When Privacy Protects but Excludes: Screening, Enforcement, and Credit Access in Digital Markets*

Sumit Agarwal Pulak Ghosh Peiyi Jin
Shohini Kundu Nishant Vats Xinbo Wang

Abstract

This paper examines how privacy regulation can inadvertently constrain financial inclusion and lender profitability. We study Google's 2019 restriction on FinTech lenders' access to call detail records (CDRs) in India as a natural experiment. Using comprehensive loan-level data and a difference-in-differences design, we find that the policy increased loan applications by 21%, consistent with heightened borrower demand for privacy, but reduced approval rates by 18 pp. However, default rates remained stable. These patterns suggest that lenders replaced social collateral based enforcement using CDR with stricter ex-ante screening to maintain portfolio quality. This adjustment disproportionately restricted access for new-to-credit, lower-income, younger, and socially disadvantaged borrowers, whose approval rates fell by 7–15 pp. Using credit bureau data, we show that rejected applicants following stricter privacy regulation were 4–5 pp less likely to obtain formal credit even four years later, indicating long-term effects of privacy regulation on exclusion from credit markets. Lender profits declined by nearly 30%, reflecting the loss of profitable but underserved borrowers. Overall, our results suggest a tension between protecting consumer privacy and promoting financial inclusion.

*Agarwal: National University of Singapore (ushakri@yahoo.com); Ghosh: Indian Institute of Management Bangalore (pulak.ghosh@iimb.ac.in); Jin: National University of Singapore (jin_peiyi@u.nus.edu); Kundu: UCLA Anderson School of Management and CEPR (shohini.kundu@anderson.ucla.edu); Vats: Olin Business School, Washington University at St Louis (vats@wustl.edu); Wang: National University of Singapore (xinbo_wang@u.nus.edu)

We thank Giorgia Barboni, Kim Fe Cramer, Nicola Limodio, Sean Higgins, Amiyatosh Purnanandam, Tyler Muir, Huan Tang, Michaela Pagel as well as seminar and conference participants at UCLA Anderson Conference and the CEPR-WEFIDEV Conference for helpful comments. We do not have any conflicts of interest to disclose. We take responsibility for all errors.

1 Introduction

The rapid expansion of artificial intelligence (AI) and digital technologies has fundamentally transformed how lenders assess borrower risk. While alternative data have been used primarily to screen loan applicants at origination, the ability of alternative data to accurately classify borrower risk is often limited, particularly for economically and socially disadvantaged individuals (Blattner and Nelson, 2021). In response, many FinTech lenders, especially in emerging economies where traditional credit infrastructure remains underdeveloped, have extended their use of alternative data beyond initial screening to include ongoing monitoring or disciplining throughout the loan lifecycle. These lenders leverage diverse data sources, such as call detail records (CDRs), mobile phone usage patterns, and other digital behavioral traces to create an implicit social collateral that reduces default risk. These innovations promise to advance financial inclusion for low-income and marginalized populations who have historically been excluded from formal credit markets.[1]

Yet, the increasing reliance on sensitive personal data has heightened regulatory concerns about privacy. As a result, a wave of privacy regulations have emerged, restricting FinTech lenders' access to alternative data sources that were previously central to their business models. Prior research has largely documented the positive effects of these regulations on consumers, highlighting benefits such as improved data security and increased consumer autonomy (Tang, 2019; Bian, Ma and Tang, 2021; Lin, 2022; Bian, Pagel, Tang and Raval, 2023). However, these studies largely overlook the potential tradeoffs: while privacy regulations increase data protection, they may also constrain lenders' ability to make profitable loans by narrowing the set of observable borrower characteristics they can rely on, which can result in a reduction in credit supply. The contraction in credit supply disproportionately affects low-income and marginalized borrowers – precisely those populations who stand to benefit most from FinTech innovation, including the use of alternative data. Additionally, these dynamics due to restrictions on data sharing may affect the competitiveness of Fintechs that specialize in using such data to make lending decisions.

This paper examines the tradeoff between consumer protection and financial inclusion, and its implications for lender profitability, leveraging a policy shock that curtailed the use of alternative data in credit markets. The setting focuses on a large Indian FinTech lender that, as of January 1, 2019, lost access to borrowers' call detail records (CDR) following a Google privacy policy directive. Prior to this change, CDR data allowed the lender to use borrowers' social networks as a form of social collateral. For instance, in cases of delinquency, collection agents could contact close friends, family members, or professional associates. This mechanism creates an ex-post threat improving repayment discipline and keeping default rates low (Karlan, Mobius, Rosenblat and Szeidl, 2009; Bryan, Karlan and Zinman, 2015). Google's policy change, thus, creates a natural experiment for assessing the economic costs of privacy restrictions and identifying the channels through which alternative data affects credit market outcomes. Using comprehensive loan-level

[1] See Thakor (2020); Allen, Gu and Jagtiani (2021) and Berg, Fuster and Puri (2022) for reviews.

administrative data spanning 2017-2022, we exploit the platform-specific nature of Google's policy – which affected Android applications but left iOS applications unaffected – to identify the causal effects of privacy regulation on both sides of the credit market and its distributional consequences across borrower groups. Importantly, the loss of CDR data eliminated the lender's primary ex-post enforcement mechanism. This credible threat of contacting borrowers' social networks in cases of delinquency had previously been central to disciplining borrowers and containing default risk.

Our empirical strategy combines difference-in-differences and high-frequency event study designs to identify the effects of privacy regulation. The difference-in-differences approach compares Android users (treated by the policy) to iOS users (control group) around the policy implementation date, leveraging the fact that both groups were served by the same lender with identical products and pricing structures prior to the policy change. We validate this identification strategy by demonstrating parallel trends in lending outcomes between Android and iOS users in the months leading up to the policy change, and by showing that the policy did not alter the relative composition of applications across platforms. We complement this with a high-frequency event study that examines daily patterns in lending outcomes for Android users around January 1, 2019, providing additional evidence on the precise timing and robustness of the policy's effects.

We begin by examining the response of borrowers to this regulation by estimating the effect on their application behavior. Following the reform, loan applications from Android users increased by roughly 21 percent relative to iOS users. The increase in the number of applications is consistent with the evidence from the existing literature which shows that consumers place a meaningful value on privacy protections (Tang, 2019; Lin, 2022). The magnitude of this response indicates that data-sharing requirements represent a meaningful factor in financial participation decisions.

Next, we turn to the lender's response, examining how the privacy regulation affected screening and contract terms. Following the reform, acceptance rates for Android applicants declined by 18 percentage points relative to iOS applicants. Among approved borrowers, Android users faced somewhat higher interest rates (by 0.25–0.35 percentage points) and shorter maturities (by 0.10–0.12 months). These findings show that the lender's response to information loss operated through multiple channels: primarily through quantity rationing (lower acceptance rates), but also through modest price and maturity adjustments for approved borrowers. The dominance of quantity rationing over pricing adjustments suggests that the social collateral embedded in CDRs was particularly valuable for lending to borrowers with high information asymmetry. By enabling credible ex-post enforcement, social collateral lowered the perceived risk of such loans and allowed the lender to extend credit to such borrowers.

Taken together, these results suggest that lenders respond to privacy regulation by fundamentally shifting their business model from social-collateral based enforcement to screening. In the pre-regulation period, the lender used CDRs to construct maps of borrowers' social networks, creating an implicit social collateral mechanism whereby the lender could contact family members,

friends, and professional associates. This enforcement technology allowed the lender to extend credit to borrowers who may appear risky on traditional metrics, knowing they could leverage social pressure to encourage repayment. The privacy regulation's removal of this enforcement technology forced the lender to compensate by substantially tightening ex-ante screening standards to maintain portfolio quality. We provide supporting evidence of this by examining the effect on default rates. We do not find any statistically significant or economically meaningful effects on default rates. Hence, our findings suggest that once deprived of post-disbursement social enforcement, the lender became far more selective at the application stage to maintain portfolio quality. As a result, many borrowers who would have been accepted under the prior system, when CDR-based social collateral provided ongoing repayment discipline, were now rejected.

We then examine the distributional consequences of this policy shift. We hypothesize that the privacy regulation disproportionately affects borrower groups most reliant on sources of alternative information. Specifically, we find that low-income applicants, younger borrowers, first-time applicants, and individuals from marginalized social backgrounds experience declines in acceptance rates that are 7-15 percentage points larger than their more advantaged counterparts. By contrast, borrowers with established credit histories or prior relationships with the lender are largely insulated from these effects. This heterogeneity highlights that privacy protections can inadvertently exacerbate financial exclusion by limiting access for precisely those groups that face the greatest barriers to traditional credit.

We further investigate whether these immediate effects translate into persistent financial exclusion by linking our dataset to comprehensive credit bureau records covering the universe of formal lending activity in India. The credit bureau linkage serves two central purposes. First, it allows us to directly test whether privacy-induced rejections represent financial exclusion versus mere substitution across lenders. Second, it enables us to assess longer-term effects of the policy on borrowers' access to and use of credit. For example, we can track whether borrowers excluded after the privacy policy remained systematically locked out of the financial system, or whether they eventually re-entered through other channels.

Using both reduced-form difference-in-differences estimates and instrumental variables analysis that exploits the policy change as an exogenous shock to loan approval decisions, we find that borrowers whose applications were rejected due to privacy restrictions remain 4-5 percentage points less likely to access formal credit even four years later. Our IV estimates suggest that loan approval by our focal lender is associated with a 10-12 percentage point increase in the likelihood of long-term credit access. This result indicates that rejection of first-time applicants can create a persistent pattern of financial exclusion, consistent with a "FinTech ladder effect" whereby initial access to digital credit facilitates subsequent access to the broader formal financial system.

Lastly, we quantify the economic effect of the stricter screening induced by privacy regulation on lender profitability. The net effect of privacy regulation on lender profitability remains ambiguous ex-ante. On the one hand, privacy restrictions force the lender to adopt a more selec-

tive screening strategy, as it can no longer rely on post-disbursement enforcement via CDRs. This shift leads to the rejection of many borrowers who would likely have been profitable under the earlier model, where CDRs provided ongoing repayment discipline through social collateral. On the other hand, the regulation simultaneously expands application volumes, potentially enlarging the pool of borrowers from which the lender could select.

To assess the economic magnitude of these effects, we construct a counterfactual framework that estimates the lender's foregone profits from excluding borrowers who, absent the privacy restrictions, would have been deemed creditworthy due to the enforcement value of social collateral. Our analysis identifies "marginal borrowers" those rejected post-policy who would have been approved under pre-policy standards and calculates the expected profitability of these foregone lending opportunities. We find that privacy regulation reduces lender profits by approximately 29% in our baseline specification, with an estimated annual effect of nearly ₹100 million ($5.3 million in PPP terms) for this single institution. Specifically, we find that the foregone profitability is primarily driven by borrowers who are economically and socially disadvantaged, including female, younger individuals, and those with lower incomes or credit scores.

Overall, our findings emphasize the fundamental tradeoff facing privacy regulation in digital finance. On one hand, consumers clearly value stronger protections, as reflected in the surge in loan applications following the policy change. On the other hand, blanket restrictions on alternative data can unintentionally deepen financial exclusion by removing precisely the information needed to discipline underbanked borrowers post-disbursement. The substantial foregone profits we document, nearly 30% of baseline earnings, concentrated among female, younger, and lower-income borrowers, demonstrate that privacy restrictions eliminate economically valuable lending opportunities precisely among the populations that innovations in FinTech were designed to serve. Our findings help inform future policy design that considers both consumer protection and financial inclusion objectives simultaneously.

1.1 Related Literature

The key contribution of this paper is to examine a key tradeoff of privacy regulation. While privacy regulations can improve consumer protection, they can also have a negative effect on financial inclusion by narrowing the set of observable borrower characteristics lenders can rely on, resulting in a reduction in credit supply. While the theoretical literature have discussed these tradeoffs, as reviewed by Acquisti, Taylor and Wagman (2016), Bergemann and Morris (2019) and Goldfarb and Tucker (2019), the extant empirical literature has primarily focused on the value consumers attribute to data privacy highlighting the benefits of privacy regulation to consumers (Tang, 2019; Bian, Ma and Tang, 2021; Lin, 2022; Bian, Pagel, Tang and Raval, 2023). We contribute to this literature by documenting the tradeoffs associated with privacy regulation in credit markets. On the borrower side, we document an increase in the number of applications consistent with the idea that consumers place a positive value on privacy. On the lender side, we document a reduction

in credit supply, primarily for socially and economically marginalized groups and new-to-credit borrowers. To highlight the importance of the latter group: 50% of the credit-eligible population in India is underserved due to being new-to-credit, accounting for a total of 572 million individuals (Cramer, Ghosh, Kulkarni and Vats, 2024).

Empirically, our paper is closest to Doerr, Gambacorta, Guiso and Sanchez del Villar (2023) which shows that privacy restrictions decreased FinTech lending, disproportionately affecting marginal borrowers. We contribute by providing the first micro-level analysis within a single lender that quantifies foregone profits from excluding creditworthy borrowers and separates demand-side effects (borrowers valuing privacy) from supply-side effects (lenders losing screening ability). Importantly, we demonstrate that the primary mechanism is tightened ex-ante screening. Our empirical results allow us to isolate ex-ante screening from ex-post enforcement effects. CDRs primarily improved borrower discipline, and the removal of lender's access to this enforcement technology lead to a reduction in credit supply via stringent ex-ante screening. We show that social information remains economically meaningful even in data-rich FinTech environments where lenders have extensive traditional data. The dramatic effects of removing CDRs, despite continued access to credit scores and income verification, demonstrate that social data addresses frictions beyond information asymmetry.

We further contribute to the work of Doerr, Gambacorta, Guiso and Sanchez del Villar (2023) by documenting the importance of access to Fintech credit. Prior work suggests that initial access to FinTech loans can serve as a stepping stone toward future credit, often described as the *FinTech ladder effect*, either by helping borrowers accumulate collateral (Beaumont, Tang and Vansteenberghe, 2022) or by generating hard information that improves creditworthiness (Balyuk, 2023; Dubey, 2024). In contrast, Chava, Ganduri, Paradkar and Zhang (2021) finds that marketplace lending can impede future credit access. We contribute to this literature in two ways. First, the privacy regulation we study induced an exogenous contraction in FinTech loan approvals, enabling identification of the effect of FinTech credit access on subsequent borrowing. Our IV estimates indicate that receiving a loan from a FinTech lender raises the probability of long-term credit access by 10–12 percentage points. Second, our results imply that privacy-induced restrictions on alternative data can amplify financial exclusion, curtailing pathways into the formal credit system and limiting the long-term inclusion benefits of FinTech innovation.

Lastly, we document the negative effect of this policy on the lender's profitability. The contraction in credit supply due to privacy regulation dominates the positive effect of this regulation on consumer demand leading to a net negative effect on lenders profitability. Put more directly, the privacy regulation led to the rejection of many borrowers who would likely have been profitable under the earlier model, where CDRs provided ongoing repayment discipline through social collateral. Prior work has primarily focused on documenting the negative effect of privacy regulation on firms' web traffic and application installation (Goldberg, Johnson and Shriver, 2019; Aridor, Che and Salz, 2020; Bessen, Impink, Reichensperger and Seamans, 2020; Bian, Ma and Tang, 2021; Jia,

Jin and Wagman, 2021; Janßen, Kesler, Kummer and Waldfogel, 2022; Kraft, Skiera and Koschella, 2023). Other works examine how firms modify their data-collection and data-sharing processes following privacy regulations (Peukert, Bechtold, Batikas and Kretschmer, 2022; Bae, Mayya and Nian, 2023; Kesler, 2023; Demirer, Hernández, Li and Peng, 2024; Ramadorai, Uettwiller and Walther, 2025). A notable exception is Abis, Tang and Bian (2025), which documents the negative effect of privacy regulation on market efficiency. We add to this literature by documenting the negative effect of privacy regulation on the profitability of Fintech platforms. Specifically, we show that privacy regulation can reduce profitability by eroding information environments that can be key for borrower discipline. These results echo the key idea that alternative data can create value for firms as in Veldkamp and Chung (2024), and our exercise closely resembles the revenue approach of quantifying the value of data outlined in Veldkamp (2023) and Farboodi, Singal, Veldkamp and Venkateswaran (2025).

The paper proceeds as follows. Section 2 describes the institutional setting and privacy policy change. Section 3 presents the data and Section 4 outlines the empirical methodology. Section 5 reports the main results, documenting a tightening of credit standards that disproportionately affects borrowers from disadvantaged backgrounds. Section 6 examines the underlying mechanisms, showing that tighter ex-ante screening substitutes for ex-post enforcement, and finds that excluded borrowers cannot obtain credit elsewhere. Section 7 quantifies the economic costs of tightened screening ability. Section 8 concludes.

2 Institutional Setting

Our analysis focuses on a prominent FinTech lender operating in India's digital lending ecosystem. The company was established in 2016 and represents a new generation of alternative lending platforms that emerged to serve India's underbanked population. The institutional setting is particularly relevant for understanding privacy regulation effects on financial inclusion, as this lender was among the first to employ alternative data to discipline borrower behavior after loan disbursement.

2.1 Business Model and Target Market

The FinTech operates primarily through a mobile application platform, providing unsecured short-term personal loans. The company specifically targets millennials and individuals from marginalized backgrounds who are often excluded from traditional banking services due to limited or no credit history. Loan amounts range from approximately ₹10,000 to ₹300,000 (roughly $120 to $3,600), with repayment tenures spanning from 15 days to 18 months. The average loan amount is approximately ₹35,000-40,000 ($420-480).

The lender's primary value proposition centers on rapid loan processing and disbursement, with the ability to approve and disburse loans within minutes of application submission. This

speed advantage is critical in the competitive Indian FinTech market, where traditional banks often require weeks for loan processing. The company processes thousands of loan applications daily and has maintained high borrower retention rates, with approximately 70% of borrowers returning for additional loans.

2.2 Call Detail Records and Social Enforcement System

Traditionally, lenders have leveraged alternative data primarily to screen loan applicants at origination. However, the ability of these models and data sources to classify borrower types may be limited, especially for economically and socially disadvantaged borrowers who lack extensive traditional credit profiles (Blattner and Nelson, 2021). As a result, many FinTechs have expanded their use of alternative data beyond screening to include ongoing enforcement capabilities.

Call detail records (CDRs) represent a particularly rich example of such datasets, enabling lenders to discipline borrower behavior through multiple channels beyond traditional screening. Our institutional setting is distinctive in how extensively the lender integrated CDRs into their credit evaluation framework. Rather than using CDRs merely as supplementary information, this FinTech developed them into a central pillar of their alternative data strategy, creating a comprehensive enforcement system that tracked borrowers' social networks and communication patterns throughout the loan lifecycle. Prior to the January 2019 privacy policy change described below, these records enabled comprehensive analysis of borrowers' social networks and communication patterns, providing the lender with ongoing insights into borrower behavior throughout the loan lifecycle.

2.2.1 Structure of CDR-Based System

The CDR-based system operated on multiple analytical dimensions. First, relationship mapping allowed the lender to construct comprehensive social network graphs for each borrower, categorizing contacts into specific relationship types including immediate family (mother, father, spouse, siblings), extended family (aunts, uncles, grandparents), professional contacts (work colleagues), and close friends. This categorization enabled the lender to gauge the borrowers' social capital and track the strength of their network, a key input into its enforcement strategy.

Second, communication pattern analysis provided dynamic insights into borrower behavior around loan events (Agarwal, Alok, Ghosh and Gupta, 2020). The system tracked communication frequency, duration, and directionality with each relationship category, both before and after loan origination. The lender analyzed various communication patterns, such as changes in calling frequency, timing shifts, missed versus completed calls, and communication behavior around repayment dates, as potential indicators of borrower circumstances. Changes in communication patterns could signal financial stress, social isolation, or attempts to avoid collections contact.

Third, the monitoring capability enabled the lender to observe real-time changes in borrower

social behavior throughout the loan lifecycle. The system continuously updated communication metrics, allowing for early warning signals of potential default risk. Borrowers who exhibited declining social connectivity or changed communication patterns could be flagged for proactive intervention or modified collection strategies.

2.2.2 Empirical Evidence of Predictive Power and Role of Social Collateral

To establish the empirical importance of CDRs, Table 1 presents the relationship between contact list size and default probability for loans originated between January 2017 and December 2018 – before the privacy policy was implemented. Our most stringent specification includes ZIP code × loan purpose × month-year fixed effects, effectively comparing borrowers in the same ZIP code applying for the same loan type in the same time period. This approach ensures that the relationship is not driven by geographic, product-specific, or temporal factors that might confound the analysis.

The results reveal a strong negative correlation between contact network size and default probability. A 1% increase in the number of contacts in a borrower's contact list is associated with a 1.45 percentage point reduction in default probability in the baseline specification. This effect strengthens to 1.71 percentage points when demographic and loan-level controls are included, representing an 8% decrease relative to the average default probability. The relationship remains robust across all specifications, demonstrating that contact data provided genuine predictive value beyond what could be captured by location, loan characteristics, or borrower demographics.

The mechanism underlying this relationship operates through several channels that explain why contact list size predicts repayment behavior. First, larger contact networks reflect greater social integration and community ties, which create reputational incentives for loan repayment. Borrowers with extensive social connections face higher social costs from potential default, as non-payment could damage relationships and social standing within their community (Karlan, Mobius, Rosenblat and Szeidl, 2009; Bryan, Karlan and Zinman, 2015). Second, the lender's explicit policy of contacting individuals from the borrower's contact list in cases of delinquency created a credible threat mechanism. Borrowers understood that default would result in social embarrassment through contact with family members, friends, and professional colleagues, thereby strengthening repayment incentives. Third, contact list characteristics serve as a proxy for unobserved borrower quality dimensions. Individuals with larger, more diverse contact networks may possess better social skills, professional networks, and support systems that correlate with financial responsibility and employment stability. The diversity and nature of contacts, including the presence of professional contacts, family members, and community connections, provide additional information about the borrower's social capital and economic prospects that traditional credit metrics fail to capture.

This system also enabled sophisticated collection strategies. Rather than relying solely on direct borrower contact, the lender could strategically engage different relationship categories

based on communication patterns and relationship strength. Family members with frequent, long-duration calls might be prioritized as potential sources of repayment assistance, while professional contacts could be leveraged to create workplace pressure for repayment. The mere knowledge that the lender possessed detailed contact information for borrowers' entire social networks could influence repayment behavior even without actual contact occurring.

The strength and consistency of the contact-default relationship across the pre-policy period validates the lender's strategic emphasis on CDR data collection. Contact information was not merely auxiliary data but constituted a central component of the credit assessment algorithm, with significant predictive power for loan performance demonstrated empirically in Table 1. This evidence demonstrates why the sudden prohibition of CDR access represented a substantial shock to the lender's underwriting capabilities, forcing significant adjustments to risk assessment and loan approval processes.

The documented relationship between contact network size and default probability provides context for interpreting the effects of the privacy policy change analyzed in subsequent sections. Given the strong predictive value of contact data demonstrated in the pre-policy period, the removal of this information source would be expected to materially impact the lender's ability to assess credit risk, potentially leading to changes in approval rates, pricing, or borrower selection criteria as the lender adapts to operating with a more limited information set.

2.3 Changes to Privacy Policy

In January 2019, Google implemented a privacy policy directive that fundamentally altered the data collection landscape for Android applications. The directive specifically prohibited apps distributed through the Google Play Store from collecting CDRs from borrowers' mobile devices. Critically, this policy change affected *only* Android users, as iOS applications distributed through Apple's App Store faced no such restriction. This differential treatment creates a natural experiment that forms the foundation of our identification strategy.

The enforcement mechanism was particularly significant given that Android devices dominated the lender's user base. Non-compliance would result in the immediate removal of the lender's application from the Google Play Store, effectively eliminating their primary borrower acquisition and service delivery channel. Faced with this ultimatum, the lender had no viable alternative but to comply immediately with the Android restrictions. Consequently, the collection of call log data and related CDR information from Android users ceased entirely on January 1, 2019, while iOS users remained unaffected.

This platform-specific policy change provides an ideal natural experiment for three key reasons. First, the timing was exogenous to both borrower behavior and the lender's business strategy, as it was imposed by an external platform provider (Google) rather than emerging from internal business decisions or broader regulatory changes. Second, the differential treatment of Android versus iOS users creates a clean treatment and control group structure. Android users

lost privacy protections (treated group) while iOS users maintained the same data collection regime (control group). Third, the binary nature of the change, from comprehensive CDR access to complete prohibition for Android users only, creates a sharp discontinuity that affected all Android loan applications processed after the implementation date.

The impact of this change was substantial for Android users. The lender lost access to what had been their most sophisticated disciplining and enforcement tool for this segment. The social collateral mechanism that had potentially suppressed default rates through implicit social pressure was eliminated for Android borrowers. Collection strategies had to be redesigned to focus more heavily on ex-ante screening rather than ex-post network-based disciplining for Android users, while iOS users experienced no such change.

Importantly, all other aspects of the lender's operations remained constant across both platforms. The core business model, target market, loan products, pricing structures, approval algorithms (aside from the CDR component), and operational procedures were identical for Android and iOS users both before and after the policy change. The only systematic difference introduced on January 1, 2019 was the loss of CDR access for Android users. This ensures that the policy change represents the primary source of variation in lending outcomes during our study period, allowing us to cleanly identify the causal effect of privacy regulation on credit access.

3 Data

Our empirical analysis leverages a uniquely rich dataset spanning 2017–2022 from one of India's largest and most innovative FinTech lenders. The scale, granularity, and multidimensional nature of this dataset allow us to assess the effect of privacy regulations on lending outcomes in one of the most rapidly expanding credit markets globally. Importantly, the dataset captures millions of loan applications from both Android and iOS users observed both before and after the January 2019 Google privacy policy change, which restricted access to CDRs only for Android applications. This platform-specific restriction creates an ideal natural experiment where iOS users serve as a control group, allowing us to isolate the causal effects of privacy regulation from other contemporaneous market trends.

Borrower Characteristics Data. The backbone of our dataset is a comprehensive record of every loan application submitted to the FinTech. Each observation includes unique borrower and loan identifiers, detailed loan-level information (loan type, requested amount, interest rate, approval/rejection outcome), time stamps, and importantly, the operating system used to submit the application. This platform indicator allows us to distinguish between treated (Android) and control (iOS) users in our difference-in-differences analysis.

The dataset also provides rich borrower demographics including income, gender, caste, age, and credit scores. These variables allow us to rigorously examine heterogeneous effects of privacy

restrictions across population subgroups, with particular attention to traditionally underserved and marginalized communities that the FinTech explicitly sought to serve. The scale and detail of these demographic data make it possible to quantify distributional consequences of privacy regulation and to test whether vulnerable groups are disproportionately affected.

Credit Bureau Data. To extend the scope of our analysis beyond the focal FinTech, we link the application-level dataset to comprehensive credit bureau records covering the universe of formal lending activity in India. This linkage enables us to track borrowers' trajectories across the entire formal credit market, not just within a single institution. With this integrated dataset, we can evaluate whether applicants rejected post-policy were able to secure credit elsewhere, or whether they remained excluded from formal finance altogether.

Essentially, the credit bureau linkage serves two central purposes. First, it allows us to directly test whether privacy-induced rejections represent financial exclusion versus mere substitution across lenders. Second, it enables us to assess longer-term effects of the policy on borrowers' access to and use of credit. For example, we can track whether borrowers excluded after the privacy policy remained systematically locked out of the financial system, or whether they eventually re-entered through other channels.

The combination of (i) large-scale loan application data with detailed demographic attributes and platform identifiers, and (ii) comprehensive credit bureau coverage of the formal lending sector, provides a valuable empirical foundation for examining privacy regulation effects. The availability of both internal lending decisions and external credit market outcomes allows us to investigate not only how privacy restrictions influence lending decisions within a single FinTech, but also whether these effects manifest as broader patterns of financial exclusion or redistribution across the credit ecosystem.

4 Empirical Methodology

Our empirical strategy exploits the sharp discontinuity created by Google's privacy policy directive implemented on January 1, 2019, which prohibited Android applications from collecting CDRs while leaving iOS applications unaffected. This platform-specific policy change provides a clean natural experiment for identifying the effects of privacy regulation on lending outcomes, as the timing was exogenous to both borrower behavior and the lender's business strategy.

We employ two complementary identification approaches that together provide robust evidence on the causal effects of privacy regulation. Our primary strategy is a difference-in-differences framework that compares Android users (treated) to iOS users (control) around the policy change. This approach addresses concerns about confounding factors—such as general market trends, macroeconomic conditions, or other contemporaneous policy changes—by using iOS users as a counterfactual for what would have happened to Android users absent the privacy policy.

As a complementary analysis, we implement a high-frequency event study that focuses exclusively on Android users in a narrow time window around the January 1, 2019 cutoff. This complementary approach addresses distinct econometric concerns: it isolates the immediate discontinuous response to the policy change, relies on weaker functional form assumptions than the monthly DiD analysis, and provides a robustness check that does not depend on iOS users serving as an appropriate control group.

After establishing the aggregate effects of privacy regulation using these two approaches, we then examine heterogeneous effects across borrower characteristics to understand which groups are most adversely affected by the loss of alternative data sources.

4.1 Difference-in-Differences Design

Our primary identification strategy relies on a difference-in-differences framework that compares lending outcomes for Android users (who lost CDR access after January 1, 2019) to iOS users (who were unaffected by the policy change). This approach exploits the platform-specific nature of Google's privacy directive to construct a natural treatment and control group structure.

We employ a within-estimator that compares Android and iOS users within the same ZIP code, loan purpose, and month-year. This approach absorbs a very high-dimensional set of confounding factors. Specifically, the three-way fixed effects control for: (i) any time-invariant differences between Android and iOS users within a given location and loan type, (ii) any location-specific time trends that affect both platforms equally, and (iii) any loan-purpose-specific temporal patterns (e.g., seasonal demand for certain loan types) that are common across platforms. The identifying variation, therefore, comes from comparing Android to iOS users who are applying for the same type of loan, in the same geographic area, at the same point in time, with the only differential shock being the loss of CDR access for Android users.

We estimate the following baseline specification:

$$Y_{i,t} = \alpha + \beta \cdot (Android_i \times Post) + \gamma \cdot Android_i + \delta \cdot Post_t + X'_{i,t}\lambda + \theta_{z,p,t} + \epsilon_{i,t}.$$

where $Y_{i,t}$ represents the outcome of interest for loan application i submitted at time t. The key parameter of interest is β, which identifies the causal effect of the privacy policy on Android users relative to iOS users. $Android_i$ is an indicator equal to one if the application was submitted through an Android device and $Post_t$ is an indicator equal to one for all applications submitted on or after January 1, 2019. The interaction term, $Android_i \times Post_t$ captures the differential effect of the policy on Android users. $X_{i,t}$ includes individual-level controls such as loan purpose, loan maturity, and borrower demographics (age, gender, income, caste, education, credit scores). Our most comprehensive specification includes $\theta_{z,p,t}$, which represents ZIP code $\times$ loan purpose $\times$ month-year fixed effects. These three-way fixed effects control for location-specific temporal trends that vary by loan type, absorbing any confounding variation from geographic differences,

product-specific dynamics, or time-varying factors that might correlate with both platform choice and lending outcomes.

Our main outcome variables include: (1) the natural logarithm of monthly loan applications to capture demand effects, (2) a binary indicator for loan approval to measure supply-side responses, and (3) subsequent default rates to examine ex-post risk. We focus our analysis on a 180-day window on either side of the policy change (roughly six months before and after) to balance sample size considerations with the need to isolate the policy effect from longer-run trends.

The key identifying assumption underlying our difference-in-differences approach is the parallel trends assumption: absent the privacy policy change, lending outcomes for Android and iOS users would have evolved similarly over time. This assumption is plausible for several reasons. First, the policy change was imposed externally by Google with no advance notice and was unrelated to underlying trends in the lending market or borrower composition across platforms. Second, Android and iOS users were served by the same lender with identical products, pricing structures, approval algorithms (aside from the CDR component), and operational procedures prior to the policy change. Third, the lender made no other changes to its business model or lending practices differentially affecting one platform versus the other around the policy implementation date. We present formal tests of the parallel trends assumption in Section 5.2, where event study estimates confirm that Android and iOS users exhibited parallel pre-treatment trends.

4.2 High-Frequency Event Study Design

While our primary DiD analysis provides the most credible identification by controlling for confounding time trends, we complement this approach with a high-frequency event study that focuses exclusively on Android users in a narrow time window around the January 1, 2019 policy implementation.

This complementary high-frequency analysis addresses several limitations that the monthly DiD approach may not fully resolve. First, by exploiting daily variation in lending outcomes, the event study captures the *instantaneous* adjustment in lender behavior when CDR access was removed on January 1, 2019. Whereas the monthly DiD design provides robustness to confounding trends, its temporal aggregation may smooth over short-run dynamics and mask how quickly lenders responded to the information shock. Second, the daily event study rests on a distinct identification assumption – continuity of potential outcomes around the exact policy threshold – which allows us to detect discontinuous jumps in outcomes without relying on iOS users as a comparison group. Third, using daily data rather than monthly aggregates enables finer-grained inference about the timing, speed, and magnitude of the policy's effects, revealing an immediate response at the daily level that would be attenuated in coarser data. Finally, because this design compares Android borrowers to themselves before and after the cutoff, it mitigates concerns that unobserved differences between Android and iOS users could bias the DiD estimates, offering an internally valid test of the policy's causal impact.

We estimate the following specification for Android users only:

$$Y_{i,t} = \alpha + \gamma \cdot Post_t + f(distance_t) + X'_{i,t}\beta + \theta_{z,p} + \theta_d + \epsilon_{i,t}$$

where $Y_{i,t}$ represents the outcome of interest for loan application i submitted on date t. The key parameter of interest is γ, which identifies the effect of the privacy policy implementation. Post_t is a binary indicator equal to one for all loan applications submitted on or after January 1, 2019. $f(\text{distance}_t)$ controls for smooth functions of time around the discontinuity threshold. Following best practices, our baseline specification employs a local linear polynomial estimated separately on each side of the threshold, using a triangular weighting kernel (Calonico, Cattaneo and Titiunik, 2014; Gelman and Imbens, 2019; Cattaneo, Idrobo and Titiunik, 2024).

$X_{i,t}$ includes individual-level controls such as requested loan amount, borrower demographics (age, gender, income, caste), credit scores, and indicators for prior relationships with the lender. $\theta_{z,p}$ represents ZIP code × loan purpose fixed effects that control for unobserved heterogeneity across geographic areas and loan types. θ_d denotes day-of-the-week fixed effects to account for systematic weekly patterns in loan applications and processing.

We estimate this specification across multiple bandwidths (10, 15, 20, 25, and 30 days on either side of the January 1, 2019 threshold) to ensure robustness. This approach balances the bias-variance tradeoff inherent in these designs: narrower bandwidths reduce potential bias from functional form misspecification but increase variance due to smaller sample sizes, while wider bandwidths provide more precise estimates but may be more susceptible to confounding factors. We verify that our results are invariant to bandwidth choice by examining coefficient stability across the full range of bandwidths.

The validity of our high-frequency event study rests on two key identifying assumptions. First, the privacy policy change must represent a true discontinuity in the lender's information environment while leaving other relevant factors unchanged. This assumption is credible given that the policy was imposed externally by Google with no advance notice, was unrelated to the lender's internal operations or market conditions, and created an immediate and complete prohibition on CDR collection.

Second, there should be no systematic manipulation of the assignment variable (application date) around the threshold. This assumption requires that borrowers could not strategically time their applications to fall before or after the policy implementation date. The lack of advance public notice about the policy change, combined with the focus on minutes-to-hours loan processing times, makes strategic timing highly unlikely. While we do observe increased application volume after the policy change, this reflects a genuine aggregate demand response to enhanced privacy protections rather than individual-level manipulation that would threaten our identification strategy.

4.3 Heterogeneous Effects Across Borrower Characteristics

After establishing the aggregate effects of privacy regulation using these two approaches, we examine heterogeneous effects across borrower characteristics, conditional on detecting significant policy impacts, to understand which groups are most adversely affected. However, understanding the distributional consequences of privacy regulation requires examining which borrower groups are most affected. This analysis is particularly important from a policy perspective: if privacy regulation disproportionately harms vulnerable populations – those whom FinTech lenders were designed to serve – then policymakers face a meaningful tradeoff between privacy protection and financial inclusion.

We hypothesize that privacy restrictions disproportionately affect borrowers who are more difficult to evaluate using traditional credit metrics, specifically, younger applicants, lower-income individuals, those with limited credit histories (sub-prime credit scores), borrowers from marginalized castes, and women. These groups may have benefited most from alternative data assessment, as CDR-based social collateral metrics could substitute for limited formal credit histories or provide additional information about creditworthiness not captured by traditional scoring models. When this alternative data source is removed, lenders may become more conservative in extending credit to these information-opaque borrowers, leading to differential impacts across demographic groups.

To examine these heterogeneous effects, we estimate a dynamic difference-in-differences specification that focuses exclusively on Android users and interacts borrower characteristics with time indicators around the policy change:

$$y_{i,t} = \sum_{k=-8, k\neq -1}^{k=+7} \beta_k \cdot \text{Characteristic}_i \times \mathbb{I}(t = k) + \theta_{z,p} + \theta_{z,t} + \theta_{p,t} + \theta_w + \varepsilon_{i,t}$$

where Characteristic_i represents various borrower attributes including low-income status, young age, sub-prime credit score, low-caste background, new borrower, and female gender. The coefficients $\{\beta_{-8}, ..., \beta_{-2}\}$ test for pre-existing differential trends across borrower types, while $\{\beta_0, ..., \beta_7\}$ trace out the dynamic treatment effects after policy implementation. Week $k = -1$ (immediately before policy implementation) serves as the omitted reference period.

This specification includes a comprehensive set of fixed effects: ZIP code × loan purpose interactions ($\theta_{z,p}$), ZIP code × week interactions ($\theta_{z,t}$), loan purpose × week interactions ($\theta_{p,t}$), and day-of-the-week effects (θ_w). These fixed effects control for time-invariant geographic and loan-type heterogeneity, location-specific time trends, loan-purpose-specific temporal patterns, and systematic weekly variation in lending activity.

By restricting attention to Android users only, this analysis identifies the differential impact of the privacy policy across borrower subgroups within the treated population. The key assumption is that, absent the privacy policy change, the relative lending outcomes for different borrower

groups would have evolved similarly across locations and loan purposes. This assumption is plausible because the policy change was imposed uniformly across all Android users regardless of borrower characteristics, and our rich set of fixed effects controls for many potential confounders. We assess this assumption by examining pre-treatment trends across borrower groups.

5 Main Results

The Google privacy policy implemented on January 1, 2019, created a sharp change that serves as a natural experiment to study how restrictions on alternative data affect lending markets. By prohibiting Android applications from collecting CDRs while leaving iOS applications unaffected, the policy provides an ideal setting to identify the causal effects of privacy regulation on credit access. Our findings reveal that while privacy regulations achieve their intended goal of protecting consumer data, they generate significant unintended consequences for credit access, particularly among vulnerable populations.

Figure 1 provides an overview of the lending market dynamics over our study period from January 2017 through December 2021. The figure plots monthly application volumes, acceptance rates, and default rates across both Android and iOS users. We observe sharp discontinuities around the January 2019 policy implementation date: loan applications increase abruptly, acceptance rates decline from roughly 80-85% to 60-70%, while default rates remain largely unchanged around 20%.

These aggregate patterns suggest that the privacy policy generated countervailing forces in the lending market. On the demand side, borrowers appear to value enhanced privacy protections, prompting more applications. On the supply side, the loss of CDRs impaired the lender's ability to engage in ex-post enforcement, resulting in a substantial tightening of screening standards. Importantly, default rates did not increase following the change, indicating that the stricter screening did not compromise portfolio quality. This finding suggests that the privacy shock primarily altered selection rather than ex-post repayment behavior.

5.1 Comparing Android and iOS Users: Raw Trends

Our empirical strategy exploits the differential impact of Google's privacy policy on Android versus iOS users. While Android and iOS users may differ along observable dimensions – for example, prior research documents that Android users tend to have lower incomes and higher default rates on average (Berg, Burg, Gombović and Puri, 2020; Bertrand and Kamenica, 2023) – the key identifying assumption is that these groups would have experienced parallel trends in lending outcomes absent the policy change. Since both groups were served by the same lender with identical products, pricing structures, and approval algorithms (aside from the CDR component), we expect their lending outcomes to comove closely prior to the policy intervention.

To provide initial evidence for this assumption and illustrate the policy's impact, Figure 2 plots unconditional mean outcomes separately for Android users (treatment group) and iOS users (control group) in the six months before and after the policy change. Panel (a) shows the natural logarithm of daily applications, Panel (b) displays acceptance rates, and Panel (c) presents default rates. The results reveal several striking patterns.

First, in the pre-policy period (months -6 through -1), Android and iOS users exhibit remarkably similar trends across all three outcomes. Application volumes, acceptance rates, and default rates move in parallel for both groups, with no systematic divergence in the months leading up to the policy change. This close comovement validates our use of iOS users as a control group and suggests that, absent the privacy regulation, the two groups would have continued on similar trajectories.

Second, the policy implementation in month 0 creates an immediate and persistent divergence between the two groups. For applications (Panel 2a), Android users show a clear upward shift relative to iOS users starting in month 0, suggesting that privacy protections increased demand for loans among the treated group. For acceptance rates (Panel 2b), Android users experience a sharp downward shift, indicating that the loss of CDR access forced the lender to tighten credit standards for this population. Notably, default rates (Panel 2c) show no systematic divergence between the groups, remaining stable throughout the entire period for both Android and iOS users.

These raw unconditional means provide compelling visual evidence of the policy's impact. The parallel pre-trends establish the credibility of our identification strategy, while the sharp breaks at the policy implementation date demonstrate that the privacy regulation had immediate and substantial effects on lending market outcomes. In the following section, we formalize these patterns through a parallel trends analysis that controls for potential confounding factors.

5.2 Validation of Research Design: Parallel Trends

While the raw trends in Figure 2 provide strong visual evidence for parallel pre-treatment trends, we now formalize this analysis by controlling for potential confounding factors. To rigorously test for pre-existing differential trends between Android and iOS users, we estimate an event study specification:

$$Y_{i,t} = \alpha + \sum_{k=-7, k\neq -1}^{k=+6} \beta_k \cdot (Android_i \times \mathbb{1}_{t=k}) + \gamma \cdot Android_i + \theta_{z,p,t} + \epsilon_{i,t} \tag{1}$$

where k indexes months relative to the policy implementation (January 2019), with $k = -1$ (December 2018) serving as the omitted reference period. The coefficients $\{\beta_{-7}, \ldots, \beta_{-2}\}$ test for pre-existing differential trends between Android and iOS users, while $\{\beta_0, \ldots, \beta_6\}$ trace out the

dynamic treatment effects after policy implementation. Importantly, this specification includes ZIP code × month × loan purpose fixed effects ($\theta_{z,p,t}$), which control for location-specific temporal trends that vary by loan type, ensuring that our results are not driven by compositional changes or differential trends across geographies or product types.

Figure 3 presents these event study estimates for the six months before and after the policy change. The results strongly support the parallel trends assumption documented in the raw data. For all three main outcomes – applications Figure 3a), acceptance rates (Figure 3b), and default rates (Figure 3c) – the pre-treatment coefficients are statistically indistinguishable from zero and exhibit no systematic trends. This indicates that Android and iOS users experienced parallel lending outcomes in the months leading up to the policy change, even after controlling for geographic, temporal, and loan-type heterogeneity.

After implementation (month 0 onward), we observe clear breaks in the time series. Android applications increase by approximately 30-40% relative to iOS applications, while Android acceptance rates decline by 20-30 percentage points relative to iOS users. Default rates remain statistically unchanged across platforms throughout the entire period. The absence of pre-trends, combined with the immediate breaks at month 0 for applications and acceptance rates, validates our identification strategy and rules out concerns about pre-existing differential trends driving our results.

The combination of raw unconditional means from Figure 2 and controlled event study estimates from Figure 3 provides robust evidence that Android and iOS users followed parallel trajectories before the policy change and diverged sharply afterward. This pattern supports a causal interpretation of our difference-in-differences estimates: the observed changes in lending outcomes stem from the privacy policy's differential impact on Android users rather than pre-existing trends or compositional differences between the groups.

5.3 Main Effects of Privacy Policy

Having established the validity of our parallel trends assumption, we now turn to our primary difference-in-differences estimates that compare Android users (treated) to iOS users (control) around the policy change. These estimates identify the causal effect of privacy regulation on credit market outcomes by exploiting the platform-specific nature of Google's privacy directive.

The DiD estimates confirm the aggregate patterns: privacy protections raised loan applications by roughly 20 percent, reduced approval rates by 18–20 pp, and left default rates unchanged. The results imply that lenders compensated for information loss through stricter screening rather than by accepting higher default risk.

5.3.1 Borrower Demand for Privacy Protection

Table 2 presents our difference-in-differences estimates examining the effect of the privacy policy on loan applications. The key coefficient of interest is Android × Post, which captures the differential change in applications for Android users relative to iOS users after the policy implementation. Across all specifications in columns (1)-(4), we consistently estimate that the privacy policy increased Android applications by approximately 21% relative to iOS applications. This effect is remarkably stable across specifications that progressively add more demanding fixed effects structures, including ZIP code × month-year fixed effects (column 2), ZIP code × loan purpose fixed effects (column 3), and the full interaction of ZIP code × month-year × loan purpose fixed effects (column 4).

This behavioral response provides compelling evidence that borrowers place meaningful value on privacy protections. The willingness of borrowers to apply for loans from a lender that had publicly committed to enhanced privacy protections suggests that data collection practices represent a meaningful factor in financial decision-making. Moreover, it suggests that borrowers are not just concerned about loan terms and accessibility but also privacy considerations.

5.3.2 Lender Supply of Credit

While borrowers responded positively to enhanced privacy protections, the policy imposed substantial costs by eliminating access to rich behavioral and social network data that had been integral in determining credit market outcomes. Table 3 demonstrates the magnitude of the lender's response. The results reveal an immediate and substantial decline in application acceptance rates for Android users relative to iOS users. In the baseline specification (column 1), Android users experienced a 45 percentage point decline in acceptance rates relative to iOS users after the policy change. When we add our most demanding fixed effects structure (column 4), which includes ZIP code × month-year × loan purpose fixed effects, the estimated effect is an 18 percentage point decline, still highly significant.

Column (5) adds comprehensive demographic controls including education, age decile, credit score decile, loan maturity, income decile, female, and low caste fixed effects, and the estimated effect remains at -18 percentage points. The stability of this coefficient across increasingly demanding specifications – particularly when controlling for detailed borrower characteristics – provides strong evidence that the policy change caused an immediate and substantial tightening of credit standards rather than reflecting compositional changes in the applicant pool.

This decline in acceptance rates is substantial, representing more than a 20% relative decline from the pre-policy baseline. This indicates that alternative data sources were fundamental to the lender's business model, and their removal forced a permanent tightening of credit standards. This suggests that CDRs provided information that was difficult or impossible to obtain through conventional credit scoring mechanisms.

5.3.3 Portfolio Quality Unchanged

A critical question is whether the tightening of credit standards successfully maintained portfolio quality or whether the loss of the enforcement technology led to deteriorated loan performance. Table 4 examines the effect of the privacy policy on default rates for approved loans. Across all specifications in columns (1)-(5), the Android × Post coefficient is small in magnitude and statistically insignificant. Even in the most parsimonious specification without fixed effects (column 1), the point estimate is only 0.93 percentage points with a standard error of 0.61 percentage points. When we add the full set of controls in column (5), the estimate is 0.55 percentage points with a standard error of 0.69 percentage points.

The stability of default rates indicates that the tightening of screening standards fully offset the loss of ex-post enforcement capabilities. Acceptance rates declined sharply, but portfolio quality was preserved. The combination of stable default rates with lower loan volumes implies that many rejected borrowers would have repaid successfully, consistent with a shift toward more conservative approval thresholds.

5.3.4 Effects on Other Credit Terms

The dramatic reduction in acceptance rates raises important questions about how the privacy policy affected other dimensions of credit supply. If lenders respond to increased information asymmetries by adjusting loan terms rather than simply rejecting more applicants, we should observe changes in interest rates or loan maturities for approved borrowers. Understanding whether the lender adjusted pricing along these margins is critical for characterizing the full scope of the policy's impact and for understanding how FinTech lenders manage information loss.

Table 5 examines whether the privacy policy affected loan terms for approved borrowers, providing evidence on whether lenders used price adjustments as a margin of response to information loss. Panel A shows the effect on annual interest rates. We find that Android users faced interest rate increases of 0.25-0.35 percentage points relative to iOS users after the policy change, with effects that are statistically significant across most specifications (columns 2-5). While this represents a modest increase relative to baseline interest rates of approximately 20-25% in this market, it indicates that the lender did make some pricing adjustments in response to increased uncertainty about borrower quality.

Panel B presents results for loan maturity measured in months. We find that Android users received loans with maturities approximately 0.10-0.12 months shorter than iOS users after the policy change, effects that are statistically significant across all specifications. Again, while economically modest in absolute terms (representing less than a 2% change relative to baseline maturities), this finding suggests the lender shortened loan durations to limit exposure to borrowers about whom they had less information.

These findings reveal that the lender's response to information loss operated through multi-

ple channels. The primary response was a large reduction in the quantity of credit extended (lower acceptance rates), but conditional on approval, the lender also made modest adjustments to loan terms – charging slightly higher interest rates and offering slightly shorter maturities. This pattern is consistent with the lender facing increased uncertainty about approved borrowers and using price and maturity adjustments to partially compensate for the loss of enforcement capabilities, though these adjustments were insufficient to fully offset the information loss. The fact that quantity rationing dominated pricing adjustments suggests that the information contained in CDRs was particularly valuable for making binary accept/reject decisions, and that once this information was lost, many borrowers became essentially unscreenable rather than simply requiring different loan terms.

This multi-dimensional response to information loss has important implications for understanding how digital lenders manage risk and how they might respond to privacy regulations more broadly. It demonstrates that while algorithmic lending platforms do adjust prices in response to information shocks, their primary margin of adjustment is through more selective lending rather than risk-based pricing.

5.4 High-Frequency Event Study: Daily Patterns Around the Policy Change

While our difference-in-differences analysis comparing Android to iOS users provides the primary identification of the policy's causal effects, we complement this evidence with a high-frequency event study that examines daily patterns in lending outcomes for Android users around January 1, 2019. This approach offers several advantages. First, by analyzing outcomes at a daily frequency rather than monthly, we can more precisely pinpoint when behavioral changes occurred and test whether they align exactly with the policy implementation date. Second, the high-frequency data allow us to assess whether there were any anticipatory effects or gradual adjustments in the days surrounding the policy change. Third, examining stability across different bandwidth choices around the cutoff provides additional evidence on the robustness of our findings.

This analysis focuses exclusively on Android users (who were affected by the policy) and examines whether outcomes changed immediately at the policy implementation date, providing additional evidence on the timing and magnitude of the policy's effects. Figure 4 presents the high-frequency event study results for three key outcomes, with each panel displaying binned daily averages and local linear trends estimated separately before and after the policy change. Panel (a) shows the natural logarithm of daily applications, Panel (b) shows application acceptance rates, and Panel (c) shows default probabilities. For each outcome, we observe smooth trends leading up to January 1, 2019, followed by clear changes in levels after the policy implementation.

The patterns in Figure 4 are consistent with our difference-in-differences findings. Applications exhibit a visible increase after January 1st, acceptance rates show a clear decline, while default rates remain relatively stable. These visual patterns are formalized in Table 6, which presents point estimates across different bandwidth choices (10, 15, 20, 25, and 30 days on either side of the cutoff).

Panel A of Table 6 examines the effect on the natural logarithm of daily applications. The results show consistently positive and statistically significant effects across all bandwidth specifications. Using the narrowest bandwidth of 10 days (column 1), we estimate that applications increased by 4.5% after the policy change. As we expand the bandwidth to include more observations further from the cutoff, the point estimates range from 3.2% to 5.7%, all significant at the 1% level. The stability of these estimates across bandwidths—with coefficients remaining between 0.032 and 0.057—provides confidence that we are capturing a genuine policy effect rather than spurious variation. The relatively high R-squared values (ranging from 0.262 to 0.407) indicate that our specification, which includes ZIP code × loan purpose and day-of-week fixed effects along with local linear trends, explains substantial variation in daily application volumes.

Panel B of Table 6 presents the effects on application acceptance rates, revealing the most economically significant impact of the privacy policy. The estimated decline in acceptance rates is remarkably consistent across bandwidth choices, ranging from 25.3 percentage points with a 10-day bandwidth (column 1) to 27.9 percentage points with a 25-day bandwidth (column 4). All estimates are statistically significant at the 1% level with relatively small standard errors (approximately 0.009-0.021), indicating precise estimation despite focusing on a narrow window around the policy change. The consistency of these estimates is particularly striking—the coefficient varies by less than 5 percentage points across specifications that range from 6,338 observations (10-day bandwidth) to 27,770 observations (30-day bandwidth). This stability suggests the effect is not sensitive to the choice of which days to include in the analysis. Moreover, the high R-squared values (0.746-0.776) indicate that our model fits the data well, with the included fixed effects capturing most of the variation in acceptance decisions. These point estimates align closely with our difference-in-differences estimates of an 18 percentage point decline when controlling for the full set of demographic characteristics, providing convergent evidence from two distinct identification strategies.

Panel C of Table 6 examines default probabilities for approved loans. Across all bandwidth specifications (columns 1-5), the point estimates are small in magnitude and statistically insignificant. The coefficients range from -0.013 (10-day bandwidth) to 0.006 (15-day bandwidth), with standard errors of approximately 0.018-0.041. Notably, even with our largest sample of 17,749 approved loans using the 30-day bandwidth (column 5), we cannot reject the null hypothesis of no effect on defaults. This null result is itself informative: despite the dramatic tightening of credit standards reflected in Panel B, the lender successfully maintained portfolio quality. The combination of sharply lower acceptance rates with stable default rates indicates that the lender became substantially more conservative in approval decisions, rejecting many borrowers who likely would have repaid their loans. The R-squared values in this panel (0.278-0.357) are lower than in Panel B, reflecting the inherently greater noise in predicting individual loan defaults compared to aggregate acceptance decisions, but remain reasonably high given the complexity of default prediction.

The stability of these estimates across different bandwidth choices is shown graphically in

Appendix Figure A.1, which plots the point estimates and confidence intervals for each outcome as we vary the bandwidth from very narrow (2 days) to wide (150 days) windows. The figure demonstrates that our findings are not sensitive to arbitrary bandwidth choices, with estimates remaining remarkably stable across a wide range of specifications. Appendix Table **??** demonstrates that the application results are robust to progressively richer fixed-effect specifications, with the estimated effect remaining positive and significant across all specifications. Appendix Table A.1 shows similar robustness for acceptance rates—the coefficient remains stable at approximately -25 to -27 percentage points even when controlling for detailed borrower demographics including education, age, credit score, gender, and caste. The high-frequency event study thus corroborates our main difference-in-differences results, showing immediate changes in borrower demand and lender supply occurring precisely at the policy implementation date, with no corresponding change in default rates.

5.4.1 Ruling Out Composition Effects

A key identifying assumption for our difference-in-differences strategy is that the privacy policy did not differentially affect the composition of applicants across platforms. If the policy caused Android users to apply at different rates than iOS users, or attracted systematically different types of borrowers to each platform, then our estimated effects on acceptance rates could reflect changes in applicant composition rather than changes in lending standards.

We test this assumption directly by examining whether the policy affected the relative probability that an application came from an Android versus iOS user. Table 7 presents these results across different bandwidth choices (10-30 days on either side of the policy change date). The dependent variable is a binary indicator equal to one if the application was submitted by an Android user. Across all specifications, the Post coefficient is small in magnitude (ranging from 0.007 to 0.013) and statistically insignificant. This provides strong evidence that the policy did not alter the relative volume or platform mix of applications, reinforcing the validity of our identification strategy.

This finding is important because it rules out a key alternative explanation for our results. If the policy had caused a surge of Android applications relative to iOS applications from borrowers with different risk profiles, the decline in Android acceptance rates might simply reflect selection rather than tightened lending standards. The stability of platform composition confirms that our main results reflect genuine changes in how the lender evaluated equivalent borrowers once alternative data became unavailable.

To further assess within-platform composition, we examine whether observable characteristics of Android applicants changed at the policy implementation date. Appendix Table A.2 tests for compositional shifts across seven key dimensions: income, age, credit scores, caste background,

gender, and whether the borrower was a new borrower. Most characteristics remain stable around the cutoff.[2]

Taken together, the daily evidence confirms that the privacy shock did not alter the underlying applicant pool, either across or within platforms. The changes observed in acceptance rates therefore reflect a tightening of lending standards in response to the sudden loss of alternative data, rather than compositional artifacts.

5.5 Heterogeneous Effects: Distributional Consequences of Privacy Regulation

This section examines the distributional consequences of privacy regulation, by assessing how the loss of alternative data affected different groups of borrowers within the treated population of Android users. While the aggregate results show that privacy restrictions tightened credit supply without affecting default risk, understanding who bore the costs of this tightening is important for evaluating the broader welfare implications of privacy policy. If the contraction in credit disproportionately affected borrowers who lack strong formal credit histories or access to traditional banking, then privacy protection may come at the expense of financial inclusion.

To understand why privacy restrictions might have differential effects across borrower types, it is helpful to consider the fundamental challenge that lenders face when serving underbanked populations. Traditional credit scoring relies heavily on established credit histories, stable employment relationships, and formal financial records. For many potential borrowers, particularly younger individuals, those with lower incomes, or people from marginalized social backgrounds, these traditional indicators may be sparse or entirely absent. Before the 2019 policy, CDRs helped fill this informational gap by providing behavioral and social network data that captured aspects of reliability and repayment capacity otherwise invisible to conventional scoring. The removal of CDR access therefore plausibly tightened screening constraints most sharply for borrowers who relied on such data to demonstrate creditworthiness. This reasoning generates clear predictions: the privacy policy should have larger negative effects on loan approval probabilities for low-income borrowers, younger applicants, first-time borrowers, and members of historically marginalized castes, while borrowers with extensive credit histories or established relationships with the lender should experience smaller impacts.

To quantify these heterogeneous effects, we exploit daily high-frequency variation around the January 1, 2019 policy cutoff, focusing exclusively on Android users. We estimate a dynamic difference-in-differences specification that interacts borrower characteristics with time indicators around the cutoff, as described in Section 4.3. This design identifies differential treatment effects across borrower subgroups within the treated population while flexibly controlling for location, loan purpose, and temporal variation through ZIP code × loan purpose, ZIP code × week, loan purpose × week, and day-of-week fixed effects. The identifying assumption is that, absent the

[2]We find only modest movements: average applicant age decreased by roughly 1.7 percent, the share of low-caste applicants rose by 3.1 percentage points, and the share of new borrowers declined by 4.9 percentage points.

policy change, relative outcomes across borrower subgroups would have evolved similarly. Because the policy was applied uniformly across all Android users, this assumption is plausible and supported by pre-trend tests.

Figure 5 presents the dynamic treatment effects for application acceptance rates across borrower subgroups. The results reveal consistent and economically meaningful heterogeneity in line with theoretical expectations. We first examine the differential effects by income level in Panel 5a. Low-income borrowers, defined as those with monthly incomes below the sample median, experienced substantially larger reductions in acceptance rates following the policy change. Estimates indicate declines of roughly 8–10 percentage points relative to higher-income applicants, with the effect emerging immediately after the policy and persisting throughout the observation window. This pattern is consistent with the idea that lower-income borrowers are more likely to have thin credit files and thus rely disproportionately on alternative data for credit evaluation. In addition, smaller loan sizes for this population may make manual underwriting economically infeasible, further amplifying their dependence on algorithmic assessment.

Age similarly predicts differential effects in Panel 5b. Younger borrowers, defined as those below the sample median age, faced declines in acceptance rates of approximately 7–10 percentage points relative to older applicants. This finding is notable because younger borrowers typically have shorter credit histories and fewer established financial relationships, making them more dependent on nontraditional data sources. The disproportionate effect on this group suggests that privacy restrictions may limit initial access to formal credit markets, with potential long-term consequences for financial inclusion.

The effects are even more pronounced when comparing new and existing borrowers in Panel 5c. First-time applicants experienced reductions in acceptance rates of 12–15 percentage points relative to borrowers with established lending relationships. Alternative data are particularly valuable for new borrowers, as lenders lack internal repayment histories or prior behavioral data. Consequently, privacy restrictions pose a significant barrier to borrower acquisition and may increase concentration in the digital lending market by favoring established platforms with large existing borrower bases.

We also find substantial differential effects across social identity groups in Panel 5d. Borrowers from historically marginalized castes—including scheduled castes, scheduled tribes, and other backward classes—experienced reductions in acceptance rates of roughly 8–10 percentage points relative to general caste borrowers. These findings suggest that privacy regulations may inadvertently exacerbate existing patterns of economic exclusion, as the removal of alternative data forces lenders to rely more heavily on conventional credit metrics that reflect historical inequities.

Gender, by contrast, appears to play a more limited role, as shown in Panel 5e. Female borrowers experienced slightly larger reductions in acceptance rates than male borrowers, though these differences are generally not statistically significant. Similarly, differences by credit score are modest, as shown in Panel 5f. Borrowers with subprime credit scores (below 680) show larger

point estimates of declines in acceptance rates, but these effects are not statistically significant, potentially reflecting the limited sample size of subprime borrowers or the fact that even borrowers with established credit histories benefited from the information provided by alternative data.

Taken together, the evidence indicates that privacy restrictions had heterogeneous effects in line with theoretical predictions: the groups most reliant on alternative data: low-income, younger, new, and socially marginalized borrowers experienced the largest reductions in access to credit, while other dimensions such as gender or subprime status were less strongly affected.

Table 8 presents the regression estimates that complement the graphical evidence from Figure 5. Across specifications, the groups most dependent on alternative data – low-income, young, new, and marginalized borrowers – experience the largest and most statistically significant declines in acceptance rates.

New borrowers experienced a 12–14 percentage point larger decline in acceptance rates, while low-income borrowers faced an 7-8 percentage point larger reduction. Similarly, younger borrowers saw a 7–10 percentage point decline, and low-caste borrowers experienced a 6-8 percentage point larger effect. All of these estimates are statistically significant and economically meaningful. Appendix Table A.4 further demonstrates the robustness of these findings to more stringent identification strategies, including the addition of ZIP code × week and loan purpose × week fixed effects, confirming that these differential effects are not driven by time-varying factors correlated with borrower characteristics. Taken together, the evidence shows that privacy regulation exacerbated informational exclusion within the digital lending market. The loss of behavioral data most strongly affected borrowers whose creditworthiness could not be easily assessed using traditional information sources. These patterns are inconsistent with explanations based on lender distress or aggregate credit tightening – which would have affected all borrowers similarly – and instead point to an information-driven mechanism.

In short, the CDR ban disproportionately constrained credit access for borrowers on the financial margins, highlighting a tradeoff between data privacy and financial inclusion.

6 Implications of the Tightening of Credit Standards

This section explores the underlying mechanism through which the loss of alternative data sources affects a lender's business model and risk management. The withdrawal of CDRs fundamentally changed how the lender managed risk. Before the privacy policy, the lender's model relied heavily on dynamic enforcement: CDR data supported post-disbursement tracking and early intervention when repayment risks emerged (Agarwal, Alok, Ghosh and Gupta, 2020). When access to CDRs was removed on January 1, 2019, this enforcement capability disappeared overnight, forcing the lender to pivot toward a model based almost entirely on ex-ante screening at the application stage.

One way of testing this hypothesis is through a comparison of default rates. If the lender successfully adapted to the loss of monitoring capabilities by tightening screening standards,

we should observe no change in portfolio risk despite the dramatic reduction in approval rates. Conversely, if the screening adjustments were insufficient to compensate for lost enforcement capabilities, default rates should increase. Figure 4c plots the regression discontinuity estimates of loan default, showing no discontinuity at the policy boundary. Defaults remained close to 20% before and after January 1, 2019.

Our earlier findings demonstrate that this shift was successful in maintaining portfolio quality. Table 4 shows that default rates remained stable across DiD specifications – the Android × Post coefficient is small and statistically insignificant, indicating that the lender's tightened screening standards fully compensated for the loss of enforcement capability. The high-frequency event study corroborates this finding, with Panel C in Table 6 showing stable default rates across different bandwidths around the policy date. Appendix Table A.3 further demonstrates that this null result is robust to the progressive inclusion of individual-level controls, even when controlling for detailed borrower demographics, education, age deciles, credit score deciles, loan maturity, and income deciles, the policy effect on default rates remains statistically insignificant and economically small.

The lender's tightening of credit standards was thus effective in maintaining risk levels, but this risk management came at a substantial cost: the exclusion of many borrowers who would have successfully repaid their loans. The combination of sharply lower acceptance rates with unchanged default rates suggests the lender became significantly more conservative in its lending decisions, rejecting many creditworthy applicants to maintain portfolio quality. This pattern implies that many rejected borrowers would have repaid successfully, consistent with a shift toward more conservative approval thresholds.

This shift from enforcement to screening has important implications for financial inclusion. The enforcement-based model enabled the lender to extend credit to borrowers with limited traditional credit profiles, as ongoing surveillance helped manage risk. The screening-based model, by contrast, requires borrowers to demonstrate creditworthiness through conventional metrics available at the time of application. For many underbanked borrowers—particularly those with limited credit histories, irregular income streams, or other characteristics that make them difficult to evaluate through traditional metrics—this shift effectively eliminated access to formal credit.

6.1 Long-Term Financial Exclusion

Having established that the privacy policy caused immediate exclusion from this lender, we now examine whether affected borrowers obtained credit elsewhere or faced persistent exclusion from formal financial markets. Our analysis of long-term outcomes is enabled by the integration of our primary dataset with comprehensive credit bureau records that capture formal lending activity across all major financial institutions in India. This external data source allows us to observe whether borrowers who applied around the policy change date subsequently accessed credit from

other lenders, providing a complete picture of their financial market participation over multiple years.

This analysis addresses an important question: did privacy restrictions simply redistribute borrowing activity across lenders, or did they generate lasting exclusion from formal credit markets? If borrowers denied credit by our focal FinTech readily obtain loans from other institutions, the welfare implications would be relatively modest. However, if obtaining a FinTech loan facilitates access to credit from conventional banks, either by generating new information (Balyuk, 2023; Dubey, 2024) or helping to build collateral (Beaumont, Tang and Vansteenberghe, 2022), then borrowers rejected due to privacy regulation might face systematic exclusion. Our setting provides a unique opportunity to quantify this "FinTech ladder effect," measuring how initial FinTech loan acceptance influences borrowers' future credit prospects across the entire financial system.

6.2 Reduced-Form Evidence

Table 9 presents reduced-form evidence on how the policy change affected borrowers' subsequent access to formal credit markets using our difference-in-differences design. We track borrowers for two-, three-, and four-year horizons following the January 1, 2019 policy implementation date. The dependent variables in columns (1)-(4), (5)-(8), and (9)-(12) are binary indicators equal to one if the borrower accessed any form of formal credit, including loans from banks, non-bank financial companies, and other FinTech platforms—within 2, 3, or 4 years respectively.

The key coefficient of interest is Android × Post, which captures whether Android users who applied after the policy change were differentially less likely to access future credit compared to iOS users. The results reveal persistent negative effects on credit access that extend far beyond the immediate policy effect. Focusing on the most demanding specifications with the full set of controls (columns 4, 8, and 12), we find that Android users who applied after the policy change were 2.5 percentage points less likely to access formal credit within 2 years, 2.7 percentage points less likely within 3 years, and 2.4 percentage points less likely within 4 years. All effects are statistically significant at conventional levels.

These estimates are economically meaningful. Given baseline credit access rates of 36% over two years in our sample, a 2.5 percentage point reduction represents approximately a 7% relative decline in formal financial market participation, consistent with prior evidence that early credit rejections can scar borrowers' long-run outcomes (Dobbie and Skiba, 2013; Dobbie, Goldsmith-Pinkham, Mahoney and Song, 2020; Liberman, Neilson, Opazo and Zimmerman, 2018). The persistence of these effects, remaining significant even four years after the policy change, suggests that being denied credit due to privacy restrictions has lasting consequences for borrowers' financial trajectories.

Appendix Table A.5 demonstrates the robustness of these findings using a high-frequency event study design with daily data around the policy cutoff. Focusing on Android users only, we find that borrowers who applied after January 1, 2019 were 4.6-4.8 percentage points less likely

to access credit within 2 years, 5.1-5.2 percentage points less likely within 3 years, and 4.6-4.7 percentage points less likely within 4 years (columns 3, 6, and 9 with fixed effects). These point estimates are larger than the DiD estimates, consistent with the idea that focusing on the immediate policy window captures the marginal borrowers most affected by the loss of alternative data.

6.3 Instrumental Variables Evidence

While the reduced-form results provide strong evidence of persistent exclusion, we further quantify the magnitude of the FinTech ladder effect by estimating the relationship between initial loan approval and long-term credit access. We employ an instrumental variables approach that exploits the privacy policy change as an exogenous shock to loan approval decisions, allowing us to estimate the elasticity of future credit access with respect to initial FinTech loan approval.

We employ an instrumental variables approach that exploits the privacy policy change as an exogenous shock to loan approval decisions. The policy change is a powerful instrument because it generates large, exogenous variation in approval rates while being unrelated to individual borrower characteristics that might independently affect future credit access. Our two-stage specification is:

$$\textbf{First Stage:} \quad \text{App Accepted}_{i,t} = \pi_0 + \pi_1 \cdot (Android_i \times Post_t) + \pi_2 \cdot Android_i + \pi_3 \cdot Post_t + X'_{i,t}\pi_4 + \theta + \nu_{i,t}$$

$$\textbf{Second Stage:} \quad \text{Credit Access}_{i,t+s} = \beta_0 + \beta_1 \cdot \widehat{\text{App Accepted}}_{i,t} + X'_{i,t}\beta_2 + \theta + \epsilon_{i,t}$$

where $\text{Credit Access}_{i,t+s}$ is a binary indicator equal to one if borrower i accessed formal credit within s years following January 1, 2019 (we examine horizons of 2, 3, and 4 years). $\text{App Accepted}_{i,t}$ indicates whether the loan application was approved by our focal FinTech. The instruments are the Android × Post interaction, the Android indicator, and the Post indicator. The specificiation includes various combinations of fixed effects (θ) and individual-level controls ($X_{i,t}$).

Conceptually, the IV estimates capture the local average treatment effect (LATE) of loan approval for borrowers whose applications would have been approved before the policy change but were rejected afterward due to privacy restrictions. This parameter is particularly policy-relevant because it identifies the credit access consequences for the marginal borrowers most directly affected by privacy regulation.

The validity of our instrumental variables approach relies on two key assumptions. First, the relevance condition requires that the privacy policy significantly affects loan approval decisions. As demonstrated in Table 3, the first-stage relationship is strong and economically meaningful—the privacy policy reduced Android approval rates by approximately 18-45 percentage points depending on the specification. Table 10 reports the Kleibergen-Paap (KP) LM statistics and F-statistics, which substantially exceed conventional thresholds for weak instruments across all specifications. The KP F-statistics range from approximately 1,300 to 15,000, providing strong evidence against weak identification.

Second, the exclusion restriction requires that the privacy policy affects future credit access only through its effect on current loan approval decisions, not through direct channels. This assumption is plausible because: (1) the policy change was specific to our focal FinTech's data collection practices and did not directly affect other lenders' operations, (2) the credit bureau data captures borrower interactions with the broader lending ecosystem beyond our focal institution, and (3) we control for observable borrower characteristics that might independently influence future credit access.

Table 10 presents the second-stage results. The estimates reveal economically large effects of loan approval on future credit access. Focusing on specifications with the full set of fixed effects and controls (columns 4, 8, and 12), we find that borrowers whose applications were approved were 21.6 percentage points more likely to access formal credit within 2 years, 20.0 percentage points more likely within 3 years, and 17.5 percentage points more likely within 4 years. All effects are highly statistically significant.

These instrumental variables estimates are substantially larger in magnitude than the reduced-form effects, which is consistent with the LATE interpretation. The IV estimates capture the effect of loan approval for marginal borrowers most affected by the policy change, precisely those for whom alternative data sources provided the most valuable information. These are borrowers with limited traditional credit profiles who were excluded when the lender shifted to conventional screening methods. For these marginal borrowers, initial loan approval appears critical for accessing the formal financial system over subsequent years.

The high-frequency event study corroborates these findings. Appendix Table A.6 presents instrumental variables estimates using daily data around the policy cutoff, focusing exclusively on Android users and instrumenting application acceptance with the Post indicator. The results show that loan approval increases credit access by 11.3-11.5 percentage points within 2 years, 12.2-12.3 percentage points within 3 years, and 11.0-11.5 percentage points within 4 years (columns 1-3, 4-6, and 7-9). These estimates are somewhat smaller than the DiD-based IV estimates but remain economically large and highly significant, reinforcing the conclusion that initial FinTech loan approval has substantial positive effects on long-term financial market participation.

6.4 Mechanisms and Interpretation

Overall, our findings suggest that privacy regulation can cause persistent financial exclusion, consistent with evidence that credit rejections can both transmit negative signals to credit bureaus (Hertzberg, Liberti and Paravisini, 2011) and discourage borrower engagement with formal markets (Banerjee, Duflo, Goldberg, Karlan, Osei, Parienté, Shapiro, Thuysbaert and Udry, 2015). Our results contrast with Chava, Ganduri, Paradkar and Zhang (2021), who document that marketplace lending leads to worse long-term credit outcomes due to information frictions and debt accumulation. The key difference is that their sample consists primarily of borrowers with existing credit histories seeking debt consolidation, whereas our lender serves predominantly new-to-credit bor-

rowers. For such borrowers, initial FinTech credit access appears to generate positive information spillovers that facilitate future access to traditional bank credit (Beaumont, Tang and Vansteenberghe, 2022; Balyuk, 2023; Dubey, 2024), rather than exacerbating information frictions. Several complementary mechanisms likely contribute to this pattern.

First, other lenders in the formal financial sector may face similar information constraints when alternative data sources become unavailable or when privacy regulations limit data collection more broadly. This interpretation gains credibility from subsequent regulatory developments. In May 2025, the Reserve Bank of India (RBI) issued Reserve Bank of India (Digital Lending) Directions requiring digital lending apps to desist from accessing mobile phone resources such as file and media, contact lists, call logs, and telephony functions. This broader shift in regulatory policy suggests that our findings may reflect the early stages of a sector-wide transition away from alternative data sources, creating systematic tightening of credit standards across multiple lenders rather than isolated effects at our focal FinTech.

Second, initial loan rejections may damage borrowers' formal credit records in ways that affect future applications. In many credit bureau systems, loan applications and rejections are recorded and may be visible to future lenders. A pattern of rejections could signal heightened risk to subsequent lenders, creating a self-reinforcing cycle of exclusion where initial privacy-induced denials lead to reduced creditworthiness in the eyes of other potential lenders.

Third, borrowers may respond to initial rejections by reducing their engagement with formal financial markets. The experience of being denied credit, particularly for borrowers who may already feel marginalized by traditional financial institutions, could lead to discouragement and withdrawal from formal lending markets. This behavioral response could generate persistent exclusion even if other lenders would be willing to extend credit, as affected borrowers may simply stop applying for formal loans after experiencing rejection.

The combination of these mechanisms suggests that the welfare costs of privacy regulation extend well beyond the immediate reduction in credit supply from a single lender. By creating barriers to initial credit access, particularly for underbanked borrowers who rely on alternative data, privacy restrictions may generate lasting exclusion from the formal financial system with potentially significant consequences for economic mobility and financial well-being.

7 Quantifying the Economic Costs of Screening

The sharp reduction in acceptance rates documented in the previous section represents a substantial tightening of credit standards that excluded many borrowers who would have successfully repaid their loans. Because we find no evidence of higher default rates among approved borrowers, the decline in lending cannot be attributed to weakened repayment discipline or reduced enforcement capacity. Instead, privacy restrictions primarily impaired the lender's ability to screen marginal applicants ex-ante. To quantify the magnitude of this tightened screening, we develop a

counterfactual framework that estimates the foregone profits from rejecting borrowers who would have been profitably served absent the policy change. This analysis provides important context for evaluating the welfare implications of privacy regulation in credit markets.

7.1 Methodology

Our approach centers on identifying "marginal borrowers" that were rejected after the policy change but would have been approved under pre-policy standards. We estimate pre-policy approval models using data from the 60 days immediately before January 1, 2019, then apply these models to predict approval probabilities for post-policy applicants. Borrowers with predicted approval probabilities above 0.5 who were actually rejected represent the marginal population whose exclusion we can directly attribute to the loss of enforcement ability induced by the privacy policy change.

Using the pre-policy data, we estimate three key relationships that governed the lender's business model. First, a logistic regression predicts loan approval as a function of credit scores, income, age, loan amount, gender, and caste. Second, linear regressions model how interest rates and loan maturities varied with borrower characteristics, capturing the lender's risk-based pricing strategy. These models enable us to construct counterfactual loan terms for rejected borrowers, approximating what their loans would have looked like under the pre-policy regime.

The profit calculation assumes a bullet loan structure where borrowers pay interest and principal at maturity. For each marginal borrower, expected profit equals the present value of interest income plus principal repayment, minus the loan disbursement and expected default losses.

$$\text{Expected Profit}_i = \frac{r_i \cdot A_i \cdot (T_i/12)}{(1+\delta)^{T_i}} - \frac{\pi \cdot \text{LGD} \cdot A_i}{(1+\delta)^{T_i}} + \frac{A_i}{(1+\delta)^{T_i}} - A_i \tag{2}$$

where r_i denotes the predicted annual interest rate for borrower i, A_i represents the loan amount, T_i is the predicted loan maturity in months, δ is the monthly discount rate, π is the baseline default rate, and LGD is the loss-given-default rate. We apply the pre-policy average default rate of 19%, pre-policy loss-given-default rate of 9%, and a monthly discount rate of 1% to all counterfactual calculations. While these assumptions are necessarily stylized, our sensitivity analysis demonstrates that conclusions remain robust across reasonable parameter ranges.

Our baseline analysis identifies marginal borrowers within the 60-day window around the policy implementation—borrowers who were rejected post-policy despite having predicted approval probabilities above 50% under pre-policy models. Table 11 summarizes the expected profit margins across different borrower characteristics.

Panel A shows pronounced heterogeneity in profitability. Margins range from 1.50% for small loan amounts to 3.58% for female borrowers, with interquartile ranges spanning between 2–5 percentage points within groups. Panel B reports group-level averages weighted by loan

size, which better capture the total economic effect of tightened screening. These dollar-weighted averages are systematically higher than simple means, consistent with larger loans tending to be more profitable. Female borrowers exhibit the highest average margin (5.54%), followed closely by low credit score (5.44%) and low caste borrowers (5.24%). Notably, most disadvantaged groups, precisely those targeted by financial inclusion policies, exhibit margins above 5%, suggesting that the loss of alternative data curtailed highly profitable lending opportunities to underserved populations.

Aggregating across groups, we estimate that tightened screening reduced profits by almost 30% relative to the lender's realized profits during our 120-day sample window. In rupee terms, this corresponds to about ₹31 million ($1.6 million in PPP terms). Scaling to annual figures implies foregone profits on the order of nearly ₹100 million ($5.3 million in PPP terms), highlighting the substantial costs imposed by privacy restrictions on a single lender. The magnitude of these losses—nearly one-third of baseline earnings—demonstrates that CDRs provided screening information that could not be readily substituted by conventional credit signals.

7.2 Heterogeneous Effects

This section examines how profit margins vary across borrower characteristics, addressing a complementary but distinct question from our earlier heterogeneous effects analysis. In Section 5.5, we documented which borrower groups experienced the largest reductions in approval rates following the policy change, identifying populations most harmed in terms of credit access. Here, we analyze which excluded borrowers would have generated the highest profit margins had they been approved—identifying populations that represented the most valuable foregone business opportunities.

These two perspectives capture different dimensions of the economic effect. A group might experience large approval rate declines (indicating heavy reliance on alternative data for access) while excluded members of that group generate relatively modest profit margins. Conversely, another group might show smaller approval effects but higher profit margins among those excluded. Understanding both patterns is essential for evaluating the full welfare consequences of stringent screening.

Figure 6 illustrates how profit margins vary across borrower characteristics. The results reveal systematic patterns in foregone profitability that complement our earlier findings about differential access effects. Panel 6a shows that profit margins increase steadily with income, with higher-earning borrowers exhibiting margins around 4% compared to under 2% for the lowest earners. This pattern suggests that alternative data sources were valuable across the income distribution, but particularly so for higher-income borrowers who may have represented larger, more profitable loan opportunities.

Margins also increase with age, as shown in Panel 6b, from under 3% among the youngest applicants to more than 3.5% for older borrowers, consistent with the view that alternative data

complemented traditional signals for those with more established credit histories. Loan size exhibits a strong gradient in Panel 6c: margins are negligible for very small loans but climb to nearly 7% for the largest. This pattern reflects both economies of scale in processing larger loans and the higher absolute returns available on substantial loan amounts.

By contrast, margins decline with credit score in Panel 6d, consistent with the idea that alternative data was most valuable for borrowers with weaker formal credit profiles. Gender differences are also notable, as female borrowers exhibit systematically higher profitability than their male counterparts in Panel 6e. For caste, the pattern is more muted. We find that non-low caste borrowers show slightly higher margins than low caste borrowers in Panel 6f, suggesting that profitability gains from alternative data were not disproportionately concentrated among marginalized caste groups.

Overall, these patterns demonstrate that tightened screening disproportionately curtailed lending to some of the most profitable and financially excluded groups. This may have undermined exactly the populations that digital lending innovations were designed to serve.

7.3 Sensitivity to Key Assumptions

Our analysis makes several critical assumptions about borrower risk characteristics, discount rates, and the definition of marginal borrowers that could potentially influence our estimates of foregone profits. To assess the robustness of our findings, we conduct a comprehensive sensitivity analysis examining how our results respond to alternative parameter values across four key dimensions: the probability threshold used to identify marginal borrowers, assumptions about their default rates, the discount rate applied to future cash flows, and the loss given default rate.

Figure 7 demonstrates that our conclusions are robust across reasonable parameter ranges. First, our results are relatively insensitive to the probability threshold used to define marginal borrowers (Figure 7a), with foregone profits declining gradually from about 31% at a 0.3 threshold to 15% at a 0.8 threshold while remaining economically meaningful throughout this range. The stability around our baseline 0.5 threshold provides confidence that our identification of marginal borrowers is not driven by arbitrary cutoff choices. Second, our findings remain positive unless we assume marginal borrowers have default rates exceeding 4.5 times the baseline rate (Figure 7b) which would imply default rates above 85%, an extreme assumption well beyond reasonable bounds for consumer lending. Third, results are robust to discount rate assumptions, remaining positive even at monthly discount rates of 2%, equivalent to 27% annually (Figure 7c, though substantially reduced at these high rates. Fourth, our conclusions hold unless loss given default rates exceed 4.5 times the baseline (Figure 7d, implying losses above 40% of loan amounts.

The sensitivity analysis reveals that our baseline estimate of 29% foregone profits represents a conservative assessment of the economic effect. Under more aggressive but still reasonable assumptions, the estimated costs of tightened screening would be even larger, reinforcing our

conclusion that alternative data sources were fundamental rather than merely supplementary to the lender's business model.

7.4 Robustness: Alternative Loan Structure

To assess the sensitivity of our findings, we replicate the analysis using an EMI (equated monthly installment) loan structure as a robustness exercise. Under EMI, borrowers make fixed monthly payments of principal and interest, which resembles many consumer lending products in India's FinTech sector. The EMI profit equation is the following.

$$\text{Expected Profit}_i = \frac{\text{EMI}_i \cdot [1 - (1+\delta)^{-T_i}]}{\delta} - A_i - \frac{\pi \cdot \text{LGD} \cdot A_i}{(1+\delta)^{T_i}} \tag{3}$$

where the monthly interest rate is calculated as $r_{m,i} = (1 + r_i)^{\frac{1}{12}}$-1 and the equated monthly installment is $EMI_i = \frac{r_{m,i} \cdot A_i}{1\text{-}(1+r_{m,i})^{\text{-}T_i}}$. All other variables are defined as in equation 2.

Neither framework is fully realistic. The bullet loan specification abstracts from amortization and concentrates all cash flows at maturity, which simplifies the analysis and makes the economic tradeoff between interest income and default losses fully transparent. The EMI approach, while closer to standard repayment patterns, introduces additional assumptions about the timing of defaults and remaining principal balances that cannot be validated without detailed loan performance data. As a result, the EMI model embeds a different set of stylized assumptions.

The EMI results yield smaller estimates of foregone profits of almost 12% of realized profits compared to almost 30% under the bullet specification, but the cross-sectional heterogeneity patterns remain remarkably consistent. Female borrowers continue to exhibit the highest profit margins at 2.17%, compared to 5.54% in the bullet model, while low amount borrowers show the lowest margins in both specifications, shown in Appendix Table A.7. The relative ranking across all demographic groups is unchanged, presented in Appendix Figure A.2, suggesting that our conclusions about which populations were most harmed by tightened screening are robust to loan structure assumptions.

The sensitivity analysis under the EMI framework in Appendix Figure A.3 reveals that foregone profits turn negative at more conservative stress scenarios, such as default or LGD rates exceeding 2.5 times baseline compared to 4.5 times in the bullet model, but results remain economically meaningful across reasonable parameter ranges. Even under the more conservative EMI assumptions, marginal borrowers would need to exhibit default rates above 47.5% ($2.5 \times 19\%$) or loss rates above 22.5% ($2.5 \times 9\%$) before becoming unprofitable, thresholds that remain outside typical bounds for consumer lending.

Overall, our findings confirm that our core conclusions about the substantial economic costs of tightened screening, and their disproportionate effect on vulnerable populations, are robust to alternative modeling frameworks.

8 Conclusion

This paper examines the unintended consequences of privacy regulation for financial inclusion by analyzing how Google's 2019 restriction on call detail record collection affected a major Indian FinTech lender. Our findings reveal a fundamental tension in digital finance policy: while privacy protections enhance consumer welfare and increase loan demand by 21%, they can simultaneously lead to substantial credit supply contraction. The policy reduced acceptance rates by 18 percentage points for affected borrowers while maintaining stable default rates, and led to modest increases in interest rates and reductions in loan maturities for approved applicants. These patterns indicate that privacy restrictions primarily affect ex-ante screening rather than ex-post enforcement mechanisms, forcing lenders to fundamentally restructure their business models from ongoing disciplining to stricter upfront evaluation. This distinction has important implications for understanding how alternative data supports lending decisions and for designing optimal privacy regulations.

The distributional consequences of privacy regulation have significant implications for financial inclusion. Our analysis demonstrates that restrictions disproportionately harm vulnerable populations who rely most heavily on alternative data for credit evaluation. Low-income borrowers, younger applicants, first-time borrowers, and individuals from marginalized castes experienced 7-15 percentage point larger declines in credit access relative to their more advantaged counterparts. These effects persist over multiple years, with initially rejected borrowers remaining 4-5 percentage points less likely to access formal credit even four years later, and our instrumental variables estimates suggest that initial loan approval increases long-term credit access by 10-12 percentage points. The economic magnitude is substantial: we estimate that privacy restrictions eliminated nearly 30% of the lender's potential profits, equivalent to approximately 100 million annually, by excluding creditworthy borrowers who would have successfully repaid their loans.

Our findings highlight the need for a more nuanced approach to privacy regulations. While limiting access to sensitive consumer data is essential for protecting privacy, blanket restrictions on alternative data usage can unintentionally widen financial disparities by disproportionately excluding borrowers without extensive credit histories. The fact that the policy did not dramatically alter the composition of applicants but dramatically changed approval decisions demonstrates that these effects reflect changes in lending standards rather than selection effects.

Thus, potential approaches include developing responsible guidelines for alternative credit scoring that specify appropriate and inappropriate uses of behavioral data, implementing consent-based data sharing mechanisms that empower consumers to control their information while maintaining access to credit, or encouraging privacy-preserving machine learning techniques that allow lenders to assess and monitor risk without compromising consumer privacy. Our findings suggest that regulations should not only safeguard data but also consider their impact on credit access for underserved populations. As digital finance continues expanding globally, particularly in

emerging markets where traditional credit infrastructure remains limited, understanding these tradeoffs becomes increasingly critical for promoting both consumer protection and inclusive economic growth. The challenge for policymakers is to design regulatory frameworks that protect consumer privacy without inadvertently excluding the very populations that FinTech innovations were designed to serve.

References

Abis, Simona, Huan Tang, and Bo Bian. 2025. "Breaking the Data Chain: The Ripple Effect of Data Sharing Restrictions on Financial Markets." *Available at SSRN 5334566*.

Acquisti, Alessandro, Curtis Taylor, and Liad Wagman. 2016. "The economics of privacy." *Journal of economic Literature*, 54(2): 442–492.

Agarwal, Sumit, Shashwat Alok, Pulak Ghosh, and Sudip Gupta. 2020. "Financial inclusion and alternate credit scoring for the millennials: role of big data and machine learning in fintech." *Business School, National University of Singapore Working Paper, SSRN*, 3507827.

Allen, Franklin, Xian Gu, and Julapa Jagtiani. 2021. "A survey of fintech research and policy discussion." *Review of Corporate Finance*, 1: 259–339.

Aridor, Guy, Yeon-Koo Che, and Tobias Salz. 2020. *The economic consequences of data privacy regulation: Empirical evidence from GDPR*. National Bureau of Economic Research Cambridge, MA, USA.

Bae, Donghwa, Raveesh Mayya, and Tingting Nian. 2023. "Privacy regulation and its unintended consequence on consumption behaviors: evidence from CCPA." *CIS 2023 proceedings*.

Balyuk, Tetyana. 2023. "FinTech lending and bank credit access for consumers." *Management Science*, 69(1): 555–575.

Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. 2015. "A multifaceted program causes lasting progress for the very poor: Evidence from six countries." *Science*, 348(6236): 1260799.

Beaumont, Paul, Huan Tang, and Eric Vansteenberghe. 2022. "Collateral effects: the role of FinTech in small business lending."

Bergemann, Dirk, and Stephen Morris. 2019. "Information design: A unified perspective." *Journal of Economic Literature*, 57(1): 44–95.

Berg, Tobias, Andreas Fuster, and Manju Puri. 2022. "Fintech lending." *Annual Review of Financial Economics*, 14(1): 187–207.

Berg, Tobias, Valentin Burg, Ana Gombović, and Manju Puri. 2020. "On the rise of fintechs: Credit scoring using digital footprints." *The Review of Financial Studies*, 33(7): 2845–2897.

Bertrand, Marianne, and Emir Kamenica. 2023. "Coming apart? Cultural distances in the United States over time." *American Economic Journal: Applied Economics*, 15(4): 100–141.

Bessen, James, Stephen Michael Impink, Lydia Reichensperger, and Robert Seamans. 2020. "GDPR and the importance of data to AI startups."

Bian, Bo, Michaela Pagel, Huan Tang, and Devesh Raval. 2023. "Consumer surveillance and financial fraud." National Bureau of Economic Research.

Bian, Bo, Xinchen Ma, and Huan Tang. 2021. "The supply and demand for data privacy: Evidence from mobile apps." *Available at SSRN 3987541*.

Blattner, Laura, and Scott Nelson. 2021. "How costly is noise? Data and disparities in consumer credit." *arXiv preprint arXiv:2105.07554*.

Bryan, Gharad, Dean Karlan, and Jonathan Zinman. 2015. "Referrals: Peer screening and enforcement in a consumer credit field experiment." *American Economic Journal: Microeconomics*, 7(3): 174–204.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik. 2014. "Robust data-driven inference in the regression-discontinuity design." *The Stata Journal*, 14(4): 909–946.

Cattaneo, Matias D, Nicolás Idrobo, and Rocío Titiunik. 2024. *A practical introduction to regression discontinuity designs: Extensions.* Cambridge University Press.

Chava, Sudheer, Rohan Ganduri, Nikhil Paradkar, and Yafei Zhang. 2021. "Impact of marketplace lending on consumers' future borrowing capacities and borrowing outcomes." *Journal of Financial Economics*, 142(3): 1186–1208.

Cramer, Kim Fe, Pulak Ghosh, Nirupama Kulkarni, and Nishant Vats. 2024. "Shadow Banks on the Rise: Evidence Across Market Segments." *Olin Business School Center for Finance & Accounting Research Paper (2024/18)*.

Demirer, Mert, Diego J Jiménez Hernández, Dean Li, and Sida Peng. 2024. "Data, privacy laws and firm production: Evidence from the GDPR." National Bureau of Economic Research.

Dobbie, Will, and Paige Marta Skiba. 2013. "Information asymmetries in consumer credit markets: Evidence from payday lending." *American Economic Journal: Applied Economics*, 5(4): 256–282.

Dobbie, Will, Paul Goldsmith-Pinkham, Neale Mahoney, and Jae Song. 2020. "Bad credit, no problem? Credit and labor market consequences of bad credit reports." *The Journal of Finance*, 75(5): 2377–2419.

Doerr, Sebastian, Leonardo Gambacorta, Luigi Guiso, and Marina Sanchez del Villar. 2023. "Privacy regulation and fintech lending." *Available at SSRN 4353798*.

Dubey, Tamanna Singh. 2024. "Home Is Where Your FinTech Loan Is." *Available at SSRN 5168570*.

Farboodi, Maryam, Dhruv Singal, Laura Veldkamp, and Venky Venkateswaran. 2025. "Valuing financial data." *The Review of Financial Studies*, 38(3): 938–980.

Gelman, Andrew, and Guido Imbens. 2019. "Why high-order polynomials should not be used in regression discontinuity designs." *Journal of Business & Economic Statistics*, 37(3): 447–456.

Goldberg, Samuel, Garrett Johnson, and Scott Shriver. 2019. "Regulating privacy online: The early impact of the GDPR on European web traffic & e-commerce outcomes." *Available at SSRN*, 3421731.

Goldfarb, Avi, and Catherine Tucker. 2019. "Digital economics." *Journal of economic literature*, 57(1): 3–43.

Hertzberg, Andrew, Jose Maria Liberti, and Daniel Paravisini. 2011. "Public information and coordination: evidence from a credit registry expansion." *The Journal of Finance*, 66(2): 379–412.

Janßen, Rebecca, Reinhold Kesler, Michael E Kummer, and Joel Waldfogel. 2022. "GDPR and the lost generation of innovative apps." National Bureau of Economic Research.

Jia, Jian, Ginger Zhe Jin, and Liad Wagman. 2021. "The short-run effects of the general data protection regulation on technology venture investment." *Marketing Science*, 40(4): 661–684.

Karlan, Dean, Markus Mobius, Tanya Rosenblat, and Adam Szeidl. 2009. "Trust and social collateral." *The Quarterly Journal of Economics*, 124(3): 1307–1361.

Kesler, Reinhold. 2023. "The Impact of Apple's App Tracking Transparency on App Monetization." *Available at SSRN 4090786*.

Kraft, Lennart, Bernd Skiera, and Tim Koschella. 2023. "Economic impact of opt-in versus opt-out requirements for personal data usage: The case of apple's app tracking transparency (att)." *Available at SSRN 4598472*.

Liberman, Andres, Christopher Neilson, Luis Opazo, and Seth Zimmerman. 2018. "The equilibrium effects of information deletion: Evidence from consumer credit markets." National Bureau of Economic Research.

Lin, Tesary. 2022. "Valuing intrinsic and instrumental preferences for privacy." *Marketing Science*, 41(4): 663–681.

Peukert, Christian, Stefan Bechtold, Michail Batikas, and Tobias Kretschmer. 2022. "Regulatory spillovers and data governance: Evidence from the GDPR." *Marketing Science*, 41(4): 746–768.

Ramadorai, Tarun, Antoine Uettwiller, and Ansgar Walther. 2025. "Privacy Policies and Consumer Data Extraction: Evidence From US Firms." *Review of Finance*, rfaf017.

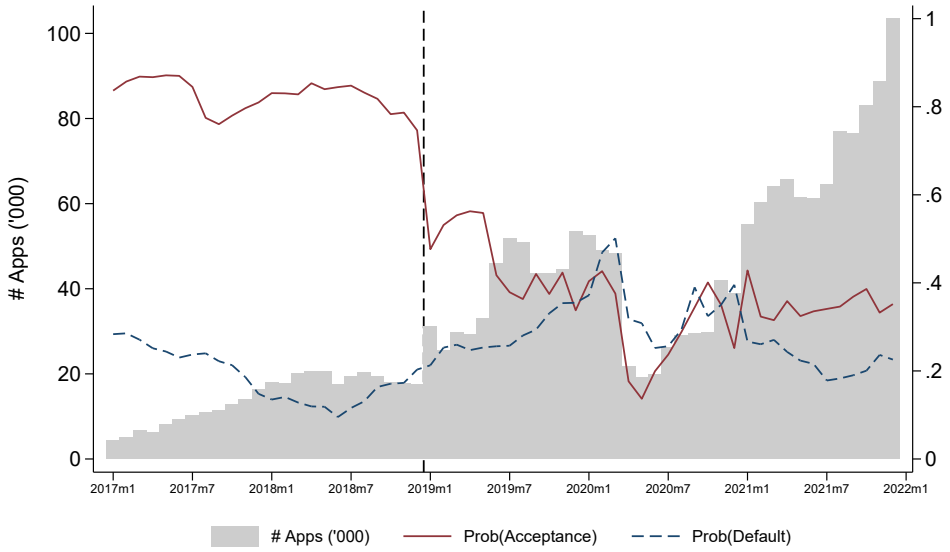
Tang, Huan. 2019. "The value of privacy: Evidence from online borrowers." *Available at SSRN*.

Thakor, Anjan V. 2020. "Fintech and banking: What do we know?" *Journal of financial intermediation*, 41: 100833.

Veldkamp, Laura. 2023. "Valuing data as an asset." *Review of Finance*, 27(5): 1545–1562.

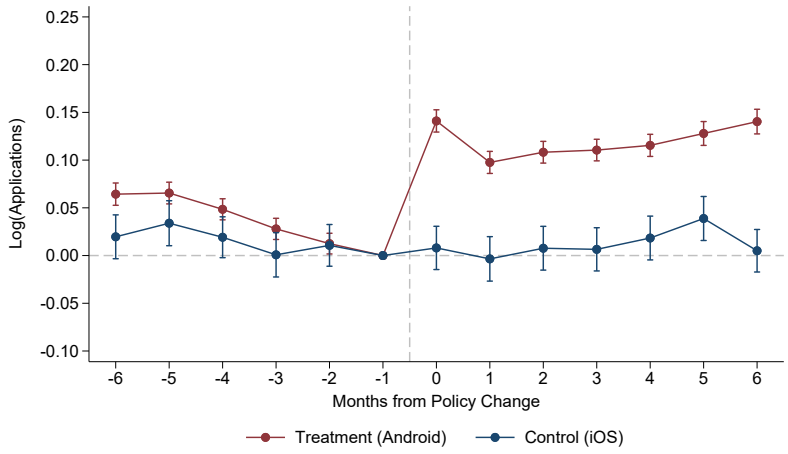
Veldkamp, Laura, and Cindy Chung. 2024. "Data and the aggregate economy." *Journal of Economic Literature*, 62(2): 458–484.

Figure 1: Number of Applications, Probability of Acceptance and Default over time

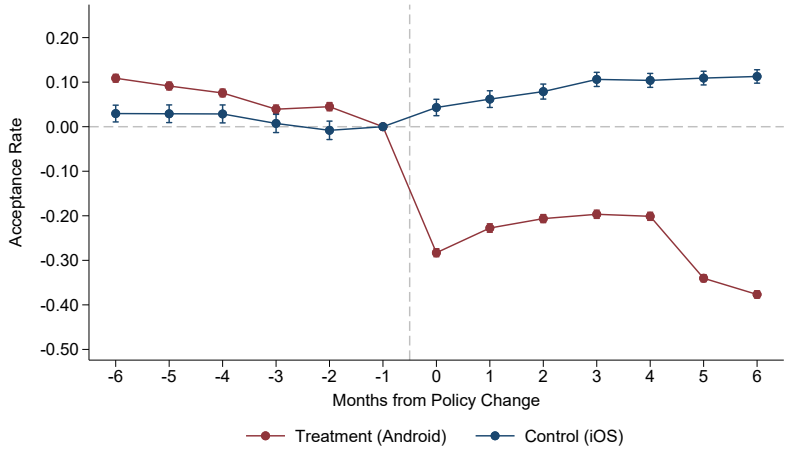


This figure illustrates the number of monthly applications, the monthly average probability of acceptance, the monthly average probability of default from January 2017 until December 2021.

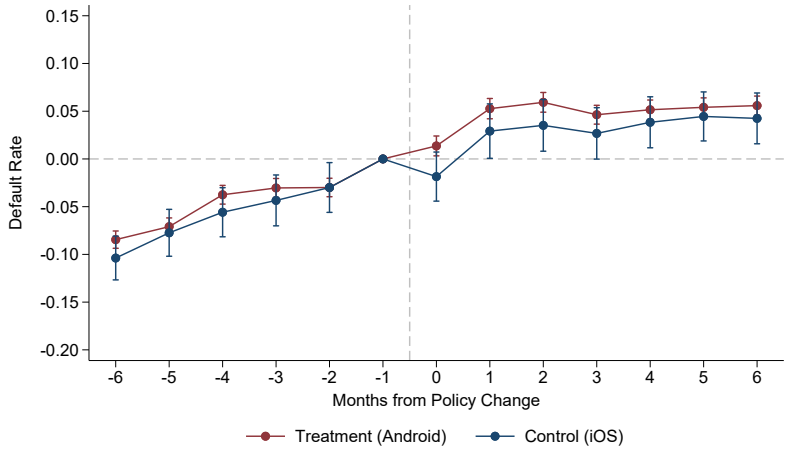
Figure 2: Unconditional Means: Treatment and Control Group Trends



(a) Number of Applications



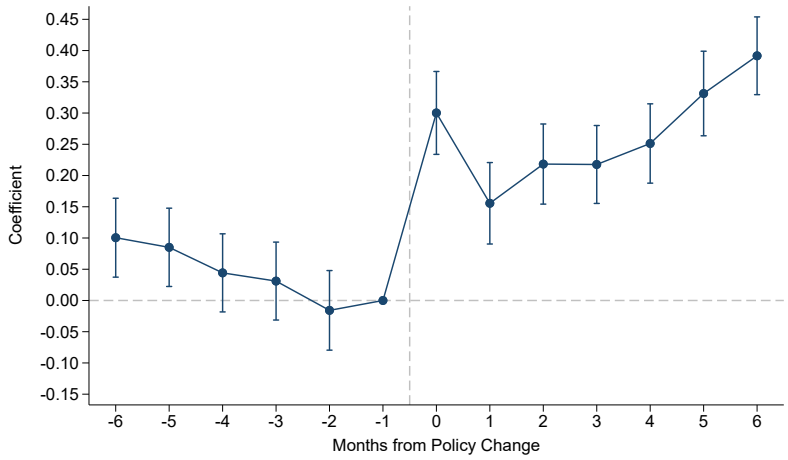
(b) Application Acceptance Rate



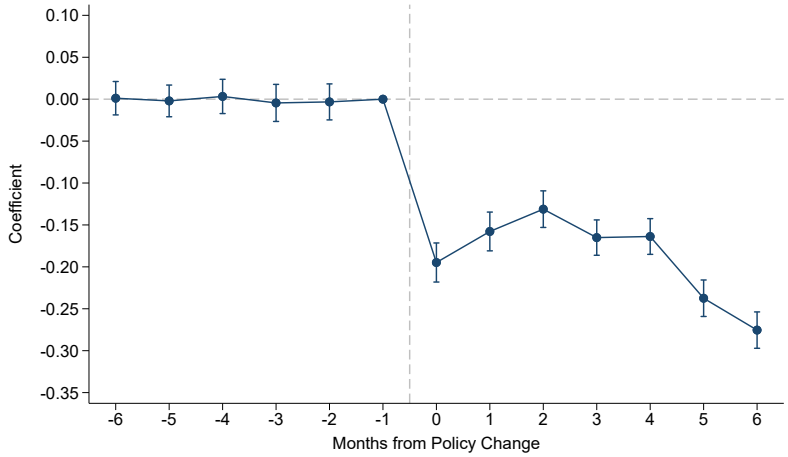
(c) Default Rate

This figure presents unconditional mean outcomes by operating system around the Google privacy policy implementation on January 1, 2019. Each panel displays coefficient estimates from separate regressions for Android users (treatment group, shown in maroon) and iOS users (control group, shown in navy) on month indicators, with month -1 as the omitted category. Panel (a) shows the natural logarithm of daily applications. Panel (b) shows application acceptance rates. Panel (c) shows default rates for approved loans. The sample is restricted to observations within 6 months on either side of the policy change. Vertical bars represent 95% confidence intervals based on standard errors clustered at the ZIP code level. The vertical dashed line indicates the policy implementation date. These unconditional trends illustrate the raw differences between Android and iOS users before controlling for fixed effects.

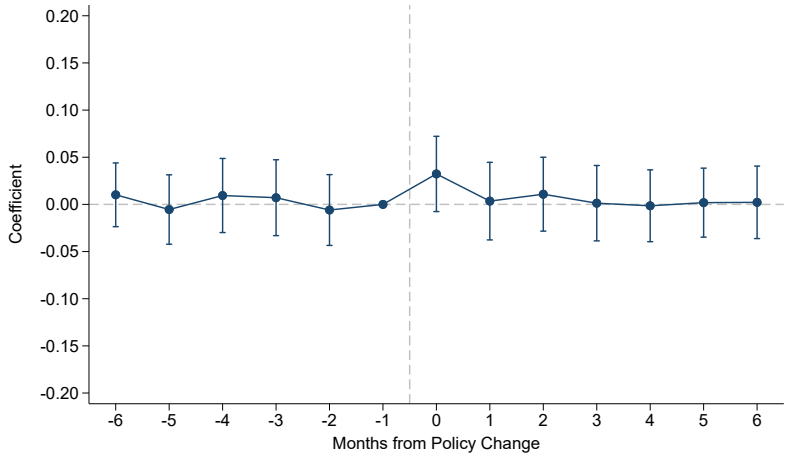
Figure 3: Parallel Trends



(a) Number of Applications



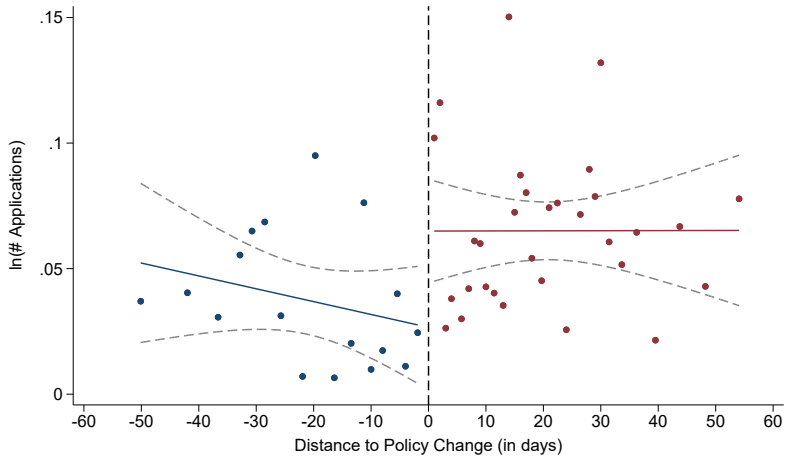
(b) Application Accepted



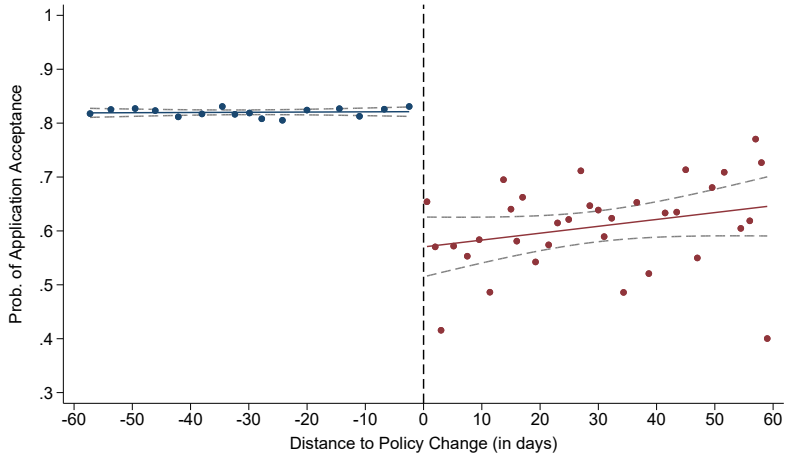
(c) Default Probability

This figure presents event study estimates examining parallel trends in loan market outcomes around the Google privacy policy implementation on January 1, 2019. Each panel plots coefficients from a difference-in-differences regression comparing Android users (treated group) to iOS users (control group), with month fixed effects relative to the policy change. Panel (a) shows the natural logarithm of daily applications. Panel (b) shows application acceptance rates. Panel (c) shows default probabilities. The specification includes ZIP code × month × loan purpose fixed effects to control for location-specific temporal trends that vary by loan type. Vertical bars represent 95% confidence intervals based on standard errors clustered at the zip-code level.

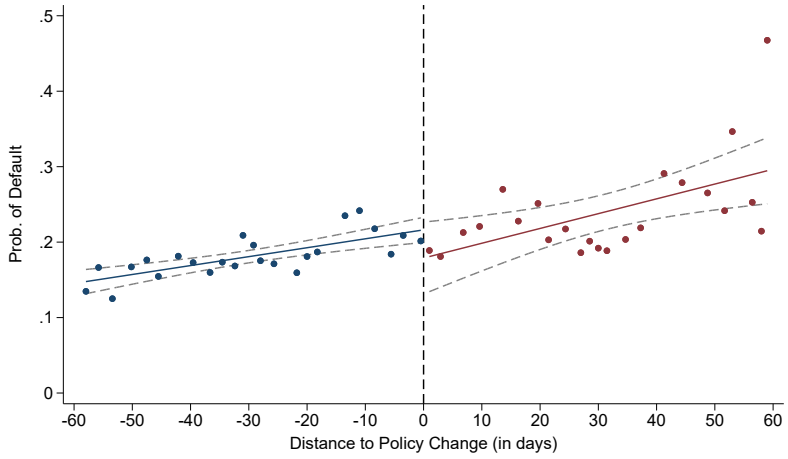
Figure 4: Main Effects of Privacy Policy Change: High-Frequency Event Study



(a) Number of Applications



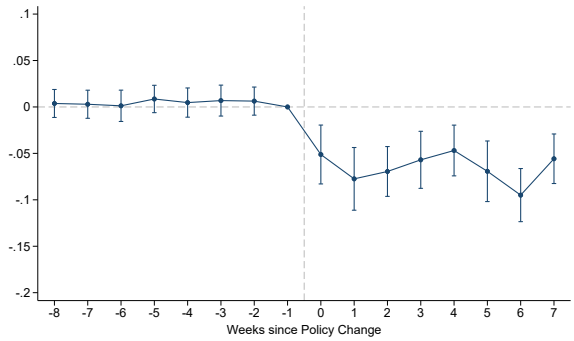
(b) Application Acceptance



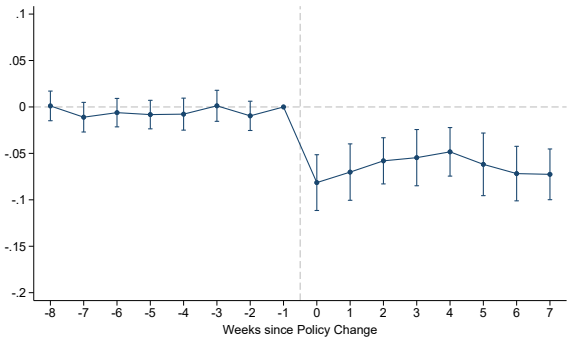
(c) Default Probability

This figure presents high-frequency event study estimates of the Google privacy policy change's effect on loan market outcomes, implemented on January 1, 2019. Panel (a) shows the natural logarithm of daily loan applications. Panel (b) shows the probability of application acceptance. Panel (c) shows the probability of loan default. Each panel displays binned averages along the running variable (distance in days to the policy change date) with local linear trends estimated separately on each side of the discontinuity, along with 95% confidence intervals. Red dots indicate days after the policy implementation and blue dots indicate days before. All estimates include ZIP code × loan purpose and day of the week fixed effects. The unit of observation is a day for panels (a) and (b), and an individual loan for panel (c). The sample is restricted to Android users only, as iOS users were unaffected by the policy.

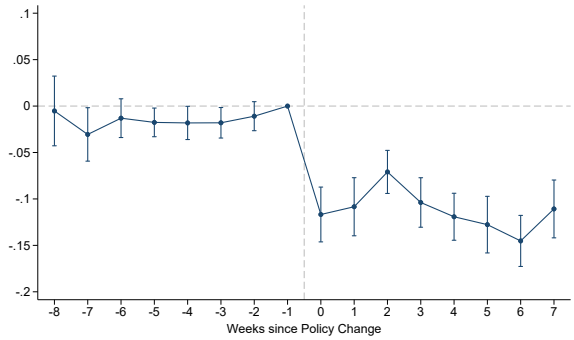
Figure 5: DiD: Heterogeneous Effects on Application Acceptance by Borrower Characteristics



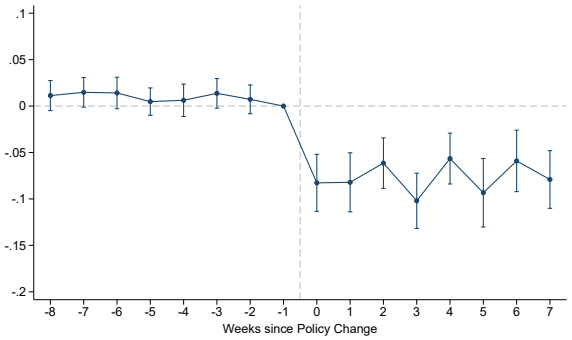
(a) Monthly Borrower Income



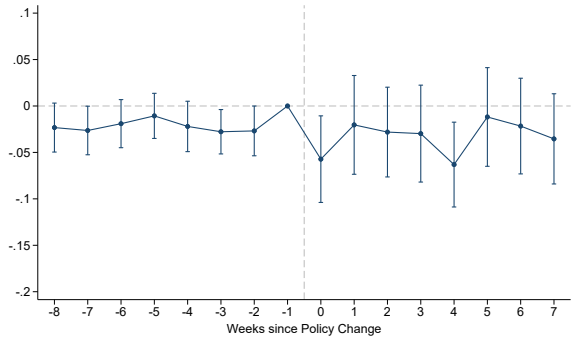
(b) Age



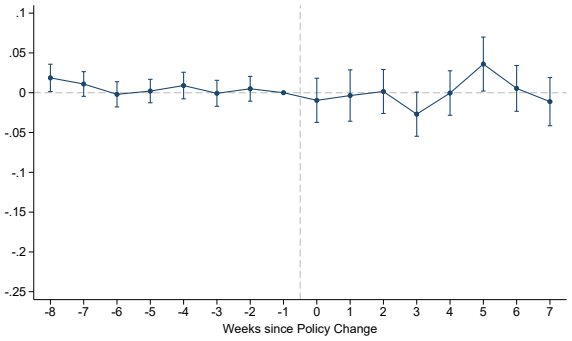
(c) Prior Relationship



(d) Caste



(e) Gender



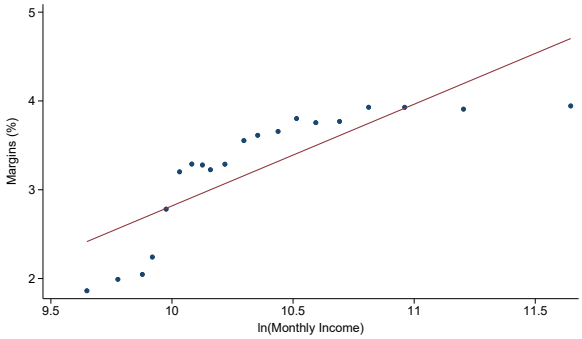
(f) Credit Scores

Notes: This figure presents the DID estimates and 95% confidence intervals for heterogeneous treatment effects on application acceptance status across different borrower characteristics. Each panel shows results from the following specification:

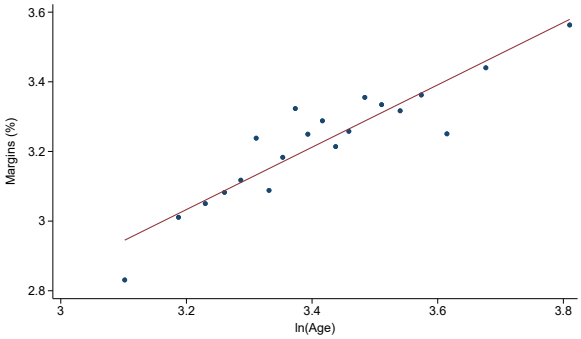
$$y_{i,t} = \sum_{k=-8, k\neq -1}^{k=+7} \beta_k \cdot \text{Characteristic}_i \times \mathbb{I}(t = k) + \theta_{z,p} + \theta_{z,t} + \theta_{p,t} + \theta_w + \varepsilon_{i,t}$$

All regressions include ZIP code × loan purpose ($\theta_{z,p}$), ZIP code × week ($\theta_{z,t}$), loan purpose × week ($\theta_{p,t}$), and day of the week (θ_w) fixed effects. Variable definitions: Panel 5a *Low Income* equals one for borrowers with monthly income below the sample median; Panel 5b *Young* equals one for borrowers with age below the sample median; Panel 5c *New borrower* equals one for first-time applicants; Panel 5d *Low Caste* equals one for applicants belonging to scheduled caste, scheduled tribe, and other backward class; Panel 5e *Female* equals one for female applicants; Panel 5f *Sub-prime Credit Score* is a binary variable equal to one for borrowers with CIBIL credit score less than 680. 95% confidence bands are calculated based on standard errors clustered at the zip-code level. The sample is restricted to Android users only, as iOS users were unaffected by the policy.

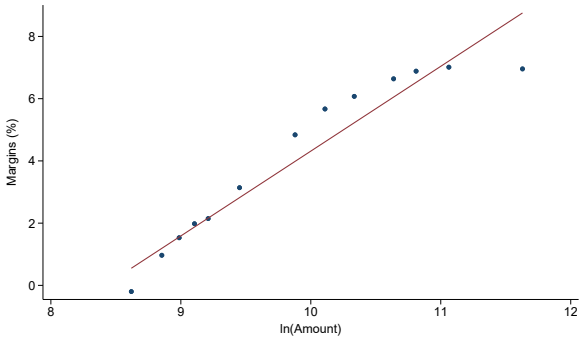
Figure 6: Profit Margins of Marginal Borrowers by Characteristics



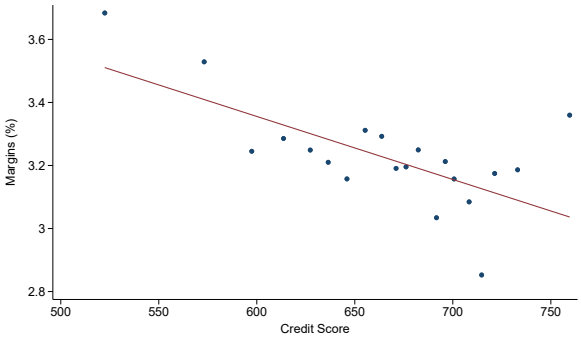
(a) By Income



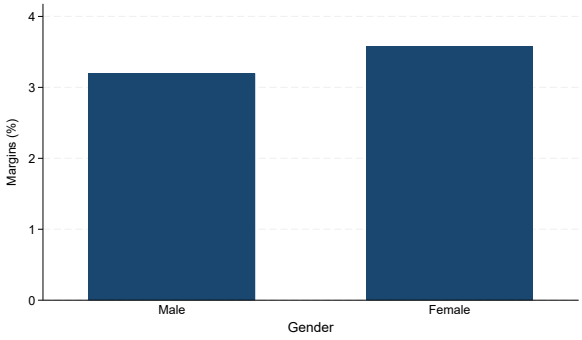
(b) By Age



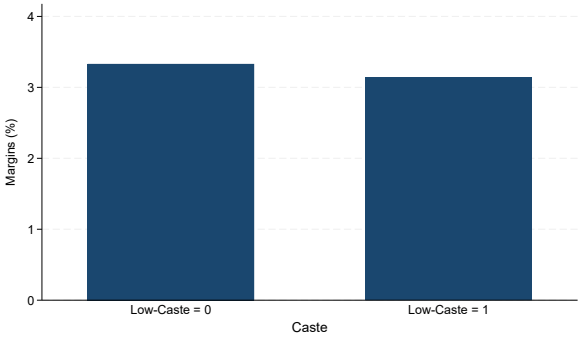
(c) By Loan Amount



(d) By Credit Score



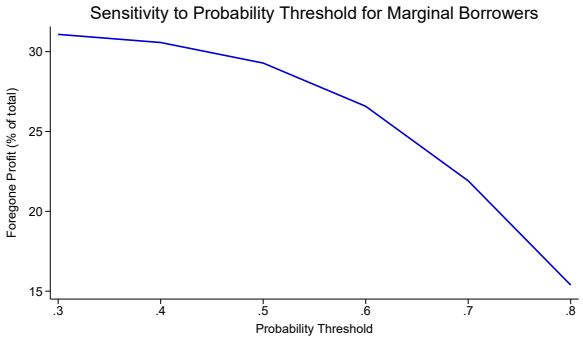
(e) By Gender



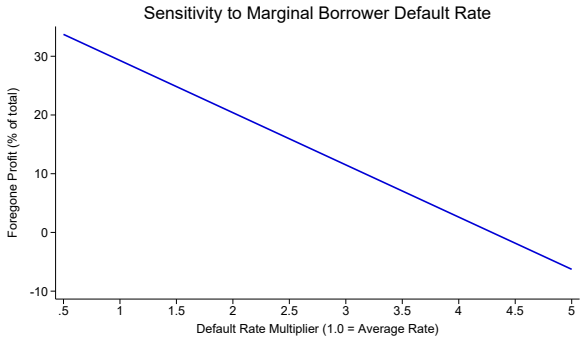
(f) By Caste

Notes: This figure shows the expected profit margins (as percentage of loan amount) for marginal borrowers—those rejected post-policy but who would have been approved pre-policy based on predicted approval probability > 0.5. Panel 6a shows margins by log monthly income with quadratic fitted trend. Panel 6b shows margins by log age with linear fitted trend. Panel 6c shows margins by log loan amount displaying a concave relationship. Panel 6d shows margins by credit score with a declining trend line. Panels 6e and 6f show average margins by gender and caste categories respectively. Profit calculations assume bullet loan structure with interest and principal repaid at maturity, 19% baseline default rate, 9% loss given default, and 1% monthly discount rate. Sample includes marginal borrowers within 60-day window around January 1, 2019 policy implementation.

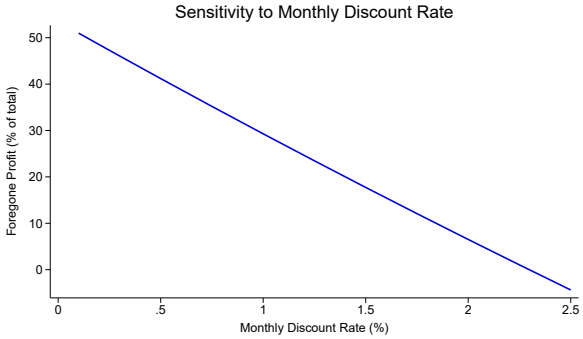
Figure 7: Sensitivity Analysis of Foregone Profits from Privacy Regulation



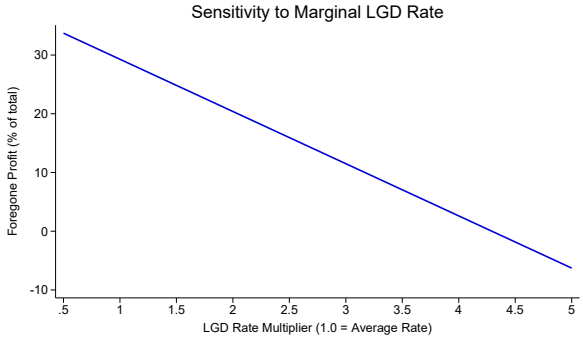
(a) Probability Threshold



(b) Default Rate



(c) Monthly Discount Rate



(d) Loss Given Default

Notes: This figure presents sensitivity analysis for the foregone profit calculations from the privacy policy change. Panel 7a shows how results vary with the probability threshold used to define marginal borrowers (those rejected post-policy but who would have been approved pre-policy). Panel 7b examines sensitivity to assumptions about marginal borrowers' default rates, expressed as multiples of the baseline 19% rate. Panel 7c shows how results depend on the monthly discount rate used in present value calculations. Panel 7d examines sensitivity to the loss given default (LGD) rate, expressed as multiples of the baseline 9% rate. All calculations assume bullet loan structure with interest and principal repaid at maturity. Baseline parameters: 0.5 probability threshold, 19% default rate, 1% monthly discount rate, 9% LGD. Sample includes borrowers within 60-day window around January 1, 2019 policy implementation.

Table 1: Number of Contacts and Default Probability: Pre-Policy Period

Dep Var: 𝟙[Default]	(1)	(2)	(3)	(4)	(5)
ln(Num. Contacts)	-0.0145***	-0.0133***	-0.0136***	-0.0127***	-0.0171***
	(0.0032)	(0.0034)	(0.0036)	(0.0037)	(0.0036)
Zip Code × Month-Year FE		✓	✓		
Zip Code × Loan Purpose FE			✓		
Zip Code × Month-Year × Loan Purpose FE				✓	✓
Education FE					✓
Age Decile FE					✓
Credit Score Decile FE					✓
Loan Maturity FE					✓
Income Decile FE					✓
Female FE					✓
Low Caste FE					✓
N	44,856	44,856	44,856	44,856	44,856
R^2	0.0011	0.2507	0.2833	0.3825	0.4340

This table examines the relationship between contact list size and default probability for loans originated between January 2017 and December 2018, prior to the Google privacy policy change. The dependent variable is a binary indicator equal to one if the borrower defaulted on their loan. The key independent variable is the natural logarithm of the number of contacts in the borrower's phone at the time of loan application. Column (1) presents the baseline specification without fixed effects. Column (2) incorporates ZIP code × month-year fixed effects to control for location-specific temporal trends. Column (3) adds ZIP code × loan purpose fixed effects to account for geographic variation in loan demand across different product types. Column (4) includes the interaction of ZIP code × month-year × loan purpose fixed effects to control for location-specific temporal trends that vary by loan type. Column (5) adds comprehensive demographic controls including education fixed effects, age decile fixed effects, credit score decile fixed effects, loan maturity fixed effects, income decile fixed effects, female fixed effects, and low caste fixed effects. Standard errors reported in parentheses are clustered at the zip-code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Effect of the Policy Change on Number of Applications

Dep Var: ln(# Applications)	(1)	(2)	(3)	(4)
Android × Post	0.2113***	0.2113***	0.2113***	0.2113***
	(0.0142)	(0.0142)	(0.0142)	(0.0142)
Android	0.8212***	0.8212***	0.8212***	0.8212***
	(0.0239)	(0.0239)	(0.0239)	(0.0239)
Post	-0.0192***			
	(0.0066)			
ZIP Code × Month-Year FE		✓	✓	
Zip Code × Loan Purpose FE			✓	
ZIP Code × Month-Year × Loan Purpose FE				✓
# Obs	23,790	23,790	23,790	23,790
R^2	0.3006	0.6710	0.7347	0.7732
Adj. R^2	0.3006	0.5415	0.5220	0.5464

This table presents difference-in-differences estimates examining the effect of the Google privacy policy change on monthly loan applications. The dependent variable is the natural logarithm of the number of applications per month. The key independent variable is Android × Post, which equals one for Android users applying after January 1, 2019, and zero otherwise. The sample includes applications within 180 days on either side of the policy change date. Column (1) presents the baseline specification without fixed effects. Column (2) incorporates ZIP code × month-year fixed effects. Column (3) adds ZIP code × loan purpose fixed effects. Column (4) includes the full interaction of ZIP code × month-year × loan purpose fixed effects. Standard errors reported in parentheses are clustered at the zip-code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Effect of the Policy Change on Application Acceptance

Dep Var: Application Acceptance	(1)	(2)	(3)	(4)	(5)
Android × Post	-0.4492***	-0.3904***	-0.1737***	-0.1752***	-0.1811***
	(0.0050)	(0.0050)	(0.0040)	(0.0046)	(0.0036)
Android	-0.0193***	-0.0133***	-0.0199***	-0.0120***	0.0000
	(0.0034)	(0.0035)	(0.0030)	(0.0029)	(0.0011)
Post	0.0453***				
	(0.0035)				
ZIP Code × Month-Year FE		✓	✓		
Zip Code × Loan Purpose FE			✓		
ZIP Code × Month-Year × Loan Purpose FE				✓	✓
Education FE					✓
Age Decile FE					✓
Credit Score Decile FE					✓
Loan Maturity FE					✓
Income Decile FE					✓
Female FE					✓
Low Caste FE					✓
# Obs	204,682	204,682	204,682	204,682	193,181
R^2	0.1812	0.3379	0.6938	0.7421	0.7914
Adj. R^2	0.1812	0.2623	0.6304	0.6609	0.7261

This table presents difference-in-differences estimates examining the effect of the Google privacy policy change on loan application acceptance rates. The dependent variable is a binary indicator equal to one if the application was accepted. The key independent variable is Android × Post, which equals one for Android users applying after January 1, 2019, and zero otherwise. The sample includes applications within 180 days on either side of the policy change date. Column (1) presents the baseline specification without fixed effects. Column (2) incorporates ZIP code × month-year fixed effects. Column (3) adds ZIP code × loan purpose fixed effects. Column (4) includes the full interaction of ZIP code × month-year × loan purpose fixed effects. Column (5) adds comprehensive demographic controls including education fixed effects, age decile fixed effects, credit score decile fixed effects, loan maturity fixed effects, income decile fixed effects, female fixed effects, and low caste fixed effects. Standard errors reported in parentheses are clustered at the zip-code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Effect of the Policy Change on Default

Dep Var: Loan Default	(1)	(2)	(3)	(4)	(5)
Android × Post	0.0093	0.0102	0.0115*	0.0049	0.0055
	(0.0061)	(0.0066)	(0.0066)	(0.0069)	(0.0069)
Android	0.0131***	0.0154***	0.0157***	0.0175***	0.0150***
	(0.0043)	(0.0046)	(0.0047)	(0.0049)	(0.0048)
Post	0.0828***				
	(0.0056)				
ZIP Code × Month-Year FE		✓	✓		
Zip Code × Loan Purpose FE			✓		
ZIP Code × Month-Year × Loan Purpose FE				✓	✓
Education FE					✓
Age Decile FE					✓
Credit Score Decile FE					✓
Loan Maturity FE					✓
Income Decile FE					✓
Female FE					✓
Low Caste FE					✓
# Obs	130,352	130,352	130,352	130,352	130,171
R^2	0.0129	0.1328	0.1626	0.2608	0.3255
Adj. R^2	0.0129	0.0187	-0.0152	0.0239	0.1089

This table presents difference-in-differences estimates examining the effect of the Google privacy policy change on loan default rates. The dependent variable is a binary indicator equal to one if the borrower defaulted on their loan (defined as 90+ days past due). The key independent variable is Android × Post, which equals one for Android users whose loans were originated after January 1, 2019, and zero otherwise. The sample includes approved loans originated within 180 days on either side of the policy change date. Column (1) presents the baseline specification without fixed effects. Column (2) incorporates ZIP code × month-year fixed effects. Column (3) adds ZIP code × loan purpose fixed effects. Column (4) includes the full interaction of ZIP code × month-year × loan purpose fixed effects. Column (5) adds comprehensive demographic controls including education fixed effects, age decile fixed effects, credit score decile fixed effects, loan maturity fixed effects, income decile fixed effects, female fixed effects, and low caste fixed effects. Standard errors reported in parentheses are clustered at the zip-code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Effect of the Policy Change on Interest Rates and Loan Maturity

Panel A: Annual Interest Rates (%)

Dep Var: Interest Rate (Annual)	(1)	(2)	(3)	(4)	(5)
Android × Post	0.3504***	0.2769**	0.2985**	0.3324**	0.2547**
	(0.1184)	(0.1223)	(0.1240)	(0.1332)	(0.1201)
Android	1.2130***	1.0297***	1.0049***	0.9743***	0.2839***
	(0.1149)	(0.1211)	(0.1225)	(0.1269)	(0.1097)
Post	1.4317***				
	(0.1115)				
ZIP Code × Month-Year FE		✓	✓		
Zip Code × Loan Purpose FE			✓		
ZIP Code × Month-Year × Loan Purpose FE				✓	✓
Demographic Controls					✓
# Obs	127,371	127,371	127,371	127,371	127,194
R^2	0.0221	0.1803	0.2161	0.3031	0.4493
Adj. R^2	0.0221	0.0709	0.0473	0.0775	0.2708

Panel B: Loan Maturity (Months)

Dep Var: Loan Maturity (Months)	(1)	(2)	(3)	(4)	(5)
Android × Post	-0.1123***	-0.1058***	-0.1141***	-0.1232***	-0.1005***
	(0.0362)	(0.0376)	(0.0381)	(0.0410)	(0.0371)
Android	-0.3628***	-0.2938***	-0.2855***	-0.2787***	-0.0664**
	(0.0339)	(0.0364)	(0.0369)	(0.0383)	(0.0332)
Post	0.6994***				
	(0.0340)				
ZIP Code × Month-Year FE		✓	✓		
Zip Code × Loan Purpose FE			✓		
ZIP Code × Month-Year × Loan Purpose FE				✓	✓
Demographic Controls					✓
# Obs	127,371	127,371	127,371	127,371	127,194
R^2	0.0250	0.1708	0.2060	0.2939	0.4248
Adj. R^2	0.0250	0.0601	0.0350	0.0653	0.2383

This table presents difference-in-differences estimates examining the effect of the Google privacy policy change on loan pricing terms for approved loans. Panel A shows effects on annual interest rates (in percentage points), while Panel B shows effects on loan maturity (in months). The key independent variable is Android × Post, which equals one for Android users whose loans were originated after January 1, 2019, and zero otherwise. The sample includes approved loans originated within 180 days on either side of the policy change date. Column (1) presents the baseline specification without fixed effects. Column (2) incorporates ZIP code × month-year fixed effects. Column (3) adds ZIP code × loan purpose fixed effects. Column (4) includes the full interaction of ZIP code × month-year × loan purpose fixed effects. Column (5) adds comprehensive demographic controls including education, age decile, credit score decile, income decile, female, and low caste fixed effects. Standard errors reported in parentheses are clustered at the zip-code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Effect of the Policy Change on Loan Market Outcomes

Panel A: Number of Applications

Dep Var: ln(# Applications)	(1)	(2)	(3)	(4)	(5)
Post	0.0453***	0.0568***	0.0347***	0.0338***	0.0324***
	(0.0173)	(0.0139)	(0.0107)	(0.0087)	(0.0075)
ZIP Code × Loan Purpose FE	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓
Bandwidth	10	15	20	25	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	6,155	11,127	15,613	21,266	27,951
R^2	0.4066	0.3367	0.2960	0.2759	0.2620
Adj. R^2	0.1034	0.0551	0.0297	0.0286	0.0339

Panel B: Application Acceptance

Dep Var: Application Accepted	(1)	(2)	(3)	(4)	(5)
Post	-0.2526***	-0.2628***	-0.2850***	-0.2790***	-0.2749***
	(0.0210)	(0.0151)	(0.0121)	(0.0104)	(0.0093)
ZIP Code × Loan Purpose FE	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓
Bandwidth	10	15	20	25	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	6,338	11,373	15,538	21,037	27,770
R^2	0.7762	0.7651	0.7522	0.7485	0.7465
Adj. R^2	0.6768	0.6791	0.6694	0.6716	0.6761

Panel C: Default Probability

Dep Var: Loan Default	(1)	(2)	(3)	(4)	(5)
Post	-0.0126	0.0063	0.0029	-0.0077	-0.0072
	(0.0407)	(0.0288)	(0.0228)	(0.0200)	(0.0178)
ZIP Code × Loan Purpose FE	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓
Bandwidth	10	15	20	25	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	3,591	7,063	9,422	12,747	17,749
R^2	0.3572	0.3302	0.2943	0.2774	0.2788
Adj. R^2	0.0573	0.0842	0.0568	0.0544	0.0770

This table presents high-frequency event study estimates of the Google privacy policy change's effect on loan market outcomes. Panel A shows effects on the natural logarithm of daily applications. Panel B shows effects on application acceptance rates. Panel C shows effects on default probability. The key independent variable is Post, which equals one for observations after January 1, 2019. Columns (1) through (5) use bandwidths of 10, 15, 20, 25, and 30 days on either side of the policy change date, respectively. All regressions include ZIP code × loan purpose and day of the week fixed effects, a local linear specification estimated separately on each side of the policy change date, and use a triangular kernel. The sample is restricted to Android users only, as iOS users were unaffected by the policy. Standard errors reported in parentheses are clustered at the zip-code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Effect of the Policy Change on Android Applications

Dep Var: Android Applicant	(1)	(2)	(3)	(4)	(5)
Post	0.0071	0.0086	0.0127	0.0101	0.0096
	(0.0252)	(0.0194)	(0.0164)	(0.0137)	(0.0120)
ZIP Code X Loan Purpose FE	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓
Bandwidth	10	15	20	25	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	7,089	12,698	17,236	23,231	30,625
R^2	0.3508	0.3224	0.2885	0.2774	0.2710
Adj. R^2	0.0769	0.0877	0.0637	0.0686	0.0810

This table presents high-frequency event study estimates examining the effect of the privacy policy change on the composition of loan applicants by operating system. The dependent variable is a binary indicator equal to one if the application was submitted by an Android user and zero if submitted by an iOS user. We use different bandwidths of 10, 15, 20, 25, and 30 days on either side of the policy change date (January 1, 2019), in columns (1), (2), (3), (4), and (5) respectively. All regressions include Zip Code × loan reason and the day of the week fixed effects, a local linear specification estimated separately on each side of the policy change date, and use a triangular kernel. Standard errors reported in parentheses are clustered at the Zip Code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Heterogeneity & Effect of the Policy Change on Application Acceptances

Dep Var: Application Accepted	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post	-0.2753***	-0.2643***	-0.2417***	-0.2052***	-0.2396***	-0.2282***	-0.2687***	-0.0812***
	(0.0094)	(0.0142)	(0.0113)	(0.0133)	(0.0114)	(0.0117)	(0.0097)	(0.0204)
Sub-Prime Credit Score × Post		-0.0180						-0.0202
		(0.0168)						(0.0163)
Sub-Prime Credit Score		0.0007						-0.0020
		(0.0102)						(0.0097)
Low Caste × Post			-0.0794***					-0.0601***
			(0.0166)					(0.0163)
Low Caste			-0.0029					-0.0037
			(0.0107)					(0.0104)
New Borrower × Post				-0.1209***				-0.1302***
				(0.0159)				(0.0160)
New Borrower				0.0029				0.0013
				(0.0105)				(0.0107)
Low Income × Post					-0.0757***			-0.0727***
					(0.0163)			(0.0163)
Low Income					-0.0067			-0.0076
					(0.0100)			(0.0099)
Young × Post						-0.0904***		-0.0796***
						(0.0162)		(0.0161)
Young						-0.0089		-0.0059
						(0.0100)		(0.0098)
Female × Post							-0.0632**	-0.0540**
							(0.0275)	(0.0268)
Female							0.0161	0.0168
							(0.0176)	(0.0175)
ZIP Code X Loan Purpose FE	✓	✓	✓	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓	✓	✓	✓
Bandwidth	30	30	30	30	30	30	30	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	27,190	27,190	27,190	27,190	27,190	27,190	27,190	27,190
R^2	0.7478	0.7479	0.7514	0.7546	0.7506	0.7509	0.7480	0.7638
Adj. R^2	0.6773	0.6773	0.6818	0.6860	0.6808	0.6812	0.6775	0.6975

This table presents high-frequency event study estimates using the application acceptance status as the dependent variable. The dependent variable takes a value of one if the application was accepted and zero otherwise. All regressions use a bandwidth of 30 days on either side of the policy change date (January 1, 2019). All regressions include Zip Code × loan reason and the day of the week fixed effects, a local linear specification estimated separately on each side of the policy change date, and use a triangular kernel. Standard errors reported in parentheses are clustered at the zip-code level. Sub-prime credit score is a binary variable that takes a value of one for borrowers with CIBIL credit score less than 680, and zero otherwise. Low caste is a binary variable that takes a value of one for applicants belonging to the scheduled caste, scheduled tribe and other backward class, and zero otherwise. New borrower is a binary variable for borrowers who are first time applicants. Low income is a binary variable that takes a value of one for borrowers with monthly income less than the sample median value, and zero otherwise. young is a binary variable that takes a value of one for borrowers with age less than the sample median value, and zero otherwise. Female is a binary variable that takes a value of one for females, and zero otherwise. The sample is restricted to Android users only, as iOS users were unaffected by the policy. Standard errors are clustered at the zip-code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Credit Access: Reduced Form Results

	2 Years				3 Years				4 Years			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Android × Post	-0.0975***	-0.0303***	-0.0313***	-0.0253**	-0.1011***	-0.0316***	-0.0355***	-0.0270***	-0.0947***	-0.0276***	-0.0323***	-0.0240**
	(0.0081)	(0.0089)	(0.0095)	(0.0098)	(0.0084)	(0.0092)	(0.0099)	(0.0102)	(0.0083)	(0.0092)	(0.0099)	(0.0100)
Android	-0.0421***	-0.0486***	-0.0467***	-0.0140	-0.0534***	-0.0567***	-0.0526***	-0.0093	-0.0587***	-0.0605***	-0.0558***	-0.0077
	(0.0093)	(0.0097)	(0.0103)	(0.0112)	(0.0096)	(0.0101)	(0.0108)	(0.0117)	(0.0093)	(0.0098)	(0.0104)	(0.0111)
Post	0.0880***				0.0871***				0.0807***			
	(0.0074)				(0.0077)				(0.0077)			
ZIP Code × Month-Year FE		✓				✓				✓		
ZIP Code × Month-Year × Loan Purpose FE			✓	✓			✓	✓			✓	✓
Education FE				✓				✓				✓
Age Decile FE				✓				✓				✓
Credit Score Decile FE				✓				✓				✓
Loan Maturity FE				✓				✓				✓
Income Decile FE				✓				✓				✓
Female FE				✓				✓				✓
Low Caste FE				✓				✓				✓
N	204,682	204,682	204,682	193,181	204,682	204,682	204,682	193,181	204,682	204,682	204,682	193,181
R^2	0.0040	0.2143	0.3075	0.3373	0.0048	0.2159	0.3104	0.3531	0.0047	0.2149	0.3100	0.3595
Adj. R^2	0.0040	0.0514	0.0896	0.1301	0.0048	0.0534	0.0934	0.1508	0.0047	0.0522	0.0929	0.1593

This table presents difference-in-differences estimates examining the effect of the privacy policy change on long-term credit access. The dependent variables are indicators for whether a borrower accessed any formal credit within two years (columns 1-3), three years (columns 4-6), or four years (columns 7-9) following January 1, 2019. The key independent variable is Android × Post, which equals one for Android users who applied after January 1, 2019, and zero otherwise. The sample includes all applications within 180 days on either side of the policy change date. Column (1), (4), and (7) present baseline specifications without fixed effects. Columns (2), (5), and (8) add ZIP code × loan purpose and ZIP code × month-year fixed effects. Columns (3), (6), and (9) include the full interaction of ZIP code × loan purpose × month-year fixed effects. Standard errors reported in parentheses are clustered at the ZIP code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Credit Access: Instrumental Variables Results

	2 Years				3 Years				4 Years			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
App Accepted	0.0798***	0.4254***	0.4288***	0.2162***	0.0973***	0.4769***	0.4841***	0.2000***	0.0984***	0.4776***	0.4856***	0.1748***
	(0.0071)	(0.0469)	(0.0491)	(0.0512)	(0.0073)	(0.0471)	(0.0492)	(0.0500)	(0.0076)	(0.0457)	(0.0477)	(0.0482)
ZIP Code × Month-Year FE		✓				✓				✓		
ZIP Code × Month-Year × Loan Purpose FE			✓	✓			✓	✓			✓	✓
Education FE				✓				✓				✓
Age Decile FE				✓				✓				✓
Credit Score Decile FE				✓				✓				✓
Loan Maturity FE				✓				✓				✓
Income Decile FE				✓				✓				✓
Female FE				✓				✓				✓
Low Caste FE				✓				✓				✓
N	204,682	204,682	204,682	193,181	204,682	204,682	204,682	193,181	204,682	204,682	204,682	193,181
KP LM Statistic	648.6040	253.1382	213.8387	217.1915	648.6040	253.1382	213.8387	217.1915	648.6040	253.1382	213.8387	217.1915
KP F Statistic	15100.1119	1344.6566	1308.1302	1508.8734	15100.1119	1344.6566	1308.1302	1508.8734	15100.1119	1344.6566	1308.1302	1508.8734

This table presents instrumental variables estimates of the effect of loan application acceptance on long-term credit access. The dependent variables are indicators for whether a borrower accessed any formal credit within two years (columns 1-3), three years (columns 4-6), or four years (columns 7-9) following January 1, 2019. The endogenous variable is App Accepted, which equals one if the loan application was approved. The instrument set includes Android × Post, Android, and Post, exploiting the exogenous variation in approval rates induced by the privacy policy change. The sample includes all applications within 180 days on either side of the policy change date. Column (1), (4), and (7) present baseline specifications without fixed effects. Columns (2), (5), and (8) add ZIP code × loan purpose and ZIP code × month-year fixed effects. Columns (3), (6), and (9) include the full interaction of ZIP code × loan purpose × month-year fixed effects. KP LM Statistic and KP F Statistic test for underidentification and weak identification, respectively. Standard errors reported in parentheses are clustered at the ZIP code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: Summary Statistics of Profit Margins by Borrower Characteristics

Panel A: Profit Margin by Borrower Group

Characteristic	25th pct	Median	75th pct	Mean	Std Dev
Low Income	1.3331	2.4980	3.5361	2.7867	2.2335
Female	1.5098	3.0379	6.2073	3.5754	2.5521
Young Age	1.4376	2.6853	5.5551	3.1167	2.4276
Low Caste	1.3010	2.6089	5.7949	3.1432	2.5120
Low Amount	0.6657	1.6841	2.4837	1.5026	1.2830
Low Credit Score	1.3567	2.6944	6.0502	3.3130	2.5752

Panel B: Average Group Margins

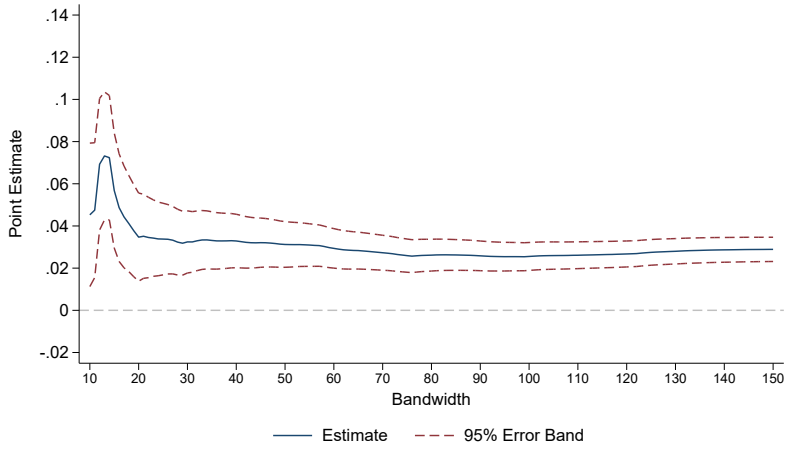
Characteristic	Average Margin (%)
Low Income	4.4620
Female	5.5389
Young Age	5.0959
Low Caste	5.2371
Low Amount	1.6113
Low Credit Score	5.4383

Notes: This table presents summary statistics for expected profit margins (as percentage of loan amount) across different borrower characteristics for marginal borrowers. Marginal borrowers are defined as those rejected post-policy but who would have been approved pre-policy based on predicted approval probability > 0.5. Low income, young age, low amount, and low credit score are defined as below-median values in the sample. Low caste includes scheduled castes, scheduled tribes, and other backward classes. Female is an indicator for female borrowers. Profit calculations assume bullet loan structure with interest and principal repaid at maturity, 19% baseline default rate, 9% loss given default, and 1% monthly discount rate. Sample restricted to 60-day window around January 1, 2019 policy implementation. Groups may overlap across categories. Panel A shows individual-level distribution statistics, while Panel B shows dollar-weighted group averages.

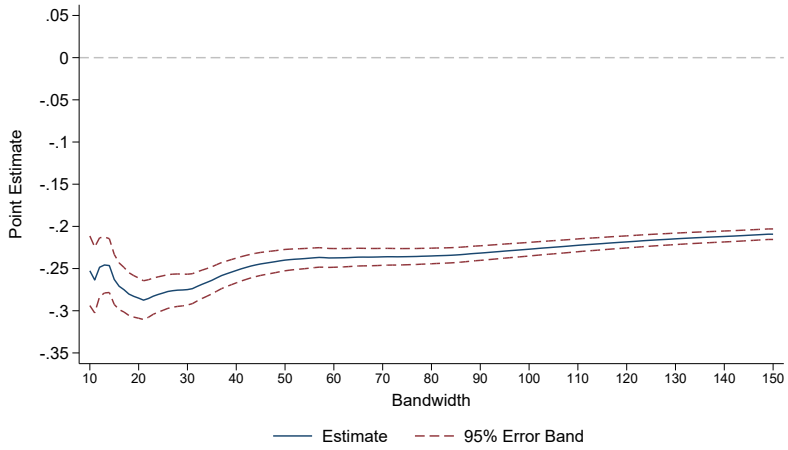
For Online Publication:

"When Privacy Protects but Excludes: Screening, Enforcement, and Credit Access in Digital Markets"

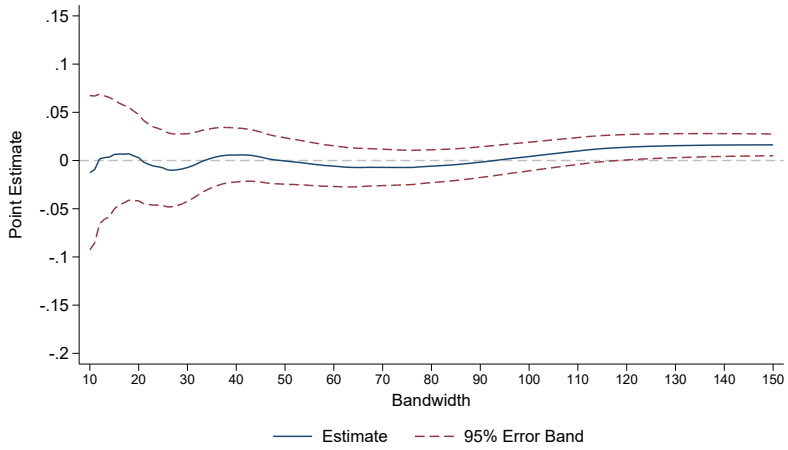
Figure A.1: Robustness to Bandwidth Choice: High-Frequency Event Study Estimates Across Multiple Bandwidths



(a) Number of Applications



(b) Application Acceptance



(c) Default Probability

This figure presents high-frequency event study estimates and 95% confidence intervals across various bandwidth choices to demonstrate robustness of the main results. Panel (a) shows estimates for the natural logarithm of daily applications (bandwidths 2-150 days). Panel (b) shows estimates for application acceptance rates (bandwidths 2-150 days). Panel (c) shows estimates for default probability (bandwidths 10-150 days). For each panel, the high-frequency event study specification is estimated at bandwidths in increments of 1 day. The Y-axis reports the high-frequency event study estimate, and the X-axis reports the bandwidth used. All regressions include ZIP code × loan purpose and day of the week fixed effects, a local linear specification estimated separately on each side of the policy change date, and use a triangular kernel. The solid red line reports the high-frequency event study estimate, and the dashed grey lines indicate the 95% confidence bands calculated based on standard errors clustered at the zip-code level. The sample is restricted to Android users only, as iOS users were unaffected by the policy.

Table A.1: Robustness to Covariates: Effect of the Policy Change on Application Acceptances

Dep Var: Application Accepted	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Post	-0.4912***	-0.4920***	-0.2687***	-0.2749***	-0.2746***	-0.2743***	-0.2720***	-0.2688***	-0.2684***	-0.2715***	-0.2711***
	(0.0079)	(0.0091)	(0.0073)	(0.0093)	(0.0093)	(0.0093)	(0.0093)	(0.0093)	(0.0094)	(0.0094)	(0.0093)
Female (=1)						-0.0241***	-0.0297***	-0.0207**	-0.0233***	-0.0236***	-0.0213**
						(0.0085)	(0.0085)	(0.0083)	(0.0085)	(0.0084)	(0.0085)
Low Caste (=1)							-0.0639***	-0.0612***	-0.0604***	-0.0583***	-0.0553***
							(0.0055)	(0.0055)	(0.0056)	(0.0055)	(0.0056)
ZIP Code FE		✓	✓								
Loan Purpose FE			✓								
ZIP Code X Loan Purpose FE				✓	✓	✓	✓	✓	✓	✓	✓
Day of the Week FE				✓	✓	✓	✓	✓	✓	✓	✓
Education FE					✓	✓	✓	✓	✓	✓	✓
Age Decile FE								✓	✓	✓	✓
Credit Score Decile FE									✓	✓	✓
Loan Maturity FE										✓	✓
Income Decile FE											✓
Bandwidth	30	30	30	30	30	30	30	30	30	30	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	30,632	29,805	29,805	27,770	27,770	27,770	27,770	27,720	27,190	27,190	27190
R^2	0.2170	0.3431	0.6650	0.7465	0.7474	0.7475	0.7501	0.7525	0.7544	0.7561	0.7583
Adj. R^2	0.2169	0.2778	0.6316	0.6761	0.6772	0.6774	0.6807	0.6836	0.6855	0.6876	0.6902

This table presents the high-frequency event study estimates using the application acceptance status as the dependent variable. The dependent variable takes a value of one if the application was accepted and zero otherwise. We use a bandwidth of 30 days on either side of the policy change date (January 1, 2019). All regressions include different set of fixed effects, a local linear specification estimated separately on each side of the policy change date, and use a triangular kernel. The sample is restricted to Android users only, as iOS users were unaffected by the policy. Standard errors reported in parentheses are clustered at the zip-code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2: Effect of the Policy Change on Composition of Applications

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(Income)	ln(Age)	Credit Score	Low Caste	Female	New Borrower
Post	-0.0089	-0.0173***	0.6544	0.0312*	0.0140	-0.0490***
	(0.0208)	(0.0055)	(2.0006)	(0.0168)	(0.0103)	(0.0165)
ZIP Code X Loan Purpose FE	✓	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓	✓
Bandwidth	30	30	30	30	30	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	30,625	30,566	30,025	30,625	30,625	30,625
R^2	0.3606	0.3443	0.3110	0.3805	0.3016	0.3031
Adj. R^2	0.1940	0.1734	0.1300	0.2190	0.1196	0.1215

This table presents the high-frequency event study estimates indicating the effect of the policy change on the composition of the applications. Column 1 the natural logarithm of the monthly borrower income as the dependent variable. Column 2 uses the natural logarithm of the applicant's age as the dependent variable. Column 3 uses the applicants CIBIL credit score as the dependent variable. Column 4 uses a binary variable as a dependent variable that takes a value of one for applicants belonging to the scheduled caste, scheduled tribe and other backward class, and zero otherwise. Column 5 uses a binary variable as the dependent variable that takes a value of one for female applicants and zero otherwise. Column 6 uses a binary variable as the dependent variable that takes a value of one for new borrowers and zero otherwise. We use a bandwidth of 30 days on either side of the policy change date (January 1, 2019). All regressions include different set of fixed effects, a local linear specification estimated separately on each side of the policy change date, and use a triangular kernel. The sample is restricted to Android users only, as iOS users were unaffected by the policy. Standard errors reported in parentheses are clustered at the zip-code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Robustness to Covariates: Effect of the Policy Change on Default

Dep Var: Loan Default	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Post	-0.0085	-0.0162	-0.0153	-0.0072	-0.0072	-0.0071	-0.0071	-0.0055	-0.0067	-0.0076	-0.0024
	(0.0126)	(0.0145)	(0.0145)	(0.0178)	(0.0178)	(0.0178)	(0.0178)	(0.0177)	(0.0173)	(0.0163)	(0.0164)
Female (=1)						-0.0060	-0.0052	-0.0063	-0.0005	-0.0072	-0.0131
						(0.0148)	(0.0148)	(0.0151)	(0.0154)	(0.0148)	(0.0147)
Low Caste (=1)							0.0069	0.0066	0.0071	0.0116	0.0083
							(0.0091)	(0.0091)	(0.0093)	(0.0086)	(0.0086)
ZIP Code FE		✓	✓								
Loan Purpose FE			✓								
ZIP Code X Loan Purpose FE				✓	✓	✓	✓	✓	✓	✓	✓
Day of the Week FE				✓	✓	✓	✓	✓	✓	✓	✓
Education FE					✓	✓	✓	✓	✓	✓	✓
Age Decile FE								✓	✓	✓	✓
Credit Score Decile FE									✓	✓	✓
Loan Maturity FE										✓	✓
Income Decile FE											✓
Bandwidth	30	30	30	30	30	30	30	30	30	30	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	20,272	19,341	19,341	17,749	17,749	17,749	17,749	17,721	17,259	17,259	17,259
R^2	0.0009	0.1403	0.1432	0.2788	0.2790	0.2790	0.2791	0.2799	0.2927	0.3484	0.3599
Adj. R^2	0.0008	0.0351	0.0381	0.0770	0.0770	0.0770	0.0770	0.0774	0.0916	0.1628	0.1770

This table presents the high-frequency event study estimates using loan default status as the dependent variable. The dependent variable takes a value of one if the loan is in default and zero otherwise. We use a bandwidth of 30 days on either side of the policy change date (January 1, 2019). All regressions include Zip Code × loan purpose and the day of the week fixed effects, a local linear specification estimated separately on each side of the policy change date, and use a triangular kernel. The sample is restricted to Android users only, as iOS users were unaffected by the policy. Standard errors reported in parentheses are clustered at the zip-code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Strict Identification: Heterogeneity & Effect of the Policy Change on Application Acceptances

Dep Var: Application Accepted	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sub-Prime Credit Score × Post	-0.0075						-0.0088
	(0.0177)						(0.0175)
Sub-Prime Credit Score	-0.0026						-0.0051
	(0.0087)						(0.0085)
Low Caste × Post		-0.0857***					-0.0672***
		(0.0182)					(0.0179)
Low Caste		0.0006					-0.0007
		(0.0090)					(0.0088)
New Borrower × Post			-0.1063***				-0.1163***
			(0.0173)				(0.0176)
New Borrower			-0.0018				-0.0013
			(0.0090)				(0.0091)
Low Income × Post				-0.0704***			-0.0655***
				(0.0181)			(0.0181)
Low Income				-0.0032			-0.0054
				(0.0090)			(0.0089)
Young × Post					-0.0842***		-0.0767***
					(0.0165)		(0.0166)
Young					-0.0055		-0.0018
					(0.0088)		(0.0089)
Female × Post						-0.0331	-0.0216
						(0.0279)	(0.0264)
Female						-0.0139	-0.0138
						(0.0159)	(0.0156)
ZIP Code X Loan Purpose FE	✓	✓	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓	✓	✓
ZIP Code X Week FE	✓	✓	✓	✓	✓	✓	✓
Loan Purpose X Week FE	✓	✓	✓	✓	✓	✓	✓
Bandwidth	30	30	30	30	30	30	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	20,952	20,952	20,952	20,952	20,952	20,952	20,952
R^2	0.7905	0.7934	0.7945	0.7930	0.7931	0.7906	0.8022
Adj. R^2	0.6470	0.6519	0.6537	0.6512	0.6513	0.6472	0.6661

This table presents the high-frequency event study estimates using the application acceptance status as the dependent variable. The dependent variable takes a value of one if the application was accepted and zero otherwise. All regressions use a bandwidth of 15 days on either side of the policy change date (January 1, 2019). All regressions include Zip Code × loan reason and the day of the week fixed effects, a local linear specification estimated separately on each side of the policy change date, and use a triangular kernel. Standard errors reported in parentheses are clustered at the zip-code level. Sub-prime credit score is a binary variable that takes a value of one for borrowers with CIBIL credit score less than 680, and zero otherwise. Low caste is a binary variable that takes a value of one for applicants belonging to the scheduled caste, scheduled tribe and other backward class, and zero otherwise. New borrower is a binary variable for borrowers who are first time applicants, and zero otherwise. Low income is a binary variable that takes a value of one for borrowers with monthly income less than the sample median value, and zero otherwise. Young is a binary variable that takes a value of one for borrowers with age less than the sample median value, and zero otherwise. Female is a binary variable that takes a value of one for females, and zero otherwise. The sample is restricted to Android users only, as iOS users were unaffected by the policy. Standard errors reported in parentheses are clustered at the zip-code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Credit Access: Reduced Form Results using High-Frequency Identification

	2 Years			3 Years			4 Years		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post	-0.0556***	-0.0484***	-0.0480***	-0.0601***	-0.0511***	-0.0519***	-0.0560***	-0.0464***	-0.0468***
	(0.0060)	(0.0059)	(0.0059)	(0.0061)	(0.0061)	(0.0061)	(0.0062)	(0.0062)	(0.0062)
ZIP Code FE		✓	✓		✓	✓		✓	✓
Day of the Week FE			✓			✓			✓
N	27,770	27,770	27,770	27,770	27,770	27,770	27,770	27,770	27,770
R^2	0.0029	0.1784	0.1787	0.0032	0.1753	0.1758	0.0027	0.1725	0.1730
Adj. R^2	0.0028	0.0963	0.0966	0.0031	0.0930	0.0933	0.0026	0.0899	0.0903

This table compares the effect of the policy change on credit access for those who applied after the policy compared to those who applied before. The dependent variables are indicators for whether a borrower accessed credit within two years (columns 1-3), three years (columns 4-6), or four years (columns 7-9) following January 1, 2019. The key independent variable is *Post*, which indicates whether the application was submitted after the policy. All regressions use a sample restricted to observations within 30 days of the policy cutoff . Column specifications progressively add fixed effects: no fixed effects (columns 1, 4, 7), Zip Code fixed effects (columns 2, 5, 8), and Zip Code plus day-of-week fixed effects (columns 3, 6, 9). The sample is restricted to Android users only, as iOS users were unaffected by the policy. Standard errors are clustered at the zip-code level. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.6: Credit Access: Instrumental Variables Results using High-Frequency Identification

	2 Years			3 Years			4 Years		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
App Accepted	0.1138***	0.1129***	0.1125***	0.1229***	0.1193***	0.1217***	0.1145***	0.1083***	0.1098***
	(0.0122)	(0.0138)	(0.0138)	(0.0125)	(0.0141)	(0.0143)	(0.0126)	(0.0144)	(0.0145)
ZIP Code FE		✓	✓		✓	✓		✓	✓
Day of the Week FE			✓			✓			✓
N	27,770	27,770	27,770	27,770	27,770	27,770	27,770	27,770	27,770
KP LM Statistic	560.5439	387.6932	388.3385	560.5439	387.6932	388.3385	560.5439	387.6932	388.3385
KP F Statistic	11863.4999	8591.9718	8029.3159	11863.4999	8591.9718	8029.3159	11863.4999	8591.9718	8029.3159

This table presents instrumental variables estimates of the effect of loan application acceptance on future credit access. The dependent variables are indicators for whether a borrower accessed credit within two years (columns 1-3), three years (columns 4-6), or four years (columns 7-9) following January 1, 2019. The endogenous variable *App Accepted* is instrumented using the post-policy indicator. All regressions use a sample restricted to observations within 30 days of the policy cutoff. Column specifications progressively add fixed effects: no fixed effects (columns 1, 4, 7), ZIP Code fixed effects (columns 2, 5, 8), and ZIP Code plus day-of-week fixed effects (columns 3, 6, 9). The sample is restricted to Android users only, as iOS users were unaffected by the policy. Standard errors are clustered at the zip-code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Summary Statistics of Profit Margins by Borrower Characteristics (EMI Model)

Panel A: Profit Margin by Borrower Group

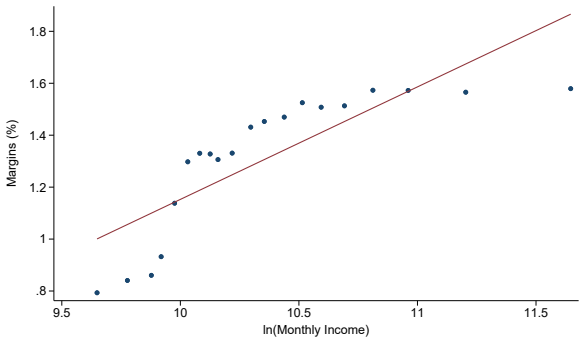
Characteristic	N	25th pct	Median	75th pct	Mean	Std Dev
Low Income	14,217	0.5899	1.0204	1.4142	1.1409	0.8420
Female	2,152	0.6515	1.2204	2.4370	1.4366	0.9634
Young Age	12,766	0.6271	1.0908	2.1830	1.2660	0.9158
Low Caste	13,564	0.5771	1.0632	2.2732	1.2758	0.9468
Low Amount	15,247	0.3446	0.7181	1.0154	0.6553	0.4720
Low Credit Score	11,755	0.5954	1.0931	2.3710	1.3383	0.9709

Panel B: Average Group Margins

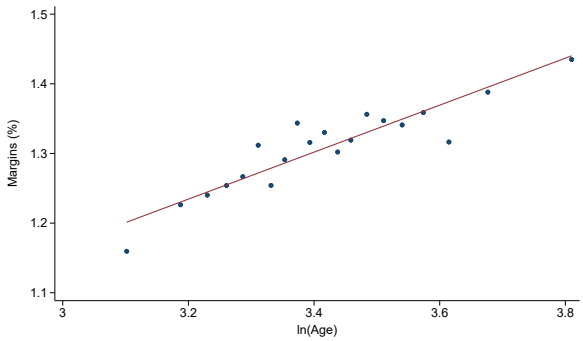
Characteristic	Average Margin (%)
Low Income	1.7747
Female	2.1726
Young Age	2.0114
Low Caste	2.0626
Low Amount	0.6911
Low Credit Score	2.1359

Notes: This table presents summary statistics for expected profit margins (as percentage of loan amount) across different borrower characteristics for marginal borrowers under EMI loan structure. Marginal borrowers are defined as those rejected post-policy but who would have been approved pre-policy based on predicted approval probability > 0.5. Low income, young age, low amount, and low credit score are defined as below-median values in the sample. Low caste includes scheduled castes, scheduled tribes, and other backward classes. Female is an indicator for female borrowers. Profit calculations assume EMI structure with monthly payments, 19% baseline default rate, 9% loss given default, and 1% monthly discount rate. Sample restricted to 60-day window around January 1, 2019 policy implementation. Groups may overlap across categories. Panel A shows individual-level distribution statistics, while Panel B shows dollar-weighted group averages.

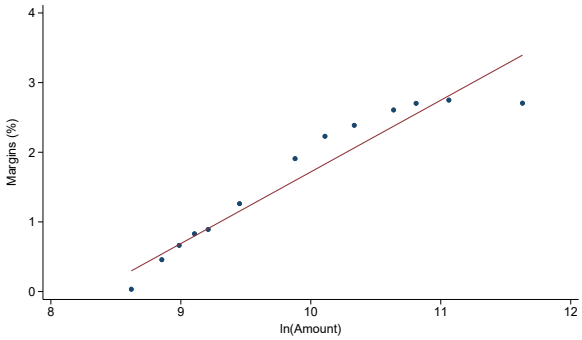
Figure A.2: Profit Margins of Marginal Borrowers by Characteristics (EMI Model)



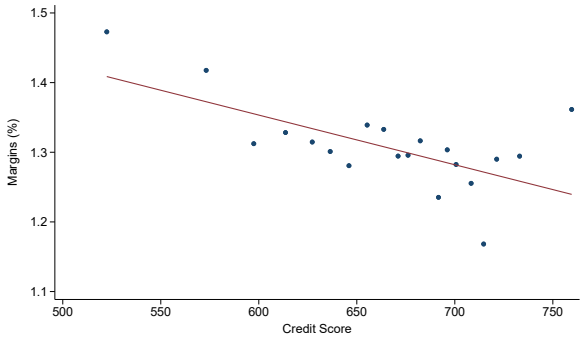
(a) By Income



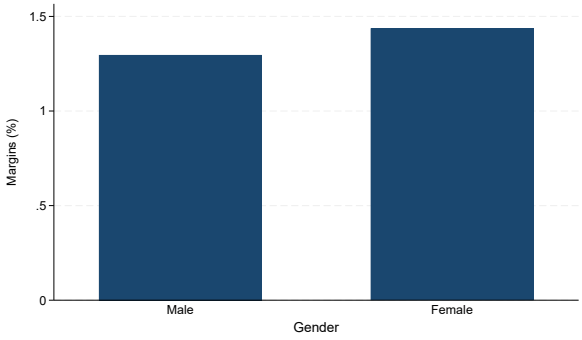
(b) By Age



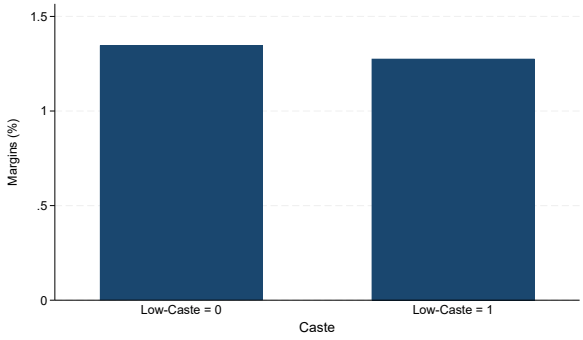
(c) By Loan Amount



(d) By Credit Score



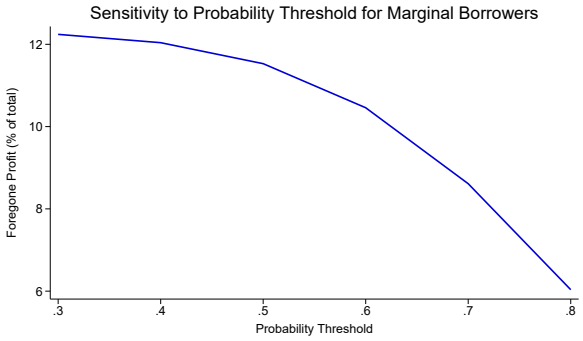
(e) By Gender



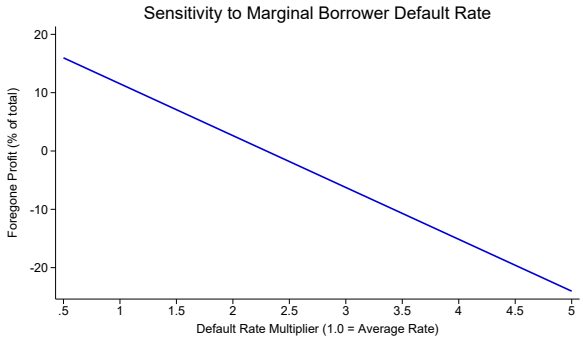
(f) By Caste

Notes: This figure shows the expected profit margins (as percentage of loan amount) for marginal borrowers under EMI loan structure. Marginal borrowers are those rejected post-policy but who would have been approved pre-policy based on predicted approval probability > 0.5. Panel A.2a shows margins by log monthly income with quadratic fitted trend. Panel A.2b shows margins by log age with linear fitted trend. A.2c shows margins by log loan amount displaying a concave relationship. Panel A.2d shows margins by credit score with a declining trend line. Panel A.2e and Panel A.2f show average margins by gender and caste categories respectively. Profit calculations assume EMI structure with monthly payments, 19% baseline default rate, 9% loss given default, and 1% monthly discount rate. Sample includes marginal borrowers within 60-day window around January 1, 2019 policy implementation.

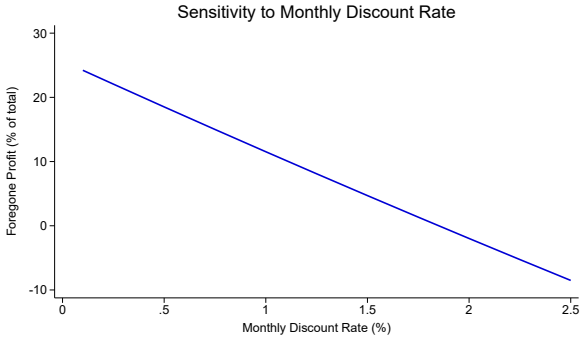
Figure A.3: Sensitivity Analysis of Foregone Profits (EMI Model)



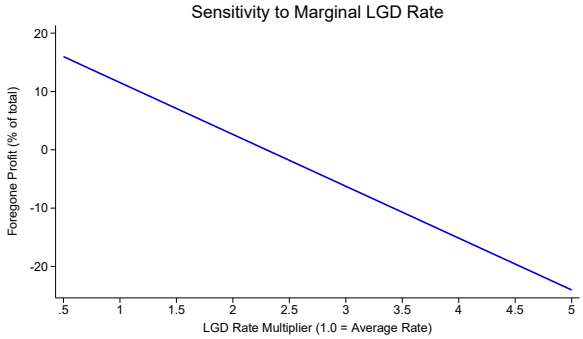
(a) Probability Threshold



(b) Default Rate



(c) Monthly Discount Rate



(d) Loss Given Default

Notes: This figure presents sensitivity analysis for the foregone profit calculations under EMI loan structure. Panel A.3a shows how results vary with the probability threshold used to define marginal borrowers (those rejected post-policy but who would have been approved pre-policy). Panel A.3b examines sensitivity to assumptions about marginal borrowers' default rates, expressed as multiples of the baseline 19% rate. Panel A.3c shows how results depend on the monthly discount rate used in present value calculations. Panel A.3d examines sensitivity to the loss given default (LGD) rate, expressed as multiples of the baseline 9% rate. All calculations assume EMI structure with monthly payments. Baseline parameters: 0.5 probability threshold, 19% default rate, 1% monthly discount rate, 9% LGD. Sample includes borrowers within 60-day window around January 1, 2019 policy implementation.