(\alpha) = q(1 - p(\alpha))$	Sun–Abraham approval and default
δ_γ	Law-induced reach loss	Solves observed decline in Q_T^{ad} and rise in Q_N^0, given δ_α, using $d(\gamma A_T) = \gamma A_T(\varepsilon_A \delta_\alpha - \delta_\gamma)$	Sun–Abraham quantities
$D''_T(\alpha)\alpha^2$	Local screening curvature	Not estimated at baseline; calibrated or bounded in robustness via quadratic dose-response when dose variation is sufficiently powered	Robustness
η	Platform take rate	Backed out from observed advertising expenditure where available; otherwise externally calibrated and varied in sensitivity analysis	External
β	First-party data intensity	Deposit or relationship intensity; used in robustness with heterogeneous first-party-data exposure	External / proxy

J.5 Overidentifying Restrictions

The primary overidentifying restriction tests the pricing structure of the model. Because posted search rates equal marginal cost plus a common differentiation markup, the posted-rate gap should equal the default-cost gap:

$$r_N^{\text{se}} - r_T^{\text{se}} = D_N - D_T. \tag{J.10}$$

This restriction is not a channel-weight identity. It links independently measured rates and default costs and therefore provides the sharpest test of the common-markup assumption.

The remaining diagnostics test quantity consistency, feasibility, and the interpretation of the passive-quote residual:

a. *Quantity consistency.* The paid share from quantities should match the paid share used in the average-rate decomposition:

$$\omega_T^{\text{ad}} = \frac{Q_T^{\text{ad}}}{Q_T^{\text{ad}} + Q_T^{\text{se}}}. \tag{J.11}$$

b. *Advertising-intensity validation.* Technology-lender passive-quote mass should be increasing in independently measured ad-tech intensity:

$$\rho > 0 \quad \text{in} \quad \bar{w}_{mt}^{\text{ad},T} = \rho Z_{mt}^{\text{ad}} + X_{mt}'\Gamma + \eta_m + \eta_t + u_{mt}. \tag{J.12}$$

The corresponding relationship should be weak or absent for non-technology lenders.

c. *Reach feasibility.* The implied targeting reach must satisfy

$$0 \le \gamma = \frac{Q_T^{\text{ad}}}{h_N A_T} \le 1. \tag{J.13}$$

d. *Captive-share feasibility.* The implied captive share must satisfy

$$0 \le \lambda = \frac{Q^0}{h_T + h_N M(\gamma)} \le 1. \tag{J.14}$$

e. *Contract-stock placebo.* The Sun–Abraham effect on predetermined incumbent rates or locked contracts should be close to zero. This is a placebo for contract repricing, not a restriction on the new-origination equilibrium.

f. *Return-home effect.* The Sun–Abraham effect on Q_N^0 should be positive when Q_T^{ad} falls:

$$\widehat{\theta}^{Q_T^{\text{ad}}} < 0 \quad \Rightarrow \quad \widehat{\theta}^{Q_N^0} > 0. \tag{J.15}$$

g. *Exclusion versus surcharge.* Conditional on score and observables, default in the passive-quote cell should not exceed default in the organic-search cell. If it does, the noisy-flag or surcharge extension is required.

h. *Search-approval test.* The fear-search probability should predict lower approval and higher default within credit-score buckets.

i. *Probability audit.* Results should be stable when using alternative overlap corrections for π_i^P and π_i^F.

j. *Passive-residual diagnostic.* The passive residual among non-technology lenders should be small or explained by brokers, referrals, or other non-advertising solicitation channels.

J.6 Inference

The generated weights w_i^0, w_i^{se}, and w_i^{ad} are first-stage objects. The preferred inference procedure is a two-layer bootstrap. In each bootstrap replication, I re-estimate the three propensity models, reconstruct the channel weights, recompute weighted rate and quantity moments, invert the model parameters, and re-run the Sun–Abraham event-study block.

At a minimum, standard errors cluster at the state-time level of treatment assignment and account for survey sampling weights. I also report sensitivity to first-stage uncertainty by comparing estimates that hold the channel weights fixed with estimates that re-estimate the weights in each bootstrap replication.