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 data-access restriction. We document a key trade-off: stronger privacy protections increase loan applications, suggesting higher demand, but lead to tighter screening and reduced credit supply. This contraction disproportionately excludes economically and socially marginalized applicants. Using bureau records, we quantify the "FinTech ladder effect," whereby early digital credit access enables long-term borrowing. Privacy-driven rejections lower the likelihood of obtaining credit by 13.7 percentage-points even after four years. A structural model decomposes the welfare effects and shows that regulation raises consumer surplus by 0.23-0.60% but reduces lender profits by 20-23%.

*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, Julie Fu, Sean Higgins, Nicola Limodio, Muhammad Meki, Kevin Mei, Tyler Muir, Shumiao Ouyang, Stavros Panageas, Michaela Pagel, Amiyatosh Purnanandam, Avanidhar Subrahmanyam, Huan Tang, Margarita Tsoutsoura, 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 populations (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). 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 lifecycle. 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 has 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' CDRs following a Google privacy policy directive. Prior to this change, CDR data allowed the lender to use borrowers' social networks as a form of social collateral. For instance, in cases

[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.

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 relies on the parallel trends assumption: absent the policy, their lending outcomes would have evolved similarly. We assess this identifying assumption. We show that average lending outcomes exhibit parallel dynamics between the two groups during the pre-regulation period. 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 $\times$ loan-purpose $\times$ month-year and ZIP code $\times$ loan-purpose $\times$ 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 consists of 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 rise by approximately 26% relative to iOS users following the reform. This increase suggests that borrowers adjust their participation decisions once data-sharing requirements are relaxed.

On the lender side, we find a significant decline in the likelihood of application acceptance after the regulation. Acceptance rates among Android applicants decrease by 16–18 percentage points relative to iOS users, corresponding to roughly a 25% decline relative to the sample mean. Importantly, this reduction is not explained by compositional changes in applicant quality, as the income and credit score distributions 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 do 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.

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), helping borrowers build collateral (Beaumont, Tang, and Vansteenberghe, 2022), or building trust and familiarizing them with the banking sector, then rejection caused by privacy regulation may lead to systematic exclusion from formal credit markets.

We find that Android users who applied 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 to 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 decide whether to repay or default (repayment and recovery). Privacy influences the borrower's problem through two channels. First, it provides a direct utility benefit from withholding sensitive communication data. Second, it endogenously shapes credit access, which in turn affects the borrower's ability to smooth consumption. 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. The nature of this tension depends on whether borrowers fully internalize the lender's tighter screening when deciding to apply, so we estimate the model under both extremes – full and no internalization of the screening response – to bound the privacy value and the estimates for the effect of privacy regulation on consumer surplus and lender profits. We find that the direct utility gain from enhanced privacy is equivalent to 20.8%-36.3% of monthly income, implying that borrowers are willing to forgo roughly one-fifth to one-third 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 it into a privacy margin and a contract margin. The privacy regulation raises consumer surplus on net by 0.23-0.60%. The privacy margin increases the consumer surplus by 0.49-0.82%, reflecting the direct utility gain from removing lender access to CDRs. The contract margin reduces the consumer surplus by 0.22-0.26%, reflecting the utility loss due to reduced access to credit. The contract margin operates

through two channels: tighter screening reduces the probability of approval on the current loan (immediate access) and lowers the option value of future loans through the FinTech-ladder effect (long-term access). Specifically, 73% of the contract margin is driven by immediate access and 27% by long-term access. Together, these results imply that privacy benefits more than offset surplus losses from reduced credit access. This increase in consumer surplus is equivalent to a 0.02-0.12% permanent increase in monthly income.

Finally, the model shows that privacy regulation reduces total lender profits by 20.95-23.08%, decomposed into three channels. The application margin increases lender profits by 3.56-5.69% because higher application rates expand the pool of potential borrowers. The approval margin increases lender profits by 6.00%. This is because without CDR-based enforcement technology, the lender tightens its screening threshold and rejects applicants whose expected loan surplus is negative, thereby pruning loss-making loans from its portfolio and raising overall profitability. The per-loan surplus margin reduces lender profits by 32.64% as removal of CDR-based enforcement lowers expected recovery and raises expected losses on approved loans. Together, the losses due to loss in enforcement technology dominate the gains due to an increase in applications and reduction in unprofitable loans, resulting in an overall decline in lender profits.

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 several ways. 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. Furthermore, 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. However, lenders' profits decline, driven primarily by the loss of CDR-based enforcement rents through the per-loan surplus margin, with gains from higher application rates and tighter approval standards only partially offsetting this decline.

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 that allows 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

[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 a significant portion of first-time applicants. Another underexplored channel through which initial access to FinTech credit may raise future credit uptake is either by building trust or by familiarizing borrowers with the banking sector. See Guiso, Sapienza, and Zingales (2008); Karlan et al. (2009); Banerjee et al. (2015); Johnson, Meier, and Toubia (2019); Bachas et al. (2021), and Park, Sarkar, and Vats (2021), among others, for work on the role of trust in the banking sector.

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), which documents the 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. The regulation raises application rates, expanding the pool of potential borrowers and increasing lender profits by 3.6-5.7% through the application margin. The approval margin also adds to profits by 6.0% as lenders tighten screening and eliminate unprofitable loans. However, these gains are outweighed by large losses on the per-loan surplus margin, as the removal of CDR-based enforcement technology lowers recovery rates, reducing lender profits by 32.6%. 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 serves 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 and provides 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 ₹10,000 to ₹300,000 (roughly $120 to $3,600), with repayment tenures spanning from 15 days to 18 months.[3]

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 India's 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

[3]These bounds refer to the lender's standard product menu for approved and disbursed loans. Table 1 reports the full distribution of requested amounts, which occasionally fall below this floor because applicants may submit requests outside the standard product range.

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. Popular media has also discussed how FinTech lenders use alternative data to improve ex post enforcement in emerging economies; see, for example, New York Times (2017); Wall Street Journal (2019); Bloomberg (2023).

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 risk management 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 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.

An important nuance is that although the lender refreshed CDR metrics throughout the loan lifecycle, the enforcement power of the system rested primarily on the static threat created by the contacts captured at loan origination, which the lender retained even after the policy change. This is consistent with the absence of a deterioration in default rates among pre-policy Android cohorts whose CDR access was severed mid-loan (see Figure 4c). Because the lender already possessed the borrower's contact list from origination, the disciplinary threat remained credible for existing loans. For new post-policy loans, however, the lender could no longer capture any contact information at origination, eliminating the enforcement mechanism entirely. Thus, the ongoing monitoring capability augmented the baseline enforcement threat but was not necessary to sustain it for loans where contact data had already been collected.

We note that although our focal lender appears to have used CDRs primarily for ex post enforcement, the qualitative interpretation of our results is unlikely to depend on whether these data were used instead for ex ante screening. In both cases, the policy eliminates a source of comparative advantage that facilitated lending to borrowers who are difficult to assess using conventional data. If CDRs were chiefly an enforcement device, their removal would weaken repayment discipline and force the lender to tighten credit supply to protect portfolio quality. If they were chiefly a screening device, their removal would impair the lender's ability to identify creditworthy borrowers at origination. In either case, the implied consequence is a contraction in approvals and a reduction in formal credit access for marginal borrowers.

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 10% increase in the number of contacts in a borrower's contact list is associated with a 0.15 percentage point reduction in default probability in column 1. This effect increases to 0.17 percentage points when demographic and loan-level controls are included, representing approximately a 0.7-0.8% decline relative to the average default probability. The model's discriminatory power is substantial: the AUC rises from 0.52 in column 1 to 0.92 with the full set of fixed effects and demographic controls in column 5.[4]

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 10% increase in contact network size is associated with a 0.60–0.73 percentage point reduction in loss given default across specifications, representing approximately a 1.4-1.7% decline relative to the mean LGD. 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

[4]We note that the in-sample AUC of 0.92 is mechanically inflated by the high-dimensional fixed effects (ZIP code × loan purpose × month-year), which absorb substantial within-cell variation. This statistic should therefore be interpreted as a measure of in-sample fit conditional on these absorbing variables, rather than as an estimate of out-of-sample discriminatory power.

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 FinTech 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. Moreover, while the privacy regulation eliminated CDR access for new loans, it did not affect access for existing loans, since existing borrowers had already granted permission for those records.

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, and management had to shift its focus toward 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 gained 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, 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 also observe the operating system used for each application, 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.[5] 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

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

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 a credit score of approximately 659. The 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 predicted from surname-based classification following Dutta et al. (2021).[6] 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

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

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 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 define treatment and control groups. Specifically, we estimate the following baseline specification:

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

where $Y_{i,t}$ denotes 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 captures ZIP Code × loan purpose × month-year fixed effects. These fixed effects control for any shocks common to Android and iOS users within the same $\{z, p, t\}$ cell, e.g., local macro conditions, lender policy shifts, loan-purpose specific demand changes, etc., so identification comes only from Android–iOS differences within the same cell over time. The identifying variation comes from comparing Android to iOS users 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 does not capture 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. 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 a credible identification strategy 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 relies 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 Gelman and Imbens (2019) and Cattaneo, Idrobo, and Titiunik (2024). $\theta_{z,p}$ represents ZIP Code $\times$ 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, borrowers cannot systematically time their applications around the policy cutoff. This assumption rules out manipulation of the application date by individuals who would have applied anyway, for example, by strategically delaying an application to fall after the policy implementation date. The absence of advance public notice about the policy and the lender's minutes-to-hours processing times makes such strategic timing highly unlikely. By contrast, a sudden increase in the number of new applicants at the policy date, reflecting higher demand when CDR-sharing is removed, is compatible with our design and does not violate this assumption.[7]

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.

[7] In standard cross-sectional RDD applications, a discontinuity in the density of the running variable at the cutoff is often interpreted as evidence of manipulation of the assignment variable. In our setting, however, the running variable is calendar day, and the observed jump in daily application volume at the policy date should reflect an extensive-margin demand response, i.e., new applicants entering when CDR-sharing requirements are relaxed, rather than driven by strategic timing of existing borrowers. Because the policy was an unannounced surprise and loans are processed within minutes to hours, borrowers could not systematically shift application dates to just before or after the cutoff. Consequently, the density jump we document is not a McCrary-type violation of our assumption but an outcome of interest: it captures the causal effect of privacy protection on credit demand.

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. 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. The coefficient of interest, Android × 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%.[8] 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 respond positively to enhanced privacy protections, the policy simultaneously imposes substantial costs on lenders by eliminating access to social collateral that had been integral to its ex post enforcement technology. In this section, we document how the removal of CDRs from the lenders' enforcement toolkit affects 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

[8] Appendix Table B.1 shows that our results are robust even after balancing the panel at ZIP code × month × Loan purpose level and estimating the model using Poisson Pseudo-Maximum Likelihood (PPML) to accommodate zeros.

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 16.4 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 the credit supply in response to the loss of alternative data that previously served as a valuable enforcement tool through its social collateral value.

5.1.3 Effect of Privacy Policy on Portfolio Quality

A natural question is whether the lender's tighter screening in response to the loss of CDR-based enforcement merely preserves portfolio quality, actively deteriorates it, or actually improves ex post loan performance by selecting a more conservative pool of borrowers.[9]

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 percentage points. When we include the full set of applicant- and application-level controls in column (5), the magnitude of the estimate is -0.25 percentage points. 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.

[9]Note that while the privacy regulation eliminated CDR access for new loans, it did not affect access for existing loans; existing borrowers had already granted permission for those records.

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 $\times$ time $\times$ loan-purpose fixed effects (red diamonds); and (iii) ZIP-code $\times$ time $\times$ loan-purpose plus ZIP-code $\times$ loan-purpose $\times$ 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 induces 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.2 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 in some specifications, the magnitude of the effect is economically modest, corresponding to a 0.9% change relative to average annual rates near 40% in this market. Panel B of Appendix Table B.2 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 do 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-type 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 (120-day total window) around the policy date. These visual patterns are formalized in Table 5, which reports point estimates from our regression-discontinuity-type 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 aligns closely 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 28.5 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 (10 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 of our empirical design is that the privacy policy does not change the composition of Android applicants in a way that could explain the estimated decline in acceptance rates. If the policy were to shift applicant composition, the DiD estimates could confound changes in lending standards with changes in borrower quality. Table 6 addresses this concern by examining whether the applicant pool shifts toward lower-income borrowers, lower-credit-score borrowers, or borrowers without observable credit scores around the policy date.

The evidence in Panels A and B suggests that there is no statistically or economically significant deterioration in observable applicant quality: average income and average credit scores do not decline after

the policy. Panel C documents a modest shift in applicant composition, as the share of applicants without a credit score declines after the policy, but this pattern is not statistically robust and is economically small. Importantly, this compositional shift would, if anything, work against our main finding: a smaller share of unscored applicants raises the observable quality of the applicant pool, making it harder to detect a post-policy decline in acceptance rates. Overall, the evidence is more consistent with a genuine contraction in credit supply induced by the loss of CDR-based enforcement than with a deterioration in applicant quality.

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. In response, lenders raise acceptance thresholds and reduce approval probabilities for applicants whose risk would have previously 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 operate 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 fully offset the loss of enforcement capacity by tightening screening. A natural question is who bears the cost of this adjustment. If the lender raised approval thresholds uniformly, the contraction in credit supply would be spread evenly across applicants. If, instead, the lender substituted screening for enforcement selectively, tightening most against borrowers whose risk was previously managed through CDR-based discipline, the burden would fall disproportionately on applicants with thin credit profiles. This section tests between these alternatives by examining how the policy differentially affected acceptance rates across applicant subgroups within the treated Android population.

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 × loan-purpose, ZIP-code × time (week), loan-purpose × time (week), and application date fixed effects. The ZIP-code × 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 × week and loan-purpose × 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. Date fixed effects remove systematic daily patterns in application and approval behavior, ensuring that our subgroup comparisons are not driven by mechanical 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 approximately 6.6 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 7.7 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 approximately 25.4 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 approximately 8.1 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 scores. 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 7 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. 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 exacerbates exclusion within the digital lending market. The loss of private alternative data most strongly affects borrowers whose creditworthiness cannot be easily assessed using traditional information sources. In short, the CDR ban disproportionately constrains 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 8 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 8 are binary indicators equal to one if the borrower accessed any form of formal credit including loans from banks, non-bank financial companies, and other FinTech platforms within 2, 3, or 4 years respectively.

The key coefficient of interest is Android × Post, which captures whether Android users who applied after the policy change were differentially less likely to access future credit compared to iOS users. The results reveal persistent negative effects on credit access that extend far beyond the immediate policy effect. Focusing on the most demanding specifications 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 34% over two years in our sample, a 2.1 percentage point reduction represents approximately a 6.2% 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$_{i,s}$ is a binary indicator equal to one if borrower i accessed formal credit within s years of January 1, 2019 (we examine horizons of 2, 3, and 4 years). Because all applicants in our sample applied within a narrow window around the policy date, this calendar anchor closely approximates a rolling window measured from each borrower's application date. 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 9 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. Panel A of Table 4 shows a sizable and precisely estimated first stage: the policy lowers Android approval rates by roughly 15.6–17.8 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 9 reports a Kleibergen–Paap rk LM statistic over 100, rejecting underidentification in all columns, and a Kleibergen–Paap rk Wald F-statistic above 1000 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.

To assess whether the FinTech ladder effect is confounded by a concurrent tightening of screening standards across the broader FinTech sector, we re-estimate both the reduced-form and 2SLS specifications after restricting the credit access outcome to traditional lenders that were plausibly unaffected by Google's privacy policy. Under the alternative hypothesis of a broad-based contraction in FinTech credit supply, excluding FinTech lenders from the outcome should attenuate or eliminate the estimated effects. Appendix Tables D.2 and D.3 present these results. The point estimates are, if anything, marginally larger across all horizons and specifications, indicating that the FinTech ladder effect operates primarily through traditional lenders and is not a byproduct of a uniform contraction in FinTech credit supply driven by privacy regulation.

The second-stage estimates presented in Table 9 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 in column (4)—ZIP code × loan purpose × month-year and ZIP code × loan purpose ×

Android fixed effects—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).

Because the IV estimate is mechanically equal to the reduced-form effect scaled by the inverse of the first stage, the IV magnitudes are necessarily larger than the reduced-form coefficients; this proportional scaling holds regardless of treatment-effect heterogeneity. More informative is the comparison with OLS: the IV estimates substantially exceed their OLS counterparts, consistent with a local average treatment effect (LATE) interpretation in which compliers, borrowers whose approval status was changed by the policy, experience larger returns to initial credit access than the average approved borrower. 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.4 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 (Guiso, Sapienza, and Zingales, 2008; Karlan et al., 2009; Banerjee et al., 2015; Johnson, Meier, and Toubia, 2019; Bachas et al., 2021; 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 a significant portion of 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 identify the direction of both the demand and supply responses to the privacy regulation but leave three quantitative questions unanswered. First, how much do borrowers value the privacy afforded by the removal of CDR access? Second, does the direct utility gain from enhanced privacy outweigh the indirect cost of reduced credit access, so that consumer surplus rises on net? Third, what are the net profits or losses borne by the lender, and through which margins – application, approval, or per-loan surplus – do they operate?

To address these questions, we develop a three-stage structural model of credit demand, lender screening, and repayment, based on the model in Adams, Einav, and Levin (2009), Einav, Jenkins, and Levin (2012), and Tang (2019). The central structural parameter is θ, the direct utility gain borrowers derive from enhanced data privacy, identified by exploiting a tension between the model's predictions and the data. The nature of this tension depends on whether borrowers fully internalize the lender's screening response when deciding to apply.

Under *full internalization*, borrowers correctly perceive that CDR removal worsens the lending environment. Absent any privacy value, the model therefore predicts that applications should have fallen after the policy. However, applications among Android users instead increase 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 effect of tighter screening on applications and match the observed increase in applications.

Under *no internalization*, borrowers do not update their beliefs about the post-policy screening environment and continue to apply as if pre-policy approval probabilities still prevail. In this case, the model predicts no change in applications at $\theta = 0$. The parameter θ is then identified as the increment to borrowers' utility that generates the observed 5.7% increase above the pre-policy baseline. Because borrowers under no internalization do not account for the deterioration in the lending conditions, the estimated θ is smaller than under full internalization, providing lower and upper bounds on the privacy value.

In both cases, θ is identified as a residual: it is the component of borrowers' willingness to apply that cannot be rationalized by lending conditions alone. Overall, this exercise allows us to quantify the

lower and upper bounds on: (1) privacy valuation, and (2) the effect of privacy regulation on consumer surplus and lender profits.

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. 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 income shock that renders the loan individually rational, modeled as a multiplicative reduction in period-0 income so that realized income is $(1 - \varepsilon_i)y_i$, where y_i is the borrower's monthly income. Under the assumption that the loan amount an applicant requests is optimal, we recover ε_i by inverting the borrower's first-order condition for optimal loan demand (equation (4) in Section 6.2). The privacy regime $s \in \{0, 1\}$ captures whether the lender has access to CDRs and is a publicly observed state variable.[10]

The interaction unfolds across three sequential stages, which we solve recursively, starting from Stage 3 and working backward: 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 idiosyncratic income shock ε_i, the prevailing privacy regime s, and the posted lending terms, and decides whether to apply. Upon application, she selects a loan amount L_i, where the repayment horizon in months T_i and the contracted interest rate r_i are taken as exogenous.

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 δ. The borrower faces the multiplicative income shock ε_i in period 0, so that realized period-0 income is

[10] 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.

$(1-\varepsilon_i)y_i$; income reverts to y_i in all subsequent periods. The utility from obtaining a loan (u_i^{loan}) and the utility from not borrowing (u_i^{noloan}) are:

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

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

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, $\frac{L_i(1+r_i)}{T_i}$ 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 $t = 1, \ldots, T_i$. Lastly, B_i indicates the future loan value that captures the FinTech ladder effect: a borrower who obtains a loan today is more likely to secure a loan in the future, defined in equation (3).

Future Loan Value. We define B_i as:

$$B_i = \Delta \cdot p_{\text{shock}} \cdot \sum_{k=1}^{\infty} \beta^k \cdot v_i^f \tag{3}$$

where $\beta = \delta^{12}$ is the annual discount factor, p_{shock} is the probability of a future liquidity shock, Δ is the increase in the probability of securing a future loan attributable to having established a credit relationship today which we refer to as the FinTech-ladder effect, and v_i^f is the utility gain from accessing that future loan. Conditional on a shock arriving, the borrower experiences a reference income shortfall of ε^f and accesses a future loan of size $L^f = \varepsilon^f y_i$ wherein the loan amount equals the income loss, with tenure T^f and total interest rate r^f. The utility gain v_i^f is defined analogously to the current-period loan surplus:

$$v_i^f = \log\left[(1-\varepsilon^f)y_i + L^f\right] + D^f \log\left(y_i - M^f\right) - \log\left[(1-\varepsilon^f)y_i\right] - D^f \log(y_i)$$

where $M^f = \frac{L^f(1+r^f)}{T^f}$ is the future monthly installment and $D^f = \frac{\delta(1-\delta^{T^f})}{1-\delta}$ is the corresponding discounted tenure sum.[11]

Inversion of the Liquidity Shock: Given the observed contract $\{L_i, T_i, r_i\}$ and income y_i, we recover ε_i via a closed-form inversion of the borrower's optimal loan demand condition. Under log utility, the

[11] The future loan is assumed to arrive no earlier than the following year, consistent with the fact that the majority of loan products in our sample have maturities of less than one year.

first-order condition for L_i is:

$$\frac{1}{(1-\varepsilon_i)y_i + L_i} = D_i \cdot \frac{c_i}{y_i - c_i L_i}$$

where $c_i = \frac{1+r_i}{T_i}$ is the installment-to-loan ratio and $D_i = \frac{\delta(1-\delta^{T_i})}{1-\delta}$ is the discounted tenure sum. Inverting the first-order condition for L_i yields the following expression:

$$\hat{\varepsilon}_i = 1 - \frac{1}{y_i} \cdot \left(\frac{y_i - c_i L_i}{D_i c_i} - L_i \right) \tag{4}$$

the unique income shock consistent with the observed loan amount L_i being the borrower's optimal choice.[12]

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 receives a utility of 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 given by the following expression:

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

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

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

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. Each alternative (apply, not apply) receives an independent Type-I Extreme Value taste shock; the borrower applies if the draw for entering the credit market exceeds that for the outside option, yielding the logistic application probability:

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

[12]For borrowers with $\hat{\varepsilon}_i \geq 1$, whose period-0 income is fully eliminated by the shock, we apply a consumption floor $\underline{c}$, reflecting government safety nets. In our sample, 30.4% of borrowers have $\hat{\varepsilon}_i \geq 1$ while the remaining 69.6% have $\hat{\varepsilon}_i \in (0, 1)$. The mean inverted shock is 0.857 and the median is 0.801.

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:

$$S_i^s = (1 - \widehat{PD}_i^s)\, R_i + \hat{\lambda}_i^s\, \widehat{PD}_i^s\, L_i - (1+\kappa) L_i \tag{5}$$

where $R_i = L_i(1 + \frac{r_i}{T_i})^{12}$ 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 $S_i^s + \epsilon_i \geq 0$, where ϵ_i is the difference between two independent Type-I Extreme Value shocks capturing 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:

$$\Pr(\text{Approve}_i \mid s) = \Lambda(S_i^s)$$

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 (5), recomputing S_i^{NoCDR}, and evaluating $\Lambda(S_i^{\text{NoCDR}})$.

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 evaluating 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 $Z_i^{\text{CDR}} =$

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

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

where μ_p, μ_t, and μ_q denote ZIP code, year-month, and loan purpose fixed effects. $\Lambda(x) = \frac{e^x}{1+e^x}$ 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. We denote this counterfactual as $\widehat{PD}_i^{\text{NoCDR}}$. For post-policy loans, the lender's information set excludes CDRs entirely, so we estimate the following specification which serves as the relevant credit risk measure in Stage 2.

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

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

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

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 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, which yields $\hat{\lambda}_i^{\text{post}}$.

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

6.5 Consumer Surplus

A key modeling choice concerns borrowers' expectations about the post-policy lending environment. In the CDR regime, there is no ambiguity: borrowers correctly internalize the lender's screening rule, so

consumer surplus is given by the standard logit inclusive value presented in the following equation where $V_i^{\text{CDR}} = \hat{P}_i^{A,\text{CDR}}(u_i^{\text{loan}} - u_i^{\text{noloan}})$.

$$CS_i^{\text{CDR}} = \log\Big(\exp\Big(u_i^{\text{noloan}}\Big) + \exp\Big(u_i^{\text{noloan}} + V_i^{\text{CDR}}\Big)\Big)$$

After the policy change, however, borrowers may not immediately recognize that the lender now faces higher predicted default probabilities and lower recovery rates. The key modeling choice is therefore the degree to which borrowers update their beliefs about the post-policy screening environment. Under *full internalization*, borrowers fully account for the lender's updated risk assessments, so perceived and actual approval probabilities coincide and consumer surplus retains the logit inclusive value form:

$$CS_i^{\text{NoCDR, FA}} = \log\Big(\exp\Big(u_i^{\text{noloan}}\Big) + \exp\Big(u_i^{\text{noloan}} + \hat{P}_i^{A,\text{NoCDR}} \cdot \big(u_i^{\text{loan}} - u_i^{\text{noloan}}\big) + \theta\Big)\Big) \tag{6}$$

Under *no internalization*, borrowers apply based on pre-policy beliefs, using V_i^{CDR} to form their application decision:

$$P_i^{\text{apply}} = \Lambda\Big(V_i^{\text{CDR}} + \theta\Big)$$

while the lender evaluates applications using the post-policy surplus S_i^{NoCDR}. Because the borrower's application decision and the lender's approval decision are governed by different probabilities, the log-sum-exp formula no longer applies, and consumer surplus is instead the expected payoff evaluated at the borrower's mistaken beliefs:

$$CS_i^{\text{NoCDR, NA}} = u_i^{\text{noloan}} + P_i^{\text{apply}} \cdot \Big(V_i^{\text{NoCDR}} + \theta\Big) + H\Big(P_i^{\text{apply}}\Big) \tag{7}$$

where $H(p) = -p \log p - (1-p)\log(1-p)$ is the binary entropy arising from integrating out the Type-I Extreme Value taste shock over the borrower's mistaken decision rule. Because the true degree of borrower sophistication lies somewhere between these two extremes, we estimate the model under each assumption and use the results to bound key outcomes: privacy valuation, consumer surplus, and lender profits.

We decompose $\Delta CS_i = CS_i^{\text{NoCDR}} - CS_i^{\text{CDR}}$ using an intermediate counterfactual CS_i^{CF} in which borrowers receive the privacy benefit θ but still face the pre-policy approval environment:

$$CS_i^{\text{CF}} = \log\Big(\exp\Big(u_i^{\text{noloan}}\Big) + \exp\Big(u_i^{\text{noloan}} + V_i^{\text{CDR}} + \theta\Big)\Big)$$

The total change then decomposes as:

$$\Delta CS_i = \underbrace{CS_i^{\text{CF}} - CS_i^{\text{CDR}}}_{\text{privacy margin}} + \underbrace{CS_i^{\text{NoCDR}} - CS_i^{\text{CF}}}_{\text{contract margin}}$$

Our decomposition of the change in consumer surplus shows that privacy regulation affects borrowers through two distinct channels. The privacy margin isolates the direct utility gain from stronger data protection, holding the approval environment fixed at its pre-policy level. This margin is positive and is larger for borrowers with higher application propensities. The contract margin captures the change in the approval environment: the loss of CDR-based enforcement tightens screening, reduces the probability that an application is approved, and thereby lowers consumer surplus along this channel. The overall sign of ΔCS_i depends on which margin dominates for each borrower.

Decomposing the contract margin. The contract margin reflects two economically distinct sources of welfare loss from tighter screening. First, borrowers lose immediate access to credit: the current loan is less likely to be approved, reducing the direct consumption-smoothing value of the loan. Second, borrowers lose long-term access: tighter screening today reduces the probability of establishing a credit relationship with the lender, lowering the option value of future loans through the FinTech-ladder effect captured by B_i. To separate these channels, we construct an additional counterfactual in which the future loan bonus is shut down by setting $\Delta = 0$ in equation (3), so that $B_i = 0$ for all borrowers. Appendix Section E.1 provides additional details on the construction of this counterfactual and the decomposition of the contract margin into its two components.

Equilvalent Variation. We follow Dávila and Schaab (2025) and translate the change in consumer surplus into a scale-invariant measure. Specifically, we convert ΔCS_i into a equivalent variation, i.e., the permanent increase in monthly income that produces the same utility gain as the CDR ban. Appendix E.2 presents the details of this computation.

6.6 Lender Profits

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

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

where $S_i^s = (1 - \widehat{PD}_i^s) R_i + \hat{\lambda}_i^s \widehat{PD}_i^s L_i - (1 + \kappa) L_i$ is the per-loan surplus. Under no internalization, $P_i^{\text{apply,NoCDR}}$ uses pre-policy beliefs about default probabilities and recovery rates based on borrower's

belief, while $\hat{P}_i^{A,\text{NoCDR}}$ is evaluated at their post-policy counterparts which reflects lender's actual decision. The effect of the privacy policy on lender profits, $\Delta\Pi = \Pi^{\text{NoCDR}} - \Pi^{\text{CDR}}$, decomposes as:

$$\Delta\Pi_i = \underbrace{(P_i^{\text{apply,NoCDR}} - P_i^{\text{apply,CDR}}) \cdot \hat{P}_i^{A,\text{NoCDR}} \cdot S_i^{\text{NoCDR}}}_{\text{application margin}} + \underbrace{P_i^{\text{apply,CDR}} \cdot (\hat{P}_i^{A,\text{NoCDR}} - \hat{P}_i^{A,\text{CDR}}) \cdot S_i^{\text{NoCDR}}}_{\text{approval margin}}$$
$$+ \underbrace{P_i^{\text{apply,CDR}} \cdot \hat{P}_i^{A,\text{CDR}} \cdot (S_i^{\text{NoCDR}} - S_i^{\text{CDR}})}_{\text{per-loan surplus margin}}$$

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 per-loan surplus margin, which is negative as the loss of enforcement technology reduces the amount of surplus for the lender to capture.

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.2 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:

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

Under the NoCDR regime, the aggregate application rate depends on the internalization assumption. Under no internalization, borrowers apply based on pre-policy beliefs. The aggregate application rate is given by the following expression:

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

Under full internalization, borrowers completely account for the lender's tighter screening standards and the aggregate application rate in this case is given by the following expression:

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

In both cases, θ is identified by matching the model-predicted percentage change in application rates to the observed 5.7% increase:

$$g(\theta) = \frac{\bar{P}^{\text{post}}(\theta) - \bar{P}^{\text{pre}}}{\bar{P}^{\text{pre}}} = 0.057 \tag{8}$$

The identification logic is as follows. With $\theta = 0$, the model predicts that application rates should have *fallen* (remained the same) after the policy under the full internalization (no internalization) assumption, because the loss of CDR-based enforcement deteriorates lending conditions: $\Delta_i = V_i^{\text{NoCDR}} - V_i^{\text{CDR}} < 0$ on average, so $\bar{P}^{\text{post}}(0) < \bar{P}^{\text{pre}}$ in the case of full internalization and $\bar{P}^{\text{post}}(0) - \bar{P}^{\text{pre}} = 0$ in the case of no internalization. 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. Under full internalization, it is the premium required to overturn the negative contract-terms effect; under no internalization, it is the increment above the pre-policy baseline. Moreover, the privacy value required to match the observed increase in application rates will be higher in the full internalization case relative to the no internalization case. The two assumptions therefore deliver lower and upper bounds on the privacy value.

6.8 Calibration & Estimation

Discount factor (δ). We set the monthly discount factor to $\delta = 0.99$, corresponding to an annual discount factor of $\delta^{12} \approx 0.886$.

Consumption floor ($\underline{c}$). For borrowers whose period-0 income is fully eliminated by the liquidity shock ($\hat{\varepsilon}_i \geq 1$), we apply a consumption floor of $\underline{c} = ₹\ 2000$, reflecting government safety nets available to low-income or unemployed households. We verify in Appendix Section E.5 that our key estimates are robust to the choice of this floor.

Future loan parameters. The reference income shortfall is set to $\varepsilon^f = 0.8$, corresponding to the median of the inverted liquidity shock distribution in our estimation sample. For simplicity, the future loan is assumed to carry a one-year tenure ($T^f = 12$ months). The total interest rate is set to $r^f = 0.3$, consistent with a monthly interest rate of 2.5%. The future loan shock probability is set to $p_{\text{shock}} = 0.5$. We conduct sensitivity analysis on p_{shock} in Appendix Section E.5, showing that the key estimates of θ and the welfare decomposition are robust to alternative values.

FinTech-ladder effect parameter (Δ). We set $\Delta = 0.137$, taken directly from Panel C, column (4) of Table 9, which estimates the increase in the probability of obtaining a future loan attributable to having established a credit relationship through the current loan.

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

$$\hat{\kappa} = \arg\max_{\kappa} \sum_i \left[\text{Approve}_i \cdot \log \Lambda(S_i) + (1 - \text{Approve}_i) \cdot \log(1 - \Lambda(S_i)) \right. \tag{9}$$

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.3.

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 10 presents the parameter estimates for θ and κ, along with

the quantitative welfare analysis, which shows that the privacy regulation increases consumer surplus by 0.23-0.60% but reduces lender profits by 20-23%.

Estimate of κ. Maximum likelihood yields $\hat{\kappa} = 0.1832$, estimated jointly on pre- and post-policy loan applications under the assumption of a common required rate of return across regimes. We direct readers to Appendix Section E.4.1 for further details on the identification of κ. The estimator implies model-predicted approval probabilities of 85.1% in the CDR period and 53.2% in the NoCDR period. These numbers are broadly consistent with empirically observed approval rates of 82.0% in the CDR period and 60.8% in the NoCDR period. The model captures the direction and broad magnitude of the approval decline reasonably well and supports the ability of our parsimonious model to capture key empirical patterns in the data.

Estimate of θ. Under no internalization, matching the model moment $g(\theta)$ to the observed 5.7% increase in Android applications yields $\hat{\theta} = 0.1891$. Under full internalization, borrowers fully account for the post-policy approval environment and must be compensated by a larger privacy premium to rationalize the same application increase, yielding $\hat{\theta} = 0.3095$. We direct readers to Appendix Section E.4.2 for further details on the identification of θ. Moreover, Appendix Figure E.3 shows that θ is identified from the steeper part of the logit cumulative distribution function (CDF), suggesting that the estimate is not artificially inflated by applicants with structural surplus configurations that place them on the upper, flatter portion of the logit CDF.

Since θ is measured in log-utility units, it is not directly interpretable in economic terms like 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 $y_i \left(e^{\hat{\theta}} - 1\right)$.[13] Economically, this corresponds to the following counterfactual thought experiment: "What one-time transfer would fully compensate this borrower for losing the privacy protection?" Alternatively, this estimate of τ can also be interpreted as the willingness to pay for privacy.

Across the two internalization assumptions, the implied privacy value is substantial: borrowers are willing to forgo roughly one-fifth to one-third of a month's income to keep their call records private from their lender. Specifically, the certainty-equivalent income transfer, the lump-sum payment that would leave a borrower just as well off as the privacy reform, averages ₹ 9,368, which corresponds

[13] 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.

to 20.8% of monthly income under no internalization. Under the full internalization assumption, the certainty-equivalent income transfer is ₹ 16,325, corresponding to 36.3% of monthly income. These two estimates provide the lower and upper bounds on willingness to pay for privacy, respectively.

6.9.1 Effect of Privacy Regulation on Consumer Surplus

Our decomposition of the change in consumer surplus shows that the privacy regulation raises consumer surplus 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 +0.23% (+0.60%) under no (full) internalization. Specifically, under no (full) internalization, the privacy margin contributes +0.49% (+0.82%), reflecting the direct utility gain from removing lender access to call-detail records. The contract margin contributes −0.26% (−0.22%) under no (full) internalization.

The contract margin channel operates through the lender's response to the privacy regulation, as well as the effect on long-term credit access through the FinTech ladder effect. The loss of CDR-based enforcement leads the lender to screen more aggressively, reducing approval probabilities and thereby lowering the expected surplus from entering the credit market and the option value of long-term credit access. We decompose the contract margin into the immediate and long-term access components as described in Appendix E.1. Under no (full) internalization, the immediate access component accounts for -0.19% (−0.16%), reflecting the direct reduction in approval probabilities for the current loan. The long-term access component contributes −0.07% (−0.06%), capturing the surplus loss from the reduced probability of establishing a credit relationship and thereby forgoing the FinTech-ladder effect. The immediate access channel dominates, accounting for roughly three-quarters of the total contract margin loss, while the long-term access channel accounts for the remainder. Overall, 73% of the contract margin is driven by immediate access and 27% by long-term access.

Table 10 also reports the equivalent variation (EV) for each component, expressed as a percentage of mean monthly income, which translates the change in expected utility into a monetary equivalent. Under no (full) internalization, the total EV is +0.02% (+0.12%) of monthly income. The privacy margin alone contributes +0.13% (+0.21%), while the contract margin contributes −0.11% (-0.09%), of which −0.08% (-0.07%) stems from reduced immediate credit access and −0.03% (-0.02%) from the erosion of long-term access via the FinTech-ladder effect.

Taken together, these results imply that the privacy margin more than offsets the contract margin, indicating that borrowers place substantial value on data privacy beyond its role in facilitating credit access.

6.9.2 Effect of Privacy Regulation on Lender Profits

The privacy regulation reduces total lender profits by 23.08% under no internalization and 20.95% under full internalization. Under no (full) internalization, the application margin contributes +3.56% (+5.69%): the privacy regulation raises application rates, expanding the pool of potential borrowers and partially offsetting the other losses. The approval margin contributes +6.00% (+6.00%): without CDR-based enforcement technology, the lender tightens its screening threshold and rejects applicants whose expected loan surplus is negative, thereby pruning loss-making loans from its portfolio and raising overall profitability. The per-loan surplus margin contributes −32.64% (−32.64%): removal of CDR-based enforcement reduces the lender's ability to discipline borrowers, lowering expected recovery and raising expected losses on approved loans. Together, the per-loan surplus margin exceeds the combined gains from the approval and application margins, resulting in an overall decline in lender profits.

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 segments FinTech lending was expected to bring into the formal financial system. These effects are also persistent over time. 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 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, from a consumer perspective, the estimated benefits of privacy outweigh these costs, with consumer surplus rising by +0.23% to +0.60%.

The model quantifies the effect of privacy regulation on lender profitability, showing that total

profits decline by 20.95-23.08%. This reduction reflects opposing forces across three margins. The regulation raises application rates, expanding the pool of potential borrowers and contributing modestly positive effects through the application margin. The approval margin also adds to profits as lenders tighten screening and eliminate unprofitable loans. However, these gains are outweighed by large losses on the per-loan surplus margin, as the removal of CDR-based enforcement technology lowers recovery rates. Consequently, the erosion of enforcement rents dominates, producing an overall decline in lender profitability.

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.

Adams, William, Liran Einav, and Jonathan Levin. 2009. "Liquidity constraints and imperfect information in subprime lending." *American Economic Review* 99 (1):49–84.

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).

Bachas, Pierre, Paul Gertler, Sean Higgins, and Enrique Seira. 2021. "How debit cards enable the poor to save more." *The Journal of finance* 76 (4):1913–1957.

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.

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)* .

Dávila, Eduardo and Andreas Schaab. 2025. "Welfare Assessments with Heterogeneous Individuals." *Journal of Political Economy* 133 (9):2918–2961.

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* .

Dutta, Harsha, Pulak Ghosh, Arkodipta Sarkar, and Nishant Vats. 2021. "Political power-sharing, firm entry, and economic growth: Evidence from multiple elected representatives." *Firm Entry, and Economic Growth: Evidence from Multiple Elected Representatives (November 21, 2021)* .

Einav, Liran, Mark Jenkins, and Jonathan Levin. 2012. "Contract pricing in consumer credit markets." *Econometrica* 80 (4):1387–1432.

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.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2008. "Trusting the stock market." *the Journal of Finance* 63 (6):2557–2600.

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.

Johnson, Eric J, Stephan Meier, and Olivier Toubia. 2019. "What's the catch? Suspicion of bank motives and sluggish refinancing." *The Review of Financial Studies* 32 (2):467–495.

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* .

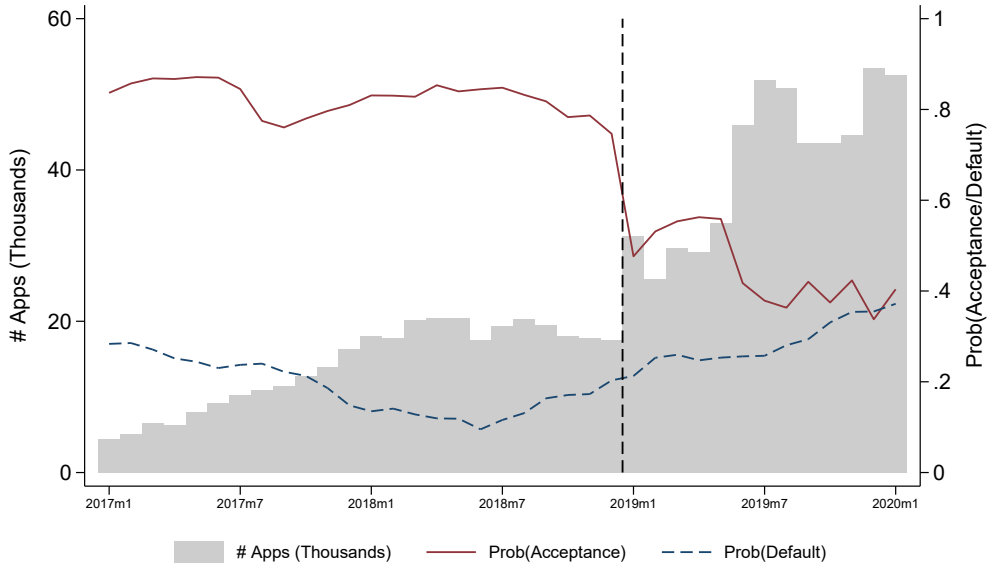
Thakor, Anjan V. 2020. "Fintech and banking: What do we know?" *Journal of financial intermediation* 41:100833.

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.

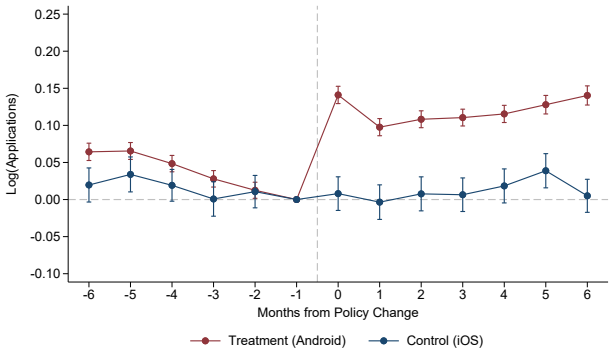
Wall Street Journal. 2019. "Silicon Valley-Backed App Lenders Use Phone Data to Hassle Borrowers." Available here (Accessed: March 17, 2026).

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

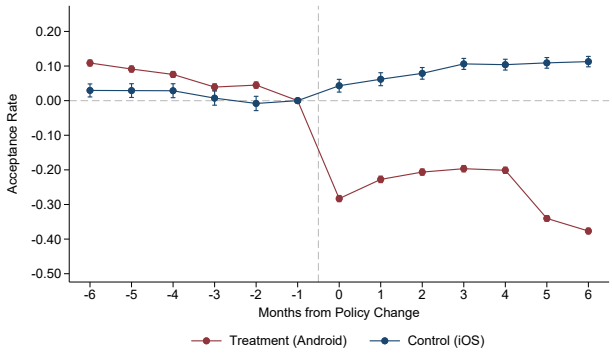


This figure illustrates the number of monthly applications (thousands, left axis), the monthly average probability of acceptance, the monthly average probability of default (right axis) from January 2017 until January 2020.

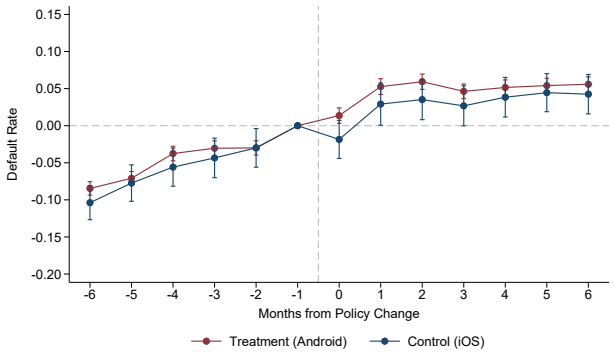
Figure 2: Unconditional Means: Treatment and Control Group Trends



(a) Number of Applications



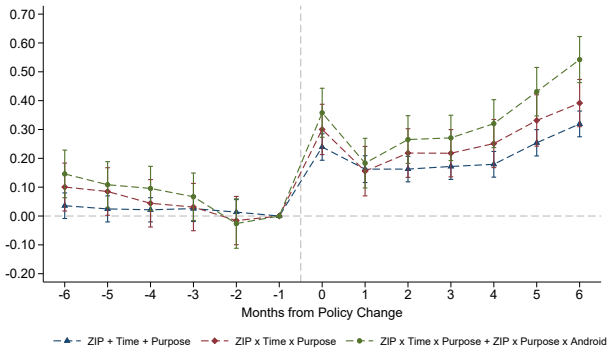
(b) Likelihood of Application Acceptance



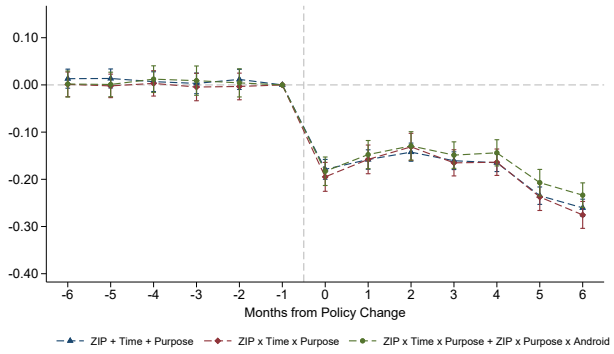
(c) Likelihood of Default

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 monthly 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.

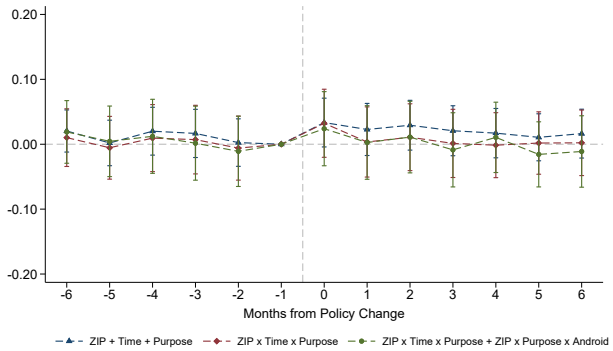
Figure 3: Assessment of Pre-Trends



(a) Number of Applications



(b) Likelihood of Application Acceptance



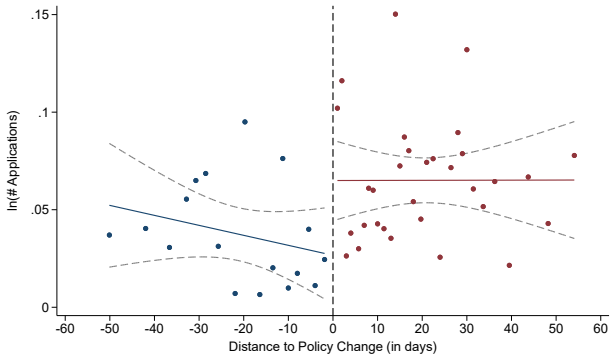
(c) Likelihood of Default

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:

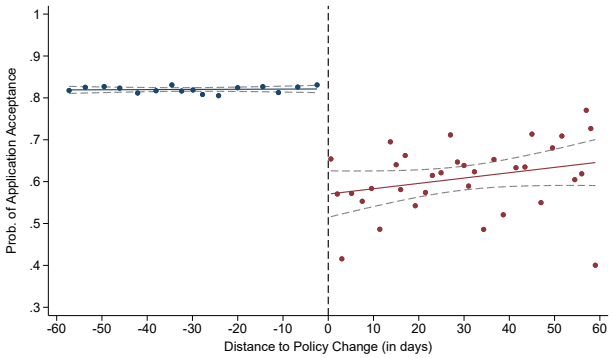
$$Y_{i,t} = \alpha + \sum_{k=-6, k\neq -1}^{k=+6} \beta_k \cdot (Android_i \times \mathbb{1}_{t=k}) + \theta_{z,p,t} + \theta_{z,p} \times Android_i + \epsilon_{i,t}$$

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 × Time (Month-Year) × Loan Purpose fixed effects ($\theta_{z,p,t}$) and ZIP Code × Loan Purpose × 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 × Time × Loan Purpose fixed effects. The green line (circle marker) includes both ZIP Code × Time × Loan Purpose and ZIP Code × Loan Purpose × 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.

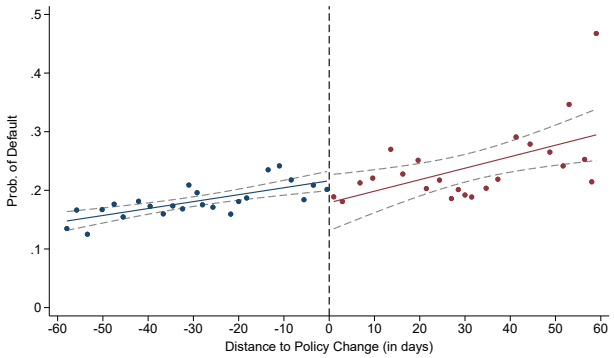
Figure 4: Main Effects of Privacy Policy Change: High-Frequency Event Study



(a) Number of Applications



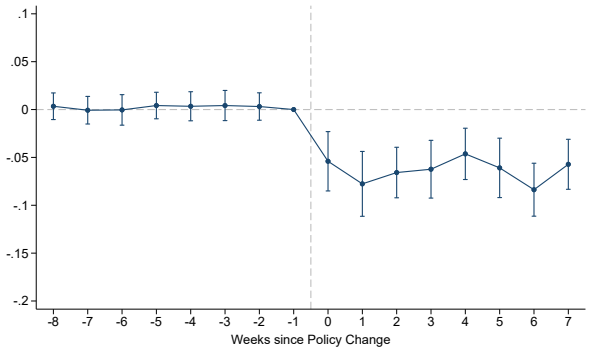
(b) Likelihood of Application Acceptance



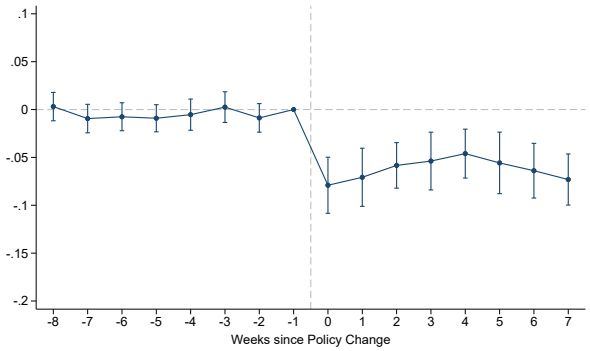
(c) Likelihood of Default

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 ZIP code-by-loan purpose-by-day in Panel (a), an application in Panel (b), and an individual loan, or accepted application, in Panel (c). The sample is restricted to Android users only, as iOS users were unaffected by the policy.

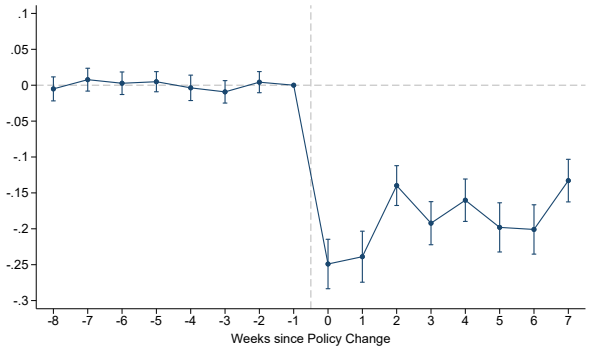
Figure 5: DiD: Heterogeneous Effects on Application Acceptance by Borrower Characteristics



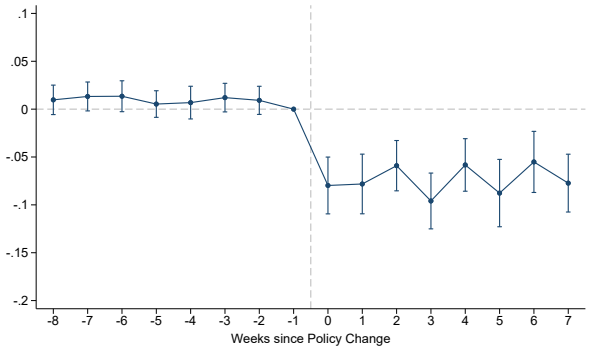
(a) Monthly Borrower Income



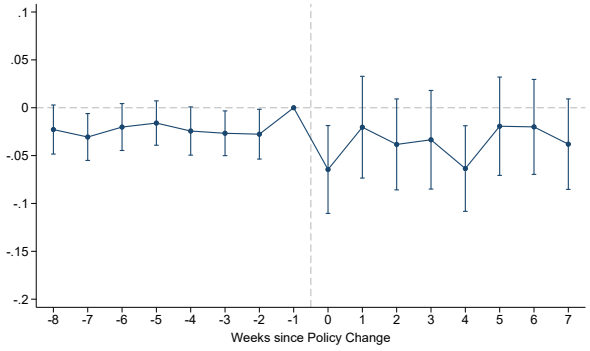
(b) Age



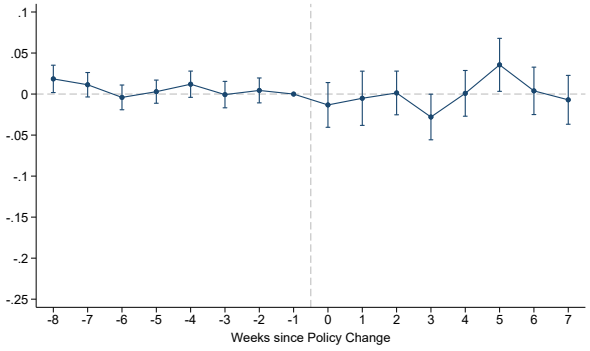
(c) First-Time Applicant



(d) Caste



(e) Gender



(f) Credit Scores

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:

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

All regressions include ZIP Code × loan purpose ($\theta_{z,p}$), ZIP Code × week ($\theta_{z,t}$), loan purpose × week ($\theta_{p,t}$), and date of application (θ_d) 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 a 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

	# Obs	p25	p50	p75	Mean	Std. Dev.
Panel A: Characteristics of Loan Applicants						
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
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) for Panels A-C. 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. For each applicant in the 180-day sample, we merge to an external credit bureau dataset and identify whether the applicant obtained any formal credit (from any lender, not just focal FinTech) within two, three, or four years of January 1, 2019. The indicators thus capture extensive-margin credit access in the broader financial system following the policy change. Monthly income and loan amount are reported in Indian Rupees (₹). Default is defined as 90+ days past due. Low caste indicates membership in the 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: $\mathbb{1}$[Default]	(1)	(2)	(3)	(4)	(5)
ln(Num. Contacts)	-0.0145***	-0.0133***	-0.0136***	-0.0127***	-0.0171***
	(0.0032)	(0.0034)	(0.0036)	(0.0037)	(0.0036)
ZIP Code × Month-Year FE		✓	✓		
ZIP Code × Loan Purpose FE			✓		
ZIP Code × Month-Year × Loan Purpose FE				✓	✓
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)	-0.0730***	-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	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: 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(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
			Panel C: No Credit Score		
Dep Var: No Credit Score	(1)	(2)	(3)	(4)	(5)
Post	-0.0113	-0.0158**	-0.0156**	-0.0124**	-0.0095**
	(0.0105)	(0.0075)	(0.0063)	(0.0052)	(0.0045)
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	0.3455	0.3509	0.3208	0.3184	0.3306

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 credit scores. Panel C shows effects on borrowers with missing 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 7: Heterogeneity & Effect of the Policy Change on Application Acceptances

Dep Var: Application Accepted	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Low Income × Post	-0.0659***						-0.0383**
	(0.0174)						(0.0171)
Low Income	-0.0013						0.0009
	(0.0085)						(0.0078)
Young × Post		-0.0770***					-0.0572***
		(0.0161)					(0.0162)
Young		-0.0092					-0.0019
		(0.0079)					(0.0079)
First-time Applicant × Post			-0.2541***				-0.2385***
			(0.0189)				(0.0184)
First-time Applicant			-0.0038				-0.0014
			(0.0082)				(0.0081)
Low Caste × Post				-0.0811***			-0.0361**
				(0.0173)			(0.0162)
Low Caste				-0.0014			-0.0005
				(0.0083)			(0.0080)
Female × Post					-0.0305		-0.0183
					(0.0255)		(0.0241)
Female					-0.0134		-0.0103
					(0.0129)		(0.0132)
Sub-Prime Credit Score × Post						-0.0155	-0.0247
						(0.0174)	(0.0161)
Sub-Prime Credit Score						-0.0003	-0.0009
						(0.0079)	(0.0071)
ZIP Code × Loan Purpose FE	✓	✓	✓	✓	✓	✓	✓
Application Date 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	0.7966	0.7968	0.8129	0.7972	0.7946	0.7955	0.8170

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 application date 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 applicants who have not previously had a loan approved by the platform, 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 a 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 8: 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)	
Post	0.0885***			
	(0.0077)			
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)	
Post	0.0868***			
	(0.0080)			
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)	
Post	0.0804***			
	(0.0081)			
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 9: 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	0.2194***	0.2036***	0.1860***	0.1268**
	(0.0187)	(0.0472)	(0.0582)	(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	294.7749	236.4380	155.4587	123.5356
KP F Statistic	8722.8835	2300.5141	1274.7938	1150.3214
Panel B: Credit Access Within 3 Years				
Dep Var: Formal Credit Access (3 Years)	(1)	(2)	(3)	(4)
App Accepted	0.2253***	0.1973***	0.2049***	0.1359**
	(0.0193)	(0.0492)	(0.0603)	(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	294.7749	236.4380	155.4587	123.5356
KP F Statistic	8722.8835	2300.5141	1274.7938	1150.3214
Panel C: Credit Access Within 4 Years				
Dep Var: Formal Credit Access (4 Years)	(1)	(2)	(3)	(4)
App Accepted	0.2110***	0.1697***	0.1876***	0.1365**
	(0.0191)	(0.0493)	(0.0600)	(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	294.7749	236.4380	155.4587	123.5356
KP F Statistic	8722.8835	2300.5141	1274.7938	1150.3214

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 10: Parameter Estimates and Welfare Decomposition

Key Quantities	No Internalization	Full Internalization
Lender's Required Rate of Return ($\hat{\kappa}$)	0.1832***	0.1832***
	(0.0011)	(0.0011)
Direct Utility Gain from Privacy ($\hat{\theta}$)	0.1891***	0.3095***
	(0.0037)	(0.0037)
WTP [% of income]	₹9,368 [20.82%]	₹16,325 [36.27%]
(1) Percent Change in Consumer Surplus		
Total	+0.23%	+0.60%
Privacy Margin	+0.49%	+0.82%
Contract Margin	−0.26%	−0.22%
Immediate Credit Access	−0.19%	−0.16%
Long-Term Credit Access	−0.07%	−0.06%
(2) Equivalent Variation (% of income)		
Total	+0.02%	+0.12%
Privacy Margin	+0.13%	+0.21%
Contract Margin	−0.11%	−0.09%
Immediate Credit Access	−0.08%	−0.07%
Long-Term Credit Access	−0.03%	−0.02%
(3) Percent Change in Lender Profits		
Total	−23.08%	−20.95%
Application Margin	+3.56%	+5.69%
Approval Margin	+6.00%	+6.00%
Per-Loan Surplus Margin	−32.64%	−32.64%

$\hat{\kappa}$ 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 (8)). Under no internalization, borrowers apply using pre-policy beliefs about default probabilities and recovery rates; consumer surplus is computed as expected utility plus binary entropy (equation (7)). Under full internalization, borrowers correctly account for the post-policy screening environment; consumer surplus is computed as a standard logit inclusive value (equation (6)). The contract margin is further decomposed into immediate and long-term credit access components by setting $\Delta = 0$ in equation (3) to shut down the future loan bonus B_i. Equivalent variation transforms the change in expected utility into a monetary equivalent, scaled as a percentage of mean monthly income. Standard errors 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

A Data **A2**

B Robustness for DiD Analysis **A3**

C Robustness for High-Frequency Event Study Analysis **A5**

D Effect of Privacy Policy on Long-Term Credit Access **A8**

E Framework **A12**

E.1 Decomposing the contract margin A12

E.2 Computing Equivalent Variation A13

E.3 Computing Standard Errors A15

E.4 Identification A16

E.4.1 Identification of κ A16

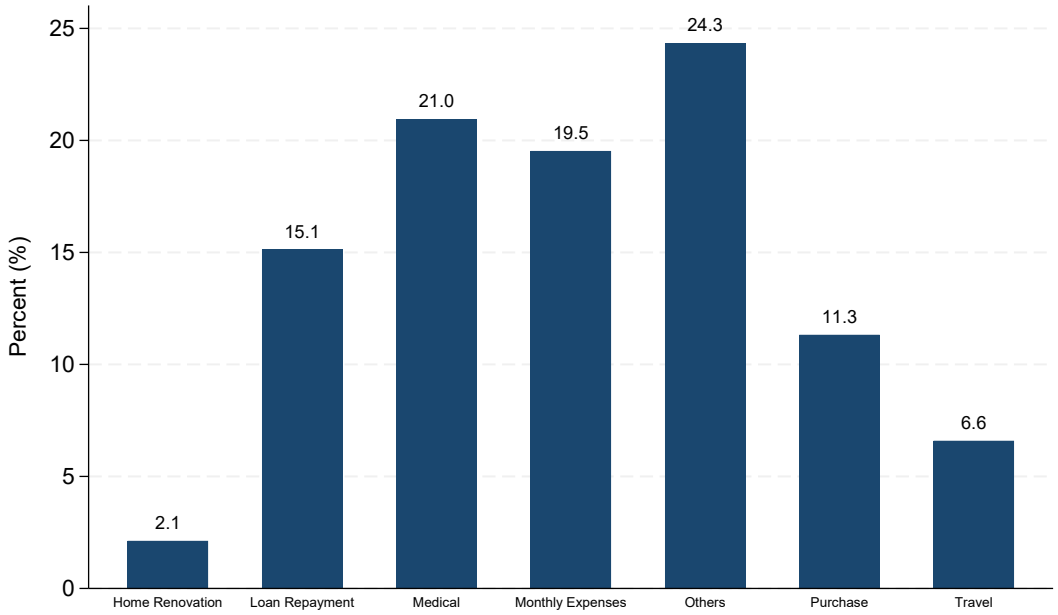
E.4.2 Identification of θ A17

E.4.3 Location of applicants on logit CDF A18

E.5 Sensitivity of estimates to consumption floor and future shock probability A19

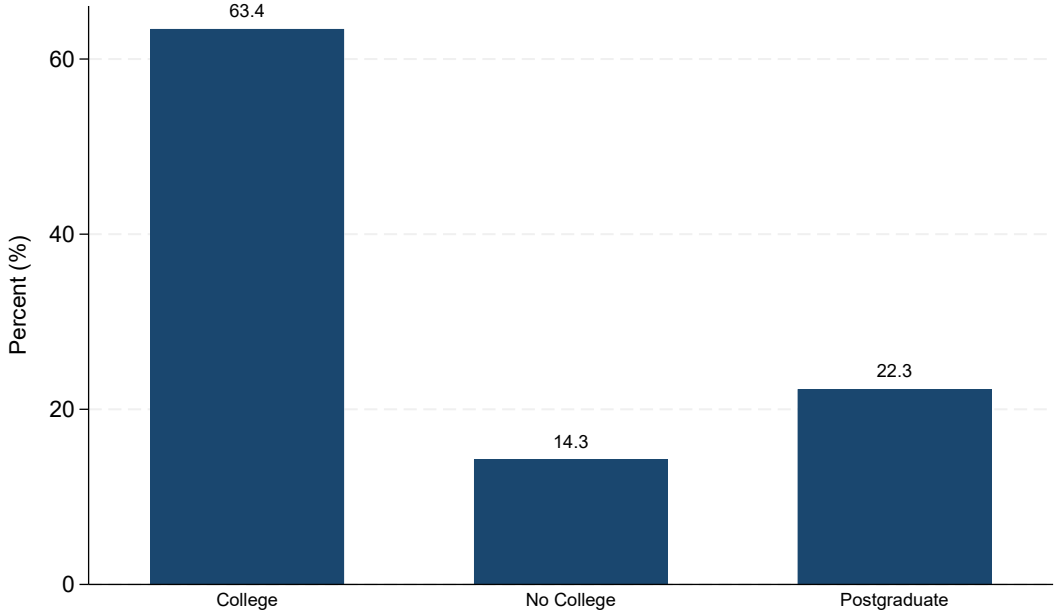
Appendix A Data

Figure A.1: Distribution of Loan Purpose



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

Figure A.2: Distribution of Applicant Education



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 Number of Applications

Dep Var: # Applications	(1)	(2)	(3)	(4)	(5)
Android × Post	0.4505***	0.4505***	0.4505***	0.4623***	0.3862***
	(0.0167)	(0.0167)	(0.0167)	(0.0174)	(0.0192)
Android	1.6159***	1.6159***	1.6159***		
	(0.0201)	(0.0201)	(0.0201)		
Post	-0.1489***				
	(0.0147)				
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	87,876	87,876	87,876	87,876	87,876
Pseudo R^2	0.2259	0.4712	0.5520	0.5662	0.5796

This table presents difference-in-differences estimates examining the effect of the Google privacy policy change on monthly loan applications using a Poisson pseudo-maximum likelihood (PPML) estimator. The dependent variable is 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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

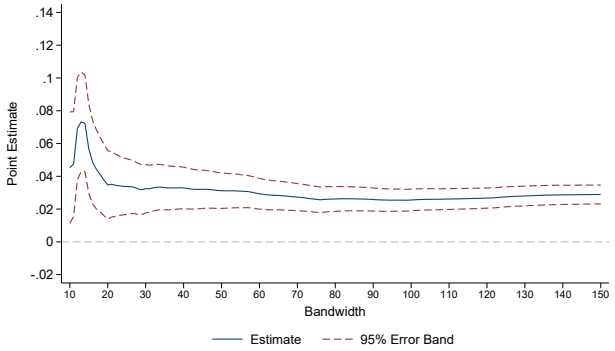
Table B.2: Effect of the Policy Change on Interest Rates and Loan Maturity

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

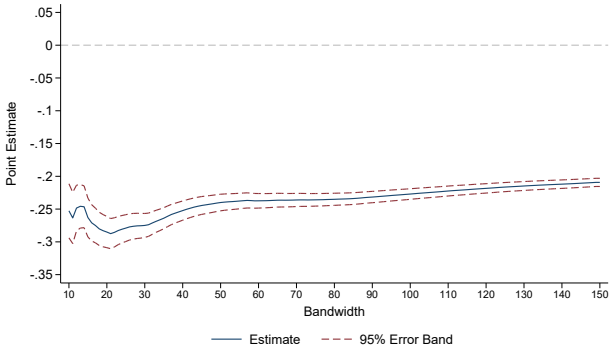
This table presents difference-in-differences estimates examining the effect of the Google privacy policy change on loan pricing terms for approved loans. Panel A shows OLS effects on annual interest rates (in percentage points), while Panel B shows Poisson pseudo-maximum likelihood (PPML) effects on 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

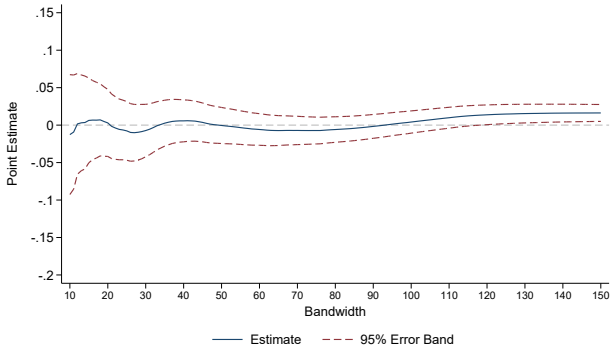
Figure C.1: Robustness to Bandwidth Choice: High-Frequency Event Study Estimates Across Multiple Bandwidths



(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 10-150 days). Panel (b) shows estimates for application acceptance rates (bandwidths 10-150 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 $\times$ 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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

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 (Excluding FinTech Lenders): Evidence from DiD Specification

Panel A: Credit Access Within 2 Years				
Dep Var: Formal Credit Access, Excl. FinTech (2 Years)	(1)	(2)	(3)	(4)
Android × Post	-0.0970***	-0.0353***	-0.0317***	-0.0217**
	(0.0085)	(0.0084)	(0.0100)	(0.0097)
Android	-0.0492***	-0.0496***	-0.0493***	
	(0.0097)	(0.0095)	(0.0108)	
Post	0.0877***			
	(0.0078)			
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	0.0964	0.3054	0.3291
Panel B: Credit Access Within 3 Years				
Dep Var: Formal Credit Access, Excl. FinTech (3 Years)	(1)	(2)	(3)	(4)
Android × Post	-0.0992***	-0.0339***	-0.0345***	-0.0231**
	(0.0087)	(0.0088)	(0.0103)	(0.0099)
Android	-0.0607***	-0.0604***	-0.0547***	
	(0.0100)	(0.0099)	(0.0113)	
Post	0.0855***			
	(0.0081)			
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	0.0049	0.0982	0.3086	0.3315
Panel C: Credit Access Within 4 Years				
Dep Var: Formal Credit Access, Excl. FinTech (4 Years)	(1)	(2)	(3)	(4)
Android × Post	-0.0932***	-0.0292***	-0.0318***	-0.0235**
	(0.0087)	(0.0088)	(0.0103)	(0.0099)
Android	-0.0657***	-0.0650***	-0.0576***	
	(0.0097)	(0.0097)	(0.0109)	
Post	0.0794***			
	(0.0081)			
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	0.0965	0.3084	0.3300

This table replicates the analysis of Table 8 after excluding FinTech lenders from the definition of formal credit access, to address the concern that competing FinTech platforms were subject to the same privacy policy shock. The dependent variables are indicators for whether a borrower accessed any formal non-FinTech 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 D.3: Effect of Privacy Policy on Long-Term Credit Access (Excluding FinTech Lenders): Evidence from 2SLS Specification

Panel A: Credit Access Within 2 Years				
Dep Var: Formal Credit Access, Excl. FinTech (2 Years)	(1)	(2)	(3)	(4)
App Accepted	0.2156***	0.1985***	0.1863***	0.1317**
	(0.0186)	(0.0472)	(0.0584)	(0.0588)
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	294.7749	236.4379	155.4587	125.3834
KP F Statistic	8722.8835	2300.5170	1274.7938	1126.6004
Panel B: Credit Access Within 3 Years				
Dep Var: Formal Credit Access, Excl. FinTech (3 Years)	(1)	(2)	(3)	(4)
App Accepted	0.2206***	0.1904***	0.2029***	0.1403**
	(0.0192)	(0.0490)	(0.0604)	(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	294.7749	236.4379	155.4587	125.3834
KP F Statistic	8722.8835	2300.5170	1274.7938	1126.6004
Panel C: Credit Access Within 4 Years				
Dep Var: Formal Credit Access, Excl. FinTech (4 Years)	(1)	(2)	(3)	(4)
App Accepted	0.2071***	0.1643***	0.1870***	0.1432**
	(0.0190)	(0.0494)	(0.0604)	(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	294.7749	236.4379	155.4587	125.3834
KP F Statistic	8722.8835	2300.5170	1274.7938	1126.6004

This table replicates the analysis of Table 9 after excluding FinTech lenders from the definition of formal credit access, to address the concern that competing FinTech platforms were subject to the same privacy policy shock. The dependent variables are indicators for whether a borrower accessed any formal non-FinTech 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 D.4: 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 Decomposing the contract margin

The contract margin reflects two economically distinct sources of consumer surplus from tighter screening. First, borrowers lose immediate access to credit: the current loan is less likely to be approved, reducing the direct consumption-smoothing value of the loan. Second, borrowers lose long-term access: tighter screening today reduces the probability of establishing a credit relationship with the lender, lowering the option value of future loans through the FinTech-ladder effect captured by B_i. To separate these channels, we construct an additional counterfactual in which the future loan bonus is shut down by setting $\Delta = 0$ in equation (3), so that $B_i = 0$ for all borrowers. This section presents the details about constructing this counterfactual.

Define $\tilde{u}_i^{\text{loan}} = u_i^{\text{loan}} - B_i$ as loan utility net of the future relationship value, and let tildes denote all downstream objects recomputed under $\Delta = 0$:

$$\tilde{V}_i^{\text{CDR}} = \hat{P}_i^{A,\text{CDR}} \cdot \left(\tilde{u}_i^{\text{loan}} - u_i^{\text{noloan}}\right)$$
$$\tilde{V}_i^{\text{NoCDR}} = \hat{P}_i^{A,\text{NoCDR}} \cdot \left(\tilde{u}_i^{\text{loan}} - u_i^{\text{noloan}}\right)$$

The corresponding counterfactual consumer surplus objects under $\Delta = 0$ are:

$$\widetilde{CS}_i^{\text{CF}} = \log\left(\exp\left(u_i^{\text{noloan}}\right) + \exp\left(u_i^{\text{noloan}} + \tilde{V}_i^{\text{CDR}} + \theta\right)\right)$$
$$\widetilde{CS}_i^{\text{NoCDR}} = \log\left(\exp\left(u_i^{\text{noloan}}\right) + \exp\left(u_i^{\text{noloan}} + \tilde{V}_i^{\text{NoCDR}} + \theta\right)\right)$$

under full internalization, and analogously using the entropy-based equation (7) under no internalization.[14]

The contract margin then decomposes as:

$$\underbrace{CS_i^{\text{NoCDR}} - CS_i^{\text{CF}}}_{\text{contract margin}} = \underbrace{\widetilde{CS}_i^{\text{NoCDR}} - \widetilde{CS}_i^{\text{CF}}}_{\text{immediate access}} + \underbrace{\left(CS_i^{\text{NoCDR}} - CS_i^{\text{CF}}\right) - \left(\widetilde{CS}_i^{\text{NoCDR}} - \widetilde{CS}_i^{\text{CF}}\right)}_{\text{long-term access}} \tag{E.1}$$

The immediate access component isolates the reduction in consumer surplus from tighter screening on the current loan, abstracting from the future relationship value by setting $\Delta = 0$. The long-term access component captures the surplus loss from the FinTech ladder channel: tighter screening today reduces the probability of securing future loans, lowering their option value.

[14] Under no internalization, the $\Delta = 0$ consumer surplus objects are $\widetilde{CS}_i^{\text{CF}} = u_i^{\text{noloan}} + \tilde{P}_i^{\text{apply}} \cdot (\tilde{V}_i^{\text{CDR}} + \theta) + H(\tilde{P}_i^{\text{apply}})$ and $\widetilde{CS}_i^{\text{NoCDR}} = u_i^{\text{noloan}} + \tilde{P}_i^{\text{apply}} \cdot (\tilde{V}_i^{\text{NoCDR}} + \theta) + H(\tilde{P}_i^{\text{apply}})$, where $\tilde{P}_i^{\text{apply}} = \Lambda(\tilde{V}_i^{\text{CDR,borrower}} + \theta)$ and the borrower's perceived approval probability $\tilde{V}_i^{\text{CDR,borrower}}$ is also recomputed under $\Delta = 0$.

E.2 Computing Equivalent Variation

We follow Dávila and Schaab (2025) and translate the change in consumer surplus into a scale-invariant measure. Specifically, we convert ΔCS_i into an equivalent variation: the permanent increase in monthly income that, if received under the pre-policy regime, would produce the same welfare gain as the CDR ban. This section presents the details of this computation.

To do so, we define α_i as the solution to

$$CS_i^{\text{CDR}}\big|_{(1+\alpha_i)\,y_i} = CS_i^{\text{NoCDR}}\big|_{y_i}$$

That is, α_i is the permanent proportional income increase that, if received in the CDR regime, would make the borrower indifferent to the policy change. We assume that the income shift does not affect the loan contract (L_i, r_i, T_i) or the lender's approval rule, so that V_i^s is unchanged. Under this assumption, consumer surplus shifts only through the no-loan utility u_i^{noloan}. When income changes from y_i to $(1+\alpha)y_i$, the no-loan utility shifts by

$$u_i^{\text{noloan}}\big|_{(1+\alpha)\,y_i} - u_i^{\text{noloan}}\big|_{y_i} = \frac{\log(1+\alpha)}{1-\delta}$$

since every period's flow utility gains $\log(1+\alpha)$ and the infinite-horizon sum is $\sum_{t=0}^{\infty}\delta^t = 1/(1-\delta)$. Because V_i^{CDR} is held fixed, consumer surplus shifts by the same amount:

$$CS_i^{\text{CDR}}\big|_{(1+\alpha)\,y_i} = CS_i^{\text{CDR}} + \frac{\log(1+\alpha)}{1-\delta}$$

Substituting into the indifference condition yields

$$\frac{\log(1+\alpha_i)}{1-\delta} = \Delta CS_i$$

which inverts to

$$\alpha_i = \exp\left((1-\delta)\,\Delta CS_i\right) - 1$$

We report the mean individual equivalent variation as the aggregate welfare measure:

$$\alpha = \frac{1}{N}\sum_{i=1}^{N}\alpha_i = \frac{1}{N}\sum_{i=1}^{N}\left[\exp\left((1-\delta)\,\Delta CS_i\right) - 1\right]$$

Decomposition. The welfare change decomposes additively in utility space: $\Delta CS_i = \Delta CS_i^{\text{priv}} + \Delta CS_i^{\text{cont}}$, and the contract margin further decomposes as $\Delta CS_i^{\text{cont}} = \Delta CS_i^{\text{imm}} + \Delta CS_i^{\text{lt}}$, where ΔCS_i^{imm} captures the surplus change from reduced approval probabilities for the current loan and ΔCS_i^{lt} captures the loss in option value from a reduced probability of establishing a long-term credit relationship via the FinTech-ladder effect. To decompose α into margin-specific components, we use the first-order approximation $\exp(x) - 1 \approx x$ for small x:

$$\alpha_i \;\approx\; (1-\delta)\,\Delta CS_i \;=\; \underbrace{(1-\delta)\,\Delta CS_i^{\text{priv}}}_{\alpha_i^{\text{priv}}} \;+\; \underbrace{(1-\delta)\,\Delta CS_i^{\text{imm}}}_{\alpha_i^{\text{imm}}} \;+\; \underbrace{(1-\delta)\,\Delta CS_i^{\text{lt}}}_{\alpha_i^{\text{long-term}}}$$

Under this approximation, the decomposition is additive: each margin's equivalent variation is obtained by replacing ΔCS_i with the corresponding margin ΔCS_i^m in the same formula. The aggregate margin-specific equivalent variations are

$$\alpha^m = \frac{1}{N}\sum_{i=1}^{N} \alpha_i^m = \frac{1}{N}\sum_{i=1}^{N}(1-\delta)\,\Delta CS_i^m \qquad m \in \{\text{priv, imm, long-term}\}$$

E.3 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 $\hat{\kappa}$, we compute the cluster-robust standard error from the scores of the pooled log-likelihood. The score for observation i in cluster c is given by the following expression:

$$s_i(\kappa) = (y_i - \Lambda(S_i))\,(-L_i) \tag{E.2}$$

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

$$\widehat{SE}(\hat{\kappa}) = \frac{\sqrt{\sum_c S_c^2}}{|\mathcal{H}|} \tag{E.3}$$

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

For $\hat{\theta}$, let $\bar{m} = 0.057$ denote the empirical moment from Panel A of Table 5. We define the individual moment contribution:

$$m_i(\theta) = \frac{\Lambda(V_i^{\text{NoCDR}} + \theta)}{\bar{P}^{\text{pre}}} - 1 - \bar{m} \tag{E.4}$$

so that $N^{-1} \sum_i m_i(\hat{\theta}) = 0$ at the solution. The cluster-robust delta method yields:

$$\widehat{SE}(\hat{\theta}) = \frac{\sqrt{\sum_c M_c^2 / N^2}}{\left|\partial g / \partial \theta\right|_{\hat{\theta}}\right|} \tag{E.5}$$

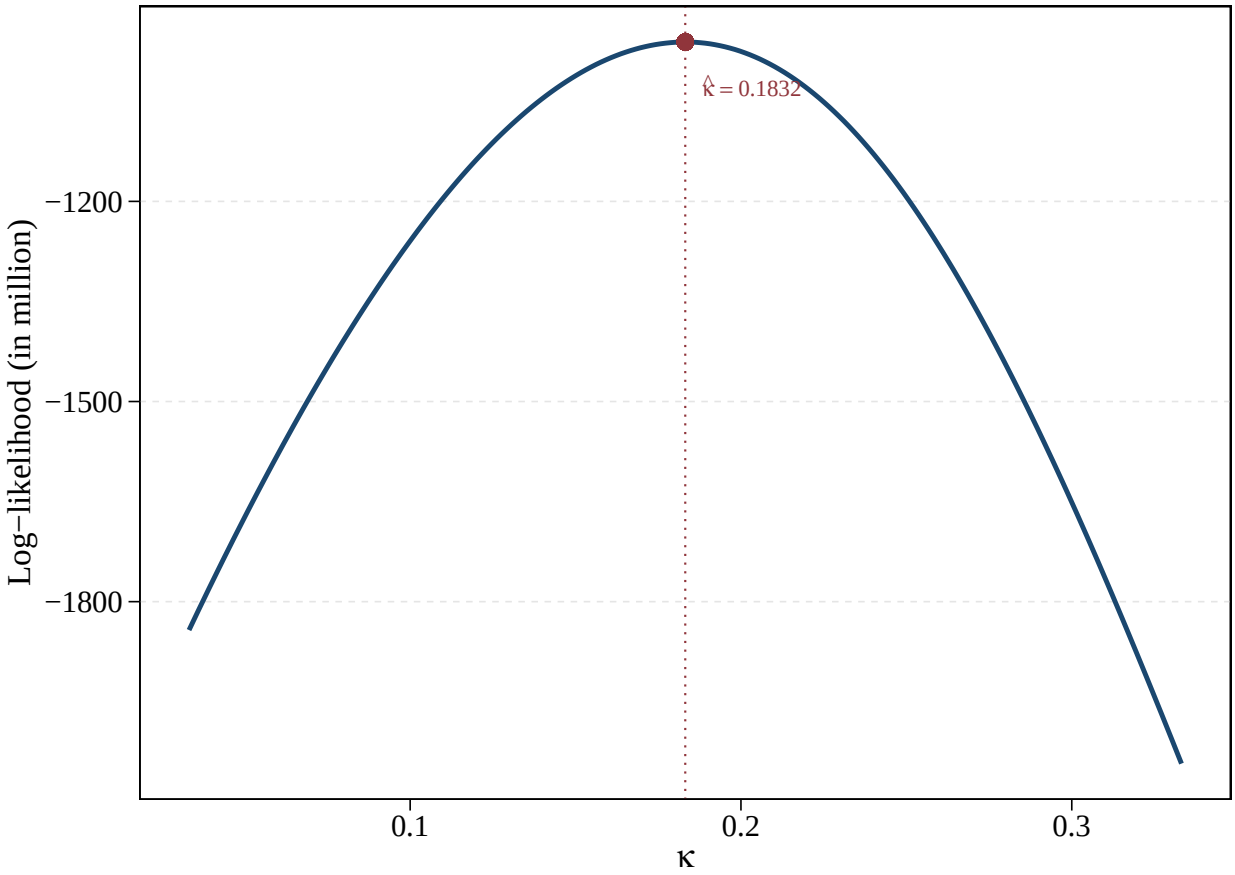
where $M_c = \sum_{i \in c} m_i(\hat{\theta})$ is the cluster sum of moment contributions and the derivative is approximated by a symmetric finite difference with step size $h = 10^{-6}$.

E.4 Identification

E.4.1 Identification of κ

Figure E.1 displays the log-likelihood profile for κ, demonstrating a unique, well-defined global maximum.

Figure E.1: Identification of κ: Log Likelihood Function

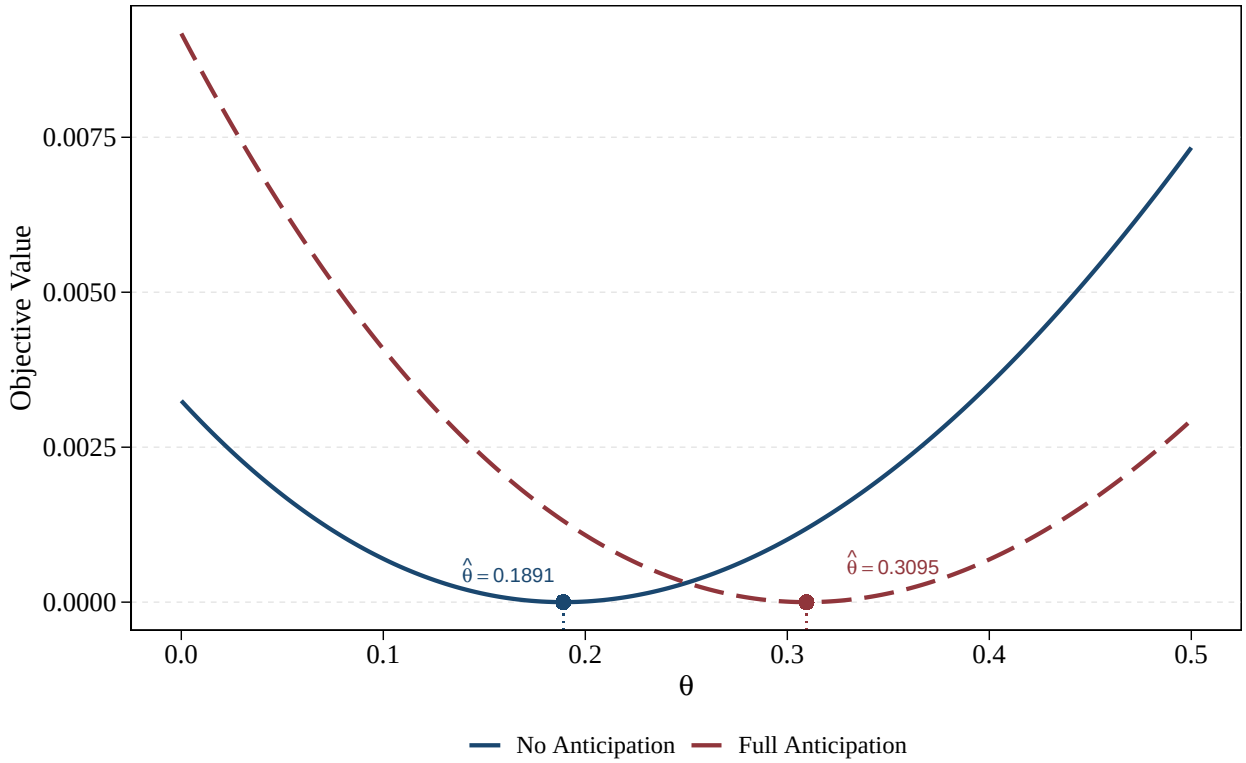


This figure plots the log-likelihood function for the per-application processing cost κ estimated via maximum likelihood. The log-likelihood is evaluated over a grid of κ values. The unique global maximum confirms that κ is sharply identified by the observed approval rates in the data.

E.4.2 Identification of θ

Figure E.2 plots the objective function used to identify θ. Both specifications yield a unique, sharply defined minimum, confirming that the 5.7% moment pins down θ precisely.

Figure E.2: Identification of θ: Objective Function

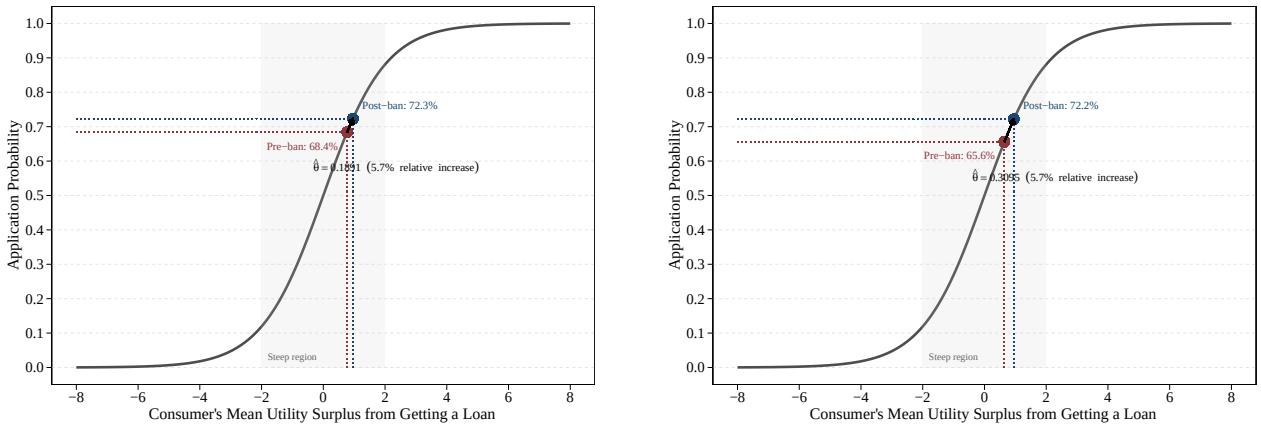


This figure plots the objective function $(g(\theta) - 0.057)^2$ against θ under both the no-internalization and full-internalization specifications, where $g(\theta) = \frac{\bar{P}^{\text{post}}(\theta) - \bar{P}^{\text{pre}}}{\bar{P}^{\text{pre}}}$ is the model-predicted percentage change in application rates. The parameter θ is identified as the unique minimizer that matches the observed 5.7% increase in Android applications documented in Panel A of Table 5. All other parameters are held at their baseline values.

E.4.3 Location of applicants on logit CDF

A potential concern is that θ may be artificially inflated if the observed applicants sit on the flat, upper portion of the logistic CDF, where the marginal response to θ is low. Figures E.3a and E.3b show that this is not the case. Under no-internalization, the mean pre-ban application probability is approximately 68%, and under full-internalization it is lower still. In both cases, the model-implied probabilities fall within the responsive region of the logistic (shaded), not on the flat upper tail. The local slope of the CDF at these points remains substantial, so θ is not mechanically inflated by low marginal responsiveness.

Figure E.3: Identification of θ under Different Internalization Assumptions



(a) Logistic CDF under No-Internalization

(b) Logistic CDF under Full-Internalization

Each panel plots the logistic application probability $\Lambda(V) = \frac{\exp(V)}{1+\exp(V)}$ against the expected utility gain V from applying for a loan. The two points mark the model-implied pre-ban and post-ban application probabilities evaluated at $\bar{V}$, with the arrow indicating the shift induced by the CDR ban. In Panel (a), the no-internalization specification yields $\hat{\theta} = 0.1891$, with application probability rising from 68.4% to 72.3%. In Panel (b), the full-internalization assumption yields a larger $\hat{\theta} = 0.3095$ to rationalize the same 5.7% relative increase. In both panels, the pre- and post-ban points lie well within the steep region of the logistic curve, confirming that identification does not rely on the flat tails.

E.5 Sensitivity of estimates to consumption floor and future shock probability

We assess the robustness of our key estimates to the two calibrated parameters: the future shock probability p_{shock} and the consumption floor C_{FLOOR}. For each combination of $p_{\text{shock}} \in \{0.1, 0.2, \ldots, 0.9\}$ and $C_{\text{FLOOR}} \in \{1250, 1500, \ldots, 3000\}$, we re-estimate θ by exact moment matching to the observed 5.7% application increase under both the no-internalization and full-internalization specifications, holding all other parameters at their baseline values. Appendix Tables E.1 and E.2 report the results.

The results are qualitatively stable across the entire parameter space and both specifications. Under no internalization, $\hat{\theta}$ ranges from 0.154 to 0.233, corresponding to a willingness to pay between 16.7% and 26.2% of mean monthly income. Under full internalization, $\hat{\theta}$ is uniformly larger, ranging from 0.228 to 0.445, with implied willingness to pay between 25.6% and 56.1% of mean monthly income. The net change in consumer surplus remains positive throughout under both specifications, ranging from +0.168% to +0.304% under no internalization and from +0.417% to +0.810% under full internalization, with the privacy margin always dominating the contract margin. The equivalent variation likewise remains positive across all cells, ranging from +0.001% to +0.039% of monthly income under no internalization and from +0.089% to +0.161% under full internalization. The change in lender profits remains negative throughout, ranging from −23.73% to −22.56% under no internalization and from −21.48% to −20.55% under full internalization.

Within each specification, $\hat{\theta}$ increases in p_{shock} and decreases in C_{FLOOR}. Both patterns reflect the shape of the logistic application probability. A higher p_{shock} raises the future loan bonus B_i, which increases borrowers' baseline expected utility from applying and pushes application probabilities into the flatter region of $\Lambda(\cdot)$; a larger θ is therefore needed to match the 5.7% moment. Conversely, a higher C_{FLOOR} reduces loan surplus for constrained borrowers, pushing their expected utility from applying closer to zero where the logistic function is steepest; application probabilities are therefore more responsive to θ, and a smaller θ suffices to match it.

Table E.1: Sensitivity Analysis: No Internalization

	p_{shock}								
C_{FLOOR}	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Panel A: Direct Utility Gain ($\hat{\theta}$)									
1250	0.1754	0.1813	0.1875	0.1940	0.2010	0.2083	0.2161	0.2243	0.2330
1500	0.1714	0.1772	0.1833	0.1898	0.1966	0.2039	0.2116	0.2197	0.2282
1750	0.1678	0.1735	0.1795	0.1859	0.1927	0.1998	0.2074	0.2154	0.2238
2000	0.1646	0.1702	0.1762	0.1824	**0.1891**	0.1961	0.2036	0.2115	0.2198
2250	0.1617	0.1673	0.1731	0.1793	0.1859	0.1928	0.2001	0.2079	0.2161
2500	0.1591	0.1646	0.1703	0.1764	0.1829	0.1897	0.1969	0.2046	0.2127
2750	0.1567	0.1620	0.1677	0.1737	0.1801	0.1868	0.1939	0.2014	0.2094
3000	0.1544	0.1597	0.1653	0.1712	0.1774	0.1841	0.1911	0.1985	0.2063
Panel B: WTP (% of monthly income)									
1250	19.18%	19.87%	20.62%	21.41%	22.26%	23.16%	24.12%	25.14%	26.24%
1500	18.70%	19.39%	20.12%	20.90%	21.73%	22.62%	23.56%	24.57%	25.64%
1750	18.27%	18.95%	19.67%	20.43%	21.25%	22.12%	23.05%	24.03%	25.08%
2000	17.89%	18.56%	19.26%	20.02%	**20.82%**	21.67%	22.58%	23.55%	24.58%
2250	17.55%	18.21%	18.90%	19.64%	20.42%	21.26%	22.16%	23.11%	24.12%
2500	17.25%	17.89%	18.57%	19.29%	20.07%	20.89%	21.77%	22.70%	23.70%
2750	16.96%	17.59%	18.26%	18.97%	19.73%	20.54%	21.40%	22.32%	23.29%
3000	16.70%	17.31%	17.97%	18.67%	19.42%	20.21%	21.05%	21.96%	22.92%
Panel C: ΔConsumer Surplus (%)									
1250	0.168%	0.177%	0.186%	0.196%	0.207%	0.220%	0.234%	0.250%	0.267%
1500	0.181%	0.189%	0.198%	0.208%	0.219%	0.232%	0.245%	0.261%	0.277%
1750	0.191%	0.199%	0.208%	0.217%	0.228%	0.240%	0.254%	0.269%	0.285%
2000	0.199%	0.207%	0.215%	0.225%	**0.236%**	0.247%	0.261%	0.275%	0.291%
2250	0.206%	0.213%	0.222%	0.231%	0.241%	0.253%	0.266%	0.280%	0.295%
2500	0.211%	0.218%	0.227%	0.236%	0.246%	0.258%	0.270%	0.284%	0.299%
2750	0.215%	0.223%	0.231%	0.240%	0.250%	0.261%	0.273%	0.287%	0.302%
3000	0.219%	0.226%	0.234%	0.243%	0.253%	0.264%	0.276%	0.289%	0.304%
Panel D: Equivalent Variation (% of income)									
1250	0.0012%	0.0013%	0.0016%	0.0022%	0.0031%	0.0043%	0.0059%	0.0078%	0.0102%
1500	0.0085%	0.0085%	0.0088%	0.0093%	0.0102%	0.0113%	0.0128%	0.0147%	0.0170%
1750	0.0143%	0.0143%	0.0146%	0.0151%	0.0159%	0.0170%	0.0185%	0.0202%	0.0224%
2000	0.0192%	0.0192%	0.0195%	0.0200%	**0.0207%**	0.0218%	0.0232%	0.0249%	0.0269%
2250	0.0234%	0.0234%	0.0236%	0.0241%	0.0248%	0.0258%	0.0271%	0.0288%	0.0308%
2500	0.0270%	0.0270%	0.0272%	0.0277%	0.0283%	0.0293%	0.0306%	0.0322%	0.0341%
2750	0.0301%	0.0301%	0.0303%	0.0307%	0.0314%	0.0323%	0.0335%	0.0351%	0.0369%
3000	0.0328%	0.0328%	0.0330%	0.0334%	0.0341%	0.0350%	0.0361%	0.0376%	0.0394%
Panel E: ΔLender Profits (%)									
1250	-23.73%	-23.65%	-23.57%	-23.50%	-23.43%	-23.36%	-23.30%	-23.24%	-23.19%
1500	-23.59%	-23.51%	-23.44%	-23.37%	-23.30%	-23.24%	-23.19%	-23.13%	-23.08%
1750	-23.45%	-23.38%	-23.31%	-23.25%	-23.19%	-23.13%	-23.08%	-23.03%	-22.99%
2000	-23.32%	-23.25%	-23.19%	-23.13%	**-23.08%**	-23.03%	-22.98%	-22.93%	-22.89%
2250	-23.19%	-23.13%	-23.08%	-23.02%	-22.97%	-22.93%	-22.88%	-22.84%	-22.80%
2500	-23.07%	-23.02%	-22.97%	-22.92%	-22.87%	-22.83%	-22.79%	-22.75%	-22.72%
2750	-22.96%	-22.91%	-22.86%	-22.82%	-22.77%	-22.74%	-22.70%	-22.67%	-22.64%
3000	-22.85%	-22.80%	-22.76%	-22.72%	-22.68%	-22.65%	-22.61%	-22.58%	-22.56%

This table summarizes the sensitivity of structural estimates to calibrated parameters under the no-internalization specification. The baseline case ($p_{\text{shock}} = 0.5, C_{\text{FLOOR}} = 2000$) is shown in **bold**.

Table E.2: Sensitivity Analysis: Full Internalization

	p_{shock}								
C_{FLOOR}	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Panel A: Direct Utility Gain ($\hat{\theta}$)									
1250	0.2906	0.3025	0.3149	0.3279	0.3415	0.3557	0.3705	0.3860	0.4023
1500	0.2792	0.2911	0.3035	0.3165	0.3300	0.3441	0.3589	0.3744	0.3905
1750	0.2687	0.2806	0.2929	0.3058	0.3193	0.3334	0.3481	0.3635	0.3796
2000	0.2591	0.2710	0.2833	0.2961	**0.3095**	0.3235	0.3381	0.3534	0.3693
2250	0.2504	0.2621	0.2744	0.2871	0.3004	0.3143	0.3288	0.3440	0.3598
2500	0.2423	0.2540	0.2661	0.2788	0.2920	0.3058	0.3202	0.3352	0.3509
2750	0.2347	0.2463	0.2583	0.2709	0.2840	0.2976	0.3119	0.3268	0.3424
3000	0.2276	0.2391	0.2510	0.2635	0.2764	0.2900	0.3041	0.3189	0.3343
Panel B: WTP (% of monthly income)									
1250	33.73%	35.33%	37.02%	38.81%	40.70%	42.71%	44.85%	47.11%	49.52%
1500	32.21%	33.79%	35.46%	37.23%	39.10%	41.08%	43.18%	45.41%	47.78%
1750	30.83%	32.39%	34.04%	35.78%	37.62%	39.57%	41.64%	43.83%	46.16%
2000	29.58%	31.12%	32.75%	34.46%	**36.27%**	38.19%	40.23%	42.39%	44.68%
2250	28.45%	29.97%	31.57%	33.26%	35.04%	36.93%	38.93%	41.06%	43.31%
2500	27.42%	28.91%	30.49%	32.15%	33.91%	35.77%	37.73%	39.82%	42.04%
2750	26.45%	27.92%	29.48%	31.11%	32.84%	34.67%	36.60%	38.65%	40.83%
3000	25.55%	27.01%	28.53%	30.14%	31.84%	33.64%	35.54%	37.56%	39.70%
Panel C: ΔConsumer Surplus (%)									
1250	0.503%	0.534%	0.567%	0.602%	0.639%	0.678%	0.719%	0.763%	0.810%
1500	0.490%	0.521%	0.553%	0.588%	0.624%	0.663%	0.704%	0.748%	0.794%
1750	0.477%	0.507%	0.539%	0.573%	0.609%	0.648%	0.688%	0.731%	0.777%
2000	0.464%	0.494%	0.525%	0.559%	**0.595%**	0.632%	0.672%	0.715%	0.760%
2250	0.452%	0.481%	0.512%	0.545%	0.580%	0.617%	0.657%	0.699%	0.743%
2500	0.440%	0.468%	0.499%	0.532%	0.566%	0.603%	0.642%	0.683%	0.727%
2750	0.428%	0.456%	0.486%	0.518%	0.552%	0.589%	0.627%	0.668%	0.711%
3000	0.417%	0.445%	0.474%	0.506%	0.539%	0.575%	0.612%	0.652%	0.695%
Panel D: Equivalent Variation (% of income)									
1250	0.0952%	0.1014%	0.1082%	0.1154%	0.1233%	0.1317%	0.1407%	0.1503%	0.1607%
1500	0.0953%	0.1015%	0.1082%	0.1154%	0.1231%	0.1314%	0.1404%	0.1499%	0.1602%
1750	0.0947%	0.1008%	0.1074%	0.1145%	0.1222%	0.1304%	0.1392%	0.1487%	0.1588%
2000	0.0937%	0.0998%	0.1063%	0.1133%	**0.1209%**	0.1290%	0.1377%	0.1471%	0.1571%
2250	0.0926%	0.0986%	0.1050%	0.1119%	0.1194%	0.1274%	0.1360%	0.1452%	0.1551%
2500	0.0914%	0.0972%	0.1036%	0.1104%	0.1177%	0.1256%	0.1341%	0.1432%	0.1529%
2750	0.0900%	0.0958%	0.1020%	0.1087%	0.1160%	0.1237%	0.1321%	0.1410%	0.1506%
3000	0.0886%	0.0943%	0.1004%	0.1070%	0.1141%	0.1218%	0.1300%	0.1388%	0.1482%
Panel E: ΔLender Profits (%)									
1250	-21.48%	-21.38%	-21.28%	-21.20%	-21.13%	-21.06%	-21.01%	-20.96%	-20.93%
1500	-21.42%	-21.31%	-21.22%	-21.14%	-21.06%	-21.00%	-20.94%	-20.89%	-20.86%
1750	-21.36%	-21.26%	-21.16%	-21.08%	-21.00%	-20.94%	-20.88%	-20.83%	-20.80%
2000	-21.31%	-21.21%	-21.11%	-21.03%	**-20.95%**	-20.88%	-20.83%	-20.78%	-20.74%
2250	-21.27%	-21.16%	-21.07%	-20.98%	-20.90%	-20.83%	-20.77%	-20.73%	-20.69%
2500	-21.23%	-21.12%	-21.02%	-20.93%	-20.86%	-20.79%	-20.73%	-20.68%	-20.64%
2750	-21.20%	-21.09%	-20.99%	-20.90%	-20.82%	-20.75%	-20.68%	-20.63%	-20.59%
3000	-21.17%	-21.05%	-20.95%	-20.86%	-20.78%	-20.71%	-20.65%	-20.59%	-20.55%

This table summarizes the sensitivity of structural estimates to calibrated parameters under the full-internalization specification. The baseline case ($p_{\text{shock}} = 0.5, C_{\text{FLOOR}} = 2000$) is shown in **bold**.