When Privacy Protects but Excludes: The Costs and Benefits of Privacy Regulation in Credit Markets*

Sumit Agarwal Pulak Ghosh Peiyi Jin

Shohini Kundu Nishant Vats Xinbo Wang

Yingze Xu

Abstract

This paper studies the consequences of privacy regulation by exploiting Google's 2019 restriction on CDR access for a major Indian FinTech lender. We show that this intervention reflects a key policy trade-off in digital credit markets: strengthened privacy protections raise loan applications, consistent with higher demand, yet simultaneously induce tighter screening, reflecting an overall contraction in credit supply. This credit contraction disproportionately excludes economically and socially marginalized applicants. Linking to economy-wide credit bureau records, we quantify the "FinTech ladder effect" whereby initial digital credit access serves as a gateway to broader formal credit. Privacy-induced rejection reduces the probability of obtaining any formal credit by 13.7 percentage points even four years later. Using a structural model, we decompose the welfare effects of privacy regulation and show that the regulation generates a 0.53% increase in consumer surplus and reduces lender profits by 15%.

*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); Xu: Olin Business School, Washington University at St Louis (x.yingze@wustl.edu).

We thank Giorgia Barboni, Kim Fe Cramer, Alexander Copestake, Andrea Eisfeldt, Sean Higgins, Nicola Limodio, Muhammad Meki, Kevin Mei, Tyler Muir, Shumiao Ouyang, Stavros Panageas, Michaela Pagel, Amiyatosh Purnanandam, Avanidhar Subrahmanyam, Huan Tang, and Paul Yoo, as well as seminar and conference participants at UCLA Anderson, the Midwest Finance Association, and the CEPR-WEFIDEV Conference, for their 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 manage their portfolio risk. Many lenders use alternative data to assess borrower risk and screen loan applicants at origination.[1] However, the ability of alternative data to accurately classify borrower risk is often limited, particularly for economically and socially disadvantaged individuals (Blattner and Nelson, 2021). This limitation is especially salient for FinTech lenders that seek to serve segments historically excluded from traditional banking (Alok et al., 2024; Cramer et al., 2025). In response, many FinTech lenders, particularly in emerging economies where traditional credit infrastructure remains underdeveloped, have extended their use of alternative data beyond initial screening to encompass ongoing monitoring and enforcement throughout the loan lifecycles. These lenders leverage diverse data, such as call detail records (CDRs), mobile phone usage patterns, and other digital behavioral traces to create implicit social collateral that reduces default risk (see New York Times, 2017; Wall Street Journal, 2019; Bloomberg, 2023). These innovations using sensitive and personal data promise to advance financial inclusion for low-income and marginalized populations who have historically been excluded from formal credit markets.

Yet, the increasing reliance on sensitive and 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 et al., 2023; Armantier et al., 2024). However, these studies largely overlook the potential tradeoffs: while privacy regulations increase data protection, they may also constrain lenders' ability to make profitable loans, 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. Additionally, these dynamics may affect the comparative advantage of FinTechs relative to traditional banks (Cramer et al., 2024).

This paper examines the tradeoff between consumer protection and financial inclusion, and its implications for lender profitability, leveraging a policy shock that curtails 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' CDR following a Google privacy policy directive. Prior to this change, CDR data

[1] See Frost et al. (2019); Thakor (2020); Allen, Gu, and Jagtiani (2021); Berg, Fuster, and Puri (2022) and Doerr et al. (2023a) for a review of this literature.

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 that improves repayment discipline and keeps default rates low (Karlan et al., 2009; Bryan, Karlan, and Zinman, 2015). The policy change 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 application-level data spanning 2017-2022, we exploit the platform-specific nature of Google's policy to identify the effect of privacy regulation on borrower and lender behavior and its distributional consequences across borrower groups.

We employ two complementary empirical strategies to identify the effects of the privacy regulation. Our primary approach is a difference-in-differences (DiD) design that compares Android users, who were directly affected by the policy, with iOS users, who serve as a control group around the implementation date. This design leverages the fact that both groups borrowed from the same lender under identical product and pricing conditions prior to the policy change. The identification does not require the two groups to be identical ex ante. Instead, it hinges on the parallel trends assumption that, absent the policy, their lending outcomes would have evolved similarly. We assess this identifying assumption along two dimensions. First, we show that average lending outcomes exhibit parallel dynamics between the two groups during the pre-regulation period. Second, we confirm that the policy's implementation did not differentially affect the composition of applications across platforms. To further reinforce identification, we demonstrate that these pre-policy trends remain indistinguishable even after saturating the specification with a rich set of fixed effects, including ZIP code × loan-purpose × month-year and ZIP code × loan-purpose × Android fixed effects.

As a complementary analysis, we implement a high-frequency event study, exploiting daily data for Android users around the policy implementation date. Conceptually akin to a regression discontinuity design in time, this specification focuses on the immediate adjustment in lending outcomes when access to CDRs was abruptly terminated on January 1, 2019. The event study supplements the DiD analysis in three important ways. First, by relying on within-Android variation over a narrow window, it isolates the short-run behavioral response of both lenders and borrowers to the information shock. Second, it is identified under a distinct assumption: continuity of potential outcomes around the policy cutoff, which allows the detection of discrete jumps in credit outcomes without using iOS users as a benchmark. Third, because the comparison is within-platform and temporally localized, the design alleviates concerns that unobserved cross-platform heterogeneity biases the difference-in-differences estimates. Collectively, these two approaches provide a consistent framework for quantifying the timing, magnitude, and persistence of the privacy regulation's impact on credit market outcomes.

To conduct our analysis, we construct a novel dataset that combines granular application-level records from one of India's largest digital lenders with economy-wide credit bureau data, covering the universe of formal borrowing. The proprietary lender data contain every loan application submitted between 2017 and 2022, including the borrower's operating system, detailed contract terms, approval status, and repayment outcomes, alongside a rich set of demographic characteristics such as income, age, gender, caste, education, and bureau credit scores. Importantly, each application record contains the operating system used to submit it, providing a clean treatment indicator at the individual level. We complement this with a linkage to the national credit bureau, which tracks all formal borrowing from banks and non-bank financial companies across India. This linkage allows us to track each applicant's access to formal credit over multi-year horizons, regardless of whether subsequent borrowing originates from the focal lender or elsewhere in the financial system, enabling us to distinguish persistent exclusion from simple reallocation across lenders and to quantify the longer-term "FinTech ladder effect."

We begin by examining the responses of both borrowers and the lender to the introduction of the privacy regulation. On the borrower side, we observe a sizable increase in credit application activity among Android users. Within the same ZIP code, loan purpose, and time period, applications by Android users rises by approximately 26% following the reform. This increase suggests a re-optimization of borrowers' participation decisions once data-sharing requirements were relaxed.

On the lender side, we find a significant decline in the likelihood of application acceptance after the regulation. Acceptance rates among Android applicants decreases by 16–18 percentage points, corresponding to roughly a 25% decline relative to the sample mean. Importantly, this reduction is not explained by compositional changes in applicant quality, as income and credit score distribution remain stable after the change. Other contract terms, such as interest rates and maturities, show no economically meaningful adjustments. Taken together, the small price adjustments and the absence of maturity effects indicate that credit markets does not materially reprice or re-term loans in response to the privacy regulation. Instead, the main margin of adjustment operates through quantities, i.e., application volumes and approval probabilities, rather than through contract terms for approved borrowers.

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. This enforcement technology allowed the lender to extend credit to borrowers who would have been rejected otherwise, 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 CDRs, the lender becomes 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, are now rejected.

The surge in loan applications following the policy reform is most consistent with an increase in credit demand among borrowers who attach positive value to privacy. Prior to the regulation, the requirement to share CDRs imposed an implicit non-pecuniary cost on potential borrowers, particularly those concerned about data access and surveillance. Once this requirement was removed, the effective "privacy cost" of applying declines, making participation in formal credit markets more attractive for such individuals.

Conceptually, application volumes represent an equilibrium outcome reflecting both the demand for credit and expectations of credit supply. While enhanced privacy should encourage more borrowers to apply, a simultaneous tightening of screening standards may discourage applications from those anticipating higher rejection probabilities. The observation that total applications, nevertheless, rise after the policy suggests that the privacy-induced increase in demand more than offsets the supply-side deterrent. In this sense, the observed increase provides a lower-bound estimate of the underlying shift in credit demand attributable to privacy protection.

This interpretation is consistent with a growing literature showing that borrowers place a measurable and economically meaningful value on privacy in financial interactions (Tang, 2019; Lin, 2022; Armantier et al., 2024). The magnitude of the application response reinforces the view that data-sharing requirements constitute a substantive friction in household credit decisions, influencing the extent of financial participation.

Next, we examine the distributional consequences of the policy change. Understanding these heterogeneous effects is important to assess broader welfare implications of the privacy policy, i.e., who bears the cost of privacy regulation. To this end, we examine the heterogeneity in the effect of privacy regulation on the likelihood of acceptance across borrowers. We hypothesize that the adjustment from ex-post enforcement to stricter ex-ante screening, induced by the privacy regulation, is likely to shift the composition of approved borrowers. In particular, we posit that the screening disproportionately tightens against applicants whose risk cannot be credibly verified using conventional data.

We document that low-income and younger applicants, first-time borrowers, and those belonging to marginalized social groups, such as scheduled castes, scheduled tribes, and other backward classes, experience the largest post-policy decline in acceptance rates. By contrast, applicants with established credit histories or pre-existing lender relationships appear largely insulated from these effects. Taken together, these findings indicate that constraining the use of alternative enforcement tools, such as CDRs, tightens credit supply primarily for borrowers on the financial margins. The privacy regulation, thus, reinforces exclusion within the digital lending market, highlighting a key trade-off in modern credit intermediation. Policies that safeguard data privacy may weaken financial inclusion by impairing lenders' ability to extend credit to the informationally opaque.

Thus far, we have shown that the privacy policy led to an immediate contraction in credit supply at our focal FinTech. We now ask whether affected borrowers substitute toward other lenders or, instead, face persistent exclusion from formal credit markets. To study these long-run effects, we link our primary loan-level data to comprehensive credit bureau records that track formal borrowing from the universe of financial institutions in India, allowing us to follow applicants' credit access over several years. To our knowledge, this is the first study to trace the long-run credit market consequences of privacy regulation using economy-wide bureau data.

This analysis allows us to address an important question: does privacy regulation merely reallocate borrowing across lenders, or does it permanently push some borrowers out of the formal credit system? If applicants rejected by the FinTech can readily obtain loans from alternative institutions, the welfare consequences of the policy may be limited. By contrast, if obtaining a FinTech loan improves access to traditional bank credit, by generating additional information (Balyuk, 2023; Dubey, 2024) or helping borrowers build collateral (Beaumont, Tang, and Vansteenberghe, 2022), then rejection induced by privacy regulation may translate into systematic exclusion from the formal credit markets.

We find that applicants whose loans were rejected following the privacy regulation are approximately 2.2 percentage points less likely to obtain any formal credit, even four years later. This effect is economically meaningful, representing about 5.2% of the sample mean. Using a two-stage least squares framework, we show that loan approval by our focal FinTech increases the probability of subsequent borrowing from formal credit markets by 13.7 percentage points. These estimates imply that initial access to digital credit plays a critical role in integrating marginal borrowers into the formal financial system. Conversely, denial of such access can lock applicants into persistent exclusion, consistent with a FinTech ladder effect, wherein early exposure to alternative credit platforms serves as a stepping-stone to mainstream finance. By quantifying this dynamic, our analysis contributes new empirical evidence on how initial interactions with digital lenders shape long-term financial inclusion.

Lastly, we develop a structural model to quantify the effects of privacy regulation on borrower surplus and lender profitability in a digital credit market where CDRs previously served as a key enforcement technology. While the reduced-form evidence is informative about the direction of the demand- and supply-side responses to the privacy regulation, the model allows us go beyond and separately identify three policy-relevant objects: (i) how much borrowers value the additional privacy created by the removal of CDR access, (ii) whether this privacy benefit is large enough to offset the indirect cost of reduced credit access and what the net effect on consumer surplus is, and (iii) how much profit the lender loses when CDR-based social collateral is removed, and through which margins – applications, approvals, or per-loan surplus – these effects operate.

Economically, the model has three stages: borrowers decide whether to apply (demand), the FinTech decides whom to approve (screening), and approved borrowers either repay or default (repayment and recovery). Privacy enters the borrower's problem as a direct utility gain from not sharing sensitive communication data, in addition to the standard consumption-smoothing benefits of credit. On the lender side, CDRs raise expected returns by lowering both default probabilities and loss-given-default via social pressure and targeted collections, so losing CDRs forces the lender to tighten screening for informationally opaque borrowers even though prices and maturities change little.

Identification hinges on a key tension: absent any value for privacy, the stricter screening induced by the CDR ban should reduce application volumes, yet we observe applications increasing. The model uses this discrepancy to recover borrowers' latent valuation of privacy that rationalizes higher demand despite lower approval odds. We estimate that the direct utility gain from enhanced privacy is equivalent to 14.3% of monthly income, implying a substantial willingness to pay. Specifically, borrowers are willing to forgo more than one-seventh of a month's income to keep their call records private from the lender.

The model allows us to quantify the net effect of the policy on consumer surplus and decompose the change in consumer surplus into three components. Overall, the privacy regulation raises borrower welfare by 0.53%. The direct privacy channel contributes +0.54%, reflecting the utility gain from removing lender access to CDRs, which accrues to all potential borrowers equally. The application margin adds +0.02%, as privacy encourages some additional borrowers to apply and obtain a positive surplus from credit. This is an indirect effect of the privacy channel. The approval margin contributes -0.03%, as tighter screening reduces approval probabilities and lowers expected surplus for applicants. Together, these results imply that the privacy benefits outweigh the welfare losses from reduced credit access.

Finally, the model shows that privacy regulation reduces total lender profits by 15.0%, and we decompose this loss into three channels. The application margin contributes +4.8% because higher application rates expand the pool of potential borrowers. The approval margin contributes -8.4%, as the lender rejects more applicants to maintain portfolio quality once CDR-based enforcement is unavailable. The surplus margin contributes -11.5%, reflecting lower surplus per approved loan when the lender can no longer rely on CDR-based social enforcement to contain defaults and losses.

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 has discussed these tradeoffs, as reviewed by Acquisti, Taylor, and Wagman (2016), Bergemann and Morris (2019); Goldfarb and Tucker (2019) and Goldfarb and Que (2023), 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 et al., 2023; Armantier et al., 2024; Chen et al., 2025). A notable exception is Doerr et al. (2023b), which shows that privacy restrictions decreased FinTech lending. We contribute to this literature in two key ways. First, using loan-level data, we document the tradeoffs associated with privacy regulation in credit markets. On the benefit side, we document an increase in the number of applications consistent with the idea that consumers place a positive value on privacy. On the cost side, we document a reduction in credit supply, primarily for socially and economically marginalized groups and first-time applicants, impeding long-term credit access. Taken together, our empirical evidence mirrors the fundamental trade-off consumers face when sharing sensitive data, as highlighted in the theoretical work of Agur, Ari, and Dell'Ariccia (2025), in the context of payment systems. Second, we develop a structural model to quantify the consequences of privacy regulation on consumer surplus. The model implies that privacy regulation raises consumer surplus through two channels: a direct utility gain from enhanced privacy, and an indirect gain via a higher propensity to apply for loans that deepens market participation. At the same time, stricter privacy reduces consumer surplus by increasing rejection rates, thereby weakening households' ability to smooth demand shocks. On net, the estimated benefits of privacy outweigh these costs. Lender profits, however, decline, driven primarily by the loss of CDR-based enforcement rents through the approval and surplus margins and only partially offset by higher application volumes.

The second contribution of this paper is to provide the first evidence tracing the long-run credit market consequences of privacy regulation using economy-wide bureau data. 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).[2] A key challenge in the literature is that identifying the FinTech ladder effect requires both a comprehensive universe of credit data and an exogenous shock to FinTech access. Our analysis is uniquely positioned to address both. We exploit a privacy-induced regulatory shock that generates exogenous variation in FinTech adoption by restricting or removing access to alternative data. This empirical setting allows us to trace the trajectories of borrowers that are either accepted or rejected by FinTech. Moreover, unlike prior studies that focus narrowly on specific market segments – such as mortgage lending or secured small-business credit – our dataset, drawn from a comprehensive credit bureau, covers the entire universe of formal credit in the economy. This includes both secured and unsecured loans across multiple product categories, enabling a granular analysis of allowing us to capture the long-term effects of initial credit access from FinTechs.

Therefore, our paper overcomes the two central empirical challenges that have limited prior research on the FinTech ladder effect: identifying exogenous variation in FinTech loan approval and observing borrowers' complete credit market trajectories across all loan products and lender types. Leveraging a privacy-induced regulatory shock together with economy-wide bureau data, we provide some of the first quantitative evidence on the magnitude of the FinTech ladder effect. We show that approval by a FinTech lender increases the probability of subsequent borrowing in formal credit markets by 13.7 percentage points. This economically large effect highlights the pivotal role of initial digital credit access in integrating marginal borrowers into the formal financial system and demonstrates how FinTech lending can serve as a gateway to broader financial inclusion.

The third contribution of this paper is that we quantify the value of alternative data to firms by documenting the effect of privacy regulation on lender profitability. Existing studies have primarily focused on the adverse consequences of privacy regulation for firms' online performance, such as declines in web traffic and application downloads (e.g., Goldberg, Johnson, and Shriver, 2019; Aridor, Che, and Salz, 2020; Bessen et al., 2020; Bian, Ma, and Tang, 2021; Jia, Jin, and Wagman, 2021; Janßen et al., 2022; Kraft, Skiera, and Koschella, 2023), or on firms' subsequent adjustments to data-collection and sharing practices (Peukert et al., 2022; Bae, Mayya, and Nian, 2023; Kesler, 2023; Demirer et al., 2024; Ramadorai, Uettwiller, and Walther, 2025). A notable exception is Abis, Tang, and Bian (2025), who document the

[2]In contrast, Chava et al. (2021) finds that marketplace lending can impede future credit access. The difference between our results and theirs can be attributed to the fact that their sample consists primarily of borrowers with existing credit histories seeking debt consolidation, whereas our lender serves predominantly first-time applicants. For such borrowers, perhaps initial FinTech credit access appears to facilitate future access to traditional bank credit.

detrimental impact of privacy regulation on market efficiency. We contribute to this emerging literature by providing novel evidence that privacy regulation significantly reduces the profitability of FinTech lenders. Employing a structural model, we decompose the total effect of privacy regulation into distinct channels. While the policy raises profits by 32.1% through an increase in loan applications, it reduces profitability by 55.8% and 76.4% through declines in approval rates and the available surplus extractable by the lender, respectively. These findings reinforce the idea that alternative data generate substantial firm value, as formalized by Veldkamp and Chung (2024). Our approach also parallels the revenue-based frameworks for valuing data proposed by Veldkamp (2023) and Farboodi et al. (2025).

We also contribute to the literature on information restrictions in credit markets. This literature studies how the removal of negative credit information, such as bankruptcy flags, medical debt, or delinquency records, affects subsequent borrowing and labor market trajectories (Musto, 2004; Bos, Breza, and Liberman, 2018; Dobbie et al., 2020; Herkenhoff, Phillips, and Cohen-Cole, 2021; Duarte et al., 2025). Our empirical setting differs in a key respect: we focus not on traditional forms of information or data but on *alternative* data, which has recently become central to the evolution of the fintech sector. As documented by Cramer et al. (2024), the use of such alternative data constitutes a key source of the FinTech sector's comparative advantage in unsecured credit markets. This segment is of considerable policy interest, as it represents a rapidly expanding component of financial intermediation worldwide, including, but not limited to, emerging market economies. Our primary contribution is to examine how recent policy debates and regulatory initiatives surrounding data privacy, particularly those aimed at restricting or redefining access to alternative data, affect FinTech firms' ability to operate efficiently in these markets and, by extension, influence competition and credit availability.

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. Section 6 presents a structural model to quantify how privacy regulation affects both borrowers' welfare and lender profitability. Section 7 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 the effects of privacy regulation on financial inclusion, as this lender was among the first to employ alternative data to enforce credit contracts and improve loan repayment.

2.1 Business Model and Target Market

The FinTech operates exclusively 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 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 (Fuster et al., 2019; Cramer et al., 2024). 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 enabled the lender to construct detailed social network graphs for each borrower. Contacts were classified by relationship type, including immediate family (mother, father, spouse, siblings), extended family (aunts, uncles, grandparents), professional contacts (work colleagues), and close friends. This structure allowed the lender either to reach out to these contacts in the event of default or to leverage the threat of such contact as an enforcement device. In effect, the lender treated the borrower's social capital as a form of collateral to support contract enforcement and promote timely repayment.

Second, communication pattern analysis generated dynamic information on borrower behavior around loan-related events (Agarwal et al., 2020). The lender monitored multiple dimensions of phone use, such as changes in call frequency, shifts in the timing of calls, the composition of missed versus completed calls, and communication behavior around repayment dates, as potential signals of borrower circumstances. Deviations from usual patterns could indicate financial stress, social isolation, or efforts to evade contact from collection agents. 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, prioritizing family members with frequent, long-duration calls as potential sources of repayment assistance or leveraging professional contacts to create workplace pressure. Importantly, the enforcement value of this system did not require actual contact to materialize — the mere knowledge that the lender held detailed information about a borrower's entire social network was itself a disciplining force.

Third, ongoing monitoring transformed these communication signals into a dynamic, real-time risk management tool. Rather than relying solely on ex post delinquency or default, the lender continuously refreshed communication metrics throughout the loan lifecycle. Persistent declines in network size or diversity, rising concentration of calls to a small subset of contacts, or sustained avoidance of calls from the lender triggered internal alerts and escalations. The lender used these alerts to initiate a range of targeted responses, such as early engagement by customer service, restructuring offers, or intensified collection efforts, even before a scheduled payment was missed. In this way, the monitoring capability enabled the lender to shift from reactive collection based on realized default to proactive intervention based on predicted default risk, potentially lowering losses while also altering the nature and timing of borrower–lender interactions.

2.2.2 Empirical Evidence of Predictive Power and Role of Social Collateral

We complement this discussion by empirically establishing the importance of CDRs, Table 2 presents the relationship between contact list size, default probability, and loss-given-default 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 purpose in the same month-year. This approach ensures that the relationship is not driven by geographic, need-specific, or temporal factors that might confound the analysis.

The results in Panel A 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 7.03 percentage point reduction in default probability in the baseline specification. This effect attenuates to 5.97 percentage points when demographic and loan-level controls are included, representing a 29% decline relative to the average default probability. The model's discriminatory power is substantial: the AUC rises from 0.52 in the baseline specification to 0.92 with the full set of fixed effects and demographic controls in column 5.

If CDRs serve primarily as an enforcement device, their value should extend beyond reducing the probability of default to also lowering losses *conditional* on default occurring. Panel B of Table 2 provides evidence consistent with this prediction. Estimated on the subsample of defaulted loans, a 1% increase in contact network size is associated with a 4.04–5.22 percentage point reduction in loss given default, representing a 9–12% decline relative to the mean LGD of 44%. This result indicates that the lender's ability to leverage borrowers' social networks facilitated recovery even after delinquency had occurred, reinforcing the interpretation that CDRs functioned as a post-disbursement enforcement tool.

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 et al., 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. While, in theory, these interpersonal contacts could have provided the lender with additional information about a borrower's creditworthiness, our discussions with the institution suggest that such information was not incorporated into their ex-ante screening or borrower selection process.

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. Importantly, 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 app from the Google Play Store. 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.

The impact of this change was substantial for Android users. The lender lost access to what had been their most sophisticated monitoring 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.

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 affects all Android loan applications processed after the implementation date.

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 effect of privacy regulation on credit access.

3 Data

We leverage a unique and large-scale application-level proprietary dataset from one of India's largest digital lenders, covering the period 2017–2022. The data span the full pre- and post-January 2019 window around Google's Android privacy directive that removed CDR access for our focal FinTech but left iOS operations unaffected, providing a natural platform-level treatment and control structure.

We observe every loan application submitted to the lender, with unique borrower and loan identifiers, timestamps, detailed contract terms, such as loan type, requested amount, interest rate, maturity, and approval status. We also observe information on default and loss given default for approved loans. We can also observe the operating system or platform used to apply, allowing us to distinguish between treated (Android) and control (iOS) users in our difference-in-differences analysis. The application records also provide information on a rich set of borrower-level characteristics, such as income, age, gender, caste, education, and bureau credit scores, which permit a granular characterization of the borrower pool and a systematic analysis of the heterogeneous effects of the privacy policy across applicants.

Credit Bureau Data: To move beyond the perspective of a single-lender, we link the application-level data to the national credit bureau, which records the universe of formal borrowing from banks and non-bank financial corporations, including both FinTech and non-tech lenders.[3] This linkage allows us to track each applicant's subsequent access to formal credit over multi-year horizons, regardless of whether future loans originate from the focal FinTech or from other institutions. This integrated dataset serves two purposes. First, it allows us to distinguish privacy-induced reallocation of borrowing across lenders from genuine exclusion from formal credit markets. Second, it enables us to quantify the longer-term FinTech ladder effect, i.e., how an initial approval at the digital lender shapes borrowers' future access to credit.

Taken together, the combination of high-frequency application data with rich demographics and platform identifiers and economy-wide credit bureau coverage provides a sharp empirical foundation to study the propagation of the privacy shock. It allows us to examine how the loss of CDRs reshapes lending decisions at the directly affected FinTech and to trace whether these adjustments translate into persistent financial exclusion or merely a redistribution of borrowing across the formal credit ecosystem.

Summary Statistics: Table 1 presents summary statistics for the sample of loan applications submitted within 180 days of the January 1, 2019 policy change.

Panel A describes the characteristics of loan applicants. The typical applicant is around 31 years old, earns a monthly income of roughly ₹43,000, and holds an Equifax credit score of approximately 659. The Equifax credit score is a bureau-level measure of creditworthiness analogous to FICO scores in the United States, ranging from 300 to 900, with higher scores indicating lower default risk. Women comprise about 10% of applicants. Low caste is an indicator for applicants belonging to historically disadvantaged groups, specifically, Scheduled Castes, Scheduled Tribes, and Other Backward Classes, as

[3]We direct readers to Mishra, Prabhala, and Rajan (2022) and Cramer et al. (2024) for a discussion on the Indian credit bureau data.

predicted from surname-based classification following Mishra, Prabhala, and Rajan (2022).[4] Nearly 47% of applicants belong to lower caste groups. College education is an indicator equal to one for applicants with an undergraduate or postgraduate degree. Approximately 86% of the sample holds at least a college education. First-time applicant is an indicator equal to one for applicants who have not previously had a loan approved by the platform. About 36% of applicants are new to the platform. We identify Android users as all applicants who did not submit their application through an iOS device, as recorded in the platform's operating system field. Android users account for roughly 92% of applications, consistent with Android's dominant market share in India. Appendix Figure A.1 shows the distribution of loan purposes across applications, while Appendix Figure A.2 displays the distribution of applicant education levels.

Panel B reports application-level characteristics. The ln(Applications) variable is constructed by collapsing the application data to the ZIP code × loan purpose × month-year × platform level and taking the natural logarithm of the count, yielding 20,508 cell-level observations used in the application volume regressions. The median requested loan amount is ₹10,000, and about 66% of applications are accepted.

Panel C reports loan-level characteristics for approved loans. The annual interest rate is annualized from the contract rate and loan tenure. Among approved loans, the average annual interest rate is approximately 39.5%, the median loan tenure is 3 months, and around 20% of borrowers default, where default is defined as 90 or more days past due.

Panel D reports long-term credit access outcomes. These indicators equal one if the applicant obtained any formal credit from any lender in the economy, not just the focal FinTech within two, three, or four years following January 1, 2019, as observed in the national credit bureau data. Roughly 34% of applicants access formal credit within two years, rising to 39% within three years and 43% within four years.

4 Empirical Methodology

Our empirical strategy exploits 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 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

[4]Applicants classified as "General" or "Economically Weaker Section" (EWS) categories are coded as zero.

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 is similar to a regression discontinuity design and 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.

4.1 Difference-in-Differences Design

Our primary identification strategy is based on a DiD framework that compares the lending outcomes of Android users (who lost access to CDR after January 1, 2019) to iOS users (who were not affected 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. Specifically, we estimate the following baseline specification:

$$Y_{i,t} = \beta \cdot (Android_i \times Post) + \theta_{z,p,t} + \theta_{z,p} \times Android_i + \epsilon_{i,t} \tag{1}$$

where $Y_{i,t}$ represents the outcome of interest for loan application i submitted at time t. $Android_i$ 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. Our preferred specification includes $\theta_{z,p,t}$, which represents ZIP Code × loan purpose × month-year fixed effects. This fixed effect controls for any shocks common to Android and iOS in that $\{z, p, t\}$ cell, e.g., local macro conditions, lender policy shifts, loan-purpose specific demand changes, etc., so the estimate of interest is identified only from Android–iOS differences within the same cell over time. The identifying variation comes from comparing Android to iOS users who are applying for a loan for the same purpose, in the same geographic area, and at the same point in time, with the only differential shock being the loss of CDR access for Android users. Moreover, we include $\theta_{z,p} \times Android_i$ which represents ZIP Code × loan purpose × Android fixed effects and accounts for all time-invariant differences between Android and iOS within each ZIP Code–loan purpose, such as local market share, risk composition, or long-run adoption differences. Thus, the estimate of interest is not picking up any static Android-iOS gap, only the additional post-policy shift for Android relative to iOS.

The key identification assumption underlying our DiD approach is the parallel trends assumption: in the absence of Google's privacy directive, lending outcomes for Android (treated) and iOS (control) users would have followed similar trajectories over time. This assumption is credible in our context for three reasons. First, the policy was an exogenous platform-specific intervention imposed unilaterally by Google on January 1, 2019, without advance notice to the lender, and was orthogonal to underlying trends in credit demand, borrower composition, or portfolio performance across operating systems. Second, prior to the policy change, the lender served Android and iOS borrowers through a common menu of products, pricing schedules, and operational processes, with risk-based approval algorithms that were identical across platforms save for the CDR input, ensuring that the two groups were subject to the same credit technology up to the moment of the shock. Third, around the implementation date, the lender did not introduce any other changes to its business model, underwriting policies, or operational practices that differentially targeted one platform, ruling out confounding platform-specific shocks that could mimic the effect of privacy regulation.

We provide empirical support for the parallel-trends assumption in Section 5.1.4, where we estimate an event-time version of equation 1 that allows the Android–iOS gap to evolve flexibly before and after January 2019. Specifically, we estimate the following dynamic specification:

$$Y_{i,t} = \sum_{k=-6,k\neq-1}^{k=+6} \beta_k \cdot (Android_i \times \mathbb{1}_{t=k}) + \theta_{z,p,t} + \theta_{z,p} \times Android_i + \epsilon_{i,t} \tag{2}$$

where k indexes months relative to the implementation of the policy (January 2019), with $k = -1$ (December 2018) serving as the omitted reference period. The coefficients $\{\beta_{-6}, \ldots, \beta_{-2}\}$ test for pre-existing differential trends between Android and iOS users, while $\{\beta_0, \ldots, \beta_6\}$ trace out the dynamic treatment effect after policy implementation. As in equation 1, this specification includes ZIP Code $\times$ month $\times$ loan purpose fixed effects ($\theta_{z,p,t}$) and ZIP Code $\times$ loan purpose $\times$ Android fixed effects ($\theta_{z,p} \times Android_i$), so the identification of the evolution of temporal outcomes comes from within $\{z, p, t\}$ cell Android-iOS differences.

4.2 High-Frequency Event Study Design

While our primary DiD analysis provides for a 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 and borrower behavior when CDR access was removed on January 1, 2019, documenting how quickly market participants responded to this 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, 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 impact. We estimate the following specification for Android users only:

$$Y_{i,t} = \gamma \cdot Post_t + f(distance_t) + \theta_{z,p} + \theta_d + \epsilon_{i,t}$$

where $Y_{i,t}$ represents the outcome of interest for loan application i submitted on date t. 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. Our baseline specification employs a local linear polynomial estimated separately on each side of the threshold, using a triangular weighting kernel à la Calonico, Cattaneo, and Titiunik (2014), Gelman and Imbens (2019) and Cattaneo, Idrobo, and Titiunik (2024). $\theta_{z,p}$ represents ZIP Code × loan purpose fixed effects that control for unobserved heterogeneity across geographic areas and loan purpose. θ_d denotes day-of-the-week fixed effects to account for systematic seasonal patterns in loan outcomes.

The key coefficient of interest is γ, which identifies the effect of the privacy policy implementation. The validity of our high-frequency event study rests on two key identifying assumptions. First, the privacy policy change represents a true discontinuity in the lender's ability to collect CDR 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 cannot 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.

5 Results

The Google privacy policy implemented on January 1, 2019 introduced a sharp, plausibly exogenous change that we exploit as a natural experiment to study how privacy restrictions on alternative data affect credit market outcomes. We begin by documenting aggregate patterns in the data. Figure 1 summarizes lending market dynamics around the policy from January 2017 to January 2020, by plotting monthly application volumes, acceptance rates, and default rates across all users. Around the January 2019 policy implementation date, we observe a pronounced increase in loan applications, a sharp decline in acceptance rates, and a continuation of default rates along their pre-existing trajectory, with no visible discontinuity at the policy cutoff.

These aggregate patterns suggest that the privacy policy set in motion offsetting forces in the lending market. On the borrower side, application volumes increased, consistent with borrowers placing value on enhanced data privacy protections. On the lender side, however, approval rates declined, indicating that the loss of CDRs weakened lenders' ability to enforce repayment ex post and induced tighter screening ex ante. Notably, default rates display no visible discontinuity around the policy implementation date, implying that stricter screening may have largely offset the reduction in enforcement capacity, thereby preserving portfolio quality at roughly pre-policy levels.

5.1 Effect of Privacy Policy: Evidence from DiD Specification

While the aggregate evidence in Figure 1 is informative, it is not sufficient to identify the privacy regulation's effect. We therefore turn to our primary difference-in-differences specification, which compares Android users (treated) to iOS users (control) around the policy change using Equation 1. This specification identifies the impact of the privacy regulation on credit market outcomes, including application volume, acceptance likelihood, default likelihood, and key contract terms such as interest rates and loan maturity.

5.1.1 Borrower-Side Response: Effect of Privacy Policy on Applications

We begin by examining borrower responses in terms of application volume. Since we do not observe individuals who choose not to apply, we measure applications in aggregate at the ZIP Code–month–loan-purpose level separately for Android and iOS users. The dependent variable is the natural logarithm of the number of applications submitted by Android or iOS users in a given ZIP Code in a given month for a given loan purpose.

Table 3 reports the difference-in-differences estimates of the effect of the privacy regulation on the volume of loan applications. Our coefficient of interest, Android $\times$ Post, captures the change in the number of applications by Android users relative to iOS users following the implementation of the policy. Across all specifications in columns (1) to (5), the estimate of interest is positive and statistically significant at 1% level. Moreover, the estimate of interest is fairly stable in magnitude despite the model R^2 increasing by approximately 55 percentage points from column (1) to (5). Our preferred specification reported in column (5) that includes ZIP Code × month-year × loan purpose and ZIP Code × loan purpose × Android fixed effects indicate that the number of applications increased by 25.6%. This response from the borrower-side to an increase in application volume following stricter privacy regulation suggests that borrowers place economically meaningful value on enhanced privacy protections.

5.1.2 Lender-Side Response: Effect of Privacy Policy on Acceptance

Next, we examine lenders' responses to the change in privacy policy. Although borrowers responded positively to enhanced privacy protections, the policy simultaneously imposed substantial costs on lenders by eliminating its access to social collateral that had been integral of its ex-post enforcement technology. In this section, we document how the removal of CDRs from the lenders' enforcement toolkit affected their behavior by analyzing the impact of the policy on the likelihood of application acceptance.

Panel A of Table 4 reports results from application-level regressions in which the dependent variable equals one if an application is accepted and zero otherwise. Our estimate of interest is the coefficient associated with the interaction term of Android and Post, capturing the change in acceptance likelihood for Android users relative to iOS users after the change in privacy policy. We estimate this interaction in a sequence of specifications with increasingly rich fixed effects and present our preferred specification in column (4). Across specifications, the coefficient of interest is negative and statistically significant, indicating a decline in the likelihood of application acceptance for Android users relative to iOS users. In our preferred specification in column (4), we estimate a 16.4 percentage points (pp) decline in the acceptance probability. As shown in column (5), this effect remains robust to controlling for a rich set of application- and applicant-level characteristics, including education type, age decile, credit score decile, income decile, gender, and caste fixed effects.

The 16.5 pp decline in acceptance likelihood is economically significant and corresponds to a 25% reduction relative to the sample average likelihood of application acceptance of 65.94%. This tightening in approval standards is most consistent with lenders contracting credit supply in response to the loss of alternative data that previously served as valuable enforcement tool through its social collateral value.

5.1.3 Effect of Privacy Policy on Portfolio Quality

A natural question is how the lender's response to the loss of CDR-based enforcement affected portfolio quality: Did tighter screening merely preserve performance? Did the removal of this enforcement technology deteriorate loan outcomes? Did a more conservative origination policy actually improve ex post loan performance?

Panel B of Table 4 examines the effect of the privacy policy on the likelihood of default among approved loans. As before, our estimate of interest is the coefficient associated with the interaction term of Android and Post, capturing the change in default likelihood for Android users relative to iOS users after the privacy policy change. Across all specifications, this coefficient is economically minute and statistically insignificant. In our preferred specification, reported in column (4), the point estimate is -0.19%. When we include the full set of applicant- and application-level controls in column (5), the magnitude of the estimate is -0.25%. Both magnitudes are economically small, representing approximately a 0.9-1.2% change relative to the average default rate of 20.5%.

The economically minute and statistically insignificant change in default likelihood suggests that lenders' tightening of screening standards largely offset the loss of ex post enforcement capabilities induced by the privacy regulation, thereby leaving portfolio quality essentially unchanged.

5.1.4 Discussion on Identifying Assumption: Assessment of Pre-Trends

Our baseline results examine the effect of the privacy policy on credit market outcomes by employing a DiD specification comparing lending outcomes between Android and iOS users. It is worth mentioning that the treatment and control groups are unlikely to be identical. For instance, the two groups may be different in several dimensions such as default rate (Berg et al., 2020) and income (Bertrand and Kamenica, 2023). However, for our DiD regressions to identify the effect of the privacy policy, it is not necessary for these groups to be identical in levels. Rather, the DiD framework relies on the assumption that, absent the policy change, both groups would have followed parallel trends in credit market outcomes. This assumption is plausible in our setting because all borrowers were served by the same lender with no other change across the two groups, so their portfolios would be expected to evolve similarly over time, with the only difference being the availability of CDRs.

We first provide non-parametric evidence on this assumption by plotting the evolution of unconditional mean outcomes over event time for treatment and control groups separately. Figure 2 displays six-month pre- and post-policy paths for the log number of applications, the approval probability, and the default probability for Android and iOS users. The pre-policy trajectories for the two platforms are remarkably similar across all three outcomes, lending visual support to the parallel-trends restriction and suggesting that post-policy differences are unlikely to be driven by pre-existing dynamics. Following the policy change, we observe a pronounced increase in application volumes and a decline in approval probabilities for Android users relative to iOS users, while default rates for the two groups continue to evolve similarly.

Next, we formalize these results using a dynamic DiD specification that flexibly controls for potential confounders, as described in Section 4.1. Figure 3 reports event-time coefficients from equation 2 for the six months before and after the policy change for applications, approvals, and defaults, estimated under three increasingly demanding sets of fixed effects: (i) ZIP-code, time, and loan-purpose fixed effects (blue triangles); (ii) ZIP-code × time × loan-purpose fixed effects (red diamonds); and (iii) ZIP-code × time × loan-purpose plus ZIP-code × loan-purpose × Android fixed effects (green circles). Figure 3a uses the natural logarithm of the number of applications as the dependent variable, Figure 3b uses an application-level acceptance indicator, and Figure 3c uses a loan-level default indicator among the originated loans. Across outcomes and specifications with different fixed effects, the dynamic estimates closely mirror the unconditional patterns documented in Figure 2.

Overall, there are two key takeaways for this analysis. First, the pre-policy coefficients are statistically indistinguishable from zero for all three outcomes, indicating parallel pre-trends and supporting the interpretation of the iOS users as a valid counterfactual for Android users in the absence of the privacy regulation. Second, the policy generates an immediate and persistent divergence in the post period: application volumes increase sharply and approval probabilities decline for Android users relative to iOS users, whereas default rates show no systematic differential shift. Taken together, these results indicate that the privacy regulation induced a discrete change in credit market behavior on both the borrower and lender sides, operating primarily through quantities rather than ex post loan performance.

5.1.5 Effects of Privacy Policy on Other Credit Terms: Interest Rates & Maturity

Next, we document the effect of the policy on other credit contract terms, namely interest rates and loan maturity for approved borrowers. Panel A of Appendix Table B.1 reports the impact on pricing. Across specifications, Android borrowers face interest rate increases of about 0.18-0.37 percentage points relative to iOS borrowers after the policy change. Although these estimates are statistically significant, the magnitude of the effect is modest in economic terms, corresponding to a 0.9% change relative to average annual rates near 40% in this market. Panel B of Appendix Table B.1 reports Poisson pseudo-likelihood estimates for loan maturity, measured in months. We find no statistically or economically meaningful change in the duration of new loans extended to Android users relative to iOS users following the policy.

Taken together, the small price adjustments and the absence of maturity effects indicate that credit markets did not materially reprice or re-term loans in response to the privacy regulation. Instead, the main margin of adjustment operates through quantities, application volumes and approval probabilities, rather than through contract terms for approved borrowers.

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

Although our DiD analysis comparing Android and iOS users provides the primary identification of the effects of the policy, we complement this evidence with a high-frequency event study that examines daily patterns in credit market outcomes for Android users around January 1, 2019. This approach offers several advantages. First, by analyzing outcomes at a daily rather than monthly frequency, we can more precisely locate when changes occur and assess whether they line up tightly with the policy implementation date. Second, the high-frequency data allow us to detect any anticipatory behavior or gradual adjustment in the days surrounding the policy change, rather than imposing a discrete shift at the monthly level. Third, examining the stability of the estimated effects across alternative bandwidth choices around the cutoff provides an additional robustness check on our baseline findings.

This analysis focuses exclusively on Android users, who were directly affected by the policy, and examines whether outcomes exhibit an immediate break at the implementation date using a regression-discontinuity-like methodology discussed in Section 4.2. 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. Specifically, Figures 4a, 4b, and 4c plot the change in the natural logarithm of daily applications, the likelihood of acceptance, and the default likelihood, respectively, within a 60-day window around the policy date. These visual patterns are formalized in Table 5, which reports point estimates from our regression-discontinuity-like event-study specifications using bandwidths of 10, 15, 20, 25, and 30 days on either side of the cutoff.

Overall, the evidence in Figure 4 and Table 5 is closely aligned with our baseline DiD findings. We document a discrete increase in applications and a sharp decline in the likelihood of acceptance at the policy date. Across bandwidths, the number of applications rises by 3.2% to 5.7%, while the likelihood of acceptance falls by 25.3 to 27.9 percentage points, with all estimates statistically significant at the 1% level. In contrast, we find no evidence of a discontinuous change in the likelihood of default for loans originated around the policy date.

The estimates are also remarkably stable with respect to bandwidth choice, as shown in Appendix Figure C.1, which plots point estimates and confidence intervals for each outcome as we vary the bandwidth from very narrow (2 days) to relatively wide (150 days) windows. Moreover, Appendix Tables C.1 and C.2 document that our results are robust to a variety of fixed effects specifications and controlling for a rich set of application- and applicant-level characteristics, including education type, age decile, credit score decile, income decile, gender, and caste fixed effects.

Taken together, the high-frequency event study corroborates our baseline DiD results, documenting immediate shifts in applications and acceptance likelihood precisely at the policy implementation date, with no corresponding change in default rates.

5.2.1 Ruling Out Composition Effects

A key identifying assumption for our empirical design is that the privacy policy did not differentially alter the composition of applicants across platforms. If the policy had induced a discrete shift in which operating system marginal borrowers chose (Android versus iOS), our DiD estimates of acceptance rates could confound changes in lending standards with changes in applicant composition.

Appendix Table C.3 addresses this concern by examining whether the policy affected the relative probability that an application came from an Android rather than an iOS user. The dependent variable is an indicator equal to one for Android applications. Across bandwidths of 10–30 days around the policy date, the Post coefficient is economically small and statistically indistinguishable from zero, implying no discrete change in the Android share of applications at the cutoff.

This result is informative because, although Table 3 shows that Android application volumes rise relative to iOS over the broader window, the platform share remains stable in the high-frequency neighborhood of the policy date. In other words, the policy does not appear to generate a discrete shift in which platform marginal applicants choose. Thus, the documented decline in Android acceptance rates is best interpreted as a tightening of lending standards rather than differential sorting across platforms.

A related concern is that lower acceptance rates might simply reflect a deterioration in borrower quality if the policy disproportionately attracted less creditworthy applicants into the pool. Appendix Table C.4 addresses this concern by examining whether the composition of applicants shifts toward lower-income or lower-credit-score borrowers around the policy date. We find no evidence of a statistically or economically meaningful shift toward applicants with weaker observable credit quality: neither average income nor average credit scores of applicants exhibit a decline after the policy. These results suggest that the observed reduction in acceptance likelihood is unlikely to be driven by a worsening of applicant quality and is instead consistent with a genuine contraction in credit supply induced by the loss of CDR-based enforcement.

5.3 Economic Interpretation of the Effect of Privacy Policy

This section discusses the responses from the borrowers' and lenders' perspectives to privacy regulation within a supply–demand framework. Specifically, we argue that the decline in acceptance rates is consistent with a tightening of credit supply, whereas the increase in applications is most consistent with higher credit demand following the change in privacy regulation.

The decline in application acceptance rates is consistent with a contraction in credit supply following the policy change. Prior to the regulation, the lender could rely on CDRs as an ex post enforcement device, using call and network information to pressure delinquent borrowers and tailor collection efforts, which effectively relaxed the need for very stringent screening at origination. Once the privacy regulation removed access to this enforcement technology, the lender faced a weaker ability to discipline borrowers after disbursement. In this environment, preserving portfolio quality requires shifting the margin of adjustment from enforcement to screening. A natural response is to raise acceptance thresholds and reduce approval probabilities for applicants whose risk would previously have been mitigated through CDR-based enforcement. The observed decline in acceptance rates is therefore consistent with a supply-side tightening in which the lender contracts the set of marginal borrowers it is willing to serve in order to offset the loss of enforcement technology.

The increase in the number of applications would be most consistent with an increase in credit demand among borrowers. Specifically, once the lender could no longer mandate the collection of call records, borrowers who place greater value on privacy had stronger incentives to apply. At the same time, because applications are an equilibrium outcome shaped by both demand and anticipated supply, the negative effect of the policy on credit supply is likely to dampen application volumes, as some borrowers may avoid applying when they expect a higher probability of rejection. These two forces work in opposite directions: greater privacy raises demand and tends to increase applications, whereas anticipated tighter screening reduces applications. The fact that we observe an overall increase in applications indicates that the privacy-driven demand effect dominates, and thus the observed rise in applications can be interpreted as a lower-bound estimate of the increase in credit demand attributable to enhanced privacy protections.

Lastly, we interpret the null effect of the privacy regulation on default rates as evidence that lenders fully compensated for the loss of CDR-based enforcement technology by tightening ex ante screening standards. In the pre-policy regime, part of portfolio quality was sustained by the option to deploy CDRs in ex post collection: call- and network-based pressure, better monitoring of strategic delinquency, and targeted recovery efforts. This technology allowed the lender to admit borrowers whose risk was hard to assess or whose observable characteristics were otherwise relatively weak. Once this enforcement technology was removed, the lender tightened ex ante screening, raising the approval threshold to maintain portfolio quality.

5.4 Heterogeneous Effects: Distributional Consequences of Privacy Regulation

The absence of a detectable change in default rates suggests that the lender re-optimized along the screening margin in a way that offset the deterioration in enforcement capacity. We hypothesize that this business-model adjustment shifted the composition of approved loans toward safer borrowers. This section therefore examines the distributional consequences of the loss of alternative data by studying how the policy differentially affected acceptance outcomes across applicant subgroups within the treated population of Android users.

Understanding these heterogeneous effects is important for two reasons. First, it informs the broader welfare implications of the privacy policy: if the contraction in credit supply falls disproportionately on borrowers with thin or non-existent formal credit histories, or with limited access to traditional banking, then stronger privacy protections may come at the cost of reduced financial inclusion. Second, it sheds light on how enforcement technology was used in practice: whether CDR-based tools primarily supported lending to applicants whose risk is difficult to observe, or whether they also relaxed constraints for borrowers who are observably less creditworthy according to standard measures such as credit scores.

We study heterogeneous effects using daily, high-frequency variation around the January 1, 2019 policy cutoff, restricting attention to Android users who are directly exposed to the loss of CDRs. Specifically, we estimate a dynamic difference-in-differences specification that interacts borrower characteristics with week-relative-to-cutoff indicators, thereby tracing out the evolution of treatment effects across applicant subgroups in event time.

Formally, this design identifies differential treatment effects across borrower groups within the treated population while flexibly absorbing rich spatio-temporal heterogeneity through ZIP-code $\times$ loan-purpose, ZIP-code $\times$ time (week), loan-purpose $\times$ time (week), and day-of-week fixed effects. The ZIP-code $\times$ loan-purpose fixed effects control for time-invariant differences in credit conditions across geographic–product cells, such as local demand, competition, and product mix. The ZIP-code $\times$ week and loan-purpose $\times$ week fixed effects non-parametrically capture location-specific and product-specific shocks at weekly frequency – such as regional macro news, marketing campaigns, or category-specific seasonality – so that identification comes from within-cell differences across borrower types over time. Day-of-week fixed effects remove systematic intra-week patterns in application and approval behavior, ensuring that our subgroup comparisons are not driven by mechanical weekly cycles. The identifying assumption is that, absent the policy change, relative outcomes across applicant subgroups would have evolved similarly over time.

Figure 5 reports dynamic treatment effects for application acceptance rates across applicant subgroups. Figure 5a documents heterogeneity by income. We find that low-income applicants, defined as those with monthly incomes below the sample median, experience substantially larger reductions in acceptance likelihood following the policy change, with declines of roughly 7.0 percentage points relative to higher-income applicants. Figure 5b shows differential effects by age. Younger applicants, defined as those below the sample median age, face reductions in acceptance rates of approximately 8.4 percentage points relative to older applicants. Figure 5c reports heterogeneity by prior relationship with the lender. First-time applicants experience reductions in acceptance rates of 10.6 percentage points relative to applicants with established lending relationships. Figure 5d examines heterogeneity by social identity. Applicants from historically marginalized castes, such as scheduled castes, scheduled tribes, and other backward classes, experience reductions in acceptance likelihood of roughly 8.6 percentage points relative to general-caste applicants. Figure 5e analyzes heterogeneity by gender. While female applicants exhibit somewhat lower acceptance likelihoods than male applicants after the policy change, these differences are generally not statistically significant. Finally, Figure 5f documents heterogeneity by credit score. We examine whether applicants with subprime credit scores (below 680) exhibit differential changes in acceptance likelihood relative to applicants with higher credit scores. The results suggest that the effect of the policy change on acceptance likelihood does not vary with credit score. This finding indicates that when lenders possess sufficient hard information to make screening decisions the additional value of CDR-based enforcement technology is limited. Table 6 presents regression estimates that complement the graphical evidence in Figure 5.

Across specifications, low-income, young, first-time, and socially marginalized applicants experience the largest and most statistically significant declines in acceptance likelihood. Moreover, the impact of stricter privacy regulation on these marginal groups emerges immediately at the policy date and persists throughout the observation window. This pattern is consistent with the notion that lower-income, younger, first-time, and socially marginalized applicants are more likely to have thin credit files, making lenders more reliant on CDR-based enforcement mechanisms to ensure repayment for these groups. Put differently, for borrowers lacking hard information the lender substitutes ex-ante screening with ex-post enforcement technologies; when these technologies are removed, these borrowers are most adversely affected by the tightening of ex-ante screening.

Overall, the evidence suggests that the privacy regulation exacerbated exclusion within the digital lending market. The loss of private alternative data most strongly affected borrowers whose creditworthiness could not be easily assessed using traditional information sources. In short, the CDR ban disproportionately constrained credit access for borrowers on the financial margins, highlighting a trade-off between data privacy and financial inclusion.

5.5 Effect of Privacy Policy on Long-Term Financial Inclusion

We have shown that the privacy policy led to an immediate contraction in credit supply at the focal FinTech; we now ask whether affected borrowers substituted into other lenders or instead faced persistent exclusion from formal credit markets. To study these long-run effects, we link our primary loan-level data to comprehensive credit bureau records that track formal borrowing from the universe of financial institutions in India, allowing us to follow applicants' credit access over several subsequent years.

This analysis allows us to address an important question: did privacy regulation merely reallocate borrowing across lenders, or did it push some borrowers out of the formal credit system? If applicants rejected by the FinTech can readily obtain loans from alternative institutions, the welfare consequences of the policy are limited. By contrast, if obtaining a FinTech loan improves access to traditional bank credit by generating additional information (Balyuk, 2023; Dubey, 2024) or helping borrowers build collateral (Beaumont, Tang, and Vansteenberghe, 2022) then rejection induced by privacy regulation may translate into systematic exclusion from formal finance. Our setting is well suited to quantify this FinTech ladder effect by asking how initial approval at the FinTech lender shapes borrowers' subsequent ability to obtain credit elsewhere in the financial system.

5.5.1 Reduced-Form Evidence

Table 7 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 Panels A, B and C of Table 7 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 $\times$ 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 reported in column (4), we find that Android users who applied after the policy change were 2.1 percentage points less likely to access formal credit within 2 years, 2.2 percentage points less likely within 3 years, and 2.2 percentage points less likely within 4 years. All effects are statistically significant.

These estimates are economically meaningful. Given baseline credit access rates of 36% over two years in our sample, a 2.2 percentage point reduction represents approximately a 6.0% 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 et al., 2020; Liberman et al., 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 D.1 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 4.8-5.6 percentage points less likely to access credit within 2 years, 5.1-6.0 percentage points less likely within 3 years, and 4.6-5.6 percentage points less likely within 4 years. These point estimates are somewhat 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.

5.5.2 Evidence from Instrumental Variables Regression

Although 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. Specifically, we employ the following two stage least squares (2SLS) specification:

$$\textbf{First Stage:}\quad \text{App Accepted}_{i,t} = \pi \cdot (Android_i \times Post_t) + \theta_{z,p,t} + \theta_{z,p} \times Android_i + \nu_{i,t}$$

$$\textbf{Second Stage:}\quad \text{Credit Access}_{i,t+s} = \delta \cdot \widehat{\text{App Accepted}}_{i,t} + \theta_{z,p,t} + \theta_{z,p} \times Android_i + \epsilon_{i,t}$$

where 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). App Accepted$_{i,t}$ indicates whether the loan application that was made at time t was approved by our focal FinTech. Our preferred specification includes $\theta_{z,p,t}$ which represents ZIP Code × loan purpose × time (month-year) fixed effects, where time refers to the time of application to our focal FinTech. This allows us to control for time-varying conditions at the ZIP code level that may have triggered the application to our focal FinTech from the applicant. Moreover, we include $\theta_{z,p} \times Android_i$ which represents ZIP Code × loan purpose × Android fixed effects and accounts for all time-invariant differences between Android and iOS within each ZIP Code–loan purpose, such as local market share, risk composition, or long-run adoption differences. Thus, the estimate of interest is not picking up any static Android-iOS gap, only the additional post-policy shift for Android relative to iOS.

The first stage uses the interaction term of Android and Post as an instrument for acceptance of the application, and the predicted probability of acceptance from the first stage is used to estimate the relationship between future credit access and the likelihood of initial acceptance. Table 8 presents the second stage estimate from the 2SLS estimation. Conceptually, the instrumental variables (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 IV strategy rests on two conditions. First, the relevance condition requires that the privacy policy have a strong effect on loan approval decisions. Table 3 shows a sizable and precisely estimated first stage: the policy lowers Android approval rates by roughly 16.4–17.8 percentage points across specifications, indicating a large economic effect on the likelihood of application acceptance. We further document instrument relevance using both an underidentification test and a weak-identification test based on the Kleibergen–Paap statistics. Table 8 reports a Kleibergen–Paap rk LM statistic of 138, rejecting underidentification in all columns, and a Kleibergen–Paap rk Wald F-statistic of 396 in our preferred specification, substantially exceeding standard weak-IV thresholds and confirming that weak identification is not a concern.

Second, the exclusion restriction requires that the privacy policy affect long-run credit access only through its impact on contemporaneous loan approval at the focal FinTech, and not via any direct channel on future borrowing outcomes. This assumption is plausible for several reasons. First, the policy was a platform-specific change to Google's data-access rules that directly constrained only our FinTech's ability to collect CDRs on Android. It did not alter the underwriting technology, product menus, or regulatory environment faced by other lenders, particularly traditional banks, during the sample period. Second, our outcome measures come from credit bureau records that aggregate borrowing from the full universe of formal lenders, so any direct impact of the policy on the focal FinTech's balance sheet, pricing, or portfolio composition is already subsumed in the first stage and does not mechanically contaminate the measurement of subsequent borrowing from other institutions. Third, the timing and scope of the policy were orthogonal to borrower-level fundamentals; it was imposed unilaterally by Google, with no advance notice and with no concurrent changes in the broader credit registry or reporting framework that could differentially affect how future lenders observe or treat Android versus iOS applicants. Moreover, we include high-dimensional fixed effects for geography, time, and loan purpose, which absorb observable factors that might independently drive both initial approval and later credit access. Taken together, these features make it credible that the instrument shifts future credit access primarily by changing who receives an initial loan from our focal FinTech, rather than by directly altering the conditions of borrowing elsewhere in the formal credit market.

The second-stage estimates presented in Table 8 indicate statistically significant and economically large effects of loan approval on future credit access. Focusing on our preferred specifications with the full set of fixed effects and controls in column (4), we find that borrowers whose applications were approved were 12.7 percentage points more likely to access formal credit within 2 years (Panel A), 13.6 percentage points more likely within 3 years (Panel B), and 13.7 percentage points more likely within 4 years (Panel C).

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 the CDR-based enforcement technology was most effective. 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 results. Appendix Table D.2 presents second stage 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 11.2-11.3 percentage points within 2 years, 11.9-12.3 percentage points within 3 years, and 10.8-11.4 percentage points within 4 years. These estimates are somewhat smaller than the DiD-based 2SLS estimates but remain economically large and highly significant, reinforcing the conclusion that initial FinTech loan approval can have substantial positive effects on long-term financial market participation.

Overall, our findings indicate that privacy regulation is associated with persistent financial exclusion, implying that the welfare costs of stronger privacy protections extend well beyond the immediate contraction in credit supply at the affected FinTech lender. The adverse long-run effects of rejection on formal credit access complement existing evidence that credit denials can have a negative effect either by transmitting negative signals to other lenders through credit bureaus (Hertzberg, Liberti, and Paravisini, 2011) or by discouraging borrowers from engaging with financial institutions (Banerjee et al., 2015; Park, Sarkar, and Vats, 2021). Our results differ from those of Chava et al. (2021), who document that marketplace lending leads to worse long-term credit outcomes due to information frictions and debt accumulation. This difference can be attributed to the fact that their sample consists primarily of borrowers with existing credit histories seeking debt consolidation, whereas our lender serves predominantly first-time applicants. 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).

6 Welfare Consequences of Privacy Regulation

Our reduced-form results establish three facts: borrowers apply more when privacy protections are strengthened, lenders tighten screening when they lose access to ex-post enforcement technology, and default rates are unchanged. These findings reveal the direction of both the demand and supply responses to the privacy regulation but cannot, on their own, answer three quantitative questions that are central to the policy debate. First, how much do borrowers value the privacy afforded by the removal of CDR access? Second, does the direct utility from enhanced privacy offset the indirect cost of reduced credit access, so that consumer surplus rises on net? Third, how large are the profit or losses borne by the lender, and through which margins – application, approval, or per-loan surplus – do they operate?

This section attempts to answer these questions by developing a three-stage structural model of credit demand, lender screening, and repayment. The model yields tractable characterizations of borrower privacy valuation, consumer surplus, and lender profits. The key structural parameter is θ, the direct utility gain borrowers derive from enhanced data privacy. The identification of θ exploits a tension between the model's prediction and the data. Absent any privacy value, the model predicts that applications should have reduced after the policy: the loss of CDR-based enforcement worsens lending conditions by tightening screening, which lowers borrowers' expected gain from applying. However, applications among Android users instead increased by 5.7% (see Panel A of Table 5). The parameter θ resolves this tension. It is the direct utility gain from enhanced privacy that, when added to borrowers' application decision, is large enough to overturn the negative contract-terms effect and match the observed increase. In this sense, θ is identified as a residual: it is the component of borrowers' willingness to apply that cannot be rationalized by lending conditions alone.

6.1 Environment and Timing

We consider a discrete-time environment in which, each period t (month), a population of potential borrowers interacts with a FinTech lender.[5] Borrowers are heterogeneous in observable characteristics X_{it}, comprising income, age, gender, caste, and credit score, as well as in idiosyncratic, unobserved liquidity shocks ε_{it} that generate demand for high-APR short-term credit. We interpret ε_{it} as an unexpected expenditure or income shortfall that renders the loan individually rational. We model $\varepsilon_i \sim \text{Beta}(\alpha, \beta)$ as a multiplicative reduction in period-0 income, so that realized income is $(1 - \varepsilon_i)y_i$. Our baseline uses a Beta(2,2) specification, which is symmetric around $1/2$ and bounded on $[0, 1]$ so that consumption remains non-negative.[6] The privacy regime $s \in \{0, 1\}$ captures whether the lender has access to CDRs and is treated as a publicly observed state variable.

The interaction unfolds across three sequential stages, which we solve recursively, starting from Stage 3 and working backward sequentially: Stage 3 (repayment) determines expected default and recovery rates; Stage 2 (screening) uses these to form approval decisions; Stage 1 (application) incorporates the equilibrium approval probability into borrowers' decision to apply for a loan.

1. **Stage 1 (Application)**: Each potential borrower observes her income shock ε_i, the prevailing privacy regime s, and the posted lending terms, and decides whether to apply. Conditional on applying, she selects a loan product defined by the tuple (L_i, T_i, r_i), where L_i denotes loan amount, T_i the repayment horizon in months, and r_i the contracted interest rate.

2. **Stage 2 (Screening and Approval)**: The lender observes (X_i, s) and solves an approval problem, extending credit only to applicants whose expected profit exceeds a required return threshold.

3. **Stage 3 (Repayment)**: The borrower either repays in full or defaults; upon default, the lender recovers a fraction of the outstanding principal through collections.

6.2 Stage 1: Application

Utility from Credit: We adopt a hand-to-mouth framework with log utility and monthly discount factor $\delta = 0.95$. The borrower faces the multiplicative income shock $\varepsilon_i \sim \text{Beta}(2, 2)$ in period 0, so that realized period-0 income is $(1 - \varepsilon_i)y_i$; income reverts to y_i in all subsequent periods. The utility from obtaining a

[5] We model the lender as a monopolist for tractability. Because our data cover a single platform and we identify borrowers' decisions to apply to *this* lender, the relevant outside option is not applying to the focal FinTech rather than borrowing from a competitor. This framing is consistent with the application-level variation we exploit, though we note that the long-run credit bureau analysis in Section 5.5 documents that rejected borrowers do not fully substitute toward other lenders.

[6] We calibrate the parameters of the Beta distribution to match the empirical mean loan-to-income ratio of 53%. Our results are robust to assuming that borrowers experience an expenditure shock equal to the amount borrowed, as shown in Appendix Section E.1.

loan and the counterfactual utility from not borrowing are:

$$u_i^{\text{loan}} = \log((1-\varepsilon_i)y_i + L_i) + \sum_{t=1}^{T_i} \delta^t \log\left(y_i - \frac{L_i(1+r_i)}{T_i}\right), \tag{3}$$

$$u_i^{\text{noloan}} = \log((1-\varepsilon_i)y_i) + \sum_{t=1}^{T_i} \delta^t \log(y_i), \tag{4}$$

where y_i is baseline monthly income, r_i is the total interest rate over the life of the loan, T_i is loan tenure in months, and $L_i(1+r_i)/T_i$ is the fixed monthly installment. The first term captures period-0 consumption utility when income is reduced by the shock; the summation captures discounted repayment-period utility over $t = 1, \ldots, T_i$. The loan raises period-0 consumption by L_i but imposes installment costs in subsequent periods. Therefore, $u_i^{\text{loan}} > u_i^{\text{noloan}}$ if and only if the liquidity benefit outweighs the repayment burden.

Application Decision: Conditional on applying, the borrower is approved with probability $P_i^{A,s}$ and receives utility u_i^{loan}, and is rejected with probability $1 - P_i^{A,s}$ and falls back to u_i^{noloan}, where $P_i^{A,s}$ is the equilibrium approval probability derived in Stage 2. The net expected gain from applying relative to the outside option is:

$$V_i^s = P_i^{A,s}\left(u_i^{\text{loan}} - u_i^{\text{noloan}}\right), \tag{5}$$

which is the approval probability times the loan surplus. Under regime s, the total expected utility from entering the credit market is:

$$EU_i^{\text{enter},s} = V_i^s + \theta \cdot \mathbb{1}\{s = \text{NoCDR}\}, \tag{6}$$

where $\theta \geq 0$ is the direct utility gain from enhanced privacy under the post-policy (NoCDR) regime, reflecting borrowers' preference for not having their call records accessed by the lender. The parameter θ is constant across borrowers; all heterogeneity in the welfare effect of the policy therefore operates through the contract-terms channel rather than through the privacy channel. We normalize the application cost $c = 0$. Each borrower draws $\nu_i \sim$ i.i.d. Type-I EV and applies if $EU_i^{\text{enter},s} + \nu_i \geq 0$, yielding the logistic application probability:

$$P_i^{\text{apply},s} = \Lambda\left(V_i^s + \theta \cdot \mathbb{1}\{s = \text{NoCDR}\}\right). \tag{7}$$

6.3 Stage 2: Screening and Approval

Conditional on application, the lender observes borrower characteristics X_i and constructs a risk signal to assess creditworthiness. Taking the estimated default probabilities $\widehat{PD}_i^s$ and recovery rates $\hat{\lambda}_i^s$ from Stage 3 as given, the lender's per-loan expected profit is:

$$S_i^s = (1 - \widehat{PD}_i^s)\, R_i + \hat{\lambda}_i^s\, \widehat{PD}_i^s\, L_i - L_i, \tag{8}$$

where $R_i = L_i(1 + r_i)^{T_i/12}$ is the present value of full repayment under annualized compounding. The three terms represent, respectively, the expected repayment conditional on no default, the expected recovery conditional on default, and the principal outlay.

The lender approves applicant i if $S_i^s + \epsilon_i \geq \kappa L_i$, where $\epsilon_i \sim$ i.i.d. Type-I EV captures unobserved heterogeneity in processing costs or risk tolerance across loan officers, and κ is the required rate of return per unit of loan that absorbs the lender's average cost of capital. This yields the logistic approval probability:

$$\Pr(\text{Approve}_i \mid s) = \Lambda(S_i^s - \kappa L_i). \tag{9}$$

The counterfactual approval probability absent CDR is obtained by replacing $\widehat{PD}_i^{\text{CDR}}$ and $\hat{\lambda}_i^{\text{CDR}}$ with their NoCDR counterparts in (8), recomputing S_i^{NoCDR}, and evaluating $\Lambda(S_i^{\text{NoCDR}} - \kappa L_i)$.

6.4 Stage 3: Repayment and Default

We abstract from the borrower's strategic repayment decision and instead treat default as an exogenous stochastic process from the lender's perspective. The lender observes borrower characteristics X_i and the prevailing privacy regime s, forms predicted default probability $\widehat{PD}_i^s$ and recovery rate $\hat{\lambda}_i^s$, and treats them as sufficient statistics for credit risk.

The economic mechanism linking CDR to repayment operates through social networks: borrowers with larger social contact networks can be more readily reached by collections agents, raising recovery rates and lowering effective default losses. CDR provides the lender with direct observation of this network, enabling ex-post enforcement through social pressure. Its removal, therefore, degrades both default and recovery because the lender loses the enforcement technology that disciplines borrower repayment.

Default Probability: Let X_i denote observable borrower characteristics, comprising ZIP code, year-month, age decile, credit score decile, loan purpose, income decile, gender, and caste, and let Z_i^{CDR} =

$\ln(\text{Num. Contacts}_i)$ denote the CDR-based social network signal. For pre-policy loans, for which CDR access was available, we estimate:

$$\Pr(\text{Default}_i = 1 \mid X_i, Z_i^{\text{CDR}}) = \Lambda\Big(\alpha_{\text{CDR}} Z_i^{\text{CDR}} + \mu_p + \mu_t + \mu_q + \gamma' X_i\Big), \tag{10}$$

where μ_p, μ_t, and μ_q denote ZIP code, year-month, and loan purpose fixed effects. $\Lambda(x) = e^x/(1 + e^x)$ is the logistic CDF, yielding $\widehat{PD}_i^{\text{CDR}}$. We expect $\alpha_{\text{CDR}} < 0$: borrowers with more social contacts are easier to reach for collections, reducing default probability through the ex-post enforcement channel, as documented in Panel A of Table 2. To construct the counterfactual default probability for the same pre-policy cohort under CDR removal, we set $\alpha_{\text{CDR}} = 0$, reflecting the assumption that losing CDR access eliminates the lender's ability to exploit social network information for enforcement. For post-policy loans, the lender's information set excludes CDRs entirely, so we estimate the following which serves as the relevant credit risk measure in Stage 2.

$$\Pr(\text{Default}_i = 1 \mid X_i) = \Lambda\big(\tilde{\mu}_p + \tilde{\mu}_t + \tilde{\mu}_q + \tilde{\gamma}' X_i\big) \equiv \widehat{PD}_i^{\text{NoCDR}}, \tag{11}$$

Recovery Rate: The recovery rate $\lambda_i \in [0, 1]$ is the fraction of outstanding principal recovered following default. For the pre-policy period, we estimate on the subsample of defaulted loans:

$$\lambda_i = \beta_0 + \beta_z Z_i^{\text{CDR}} + \beta_x' X_i + \mu_p + \mu_t + \mu_q + \varepsilon_i, \tag{12}$$

where μ_p, μ_t, μ_q denote ZIP code, year-month, and loan-purpose fixed effects, respectively. We expect $\beta_z > 0$: borrowers with stronger social networks are easier to contact, facilitating collections, again reflecting the enforcement rather than screening channel. The predicted recovery under CDR is $\hat{\lambda}_i^{\text{CDR}}$; the counterfactual $\hat{\lambda}_i^{\text{NoCDR}}$ is obtained by setting $\beta_z = 0$. For the post-policy period, we estimate the following specification, without the social network signal, that yields $\hat{\lambda}_i^{\text{post}}$.

$$\lambda_i = \tilde{\beta}_0 + \tilde{\beta}_x' X_i + \mu_p + \mu_t + \mu_q + \varepsilon_i, \tag{13}$$

6.5 Consumer Surplus

Expected consumer surplus per potential borrower under each regime s is computed as follows:

$$CS_i^s = P_i^{\text{apply},s}\left[\hat{P}_i^{A,s}\, u_i^{\text{loan}} + (1 - \hat{P}_i^{A,s})\, u_i^{\text{noloan}}\right] + (1 - P_i^{\text{apply},s}) u_i^{\text{noloan}} + \theta \cdot \mathbb{1}\{s = \text{NoCDR}\}, \tag{14}$$

where $\hat{P}_i^{A,s} = \Lambda(S_i^s - \kappa L_i)$ is the estimated approval probability. The welfare effect of the privacy regulation on individuals is given by the change in consumer surplus, denoted by $\Delta CS_i = CS_i^{\text{NoCDR}} - CS_i^{\text{CDR}}$, and is calculated as follows:

$$\begin{aligned}\Delta CS_i = \underbrace{\theta}_{\text{privacy}} &+ \underbrace{\left(P_i^{\text{apply,NoCDR}} - P_i^{\text{apply,CDR}}\right) \hat{P}_i^{A,\text{NoCDR}} \left(u_i^{\text{loan}} - u_i^{\text{noloan}}\right)}_{\text{application margin}} \\ &+ \underbrace{P_i^{\text{apply,CDR}} \left(\hat{P}_i^{A,\text{NoCDR}} - \hat{P}_i^{A,\text{CDR}}\right) \left(u_i^{\text{loan}} - u_i^{\text{noloan}}\right)}_{\text{approval margin}}.\end{aligned} \tag{15}$$

Our decomposition of the change in consumer surplus shows that privacy regulation affects borrowers through three distinct channels. First, there is a privacy channel, captured by the parameter θ, which is positive and represents the direct utility gain from stronger data protection that accrues uniformly to all borrowers. Second, there is an application margin channel, which is also positive as tighter privacy rules increase the probability that consumers apply for credit, so θ indirectly raises consumer surplus by encouraging more borrowers to enter the market. Third, there is an approval margin channel, which is negative, as stronger privacy regulation induces the lender to screen more intensively, reducing the probability that an application is approved and thereby lowering consumer surplus along this margin.

6.6 Lender Profits

Expected lender profit per potential borrower under regime s is given by the following expression:

$$\Pi^s = \frac{1}{N} \sum_{i=1}^{N} P_i^{\text{apply},s} \, \hat{P}_i^{A,s} \, S_i^s, \tag{16}$$

where $S_i^s = (1 - \widehat{PD}_i^s) R_i + \hat{\lambda}_i^s \widehat{PD}_i^s L_i - L_i$ is the per-loan surplus. The effect of the privacy policy on lender profits, denoted by $\Delta\Pi = \Pi^{\text{NoCDR}} - \Pi^{\text{CDR}}$, is given by the following expression:

$$\begin{aligned}\Delta\Pi_i = &\underbrace{\left(P_i^{\text{apply,NoCDR}} - P_i^{\text{apply,CDR}}\right) \hat{P}_i^{A,\text{NoCDR}} \, S_i^{\text{NoCDR}}}_{\text{application margin}} + \underbrace{P_i^{\text{apply,CDR}} \left(\hat{P}_i^{A,\text{NoCDR}} - \hat{P}_i^{A,\text{CDR}}\right) S_i^{\text{NoCDR}}}_{\text{approval margin}} \\ &+ \underbrace{P_i^{\text{apply,CDR}} \, \hat{P}_i^{A,\text{CDR}} \left(S_i^{\text{NoCDR}} - S_i^{\text{CDR}}\right)}_{\text{surplus margin}}.\end{aligned} \tag{17}$$

Our decomposition of the change in lender profits shows that privacy regulation affects lenders through three distinct channels. First, there is an application margin channel, which is positive as stricter privacy rules lead more borrowers to apply for credit, thereby raising expected profits through higher application volume. Second, there is an approval margin channel, which is negative because the lender becomes more cautious and approval probabilities fall when privacy regulation limits access to enforcement technologies, reducing expected profits along this margin. Third, there is a surplus channel, which is negative as default risk rises after the privacy reform, so the lender captures less surplus.

6.7 Identification

Required rate of return (κ). The parameter κ is identified from the observed cross-sectional variation in the likelihood of application approval, i.e., the lender approves borrower i if $S_i^s + \epsilon_i \geq \kappa L_i$. Therefore, the distribution of approvals across the surplus distribution pins down κ via maximum likelihood. We impose that κ is common across the CDR and NoCDR regimes. This assumption is likely to hold because this parameter is driven by lender's cost of capital, which is determined by external funding markets, and is unlikely to have shifted at the exact policy date. Moreover, the reduced-form estimates presented in Appendix Table B.1 show that interest rates and maturities are largely unchanged by the policy, indicating that the lender did not reprice loans in an economically meaningful way to compensate for the loss of enforcement capacity. This empirical observation further supports the assumption of holding κ constant across the two regimes.

Privacy value (θ). The parameter θ is identified from the aggregate change in Android application rates around the policy date. We use the high-frequency event study estimate from Panel A of Table 5, which documents a 5.7% increase in Android applications in a narrow window around the policy date. To map this reduced-form estimate to θ, we aggregate individual application probabilities to construct model-predicted application rates. Under the CDR regime, the aggregate application rate is given by the following expression:

$$\bar{P}^{\text{pre}} = \frac{1}{N}\sum_{i=1}^{N} \Lambda\left(V_i^{\text{CDR}}\right), \tag{18}$$

where the average is taken over the N pre-policy Android applicants and $V_i^{\text{CDR}} = P_i^{A,\text{CDR}}(u_i^{\text{loan}} - u_i^{\text{noloan}})$ is the individual net gain from applying. Under the NoCDR regime, borrowers face the same lending conditions but gain privacy value, denoted by θ. Therefore, under the NoCDR regime, the aggregate application rate is given by the following expression:

$$\bar{P}^{\text{post}}(\theta) = \frac{1}{N}\sum_{i=1}^{N} \Lambda\left(V_i^{\text{NoCDR}} + \theta\right), \tag{19}$$

where V_i^{NoCDR} uses the NoCDR approval probability evaluated at the pre-policy borrower's contract terms (L_i, T_i, r_i) and observable characteristics X_i. This is a pure counterfactual exercise: contract terms (L_i, T_i, r_i) and characteristics X_i are held at their pre-policy values, so the only thing that changes between $\bar{P}^{\text{pre}}$ and $\bar{P}^{\text{post}}(\theta)$ is the privacy value θ. The model-predicted aggregate percentage change in application rates is then:

$$g(\theta) = \frac{\bar{P}^{\text{post}}(\theta) - \bar{P}^{\text{pre}}}{\bar{P}^{\text{pre}}}. \tag{20}$$

The identification logic is as follows. With $\theta = 0$, the model predicts that application rates should have *fallen* after the policy, because the loss of CDR-based enforcement deteriorates lending conditions: $\Delta_i = V_i^{\text{NoCDR}} - V_i^{\text{CDR}} < 0$ on average, so $\bar{P}^{\text{post}}(0) < \bar{P}^{\text{pre}}$. The fact that Android applications instead increase by 5.7% implies that borrowers derive a positive direct utility from the privacy protection afforded by CDR removal. θ is the unique value that rationalizes this increase. It is the privacy premium required to overturn the negative contract-terms effect and match the observed application response following privacy regulation.

6.8 Estimation

The required rate of return κ is estimated by maximum likelihood, stacking pre- and post-policy applicants and maximizing:

$$\hat{\kappa} = \arg\max_{\kappa} \sum_i \left[\text{Approve}_i \cdot \log \Lambda(S_i - \kappa L_i) + (1 - \text{Approve}_i) \cdot \log(1 - \Lambda(S_i - \kappa L_i))\right]. \tag{21}$$

Taking $\hat{\kappa}$ as given, we estimate θ by exact moment matching. Specifically, $\hat{\theta}$ is the unique value satisfying $g(\hat{\theta}) = 0.057$, based on the observed 5.7% increase in Android applications from the high-frequency event study reported in Panel A of Table 5. Lastly, we obtain standard errors for $\hat{\kappa}$ and $\hat{\theta}$ clustered by ZIP code to allow for within-location correlation in both approval decisions and application behavior using the methodology outlined in Appendix Section E.2.

6.9 Results

We begin by first presenting the estimates for θ and κ along with an economic interpretation of the magnitudes of these estimates. We then discuss our estimates for the effect of the privacy regulation on consumer surplus and lender profits. Table 9 presents the parameter estimates for θ and κ, along with the quantitative welfare analysis, which shows that the privacy regulation increases consumer surplus by 0.53% but reduces lender profits by 15.00%.

Estimate of κ: Maximum likelihood yields $\hat{\kappa} = 0.1678$, estimated jointly on pre- and post-policy loan applications under the maintained assumption of a common required rate of return across regimes. The robust standard error for the estimate clustered at the ZIP code level is 0.0009 and the associated 95% confidence interval is $[0.1661, 0.1695]$, indicating that the estimate is statistically different from zero.

The estimator implies model-predicted approval probabilities of 87.4% in the CDR period and 55.5% in the NoCDR period, against observed rates of 95.6% and 40.5% respectively. The model captures the direction and broad magnitude of the approval decline induced by the policy, though it somewhat compresses the gap between regimes: the model predicts a 31.9 percentage point decline versus the 55.1 percentage points observed in the data. The difference likely reflects dimensions of adjustments that our parsimonious structural model cannot fully accommodate. Despite this, the model's predicted approval rates are sufficient for identifying θ through the application moment, which depends on $P_i^{A,s}$ only through its effect on the expected gain from applying.

Estimate of θ: Matching the model moment, denoted by $g(\theta)$ to the observed 5.7% increase in Android applications yields $\hat{\theta} = 0.1334$. The robust standard error for the estimate clustered at the ZIP code level is 0.0009 and the associated 95% confidence interval is $[0.1317, 0.1352]$, indicating that the estimate is statistically different from zero.

Since θ is measured in log-utility units, it is not directly interpretable in economic terms like in rupees or percent of income. To make the magnitude meaningful, we translate this utility gain into a certainty-equivalent income change: the size of a sure, one-time transfer that would leave the borrower just as well off as the privacy reform. Under log utility, a utility gain of θ is equivalent to a lump-sum income transfer, denoted by τ_i, equal to $y_i\left(e^{\hat{\theta}} - 1\right)$.[7] Economically, this corresponds to the following counterfactual thought experiment: "By how much would we need to raise this borrower's monthly income, holding everything else constant, to compensate them for losing the privacy protection?" Alternatively, this estimate of τ can also be interpreted as the willingness to pay for privacy.

Our estimate of τ_i indicates that borrowers, on average, would be willing to pay ₹6,435 for the additional privacy protection. This willingness to pay is economically significant and is equivalent to 14.3% of their monthly income. This estimate suggests a substantial privacy value, i.e., borrowers are

[7] Under log utility, $\log(y_i + \tau_i) - \log(y_i) = \hat{\theta}$ implies $\tau_i = y_i(e^{\hat{\theta}} - 1)$, so the privacy gain is equivalent to an increase in monthly income of $\left(e^{\hat{\theta}} - 1\right) \times 100$ percent.

willing to forgo more than one-seventh of a month's income to have their call records protected from lender access.

6.9.1 Effect of Privacy Regulation on Consumer Surplus

Our decomposition of the change in consumer surplus shows that the privacy regulation raises borrower welfare on net, and that this effect is driven mainly by the direct privacy channel. Relative to the average level under CDR access, consumer surplus increases by 0.53%.

First, the privacy channel contributes +0.54%, This reflects the direct utility gain from removing lender access to call-detail records, which accrues to all potential borrowers, independent of their credit risk or loan terms. Second, the application margin contributes +0.02%. Tighter privacy rules encourage some additional borrowers to apply for credit, and these marginal applicants earn a positive surplus from taking up loans. This is an indirect effect of the privacy channel that operates through borrower behavior and is therefore heterogeneous across borrowers, depending on their characteristics. Third, the approval margin contributes -0.03%. This channel operates through the lenders' response to privacy regulation. The loss of CDR-based enforcement leads the lender to screen more aggressively, reducing approval probabilities and thereby lowering expected surplus for applicants. Taken together, these results imply that the privacy channel more than offsets the negative approval effect, indicating that borrowers place substantial value on data privacy, over and above its purely instrumental role in facilitating credit access.

6.9.2 Effect of Privacy Regulation on Lender Profits

The privacy regulation reduces total lender profits by 15.0%. We decompose the effect of the privacy regulation on the lender's profitability into its three distinct channels. This exercise provides a quantitative analysis of the channels through which the privacy regulation affects lender profits. The application margin contributes +4.8%: the privacy regulation raises application rates, expanding the pool of potential borrowers and partially offsetting the other losses. The approval margin contributes −8.4%: without the enforcement capacity provided by CDR, the lender tightens screening and rejects more applicants to maintain portfolio quality. The surplus margin contributes −11.5%: even conditional on approval, the per-loan surplus S_i^{NoCDR} is lower than S_i^{CDR} because the lender's default predictions are less accurate without CDR-based enforcement, leading to higher realized default rates on approved loans. Together, the approval and surplus margins sum to −19.9%, far exceeding the +4.8% application margin gain.

7 Conclusion

This paper studies the consequences of privacy regulation by exploiting Google's 2019 restriction on CDR access for a major Indian FinTech lender. We show that this intervention reflects a key policy trade-off in digital credit markets: strengthened privacy protections raise loan applications, consistent with higher demand, yet simultaneously induce tighter screening, reflecting an overall contraction in credit supply. The contraction is far from symmetric, as it falls disproportionately on low-income, younger, first-time, and socially marginalized borrowers – the very clientele FinTech lending was expected to bring into the formal financial system. These effects are also persistent. Linking our setting to economy-wide credit bureau data, we document a FinTech ladder effect, whereby initial digital credit access serves as a gateway to broader formal credit. Privacy-induced rejections reduce the probability of obtaining any formal credit by 13.7 percentage points even four years later, indicating that exclusion accumulates over time rather than self-correcting through substitution to other lenders.

We further develop a structural model to quantify the welfare effects of privacy regulation in a digital credit market where CDRs previously played a central enforcement role. The model implies that privacy regulation raises consumer surplus through two channels: a direct utility gain from enhanced privacy, and an indirect gain via a higher propensity to apply for loans that deepens market participation. At the same time, stricter privacy reduces consumer surplus by increasing rejection rates, thereby weakening households' ability to smooth demand shocks. On net, the estimated benefits of privacy outweigh these costs, with consumer surplus rising by 0.53%.

The model also allows us to quantify the impact of privacy regulation on lender profitability. We estimate that lender profits decline by 15%, and decompose this effect into three margins: applications, approvals, and per-loan surplus. We document that the decline in lender profits is driven primarily by the loss of CDR-based enforcement rents through the approval and surplus margins, only partially offset by higher application volumes.

Our findings speak directly to a rapidly expanding global wave of data-privacy regulation in digital credit markets. Following Google's 2019 privacy directive, the Reserve Bank of India's 2022 Guidelines on Digital Lending – and their subsequent consolidation and strengthening in the 2025 Digital Lending Directions – explicitly bar lending apps from accessing borrowers' call logs, contact lists, and telephony functions (Reserve Bank of India, 2022, 2025). Similar restrictions are emerging elsewhere: the European Union's General Data Protection Regulation and ePrivacy Regulation limits the use of personal behavioral data in automated credit decisions, while regulators in emerging markets, such as Brazil, Indonesia, Philippines, Ghana, Kenya, and Nigeria, have moved against mobile lending apps that rely on invasive access to contacts, call logs, and photos amid mounting evidence of predatory data collection and debt-shaming practices. Our results speak directly to the consequences of these interventions: although such rules curtail borrower surveillance, they may simultaneously eliminate enforcement technologies that help lenders extend credit to informationally opaque borrowers.

Taken together, our findings point to the need for a more nuanced design of privacy regulation in digital credit markets. Limiting access to sensitive consumer data is essential for protecting privacy and curbing abusive collection practices, but blanket prohibitions on alternative data can unintentionally widen financial disparities by disproportionately excluding borrowers without extensive credit histories or collateral. The evidence we present suggests that privacy frameworks should be evaluated not only on the safeguards they provide but also on their equilibrium implications for credit access among underserved populations. As digital finance continues to expand globally, particularly in emerging markets where traditional credit infrastructure remains thin, understanding these trade-offs is critical for achieving both consumer protection and inclusive growth. The central challenge for policymakers is to craft regulations that protect privacy and rein in predatory practices without inadvertently shutting out precisely those households that FinTech innovations were intended to bring into the formal financial system.

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.

Agur, Itai, Anil Ari, and Giovanni Dell'Ariccia. 2025. "Bank competition and household privacy in a digital payment monopoly." *Journal of Financial Economics* 166:104019.

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

Alok, Shashwat, Pulak Ghosh, Nirupama Kulkarni, and Manju Puri. 2024. "Breaking barriers to financial access: Cross-platform digital payments and credit markets." Tech. rep., National Bureau of Economic Research.

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.

Armantier, Olivier, Sebastian Doerr, Jon Frost, Andreas Fuster, and Kelly Shue. 2024. "Nothing to hide? Gender and age differences in willingness to share data." *Swiss Finance Institute Research Paper* (24-99).

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." In *Proceedings of the EUROFIDAI-ESSEC Paris December Finance Meeting*.

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.

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

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

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

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

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

Bloomberg. 2023. "Predatory Lending Schemes Are Exploding Across Asia." Available here (Accessed: March 17, 2026).

Bos, Marieke, Emily Breza, and Andres Liberman. 2018. "The labor market effects of credit market information." *The Review of Financial Studies* 31 (6):2005–2037.

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.

Chen, Long, Yadong Huang, Shumiao Ouyang, and Wei Xiong. 2025. "Data Privacy, Data Sharing and Credit Access." Tech. rep., Working Paper.

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)* .

———. 2025. *Shadow banks on the rise: Evidence across market segments*. Financial Markets Group.

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

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.

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.

Doerr, Sebastian, Jon Frost, Leonardo Gambacorta, and Vatsala Shreeti. 2023a. "Big techs in finance." In *The Oxford Handbook of Banking*. Oxford University Press Oxford, 296–325.

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

Duarte, Victor, Julia Fonseca, Divij Kohli, and Julian Reif. 2025. "The effects of deleting medical debt from consumer credit reports." Tech. rep., National Bureau of Economic Research.

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.

Frost, Jon, Leonardo Gambacorta, Yi Huang, Hyun Song Shin, and Pablo Zbinden. 2019. "BigTech and the changing structure of financial intermediation." *Economic policy* 34 (100):761–799.

Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery. 2019. "The role of technology in mortgage lending." *The review of financial studies* 32 (5):1854–1899.

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 Verina F Que. 2023. "The economics of digital privacy." *Annual Review of Economics* 15:267–286.

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

Herkenhoff, Kyle, Gordon M Phillips, and Ethan Cohen-Cole. 2021. "The impact of consumer credit access on self-employment and entrepreneurship." *Journal of financial economics* 141 (1):345–371.

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." Tech. rep., 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." Tech. rep., National Bureau of Economic Research.

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

Mishra, Prachi, Nagpurnanand Prabhala, and Raghuram G Rajan. 2022. "The relationship dilemma: Why do banks differ in the pace at which they adopt new technology?" *The Review of Financial Studies* 35 (7):3418–3466.

Musto, David K. 2004. "What happens when information leaves a market? Evidence from postbankruptcy consumers." *The Journal of Business* 77 (4):725–748.

New York Times. 2017. "China's New Lenders Collect Invasive Data and Offer Billions. Beijing Is Worried." Available here (Accessed: March 17, 2026).

Park, Seongjin, Arkodipta Sarkar, and Nishant Vats. 2021. "Political Voice and (Mortgage) Market Participation: Evidence from the Dilution of Voting Rights Act." *Available at SSRN 3891961* .

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.

Reserve Bank of India. 2022. "Guidelines on Digital Lending." Tech. Rep. DOR.CRE.REC.66/21.07.001/2022-23, Reserve Bank of India. URL https://rbi.org.in/Scripts/NotificationUser.aspx?Id=12382.

———. 2025. "Reserve Bank of India (Digital Lending) Directions, 2025." Tech. rep., Reserve Bank of India. URL https://www.rbi.org.in/Scripts/NotificationUser.aspx?Id=12848.

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

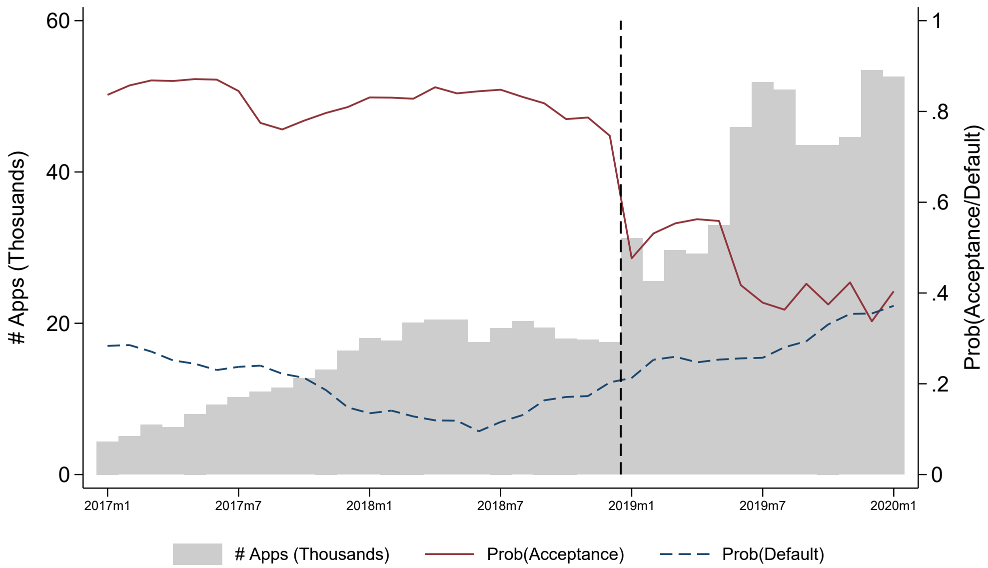
Thakor, Anjan V. 2020. "Fintech and banking: What do we know?" *Journal of financial intermediation* 41:100833.

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.

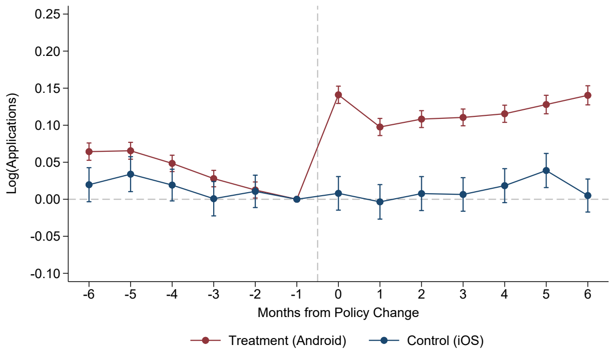
Wall Street Journal. 2019. "Silicon Valley-Backed App Lenders Use Phone Data to Hassle Borrowers." Available here (Accessed: March 17, 2026).

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

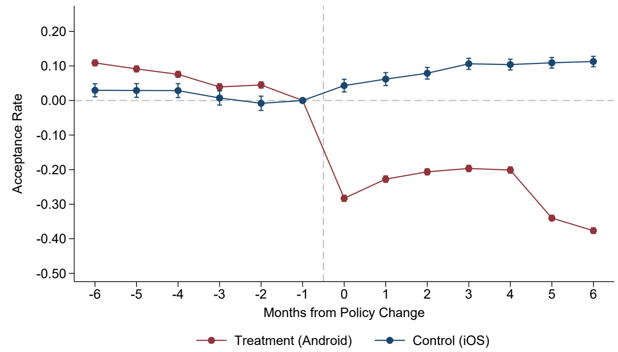


This figure illustrates the number of monthly applications (thousands, left axis), the monthly average probability of acceptance, the monthly average probability of default (right axis) from January 2017 until January 2020.

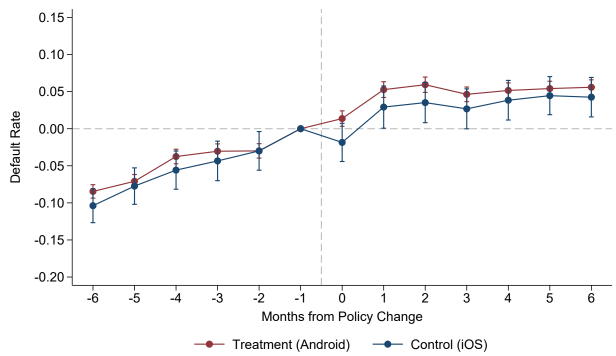
Figure 2: Unconditional Means: Treatment and Control Group Trends



(a) Number of Applications



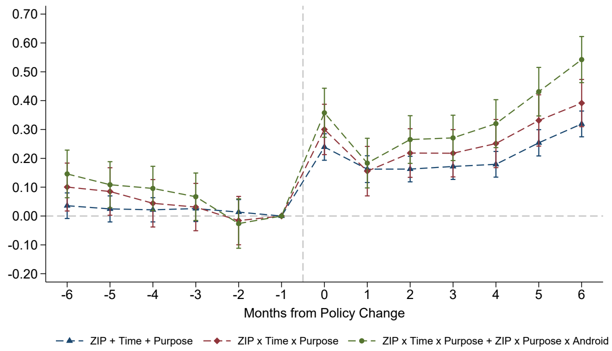
(b) Likelihood of Application Acceptance



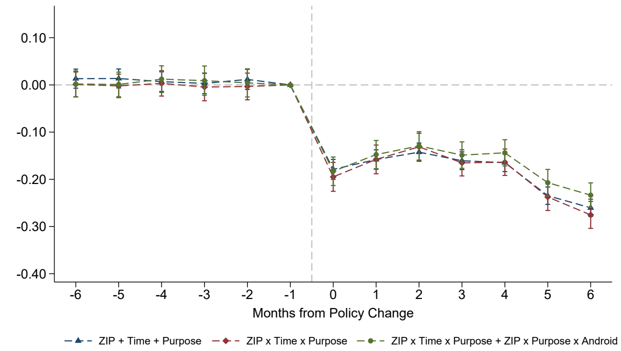
(c) Likelihood of Default

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. Figure 2a shows the natural logarithm of daily applications. Figure 2b shows application acceptance rates. Figure 2c 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.

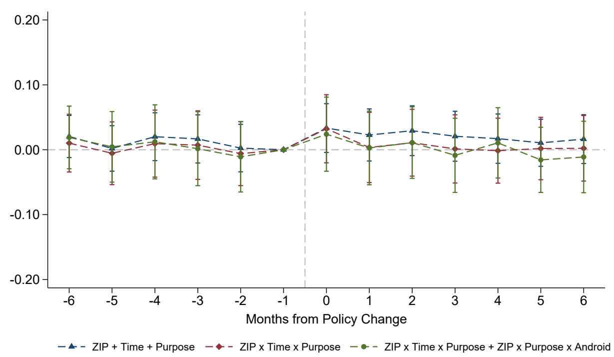
Figure 3: Assessment of Pre-Trends



(a) Number of Applications



(b) Likelihood of Application Acceptance



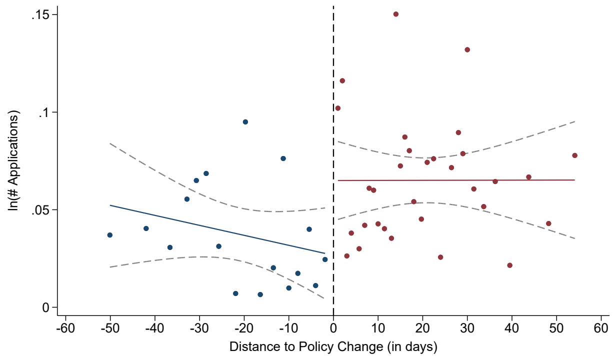
(c) Likelihood of Default

This figure presents estimates from a dynamic difference-in-differences specification that examines pre-trends and the temporal evolution of the policy's effect on credit market outcomes around the Google privacy policy implementation on January 1, 2019. Specifically, each panel plots the coefficients from a dynamic specification that traces the evolution of outcomes for Android users (treated group) relative to iOS users (control group) over event time according to the following regression specification:

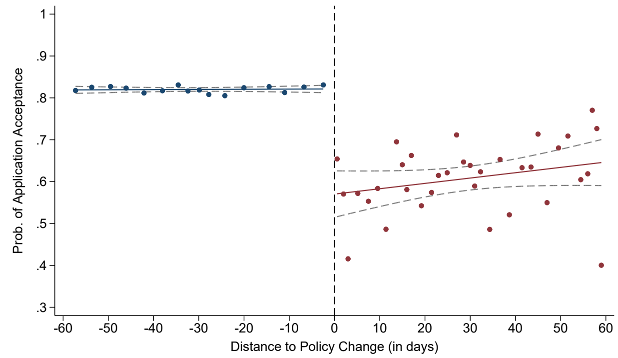
$$Y_{i,t} = \alpha + \sum_{k=-6,k\neq-1}^{k=+6} \beta_k \cdot (Android_i \times \mathbb{1}_{t=k}) + \theta_{z,p,t} + \theta_{z,p} \times Android_i + \epsilon_{i,t}$$

where k indexes months relative to the policy implementation (January 2019), with $k = -1$ (December 2018) serving as the omitted reference period. The coefficients $\{\beta_{-6}, \ldots, \beta_{-2}\}$ 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. This specification includes ZIP Code $\times$ Time (Month-Year) $\times$ Loan Purpose fixed effects ($\theta_{z,p,t}$) and ZIP Code $\times$ Loan Purpose $\times$ Android fixed effects ($\theta_{z,p} \times Android_i$). We plot the temporal evolution of effects using three speciations with different combinations of fixed effects. The blue line (triangle marker) includes ZIP Code, time, and loan purpose fixed effects. The red line (diamond marker) includes ZIP Code $\times$ Time $\times$ Loan Purpose fixed effects. The green line (circle marker) includes both ZIP Code $\times$ Time $\times$ Loan Purpose and ZIP Code $\times$ Loan Purpose $\times$ Android fixed effects. Vertical bars represent 99% confidence intervals based on standard errors clustered at the zip-code level. Figure 3a presents the results using natural logarithm of the number of applications as the dependent variable. Figure 3b presents the results using application-level data where the dependent variable takes a value of one if the application is accepted, and zero otherwise. Figure 3c presents the results using the data on originated loans where the dependent variable takes a value of one if the loan is in default, and zero otherwise.

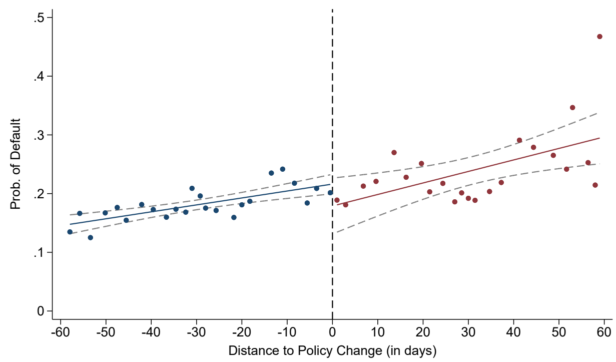
Figure 4: Main Effects of Privacy Policy Change: High-Frequency Event Study



(a) Number of Applications



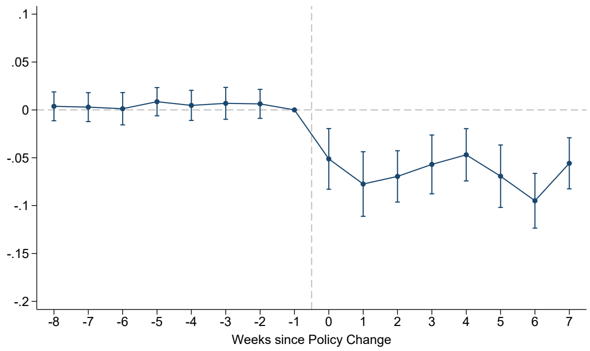
(b) Likelihood of Application Acceptance



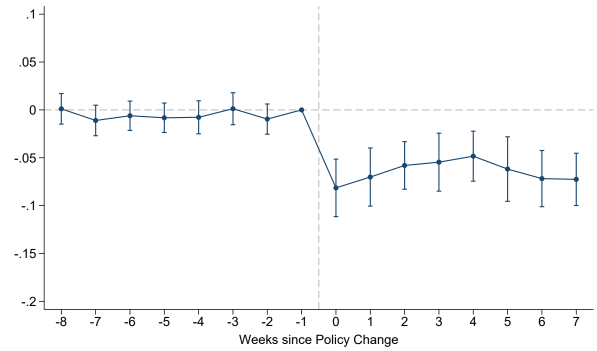
(c) Likelihood of Default

This figure presents high-frequency event study estimates of the Google privacy policy change's effect on credit market outcomes, implemented on January 1, 2019. Figure 4a shows the natural logarithm of daily loan applications. Figure 4b shows the probability of application acceptance. Figure 4c 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.

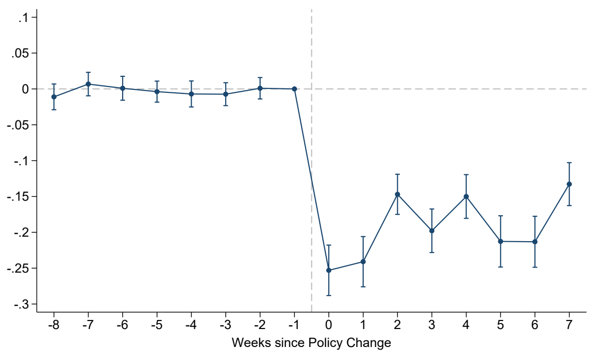
Figure 5: DiD: Heterogeneous Effects on Application Acceptance by Borrower Characteristics



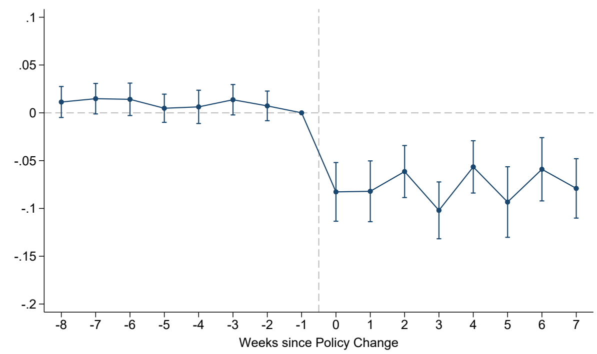
(a) Monthly Borrower Income



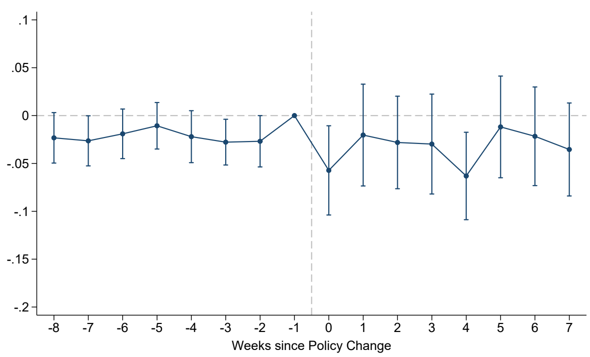
(b) Age



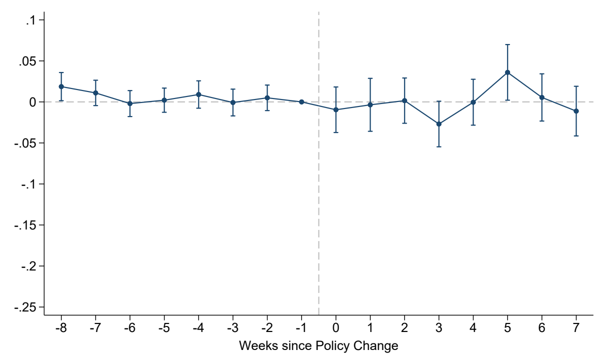
(c) First-Time Applicant



(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: Figure 5a *Low Income* equals one for borrowers with monthly income below the sample median; Figure 5b *Young* equals one for borrowers with age below the sample median; Figure 5c *First-time Applicant* equals one for first-time applicants; Figure 5d *Low Caste* equals one for applicants belonging to scheduled caste, scheduled tribe, and other backward class; Figure 5e *Female* equals one for female applicants; Figure 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.

Table 1: Summary Statistics

Panel A: Characteristics of Loan Applicants						
	# Obs	p25	p50	p75	Mean	Std. Dev.
Age (years)	202,011	27.1781	30.5205	34.5671	31.2827	5.7278
Monthly Income (₹)	202,231	23,000	32,000	52,000	43,164	29,787
Equifax Credit Score	198,946	629.0000	671.0000	702.0000	659.0747	61.2229
1[Female]	202,231	0.0000	0.0000	0.0000	0.1001	0.3001
1[Low Caste]	202,231	0.0000	0.0000	1.0000	0.4689	0.4990
1[College Education]	202,231	1.0000	1.0000	1.0000	0.8573	0.3497
1[First-Time Applicant]	202,231	0.0000	0.0000	1.0000	0.3605	0.4801
1[Android User]	202,231	1.0000	1.0000	1.0000	0.9185	0.2736
Panel B: Application-level Characteristics						
	# Obs	p25	p50	p75	Mean	Std. Dev.
ln(Applications)	20,508	0.0000	0.6931	1.3863	0.7943	0.8812
Loan Amount Requested (₹)	202,231	8,000	10,000	40,000	29,179	34,805
1[Application Accepted]	202,231	0.0000	1.0000	1.0000	0.6594	0.4739
Panel C: Loan-level Characteristics						
	# Obs	p25	p50	p75	Mean	Std. Dev.
1[Default]	128,364	0.0000	0.0000	0.0000	0.2047	0.4034
Annual Interest Rate (%)	125,393	32.2500	37.3123	46.7847	39.5200	6.6440
Loan Tenure (months)	125,393	1.0000	3.0000	6.0000	2.7508	2.0985
Panel D: Long-term Credit Access						
	# Obs	p25	p50	p75	Mean	Std. Dev.
1[Two Years]	202,231	0.0000	0.0000	1.0000	0.3412	0.4741
1[Three Years]	202,231	0.0000	0.0000	1.0000	0.3911	0.4880
1[Four Years]	202,231	0.0000	0.0000	1.0000	0.4273	0.4947

This table presents summary statistics for the sample of loan applications spanning 180 days around the Google privacy policy change date (January 1, 2019). Panel A reports borrower demographic and credit characteristics. Panel B presents application-level characteristics including loan amount requested and acceptance rates. Panel C reports loan-level characteristics for approved loans, including default rates, interest rates, and loan tenure. Panel D reports long-term credit access indicators, measuring whether the applicant obtained any formal credit within two, three, or four years following January 1, 2019. Monthly income and loan amount are reported in Indian Rupees (₹). Default is defined as 90+ days past due. Low caste indicates membership in scheduled caste, scheduled tribe, or other backward class. First-time Applicant indicates first-time applicants to the platform.

Table 2: Number of Contacts, Default Probability, and Loss Given Default: Pre-Policy Period

Panel A: Default Probability					
Dep Var: 𝟙[Default]	(1)	(2)	(3)	(4)	(5)
ln(Num. Contacts)	-0.0730***	-0.0703***	-0.0682***	-0.0621***	-0.0597***
	(0.0159)	(0.0179)	(0.0178)	(0.0173)	(0.0169)
ZIP Code × Month-Year FE		✓	✓		
ZIP Code × Loan Purpose FE			✓		
ZIP Code × Month-Year × Loan Purpose FE				✓	✓
Demographic Controls					✓
# Obs	44,856	44,856	44,856	44,856	44,856
R^2	0.0011	0.2507	0.2833	0.3825	0.4340
AUC	0.5204	0.8279	0.8477	0.8943	0.9214

Panel B: Loss Given Default					
Dep Var: Loss Given Default	(1)	(2)	(3)	(4)	(5)
ln(Num. Contacts)	-0.0404***	-0.0522***	-0.0493***	-0.0417***	-0.0462***
	(0.0050)	(0.0075)	(0.0083)	(0.0113)	(0.0125)
ZIP Code × Month-Year FE		✓	✓		
ZIP Code × Loan Purpose FE			✓		
ZIP Code × Month-Year × Loan Purpose FE				✓	✓
Demographic Controls					✓
# Obs	4,788	4,788	4,788	4,788	4,788
R^2	0.0054	0.4377	0.4626	0.5055	0.5195

This table examines the relationship between contact list size and loan performance for loans originated between January 2017 and December 2018, prior to the Google privacy policy change. In Panel A, the dependent variable is a binary indicator equal to one if the borrower defaulted on their loan (defined as 90+ days past due). In Panel B, the dependent variable is loss given default, defined as principal due divided by the total loan obligation (loan amount grossed up by the annualized interest rate) and estimated on the subsample of defaulted loans. 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. Column (3) adds ZIP Code × loan purpose fixed effects. Column (4) includes ZIP Code × month-year × loan purpose fixed effects. Column (5) adds demographic controls including education, age decile, credit score decile, loan maturity, income decile, female, and low caste fixed effects. AUC reports the area under the receiver operating characteristic curve. 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 Number of Applications

Dep Var: ln(# Applications)	(1)	(2)	(3)	(4)	(5)
Android × Post	0.2349***	0.2349***	0.2349***	0.2600***	0.2557***
	(0.0153)	(0.0153)	(0.0153)	(0.0139)	(0.0140)
Android	0.8977***	0.8977***	0.8977***		
	(0.0248)	(0.0249)	(0.0248)		
Post	-0.0111				
	(0.0075)				
Month-Year FE		✓			
ZIP Code FE		✓			
Loan Purpose FE		✓			
ZIP Code × Month-Year × Loan Purpose FE			✓	✓	✓
ZIP Code × Android FE				✓	
ZIP Code × Loan Purpose × Android FE					✓
# Obs	20,508	20,508	20,508	20,508	20,508
R^2	0.3423	0.6648	0.7868	0.8639	0.9026

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) reports the baseline specification without fixed effects. Column (2) adds Month-Year, ZIP Code, and Loan Purpose fixed effects. Column (3) replaces the additive fixed effects with fully interacted ZIP Code × Month-Year × Loan Purpose fixed effects. Column (4) includes ZIP Code × Android fixed effects in addition to ZIP Code × Month-Year × Loan Purpose fixed effects. Column (5) further saturates the specification by including ZIP Code × Loan Purpose × Android fixed effects along with 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 4: Effect of the Policy Change on Application Acceptance and Default

Panel A: Application Acceptance

Dep Var: Application Accepted	(1)	(2)	(3)	(4)	(5)
Android × Post	-0.1779***	-0.1699***	-0.1706***	-0.1644***	-0.1557***
	(0.0037)	(0.0048)	(0.0049)	(0.0048)	(0.0048)
Android	-0.0249***	-0.0138***			
	(0.0030)	(0.0029)			
ZIP Code FE	✓				
Month-Year FE	✓				
Loan Purpose FE	✓				
ZIP Code × Month-Year × Loan Purpose FE		✓	✓	✓	✓
ZIP Code × Android FE			✓	✓	✓
ZIP Code × Loan Purpose × Android FE				✓	✓
Demographic Controls					✓
# Obs	202,231	202,231	202,231	202,231	201,970
R^2	0.6385	0.7426	0.7440	0.7477	0.7593

Panel B: Loan Default

Dep Var: Loan Default	(1)	(2)	(3)	(4)	(5)
Android × Post	0.0100	0.0048	0.0019	-0.0019	-0.0025
	(0.0064)	(0.0070)	(0.0072)	(0.0071)	(0.0072)
Android	0.0197***	0.0198***			
	(0.0046)	(0.0050)			
ZIP Code FE	✓				
Month-Year FE	✓				
Loan Purpose FE	✓				
ZIP Code × Month-Year × Loan Purpose FE		✓	✓	✓	✓
ZIP Code × Android FE			✓	✓	✓
ZIP Code × Loan Purpose × Android FE				✓	✓
Demographic Controls					✓
# Obs	128,364	128,364	128,364	128,364	128,182
R^2	0.0480	0.2583	0.2655	0.2762	0.2873

This table presents difference-in-differences estimates examining the effect of the Google privacy policy change on loan application acceptance rates (Panel A) and loan default rates (Panel B). In Panel A, the dependent variable is a binary indicator equal to one if the application was accepted. In Panel B, 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 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) includes separate ZIP Code, month-year, and loan purpose fixed effects. Column (2) includes ZIP Code × month-year × loan purpose fixed effects. Column (3) adds ZIP Code × Android fixed effects. Column (4) further adds ZIP Code × Loan Purpose × Android fixed effects. Column (5) adds 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 5: Effect of the Policy Change on Credit 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

Panel B: Likelihood of 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

Panel C: Likelihood of Default

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

This table presents high-frequency event study estimates of the Google privacy policy change's effect on credit 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 6: Heterogeneity & Effect of the Policy Change on Application Acceptances

Dep Var: Application Accepted	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Low Income × Post	-0.0596***						-0.0326*
	(0.0175)						(0.0172)
Low Income	-0.0074						-0.0049
	(0.0090)						(0.0082)
Young × Post		-0.0790***					-0.0610***
		(0.0163)					(0.0165)
Young		-0.0062					0.0019
		(0.0084)					(0.0083)
First-time Applicant × Post			-0.2514***				-0.2358***
			(0.0194)				(0.0189)
First-time Applicant			-0.0090				-0.0067
			(0.0085)				(0.0085)
Low Caste × Post				-0.0830***			-0.0381**
				(0.0178)			(0.0167)
Low Caste				-0.0013			0.0008
				(0.0087)			(0.0083)
Female × Post					-0.0281		-0.0168
					(0.0267)		(0.0252)
Female					-0.0150		-0.0101
					(0.0145)		(0.0144)
Sub-Prime Credit Score × Post						-0.0078	-0.0180
						(0.0177)	(0.0163)
Sub-Prime Credit Score						-0.0025	-0.0029
						(0.0087)	(0.0078)
ZIP Code×Loan Purpose FE	✓	✓	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓	✓	✓
ZIP Code×Week FE	✓	✓	✓	✓	✓	✓	✓
Loan Purpose×Week FE	✓	✓	✓	✓	✓	✓	✓
Bandwidth	30	30	30	30	30	30	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	21,402	21,358	21,402	21,402	21,402	20,989	20,952
R^2	0.7917	0.7918	0.8084	0.7925	0.7896	0.7905	0.8125

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 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. 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. First-time Applicant is a binary variable for borrowers who are first time applicants, 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. Female is a binary variable that takes a value of one for females, and zero otherwise. 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. 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 Privacy Policy on Long-Term Credit Access: Evidence from DiD Specification

Panel A: Credit Access Within 2 Years				
Dep Var: Formal Credit Access (2 Years)	(1)	(2)	(3)	(4)
Android × Post	-0.0987***	-0.0362***	-0.0316***	-0.0209**
	(0.0085)	(0.0084)	(0.0099)	(0.0097)
Android	-0.0468***	-0.0477***	-0.0479***	
	(0.0097)	(0.0095)	(0.0108)	
ZIP Code FE		✓		
Month-Year FE		✓		
Loan Purpose FE		✓		
ZIP Code × Month-Year × Loan Purpose FE			✓	✓
ZIP Code × Loan Purpose × Android FE				✓
# Obs	202,231	202,231	202,231	202,231
R^2	0.0041	0.0968	0.3061	0.3298
Panel B: Credit Access Within 3 Years				
Dep Var: Formal Credit Access (3 Years)	(1)	(2)	(3)	(4)
Android × Post	-0.1014***	-0.0351***	-0.0348***	-0.0223**
	(0.0088)	(0.0088)	(0.0103)	(0.0099)
Android	-0.0578***	-0.0581***	-0.0528***	
	(0.0100)	(0.0100)	(0.0113)	
ZIP Code FE		✓		
Month-Year FE		✓		
Loan Purpose FE		✓		
ZIP Code × Month-Year × Loan Purpose FE			✓	✓
ZIP Code × Loan Purpose × Android FE				✓
# Obs	202,231	202,231	202,231	202,231
R^2	0.0048	0.0986	0.3091	0.3320
Panel C: Credit Access Within 4 Years				
Dep Var: Formal Credit Access (4 Years)	(1)	(2)	(3)	(4)
Android × Post	-0.0949***	-0.0302***	-0.0319***	-0.0224**
	(0.0087)	(0.0088)	(0.0103)	(0.0099)
Android	-0.0630***	-0.0628***	-0.0559***	
	(0.0098)	(0.0097)	(0.0108)	
ZIP Code FE		✓		
Month-Year FE		✓		
Loan Purpose FE		✓		
ZIP Code × Month-Year × Loan Purpose FE			✓	✓
ZIP Code × Loan Purpose × Android FE				✓
# Obs	202,231	202,231	202,231	202,231
R^2	0.0047	0.0967	0.3087	0.3300

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 (Panel A), three years (Panel B), or four years (Panel C) 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) presents the baseline specification without fixed effects. Column (2) adds ZIP Code, month-year, and loan purpose fixed effects. Column (3) includes the full interaction of ZIP Code × month-year × loan purpose fixed effects. Column (4) adds ZIP Code × loan purpose × Android fixed effects. In column (4), the Android main effect is absorbed by the ZIP Code × loan purpose × Android fixed effects. Standard errors reported in parentheses are clustered at the ZIP code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Effect of Privacy Policy on Long-Term Credit Access: Evidence from 2SLS Specification

Panel A: Credit Access Within 2 Years

Dep Var: Formal Credit Access (2 Years)	(1)	(2)	(3)	(4)
App Accepted	0.0812***	0.4318***	0.4477***	0.1268**
	(0.0071)	(0.0461)	(0.0537)	(0.0589)
ZIP Code FE		✓		
Month-Year FE		✓		
Loan Purpose FE		✓		
ZIP Code × Month-Year × Loan Purpose FE			✓	✓
ZIP Code × Loan Purpose × Android FE				✓
# Obs	202,231	202,231	202,231	202,231
KP LM Statistic	687.8414	275.9936	177.1940	138.0451
KP F Statistic	4619.7008	1758.7024	799.6141	396.2597

Panel B: Credit Access Within 3 Years

Dep Var: Formal Credit Access (3 Years)	(1)	(2)	(3)	(4)
App Accepted	0.0977***	0.4821***	0.4935***	0.1359**
	(0.0073)	(0.0460)	(0.0534)	(0.0599)
ZIP Code FE		✓		
Month-Year FE		✓		
Loan Purpose FE		✓		
ZIP Code × Month-Year × Loan Purpose FE			✓	✓
ZIP Code × Loan Purpose × Android FE				✓
# Obs	202,231	202,231	202,231	202,231
KP LM Statistic	687.8414	275.9936	177.1940	138.0451
KP F Statistic	4619.7008	1758.7024	799.6141	396.2597

Panel C: Credit Access Within 4 Years

Dep Var: Formal Credit Access (4 Years)	(1)	(2)	(3)	(4)
App Accepted	0.0985***	0.4835***	0.4953***	0.1365**
	(0.0076)	(0.0446)	(0.0518)	(0.0596)
ZIP Code FE		✓		
Month-Year FE		✓		
Loan Purpose FE		✓		
ZIP Code × Month-Year × Loan Purpose FE			✓	✓
ZIP Code × Loan Purpose × Android FE				✓
# Obs	202,231	202,231	202,231	202,231
KP LM Statistic	687.8414	275.9936	177.1940	138.0451
KP F Statistic	4619.7008	1758.7024	799.6141	396.2597

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 (Panel A), three years (Panel B), or four years (Panel C) following January 1, 2019. The endogenous variable is App Accepted, which equals one if the loan application was approved. The instrument is Android × 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) presents the baseline specification without fixed effects. Column (2) adds ZIP Code, month-year, and loan purpose fixed effects. Column (3) includes the full interaction of ZIP Code × month-year × loan purpose fixed effects. Column (4) adds ZIP Code × loan purpose × Android 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 9: Parameter Estimates and Welfare Decomposition

Key Quantities	Estimate
Direct Utility Gain from Privacy ($\hat{\theta}$)	0.1334***
	(0.0009)
Lender's Required Rate of Return, ($\hat{\kappa}$)	0.1678***
	(0.0009)
(1) Percent Change in Consumer Surplus	
Total	+0.53%
Privacy Value	+0.54%
Application Margin	+0.02%
Approval Margin	-0.03%
(2) Percent Change in Lender Profits	
Total	-15.00%
Application Margin	+4.84%
Approval Margin	-8.40%
Surplus Margin	-11.50%

$\hat{\kappa}$ is estimated by MLE on pre- and post-policy applicants (equation (9)). $\hat{\theta}$ is estimated by exact moment matching to the observed 5.7% increase in Android applications (equation (20)). Changes in consumer surplus and lender profits are computed using equations (14) and (16), respectively, and expressed as percentages of the pre-policy (CDR) baseline. Standard errors reported in parentheses are clustered at the ZIP code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

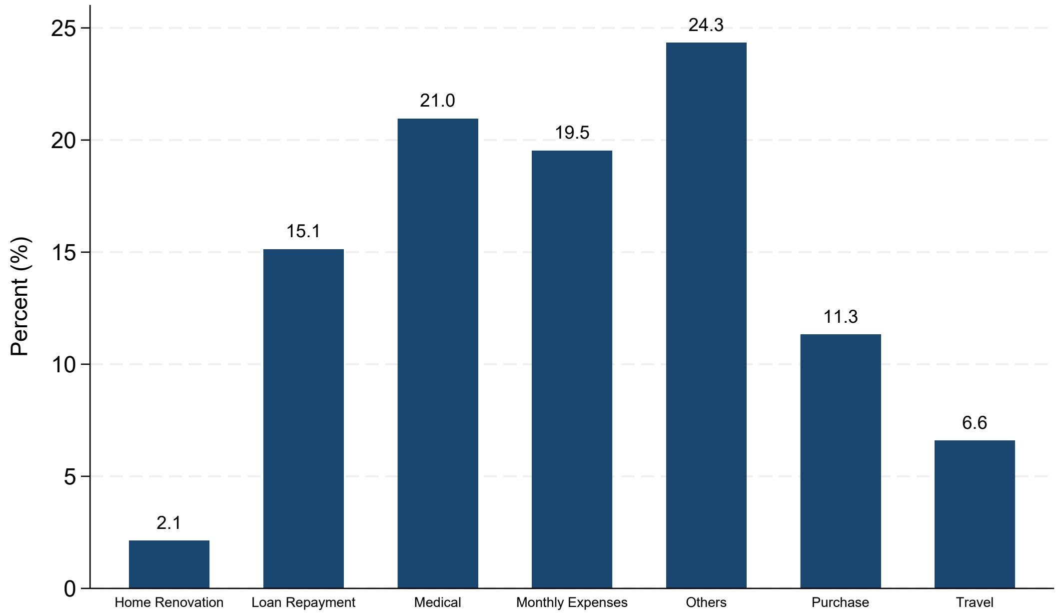
Internet Appendix for:

"When Privacy Protects but Excludes: The Costs and Benefits of Privacy Regulation in Credit Markets"

Contents

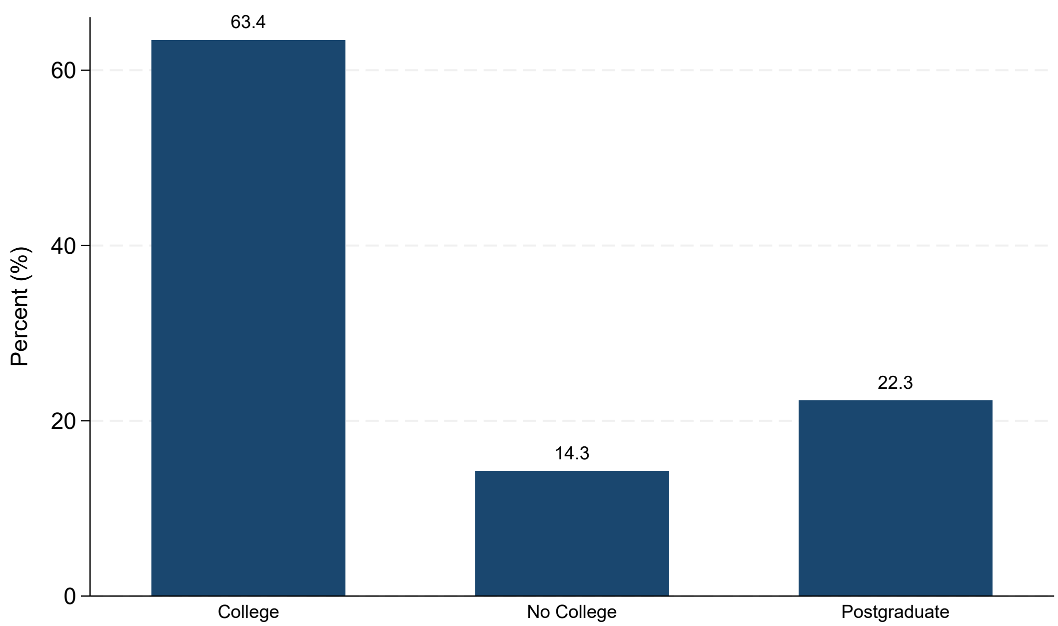
Appendix A Data

Figure A.1: Distribution of Loan Purpose



Notes: This figure shows the distribution of loan purposes across all applications in the sample.

Figure A.2: Distribution of Applicant Education



Notes: This figure shows the distribution of education levels across all applicants in the sample.

Appendix B Robustness for DiD Analysis

Table B.1: Effect of the Policy Change on Interest Rates and Loan Maturity

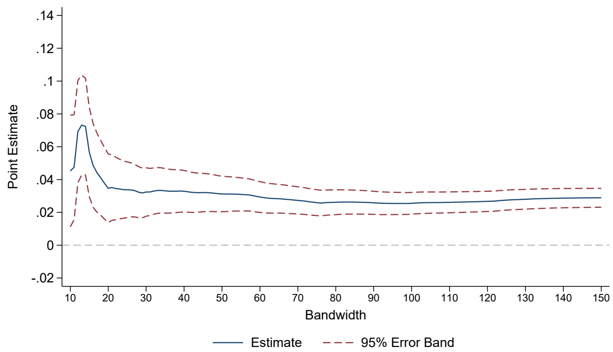
Panel A: Annual Interest Rates (%)					
Dep Var: Interest Rate (Annual)	(1)	(2)	(3)	(4)	(5)
Android × Post	0.1786	0.2477*	0.3307**	0.3747***	0.2888**
	(0.1212)	(0.1369)	(0.1356)	(0.1397)	(0.1305)
Android	0.9462***	0.9172***			
	(0.1208)	(0.1312)			
ZIP Code FE	✓				
Month-Year FE	✓				
Loan Purpose FE	✓				
ZIP Code × Month-Year × Loan Purpose FE		✓	✓	✓	✓
ZIP Code × Android FE			✓	✓	✓
ZIP Code × Android × Loan Purpose FE				✓	✓
Demographic Controls					✓
# Obs	125,393	125,393	125,393	125,393	125,215
R^2	0.1027	0.3014	0.3127	0.3254	0.4660

Panel B: Loan Maturity (Months, PPML)					
Dep Var: Loan Maturity (Months)	(1)	(2)	(3)	(4)	(5)
Android × Post	0.0038	-0.0077	-0.0160	-0.0187	-0.0078
	(0.0124)	(0.0143)	(0.0143)	(0.0148)	(0.0139)
Android	-0.1065***	-0.0997***			
	(0.0135)	(0.0147)			
ZIP Code FE	✓				
Month-Year FE	✓				
Loan Purpose FE	✓				
ZIP Code × Month-Year × Loan Purpose FE		✓	✓	✓	✓
ZIP Code × Android FE			✓	✓	✓
ZIP Code × Android × Loan Purpose FE				✓	✓
Demographic Controls					✓
# Obs	125,393	125,393	125,393	125,393	125,215
Pseudo R^2	0.0373	0.1146	0.1186	0.1233	0.1783

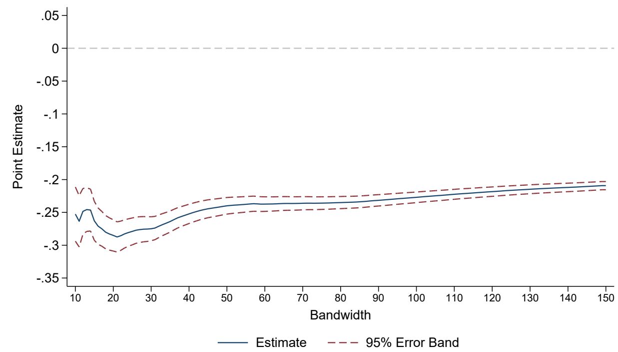
This table presents difference-in-differences estimates examining the effect of the Google privacy policy change on loan pricing terms for approved loans. Panel A shows OLS effects on annual interest rates (in percentage points), while Panel B shows Poisson pseudo-maximum likelihood (PPML) effects on 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) includes separate ZIP code, month-year, and loan purpose fixed effects. Column (2) includes ZIP Code × month-year × loan purpose fixed effects. Column (3) adds ZIP Code × Android fixed effects. Column (4) further adds ZIP Code × Android × loan purpose fixed effects. Column (5) adds demographic controls including education, age decile, credit score decile, income decile, female, and low caste fixed effects. In columns (3)–(5), the Android main effect is absorbed by the ZIP Code × Android fixed effects. Standard errors reported in parentheses are clustered at the ZIP code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix C Robustness for High-Frequency Event Study Analysis

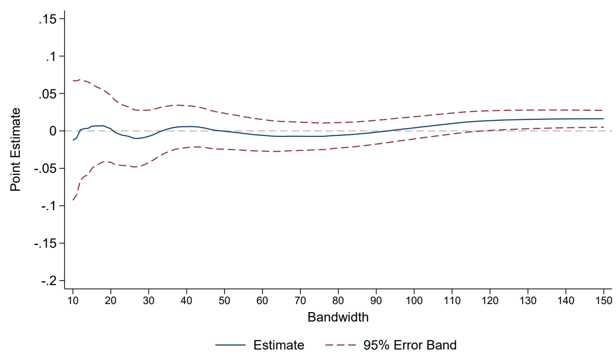
Figure C.1: Robustness to Bandwidth Choice: High-Frequency Event Study Estimates Across Multiple Bandwidths



(a) Number of Applications



(b) Likelihood of Application Acceptance



(c) Likelihood of Default

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 C.1: Robustness to Covariates: Effect of the Policy Change on Application Acceptances in High-Frequency Event Study Design

Dep Var: Application Accepted	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Post	-0.4912***	-0.4920***	-0.2687***	-0.2749***	-0.2746***	-0.2743***	-0.2720***	-0.2688***	-0.2684***	-0.2676***
	(0.0079)	(0.0091)	(0.0073)	(0.0093)	(0.0093)	(0.0093)	(0.0093)	(0.0093)	(0.0094)	(0.0093)
Female (=1)						-0.0241***	-0.0297***	-0.0207**	-0.0233***	-0.0220***
						(0.0085)	(0.0085)	(0.0083)	(0.0085)	(0.0085)
Low Caste (=1)							-0.0639***	-0.0612***	-0.0604***	-0.0570***
							(0.0055)	(0.0055)	(0.0056)	(0.0056)
ZIP Code FE		✓	✓							
Loan Purpose FE			✓							
ZIP Code×Loan Purpose FE				✓	✓	✓	✓	✓	✓	✓
Day of the Week FE				✓	✓	✓	✓	✓	✓	✓
Education FE					✓	✓	✓	✓	✓	✓
Age Decile FE								✓	✓	✓
Credit Score Decile FE									✓	✓
Income Decile FE										✓
Bandwidth	30	30	30	30	30	30	30	30	30	30
Kernel	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
R^2	0.2170	0.3431	0.6650	0.7465	0.7474	0.7475	0.7501	0.7525	0.7544	0.7563

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 C.2: Robustness to Covariates: Effect of the Policy Change on Default in High-Frequency Event Study Design

Dep Var: Loan Default	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Post	-0.0085	-0.0162	-0.0153	-0.0072	-0.0072	-0.0071	-0.0071	-0.0055	-0.0067	-0.0058
	(0.0126)	(0.0145)	(0.0145)	(0.0178)	(0.0178)	(0.0178)	(0.0178)	(0.0177)	(0.0173)	(0.0174)
Female (=1)						-0.0060	-0.0052	-0.0063	-0.0005	-0.0020
						(0.0148)	(0.0148)	(0.0151)	(0.0154)	(0.0154)
Low Caste (=1)							0.0069	0.0066	0.0071	0.0055
							(0.0091)	(0.0091)	(0.0093)	(0.0093)
ZIP Code FE		✓	✓							
Loan Purpose FE			✓							
ZIP Code×Loan Purpose FE				✓	✓	✓	✓	✓	✓	✓
Day of the Week FE				✓	✓	✓	✓	✓	✓	✓
Education FE					✓	✓	✓	✓	✓	✓
Age Decile FE								✓	✓	✓
Credit Score Decile FE									✓	✓
Income Decile FE										✓
Bandwidth	30	30	30	30	30	30	30	30	30	30
Kernel	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
R^2	0.0009	0.1403	0.1432	0.2788	0.2790	0.2790	0.2791	0.2799	0.2927	0.2949

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 C.3: Did the Policy Change which Platform (Android or iOS) Applicants Choose?

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×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

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 C.4: Effect of the Policy on the Composition of Applicant Pool: Credit Quality

Panel A: ln(Income)

Dep Var: ln(Income)	(1)	(2)	(3)	(4)	(5)
Post	-0.0555	-0.0246	-0.0173	-0.0087	-0.0089
	(0.0428)	(0.0326)	(0.0277)	(0.0237)	(0.0208)
ZIP Code × 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	0.4106	0.3928	0.3696	0.3605	0.3606

Panel B: ln(Equifax Credit Score)

Dep Var: ln(Credit Score)	(1)	(2)	(3)	(4)	(5)
Post	-0.0016	-0.0026	-0.0005	0.0021	0.0014
	(0.0076)	(0.0056)	(0.0047)	(0.0038)	(0.0033)
ZIP Code × Loan Purpose FE	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓
Bandwidth	10	15	20	25	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	6,933	12,426	16,892	22,774	30,025
R^2	0.3798	0.3507	0.3209	0.3062	0.3035

This table presents high-frequency event study estimates examining the effect of the privacy policy change on the characteristics of loan applicants. Panel A shows effects on the natural logarithm of applicant income. Panel B shows effects on the natural logarithm of Equifax credit scores. Columns (1) through (5) use bandwidths of 10, 15, 20, 25, and 30 days on either side of the policy change date (January 1, 2019), 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 C.5: Effect of the Policy Change on Composition of Applications

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(Income)	ln(Age)	Credit Score	Low Caste	Female	First-time Applicant
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×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

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 first-time applicants 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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix D Effect of Privacy Policy on Long-Term Credit Access

Table D.1: Effect of Privacy Policy on Long-Term Credit Access: Robustness using High-Frequency Identification

Panel A: Credit Access Within 2 Years

Dep Var: Formal Credit Access (2 Years)	(1)	(2)	(3)
Post	-0.0556***	-0.0484***	-0.0480***
	(0.0060)	(0.0059)	(0.0059)
ZIP Code FE		✓	✓
Day of the Week FE			✓
# Obs	27,770	27,770	27,770
R^2	0.0029	0.1784	0.1787

Panel B: Credit Access Within 3 Years

Dep Var: Formal Credit Access (3 Years)	(1)	(2)	(3)
Post	-0.0601***	-0.0511***	-0.0519***
	(0.0061)	(0.0061)	(0.0061)
ZIP Code FE		✓	✓
Day of the Week FE			✓
# Obs	27,770	27,770	27,770
R^2	0.0032	0.1753	0.1758

Panel C: Credit Access Within 4 Years

Dep Var: Formal Credit Access (4 Years)	(1)	(2)	(3)
Post	-0.0560***	-0.0464***	-0.0468***
	(0.0062)	(0.0062)	(0.0062)
ZIP Code FE		✓	✓
Day of the Week FE			✓
# Obs	27,770	27,770	27,770
R^2	0.0027	0.1725	0.1730

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 (Panel A), three years (Panel B), or four years (Panel C) 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 (1) presents the baseline specification without fixed effects. Column (2) adds ZIP Code fixed effects. Column (3) adds day-of-week fixed effects. 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 D.2: Effect of Privacy Policy on Long-Term Credit Access: Robustness using 2SLS High-Frequency Identification

Panel A: Credit Access Within 2 Years			
Dep Var: Formal Credit Access (2 Years)	(1)	(2)	(3)
App Accepted	0.1138***	0.1129***	0.1125***
	(0.0122)	(0.0138)	(0.0138)
ZIP Code FE		✓	✓
Day of the Week FE			✓
# Obs	27,770	27,770	27,770
KP LM Statistic	560.5439	387.6932	388.3385
KP F Statistic	11863.4999	8591.9718	8029.3159
Panel B: Credit Access Within 3 Years			
Dep Var: Formal Credit Access (3 Years)	(1)	(2)	(3)
App Accepted	0.1229***	0.1193***	0.1217***
	(0.0125)	(0.0141)	(0.0143)
ZIP Code FE		✓	✓
Day of the Week FE			✓
# Obs	27,770	27,770	27,770
KP LM Statistic	560.5439	387.6932	388.3385
KP F Statistic	11863.4999	8591.9718	8029.3159
Panel C: Credit Access Within 4 Years			
Dep Var: Formal Credit Access (4 Years)	(1)	(2)	(3)
App Accepted	0.1145***	0.1083***	0.1098***
	(0.0126)	(0.0144)	(0.0145)
ZIP Code FE		✓	✓
Day of the Week FE			✓
# Obs	27,770	27,770	27,770
KP LM Statistic	560.5439	387.6932	388.3385
KP F Statistic	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 (Panel A), three years (Panel B), or four years (Panel C) 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 (1) presents the baseline specification without fixed effects. Column (2) adds ZIP Code fixed effects. Column (3) adds day-of-week fixed effects. The sample is restricted to Android users only, as iOS users were unaffected by the policy. KP LM Statistic and KP F Statistic test for underidentification and weak identification, respectively. Standard errors are clustered at the ZIP code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix E Framework

E.1 Alternative Assumption: Expenditure Shock

As a robustness check, we consider an alternative interpretation of the liquidity shock. Rather than a multiplicative reduction in period-0 income, we assume the borrower faces an unexpected required expenditure $e_i > 0$ in period 0, representing, for example, a medical bill, school fee, or equipment repair that must be paid regardless of whether a loan is obtained. Period-0 consumption is therefore $y_i - e_i + L_i$ if the borrower obtains a loan and $y_i - e_i$ otherwise, with income reverting to y_i in all subsequent repayment periods.

We assume $e_i = L_i$: the expenditure shock is exactly equal to the loan amount the borrower chooses to request. Under this assumption, the borrower without a loan would be unable to cover the required expenditure from income alone, making the loan's role purely to finance the shock. Period-0 consumption collapses to y_i with the loan and $y_i - L_i$ without it, and equations (3) and (4) become:

$$u_i^{\text{loan}} = \log(y_i) + \sum_{t=1}^{T_i} \delta^t \log\left(y_i - \frac{L_i(1+r_i)}{T_i}\right), \tag{E.1}$$

$$u_i^{\text{noloan}} = \log(y_i - L_i) + \sum_{t=1}^{T_i} \delta^t \log(y_i). \tag{E.2}$$

All downstream stages of the model are unchanged. This specification is more parsimonious than the baseline and requires no additional calibration beyond the observed loan amounts in the data.

Estimate of θ: Under the expenditure-shock specification, moment matching yields $\hat{\theta} = 0.1601$ (clustered SE = 0.0017; 95% CI [0.1569, 0.1634]), somewhat larger than the baseline estimate of 0.1334. The implied willingness to pay is ₹ 7,875, equivalent to 17.4% of mean monthly income, compared to ₹ 6,435 (14.3%) in the baseline.

Effect of Privacy Regulation on Consumer Surplus: The decomposition of the change in consumer surplus is qualitatively identical to the baseline. The privacy regulation raises consumer surplus by 0.60%, slightly above the baseline 0.53%. The privacy channel contributes +0.67%, reflecting the direct utility gain from removing lender access to call-detail records, which accrues to all potential borrowers independent of their credit risk or loan terms. The application margin contributes +0.01%: tighter privacy rules encourage additional borrowers to apply, and these marginal applicants earn a positive surplus evaluated at NoCDR approval rates. The approval margin contributes −0.08%: the loss of CDR-based enforcement leads the lender to screen more aggressively, reducing approval probabilities and thereby lowering expected surplus for applicants. As in the baseline, the privacy channel more than offsets the negative approval effect.

Effect of Privacy Regulation on Lender Profits: The privacy regulation reduces total lender profits by 12.4%, compared to 15.0% in the baseline. The decomposition preserves the same sign pattern. The application margin contributes +5.5%: the privacy regulation raises application rates, expanding the pool of potential borrowers and partially offsetting losses elsewhere. The approval margin contributes −7.0%: without CDR-based enforcement, the lender tightens screening and rejects more applicants to maintain portfolio quality. The surplus margin contributes −10.9%: even conditional on approval, per-loan surplus S_i^{NoCDR} is lower than S_i^{CDR} because the lender's default predictions are less accurate without CDR-based enforcement, leading to higher realized default rates on approved loans. Together, the approval and surplus margins sum to −17.9%, far exceeding the +5.5% application margin gain.

Table E.1: Parameter Estimates and Welfare Decomposition: Expenditure Shock Specification

Key Quantities	Estimate
Direct Utility Gain from Privacy ($\hat{\theta}$)	0.1601***
	(0.0017)
Lender's Required Rate of Return, ($\hat{\kappa}$)	0.1678***
	(0.0009)
(1) Percent Change in Consumer Surplus	
Total	+0.60%
Privacy Value	+0.67%
Application Margin	+0.01%
Approval Margin	-0.08%
(2) Percent Change in Lender Profits	
Total	-12.40%
Application Margin	+5.50%
Approval Margin	-7.00%
Surplus Margin	-10.90%

This table replicates Table 9 under an alternative expenditure-shock specification for the liquidity shock. Rather than a multiplicative income reduction, borrowers face an unexpected required expenditure equal to the loan amount requested. $\hat{\kappa}$ is estimated by MLE on pre- and post-policy applicants (equation (9)). $\hat{\theta}$ is estimated by exact moment matching to the observed 5.7% increase in Android applications (equation (20)). Changes in consumer surplus and lender profits are computed using equations (14) and (16), respectively, and expressed as percentages of the pre-policy (CDR) baseline. Standard errors reported in parentheses are clustered at the ZIP code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

E.2 Computing Standard Errors

We obtain standard errors for $\hat{\kappa}$ and $\hat{\theta}$ clustered by ZIP code to allow for within-location correlation in both approval decisions and application behavior. For $\hat{\kappa}$, the cluster-robust standard error is computed from the scores of the pooled log-likelihood. The score for observation i in cluster c is given by the following expression:

$$s_i(\kappa) = (y_i - \Lambda(S_i - \kappa L_i))\,(-L_i), \tag{E.3}$$

and the cluster-robust variance estimator is given by the following expression:

$$\widehat{SE}(\hat{\kappa}) = \frac{\sqrt{\sum_c S_c^2}}{|\mathcal{H}|}, \tag{E.4}$$

where $S_c = \sum_{i \in c} s_i(\hat{\kappa})$ is the cluster score sum and $\mathcal{H} = \sum_i \partial^2 \log L_i / \partial \kappa^2$ is the total Hessian evaluated at $\hat{\kappa}$.

For $\hat{\theta}$, define the individual moment contribution:

$$m_i(\theta) = \frac{\Lambda(V_i^{\text{NoCDR}} + \theta)}{\bar{P}^{\text{pre}}} - 1, \tag{E.5}$$

so that $N^{-1} \sum_i m_i(\hat{\theta}) = \tau$ at the solution. The cluster-robust delta method yields:

$$\widehat{SE}(\hat{\theta}) = \frac{\sqrt{\sum_c M_c^2 / N^2}}{\left|\partial g / \partial \theta\big|_{\hat{\theta}}\right|}, \tag{E.6}$$

where $M_c = \sum_{i \in c} m_i(\hat{\theta})$ is the cluster sum of moment contributions and the derivative is approximated by a symmetric finite difference with step size $h = 10^{-6}$.