Privacy Regulation and the Economics of Screening versus Monitoring in Digital Credit Markets

Sumit Agarwal* Pulak Ghosh† Shohini Kundu‡ Nishant Vats§

Preliminary & Incomplete: Please do not circulate

Abstract

This paper examines the impact of privacy regulation on financial inclusion and lender profitability. Combining loan-level data with a regression discontinuity design around Google's 2019 ban on FinTech lenders' access to call detail records (CDR) in India, we document two key forces. On the demand side, loan applications increased by 4%, consistent with borrowers valuing enhanced privacy. On the supply side, approval rates fell by 25% with no change in default rates, indicating that lenders substituted away from CDR-based monitoring toward stricter ex-ante screening. This shift disproportionately affected new-to-credit, lower income, younger, and socially disadvantaged borrowers. Using credit bureau data, we show these effects persist: rejected applicants were 4–5 percentage points less likely to access formal credit even four years later. For lenders, foregone profits were substantial, amounting to nearly 30% of potential earnings. Our results highlight a key trade-off of privacy regulation: while consumers value privacy protections, these regulations can deepen financial exclusion, particularly among the very groups that FinTech innovations were designed to serve.

*National University of Singapore, email: ushakri@yahoo.com
†Indian Institute of Management Bangalore, email: pulak.ghosh@iimb.ac.in
‡UCLA Anderson School of Management, email: shohini.kundu@anderson.ucla.edu
§Olin Business School, Washington University at St Louis, email: vats@wustl.edu

1 Introduction

The rapid expansion of artificial intelligence (AI) and digital technologies in credit markets has fundamentally transformed how lenders assess borrower risk, particularly in emerging economies where traditional credit infrastructure remains underdeveloped. FinTech lenders have developed sophisticated credit scoring models that leverage alternative data sources, including call detail records (CDRs), mobile phone usage patterns, and digital behavioral traces, to evaluate borrowers traditionally excluded from formal financial systems due to thin or nonexistent credit histories. These innovations promise to advance financial inclusion by reducing information asymmetries that historically limited credit access for low-income and marginalized populations.[1]

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

This paper attempts to fill this gap and examines the tradeoff between consumer protection and financial inclusion, alongside the effect on lender profitability, through the lens of a large Indian FinTech lender that lost access to call detail records due to Google's privacy policy directive implemented on January 1, 2019. Our setting provides a unique natural experiment to quantify the economic costs of privacy restrictions, and to disentangle the mechanisms through which alternative data supports lending. Using comprehensive loan-level administrative data spanning 2017-2022, we exploit the sharp discontinuity created by the policy change to examine its effects on both sides of the credit market and its distributional consequences across borrower groups.

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

Our empirical strategy combines regression discontinuity and difference-in-differences designs to identify the effects of privacy regulation. The results reveal sharp and divergent responses between borrowers and lenders. On the demand side, we find that loan applications rise by approximately 4% following the privacy reform. This result is consistent with the evidence from the existing literature which shows that consumers place a meaningful value on privacy protections (Tang, 2019; Lin, 2022). However, the modest effect of privacy regulation on applications is in line with the conclusions of Athey, Catalini and Tucker (2017) and Prince and Wallsten (2022), which demonstrate that consumers are generally willing to accept small monetary value in exchange for sharing data on their social networks.

On the supply side, we find that acceptance rates fall by 25%, while default rates remain unchanged. Taken together, these results suggest that lenders respond to the regulation by tightening their ex-ante screening standards. In the pre-regulation period, the lender used CDRs to construct maps of borrowers' social networks. This created an implicit social collateral mechanism, whereby the lender could contact family members, friends, and professional associates. Specifically, this monitoring or disciplining technology allowed the lender to extend credit to borrowers who may appear risky on traditional metrics, knowing they could leverage social pressure to encourage repayment. Our findings suggest that the privacy regulation's removal of this monitoring technology forced the lender to compensate by substantially tightening ex-ante screening standards to maintain portfolio quality. That is, unable to rely on post-disbursement social enforcement, the lender becomes much more selective upfront, rejecting many borrowers who would have been profitable under the previous model where social collateral in the form of CDRs provided ongoing repayment discipline.

We then examine the distributional consequences of this policy shift. The loss of alternative data disproportionately affects borrower groups most reliant on nontraditional information sources. Low-income applicants, younger borrowers, first-time customers, and individuals from marginalized social backgrounds experience declines in acceptance rates that are 7-15 percentage points larger than their more advantaged counterparts. By contrast, borrowers with established credit histories or prior relationships with the lender are largely insulated from these effects. This heterogeneity highlights a fundamental tension. While privacy protections may benefit all consumers in principle, they can inadvertently exacerbate patterns of financial exclusion by limiting access for precisely those groups that face the greatest barriers to traditional credit.

We further investigate whether these immediate effects translate into persistent financial

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

Using both reduced-form regression discontinuity estimates and instrumental variables analysis that exploits the policy change as an exogenous shock to loan approval decisions, we find that borrowers whose applications are rejected due to privacy restrictions remain 4-5 percentage points less likely to access formal credit even four years later. Our IV estimates suggest that loan approval by our focal lender is associated with a 10-12 percentage points increase in the likelihood of long-term credit access. This result indicates that rejection of first-time applicants can create a persistent pattern of financial exclusion. These long-term effects are consistent with the prior results on the effect of Fintech credit on long-term access to bank credit (Beaumont, Tang and Vansteenberghe, 2022; Balyuk, 2023; Dubey, 2024).

Lastly, we quantify the economic effect of the stricter screening induced by privacy regulation for lender profitability. Although we observe an increase in applications, a decline in acceptance rates, and no change in ex-post default outcomes, the net effect on profitability remains ambiguous ex-ante. On the one hand, privacy restrictions force the lender to adopt a more selective screening strategy, as it can no longer rely on post-disbursement monitoring via CDRs. This shift lead to the rejection of many borrowers who would likely have been profitable under the earlier model, where CDRs provided ongoing repayment discipline through social collateral. On the other hand, the regulation simultaneously expands application volumes, potentially enlarging the pool of borrowers from which the lender could select.

To quantify the economic magnitude of these effects, we develop a counterfactual framework that estimates foregone profits from excluding creditworthy borrowers due to privacy restrictions. Our analysis identifies "marginal borrowers" those rejected post-policy who would have been approved under pre-policy standards and calculates the expected profitability of these foregone lending opportunities. We find that privacy regulation reduces lender profits by approximately 29% in our baseline specification, with an estimated annual effect of nearly ₹100 million ($5.3 million in PPP terms) for this single institution. Specifically, we find that that the foregone

profitability primarily stem from groups belonging to low social strata, including female and younger applicants with lower incomes and lower credit scores.

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

1.1 Related Literature

The key contribution of this paper is to examine a key tradeoff of privacy regulation. While privacy regulations can improve consumer protection, they can also have a negative effect on financial inclusion by narrowing the set of observable borrower characteristics lenders can rely on, resulting in a reduction in credit supply. While the theoretical literature have discussed these tradeoffs, as reviewed by Acquisti, Taylor and Wagman (2016), Bergemann and Morris (2019) and Goldfarb and Tucker (2019), the extant empirical literature has primarily focused on the value consumers attribute to data privacy highlighting the benefits of privacy regulation to consumers (Tang, 2019; Bian, Ma and Tang, 2021; Lin, 2022; Bian, Pagel, Tang and Raval, 2023). A notable exception is Miller and Tucker (2009), which documents that state privacy regulations can impose a cost on patients by deterring hospitals from adopting electronic medical records technology.[2] We contribute to this literature by documenting the tradeoffs associated with privacy regulation in credit markets. On the borrower side, we document an increase in the number of applications consistent with the idea that consumers place a positive value on privacy. On the lender side, we document a reduction in credit supply, primarily for socially and economically marginalized groups and new-to-credit borrowers. To highlight the importance of the latter group: 50% of the

[2]Relatedly, Dubé, Lynch, Bergemann, Demirer, Goldfarb, Johnson, Lambrecht, Lin, Tuchman and Tucker (2025) present a review of costs and benefits of data sharing for marketing purposes.

credit-eligible population in India is underserved due to being new-to-credit, accounting for a total of 572 million individuals (Cramer, Ghosh, Kulkarni and Vats, 2024).

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

We further contribute by documenting the importance of access to Fintech credit, especially for new-to-credit borrowers, for their long-term access to bank credit. We find that loan approval by our focal Fintech lender is associated with a 10-12 percentage points increase in the likelihood of long-term credit access. This result indicates that rejection of first-time applicants can create a persistent pattern of financial exclusion. These long-term effects are consistent with Fitench credit aiding borrowers to generate collateral (Beaumont, Tang and Vansteenberghe, 2022) as well as hard information (Balyuk, 2023; Dubey, 2024). Our findings complement these papers by demonstrating that when privacy regulation restricts access to FinTech loans for marginal borrowers, it can trigger persistent exclusion from the broader formal credit system. This finding suggests that FinTech lending plays an intermediating role in financial inclusion, making regulatory restrictions on alternative data particularly consequential for underserved populations' long-term economic prospects.

Lastly, we document the negative effect of this policy on the lender's profitability. The contraction in credit supply due to privacy regulation dominates the positive effect of this regulation on consumer demand leading to a net negative effect on lenders profitability. Put more directly, the privacy regulation led to the rejection of many borrowers who would likely have been profitable under the earlier model, where CDRs provided ongoing repayment discipline through social col-

lateral. Prior work has primairly focused on documenting the negative effect of privacy regulation on firms' web traffic and application installation Goldberg, Johnson and Shriver (2019); Aridor, Che and Salz (2020); Bessen, Impink, Reichensperger and Seamans (2020); Bian, Ma and Tang (2021); Jia, Jin and Wagman (2021); Janßen, Kesler, Kummer and Waldfogel (2022); Kraft, Skiera and Koschella (2023). Other works examine the how firms modify their data-collection and data-sharing processes following privacy regulations (Peukert, Bechtold, Batikas and Kretschmer, 2022; Bae, Mayya and Nian, 2023; Kesler, 2023; Demirer, Hernández, Li and Peng, 2024; Ramadorai, Uettwiller and Walther, 2025). A notable exception is Abis, Tang and Bian (2025), which documents the negative effect of privacy regulation on market efficiency. We add to this literature by documenting the negative effect of privacy regulation on the profitability of Fintech platforms. Specifically, we show that privacy regulation can reduce profitability by eroding information environments that can be key for borrower discipline.

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

2 Institutional Setting

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

2.1 Business Model and Target Market

The FinTech operates primarily through a mobile application platform, providing unsecured short-term personal loans. The company specifically targets millennials and individuals from marginal-

ized backgrounds who are often excluded from traditional banking services due to limited or no credit history. Loan amounts range from approximately ₹10,000 to ₹300,000 (roughly $120 to $3,600), with repayment tenures spanning from 15 days to 18 months. The average loan amount is approximately ₹35,000-40,000 ($420-480).

The lender's primary value proposition centers on rapid loan processing and disbursement, with the ability to approve and disburse loans within minutes of application submission. This speed advantage is crucial in the competitive Indian FinTech market, where traditional banks often require weeks for loan processing. The company processes thousands of loan applications daily and has maintained high customer retention rates, with approximately 70% of borrowers returning for additional loans.

2.2 Call Detail Records and Social Monitoring System

The most distinctive feature of this institutional setting is the lender's use of call detail records (CDRs) as an implicit borrower monitoring system. Prior to the January 2019 privacy policy change described below, CDRs were central to the lender's alternative credit evaluation framework, enabling analysis of borrowers' social networks and communication patterns.

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

Second, communication pattern analysis provided dynamic insights into borrower behavior around loan events (Agarwal, Alok, Ghosh and Gupta, 2020). The system tracked communication frequency, duration, and directionality with each relationship category, both before and after loan origination. Changes in communication patterns could signal financial stress, social isolation, or attempts to avoid collections contact. For instance, decreased communication with family members post-loan might indicate relationship strain due to financial difficulties.

Third, the temporal monitoring capability enabled the lender to observe real-time changes in borrower social behavior throughout the loan lifecycle. The system continuously updated communication metrics, allowing for early warning signals of potential default risk. Borrowers

who exhibited declining social connectivity or changed communication patterns could be flagged for proactive intervention or modified collection strategies.

Most critically, the CDR system created an implicit social collateral mechanism. The comprehensive mapping of social networks served as an implicit threat that non-payment could result in contact with family, friends, and professional associates. This social pressure mechanism potentially reduced default rates by leveraging reputational concerns and social shame as compliance incentives. The mere knowledge that the lender possessed detailed contact information for their entire social network could influence borrower repayment behavior even without actual contact occurring.

The monitoring system also enabled sophisticated collection strategies. Rather than relying solely on direct borrower contact, the lender could strategically engage different relationship categories based on communication patterns and relationship strength. Family members with frequent, long-duration calls might be prioritized as potential sources of repayment assistance, while professional contacts could be leveraged to create workplace pressure for repayment.

2.3 Changes to Privacy Policy Change

In January 2019, Google implemented a privacy policy directive that fundamentally altered this data collection landscape. The directive specifically prohibited the lender from continuing to collect CDRs from their customers' mobile devices through their Android application. This policy change was enforced through Google's platform control mechanism. Non-compliance would result in the immediate removal of the lender's application from the Google Play Store, effectively eliminating their primary customer acquisition and service delivery channel.

The enforcement mechanism was particularly significant given that the vast majority of the lender's user base accessed services through Android devices via the Google Play Store. Faced with this ultimatum, the lender had no viable alternative but to comply immediately. Consequently, the collection of call log data and related CDR information ceased entirely on January 1, 2019.

The impact of this change was significant. The lender lost access to what had been their most sophisticated monitoring and disciplining tool. The social collateral mechanism that had potentially suppressed default rates through implicit social pressure was eliminated. As a result, collection strategies had to be redesigned to focus more heavily on ex-ante screening rather than ex-post network-based monitoring.

This policy implementation created several important conditions for empirical analysis.

First, the timing was exogenous to both borrower behavior and the lender's business strategy, as it was imposed by an external platform provider rather than emerging from internal business decisions or broader regulatory changes. Second, the binary nature of the change – from comprehensive CDR access to complete prohibition – creates a sharp discontinuity in data availability that affected all loan applications processed after the implementation date. Third, the lender's core business model, target market, and operational procedures remained otherwise unchanged, ensuring that the policy change represents the primary source of variation in lending outcomes during the study period. Overall, the research design benefits from the clean identification provided by this regulatory discontinuity, which allows us to identify the relationship between privacy protection measures and financial inclusion outcomes.

It is worth noting that while iOS users were unaffected by Google's privacy policy directive, we cannot exploit differences between iOS and Android users as an alternative identification strategy due to insufficient variation in platform usage across our sample.

3 Data

Our empirical analysis leverages a uniquely rich dataset spanning 2017–2022 from one of India's largest and most innovative FinTech lenders. The scale, granularity, and multidimensional nature of this dataset allow us to assess the effect of privacy regulations on lending outcomes in one of the most rapidly expanding credit markets globally. The dataset captures millions of loan applications observed both before and after the January 2019 Google privacy policy change, which restricted access to call detail records. This exogenous shock allows us to study, at scale, how the removal of alternative data reshaped both borrower and lender behavior.

Customer Characteristics Data. The backbone of our dataset is a comprehensive record of every loan application submitted to the FinTech. Each observation includes unique customer and loan identifiers, detailed loan-level information (loan type, requested amount, interest rate, approval/rejection outcome), and time stamps. Importantly, the dataset also provides rich borrower demographics including income, gender, caste, age, and credit scores. These variables allow us to rigorously examine heterogeneous effects of privacy restrictions across population subgroups, with particular attention to traditionally underserved and marginalized communities that the Fin-

Tech explicitly sought to serve. The scale and detail of these demographic data make it possible to quantify distributional consequences of privacy regulation.

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

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

The combination of (i) large-scale loan application data with detailed demographic attributes, and (ii) comprehensive credit bureau coverage of the formal lending sector, provides a valuable empirical foundation for examining privacy regulation effects. The availability of both internal lending decisions and external credit market outcomes allows us to investigate not only how privacy restrictions influence lending decisions within a single FinTech, but also whether these effects manifest as broader patterns of financial exclusion or redistribution across the credit ecosystem. This dual-perspective approach contributes to our understanding of how regulatory changes in the digital lending space may affect overall financial inclusion outcomes.

4 Empirical Methodology

Our empirical strategy exploits the sharp discontinuity created by Google's privacy policy directive implemented on January 1, 2019, which prohibited our FinTech lender from collecting CDRs through their Android applications. This 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 to examine the effect of this pri-

vacy shock: (1) a regression discontinuity design that exploits the sharp temporal cutoff in data collection capabilities, and (2) a difference-in-differences framework that examines heterogeneous effects across borrower characteristics. Together, these approaches allow us to identify both the aggregate effects of privacy regulation on credit markets and the differential effects across demographic groups.

4.1 Regression Discontinuity Design

Our primary identification strategy relies on a sharp regression discontinuity design around the January 1, 2019 policy implementation date. We compare loan application outcomes immediately before and after this cutoff date, when the lender's ability to collect alternative data was abruptly terminated. The sharp temporal discontinuity allows us to isolate the effect of privacy regulation from other factors that might influence lending decisions.

We estimate the following baseline regression discontinuity specification:

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

where $Y_{i,t}$ represents the outcome of interest for loan application i submitted on date t. Our main outcome variables include: (1) the natural logarithm of daily loan applications to capture demand effects, (2) a binary indicator for loan approval to measure supply-side responses, and (3) subsequent default rates to examine ex-post risk.

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

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

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

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

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

We verify these assumptions through several diagnostic tests. We examine the density of loan applications around the threshold to detect potential manipulation. We also estimate our specification on predetermined covariates (borrower characteristics, loan amounts, geographic distribution) to verify that these variables do not exhibit discontinuous changes at the threshold, which would suggest sorting or other confounding factors.

4.2 Difference-in-Differences Design

While our RD design identifies the average treatment effect of privacy regulation, we are particularly interested in understanding which borrower groups are most affected by the loss of alternative data sources. We hypothesize that privacy restrictions disproportionately affect borrowers who are

more difficult to evaluate using traditional credit metrics, specifically, younger applicants, lower-income individuals, those with limited credit histories, borrowers from marginalized castes, and women.

To examine these heterogeneous effects, we supplement our RD analysis with a difference-in-differences framework that interacts the policy change with borrower characteristics:

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

where Characteristic_i represents various borrower attributes including low-income status, young age, limited credit history, low-caste background, and female gender. The coefficients β_k capture the dynamic effect of the privacy policy on borrowers with each characteristic relative to a baseline group, with week k=-1 (immediately before policy implementation) serving as the omitted category.

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

The difference-in-differences approach allows us to examine whether the privacy policy's effect varies systematically across borrower groups while controlling for general time trends and location-specific factors. The key identifying assumption is that, absent the privacy policy change, the relative lending outcomes for different borrower groups would have evolved similarly across locations and loan purposes.

The parallel trends assumption is plausible in our setting for three main reasons. First, the policy change was imposed externally and uniformly affected all borrowers simultaneously. Second, the short time window around the policy change minimizes concerns about differential trends. Third, our rich set of fixed effects controls for many potential confounders. Although this assumption is empirically untestable, we present an assessment of the pre-trends that support this assumption.

5 Main Results

Based on the framework detailed in Section 4.1, the Google privacy policy, implemented on January 1, 2019, created a sharp and well-defined discontinuity that serves as a natural experiment to study how restrictions on alternative data affect lending markets. Our findings reveal a nuanced set of borrower and lender responses that fundamentally reshaped market dynamics. The results demonstrate that while privacy regulations may achieve their intended goal of protecting consumer data, they can generate significant unintended consequences for credit access, particularly among the most vulnerable populations.

5.1 Baseline Results

We begin by conducting an aggregate analysis of the credit market around the implementation of the policy directive. Figure 1 plots monthly application volumes, acceptance rates, and default rates between 2017 and 2021. We observe sharp discontinuities around the policy implementation date: loan applications increase abruptly, acceptance rates decline from roughly 80–85% to 60–70%, while default rates remain largely unchanged.

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

5.1.1 Borrower Demand for Privacy Protection

The increase in loan applications provides compelling evidence that borrowers place meaningful value on privacy. Figure 2 presents the regression discontinuity estimate for the logarithm of the daily number of applications, showing a clear discontinuous increase at the policy threshold. Specifically, applications exhibit a smooth trend leading up to the policy change, followed by an immediate and persistent increase afterwards.

The robustness of this demand response is demonstrated in Figure 3, which plots regression

discontinuity estimates across bandwidth choices ranging from 10 to 150 days. The effect remains consistently positive and statistically significant across this entire range, with point estimates stabilizing around 4% increases in daily applications. This stability across different bandwidths around the cutoff date provides confidence that we are capturing the true behavioral response of borrowers rather than spurious variation around the threshold.

We test these responses rigorously in Table 1 through a regression discontinuity estimation using various bandwidth specifications. Across bandwidths from 10 to 30 days, we consistently estimate that the privacy policy increased daily applications by 3.4-6.3%, with all estimates being statistically significant. Appendix Table A.1 demonstrates the robustness of this finding under progressively richer fixed-effect specifications. In the most parsimonious model without any fixed effects, the estimated effect remains positive and statistically significant at 2.9%. Incorporating the full set of ZIP code × loan purpose and day of the week fixed effects further increases the estimate to 3.4%. The consistency of results across these specifications confirms that the increase in applications cannot be explained by compositional shifts in loan demand across geography, product type, or seasonal variation.

This behavioral response has important implications for understanding consumer preferences in digital financial services. The willingness of borrowers to apply for loans from a lender that had publicly committed to enhanced privacy protections suggests that data collection practices represent a meaningful factor in financial decision-making. Moreover, it suggests that borrowers are not just concerned about loan terms and accessibility but privacy considerations as well.

5.1.2 Lender Supply of Credit

While borrowers responded positively to enhanced privacy protections, the policy imposed substantial costs by eliminating access to rich behavioral and social network data that had been integral in determining credit market outcomes. Figure 4 demonstrates the magnitude of the lender's response, showing application acceptance rates before and after the policy change. The discontinuity is stark and immediate: acceptance rates drop precipitously from around 80-85% to approximately 60-70%, representing a decline of up to 25 percentage points. This decline in the acceptance rate is substantial, representing more than a 30% relative decline from the pre-policy baseline, indicating that alternative data sources were fundamental to the lender's business model. Their removal forced a permanent tightening of credit standards. This suggests that call detail

records provided information that was difficult or impossible to obtain through conventional credit scoring mechanisms.

Table 2 provides formal regression discontinuity estimates of this supply response across different bandwidth specifications. The estimates are remarkably consistent, showing reductions in acceptance rates of 24-26 percentage points across all bandwidths from 10 to 30 days. The consistency of these estimates provides strong evidence that the policy change caused an immediate and substantial tightening of credit standards. Appendix Table A.2 demonstrates that this result is not sensitive to the inclusion of progressively richer controls. The coefficient remains stable at approximately -25 percentage points even when controlling for borrower demographics (gender, caste), education, age deciles, credit score deciles, loan maturity, and income deciles. This stability across increasingly demanding specifications provides strong evidence that the policy change caused an immediate and substantial tightening of credit standards.

Figure 8 further demonstrates the robustness of this finding by plotting estimates across a wide range of bandwidth choices. The supply response remains stable and statistically significant across bandwidths from 10 to 150 days, with confidence intervals that exclude zero throughout this range. This temporal stability suggests that the effect represents a permanent shift in lending standards rather than a temporary adjustment period.

5.1.3 Ruling Out Composition Effects

A critical concern for our identification strategy is whether the observed changes in acceptance rates reflect genuine shifts in lending standards or merely changes in the composition of loan applicants. If the privacy policy attracted different types of borrowers to apply for loans, the decline in acceptance rates might simply reflect the screening of a riskier applicant pool rather than tighter lending standards for equivalent borrowers.

Table 3 addresses this concern by examining whether observable borrower characteristics changed discontinuously at the policy implementation date. We test for compositional shifts across seven key dimensions: requested loan amounts, borrower income, debt-to-income ratios, credit scores, caste background, age, and gender. The results provide reassuring evidence that the privacy policy did not substantially alter the composition of loan applicants.

Most borrower characteristics show no statistically significant changes at the policy threshold. Requested loan amounts, income levels, debt-to-income ratios, credit scores, and gender composition all remain stable across the discontinuity. We do observe two modest compositional

changes: a 3.1 percentage point increase in the proportion of applicants from low-caste backgrounds and a small decrease in average applicant age. However, these effects are economically small relative to the large changes in acceptance rates, suggesting that composition effects cannot explain our main findings.

The absence of substantial composition effects strengthens the interpretation of our results and supports the validity of our regression discontinuity design. The dramatic decline in acceptance rates appears to reflect genuine changes in lending standards rather than shifts in applicant quality, providing clear evidence that the loss of alternative data sources forced the lender to become substantially more selective in its approval decisions.

5.1.4 Effects on Other Credit Terms

The dramatic reduction in acceptance rates raises important questions about how the privacy policy affected other dimensions of credit supply. Standard economic theory suggests that if lenders face an upward-sloping supply curve for credit, tighter screening should materialize not only through quantity restrictions but also through price adjustments (Grossman and Stiglitz, 1980). Under this framework, lenders facing increased information asymmetries following the loss of alternative data might respond by charging higher interest rates or offering shorter maturities to limit exposure to risky borrowers.

However, our institutional setting suggests a different supply structure more consistent with a step function: borrowers either receive credit at market rates or are rejected entirely and face an infinite effective interest rate. This binary outcome reflects the FinTech's algorithmic underwriting process, where loan terms are primarily determined by standardized pricing models based on risk categories rather than continuous price discrimination across individual borrowers.

Appendix Table A.3 examines whether the privacy policy affected loan terms for approved borrowers. Panel A shows the effect on annual interest rates, while Panel B presents results for loan maturity measured in months. Across all bandwidth specifications, we find no statistically significant changes in either interest rates or loan maturities following the policy implementation. Point estimates are small in magnitude and generally not significant, suggesting that conditional on approval, the lender did not adjust pricing to compensate for lost information.

Having established that the policy primarily operated through quantity rather than price adjustments, we next explore the distributional consequences of this exclusion across different borrower groups.

5.2 Heterogeneous Effects: Distributional Consequences of Privacy Regulation

This section examines the distributional consequences of privacy regulation, focusing on how these policies may contribute to financial exclusion. While the aggregate results demonstrate significant market-level responses to privacy regulation, it is equally important to identify which borrowers bore the costs of these changes. The welfare implications of such policies depend not only on overall shifts in credit supply but also on their distributional consequences.

To understand why privacy restrictions might have differential effects across borrower types, it is helpful to consider the fundamental challenge that lenders face when serving underbanked populations. Traditional credit scoring relies heavily on established credit histories, stable employment relationships, and formal financial records. For many potential borrowers, particularly younger individuals, those with lower incomes, or people from marginalized social backgrounds, these traditional indicators may be sparse or entirely absent.

For borrowers with extensive traditional credit histories, these alternative data sources might provide only marginal additional information. However, for borrowers with limited formal financial footprints, alternative data could be the primary basis for credit evaluation.

This leads to clear predictions about which borrowers should be most affected by privacy restrictions. We should expect larger negative effects for borrowers with limited credit histories, those without prior lending relationships, younger applicants, lower-income individuals, and members of marginalized social groups. We test this hypothesis by examining how the same privacy regulation can affect the availability of credit for these vulnerable borrower segments.

Our empirical analysis provides strong evidence consistent with this prediction. Using a difference-in-differences framework from Section 4.2 that examines how the policy's effects varied across borrower characteristics, we document systematic patterns of differential effects that reveal the privacy regulation's unintended adverse consequences for financial inclusion. The results indicate that the borrowers most negatively affected by privacy restrictions are exactly those who face the greatest barriers to traditional credit access.

Figure 6 presents dynamic difference-in-differences estimates of the policy's effect on loan acceptance rates over time, highlighting substantial heterogeneity across borrower groups. We first examine the differential effects by income level in Panel 6a. Low-income borrowers, defined as those with monthly incomes below the sample median, experienced substantially larger reductions in acceptance rates following the policy change. Estimates indicate declines of roughly 8–10

percentage points relative to higher-income applicants, with the effect emerging immediately after the policy and persisting throughout the observation window. This pattern is consistent with the idea that lower-income borrowers are more likely to have thin credit files and thus rely disproportionately on alternative data for credit evaluation. In addition, smaller loan sizes for this population may make manual underwriting economically infeasible, further amplifying their dependence on algorithmic assessment.

Age similarly predicts differential effects in Panel 6b. Younger borrowers, defined as those below the sample median age, faced declines in acceptance rates of approximately 7–10 percentage points relative to older applicants. This finding is notable because younger borrowers typically have shorter credit histories and fewer established financial relationships, making them more dependent on nontraditional data sources. The disproportionate effect on this group suggests that privacy restrictions may limit initial access to formal credit markets, with potential long-term consequences for financial inclusion.

The effects are even more pronounced when comparing new and existing customers in Panel 6c. First-time applicants experienced reductions in acceptance rates of 12–15 percentage points relative to borrowers with established lending relationships. Alternative data are particularly valuable for new customers, as lenders lack internal repayment histories or prior behavioral data. Consequently, privacy restrictions pose a significant barrier to customer acquisition and may increase concentration in the digital lending market by favoring established platforms with large existing customer bases.

We also find substantial differential effects across social identity groups in Panel 6d. Borrowers from historically marginalized castes—including scheduled castes, scheduled tribes, and other backward classes—experienced reductions in acceptance rates of roughly 8–10 percentage points relative to general caste borrowers. These findings suggest that privacy regulations may inadvertently exacerbate existing patterns of economic exclusion, as the removal of alternative data forces lenders to rely more heavily on conventional credit metrics that reflect historical inequities.

Gender, by contrast, appears to play a more limited role, as shown in Panel 6e. Female borrowers experienced slightly larger reductions in acceptance rates than male borrowers, though these differences are generally not statistically significant. Similarly, differences by credit score are modest, as shown in Panel 6f. Borrowers with subprime credit scores (below 680) show larger point estimates of declines in acceptance rates, but these effects are not statistically significant, potentially reflecting the limited sample size of subprime borrowers or the fact that even borrowers

with established credit histories benefited from the information provided by alternative data.

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

Table 4 presents the regression estimates that complement the graphical evidence from Figure 6. The regression results confirm the heterogeneity in the policy's effect across different borrower groups. Specifically, the baseline reduction in application acceptance rates documented in Table 2 is primarily driven by marginal borrowers. New customers experienced a 12–14 percentage point larger decline in acceptance rates, while low-income borrowers faced an 8–9 percentage point larger reduction. Similarly, younger borrowers saw a 7–8 percentage point decline, and low-caste borrowers experienced a 7–9 percentage point larger effect. All of these estimates are statistically significant and economically meaningful. Appendix Table A.4 further demonstrates the robustness of these findings to more stringent identification strategies, including the addition of ZIP code × week and loan purpose × week fixed effects, confirming that these differential effects are not driven by time-varying factors correlated with borrower characteristics.

Lastly, one may be concerned that several alternative explanations could potentially account for the observed changes in lending behavior beyond privacy regulation effects, including deteriorating financial conditions of the lender, legal challenges or litigation losses, or media coverage following the Google directive, among other effects. Multiple factors support our interpretation that privacy regulation drove the observed effects. The lender continued operating and expanding throughout our study period with no evidence of financial distress. The sharp timing of behavioral changes, occurring immediately on January 1, 2019, coincides precisely with policy implementation rather than gradual business deterioration or litigation timelines. The lender maintained stable default rates post-policy, indicating preserved credit quality standards rather than relaxed screening due to financial pressure. Finally, the heterogeneous effects across borrower groups are more consistent with information-driven selection than with across-the-board credit tightening that would accompany financial distress or publicity-driven application surges that would affect all borrower types similarly.

6 Implications of the Tightening of Credit Standards

This section explores the underlying mechanism through which the loss of alternative data sources affects a lender's business model and risk management. The withdrawal of CDRs fundamentally altered the lender's approach. Prior to the policy change, loan performance was supported by a business model that primarily relied on monitoring. CDRs enabled dynamic risk assessment, after the loan disbursement, and early intervention strategies throughout the loan lifecycle (Agarwal, Alok, Ghosh and Gupta, 2020). With this data no longer available, the lender was forced to shift to a model that relied almost exclusively on upfront screening at the time of application.

One way of testing this hypothesis is through a comparison of default rates. If the lender successfully adapted to the loss of monitoring capabilities by tightening screening standards, we should observe no change in portfolio risk despite the dramatic reduction in approval rates. Conversely, if the screening adjustments were insufficient to compensate for lost monitoring capabilities, default rates should increase. Figure 7 plots the regression discontinuity estimates of loan default, showing no discontinuity at the policy boundary. Defaults remained close to 20% before and after January 1, 2019.

Table 5 formalizes this finding through regression discontinuity estimation across different bandwidth specifications. The point estimates for the policy's effect on default rates are small in magnitude and statistically insignificant across all specifications. This null result confirms that the lender successfully maintained portfolio quality despite the loss of alternative data sources. Appendix Table A.5 demonstrates the robustness of this null result to the progressive inclusion of individual-level controls. Even when controlling for borrower demographics, education, age deciles, credit score deciles, loan maturity, and income deciles, the policy effect on default rates remains statistically insignificant and economically small across all specifications.

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

This stability indicates that the tightening of screening standards fully offset the loss of ex-post monitoring. Acceptance rates declined sharply, but portfolio quality was preserved. The combination of stable default rates with lower loan volumes implies that many rejected borrow-

ers would have repaid successfully, consistent with a shift toward more conservative approval thresholds.

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

Our analysis of long-term outcomes is enabled by the integration of our primary dataset with comprehensive credit bureau records that capture formal lending activity across all major financial institutions in India. This external data source allows us to observe whether borrowers who applied around the policy change date subsequently accessed credit from other lenders, providing a complete picture of their financial market participation. We examine whether the privacy policy change leads to persistent financial exclusion, rather than temporary displacement to alternative lenders. If privacy restrictions simply redistributes borrowing activity across lenders, i.e., borrowers denied credit by our focal FinTech obtain loans from other institutions, then the welfare implications would be relatively modest. However, if obtaining a Fintech loan facilitates access to credit from conventional banks – either by generating new information (Balyuk, 2023; Dubey, 2024) or helping to build collateral (Beaumont, Tang and Vansteenberghe, 2022) – then rejected borrowers due to privacy regulation might face systematic exclusion from formal credit markets. Our setting provides a unique opportunity to quantify this "FinTech ladder effect," measuring how initial FinTech loan acceptance influences borrowers' future credit prospects across the entire financial system.

Table 6 presents reduced-form evidence on how the policy change affected borrowers' subsequent access to formal credit markets. We track borrowers for 2, 3, and 4 years following the January 1, 2019 policy implementation date, measuring whether they accessed any form of formal credit, including loans from banks, non-bank financial companies, and other FinTech platforms, during these extended periods. The results reveal persistent negative effects on credit access that extend far beyond the immediate policy effect. Borrowers who applied after the policy change were 3.9-4.7 percentage points less likely to access formal credit within 2 years compared to those

who applied before the policy. These effects persist and even strengthen over longer time horizons: applicants were 4.5-5.3 percentage points less likely to access credit within 3 years and 4.1-5.0 percentage points less likely within 4 years. Appendix Table A.6 shows that these reduced-form results are robust to alternative bandwidth specifications, with effects ranging from 1.8-3.8 percentage points when using a 15-day bandwidth.

The persistence and magnitude of these effects are striking. Given baseline credit access rates of 36% in two-years in our sample, these reductions represent substantial relative declines in formal financial market participation, consistent with prior evidence that early credit rejections can scar borrowers' long-run outcomes (Dobbie and Skiba, 2013; Dobbie, Goldsmith-Pinkham, Mahoney and Song, 2020; Liberman, Neilson, Opazo and Zimmerman, 2018). The fact that effects persist and remain statistically significant even four years after the policy change suggests that being denied credit due to privacy restrictions has lasting consequences for borrowers' financial trajectories.

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

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

$$\textbf{First Stage:} \quad \text{App Accepted}_{i,t} = \pi_0 + \pi_1 \cdot Post_t + X'_{i,t}\pi_2 + \theta_z + \theta_w + \nu_{i,t}$$

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

where $\text{Credit Access}_{i,t+s}$ is a binary indicator equal to one if borrower i accessed formal credit within s years following January 1, 2019 (we examine horizons of 2, 3, and 4 years). $\text{App Accepted}_{i,t}$ indicates whether the loan application was approved by our focal FinTech. The instrument Post_t indicates whether the loan application was submitted after the privacy policy.

Conceptually, the IV estimates capture the local average treatment effect (LATE) of loan

approval for borrowers whose applications would have been approved before the policy change but were rejected afterward due to privacy restrictions. This parameter is particularly policy-relevant because it identifies the credit access consequences for the marginal borrowers most directly affected by privacy regulation.

The validity of our instrumental variables approach relies on two key assumptions. The relevance condition requires that the privacy policy significantly affects loan approval decisions. The first-stage relationship between the policy change and loan approval is strong and economically meaningful, as the privacy policy reduced approval rates by approximately 25 percentage points (see Table 2). This specification yields first-stage F-statistics that substantially exceed conventional thresholds for weak instruments, as indicated in Table 7. This strong first stage provides confidence that our instrumental variables estimates are not subject to weak instrument bias.

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

The second-stage results reveal economically large effects of loan approval on future credit access, as shown in Table 7. Borrowers whose applications were approved were 10-12 percentage points more likely to access formal credit over the subsequent 2-4 years compared to similar borrowers whose applications were rejected. These effects are highly statistically significant and economically meaningful, representing substantial improvements in long-term financial market participation. Appendix Table A.7 shows that these IV results are robust to alternative bandwidth specifications, with effects ranging from 6.5-8.4 percentage points when using a 15-day bandwidth.

The instrumental variables estimates are larger in magnitude than the reduced-form effects, which is consistent with local average treatment effects that capture the effect of loan approval for the marginal borrowers most affected by the policy change. These are precisely the borrowers for whom alternative data sources provided the most valuable information – those with limited traditional credit profiles who were excluded when the lender shifted to conventional screening methods.

Overall, our findings suggest that privacy regulation can cause persistent financial exclu-

sion, in line with evidence that credit rejections can both transmit negative signals to credit bureaus (Hertzberg, Liberti and Paravisini, 2011) and discourage borrower engagement with formal markets (Banerjee, Duflo, Goldberg, Karlan, Osei, Parienté, Shapiro, Thuysbaert and Udry, 2015). These findings raise important questions about the mechanisms through which initial credit denials generate long-term consequences. Several complementary factors likely contribute to this pattern. First, other lenders in the formal financial sector may face similar information constraints when alternative data sources become unavailable or when privacy regulations limit data collection more broadly. This interpretation gains credibility from subsequent regulatory developments: in May 2025, the Reserve Bank of India (RBI) issued Reserve Bank of India (Digital Lending) Directions requiring digital lending apps to desist from accessing mobile phone resources such as file and media, contact lists, call logs, and telephony functions. This broader shift in regulatory policy suggests that our findings may reflect the early stages of a sector-wide transition away from alternative data sources, creating systematic tightening of credit standards across multiple lenders rather than isolated effects at our focal FinTech. Second, initial loan rejections may damage borrowers' formal credit records in ways that affect future applications. In many credit bureau systems, loan applications and rejections are recorded and may be visible to future lenders. A pattern of rejections could signal heightened risk to subsequent lenders, creating a self-reinforcing cycle of exclusion where initial privacy-induced denials lead to reduced creditworthiness in the eyes of other potential lenders. Third, borrowers may respond to initial rejections by reducing their engagement with formal financial markets. The experience of being denied credit, particularly for borrowers who may already feel marginalized by traditional financial institutions, could lead to discouragement and withdrawal from formal lending markets. This behavioral response could generate persistent exclusion even if other lenders would be willing to extend credit, as affected borrowers may simply stop applying for formal loans after experiencing rejection.

7 Quantifying the Economic Costs of Screening

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

marginal applicants ex-ante. To quantify the magnitude of this tightened screening, we develop a counterfactual framework that estimates the foregone profits from rejecting borrowers who would have been profitably served absent the policy change. This analysis provides important context for evaluating the welfare implications of privacy regulation in credit markets.

7.1 Methodology

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

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

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

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

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

Our baseline analysis identifies marginal borrowers within the 60-day window around the policy implementation—borrowers who were rejected post-policy despite having predicted approval probabilities above 50% under pre-policy models. Table 8 summarizes the expected profit margins across different borrower characteristics.

Panel A shows pronounced heterogeneity in profitability. Margins range from 1.50% for small loan amounts to 3.58% for female borrowers, with interquartile ranges spanning between 2–5 percentage points within groups. Panel B reports group-level averages weighted by loan size, which better capture the total economic effect of tightened screening. These dollar-weighted averages are systematically higher than simple means, consistent with larger loans tending to be more profitable. Female borrowers exhibit the highest average margin (5.54%), followed closely by low credit score (5.44%) and low caste borrowers (5.24%). Notably, most disadvantaged groups, precisely those targeted by financial inclusion policies, exhibit margins above 5%, suggesting that the loss of alternative data curtailed highly profitable lending opportunities to underserved populations.

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

7.2 Heterogeneous Effects

This section examines how profit margins vary across borrower characteristics, addressing a complementary but distinct question from our earlier heterogeneous effects analysis. In Section 5.2, we documented which borrower groups experienced the largest reductions in approval rates following the policy change, identifying populations most harmed in terms of credit access. Here, we analyze which excluded borrowers would have generated the highest profit margins had they been approved—identifying populations that represented the most valuable foregone business opportunities.

These two perspectives capture different dimensions of the economic effect. A group might experience large approval rate declines (indicating heavy reliance on alternative data for access)

while excluded members of that group generate relatively modest profit margins. Conversely, another group might show smaller approval effects but higher profit margins among those excluded. Understanding both patterns is essential for evaluating the full welfare consequences of stringent screening.

Figure 9 illustrates how profit margins vary across borrower characteristics. The results reveal systematic patterns in foregone profitability that complement our earlier findings about differential access effects. Panel 9a shows that profit margins increase steadily with income, with higher-earning borrowers exhibiting margins around 4% compared to under 2% for the lowest earners. This pattern suggests that alternative data sources were valuable across the income distribution, but particularly so for higher-income borrowers who may have represented larger, more profitable loan opportunities.

Margins also increase with age, as shown in Panel 9b, from under 3% among the youngest applicants to more than 3.5% for older borrowers, consistent with the view that alternative data complemented traditional signals for those with more established credit histories. Loan size exhibits a strong gradient in Panel 9c: margins are negligible for very small loans but climb to nearly 7% for the largest. This pattern reflects both economies of scale in processing larger loans and the higher absolute returns available on substantial loan amounts.

By contrast, margins decline with credit score in Panel 9d, consistent with the idea that alternative data was most valuable for borrowers with weaker formal credit profiles. Gender differences are also notable, as female borrowers exhibit systematically higher profitability than their male counterparts in Panel 9e. For caste, the pattern is more muted. We find that non-low caste borrowers show slightly higher margins than low caste borrowers in Panel 9f, suggesting that profitability gains from alternative data were not disproportionately concentrated among marginalized caste groups.

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

7.3 Sensitivity to Key Assumptions

Our analysis makes several critical assumptions about borrower risk characteristics, discount rates, and the definition of marginal borrowers that could potentially influence our estimates of foregone profits. To assess the robustness of our findings, we conduct a comprehensive sensitivity analysis

examining how our results respond to alternative parameter values across four key dimensions: the probability threshold used to identify marginal borrowers, assumptions about their default rates, the discount rate applied to future cash flows, and the loss given default rate.

Figure 10 demonstrates that our conclusions are robust across reasonable parameter ranges. First, our results are relatively insensitive to the probability threshold used to define marginal borrowers (Figure 10a), with foregone profits declining gradually from about 31% at a 0.3 threshold to 15% at a 0.8 threshold while remaining economically meaningful throughout this range. The stability around our baseline 0.5 threshold provides confidence that our identification of marginal borrowers is not driven by arbitrary cutoff choices. Second, our findings remain positive unless we assume marginal borrowers have default rates exceeding 4.5 times the baseline rate (Figure 10b) which would imply default rates above 85%, an extreme assumption well beyond reasonable bounds for consumer lending. Third, results are robust to discount rate assumptions, remaining positive even at monthly discount rates of 2%, equivalent to 27% annually (Figure 10c, though substantially reduced at these high rates. Fourth, our conclusions hold unless loss given default rates exceed 4.5 times the baseline (Figure 10d, implying losses above 40% of loan amounts.

The sensitivity analysis reveals that our baseline estimate of 29% foregone profits represents a conservative assessment of the economic effect. Under more aggressive but still reasonable assumptions, the estimated costs of tightened screening would be even larger, reinforcing our conclusion that alternative data sources were fundamental rather than merely supplementary to the lender's business model.

7.4 Robustness: Alternative Loan Structure

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

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

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

Neither framework is fully realistic. The bullet loan specification abstracts from amorti-

zation and concentrates all cash flows at maturity, which simplifies the analysis and makes the economic tradeoff between interest income and default losses fully transparent. The EMI approach, while closer to standard repayment patterns, introduces additional assumptions about the timing of defaults and remaining principal balances that cannot be validated without detailed loan performance data. As a result, the EMI model embeds a different set of stylized assumptions.

The EMI results yield smaller estimates of foregone profits of almost 12% of realized profits compared to almost 30% under the bullet specification, but the cross-sectional heterogeneity patterns remain remarkably consistent. Female borrowers continue to exhibit the highest profit margins at 2.17%, compared to 5.54% in the bullet model, while low amount borrowers show the lowest margins in both specifications, shown in Appendix Table A.8. The relative ranking across all demographic groups is unchanged, presented in Appendix Figure A.1, suggesting that our conclusions about which populations were most harmed by tightened screening are robust to loan structure assumptions.

The sensitivity analysis under the EMI framework in Appendix Figure A.2 reveals that foregone profits turn negative at more conservative stress scenarios, such as default or LGD rates exceeding 2.5 times baseline compared to 4.5 times in the bullet model, but results remain economically meaningful across reasonable parameter ranges. Even under the more conservative EMI assumptions, marginal borrowers would need to exhibit default rates above 47.5% ($2.5 \times 19\%$) or loss rates above 22.5% ($2.5 \times 9\%$) before becoming unprofitable, thresholds that remain outside typical bounds for consumer lending.

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

8 Conclusion

This paper examines the unintended consequences of privacy regulation for financial inclusion by analyzing how Google's 2019 restriction on call detail record collection affected a major Indian FinTech lender. Our findings reveal a fundamental tension in digital finance policy: while privacy protections enhance consumer welfare and increase loan demand, they can simultaneously lead to substantial credit supply contraction. The policy reduced acceptance rates by 25 percentage points while maintaining stable default rates, indicating that privacy restrictions primarily affect

ex-ante screening rather than ex-post monitoring mechanisms. This distinction has important implications for understanding how alternative data supports lending decisions and for designing optimal privacy regulations.

The distributional consequences of privacy regulation have significant effects on financial inclusion. Our analysis demonstrates that restrictions disproportionately harm vulnerable populations. Low-income borrowers, younger applicants, first-time customers, and individuals from marginalized castes experienced 7-15 percentage point larger declines in credit access. These effects persist over multiple years, with initially rejected borrowers remaining 4-5 percentage points less likely to access formal credit even four years later. The economic magnitude is substantial: we estimate that privacy restrictions eliminated nearly 30% of the lender's potential profits, equivalent to approximately ₹100 million annually, by excluding creditworthy borrowers who would have successfully repaid their loans.

Our findings highlight the need for a more nuanced approach to privacy regulations. While limiting access to sensitive consumer data is essential for protecting privacy, blanket restrictions on alternative data usage can unintentionally widen financial disparities by disproportionately excluding borrowers without extensive credit histories. Policymakers should explore frameworks that balance data protection with financial inclusion, such as responsible guidelines for alternative credit scoring, consent-based data sharing mechanisms, or privacy-preserving machine learning techniques that allow lenders to monitor risk without compromising consumer privacy. Regulations should not only safeguard data but also ensure that they do not unintentionally restrict access to credit for those most in need. As digital finance continues expanding globally, understanding these tradeoffs becomes increasingly critical for promoting both consumer protection and inclusive economic growth.

References

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

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

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

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

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

Athey, Susan, Christian Catalini, and Catherine Tucker. 2017. "The digital privacy paradox: Small money, small costs, small talk." National Bureau of Economic Research.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Dubé, Jean-Pierre, John G Lynch, Dirk Bergemann, Mert Demirer, Avi Goldfarb, Garrett Johnson, Anja Lambrecht, Tesary Lin, Anna Tuchman, and Catherine E Tucker. 2025. "The intended and unintended consequences of privacy regulation for consumer marketing."

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

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

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

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

Grossman, Sanford J, and Joseph E Stiglitz. 1980. "On the impossibility of informationally efficient markets." *The American economic review*, 70(3): 393–408.

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

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

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

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

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

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

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

Miller, Amalia R, and Catherine Tucker. 2009. "Privacy protection and technology diffusion: The case of electronic medical records." *Management science*, 55(7): 1077–1093.

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.

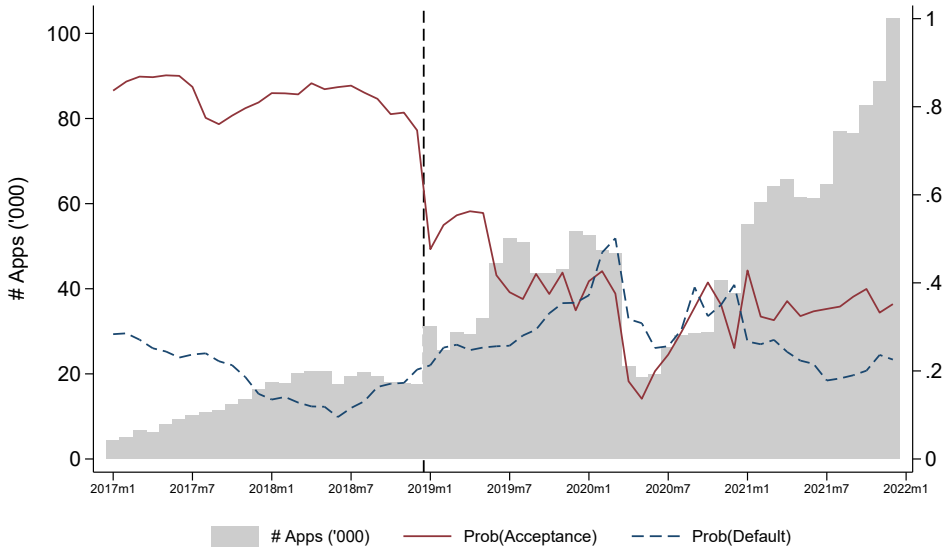
Prince, Jeffrey T, and Scott Wallsten. 2022. "How much is privacy worth around the world and across platforms?" *Journal of Economics & Management Strategy*, 31(4): 841–861.

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

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

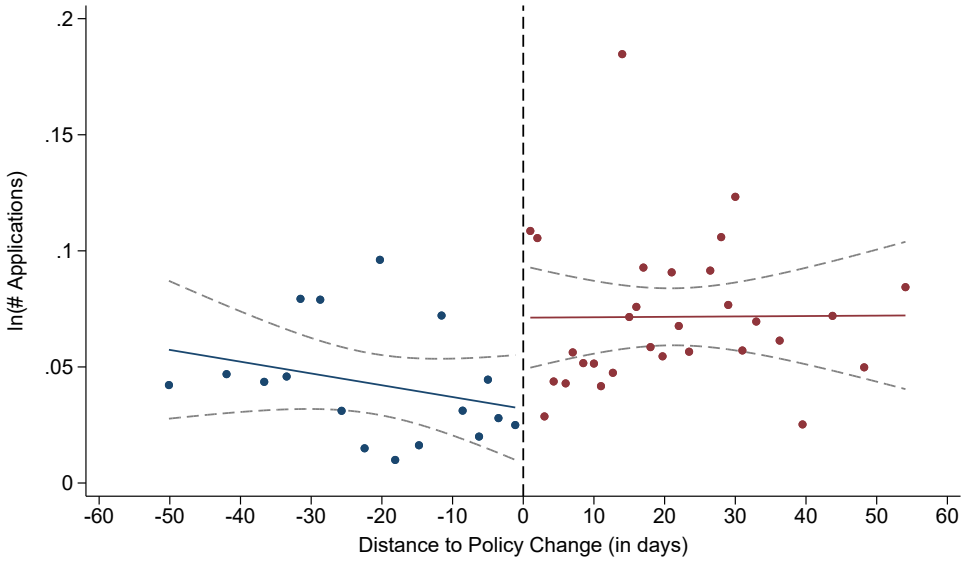
Thakor, Anjan V. 2020. "Fintech and banking: What do we know?" *Journal of financial intermediation*, 41: 100833.

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



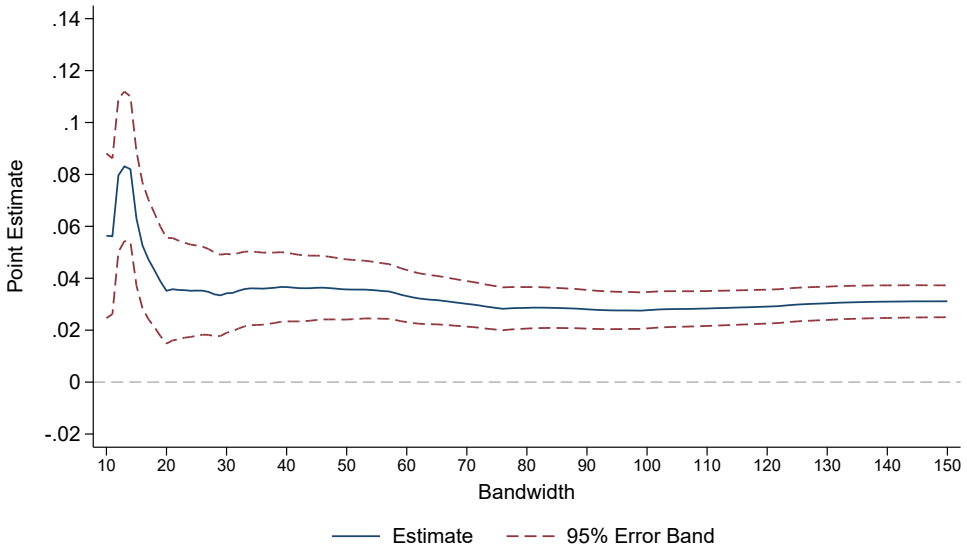
This figure illustrates the number of monthly applications, the monthly average probability of acceptance, the monthly average probability of default from January 2017 until December 2021.

Figure 2: Effect on Number of Applications: RD Estimate



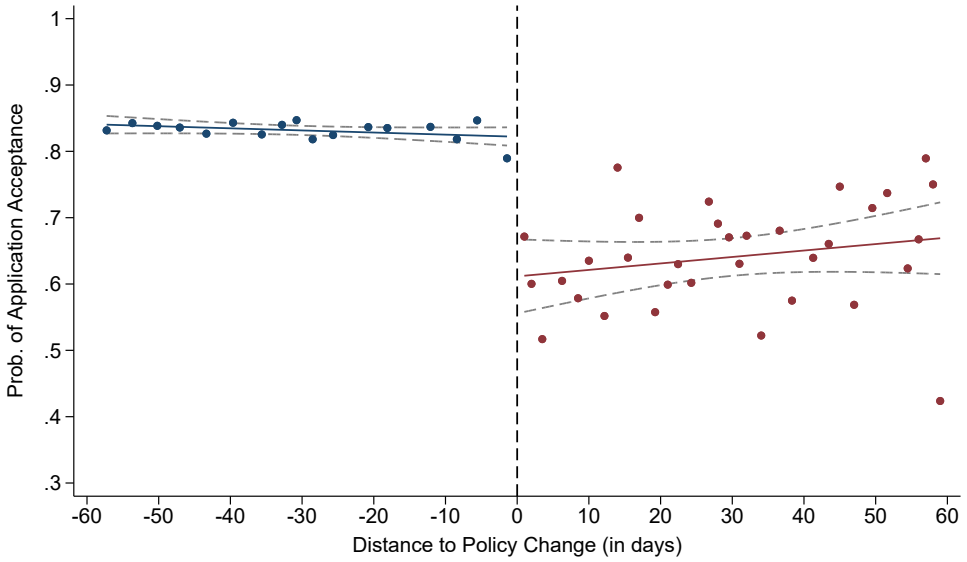
The figure presents RD plots for natural logarithm of the number of applications during the day at different bins along the running variable (distance to policy change) as well as with a local linear trend estimated separately, along with 95% confidence intervals, on each side of the discontinuity. The figure is estimated after including Zip Code × loan reason and the day of the week fixed effects. The unit of observation is a day. The red dots indicate days after the call log collection ban policy and the blue dots indicate days before the call log collection ban policy was implemented.

Figure 3: Number of Application: RD Estimates for Various Bandwidth Choices



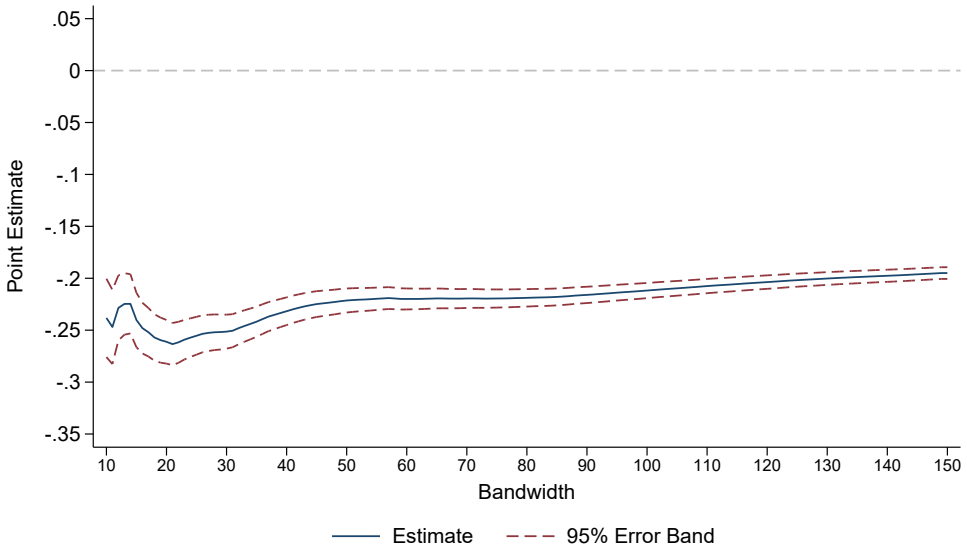
This figure presents RD estimates and 95% confidence interval for the RD specification using the natural logarithm of number of applications during the day as the dependent variable. We estimate the RD specification for several bandwidths between 2 days and 150 days in increments of 1 day. The Y-axis reports the RD estimate, and the X-axis reports the bandwidth used to estimate the RD coefficient. The unit of observation is a day. All regressions include Zip Code × loan reason and the day of the week fixed effects, a local linear specification estimated separately on each side of the policy change date, and use a triangular kernel. The solid red line reports the RD estimate, and the dashed grey lines indicate the 95% confidence bands calculated based on standard errors clustered at the block level.

Figure 4: Effect on Acceptance: RD Estimate



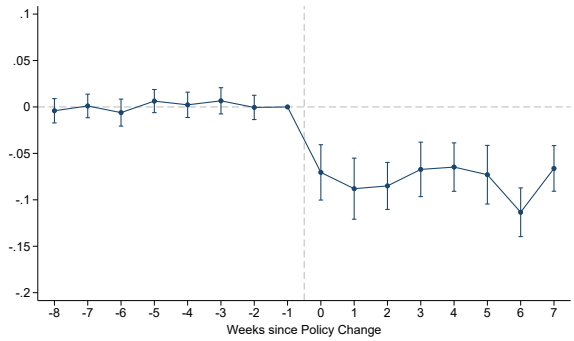
The figure presents RD plots for the probability of acceptance at different bins along the running variable (distance to policy change) as well as with a local linear trend estimated separately, along with 95% confidence intervals, on each side of the discontinuity. The figure is estimated after including Zip Code × loan reason and the day of the week fixed effects. The unit of observation is a day. The red dots indicate days after the call log collection ban policy and the blue dots indicate days before the call log collection ban policy was implemented.

Figure 5: Application Acceptance: RD Estimates for Various Bandwidth Choices

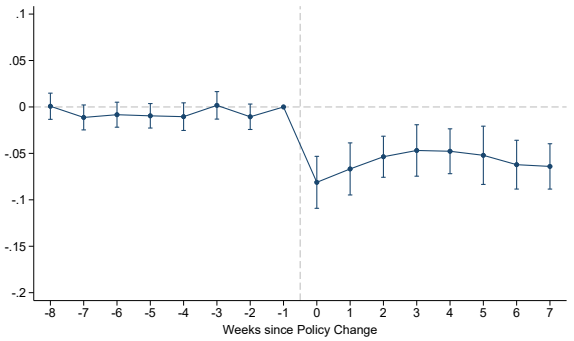


This figure presents RD estimates and 95% confidence interval for the RD specification using application acceptance status as the dependent variable. We estimate the RD specification for several bandwidths between 2 days and 150 days in increments of 1 day. The Y-axis reports the RD estimate, and the X-axis reports the bandwidth used to estimate the RD coefficient. The unit of observation is a day. All regressions include Zip Code × loan reason and the day of the week fixed effects, a local linear specification estimated separately on each side of the policy change date, and use a triangular kernel. The solid red line reports the RD estimate, and the dashed grey lines indicate the 95% confidence bands calculated based on standard errors clustered at the block level.

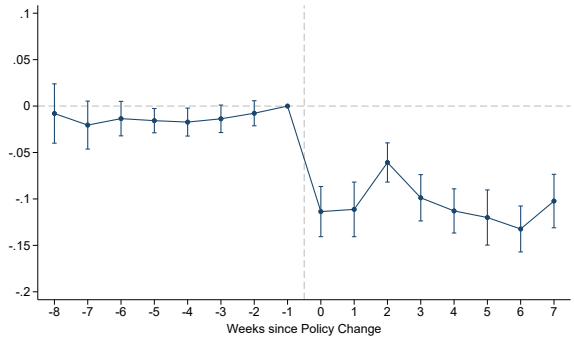
Figure 6: DiD: Heterogeneous Effects on Application Acceptance by Borrower Characteristics



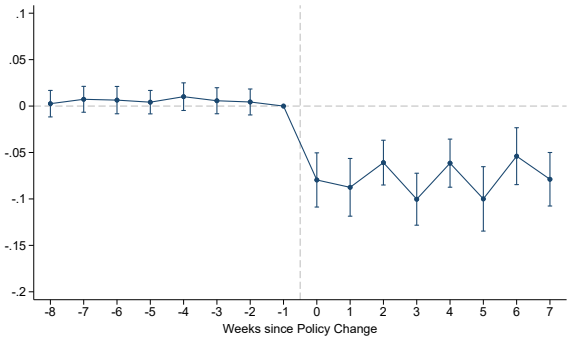
(a) Monthly Customer Income



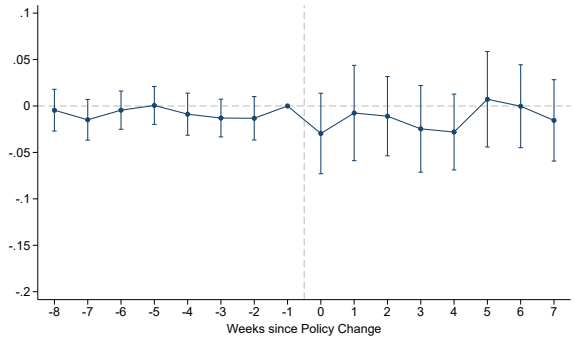
(b) Age



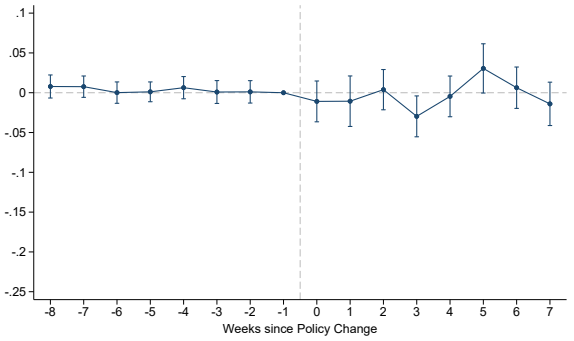
(c) Prior Relationship



(d) Caste



(e) Gender



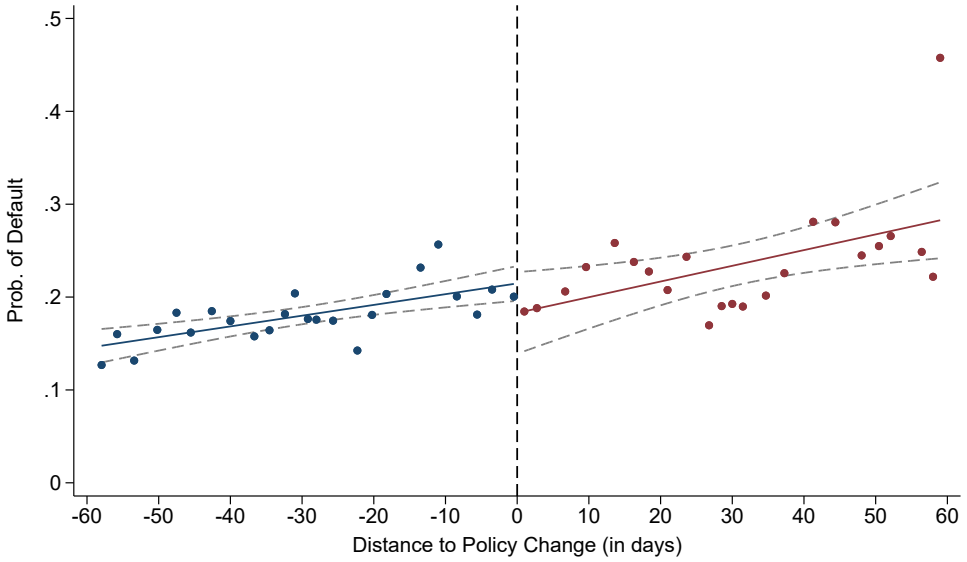
(f) Credit Scores

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

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

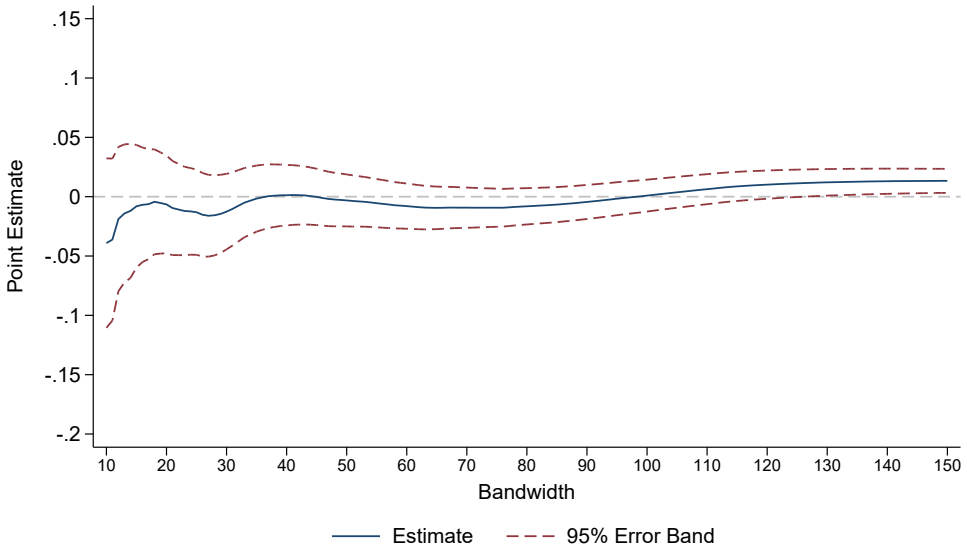
All regressions include ZIP code × loan purpose ($\theta_{z,p}$), ZIP code × week ($\theta_{z,t}$), loan purpose × week ($\theta_{p,t}$), and day of the week (θ_w) fixed effects. Variable definitions: Panel 6a *Low Income* equals one for customers with monthly income below the sample median; Panel 6b *Young* equals one for customers with age below the sample median; Panel 6c *New Customer* equals one for first-time applicants; Panel 6d *Low Caste* equals one for applicants belonging to scheduled caste, scheduled tribe, and other backward class; Panel 6e *Female* equals one for female applicants; Panel 6f *Sub-prime Credit Score* is a binary variable equal to one for customers with CIBIL credit score less than 680. 95% confidence bands are calculated based on standard errors clustered at the ZIP code level.

Figure 7: Effect on Default: RD Estimate



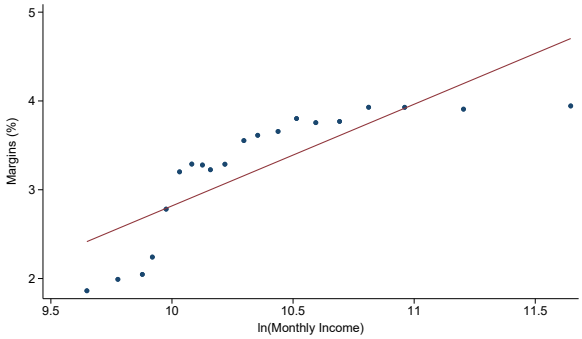
The figure presents RD plots for the probability of default at different bins along the running variable (distance to policy change) as well as with a local linear trend estimated separately, along with 95% confidence intervals, on each side of the discontinuity. The unit of observation is a day. The red dots indicate days after the call log collection ban policy and the blue dots indicate days before the call log collection ban policy was implemented.

Figure 8: Default: RD Estimates for Various Bandwidth Choices

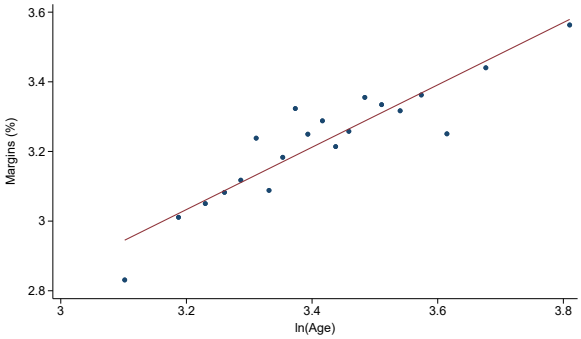


This figure presents RD estimates and 95% confidence interval for the RD specification using loan default status as the dependent variable. We estimate the RD specification for several bandwidths between 10 days and 150 days in increments of 1 day. The Y-axis reports the RD estimate, and the X-axis reports the bandwidth used to estimate the RD coefficient. The unit of observation is a day. All regressions include Zip Code × loan reason and the day of the week fixed effects, a local linear specification estimated separately on each side of the policy change date, and use a triangular kernel. The solid red line reports the RD estimate, and the dashed grey lines indicate the 95% confidence bands calculated based on standard errors clustered at the block level.

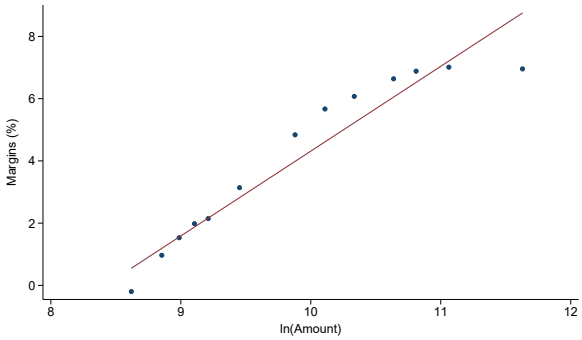
Figure 9: Profit Margins of Marginal Borrowers by Characteristics



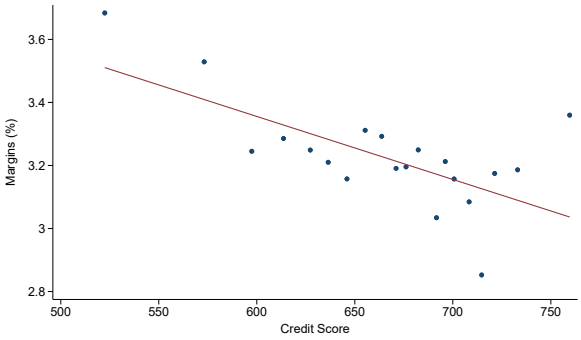
(a) By Income



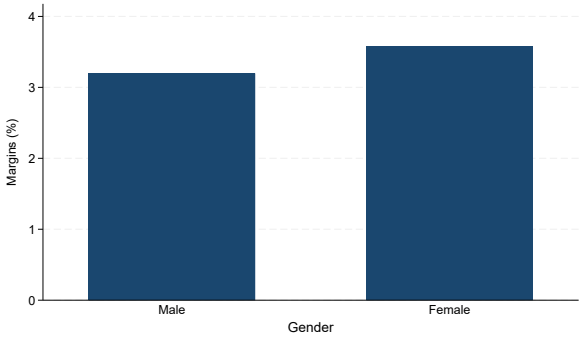
(b) By Age



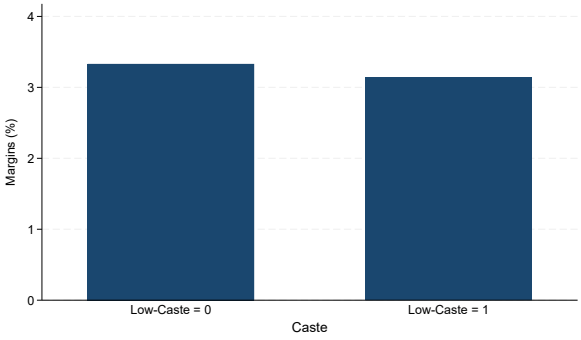
(c) By Loan Amount



(d) By Credit Score



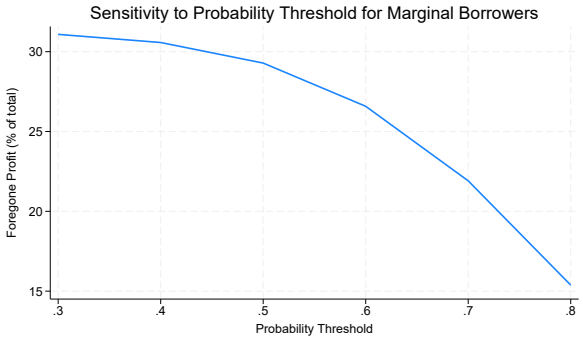
(e) By Gender



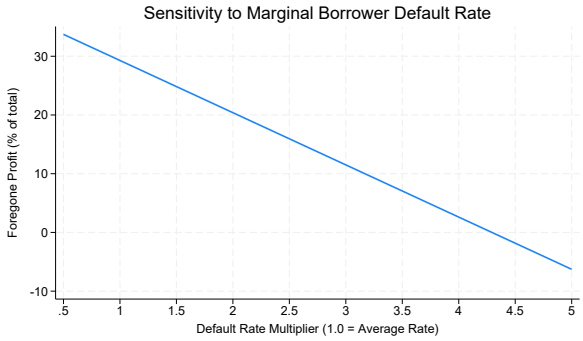
(f) By Caste

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

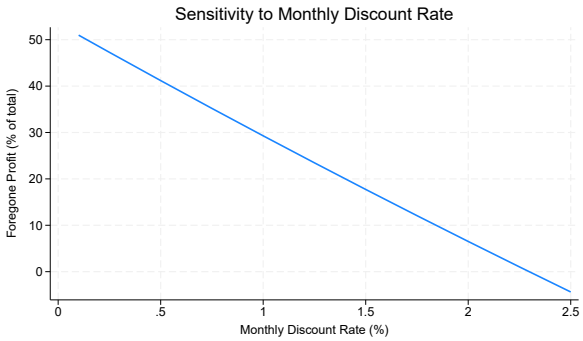
Figure 10: Sensitivity Analysis of Foregone Profits from Privacy Regulation



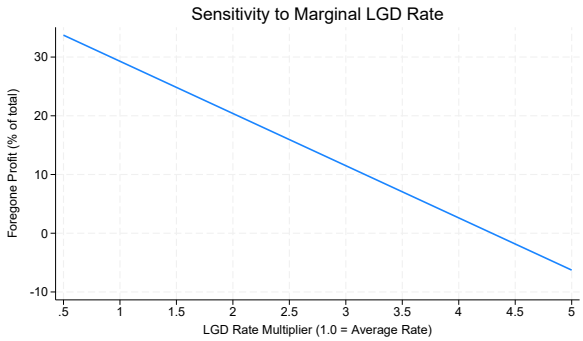
(a) Probability Threshold



(b) Default Rate



(c) Monthly Discount Rate



(d) Loss Given Default

Notes: This figure presents sensitivity analysis for the foregone profit calculations from the privacy policy change. Panel 10a shows how results vary with the probability threshold used to define marginal borrowers (those rejected post-policy but who would have been approved pre-policy). Panel 10b examines sensitivity to assumptions about marginal borrowers' default rates, expressed as multiples of the baseline 19% rate. Panel 10c shows how results depend on the monthly discount rate used in present value calculations. Panel 10d examines sensitivity to the loss given default (LGD) rate, expressed as multiples of the baseline 9% rate. All calculations assume bullet loan structure with interest and principal repaid at maturity. Baseline parameters: 0.5 probability threshold, 19% default rate, 1% monthly discount rate, 9% LGD. Sample includes borrowers within 60-day window around January 1, 2019 policy implementation.

Table 1: Effect of the Policy Change on Number of Applications

Dep Var: ln(# Applications)	(1)	(2)	(3)	(4)	(5)
Post	0.0564***	0.0630***	0.0352***	0.0353***	0.0342***
	(0.0162)	(0.0132)	(0.0104)	(0.0089)	(0.0077)
ZIP Code X Loan Purpose FE	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓
Bandwidth	10	15	20	25	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	6,805	12,280	17,130	23,230	30,494
R^2	0.4162	0.3449	0.304	0.2831	0.2689

This table presents the RD estimates using the natural logarithm of the number of applications during the day as the dependent variable. We use different bandwidths of 10, 15, 20, 25, and 30 days on either side of the policy change date (January 1, 2019), in columns (1), (2), (3), (4), and (5) respectively. All regressions include Zip Code × loan 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. Standard errors reported in parentheses are clustered at the Zip Code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Effect of the Policy Change on Application Acceptances

Dep Var: Application Accepted	(1)	(2)	(3)	(4)	(5)
Post	-0.2382***	-0.2402***	-0.2611***	-0.2554***	-0.2514***
	(0.0192)	(0.0134)	(0.0107)	(0.0093)	(0.0083)
ZIP Code X Loan Purpose FE	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓
Bandwidth	10	15	20	25	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	7,089	12,698	17,236	23,231	30,625
R^2	0.7652	0.7555	0.7438	0.741	0.7392

This table presents the RD 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 different bandwidths of 10, 15, 20, 25, and 30 days on either side of the policy change date (January 1, 2019), in columns (1), (2), (3), (4), and (5) respectively. All regressions include Zip Code × loan reason and the day of the week fixed effects, a local linear specification estimated separately on each side of the policy change date, and use a triangular kernel. Standard errors reported in parentheses are clustered at the Zip Code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Effect of the Policy Change on Composition of Applications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	ln(Amount Asked)	ln(Income)	ln(DTI)	Credit Score	Low Caste	ln(Age)	Female
Post	-0.0129	-0.0089	-0.0040	0.6544	0.0312*	-0.0173***	0.0140
	(0.0356)	(0.0208)	(0.0269)	(2.0006)	(0.0168)	(0.0055)	(0.0103)
ZIP Code X Loan Purpose FE	✓	✓	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓	✓	✓
Bandwidth	30	30	30	30	30	30	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	30,625	30,625	30,625	30,025	30,625	30,566	30,625
R^2	0.3480	0.3606	0.3128	0.3110	0.3805	0.3443	0.3016

This table presents the RD estimates indicating the effect of the policy change on the composiiton of the applications. Column 1 uses the natural logarithm of the application loan amount as the dependent variable. Column 2 uses the natural logarithm of the monthly customer income as the dependent variable. Column 3 uses the natural logarithm of the ratio of loan amount and monthly income (DTI, debt to income ratio) as the dependent variable. Column 4 uses the applicants CIBIL credit score as the dependent variable. Column 5 uses a binary variable as a dependent variable that takes a value of one for applicants belonging to the scheduled caste, scheduled tribe and other backward class, and zero otherwise. Column 6 uses the natural logarithm of the applicants age as the dependent variable. Column 7 uses a binary variable as the dependent variable that takes a value of one for female applicants and zero otherwise. We use a bandwidth of 30 days on either side of the policy change date (January 1, 2019). All regressions include different set of fixed effects, a local linear specification estimated separately on each side of the policy change date, and use a triangular kernel. 95% confidence bands calculated based on standard errors clustered at the Zip Code level. 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 4: Heterogeneity & Effect of the Policy Change on Application Acceptances

Dep Var: Application Accepted	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post	-0.2514***	-0.2393***	-0.2161***	-0.1818***	-0.2147***	-0.2079***	-0.2472***	-0.0497***
	(0.0083)	(0.0128)	(0.0100)	(0.0116)	(0.0100)	(0.0106)	(0.0088)	(0.0188)
Sub-Prime Credit Score × Post		-0.0207						-0.0240
		(0.0160)						(0.0155)
Sub-Prime Credit Score		0.0025						0.0016
		(0.0090)						(0.0086)
Low Caste × Post			-0.0824***					-0.0638***
			(0.0150)					(0.0151)
Low Caste			0.0012					0.0029
			(0.0093)					(0.0092)
New Customer × Post t				-0.1220***				-0.1356***
				(0.0144)				(0.0147)
New Customer				0.0022				0.0038
				(0.0088)				(0.0092)
Low Income × Post					-0.0817***			-0.0901***
					(0.0150)			(0.0156)
Low Income					-0.0046			-0.0047
					(0.0086)			(0.0090)
Young × Post						-0.0829***		-0.0738***
						(0.0148)		(0.0152)
Young						-0.0055		-0.0046
						(0.0089)		(0.0090)
Female × Post							-0.0374	-0.0387
							(0.0250)	(0.0248)
Female							0.0070	0.0111
							(0.0144)	(0.0151)
ZIP Code X Loan Purpose FE	✓	✓	✓	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓	✓	✓	✓
Bandwidth	30	30	30	30	30	30	30	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	30,625	30,025	30,625	30,625	30,625	30,566	30,625	29,972
R^2	0.7392	0.7406	0.7428	0.7464	0.7425	0.7416	0.7393	0.7572

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

Table 5: Effect of the Policy Change on Default

Dep Var: Loan Default	(1)	(2)	(3)	(4)	(5)
Post	-0.0391	-0.0081	-0.0065	-0.0131	-0.0126
	(0.0364)	(0.0264)	(0.0210)	(0.0183)	(0.0162)
ZIP Code X Loan Purpose FE	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓
Bandwidth	10	15	20	25	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	4,293	8,341	11,068	14,893	20,568
R^2	0.3427	0.3097	0.2726	0.2543	0.2565

This table presents the RD estimates using the default status of a loan as the dependent variable. The dependent variable takes a value of one if the loan is in default and zero otherwise. We use different bandwidths of 10, 15, 20, 25, 30, 40, 50, and 60 days on either side of the policy change date (January 1, 2019), in columns (1), (2), (3), (4), (5), (6), (7) and (4), respectively. All regressions include Zip Code × loan reason and the day of the week fixed effects, a local linear specification estimated separately on each side of the policy change date, and use a triangular kernel. Standard errors reported in parentheses are clustered at the Zip Code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Credit Access: Reduced Form Results

	2 Years			3 Years			4 Years		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post	-0.0469***	-0.0398***	-0.0391***	-0.0528***	-0.0446***	-0.0451***	-0.0496***	-0.0404***	-0.0407***
	(0.0056)	(0.0056)	(0.0056)	(0.0058)	(0.0057)	(0.0058)	(0.0059)	(0.0058)	(0.0059)
ZIP Code FE		✓	✓		✓	✓		✓	✓
Day of the Week FE			✓			✓			✓
N	30,625	30,625	30,625	30,625	30,625	30,625	30,625	30,625	30,625
R^2	0.0021	0.1642	0.1649	0.0025	0.1617	0.1623	0.0022	0.1599	0.1606

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 customer accessed credit within two years (columns 1-3), three years (columns 4-6), or four years (columns 7-9) following January 1, 2019. The key independent variable is *Post*, which indicates whether the application was submitted after the policy. All regressions use a sample restricted to observations within 30 days of the policy cutoff . Column specifications progressively add fixed effects: no fixed effects (columns 1, 4, 7), Zip Code fixed effects (columns 2, 5, 8), and Zip Code plus day-of-week fixed effects (columns 3, 6, 9). Standard errors are clustered at the Zip Code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Credit Access: Instrumental Variables Results

	2 Years			3 Years			4 Years		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
App Accepted	0.1031***	0.0998***	0.0984***	0.1161***	0.1117***	0.1135***	0.1091***	0.1014***	0.1024***
	(0.0123)	(0.0139)	(0.0140)	(0.0126)	(0.0142)	(0.0144)	(0.0128)	(0.0146)	(0.0147)
ZIP Code FE		✓	✓		✓	✓		✓	✓
Day of the Week FE			✓			✓			✓
N	30,625	30,625	30,625	30,625	30,625	30,625	30,625	30,625	30,625
KP LM Statistic	570.35	399.27	398.88	570.35	399.27	398.88	570.35	399.27	398.88
KP F Statistic	10112.296	8278.0594	7858.4592	10112.296	8278.0594	7858.4592	10112.296	8278.0594	7858.4

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 customer accessed credit within two years (columns 1-3), three years (columns 4-6), or four years (columns 7-9) following January 1, 2019. The endogenous variable *App Accepted* is instrumented using the post-policy indicator. All regressions use a sample restricted to observations within 30 days of the policy cutoff. Column specifications progressively add fixed effects: no fixed effects (columns 1, 4, 7), ZIP Code fixed effects (columns 2, 5, 8), and ZIP Code plus day-of-week fixed effects (columns 3, 6, 9). Standard errors are clustered at the ZIP Code level. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 8: Summary Statistics of Profit Margins by Borrower Characteristics

Panel A: Profit Margin by Borrower Group

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

Panel B: Average Group Margins

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

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

For Online Publication:

"Privacy Regulation and the Economics of Screening versus Monitoring in Digital Credit Markets"

Appendix A

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

Dep Var: ln(# Applications)	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.0286***	0.0347***	0.0324***	0.0387***	0.0482***	0.0342***
	(0.0044)	(0.0056)	(0.0046)	(0.0057)	(0.0082)	(0.0077)
ZIP Code FE		✓		✓		
Loan Purpose FE			✓	✓		
ZIP Code X Loan Purpose FE					✓	✓
Day of the Week FE						✓
Bandwidth	30	30	30	30	30	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	39,534	36,928	39,534	36,928	30,494	30,494
R^2	0.0021	0.161	0.0091	0.1705	0.2614	0.2689

This table presents the RD estimates using the natural logarithm of the number of applications during the day as the dependent variable. 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. 95% confidence bands calculated based on standard errors clustered at the Zip Code level. Standard errors reported in parentheses are clustered at the Zip Code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2: Robustness to Covariates: Effect of the Policy Change on Application Acceptances

Dep Var: Application Accepted	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Post	-0.4565***	-0.4530***	-0.2450***	-0.2514***	-0.2511***	-0.2509***	-0.2489***	-0.2471***	-0.2472***	-0.2502***	-0.2504***
	(0.0076)	(0.0087)	(0.0068)	(0.0083)	(0.0083)	(0.0084)	(0.0083)	(0.0083)	(0.0083)	(0.0083)	(0.0082)
Female (=1)						-0.0178**	-0.0229***	-0.0148*	-0.0174**	-0.0177**	-0.0151*
						(0.0080)	(0.0080)	(0.0079)	(0.0081)	(0.0081)	(0.0080)
Low Caste (=1)							-0.0601***	-0.0578***	-0.0573***	-0.0551***	-0.0515***
							(0.0052)	(0.0052)	(0.0053)	(0.0052)	(0.0053)
ZIP Code FE		✓	✓								
Loan Purpose FE			✓								
ZIP Code X Loan Purpose FE				✓	✓	✓	✓	✓	✓	✓	✓
Day of the Week FE				✓	✓	✓	✓	✓	✓	✓	✓
Education FE					✓	✓	✓	✓	✓	✓	✓
Age Decile FE								✓	✓	✓	✓
Credit Score Decile FE									✓	✓	✓
Loan Maturity FE										✓	✓
Income Decile FE											✓
Bandwidth	30	30	30	30	30	30	30	30	30	30	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	33,584	32,733	32,733	30,625	30,625	30,625	30,625	30,566	29,972	29,972	29,972
R^2	0.2020	0.3263	0.6606	0.7392	0.7402	0.7403	0.7428	0.7447	0.7466	0.7485	0.7513

This table presents the RD 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. 95% confidence bands calculated based on standard errors clustered at the Zip Code level. Standard errors reported in parentheses are clustered at the Zip Code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: 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)
Post	0.2824	0.3781	0.0264	-0.1485	-0.2635
	(0.5480)	(0.4046)	(0.3445)	(0.3011)	(0.2637)
ZIP Code X Loan Purpose FE	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓
Bandwidth	10	15	20	25	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	4,212	8,182	10,831	14,573	20,188
R^2	0.3733	0.3489	0.3141	0.3039	0.3105

Panel B: Loan Maturity (Months)

Dep Var: Loan Maturity (Months)	(1)	(2)	(3)	(4)	(5)
Post	-0.0438	-0.0947	0.0031	0.0495	0.0961
	(0.1794)	(0.1350)	(0.1149)	(0.1002)	(0.0875)
ZIP Code X Loan Purpose FE	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓
Bandwidth	10	15	20	25	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	4,212	8,182	10,831	14,573	20,188
R^2	0.3695	0.3446	0.3081	0.2975	0.3056

This table presents RD estimates of the policy change's effect on loan terms for approved applications. Panel A shows effects on annual interest rates, while Panel B shows effects on loan maturity in months. We use different bandwidths of 10, 15, 20, 25, and 30 days on either side of the policy change date (January 1, 2019), in columns (1), (2), (3), (4), and (5) respectively. All regressions include Zip Code × loan reason and day of the week fixed effects, a local linear specification estimated separately on each side of the policy change date, and use a triangular kernel. Standard errors reported in parentheses are clustered at the Zip Code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

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

Dep Var: Application Accepted	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sub-Prime Credit Score × Post	-0.0099						-0.0120
	(0.0165)						(0.0164)
Sub-Prime Credit Score	-0.0046						-0.0049
	(0.0077)						(0.0075)
Low Caste × Post		-0.0888***					-0.0700***
		(0.0164)					(0.0166)
Low Caste		0.0092					0.0083
		(0.0074)					(0.0075)
New Customer × Post			-0.1167***				-0.1277***
			(0.0152)				(0.0156)
New Customer			0.0025				0.0020
			(0.0076)				(0.0078)
Low Income × Post				-0.0718***			-0.0804***
				(0.0164)			(0.0169)
Low Income				-0.0078			-0.0078
				(0.0072)			(0.0074)
Young × Post					-0.0813***		-0.0779***
					(0.0152)		(0.0157)
Young					-0.0026		0.0016
					(0.0075)		(0.0079)
Female × Post						-0.0179	-0.0121
						(0.0248)	(0.0246)
Female						-0.0121	-0.0120
						(0.0108)	(0.0117)
ZIP Code X Loan Purpose FE	✓	✓	✓	✓	✓	✓	✓
Day of the Week FE	✓	✓	✓	✓	✓	✓	✓
ZIP Code X Week FE	✓	✓	✓	✓	✓	✓	✓
Loan Purpose X Week FE	✓	✓	✓	✓	✓	✓	✓
Bandwidth	30	30	30	30	30	30	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	23,562	24,034	24,034	24,034	23,985	24,034	23,522
R^2	0.7818	0.7838	0.7856	0.7840	0.7830	0.7810	0.7946

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

Table A.5: Robustness to Covariates: Effect of the Policy Change on Default

Dep Var: Loan Default	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Post	-0.0132	-0.0205	-0.0197	-0.0126	-0.0124	-0.0122	-0.0121	-0.0112	-0.0129	-0.0143	-0.0068
	(0.0116)	(0.0132)	(0.0132)	(0.0162)	(0.0162)	(0.0162)	(0.0162)	(0.0162)	(0.0158)	(0.0149)	(0.0149)
Female (=1)						-0.0129	-0.0120	-0.0126	-0.0086	-0.0159	-0.0209
						(0.0131)	(0.0131)	(0.0133)	(0.0135)	(0.0131)	(0.0130)
Low Caste (=1)							0.0094	0.0094	0.0113	0.0132	0.0083
							(0.0085)	(0.0086)	(0.0087)	(0.0084)	(0.0083)
ZIP Code FE		✓	✓								
Loan Purpose FE			✓								
ZIP Code X Loan Purpose FE				✓	✓	✓	✓	✓	✓	✓	✓
Day of the Week FE				✓	✓	✓	✓	✓	✓	✓	✓
Education FE					✓	✓	✓	✓	✓	✓	✓
Age Decile FE								✓	✓	✓	✓
Credit Score Decile FE									✓	✓	✓
Loan Maturity FE										✓	✓
Income Decile FE											✓
Bandwidth	30	30	30	30	30	30	30	30	30	30	30
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
# Obs	23,184	22,226	22,226	20,568	20,568	20,568	20,568	20,531	20,004	20,004	20,004
R^2	0.0007	0.1256	0.1291	0.2565	0.2566	0.2567	0.2568	0.2576	0.2714	0.3266	0.3400

This table presents the RD 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. 95% confidence bands calculated based on standard errors clustered at the Zip Code level. Standard errors reported in parentheses are clustered at the Zip Code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Credit Access: Reduced Form Results Robust to 15-day Bandwidth

	2 Years			3 Years			4 Years		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post	-0.0295***	-0.0285***	-0.0362***	-0.0291***	-0.0270**	-0.0375***	-0.0232**	-0.0183*	-0.0304***
	(0.0098)	(0.0103)	(0.0108)	(0.0101)	(0.0106)	(0.0114)	(0.0101)	(0.0106)	(0.0114)
ZIP Code FE		✓	✓		✓	✓		✓	✓
Day of the Week FE			✓			✓			✓
N	12,698	12,698	12,698	12,698	12,698	12,698	12,698	12,698	12,698
R^2	0.0008	0.1558	0.1572	0.0007	0.1551	0.1565	0.0005	0.1557	0.1573

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 customer accessed credit within two years (columns 1-3), three years (columns 4-6), or four years (columns 7-9) following January 1, 2019. The key independent variable is *Post*, which indicates whether the application was submitted after the policy. All regressions use a sample restricted to observations within 15 days of the policy cutoff. Column specifications progressively add fixed effects: no fixed effects (columns 1, 4, 7), Zip Code fixed effects (columns 2, 5, 8), and Zip Code plus day-of-week fixed effects (columns 3, 6, 9). Standard errors are clustered at the Zip Code level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Credit Access: Instrumental Variables Results Robust to 15-day Bandwidth

	2 Years			3 Years			4 Years		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
App Accepted	0.0650***	0.0717***	0.0809***	0.0641***	0.0679**	0.0836***	0.0509**	0.0460*	0.0678***
	(0.0216)	(0.0260)	(0.0241)	(0.0222)	(0.0266)	(0.0254)	(0.0222)	(0.0266)	(0.0253)
ZIP Code FE		✓	✓		✓	✓		✓	✓
Day of the Week FE			✓			✓			✓
N	12,698	12,698	12,698	12,698	12,698	12,698	12,698	12,698	12,698
KP LM Statistic	404.1549	263.9400	266.4229	404.1549	263.9400	266.4229	404.1549	263.9400	266.4229
KP F Statistic	4729.8063	3283.6020	3119.5205	4729.8063	3283.6020	3119.5205	4729.8063	3283.6020	3119.5205

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 customer accessed credit within two years (columns 1-3), three years (columns 4-6), or four years (columns 7-9) following January 1, 2019. The endogenous variable *App Accepted* is instrumented using the post-policy indicator. All regressions use a sample restricted to observations within 15 days of the policy cutoff. Column specifications progressively add fixed effects: no fixed effects (columns 1, 4, 7), ZIP Code fixed effects (columns 2, 5, 8), and ZIP Code plus day-of-week fixed effects (columns 3, 6, 9). Standard errors are clustered at the ZIP Code level. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.8: Summary Statistics of Profit Margins by Borrower Characteristics (EMI Model)

Panel A: Profit Margin by Borrower Group

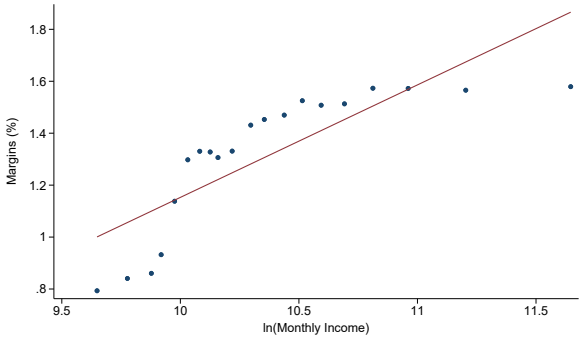
Characteristic	N	25th pct	Median	75th pct	Mean	Std Dev
Low Income	14,217	0.5899	1.0204	1.4142	1.1409	0.8420
Female	2,152	0.6515	1.2204	2.4370	1.4366	0.9634
Young Age	12,766	0.6271	1.0908	2.1830	1.2660	0.9158
Low Caste	13,564	0.5771	1.0632	2.2732	1.2758	0.9468
Low Amount	15,247	0.3446	0.7181	1.0154	0.6553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

Panel B: Average Group Margins

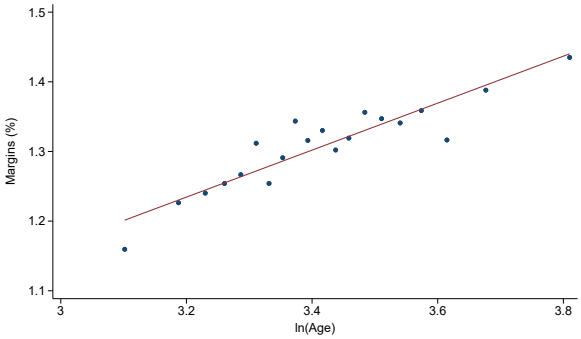
Characteristic	Average Margin (%)
Low Income	1.7747
Female	2.1726
Young Age	2.0114
Low Caste	2.0626
Low Amount	0.6911
Low Credit Score	2.1359

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

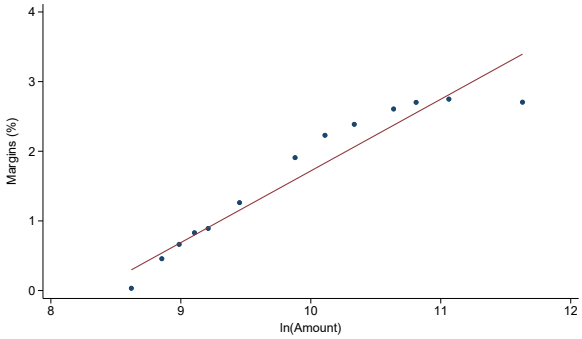
Figure A.1: Profit Margins of Marginal Borrowers by Characteristics (EMI Model)



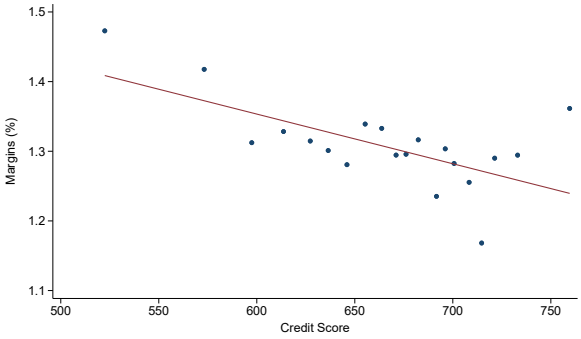
(a) By Income



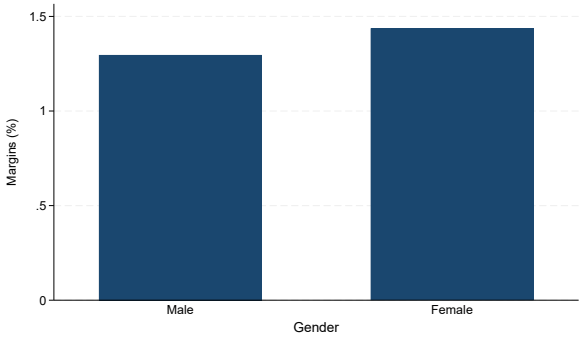
(b) By Age



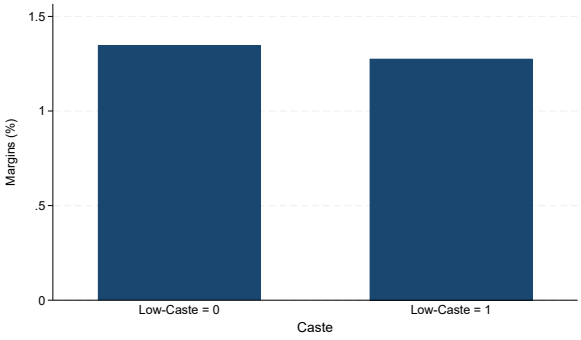
(c) By Loan Amount



(d) By Credit Score



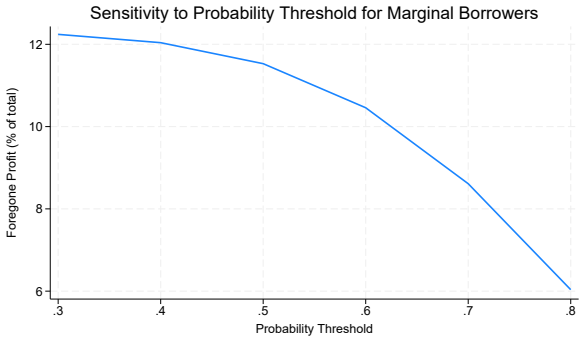
(e) By Gender



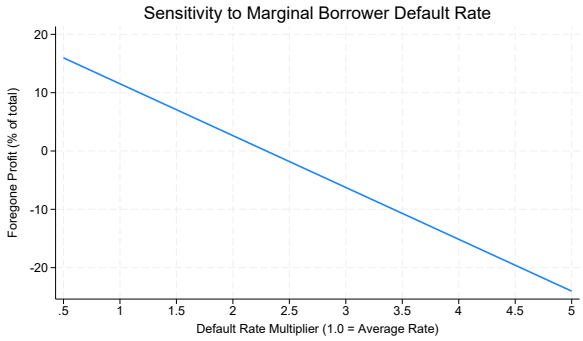
(f) By Caste

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

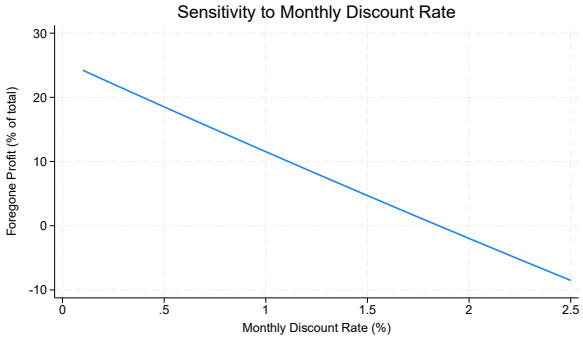
Figure A.2: Sensitivity Analysis of Foregone Profits (EMI Model)



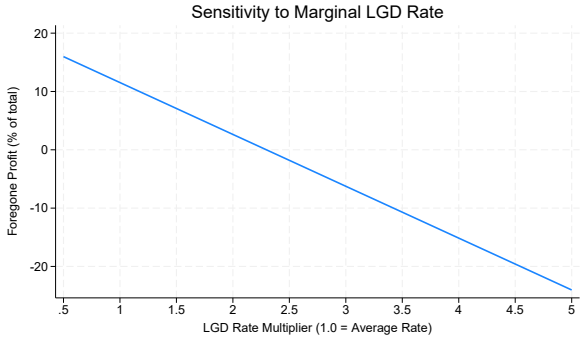
(a) Probability Threshold



(b) Default Rate



(c) Monthly Discount Rate



(d) Loss Given Default

Notes: This figure presents sensitivity analysis for the foregone profit calculations under EMI loan structure. Panel A.2a shows how results vary with the probability threshold used to define marginal borrowers (those rejected post-policy but who would have been approved pre-policy). Panel A.2b examines sensitivity to assumptions about marginal borrowers' default rates, expressed as multiples of the baseline 19% rate. Panel A.2c shows how results depend on the monthly discount rate used in present value calculations. Panel A.2d examines sensitivity to the loss given default (LGD) rate, expressed as multiples of the baseline 9% rate. All calculations assume EMI structure with monthly payments. Baseline parameters: 0.5 probability threshold, 19% default rate, 1% monthly discount rate, 9% LGD. Sample includes borrowers within 60-day window around January 1, 2019 policy implementation.