Referral Contracts:

Harnessing Social Ties in Lending Markets*

Ehsan Azarmsa† Pulak Ghosh‡ Nishant Vats§ Alan L. Zhang¶

Abstract

Referral programs are commonly treated as customer-acquisition tools. We show that, in lending markets, they can also embed loans in social relationships that support repayment after origination. Using data from a large digital lender in India, we find that referred applicants are more likely to receive credit and, conditional on approval, less likely to default. Comparing the same borrower across lenders in credit-bureau data, we find no relative approval advantage at the referring lender, but a 3.8-percentage-point reduction in default. Call-detail records show that referrer-borrower communication rises around repayment dates. Together, these findings suggest that referrals activate relationship-specific disciplining that improves repayment. We develop a model of referral contracts in which the lender screens for referrers with greater disciplining capacity and incentivizes them to exert effort. Consistent with data, the optimal contract uses default-contingent reductions in future credit access to screen and motivate referrers, with compensation paid upfront. The model yields sufficient statistics for program scale and the mix of referral incentives.

*We thank . We take full responsibility for any errors. We have no conflicts of interest to disclose.

†College of Business Administration, University of Illinois Chicago. Email: azarmsa@uic.edu

‡Indian Institute of Management Bangalore. Email: pulak.ghosh@iimb.ac.in

§Olin Business School, Washington University in St Louis. Email: vats@wustl.edu

¶Ivy College of Business, Iowa State University, Email: alanz@iastate.edu

1 Introduction

Social networks have long shaped the reach and operation of lending markets. When lenders had limited information about prospective borrowers and little ability to verify their circumstances directly, they relied on local relationships, personal reputation, and introductions from existing customers to form new lending relationships.[1] Digital technology has changed this role. Government-issued identifiers, credit bureaus, digital Know-Your-Customer procedures, and automated data systems now allow lenders to verify identity, recover financial histories, and process applications at scale.[2] These technologies can substitute for social connections in screening borrowers before origination.

Yet the informational role of social networks is only one part of lending relationships. Even a lender with rich data may have limited ability to observe a borrower's circumstances or influence repayment after a loan is disbursed. A person who knows the borrower may be better positioned to remind her about repayment, monitor her financial behavior, or impose reputational and interpersonal costs when she defaults. Social relationships may therefore affect credit outcomes not only by revealing information about borrower quality, but also by changing repayment behavior after credit has been extended.

Referral programs provide a useful setting for studying this distinction. Many digital lenders compensate existing customers for introducing new borrowers to their platforms. These programs are typically viewed as customer-acquisition tools. But a referral can do more than expand the lender's applicant pool: it can embed a digitally originated loan in a pre-existing social relationship. The lender may thereby acquire a new customer while also drawing on a relationship that can affect repayment after disbursement.

Whether referrals create value through information or through discipline has different implications for the economics and design of digital credit. Under a selection mechanism, referrers use private information to recommend borrowers with stronger unobserved credit quality. Referrals then provide an additional underwriting signal, and their value should decline as lenders obtain more comprehensive administrative, transactional, and alternative data. Under a discipline mechanism, referrers affect repayment after origination by monitoring borrowers, encouraging repayment, or creating social costs of default. In that case, referrals can remain valuable even in data-rich lending environments because they complement, rather than substitute for, algorithmic underwriting.

The discipline mechanism also creates a contracting problem. A lender does not observe the

[1] See, among others, Petersen and Rajan (1994); Berger and Udell (1995, 2002); Cole (1998); Elsas and Krahnen (1998); Stein (2002); Berger et al. (2005, 2024); Jimenez, Salas, and Saurina (2006); Bharath et al. (2007, 2011); Agarwal and Hauswald (2010); Ioannidou and Ongena (2010); Schenone (2010); Berger, Frame, and Ioannidou (2011); Berger, Bouwman, and Kim (2017); Hombert and Matray (2017); Agarwal and Ben-David (2018); Liberti and Petersen (2019); Granja, Leuz, and Rajan (2022); Bird et al. (2026). We direct readers to Elyasiani and Goldberg (2004); Boot (2000); Kysucky and Norden (2016); Duqi, Tomaselli, and Torluccio (2018) and Boot and Thakor (2019) for a detailed survey of this literature.

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

strength of the relationship between a prospective referrer and borrower, nor does it observe the referrer's capacity to influence repayment. A referral program may thus attract both referrers who can meaningfully discipline borrowers and referrers whose relationships provide little such capacity. The lender should therefore do more than reward referrals: it should design incentives that screen for effective referrers and incentivize them to remain engaged after the loan is made.

These observations raise two key questions that we attempt to answer in this paper. First, can referrals improve loan performance by affecting borrowers' repayment behavior after origination? Second, when referrers differ in their ability to influence repayment, how should lenders design referral programs to screen for effective referrers and provide incentives for discipline?

We study these questions using data from a large digital lender in India that offers unsecured loans nationwide and pays existing customers when a referred applicant receives a loan. We combine the lender's application-level records with comprehensive credit-bureau data and cellphone call-detail records. The lender data identify referral status, borrower and referrer characteristics, applications, approval decisions, loan terms, and default outcomes; the credit-bureau data track the same borrowers across other formal lenders; and the call records measure communication within referrer-borrower pairs around scheduled repayment dates. Together, these data allow us to distinguish whether referrals reveal borrower quality or alter repayment behavior after origination.

We first document that referred applicants have better outcomes than observably similar non-referred applicants. Referred applicants are 3.1 percentage points more likely to receive approval. Conditional on receiving a loan, they are 6.1 percentage points less likely to default. These effects correspond to an approximately 12 percent increase in approval relative to the sample mean and an approximately 23 percent decline in default. The estimates condition on detailed borrower characteristics, including credit score, income, age, gender, education, loan purpose, requested amount, and debt-to-income ratio, as well as month-by-zipcode fixed effects.

These comparisons, however, do not by themselves distinguish selection from discipline. Referred borrowers may have stronger unobserved credit quality, or the referral relationship may alter repayment after origination. We attempt to distinguish between these explanations using credit-bureau data that track borrowers across lenders. The key intuition is that only the focal fintech observes the referral. If the referral provides incremental information about borrower quality, the focal fintech can use that signal in underwriting, generating an approval advantage relative to other lenders that do not observe the signal. If the referral provides little additional information, other lenders would make similar approval decisions using the information already available to them. Alternatively, if referrals operate primarily through repayment discipline, their effect should arise after origination. The referral embeds the focal-fintech loan in the referrer-borrower relationship, allowing the referrer to monitor the borrower, encourage repayment, or impose social costs of default. The relevant empirical implication

is therefore a lower default rate at the focal fintech, where the referral relationship is active, rather than at other lenders.

We therefore compare each borrower's outcomes at the focal fintech, where the referral relationship is active, with her outcomes at other lenders, where it is not. Our preferred specification includes borrower-by-year fixed effects, which compare the same borrower across lenders within a year; month-by-zipcode fixed effects, which absorb local credit-market conditions; and lender-by-year-by-zipcode fixed effects, which account for lender-specific local conditions and geographic specialization. This within-borrower comparison tests whether referrals generate a focal-lender-specific approval advantage, consistent with an incremental informational role, or a focal-lender-specific reduction in default, consistent with repayment discipline.

The results are consistent with a repayment-discipline channel. The within-borrower estimates show no focal-fintech approval advantage for referred borrowers, providing little support for the view that referrals supply incremental information used in underwriting. The same borrower is, however, 3.8 percentage points less likely to default on a referred loan at the focal fintech than on loans from other lenders. This pattern indicates that the lower default rate associated with referrals is concentrated in the lending relationship in which the referral is active.

We next use cellphone call-detail records to examine communication between referrers and borrowers around repayment. These data provide unusually direct behavioral evidence consistent with the proposed mechanism. Whereas most studies of social collateral infer monitoring or peer pressure from loan outcomes, our data allow us to observe communication between the specific referrer and referred borrower. We examine daily interactions around scheduled repayment dates, the period in which a disciplining mechanism should be most active. The descriptive pattern is sharp. Communication between referrers and borrowers rises as repayment approaches, remains elevated on the repayment date, and declines afterward. The number of calls follows the same pattern. This repayment-centered concentration of communication is consistent with referrers becoming more engaged when borrowers must make a payment, rather than communication simply reflecting a stable social relationship between the two parties.

The descriptive evidence could, however, reflect persistent differences across borrowers, local conditions at the time of loan origination, or shocks that affect communication on particular calendar dates. We therefore estimate a daily event study that includes referred-borrower fixed effects, zipcode-by-origination-month fixed effects, and calendar-date fixed effects. Over the month surrounding repayment, communication between referrers and borrowers is broadly stable until approximately ten days before the due date. It then rises as repayment approaches, reaches its peak immediately before and around the repayment date, and declines sharply thereafter. Relative to three days after repayment, the omitted reference date, the probability of a call between a referrer and referred

borrower is approximately 28 percent higher during this peak period. Within several days after repayment, communication becomes statistically indistinguishable from its post-repayment reference level. Although call-detail records do not reveal the content of these interactions, this sharp, temporally localized pattern provides unusually direct behavioral evidence consistent with referrers becoming actively engaged when repayment is due.

The repayment advantage associated with referrals is also larger for relationships that plausibly give referrers greater social leverage. Default falls by more when the referrer is older than the borrower, earns more, works at the same firm, or belongs to the same social group. These characteristics do not directly measure relationship strength, but they are consistent with the idea that referrers are more effective when they occupy a position that allows them to influence the borrower.

The evidence thus suggests that referrals can improve repayment through active social discipline. This discipline is neither automatic nor uniform. A referrer must have a relationship that gives her the capacity to influence the borrower, and she must have an incentive to use that capacity. Both margins depend on the lender's referral program. Monetary rewards can affect who participates in the program and which borrowers referrers choose to bring to the platform. Non-monetary rewards can affect whether referrers remain engaged after disbursement by linking their own future relationship with the lender to the repayment behavior of the borrowers they refer. The value of a referral program therefore depends not only on the number of referrals it generates, but also on whether its incentives attract relationships capable of supporting repayment and induce referrers to remain engaged after origination.

We first study monetary incentives. In October 2021, the lender increased its payment per successful referral from ₹1,250 to ₹3,500. The policy substantially expanded referral activity: the probability that a new applicant was referred increased by 1.75 percentage points, from approximately 2 percent to approximately 4 percent. However, the credit advantage associated with referred applicants weakened. Their relative approval advantage declined by 2.1 percentage points, and their relative default advantage declined by 4.2 percentage points. In the within-borrower analysis, the default reduction associated with referred loans also became substantially smaller after the bonus increase. The evidence is consistent with larger cash rewards expanding referrals toward relationships in which referrers have less capacity to influence repayment.

A similar pattern appears among frequent referrers. Most referred borrowers are introduced by customers who make only one referral, but a meaningful share are introduced by customers who refer many borrowers. The repayment benefit associated with a referral is substantially smaller when the referrer has made at least five referrals, and it is statistically indistinguishable from zero for frequent referrers. The attenuation becomes larger under higher thresholds for frequent referral activity. These results are consistent with referrers being able to discipline a limited set of close contacts more

effectively than a broad set of weaker relationships.

The preceding evidence shows that monetary rewards affect the composition of referral relationships. We next examine the non-monetary incentives that may induce referrers to use their capacity to influence repayment. In particular, a referrer may care about the repayment performance of borrowers she refers if she expects to seek credit from the platform again and if the platform conditions her future access to credit on those borrowers' outcomes. Consistent with the importance of this dynamic relationship, customers who refer new borrowers are substantially more likely to return to the platform for additional credit. Specifically, we document that referrers are 18.6-22.2 percent more likely to submit a subsequent application and 14.2-22.9 percent more likely to receive a subsequent approval.

We then examine whether the platform links referrers' future credit access to the performance of the borrowers they refer. Among customers who make a referral, those whose referred borrowers repay are 5.8-8.6 percent more likely to have subsequent applications approved than referrers whose referred applicants were not approved. In contrast, referrers whose referred borrowers default are 13.2-15.6 percent less likely to obtain subsequent approval than that same benchmark group. These results are consistent with the platform rewarding referrers whose borrowers repay and penalizing those whose borrowers default.

We develop a model of referral contracts to formalize these forces and characterize optimal program design when referrers differ in their capacity to influence repayment. A prospective referrer privately knows whether she can discipline a borrower and, if so, the magnitude of the resulting improvement in repayment; the lender observes neither. This asymmetry creates a screening problem. A contract that offers only an upfront referral payment attracts both effective referrers and referrers who cannot improve repayment. Screening therefore requires that referrers bear a cost when the borrowers they refer default.

The model shows that this cost must take the form of a penalty imposed through the referrer's future borrowing relationship with the lender. When direct monetary penalties are infeasible, the lender can reduce the referrer's future access to credit following borrower default. This penalty deters referrers who cannot meaningfully influence repayment from participating in the program. A penalty alone, however, is not sufficient. Even an effective referrer can reduce default risk but cannot eliminate it, so a contract that only exposes referrers to the downside of default would make referral unattractive. The lender therefore needs to offset the expected penalty with compensation, either through an upfront payment or through a bonus when the borrower repays, which may take the form of a monetary payment or greater future access to credit.

The model characterizes the optimal mix of these instruments and the efficient scale of the program. Upfront compensation helps induce participation but does not itself strengthen incentives

to discipline the borrower. By contrast, greater future access to credit following repayment both compensates the referrer for bearing default risk and increases the return to successful discipline. When future credit is sufficiently valuable to referrers relative to its cost to the lender, this repayment-contingent form of compensation is less costly than an upfront cash reward. The optimal contract combines a penalty following borrower default with either upfront compensation or greater future access to credit following repayment, and it characterizes when each form of compensation is preferred.

The model characterizes the optimal mix of these instruments and the efficient scale of the program. The optimal contract combines a penalty following borrower default with either upfront compensation or a repayment-contingent reward. Upfront compensation induces participation but does not itself strengthen incentives to discipline the borrower. By contrast, a repayment-contingent reward both compensates the referrer for bearing default risk and increases the return to successful discipline. When future credit is sufficiently valuable to referrers relative to its cost to the lender, rewarding repayment with greater future credit access is therefore less costly than providing upfront cash compensation.

The model further delivers a sufficient statistic for program scale. It evaluates whether a referral program is too large or too small by comparing the value generated by marginal referrers attracted by an expansion with the increase in incentive costs required to sustain participation by referrers already in the program. Crucially, the statistic can be recovered from observable responses of referral activity and repayment outcomes to changes in program generosity, without requiring the lender to observe referrers' latent disciplining capacity or to estimate its distribution. In particular, a change in referral volume together with a change in the average default reduction associated with referrals identifies the disciplining capacity of referrers induced to enter at the margin.

We apply this approach to the lender's 2021 increase in the upfront referral bonus. The bonus increased the referred share of approved borrowers from approximately 2.3 percent to 4.1 percent, while reducing the average default reduction associated with referrals from 7.94 percentage points to 3.71 percentage points. These estimates imply that the marginal referrers attracted by the expansion increased repayment by only 1.26 percentage points. In the model, a referral must increase repayment by at least 4.22 percentage points for the lender to break even before accounting for incentive costs. The marginal referrals therefore fail even this break-even test, implying a negative gross return. The model thus implies that the increase in the cash bonus expanded the referral program beyond its efficient scale. Our analysis further reveals that the lender can reduce the cost of attracting effective referrers by shifting compensation away from upfront cash and toward repayment-contingent future access to credit.

We contribute to the literature on bank monitoring, relationship lending, and technology in credit markets. A central rationale for financial intermediaries is that they economize on the costs

of producing information and monitoring borrowers on behalf of dispersed investors (Diamond, 1984, 1991; Ramakrishnan and Thakor, 1984; James, 1987; Rajan, 1992; Holmström and Tirole, 1997; Boot and Thakor, 2000). Consistent with this view, a large empirical literature shows that lending relationships affect credit availability, loan terms, and the production of soft information.[3] Although this empirical literature has primarily emphasized how lender-borrower relationships generate information that improves credit decisions at origination, these relationships can also give a lender opportunities to monitor repayment behavior over the life of a loan (Heitz, Martin, and Ufier, 2026). Digital lending changes the environment in which these relationship-based functions operate. The fintech literature shows that algorithms, digital footprints, credit-bureau records, and alternative data can improve screening and expand credit access, particularly where traditional lenders have limited capacity to produce information.[4] By improving the information available at origination, these technologies may substitute for some traditional functions of relationship lending. Yet their ability to screen borrowers need not imply an equal ability to monitor them or influence their behavior after a loan is made. A person embedded in the borrower's social network may have access to information, communication channels, and interpersonal influence that a remote digital lender cannot readily replicate. The central question is therefore not whether technology replaces relationships, but whether the two perform different and potentially complementary functions over the life of a loan.

Our paper makes two contributions to this literature. First, we show that social relationships can reduce default even in a data-rich fintech environment and provide evidence consistent with referrers influencing repayment after loan origination. These findings suggest that social ties can complement automated underwriting systems by providing monitoring and repayment discipline that administrative data and algorithms may not replicate. Second, we identify a new form of relationship-based monitoring. Rather than monitoring borrowers directly, the lender uses referral contracts to draw on the monitoring and enforcement capacity embedded in borrowers' social networks. Referral programs can therefore be viewed as a contractual technology through which a digital lender harnesses relationship-based discipline and screens for referrers with the ability to influence borrower repayment.[5]

This contractual view builds on the literature on incentive provision under limited liability. A key constraint facing lenders is that referrers cannot be fined. This creates two challenges: deterring

[3] See, among others, Petersen and Rajan (1994); Berger and Udell (1995, 2002); Cole (1998); Elsas and Krahnen (1998); Stein (2002); Berger et al. (2005, 2024); Jimenez, Salas, and Saurina (2006); Bharath et al. (2007, 2011); Agarwal and Hauswald (2010); Ioannidou and Ongena (2010); Schenone (2010); Berger, Frame, and Ioannidou (2011); Berger, Bouwman, and Kim (2017); Hombert and Matray (2017); Agarwal and Ben-David (2018); Liberti and Petersen (2019); Granja, Leuz, and Rajan (2022); Bird et al. (2026). See Elyasiani and Goldberg (2004); Boot (2000); Kysucky and Norden (2016); Duqi, Tomaselli, and Torluccio (2018); Boot and Thakor (2019) for surveys.

[4] See, among others, Buchak et al. (2018); Vallée and Zeng (2019); Berg et al. (2020); Fuster et al. (2019, 2022); Di Maggio and Yao (2021); Gambacorta et al. (2023); Chioda et al. (2024); Erel and Liebersohn (2022); Alok et al. (2024); Agarwal et al. (2024); Cramer et al. (2024); Ghosh, Vallee, and Zeng (2026). See Frost et al. (2019); Thakor (2020); Allen, Gu, and Jagtiani (2021); Berg, Fuster, and Puri (2022); Doerr et al. (2023) for reviews.

[5] A related contractual approach appears in Heath (2018) in the context of employee referral programs.

referrers with low disciplining ability and incentivizing effective referrers to discipline borrowers after loan origination. When agents cannot be fined, optimal contracts are constrained in their ability to impose downside risk and often rely on payments concentrated in favorable states (Sappington, 1983; Innes, 1990).In lending relationships, lenders can use the threat of exclusion or reduced future access to credit to discipline borrowers (Stiglitz and Weiss, 1981; Bolton and Scharfstein, 1990, 1996; Diamond, 1989; Albuquerque and Hopenhayn, 2004; Clementi and Hopenhayn, 2006; Padilla and Pagano, 1997; Vercammen, 1995; Tedeschi, 2006). We show that conditioning referrers' future access to credit on borrower repayment can arise as an optimal feature of referral contracts. Moreover, because such contracts expose referrers to the risk of losing future credit access, they can optimally include upfront compensation for bearing this risk.

Furthermore, our contractual framework enables us to develop a sufficient-statistics approach for evaluating the scale of a referral program. This approach follows a broader literature that uses observable behavioral responses to evaluate policy without recovering all underlying model primitives (Chetty, 2009; Saez, 2001; Dávila and Korinek, 2018; Hendren and Sprung-Keyser, 2020; Kleven, 2021). Our framework shows that a lender can assess whether its referral program is too large or too small using two observable responses to a change in program generosity: the change in referral volume and the change in the average repayment benefit associated with referrals. The lender need not observe referrers' latent disciplining capacity or estimate its distribution.

This paper contributes to the literature on social collateral and peer intermediation in credit markets. Since the seminal work of Karlan et al. (2009), this literature has examined how social relationships can relax credit constraints by substituting for limited borrower collateral, imperfect lender information, and weak formal enforcement (see, e.g., Jackson, Rodriguez-Barraquer, and Tan, 2012; Ambrus, Möbius, and Szeidl, 2014; Karlan et al., 2017; Chandrasekhar, Kinnan, and Larreguy, 2018; Breza and Chandrasekhar, 2019). A central focus is joint-liability lending, in which peers can screen prospective borrowers, monitor their behavior, and enforce repayment (see, among others, Stiglitz, 1990; Besley and Coate, 1995; Banerjee, Besley, and Guinnane, 1994; Ghatak and Guinnane, 1999; Rai and Sjöström, 2004; Giné et al., 2010; Giné and Karlan, 2014; Attanasio et al., 2015; Feigenberg, Field, and Pande, 2013; de Quidt, Fetzer, and Ghatak, 2018; Maitra et al., 2017; Hussam, Rigol, and Roth, 2022). These same functions can also arise in individual-liability arrangements, such as cosigning and guarantees, in which another individual supports the borrower's application and may share formal responsibility for repayment (Bond and Rai, 2008; Tomy and Wittenberg-Moerman, 2024). More broadly, formal credit can reshape or displace the informal financial relationships that support these arrangements (Banerjee et al., 2024). Closely related to our context, referrals also appear in the online-lending literature as a source of soft information (Iyer et al., 2016; Lin, Prabhala, and Viswanathan, 2013; Hildebrand, Puri, and Rocholl, 2017; Freedman and Jin, 2017).

We contribute to this literature in three ways. First, our evidence suggests that referrals create value not only by conveying information about borrower quality, but also by strengthening repayment discipline after origination. Second, unlike in group borrowing, cosigning, or guarantee arrangements, referrers in our setting bear no contractual liability for the referred borrower's debt. Instead, the lender creates a repayment-contingent social stake by linking referrers' future access to credit to their referrals' repayment performance. This feature allows the lender to use its ongoing relationship with referrers both to screen for effective referrers capable of supporting repayment and to induce discipline after origination. Third, whereas much of the existing evidence infers peer enforcement from repayment outcomes, we combine loan records with call-detail data to observe communication within referrer-borrower pairs. These data allow us to document a sharp increase in communication as scheduled repayment dates approach, followed by a rapid decline afterwards. This repayment-centered pattern provides behavioral evidence consistent with referrers becoming actively engaged when repayment is due, and allows us to study peer intermediation not only through its effects on loan performance but also through the behavior that may generate those effects.

Our paper is closely related to Bryan, Karlan, and Zinman (2015), who use a two-stage field experiment with a South African consumer lender to distinguish peer screening from peer enforcement under individual liability. Their experiment varies whether a referrer's one-time cash bonus is contingent on the referred borrower's approval or repayment. They find large enforcement effects and no statistically precise evidence that referrers select borrowers with stronger repayment prospects. The key difference is that their referral contract is static, whereas referrals in our setting are embedded in a dynamic lender-referrer relationship. Because referrers are existing customers who may return to the platform for future credit, the lender can condition their future credit access on the repayment performance of their referrals. This continuation value gives the lender an additional incentive instrument: it can condition the referrer's future access to credit on the borrower's repayment outcome, giving the referrer a stake in repayment without making the referrer liable for the borrower's loan. We leverage this dynamic contracting feature to study the optimal design of referral programs.

The remainder of the paper proceeds as follows. Section 2 describes the institutional setting and data. Section 3 presents the empirical analysis of referrals, credit outcomes, and the disciplining mechanism. Section 4 examines the role of monetary and non-monetary incentives. Section 5 develops the model of referral contracts and evaluates the efficient scale of the program. Section 7 concludes.

2 Data

We obtain detailed application-level data from a leading fintech lender in India, which includes information on gender, age, educational attainment, income, credit score, loan amount, loan purpose, loan term, application acceptance, and subsequent default status. Furthermore, we obtain complete

credit records for these applicants from the Indian credit bureau Equifax, providing a complete record of their formal credit journey.

2.1 Institutional Background

The fintech lender from whom we source data operates a digital platform that specializes in unsecured loans nationwide. Utilizing advanced technology, data analytics, and a user-friendly interface, the platform offers a streamlined application process, often disbursing funds within minutes of approval. Its primary focus is on underserved segments, including small business owners and individuals with limited access to traditional banking services. Since its establishment in 2017, the firm has experienced rapid expansion, as depicted in Figure 1, which shows the increasing volume of applications over time. The growth has been predominantly exponential, though it encountered a notable disruption during the COVID-19 pandemic. Since then, the platform has demonstrated resilience, rebounding from the downturn and resuming its rapid growth trajectory.

Most importantly, the fintech lender implements a referral program that allows individuals with a prior borrowing relationship with the lender to recommend new borrowers for loan applications. When a referred individual successfully applies, is accepted and receives a loan, the existing borrower – *referrer* hereafter – receives a cash bonus as a reward. This bonus serves as a monetary incentive for current customers to actively promote the platform within their social network. Such referral systems are quite common among platforms that rely on network effects to drive growth, where existing users might share the platform with friends or family members, knowing they will earn rewards if those referrals lead to new connections for the platform. The solid red line in Figure 1 shows the fraction of monthly applications originating from referrals. In 2017, a significant portion of applications resulted from this process, a fraction that subsequently declined. This pattern aligns with the view that referral programs are critical for fueling initial growth and establishing a robust network effect, while their importance diminishes as the platform matures and establishes a firmer market presence.

Although the firm had a referral program from its start, a significant policy change was implemented in October 2021. The fintech lender substantially increased the bonus paid to referrers, from ₹1,250 ($15) to ₹3,500 ($42) per successful referral. The primary objective was to increase the volume of applications received by the platform. This shift is evident in Figure 1, which shows an increase in both the total number of applications and the proportion originating from referrals after October 2021.

2.2 Summary Statistics

The paper aims to examine the relationship between referrals and credit market outcomes. Therefore, we restrict the analysis to a sample of first-time applicants to the platform. This allows for a comparison between potential borrowers who were referred and those who were not, within a sample of loan

applicants lacking any prior relationship with the platform. The resulting dataset comprises of 1,216,524 loan applications from first-time customers to the fintech lender between 2017 and 2022.

Table 1 presents summary statistics for the applicant pool. The demographic data in Panel A indicates that approximately 12.4% of applicants are female. The average applicant is approximately 30 years old, with an interquartile range from 24.9 to 33.7 years old. Moreover, a substantial proportion of applicants have completed an undergraduate degree (57.6%), while a smaller fraction hold a Master's degree (13.6%). Overall, the fintech's clientele primarily comprises of young college-educated men

The economic characteristics in Panel B highlight that the average monthly income is approximately ₹31,035 ($375), with a high standard deviation of ₹22,052. Notably, 65.7% of first-time applicants lack a credit score at the time of the application. This suggests a significant role for the fintech lender in attracting traditionally underserved populations who lack a credit history and have limited access to conventional banking services. The importance of fintechs in reaching new-to-credit borrowers in the Indian context has been previously emphasized by Cramer, Ghosh, Kulkarni, and Vats (2024). Among those with credit scores, the average is 676 and the median is 694, corresponding to a *Good* rating.[6] The interquartile range, from 642 to 724, places most scored borrowers in the *Fair* to *Good* credit tiers.

Panel C presents the summary statistics of the loan applications received by the platform. The average loan amount requested is ₹35,703.93 (≈ $ 415), with a significant standard deviation of ₹45,924.02 (≈ $ 535), suggesting a wide range in the amounts borrowers seek. The debt-to-income ratio, with a mean of 1.1, indicates that borrowers, on average, seek loan amounts that are broadly commensurate with their monthly income. The mean loan maturity of approximately 8 months suggests a preference for short-term credit, while applying to this lender.

The acceptance rate across all applications is 26.04%, equivalent to 316,832 approved loans. Conditional on approval, 26.09% of them subsequently default. Finally, out of 1,216,524 loan applications to the fintech lender, 34,665 originate from referrals, representing 2.85% of the sample. However, the fraction of referred applications exhibits significant variation over time, as shown in Figure 1.

In addition, to these variables we observe the location of these applicants marked by Pincode, which is a six-digit number for the postal system in India and equivalent to Zipcode in the US. We also observe the reason of loan application or loan purpose, including education (4.78%), home renovation (9.27%), loan repayment (4.95%), medical (20.52%), monthly expenses (10.73%), purchases (4.49%), travel (2.87%), wedding (4.93%), or others (37.47%).

[6] In terms of Equifax ratings, 300-579 is Poor, 580-669 is Fair, 670-739 is Good, 740-799 is Very Good, and 800-900 is Excellent.

2.3 Credit Bureau Data

We obtain complete credit records for our sample applicants from the Indian credit bureau Equifax, providing a comprehensive record of their formal credit history. Specifically, this dataset contains information on all inquiries conducted by financial institutions, loans disbursed, and the subsequent default status of those loans. Moreover, our data allows us to identify the lender.

Our credit bureau data specifies loan types for disbursed loans, including personal, consumer, credit card, gold, two-wheeler, auto, education, and other categories. Following Ghosh and Vats (2022), we exclude credit cards from our analysis due to complexities in accurately determining their origination date and default status within the credit bureau data, stemming from flexible reporting standards. For example, when a credit card expires and a replacement is issued, the new card is often recorded as a new loan in the credit bureau data. However, in most instances, a new credit card is automatically mailed as the expiration date of the current credit card approaches, indicating neither a new application from the borrower nor a formal approval process by the lender.

The credit bureau data lacks a direct indicator of whether an inquiry resulted in a loan, hindering precise identification of application acceptance. To address this, we first remove all inquiries related to risk management. Then, we classify any non-risk-management inquiry as resulting in a loan if a disbursed loan is observed for the same applicant and lender within six months of the inquiry. This methodology for estimating loan acceptance probability using credit bureau data has been previously employed by Mishra, Prabhala, and Rajan (2022) and Cramer, Ghosh, Kulkarni, and Vats (2024) in the context of India, and by Jiménez, Ongena, Peydró, and Saurina (2014) in the case of Spain.

Therefore, our credit bureau data allows us to track each application – measured by inquiry – to acceptance and subsequent default, providing us with a complete record of formal credit market transactions of our sample applicants.

2.4 Call Detail Records (CDR) of customers

We collect detailed cellphone Call Detail Records (CDR) for a subsample of first-time customers, allowing us to precisely track their communication activities. This dataset includes anonymized customer identifiers, exact timestamps for each incoming and outgoing call, and the anonymized phone numbers of the other parties involved. Furthermore, the anonymized phone numbers of other parties are mapped to the respective anonymized customer identifiers if those parties are also customers of the FinTech lender.

These records are obtained with the explicit consent of customers who authorized our FinTech lender to access their digital footprint data. The data was collected over a brief window prior to the implementation of Google's privacy policy, and we are able to recover a portion of these records from

the FinTech's digital data archives. The sample period for this data covers the early part of our sample period and spans from January 2017 to April 2018.

The detailed digital information allows us to identify calls made between customers and their referrers, providing direct insight into the strength and evolution of their ongoing relationship. The temporal dimension of the data further allows us to examine communication patterns around critical periods, such as the loan repayment due dates.

3 Baseline Results

The objective of this paper is to examine the relationship between referrals and credit market outcomes. Specifically, we examine whether an application referred by an existing customer affects the likelihood of its approval and, if accepted, the probability of default. To this end, we estimate the regression specification 1 using the sample of first-time applicants to the platform. As noted earlier, this sample restriction allows for a comparison between potential borrowers who were referred by existing customers and those who were not, within a sample of loan applicants lacking any prior relationship with the platform.

$$y_i = \beta \cdot \text{Referred}_i + X_i + \varepsilon_i \tag{1}$$

where, y_i represents the key outcomes of interest: (1) loan acceptance, taking a value of one if the application is approved and zero otherwise; and (2) default, taking a value of one if default occurs for the approved loan and zero otherwise. The indicator Referred_i takes a value of one if applicant i was referred by an existing customer, and zero otherwise. The vector X_i includes a comprehensive set of fixed effects, including month $\times$ zipcode, credit score decile, age decile, gender, education, loan purpose, product type, loan amount decile, and debt-to-income decile, which control for observable heterogeneity. Since lending decisions are made digitally without face-to-face interaction, these controls allow us to account for the exhaustive set of observable characteristics influencing the lender's decision.[7] Lastly, we cluster standard errors at the zip code level to address potential correlation of error terms within geographic areas.

A necessary preliminary step before examining the relationship between referrals and credit market outcomes is a comparison of key variables across the referred and non-referred applicant groups. Table 2 presents this comparison. Referred applicants exhibit similar gender, educational attainment, monthly income, and credit score profiles compared to their non-referred counterparts. While the differences in means between the two groups are statistically significant for some variables

[7] A caveat to this is that we do not have access to the proprietary model employed by the platform to inform its decision-making processes.

– such as age, loan amount requested, debt-to-income ratio, and loan maturity – the magnitude of these differences is often economically small.[8] For instance, the 0.66-year age difference between the two groups is 2.2% of the sample mean. These instances of statistically significant but economically insignificant differences are common in large samples as statistical power increases.

However, the two groups do exhibit minor economic differences across some variables. First, referred applicants are 8.6% more likely to possess a credit score relative to non-referred applicants. However, conditional on having a credit score, the two groups have similar credit scores. Specifically, the difference in credit scores between the groups is 3.35 points, a negligible amount. Finally, although referred applicants request slightly lower loan amounts, the debt-to-income ratio remains broadly similar across the two groups.

In summary, while statistically significant differences exist between referred and non-referred applicants, most are economically negligible, with the exception of slight variations in requested loan amount and credit score availability.

3.1 Referral Networks & Credit Market Outcomes

This section examines the relationship between referrals and key credit market outcomes, specifically focusing on application acceptance and default among approved loans.

We begin our analysis by examining the relationship between loan acceptance and whether the applicant is referred. Table 3 presents the results. The key coefficient of interest is β, associated with Referred_i in equation 1. We estimate a series of specifications, progressively adding fixed effects, as we move from column 1 to column 11. Column 11 reports the results of our preferred specification, which includes the full set of observable characteristics. Across all specifications, the estimate of β remains positive and statistically significant, indicating that referred applicants are more likely to be accepted by the platform. In particular, the estimate from column 11 implies that referred applicants are 3.1% more likely to be accepted compared to non-referred first-time applicants residing in the same zipcode and applying in the same month. This effect is economically meaningful, representing an 11.9% increase over the average acceptance rate of 26% in our sample.

Notably, the estimated effect remains relatively stable in magnitude across different specifications presented in Table 3, even as the model R^2 increases substantially from 20% in column 1 to 80% in column 11. Specifically, under the Oster (2019) framework, it is unlikely that our estimate of interest is driven by omitted variable bias.[9]

[8] Because our sample size is very large, a standard two-sample t-test would almost certainly find a statistically significant difference, which offers limited insight into actual inference. We address this issue by employing a bootstrap method with 10,000 iterations. Given that the ratio of non-referred to referred applicants is approximately 100:3, each iteration involves randomly sampling 10,000 non-referred and 300 referred applicants for each variable of interest. We then compute the difference in means between referred and non-referred applications for all the variables. We repeat this process 10,000 times and produce a distribution of mean differences for each variable. We use these distributions to derive the p-value to assess whether the observed difference is statistically distinguishable from zero.

[9] The Oster (2019) identified set for the coefficient, as we progress from column 1 to column 11, is $[0.0230, 0.0306]$, which does not

Next, we examine the relationship between loan default and whether the applicant is referred. Table 4 presents the results across similar specifications, with column 11 presenting our preferred specification. The estimate of β remains negative and statistically significant throughout, indicating that referred applicants are less likely to default on their loans. Specifically, referred applicants are 6.1% less likely to default than their non-referred counterparts in the same zip code and month. This corresponds to a 23.3% reduction relative to the average default rate in our sample. In particular, the magnitude of the estimate increases as we move from column 1 to column 11, suggesting that omitted variables tend to bias the effect downward under the Oster (2019) framework.

3.2 Heterogeneous Effect of Referrals

We examine the heterogeneity in the relationship between referrals and credit market outcomes by applicant characteristics. Specifically, we employ the machine learning methodology outlined in Wager and Athey (2018) to flexibly estimate how referral effects differ according to borrower characteristics such as credit score, age, income, gender, education level, loan amount requested, and loan purpose. Appendix Figures B.1 and B.2 present the heterogeneous effect of referrals. Overall, we do not observe any economically meaningful or statistically significant heterogeneity across all these applicant characteristics.

There are two key takeaways from this observation that the impact of referrals remains consistent across different segments of the population. First, from a policy perspective, this simplifies decision-making since a universal approach could be effective, reducing the need for targeted strategies based on subgroup differences. Second, it improves the generalizability of our findings, making them more applicable beyond the specific setting examined, in line with the external validity framework outlined in List (2020).

3.3 Mechanism: Selection or Disciplining

Next, we investigate the mechanisms underlying the observed relationship between loan acceptance or default and whether the applicant was referred by an existing customer. Conceptually, this effect could operate through two distinct channels. The first is the selection channel, whereby being referred provides the platform with a positive information signal about the borrower's quality that cannot be fully captured by observable characteristics. In this case, referrals act as a screening filter, with referrers choosing to recommend higher-quality borrowers. The second is the disciplining channel, which suggests that once referred, individuals may face greater accountability from their referrers, leading to improved ex-post repayment behavior. This form of discipline implies that the act of being referred influences borrowers' subsequent actions, rather than simply reflecting their inherent

include zero. This indicates that omitted variables are unlikely to explain away the observed effect. Furthermore, this identified set remains valid for all values of $\lambda > 1.25$, which is below the commonly used threshold of 1.5 suggested by Oster (2019).

creditworthiness. These two channels have different implications: a selection effect indicates that referred borrowers are more creditworthy, while a disciplining effect suggests that the referral process alters borrower behavior after the application.

We disentangle the two channels by leveraging comprehensive credit bureau data that captures the full credit history of individuals in our sample. This dataset enables us to observe acceptance and default outcomes across multiple financial institutions – some where the borrower was referred and others where it was not. We combine this dataset with the Khwaja and Mian (2008) identification strategy, which exploits within-borrower variation across lenders to isolate the disciplining effect from borrower-specific fixed characteristics. Specifically, we include borrower × year fixed effects, allowing us to compare acceptance and default outcomes for the same borrower across loans with and without a referral and lender × year fixed effects that account for all preferences and shocks at the lender level.

This approach helps us differentiate between the two channels: if referrals serve as an information signal about borrower quality, we should see higher acceptance rates when lenders can observe the referral compared to situations where it is not. Alternatively, if disciplining is the primary mechanism, we should not observe a difference in acceptance rates but rather a lower default probability among referred borrowers, reflecting improved ex-post repayment behavior driven by accountability.

Table 5 presents the results from the analysis combining the credit bureau data with the Khwaja and Mian (2008) identification strategy. The estimate of interest is the interaction term of Fintech × Referred, which compares the loan outcomes for customers that were referred to the fintech platform – which is the key source of our data – with the loan outcomes for the applications by the same customer to other lenders. We estimate this interaction across various fixed effects specifications, with column 4 presenting our preferred specification. Our preferred specification includes: (1) customer × year fixed effect to leverage within-borrower variation, (2) month × zipcode to account for local demand fluctuations or investment shocks, and (3) lender × year × zipcode fixed effect to address possible lender-specific shocks or preferences within regions. The third set of fixed effects controls for endogenous matching between lenders and zip codes, thereby mitigating concerns highlighted by Paravisini, Rappoport, and Schnabl (2023) that the classical Khwaja and Mian (2008) estimator may be biased in the presence of region-specific lender specialization.

Panel A of Table 5 reports the results for application acceptance. Across all specifications, the estimate of interest is economically small and statistically insignificant. This suggests that the referred borrowers are equally likely to be accepted by lenders who do not observe the referral signal. Panel B of Table 5 reports the results for loan default. Across all specifications, the estimate of interest is negative and statistically significant, indicating that the same borrower is less likely to default on loans for which they were referred relative to loans for which they were not. Economically, they are 3.8% less likely to default on their referred loans relative to non-referred loans.

Taken together, the results on negligible difference in application acceptance and a selective lower default among referred loans for the same customer, indicates that disciplining is the primary mechanism driving our results. Particularly, these findings suggest that the improved loan performance of referred borrowers is not merely a function of initial screening by their referrers, but rather reflects their behavioral changes due to the referral.

3.4 Disciplining: Evidence from Cellphone Communication

Next, we exploit detailed cellphone Call Detail Records (CDRs) to analyze communication patterns between customers and referrers over time. Specifically, we focus on periods surrounding critical events, such as loan repayment due dates. Our hypothesis is that if referrers influence repayment behavior primarily through disciplining, we should observe increased communication around these key dates.

To investigate this, we first identify loan due dates within our sample. Most loans (94%) are structured with recurring 30-day repayment intervals until maturity, while the remaining 6% involve a single repayment at maturity. We then limit our analysis to referred customers and examine daily calls between referrers and referred individuals within a 30-day window around the due date.

We begin by presenting descriptive evidence on the extensive margin, i.e., the likelihood of a call between the referred customer and the referrer on a given day. Figure 2a presents the results. Within the [-20, 10] day window around the due date, the baseline probability of a call is 8.7%. This probability gradually increases as the due date approaches, reaching 10% two days prior and 9.7% on the due date itself. Post-repayment, this probability falls dramatically to 7.7% within ten days of the repayment date. We find similar temporal trajectory for the intensive margin – the number of calls between a referrer and the referred on a given day. Figure 2b presents the results. On average, there are 0.33 calls twenty days before the due date, rising to a peak of 0.39 on the repayment day. Following repayment, calling frequency declines, falling below 0.30 by day ten. These patterns suggest that referrers increase communication and exert discipline shortly before repayment, likely to ensure timely payments.

While the results reported in Figure 2 are informative, the temporal variation in the communications between the referrer and the referred could be driven by customer-specific characteristics or local shocks. We address these concerns by estimating the following regression specification using Poisson pseudo-likelihood regression with customer and zipcode × month fixed effects:

$$y_{i,t} = \sum_{k=-20, k\neq 3}^{k=10} \beta_k \times \text{Days until Repayment}_k + \theta_i + \theta_{z,m} + \theta_t + \varepsilon_{i,t} \tag{2}$$

where θ_i denotes referred customer fixed effects and $\theta_{z,m}$ denotes zipcode × month fixed effect, where

month refers to the time (month-year) of the origination of the loan for which the customer i was referred. $t = 0$ refers to the date when the loan installment or repayment is due. θ_t denotes date fixed effect for the date on which $y_{i,t}$ is observed.

Figure 3a reports coefficients for β_k using a binary variable which takes a value of one if there was a phone call between the referrer and the referred on that day and zero otherwise as the key dependent variable ($y_{i,t}$). Consistent with the descriptive evidence in Figure 2a, the likelihood of a communication between the referrer and the referred increases as the repayment date approaches. We observe similar results when using the number of calls as the dependent variable. Figure 3b presents the results on the intensive margin.

Overall, we find robust evidence of an increase in communication between the referrer and the referred around the repayment date. Taken together, these results lend credence to our results in Section 3.3 that the referrers play an active disciplining role to ensure timely repayment.

3.5 When are Referrers More Effective?

So far, we have established that referrals influence lending outcomes because referrers can discipline referred borrowers, thereby lowering the likelihood of default. In this section, we examine how the relationship between the referrer and the referred borrower impacts the probability of default. This exercise is crucial for identifying when referrers can effectively exercise their influence to discipline borrowers.

To this end, we examine the importance of features that characterize the relationship between the referrer and the referred borrower. These features essentially capture the social collateral or the position of the referrer within the social network relative to the referred borrower. Specifically, the features we consider include whether the referrer is older than the borrower, if they work at the same firm, whether they belong to the same social group (defined by caste), if the referrer's income is higher, and the various gender pairings between them. These factors collectively influence the strength of the social ties and, consequently, the ability of the referrer to discipline the borrower and lower the risk of default.

To this end, we examine the importance of different features that capture the relationship between the referrer and the referred borrower to lower default. Fundamentally, these features define referrers' social collateral or their location in the social network relative to the referred borrower. These features include whether the referrer is older than the referred borrower, if they work at the same firm, whether they share the same social group defined using caste, if the referrer's income is higher, and the various gender pairings between them.

We estimate eleven machine learning models to understand how each feature shapes the prediction of default. We use machine learning models – instead of simple linear models – to identify the importance of each feature in predicting default, as machine learning approaches enable us to

capture not only the direct effects of individual features but also their potential interactions with other features. This allows for a more comprehensive understanding of how different features influence the default risk.

The model we consider includes linear regression, logistic regression, decision tree, random forest, extra trees, gradient boosting, light gradient boosting machine (LightGBM), support vector machine (SVM), multilayer perception (MLP), TabNet, and wide & deep net. We estimate each of these models using the sample of accepted referred applications and then compute the feature importance for each feature across all models, assessing how well they contribute to predicting default. Panel A of Appendix Table A.1 presents the feature importance for each feature across models, and Panel B of Appendix Table A.1 presents the direction of effect for each feature on default across different models.

We take the average value of feature importance across all models to compute the overall importance of each feature. Figure 4 presents the average feature importance associated with each feature. Our findings indicate that referrers are most effective at reducing default when they are older, work at the same firm, belong to the same social group, and have a higher income. These characteristics suggest that a stronger social or economic connection increases the referrer's ability to influence the borrower's behaviour. Specifically, older referrers or those with higher income may have more experience, credibility, and willingness to offer financial support. Furthermore, sharing the same workplace or social background, such as caste, can also foster better communication and a sense of mutual responsibility. Together, these results suggest that referrals are associated with a lower likelihood of default when referrers possess higher social collateral – similar to the concept discussed in Karlan, Mobius, Rosenblat, and Szeidl (2009) – or when they occupy central positions within the network, as described in Breza and Chandrasekhar (2019).

3.6 Robustness

The results presented in Table 5 allow us to distinguish between the two channels, under the assumption that the new customer referred to our fintech platform was not simultaneously referred to other lenders. However, customers who are referred may have wide social networks that increase the likelihood of multiple referrals across different lenders. Since the credit bureau data does not explicitly observe whether an applicant was referred, we cannot directly account for this behavior. We address this potential bias by conducting a robustness check to confirm that our findings are not driven by this unobserved referral behavior. Specifically, we hand-collect information on referral programs implemented by each lender in our dataset, and exclude lenders that had any referral program during our sample period. This sample restriction reduces the probability that customers were referred to other lenders outside of our fintech, thus strengthening the validity of our counterfactual.

Panel A of Appendix Table B.1 presents the results comparing the acceptance and default

outcomes among referred customers applying to our fintech and to other lenders that did not have a referral program. The results are qualitatively consistent with those in Table 5, suggesting that the unobserved nature of referral programs at other lenders does not substantially bias our conclusions regarding the mechanism through which referral networks affect credit market outcomes.

Another concern with the analysis in Table 5 is that we compare outcomes between our fintech and all other lenders. However, shadow banks possess specific comparative advantages in certain market segments, as documented in Cramer, Ghosh, Kulkarni, and Vats (2024). These differences in comparative advantages or business models across lender types could confound our conclusions. We address this by restricting the comparison to only include fintech shadow banks among our other lenders, thus constructing a more appropriate counterfactual with lenders that have a similar behavior. Panel B of Appendix Table B.1 presents the results comparing acceptance and default rates among referred customers applying to our fintech and to other fintech shadow banks.[10] The findings are qualitatively similar to those in Table 5, indicating that differences in comparative advantages across lender types are unlikely to bias our conclusions regarding the mechanism through which referral networks influence credit market outcomes.

The results in Panel B of Table B.1 also help address concerns regarding the assumption that credit bureau inquiries represent actual loan applications. Mishra, Prabhala, and Rajan (2022) note that traditional banks, particularly state-owned banks, are less likely to conduct inquiries, meaning that not all approved or rejected loans are visible as inquiries on credit reports. Conversely, they argue that shadow banks almost always inquire in the credit bureau. Using more recent data, Cramer, Ghosh, Kulkarni, and Vats (2024) note that fintech shadow banks consistently make inquiries. Therefore, it is reasonable to conclude that, for the sample in Panel B, credit bureau inquiries reliably reflect actual loan applications.

Next, we examine the internal validity of the estimate presented in Table 5. The identification strategy relies on the assumption that applicants apply to multiple lenders. However, this requirement may induce selection bias if only certain types of applicants are willing or able to submit multiple applications. To assess this concern, we compare the observable characteristics of the full sample with those of applicants who applied to multiple lenders. Appendix Table B.2 presents this comparison. The results suggest that the two groups are similar along several observable characteristics such as gender, age, education, and income. However, the two groups exhibit a statistically significant difference in credit scores, with the Khwaja-Mian sample displaying higher scores. This difference is consistent with the idea that Khwaja-Mian borrowers are generally more active in credit markets. However, the economic significance of the difference is small: the 10-point difference between the two groups represents just a 1.5% difference relative to the overall sample. Most importantly, applicants in

[10] Note that this restriction considerably reduces our sample size, which reduces the statistical power of the test.

both samples have a similar likelihood of being referred. This similarity suggests that our estimates, despite being based on a selectively drawn sample, are unlikely to be biased by systematic selection and are therefore potentially internally valid.

Lastly, we conduct a battery of robustness checks to validate the results presented in Tables 3 and 4. Specifically, Appendix Table B.3 shows that our results are consistent when excluding data prior to January 2018 or January 2019, periods during which our fintech was in its early stages and heavily reliant on referrals for customer acquisition. Additionally, the results remain robust when excluding the COVID-19 period, which was marked by increased global uncertainty, health risks, and a surge in credit market interventions in India, such as debt moratorium, that could independently influence credit market outcomes (Fiorin, Hall, and Kanz, 2023).

4 Role of Incentives

So far, we have shown that that referred borrowers are less likely to default and that communication between referrers and referred borrowers increases around repayment due dates. This suggests that referrals improve repayment primarily because referrers can discipline referred borrowers. The effectiveness of this mechanism, however, may depend on the incentives referrers face. Monetary rewards can encourage customers to make more referrals, but may also expand referrals to borrowers with whom referrers have weaker relationships and less ability to influence repayment. At the same time, referrers may have non-monetary incentives to monitor referred borrowers if their own future relationship with the platform depends on those borrowers' performance. We next examine how these monetary and non-monetary incentives shape referral activity and credit-market outcomes.

4.1 Role of Monetary Incentives

We begin our analysis by examining the role of monetary incentives, specifically focusing on the impact of a referral bonus provided by the fintech firm to referrers. A referral bonus is a financial reward offered to existing customers to motivate them to refer new borrowers. When the person they refer takes out a loan, the referrer receives a monetary reward.

We hypothesize that while a larger referral bonus may encourage existing customers to refer more potential borrowers, it may also change the composition of referrals. In particular, stronger monetary incentives may induce referrers to recommend individuals with whom they have weaker social ties and, therefore, less ability to monitor or influence repayment behavior.

We exploit a policy change in the referral bonus amount implemented by the fintech firm to assess how monetary incentives influence referral behavior and associated credit market outcomes. Specifically, the fintech firm introduced a significant change to its referral bonus policy in October 2021 wherein it substantially increased the bonus paid to referrers, from ₹1,250 ($15) to ₹3,500 ($42) per successful referral.

The primary objective of this policy change was to increase the volume of applications received by the platform. This shift is evident in Figure 1, which shows an increase in both the total number of applications and the proportion originating from referrals after October 2021. Moreover, consistent with our framework, we find an increase in the likelihood of a referred application. Figure 5 presents the regression discontinuity plot illustrating the probability that an application is referred, before and after the change in the referral bonus policy in a 150-day window on each side of the policy change. We find that the likelihood of a new applicant being referred increases by 1.75 percentage points after the increase in referral bonus. Specifically, the referral probability increases to approximately 4% within 150 days following the policy change, compared to a stable rate of about 2% before the change. This increase represents a 50% increase over the pre-policy average value.

4.1.1 Effect on Credit Market Outcomes

We examine the effect of the change in referral bonus on application acceptance and loan default rates, respectively. This analysis extends the baseline empirical specification 1 by including an interaction term between the referral indicator and the post-policy period, where the post variable equals one for observations after October 2021.

Table 6 presents the results for the impact of an increase in referral bonus on application acceptance. The coefficient of the interaction term is negative and statistically significant, indicating that the likelihood of acceptance for referred applicants decreases after the increase in referral bonus. Specifically, referred applications were 4.3% more likely to be accepted before the bonus increase. However, their likelihood of acceptance decreases by 2.1 percentage points after the increase in referral bonus. This indicates a substantial decline in the acceptance rate for referred applications after the bonus increase, representing approximately a 50% reduction compared to the pre-change probability of acceptance for referred applicants relative to other applicants.

Table 7 presents the results for the impact of an increase in referral bonus on loan default. The coefficient of the interaction term is positive and statistically significant, indicating that the likelihood of default for referred applicants increases after an increase in referral bonus. Specifically, referred applications were 7.9% less likely to be default before the bonus increase. However, their likelihood of default increases by 4.2 percentage points after the increase in referral bonus. This indicates a substantial increase in default probability for referred applications after the bonus increase, representing approximately a 50% increase compared to the pre-change probability of default for referred applicants relative to other applicants.

The key identifying assumption of this analysis is that the relative credit outcomes for the referred and non-referred applicants would have evolved similarly in the absence of the change in referral bonus. While this assumption is inherently untestable, we present pre-trends analysis in Figure

6, suggesting that our identifying assumption is likely to be valid and our results are unlikely to be driven by pre-existing trends.

Overall, the results indicate that an increase in referral bonus is associated with an increase in default and a decline in application acceptance.

4.1.2 Effect on Selection & Disciplining

Next, we extend our analysis to examine the role of referral bonuses in affecting the underlying mechanism through which referrals affect credit market outcomes. Specifically, we augment the Khwaja and Mian (2008) analysis presented in Section 3.3 to compare acceptance and default outcomes for the same applicant across lenders, before and after the policy change.

Panel A of Table 8 presents the results for the effect of a change in referral bonus on application acceptance. We do not observe a statistically significant or economically meaningful change in the probability of acceptance following the change in the referral bonus. This result is consistent with the argument that referrals do not aid in selection or screening, and therefore, a change in referral bonus is unlikely to affect the probability of acceptance.

Panel B of Table 8 presents the results for the effect of a change in referral bonus on the likelihood of default. The results indicate an increase in default probability for customers who were referred after the increase in referral bonus relative to the customers referred before the increase. Specifically, before the increase, a referred borrower was 6.5% less likely to default on loans for which they were referred relative to loans of other lender where they were not referred. However, this relative difference in default probability for the same borrower shrinks by 4.5 percentage points after the increase in referral bonus. These results suggest that the ability of the referrers to discipline borrowers reduces after the increase in referral bonus.

Overall, these results suggest that increasing the referral bonus may incentivize referrers to recommend a broader pool of candidates, including those less familiar or closely connected to them. Consequently, this can undermine the referrer's capacity to effectively exercise oversight or impose discipline over the referred borrowers.

4.1.3 Frequent Referrers & the Disciplining Effect

The evidence above suggests that a larger referral bonus expands the set of borrowers who are referred to the platform while weakening the repayment advantage associated with referrals. A related possibility is that referrers who make many referrals are, on average, less closely connected to the borrowers they refer. A referrer may be able to monitor or influence repayment for a small number of close contacts, but this ability may diminish as the number of referred borrowers increases.

We examine this possibility by studying the distribution of referrals across referrers. Figure 7 illustrates the distribution of applications based on the number of unique referrals per referrer. For

example, the bar labeled "1" on the X-axis shows the proportion of referred applications that were referred by referrers who made exactly one referral during our sample period. Overall, 58.1% of all referred applications are referred by referrers who made only one referral. While most referred applications come from referrers with fewer than two referrals, a notable share is referred by those who refer many applicants. Specifically, 16.8% of referred applications are from referrers who refer more than ten applicants. We refer to these referrers, who refer multiple applicants to the platform, as frequent referrers. We posit that these frequent referrers tend to refer individuals with whom they have weaker ties or less familiarity, and as a result, are less able to enforce discipline over these borrowers.

We define a referrer to be a frequent referrer if they refer five or more applicants to the platform. Appendix Tables B.4 and B.5 show that applicants referred by these frequent referrers are less likely to be accepted and more likely to default. These results are consistent with the possibility that frequent referrers draw from a broader set of relationships that are, on average, less effective at supporting repayment.

We provide more direct evidence by extending the within-borrower, across-lender specification in Table 5. Specifically, we interact the FinTech × Referred indicator with an indicator for whether the borrower was referred by a frequent referrer. This interaction captures whether the effect of a referral differs when the referrer has made at least five referrals.

Panel A of Table 9 presents the results for application acceptance. It shows that neither being referred nor being referred by a frequent referrer significantly impacts the likelihood of application approval. This reinforces the idea that referrers are unlikely to influence credit market outcomes through borrower selection.

Panel B of Table 9 presents the results for loan default. The interaction between the FinTech × Referred indicator and the frequent-referrer indicator is positive and statistically significant. Thus, the reduction in default associated with referrals is substantially smaller when the borrower is referred by a frequent referrer. Moreover, the total effect of being referred by a frequent referrer on default is statistically indistinguishable from zero. In other words, the lower default probability associated with referrals disappears for borrowers referred by customers who make many referrals.

This result is robust to alternative definitions of frequent referrers. Appendix Table B.6 shows that the attenuation of the disciplining effect becomes larger as the threshold used to define frequent referrers increases. For example, the relevant interaction coefficient rises from 0.07 when a frequent referrer is defined as someone who refers more than three borrowers to 0.11 when the threshold is more than ten borrowers. Overall, these findings suggest that referrers are most effective when they refer a limited number of borrowers, consistent with the importance of close relationships for monitoring and discipline.

4.2 Role of Non-Monetary Incentives

This section examines the role of non-monetary incentives in motivating referrers to discipline referred borrowers. In particular, we discuss the role of dynamic interactions between customers and the platform – such as the intention to seek additional loans – which can serve as a motivation for referrers to discipline referred borrowers, thereby increasing their chances of loan approval.

We investigate the relationship between the likelihood of reapplying for a loan and the propensity to refer new customers to the platform. We define all borrowers who got their first loan from the platform in month m, as belonging to cohort m. We then track these cohorts over time to assess whether borrowers who reapply for a new loan within one, two, and three years of their initial loan are also more likely to refer new customers. This cohort-based approach follows the methodology of Kundu and Vats (2024) and helps mitigate concerns related to right-censoring bias. Additionally, we incorporate cohort $\times$ zipcode fixed effects to ensure comparisons are made among borrowers from the same cohort and exposed to similar demand shocks.

Table 10 presents the relationship between reapplication behavior and referral activity. Specifically, column 1 models the likelihood of reapplying for a loan as the dependent variable; column 2 examines the number of new applications submitted; and column 3 considers the number of approvals received. Panel A focuses on reapplication and referral activity within twelve months of the initial loan. We document that customers who refer new borrowers are more likely to reapply and to be approved for subsequent loans. These findings are robust to the inclusion of cohort $\times$ zipcode fixed effects and a range of individual-level controls, including credit score, income, age, gender, and education. Similar patterns persist when extending the analysis to two-year (Panel B) and three-year (Panel C) horizons. Overall, customers who refer new borrowers are 18.6–22.2% more likely to reapply and 14.2–22.9% more likely to secure approval for new loans, suggesting a positive association between referral behavior and future borrowing activity.

Next, we examine the subset of borrowers who referred a new customer to the platform to examine the relationship between the likelihood of approval of their subsequent application and whether the referred borrower defaults on their loan. The purpose of this analysis is to investigate the role of disciplining behavior, which involves costs, in this dynamic relationship between the referrer and the platform. Table 11 presents the results, documenting how the default behavior of referred borrowers influences whether the subsequent applications by the original referrer are accepted. Specifically, we compare the approval rates of applications submitted by borrowers who referred a new customer who defaulted on their loan (and those who did not) to the approval rates of applications by borrowers who referred a new customer whose loan was ultimately not accepted.

Our results show that referrers whose referred borrowers did not default are 5.8–8.6% more likely to have their applications accepted compared to referrers whose referred borrowers were not

accepted. Conversely, referrers whose referred borrowers defaulted are 13.2–15.6% less likely to have their applications accepted than those whose referred borrowers were not accepted. The difference between these two coefficients is statistically significant, as indicated by the χ^2 statistic. Overall, these findings suggest that referrers who refer borrowers that default are more likely to be rejected in subsequent applications, whereas those whose referrals do not default are more likely to be approved. These results remain robust even after accounting for cohort $\times$ zipcode fixed effects and controlling for individual characteristics such as credit score, income, age, gender, and education. This implies that referrers who effectively discipline their referred borrowers are more likely to be rewarded with future loan approvals, highlighting the importance of the dynamic relationship between the customer and the platform in fostering incentives to discipline.

5 A Model of Referrals

The empirical evidence shows that referrals are associated with lower default rates, primarily through a disciplining channel rather than borrower selection. This disciplining effect is stronger when referrers appear to have closer or more influential relationships with referred borrowers, but it weakens when the referral program is expanded through larger cash bonuses or when referrers make many referrals. These findings suggest that referral relationships differ in their capacity to improve repayment and that the lender's incentive scheme affects which referrers participate and how effectively they can discipline borrowers. We develop a model that captures these features and characterizes the optimal design and scale of a referral program.

Specifically, this section develops a theory of optimal referral contracts. In our model, a referrer can improve the outcome of a loan by disciplining the borrower, potentially by imposing a social cost on default. For example, a borrower may feel more obligated to repay when the loan application is supported by someone they respect. However, disciplining is possible only in some borrower-referrer relationships, and its effectiveness varies across relationships. A potential referrer privately knows whether they can discipline the borrower and, if so, how effective that discipline would be. The lender observes neither dimension.

This private information creates a screening problem. The lender would like to attract referrers who can meaningfully improve repayment while discouraging those who cannot. To do so, the lender offers a referral contract with four possible components: an upfront payment, a monetary bonus paid when the borrower repays, an increase in the approval probability of the referrer's future loan application following repayment, and a reduction in that probability following default.

We characterize the optimal referral contract. We first show that the optimal contract needs to include a penalty for the referrer when the borrower defaults. Because the referrer is protected by limited liability, the lender cannot impose a direct monetary penalty. Instead, the penalty takes the

form of a lower approval probability for the referrer's future loan applications. This penalty is essential for screening. Without a downside from default, even ineffective referrers would find it profitable to participate, leading the lender to pay for a large volume of referrals that do not sufficiently improve repayment. Put differently, referrers must bear some cost when they refer borrowers whom they cannot meaningfully discipline.

The penalty, however, also exposes effective referrers to risk. Even an effective referrer can only reduce the probability of default, not eliminate it. The lender therefore needs to compensate effective referrers for participating. This compensation can take the form of an upfront payment, an ex post bonus, or a higher future approval probability following repayment. Consistent with the data, we show that the optimal contract does not use an ex post bonus when referrers discount future payoffs more heavily than the lender. In this case, compensating the referrer through a higher future approval probability is cheaper than providing the same incentive through an ex post bonus, because referrers value access to future credit more than it costs the lender to provide it. In fact, we show that the optimal contract combines a penalty in the form of a lower future approval probability following default with either an upfront cash payment or a higher future approval probability following repayment. We fully characterize when each of these two contractual regimes is optimal.

Finally, we derive a scalar sufficient statistic for evaluating the size of an existing referral program, including programs of the type studied in our empirical analysis. The statistic compares the benefit of expanding the program and attracting additional referrers at the margin, who are less effective than the existing pool of referrers, with the cost of strengthening incentives for all participating referrers. It therefore allows a lender to assess, using observable data, whether a referral program is too large or too small. The model also help identify ways to reduce program costs without compromising effectiveness, for example, by shifting compensation away from upfront cash payments and toward future approval benefits.

5.1 Setup

There are three dates, $t \in \{0, 1, 2\}$. A borrower requires $K_B > 0$ units of capital at $t = 0$ and may be referred to a risk-neutral lender by a referrer. With probability $\rho \in (0, 1]$, the referrer requires $K_F > 0$ units of capital at $t = 1$. The borrower and the referrer discount future payoffs by $\beta \in (0, 1)$, whereas the lender does not discount. All loans have a maturity of one period. The borrower's repayment is due at the beginning of $t = 1$, before the referrer learns whether she needs to borrow. The gross interest rate on all loans is $R \in (1, \beta^{-1})$ and is taken as exogenous.

The restriction $R < \beta^{-1}$ ensures that a borrower who repays with certainty still strictly prefers to borrow, so that differences in discount rates generate a motive for lending. For simplicity, R is taken as exogenous. This assumption can be justified by the lender's technological advantage in collecting and processing unconventional borrower information. Because this technology gives the lender superior

information about the segment of borrowers it serves, these borrowers may have limited access to comparable credit from competing lenders. The lender may therefore exercise market power and charge an interest rate close to the monopoly rate.

Borrowers are of two types. A high type repays with probability $q_h \in (R^{-1}, 1)$ and a low type with probability $q_l \in (0, R^{-1})$. Types are observed by the lender before lending.

A key feature of our model is that it allows for peer discipline in debt repayment. The disciplining technology is as follows. With probability $\eta \in (0, 1)$ the referrer can raise the borrower's repayment probability by exerting effort $e > 0$ at $t = 0$; we call such a borrower disciplinable. For a disciplinable borrower the increment to the repayment probability is $\Delta \in [0, \bar{\Delta}]$, drawn from a continuous distribution with cumulative distribution function G and density g. With probability $1 - \eta$ the borrower is undisciplinable and effort accomplishes nothing. The referrer privately observes both whether the borrower is disciplinable and, when she is, the realization of Δ. The lender observes neither.

Note that absent referrals, the lender extends credit to high types only, since $q_h R > 1 > q_l R$. We assume that only low-type borrowers seek a referral (i.e., $q_B = q_l$), as high-type borrowers do not require a referral to obtain credit. At $t = 1$, the referrer is a high type with probability $s_h \in (0, 1)$ and a low type with probability $s_l = 1 - s_h$. All probabilities are common knowledge.

5.2 Referral Contracts

A referral contract is a quadruple (π_0, π_r, a_h, a_l). The lender pays an upfront payment $\pi_0 \geq 0$ at $t = 0$ for the referral.[11] The lender pays $\pi_r \geq 0$ at $t = 1$ if the referred borrower repays. In addition, the lender commits to raise the referrer's own approval probability at $t = 1$ by $a_h \in [0, 1]$ in the event that the borrower repays and the referrer applies as a low type, and to reduce it by $a_l \in [0, 1]$ in the event that the borrower defaults and the referrer applies as a high type. For simplicity, we assume that the referrer has no referrer of his own at $t = 1$.

Our contracting space assumes that the lender can implement and commit to the state-contingent approval rules specified in the referral contract. For instance, in a digital lending environment, the lender can incorporate the performance of referred borrowers into an internal score that affects the referrer's future eligibility for credit, with repayment improving the score and default lowering it. Because the lender interacts with a large pool of potential referrers over time, failing to honor these rules would weaken the credibility and incentive value of the referral program more broadly. The absence of such commitment could therefore undermine the effectiveness of the referral program. In fact, as discussed later, some degree of state-contingent approval is necessary, as monetary payments alone cannot support a successful referral program.

[11] The lender approves all referred applications in equilibrium, so this is equivalent to making the upfront payment contingent on the borrower being approved.

To specify the contracting problem, it is convenient to define the date-1 value of the approval-based reward to the referrer, together with the associated cost to the lender, as follows

$$W_h := \rho\, s_l K_F\,(1-\beta q_l R), \qquad\qquad c_h := \rho\, s_l K_F\,(1-q_l R), \tag{3}$$

Here, W_h is the referrer's date-1 value of a one-unit increase in a_h, and c_h is the cost to the lender of making this commitment. This instrument is activated with probability ρs_l, when the referrer needs to borrow at $t = 1$ and is a low type, so that the loan would otherwise be denied. Conditional on this event, the value of receiving the loan to the referrer is $K_F(1-\beta q_l R)$, while the cost to the lender is $K_F(1-q_l R)$. Therefore, when the borrower repays, the referrer receives $a_h W_h$ in the form of a higher loan approval probability, at a cost of $a_h c_h$ to the lender.

Likewise, we define the date-1 loss to the referrer from a one-unit increase in a_l, together with the associated cost to the lender, as follows:

$$W_l := \rho\, s_h K_F\,(1-\beta q_h R), \qquad\qquad c_l := \rho\, s_h K_F\,(q_h R-1). \tag{4}$$

The shadow cost of each instrument is the ratio of its cost to the lender to its value or loss to the referrer

$$\kappa_h \;:=\; \frac{c_h}{W_h} \;=\; \frac{1-q_l R}{1-\beta q_l R}, \qquad\qquad \kappa_l \;:=\; \frac{c_l}{W_l} \;=\; \frac{q_h R-1}{1-\beta q_h R}. \tag{5}$$

Furthermore, we denote by Δ_m the value of Δ that makes the lender break even:

$$\Delta_m \;:=\; R^{-1} - q_B. \tag{6}$$

To ensure that the referral can be profitable, we assume that $\bar{\Delta} > \Delta_m$. We further assume that η is sufficiently small that it is unprofitable to offer a contract that induces participation by referrers who lack the ability to discipline the borrower. Thus, screening for effective referrals is essential. These assumptions are stated formally below.

Assumption 1.

$$(q_B + \bar{\Delta})R - 1 > 0 > (q_B + \eta\mathbb{E}[\Delta])R - 1. \tag{7}$$

5.3 Incentive Constraints

Suppose that the borrower is disciplinable and that the realization of the disciplining effect is Δ. Given the referral contract (π_0, π_r, a_h, a_l), two constraints ensure that the referrer both participates in the referral program and exerts effort to discipline the borrower. First, the individual-rationality constraint requires that the referrer prefers referring the borrower and exerting effort to not referring,

as formalized in Equation (8).

$$(\mathrm{IR}^D): \quad \pi_0 - e + \beta\Big\{(q_B + \Delta)\big(\pi_r + a_h W_h\big) - (1 - q_B - \Delta)\, a_l W_l\Big\} \;\geq\; 0. \tag{8}$$

Second, the incentive-compatibility constraint requires that, conditional on referring the borrower, the referrer prefers exerting effort to shirking, as stated in Equation (9).

$$(\mathrm{IC}^D): \quad \beta\Delta\big(\pi_r + a_h W_h + a_l W_l\big) \;\geq\; e. \tag{9}$$

Equation (9) highlights three ways in which the lender can incentivize the referrer to discipline the borrower after receiving the upfront payment. First, the lender can offer an ex-post bonus (π_r) that is paid following repayment. Second, the lender can reward repayment by increasing the referrer's future loan-approval probability, as captured by a_h. Third, the lender can punish default by reducing the referrer's future approval probability, as captured by a_l.

The strength of the latter two incentive channels depends on the value of the referrer's future access to credit. In particular, both W_h and W_l are proportional to expected future loan demand, ρK_F (see Equations (3) and (4)). Thus, when the referrer expects to borrow more in the future, the lender can provide stronger incentives through future approval decisions. We focus on environments in which this channel is sufficiently strong that either the repayment reward or the default penalty, when used at its maximum feasible level, is by itself sufficient to induce effort when $\Delta \geq \Delta_m$. Assumption 2 formalizes this condition.

Assumption 2. *The referrer's expected future loan demand is sufficiently large that setting $a_h = 1$ or $a_l = 1$ each incentivizes a referrer with the break-even type (i.e., $\Delta = \Delta_m$) to discipline:*

$$\beta\Delta_m \min\{W_l, W_h\} \geq e \quad \Longrightarrow \quad \rho K_F \geq \max\Big\{\frac{1}{s_h(1-\beta q_h R)}, \frac{1}{s_l(1-\beta q_l R)}\Big\}\frac{e}{\Delta_m \beta}. \tag{10}$$

Assumption 2 does not imply that setting either the reward or the penalty instrument at its maximum is optimal. It only requires that expected future borrowing demand be sufficiently large for approval-based instruments to provide enough incentives to discipline the referrer. The assumption can be relaxed at the cost of a more involved characterization. In particular, when expected future borrowing demand is low, approval-based incentives may be insufficient on their own, and monetary incentives may also enter the optimal contract. Assumption 2 therefore allows us to focus on the design of the optimal referral contract in settings where the prospect of future access to credit can play an important role in providing incentives.

The lender also wants to discourage the referral when the borrower is undisciplinable, by

Assumption 1. This requires

$$(\mathrm{IR}^{UD}): \quad \pi_0 + \beta\Big\{q_B\big(\pi_r + a_h W_h\big) - (1 - q_B)\, a_l W_l\Big\} \;\leq\; 0. \tag{11}$$

Note that this condition cannot hold if the referral contract uses only monetary incentives, that is, if $a_h = a_l = 0$ and either $\pi_0 > 0$ or $\pi_r > 0$, because the left-hand side is then strictly positive. Put differently, a purely monetary referral program attracts all referrers and is therefore unprofitable, since only a small fraction of referrers can effectively discipline the borrower. Specifically, the condition in Equation (11) requires $a_l > 0$, meaning that the referral contract must include a penalty in the form of a lower future loan approval probability to discourage ineffective referrals.

The three constraints in Equations (8), (9), and (11) depend on the four contractual instruments only through two scalars. Define the *incentive slope*

$$S \;:=\; \pi_r + a_h W_h + a_l W_l \;\geq\; 0, \tag{12}$$

and the *participation intercept*

$$T \;:=\; \pi_0 + \beta\Big\{q_B\big(\pi_r + a_h W_h\big) - (1 - q_B)\, a_l W_l\Big\}, \tag{13}$$

which is also the payoff to a referrer who refers an undisciplinable borrower. Then, the three constraints are equivalent to

$$(\mathrm{IC}^{D}): \beta\Delta S \geq e, \qquad (\mathrm{IR}^{D}): T + \beta\Delta S \geq e, \qquad (\mathrm{IR}^{UD}): T \leq 0. \tag{14}$$

5.4 The Optimal Referral Contract

This section characterizes the lender's problem and derives the optimal referral contract. The lender's expected profit from offering the referral contract (π_0, π_r, a_h, a_l), assuming that the conditions in Equation (14) hold for $\Delta \geq \underline{\Delta}$, is

$$V \;=\; \eta \int_{\underline{\Delta}}^{\bar{\Delta}} \Big\{(q_B + \Delta) R K_B - K_B - C(\Delta)\Big\} g(\Delta)\, d\Delta, \tag{15}$$

where $C(\Delta)$ denotes the expected cost of the incentives associated with the participation of a referrer of type Δ:

$$C(\Delta) \;=\; \pi_0 + (q_B + \Delta)\big(\pi_r + a_h c_h\big) + (1 - q_B - \Delta) a_l c_l. \tag{16}$$

In Equation (15), $(q_B + \Delta) R K_B - K_B$ is the gross profit from lending to the referred borrower when the referrer increases the borrower's repayment probability by Δ. Without the referral, this

profit is zero because the lender does not lend to low-type borrowers. The term $C(\Delta)$ captures all costs of the referral contract incurred by the lender. The lender pays π_0 upfront for the referral. In the repayment state, which occurs with probability $q_B + \Delta$, it pays π_r and provides a reward in the form of a higher future loan approval probability, at a cost of c_h per unit. In the default state, which occurs with probability $1 - q_B - \Delta$, it incurs a cost of c_l per unit of the penalty, which takes the form of a lower future loan approval probability.

The lender maximizes (15) over $(\pi_0, \pi_r, a_h, a_l) \geq 0$, with $a_h, a_l \in [0, 1]$, subject to $T \leq 0$. We solve the lender's problem in two steps. First, for any $d \in [\Delta_m, \bar{\Delta}]$, we characterize the contract that induces participation by referrers with type $\Delta \in [d, \bar{\Delta}]$ with the lowest cost. Then, we use the characterization to find the optimal value of $\underline{\Delta} = d$.

5.4.1 Cost-Minimizing Contracts

The objective of this section is to characterize, for a given threshold type d, the values of (π_0, π_r, a_h, a_l) that minimize the average cost among all contracts satisfying the conditions in Equation (14) for $\Delta \in [d, \bar{\Delta}]$, where the average cost can be obtained from Equation (16):

$$\bar{C}(d) = \pi_0 + q^*(d)(\pi_r + a_h c_h) + (1 - q^*(d))a_l c_l, \tag{17}$$

where

$$q^*(d) = q_B + \mathbb{E}[\Delta \mid \Delta \geq d]. \tag{18}$$

Lemma 1 establishes two necessary conditions for all cost-minimizing contracts.

Lemma 1. *If the contract specified by (π_0, π_r, a_h, a_l) minimizes $\bar{C}(d)$ among all contracts satisfying the conditions in Equation (14) for $\Delta \in [d, \bar{\Delta}]$, then*

$$T = 0, \qquad S = \frac{e}{\beta d}. \tag{19}$$

Lemma 1 states that a cost-minimizing contract sets $T = 0$, so that the individual-rationality constraint for referring an undisciplinable borrower binds. The intuition is that if instead $T < 0$, the lender would have to provide a larger incentive slope S to induce all referrers with $\Delta \geq d$ to participate:

$$T + \beta \Delta S \geq e \qquad \forall\, \Delta \geq d \qquad \Longrightarrow \qquad S \geq \frac{e - T}{\beta d}. \tag{20}$$

The larger incentive slope raises the cost of implementing the contract for all inframarginal participating types, $\Delta > d$. A cost-minimizing contract therefore sets $T = 0$ and implements the minimum incentive

slope consistent with participation by the threshold type, that is

$$S(d) = \frac{e}{\beta d}. \tag{21}$$

Consequently, the screening constraint for undisciplinable borrowers binds, and both the participation and incentive-compatibility constraints bind for the marginal disciplinable type $\Delta = d$.

This intermediate step narrows the set of cost minimizing contracts to those that satisfy the conditions in Lemma 1. Proposition 1 characterizes the cost-minimizing contract for each $d \in [\Delta_m, \bar{\Delta})$ and reports its associated average cost.

Proposition 1 (Cost-Minimizing Referral Contracts)**.** *Suppose Assumption 1 and 2 hold, and define*

$$\sigma(d) \;:=\; -\beta + q^*(d)\,\kappa_h - (1 - q^*(d))\,\kappa_l. \tag{22}$$

a) *If $\sigma(d) \neq 0$, there is a unique cost-minimizing contract $(\pi_0^*(d), \pi_r^*(d), a_h^*(d), a_l^*(d))$ that induces participation for referrers with type $\Delta \geq d$ with the following form:*

(i) [Cash Regime] If $\sigma(d) > 0$, then the cost-minimizing contract uses a mix of upfront bonus and ex-post penalty in the form of lower future approval probabilities. Specifically,

$$\pi_0^*(d) = \frac{(1 - q_B)e}{d}, \quad \pi_r^*(d) = 0, \quad a_h^*(d) = 0, \quad a_l^*(d) = \frac{S(d)}{W_l}. \tag{23}$$

(ii) [Credit Access Regime] If $\sigma(d) < 0$, then the cost-minimizing contract offers no upfront or ex post bonuses and all incentives are provided through future approval probabilities. Specifically,

$$\pi_0^*(d) = 0, \quad \pi_r^*(d) = 0, \quad a_h^*(d) = \frac{(1 - q_B)S(d)}{W_h}, \quad a_l^*(d) = \frac{q_B S(d)}{W_l}. \tag{24}$$

(iii)If $\sigma(d) = 0$, the contracts specified in Equations (23) and (24), as well as any convex combination of these two contracts, are cost-minimizing.

b) *The average cost under a cost-minimizing contract is*

$$\bar{C}^*(d) = \left[\beta(1 - q_B) + \left(1 - q^*(d)\right)\kappa_l + (1 - q_B)\min\{\sigma(d), 0\}\right] S(d). \tag{25}$$

Part (a) of Proposition 1 shows that every cost-minimizing referral contract, and hence the optimal contract, includes a penalty following the borrower's default in the form of a reduction in the referrer's future loan-approval probability. As discussed above, this penalty is necessary for screening:

without a downside from default, referrers who cannot discipline the borrower would also find it attractive to participate.

The penalty, however, exposes even effective referrers to the risk of losing access to future credit, because disciplining reduces but does not eliminate the probability of default. The lender therefore needs to compensate the referrers he aims to attract. Part (a) of Proposition 1 shows that this compensation can take the form of either an upfront payment or a higher future approval probability following repayment. Which instrument is cost-minimizing is determined by the sign of $\sigma(d)$ in Equation (22).

The relative attractiveness of upfront compensation stems, in part, from the difference in discounting between the referrer and the lender. An upfront dollar is more valuable to the referrer than a dollar paid in the future, whereas the lender does not discount these payments. Thus, holding fixed the shadow costs of the approval instruments (κ_l and κ_h) a lower β makes upfront compensation relatively more attractive.

Upfront payments, however, provide compensation without directly strengthening the referrer's incentive to discipline the borrower. In particular, π_0 affects the participation intercept T but does not enter the incentive slope S. By contrast, repayment-contingent rewards enter the incentive slope and can therefore both compensate the referrer and provide incentives for discipline. Replacing upfront compensation with an ex-post reward can consequently reduce the penalty required to implement a given incentive slope.

Among ex-post rewards, the model implies that a monetary bonus is dominated by a reward in the form of a higher future loan-approval probability. To deliver one unit of date-0 value to the referrer, an ex-post monetary bonus costs the lender β^{-1}, whereas an approval-based reward costs $\beta^{-1}\kappa_h$, which is less than β^{-1}. The intuition is that, due to the difference in the discount rates, the value of a loan to the referrer is greater than its cost to the lender. As a result, as long as the referrer expects to borrow a sufficiently large amount, as imposed by Assumption 2, an ex-post monetary bonus is never part of a cost-minimizing contract. Figure 8 illustrates how the form of the cost-minimizing contract varies with the model parameters.

Part (b) of Proposition 1 characterizes the minimum average cost of implementing a given threshold d. The average implementation cost therefore depends both on the required strength of incentives ($S(d) = \frac{e}{\beta d}$) and on how these incentives are allocated across contractual instruments.

When $\sigma(d) \geq 0$, the lender provides the entire incentive slope through the default penalty, so that $a_l W_l = S(d)$. The first term inside the brackets in Equation (25) captures the upfront compensation required to offset the expected penalty, while the second captures the lender's expected cost of imposing the penalty when the referred borrower defaults. When $\sigma(d) < 0$, it is cheaper to replace the upfront payment with a reward in the form of a higher future approval probability following repayment. The

resulting cost savings are captured by the last term inside the brackets in Equation (25).

5.5 A Sufficient Statistic for the Size of Referral Program

The previous part characterizes the cost-minimizing contract for any given referral threshold $d \in [\Delta_m, \bar{D})$. We now determine the optimal threshold. Using the definition of $q^*(d)$ in Equation (18), the lender's expected payoff in Equation (15) can be written as

$$V(d) = \eta\big(1 - G(d)\big)\left\{[q^*(d)R - 1]\, K_B - \bar{C}^*(d)\right\}. \tag{26}$$

The first term inside the brackets is the average value generated by a participating referral, while $\bar{C}^*(d)$ is the minimum average cost of implementing the threshold d, characterized in Equation (25).

To characterize the optimal threshold, define

$$\Lambda(d) := \frac{\lambda(d)v(d)}{\bar{C}^*(d)}, \tag{27}$$

where $v(d)$ is the net value generated by the marginal participating referrer, and

$$\lambda(d) := \frac{dg(d)}{1 - G(d)}, \tag{28}$$

is the elasticity of the referral rate with respect to the threshold d.

Proposition 2 (Sufficient Statistic for the Optimal Referral Threshold)**.** *Suppose Assumptions 1 and 2 hold. Any interior optimal threshold d^* satisfies*

$$\Lambda(d^*) = 1. \tag{29}$$

Moreover, $\Lambda(d) > 1$ (< 1) implies that locally lowering (raising) d increases the lender's value $V(d)$.

Proposition 2 highlights the tradeoff that determines the optimal scale of the referral program. To understand the optimality condition in Equation 29, consider a reduction in d by εd. The lower threshold attracts additional referrers with disciplining ability close to d. Each marginal referral generates net value $v(d)$, while the increase in the number of referrals is $\varepsilon\lambda(d)(1 - G(d))$. Thus, the benefit of expanding the program is $\varepsilon\lambda(d)v(d)(1 - G(d))$.

Expanding the program, however, also requires stronger incentives for all referrers who would have participated under the original threshold. Holding the set of inframarginal referrers fixed, their contractual costs are proportional to the required incentive slope, $S(d) = e/(\beta d)$. An ε-fraction reduction in d therefore raises the average contractual cost of these inframarginal referrals

by approximately $\varepsilon\bar{C}^*(d)$, thereby increasing the total expected cost by $\varepsilon\bar{C}^*(d)(1-G(d))$. At an interior optimum, these marginal benefits and costs are equal, which yields $\Lambda(d^*) = 1$.

Accordingly, $\Lambda(d)$ is a sufficient statistic that measures whether the referral program is too selective or too broad. When $\Lambda(d) > 1$, the value generated by attracting additional referrals exceeds the associated increase in incentive costs, so the lender benefits from lowering d and expanding the program. When $\Lambda(d) < 1$, the opposite is true, and the lender benefits from raising d and reducing the size of the program.

The sufficient statistic $\Lambda(d)$ can be estimated using local variation in the generosity of the referral contract. Such variation shifts both referral participation and the average disciplining effect among participating referrers. The joint response of these two outcomes identifies the disciplining ability of the marginal referrer (d) which in turn determines the net value of admitting the marginal referral ($v(d)$). The local response of the referral rate to d identifies $\lambda(d)$. Importantly, this approach does not require specifying or estimating the distribution G: the distribution of disciplining ability enters $\Lambda(d)$ only through the locally identified elasticity $\lambda(d)$.

6 Empirical Implications for the Referral Program

Our model provides predictions for the optimal size of the referral program and the composition of the referral contract. We use these predictions to evaluate the lender's program and its response to the increase in the referral bonus. For this analysis, we use parameter estimates constructed from our empirical results and reported in Table 12. The table also describes how these estimates are obtained.

Was the increase in the referral bonus desirable? The October 2021 policy change increased the upfront referral bonus from Rs. 1,250 to Rs. 3,500 and substantially expanded referrals. Figure 5 shows that the referred share of applications rose from about 2% to 3.75–4%. At the same time, Table 7 shows that the average disciplining effect fell from 0.0794 to 0.0371.

Because our model concerns referrals that ultimately receive loans, the relevant measure of program size is the share of approved borrowers who are referred. Combining Figure 5 with the acceptance estimates in Table 6 and the overall acceptance rate in Table 1, we estimate that this share increased from about 2.3% to 4.1–4.3%.[12]

The total disciplining generated by the referral program is therefore approximately

$$0.023 \times 0.0794 \simeq 0.00184$$

[12] Let s denote the referred share of applications, A_R the acceptance rate for referred applications, and A_N that for non-referred applications. The referred share among approved borrowers is

$$\frac{sA_R}{sA_R + (1-s)A_N}.$$

We use $s \simeq 0.02$ before the policy and 0.0375–0.04 afterward from Figure 5, the overall acceptance rate of 26.04% from Table 1, and the acceptance effects in Table 6, column (11): 0.0429 before the policy and a 0.0206 decline afterward.

before the policy and

$$(0.041\text{–}0.043) \times 0.0371 \simeq 0.00152\text{–}0.00160$$

afterward. Thus, although the share of approved borrowers coming through referrals increased sharply, the total disciplining generated by the program did not increase. This implies that the additional referrals attracted by the larger bonus contributed very little, if anything, to repayment discipline. As a result, the marginal referrals after the expansion had no disciplining value, i.e., $v(d_{after}) \leq 0$. Since

$$\Lambda(d_{after}) = \frac{\lambda(d_{after})v(d_{after})}{\bar{C}^*(d_{after})},$$

with $\lambda(d_{after}) > 0$ and $\bar{C}^*(d_{after}) > 0$, this implies $\Lambda(d_{after}) \leq 0$. Hence, according to Proposition 2, the expansion associated with the bonus increase appears to have pushed the referral program beyond its efficient size.

Is the mix of upfront and approval-based compensation efficient? Proposition 1 gives a strong prediction for the choice between an upfront bonus and approval-based compensation. Table 12 reports $\kappa_h = 0.1857$ and $\beta = 0.75$, so that $\kappa_h < \beta$. Because $q^*(d)\kappa_h \leq \kappa_h < \beta$ and $\kappa_l > 0$, it follows that $\sigma(d) < 0$ for any composition of the referral pool. This means that future loan approval provides relatively inexpensive compensation to the referrer, making it preferable to upfront cash. The optimal contract therefore compensates the referrer through future approval probabilities rather than an upfront payment.

The model consequently predicts $\pi_0^* = 0$ conditional on any target referral threshold. In contrast, the observed program increased the upfront bonus from Rs. 1,250 to Rs. 3,500, approximately from 3.5% to 9.8% of requested loan principal. From the perspective of the model, the policy therefore shifted compensation further away from the cost-minimizing mix. Moreover, as discussed above, the resulting expansion of the referral program appears to have been inefficient. A more effective policy would have been to strengthen future approval rewards rather than increase the upfront bonus.

The absence of an ex-post monetary bonus. One feature of the observed contract is consistent with our results in Proposition 1: the program does not pay a monetary bonus following repayment. Both an ex-post cash bonus and an approval reward incentivize disciplining effort by referrers, but the approval reward is cheaper for the lender because future credit is more valuable to the referrer than it costs the lender to provide. This difference arises because the referrer discounts future payoffs more heavily than the lender. The absence of an ex-post monetary bonus is therefore consistent with the cost-minimizing contract.

Is the balance between approval rewards and penalties efficient? Proposition 1 also gives a prediction for the relative strength of the approval reward following repayment and the approval penalty

following default.As discussed earlier, given the parameter values in Table 1, the optimal contract provides incentives exclusively through future access to credit. According to Part (a.ii) in Proposition 1, this implies that

$$\frac{a_l^* W_l}{a_h^* W_h} = \frac{q_B}{1 - q_B}.$$

Using $q_B = 0.7391$ from Table 12, the model therefore predicts a penalty-to-reward ratio of approximately 2.83 in referrer-value units.

Because repayment is much more likely than default, the reward is received more frequently than the penalty is imposed. The penalty must therefore be stronger per occurrence to prevent ineffective referrers from finding participation attractive.

The data exhibit the same qualitative asymmetry. Table 11 shows that a default by the referred borrower is associated with a substantially larger reduction in the referrer's subsequent loan approvals than the increase associated with repayment. For example, over the three-year horizon, the coefficients are -0.1449 following default and 0.0803 following no default, giving a penalty-to-reward ratio of about 1.80. Thus, the observed contract places more weight on the penalty than on the reward, as predicted by the model. Quantitatively, however, the observed asymmetry appears somewhat weaker than the model predicts.

7 Conclusion

Referral programs are typically viewed as a way for firms to acquire customers and, in credit markets, potentially to obtain information about new borrowers. This paper suggests that referral programs can serve an additional function. By embedding a loan in a pre-existing social relationship, a referral may create repayment incentives after credit is extended.

We study a large digital lender in India that pays existing customers to refer new borrowers. Using detailed application-level data, we document that referred applicants are more likely to receive credit and, conditional on approval, are less likely to default than observably similar non-referred applicants. These differences may arise through two key mechanisms. First, referrers may relay information about borrowers that is not fully available in administrative data, allowing lenders to identify applicants with stronger repayment prospects. Second, referrers may affect borrower behavior after origination by reminding borrowers about payments, monitoring their circumstances, or creating interpersonal costs of default.

Our analysis suggests that post-origination discipline is an important component of the value created by referrals. Using credit-bureau data, we compare the same borrower's outcomes across lenders. If referrals primarily provided the focal fintech with information about borrower quality that was not available to other lenders, referred borrowers should be relatively more likely to receive credit from the focal fintech. We find no evidence of such an approval advantage. Yet the same borrower is

substantially less likely to default on a referred loan from the focal fintech than on loans from other lenders. This result suggests that referrals do more than relay information at origination: they activate social relationships that support repayment after the loan is made.

We provide complementary evidence on this mechanism using detailed call records that allow us to observe communication between each referred borrower and the specific customer who referred her. Whereas most studies infer peer monitoring or enforcement from credit outcomes alone, these data allow us to examine how communication within the referrer-borrower relationship changes around scheduled repayment dates. Communication rises as repayment approaches and declines shortly afterward. This repayment-centered pattern is difficult to reconcile with referrals operating only by relaying information at origination and is consistent with referrers becoming more engaged when repayment is due.

Finally, the repayment advantage associated with referrals is larger in relationships in which referrers plausibly have greater capacity to influence borrowers, including relationships in which the referrer is older, earns more, works at the same firm, or belongs to the same social group. Taken together, the within-borrower comparison across lenders, the repayment-centered pattern in referrer-borrower communication, and the stronger repayment effects in relationships characterized by greater potential social influence are most consistent with referrals improving loan performance by supporting repayment through post-origination discipline.

The evidence has direct implications for the design of referral programs. If referrals improve repayment by activating social relationships after origination, their value depends on whether referrers both have the capacity to influence borrowers and the incentive to use that capacity when repayment is due. A referral therefore does more than generate a lead or provide an additional input into underwriting: it gives the lender access to a relationship that may shape borrower behavior after the loan is made. Yet the lender does not observe either the strength of that relationship or the referrer's ability to affect repayment. Referral design must therefore address two linked problems. The lender must attract referrers whose relationships can support repayment and provide them with incentives to remain engaged after origination.

The model formalizes the resulting contracting problem. A prospective referrer privately knows whether she can influence a borrower's repayment and, if so, by how much. An upfront payment for a successful referral alone therefore attracts both referrers whose relationships can meaningfully improve repayment and those whose participation adds little value. The contract must therefore include a consequence for borrower default that screens out referrers who cannot meaningfully support repayment. When direct monetary penalties are infeasible, it can give referrers a stake in repayment by reducing their future access to credit following borrower default. Even effective referrers cannot fully eliminate default risk, however, so this downside incentive must be accompanied by

compensation, either through an upfront payment or through greater future credit access when the borrower repays. The lender's ongoing relationship with referrers thus provides instruments to both screen for relationships that can support repayment and encourage referrers to remain engaged after origination.

Our model also provides a sufficient-statistics approach for evaluating the design of referral programs. The sufficient statistic uses the observed increase in referral volume and the accompanying decline in the average repayment benefit following a change in program generosity to assess whether the program is too large or too small. The model also characterizes the optimal mix of incentives, including the tradeoff between upfront compensation and repayment-contingent rewards through future access to credit.

The evidence and model point to a broader implication for digital credit markets. Improvements in credit bureaus, digital identification, alternative data, and algorithmic underwriting may reduce the informational value of local relationships at the time of origination. Yet even if lenders can recover increasingly rich information about borrower quality, they may still face limits in influencing borrower behavior after a loan is made. Social relationships can remain valuable along this margin. They can provide monitoring, communication, and interpersonal accountability at moments when repayment decisions are made, functions that a remote lender may find difficult to replicate through data and automated underwriting alone. In this sense, formal lending technologies and informal social relationships can be complements in originating and administering credit at scale.

References

Agarwal, Sumit and Itzhak Ben-David. 2018. "Loan prospecting and the loss of soft information." *Journal of Financial Economics* 129 (3):608–628.

Agarwal, Sumit, Pulak Ghosh, Peiyi Jin, Shohini Kundu, Nishant Vats, Xinbo Wang, and Yingze Xu. 2024. "When Privacy Protects but Excludes: The Costs and Benefits of Privacy Regulation in Credit Markets." *Olin Business School Center for Finance & Accounting Research Paper* (2026/01).

Agarwal, Sumit and Robert Hauswald. 2010. "Distance and private information in lending." *The Review of Financial Studies* 23 (7):2757–2788.

Albuquerque, Rui and Hugo A. Hopenhayn. 2004. "Optimal Lending Contracts and Firm Dynamics." *Review of Economic Studies* 71 (2):285–315.

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

Alok, Shashwat, Pulak Ghosh, Nirupama Kulkarni, and Manju Puri. 2024. "Open Banking and Digital Payments: Implications for Credit Access." NBER Working Paper 33259, National Bureau of Economic Research.

Ambrus, Attila, Markus Möbius, and Adam Szeidl. 2014. "Consumption Risk-Sharing in Social Networks." *American Economic Review* 104 (1):149–182.

Attanasio, Orazio, Britta Augsburg, Ralph De Haas, Emla Fitzsimons, and Heike Harmgart. 2015. "The Impacts of Microfinance: Evidence from Joint-Liability Lending in Mongolia." *American Economic Journal: Applied Economics* 7 (1):90–122.

Banerjee, Abhijit, Emily Breza, Arun G. Chandrasekhar, Esther Duflo, Matthew O. Jackson, and Cynthia Kinnan. 2024. "Changes in Social Network Structure in Response to Exposure to Formal Credit Markets." *Review of Economic Studies* 91 (3):1331–1372.

Banerjee, Abhijit V., Timothy Besley, and Timothy W. Guinnane. 1994. "Thy Neighbor's Keeper: The Design of a Credit Cooperative with Theory and a Test." *Quarterly Journal of Economics* 109 (2):491–515.

Berg, Tobias, Valentin Burg, Ana Gombović, and Manju Puri. 2020. "On the Rise of FinTechs: Credit Scoring Using Digital Footprints." *Review of Financial Studies* 33 (7):2845–2897.

Berg, Tobias, Andreas Fuster, and Manju Puri. 2022. "FinTech Lending." *Annual Review of Financial Economics* 14:187–207.

Berger, Allen N, Christa HS Bouwman, and Dasol Kim. 2017. "Small bank comparative advantages in alleviating financial constraints and providing liquidity insurance over time." *The Review of Financial Studies* 30 (10):3416–3454.

Berger, Allen N, Christa HS Bouwman, Lars Norden, Raluca A Roman, Gregory F Udell, and Teng Wang. 2024. "Piercing through opacity: Relationships and credit card lending to consumers and small businesses during normal times and the COVID-19 crisis." *Journal of Political Economy* 132 (2):484–551.

Berger, Allen N, Marco A Espinosa-Vega, W Scott Frame, and Nathan H Miller. 2005. "Debt maturity, risk, and asymmetric information." *The Journal of Finance* 60 (6):2895–2923.

Berger, Allen N, W Scott Frame, and Vasso Ioannidou. 2011. "Tests of ex ante versus ex post theories of collateral using private and public information." *Journal of Financial Economics* 100 (1):85–97.

Berger, Allen N. and Gregory F. Udell. 1995. "Relationship Lending and Lines of Credit in Small Firm Finance." *Journal of Business* 68 (3):351–381.

———. 2002. "Small Business Credit Availability and Relationship Lending: The Importance of Bank Organisational Structure." *Economic Journal* 112 (477):F32–F53.

Besley, Timothy and Stephen Coate. 1995. "Group Lending, Repayment Incentives and Social Collateral." *Journal of Development Economics* 46 (1):1–18.

Bharath, Sreedhar, Sandeep Dahiya, Anthony Saunders, and Anand Srinivasan. 2007. "So what do I get? The bank's view of lending relationships." *Journal of financial Economics* 85 (2):368–419.

Bharath, Sreedhar T, Sandeep Dahiya, Anthony Saunders, and Anand Srinivasan. 2011. "Lending relationships and loan contract terms." *The Review of Financial Studies* 24 (4):1141–1203.

Bird, Andrew, Michael Hertzel, Stephen A Karolyi, and Thomas G Ruchti. 2026. "What do covenant violations reveal about lending relationships?" *Journal of Financial Intermediation* :101220.

Bolton, Patrick and David S. Scharfstein. 1990. "A Theory of Predation Based on Agency Problems in Financial Contracting." *American Economic Review* 80 (1):93–106.

———. 1996. "Optimal Debt Structure and the Number of Creditors." *Journal of Political Economy* 104 (1):1–25.

Bond, Philip and Ashok S. Rai. 2008. "Cosigned vs. Group Loans." *Journal of Development Economics* 85 (1–2):58–80.

Boot, Arnoud W.A. 2000. "Relationship Banking: What Do We Know?" *Journal of Financial Intermediation* 9 (1):7–25. URL https://www.sciencedirect.com/science/article/pii/S1042957300902821.

Boot, Arnoud WA and Anjan V Thakor. 2000. "Can relationship banking survive competition?" *The journal of Finance* 55 (2):679–713.

———. 2019. "Commercial banking and shadow banking: The accelerating integration of banks and markets and its implications for regulation." .

Breza, Emily and Arun G Chandrasekhar. 2019. "Social networks, reputation, and commitment: evidence from a savings monitors experiment." *Econometrica* 87 (1):175–216.

Bryan, Gharad, Dean Karlan, and Jonathan Zinman. 2015. "Referrals: Peer Screening and Enforcement in a Consumer Credit Field Experiment." *American Economic Journal: Microeconomics* 7 (3):174–204.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru. 2018. "Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks." *Journal of Financial Economics* 130 (3):453–483.

Chandrasekhar, Arun G., Cynthia Kinnan, and Horacio Larreguy. 2018. "Social Networks as Contract Enforcement: Evidence from a Lab Experiment in the Field." *American Economic Journal: Applied Economics* 10 (4):43–78.

Chetty, Raj. 2009. "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods." *Annual Review of Economics* 1:451–488.

Chioda, Laura, Paul Gertler, Sean Higgins, and Paolina C. Medina. 2024. "FinTech Lending to

Borrowers with No Credit History." NBER Working Paper 33208, National Bureau of Economic Research.

Clementi, Gian Luca and Hugo A. Hopenhayn. 2006. "A Theory of Financing Constraints and Firm Dynamics." *Quarterly Journal of Economics* 121 (1):229–265.

Cole, Rebel A. 1998. "The importance of relationships to the availability of credit." *Journal of Banking & Finance* 22 (6-8):959–977.

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

Dávila, Eduardo and Anton Korinek. 2018. "Pecuniary Externalities in Economies with Financial Frictions." *Review of Economic Studies* 85 (1):352–395.

de Quidt, Jonathan, Thiemo Fetzer, and Maitreesh Ghatak. 2018. "Commercialization and the Decline of Joint Liability Microcredit." *Journal of Development Economics* 134:209–225. [VERIFY] page range from RePEc.

Di Maggio, Marco and Vincent Yao. 2021. "Fintech Borrowers: Lax Screening or Cream-Skimming?" *Review of Financial Studies* 34 (10):4565–4618.

Diamond, Douglas W. 1984. "Financial Intermediation and Delegated Monitoring." *Review of Economic Studies* 51 (3):393–414.

———. 1989. "Reputation Acquisition in Debt Markets." *Journal of Political Economy* 97 (4):828–862.

———. 1991. "Monitoring and Reputation: The Choice between Bank Loans and Directly Placed Debt." *Journal of Political Economy* 99 (4):689–721.

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

Duqi, Andi, Angelo Tomaselli, and Giuseppe Torluccio. 2018. "Is relationship lending still a mixed blessing? A review of advantages and disadvantages for lenders and borrowers." *Journal of Economic Surveys* 32 (5):1446–1482.

Elsas, Ralf and Jan Pieter Krahnen. 1998. "Is relationship lending special? Evidence from credit-file data in Germany." *Journal of Banking & Finance* 22 (10-11):1283–1316.

Elyasiani, Elyas and Lawrence G Goldberg. 2004. "Relationship lending: a survey of the literature." *Journal of Economics and Business* 56 (4):315–330.

Erel, Isil and Jack Liebersohn. 2022. "Can FinTech Reduce Disparities in Access to Finance? Evidence from the Paycheck Protection Program." *Journal of Financial Economics* 146 (1):90–118.

Feigenberg, Benjamin, Erica Field, and Rohini Pande. 2013. "The Economic Returns to Social Interaction: Experimental Evidence from Microfinance." *Review of Economic Studies* 80 (4):1459–1483.

Fiorin, Stefano, Joseph Hall, and Martin Kanz. 2023. *How Do Borrowers Respond to a Debt Moratorium?: Experimental Evidence from Consumer Loans in India*.

Freedman, Seth and Ginger Zhe Jin. 2017. "The Information Value of Online Social Networks: Lessons from Peer-to-Peer Lending." *International Journal of Industrial Organization* 51:185–222.

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

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther. 2022. "Predictably Unequal? The Effects of Machine Learning on Credit Markets." *Journal of Finance* 77 (1):5–47.

Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery. 2019. "The Role of Technology in Mortgage Lending." *Review of Financial Studies* 32 (5):1854–1899.

Gambacorta, Leonardo, Yiping Huang, Zhenhua Li, Han Qiu, and Shu Chen. 2023. "Data versus Collateral." BIS Working Paper 881, Bank for International Settlements. Published version in *Review of Finance*; [VERIFY] final journal details.

Ghatak, Maitreesh and Timothy W. Guinnane. 1999. "The Economics of Lending with Joint Liability: Theory and Practice." *Journal of Development Economics* 60 (1):195–228.

Ghosh, Pulak, Boris Vallee, and Yao Zeng. 2026. "FinTech lending and cashless payments." *The Journal of Finance* 81 (2):1053–1101.

Ghosh, Pulak and Nishant Vats. 2022. "Safety nets, credit, and investment: Evidence from a guaranteed income program." *Credit, and Investment: Evidence from a Guaranteed Income Program (November 1, 2022)* .

Giné, Xavier, Pamela Jakiela, Dean Karlan, and Jonathan Morduch. 2010. "Microfinance games." *American Economic Journal: Applied Economics* 2 (3):60–95.

Giné, Xavier and Dean S. Karlan. 2014. "Group versus Individual Liability: Short and Long Term Evidence from Philippine Microcredit Lending Groups." *Journal of Development Economics* 107:65–83.

Granja, João, Christian Leuz, and Raghuram G Rajan. 2022. "Going the extra mile: Distant lending and credit cycles." *The Journal of Finance* 77 (2):1259–1324.

Heath, Rachel. 2018. "Why Do Firms Hire Using Referrals? Evidence from Bangladeshi Garment Factories." *Journal of Political Economy* 126 (4):1691–1746.

Heitz, Amanda, Christopher Martin, and Alexander Ufier. 2026. "Bank Monitoring with On-Site Inspections." *Journal of Finance* 81 (2):687–737.

Hendren, Nathaniel and Ben Sprung-Keyser. 2020. "A Unified Welfare Analysis of Government Policies." *Quarterly Journal of Economics* 135 (3):1209–1318.

Hildebrand, Thomas, Manju Puri, and Jörg Rocholl. 2017. "Adverse Incentives in Crowdfunding." *Management Science* 63 (3):587–608.

Holmström, Bengt and Jean Tirole. 1997. "Financial Intermediation, Loanable Funds, and the Real Sector." *Quarterly Journal of Economics* 112 (3):663–691.

Hombert, Johan and Adrien Matray. 2017. "The real effects of lending relationships on innovative firms and inventor mobility." *The Review of Financial Studies* 30 (7):2413–2445.

Hussam, Reshmaan, Natalia Rigol, and Benjamin N. Roth. 2022. "Targeting High Ability Entrepreneurs Using Community Information: Mechanism Design in the Field." *American Economic Review* 112 (3):861–898.

Innes, Robert D. 1990. "Limited Liability and Incentive Contracting with Ex-ante Action Choices." *Journal of Economic Theory* 52 (1):45–67.

Ioannidou, Vasso and Steven Ongena. 2010. ""Time for a change": Loan conditions and bank behavior when firms switch banks." *The Journal of Finance* 65 (5):1847–1877.

Iyer, Rajkamal, Asim Ijaz Khwaja, Erzo F. P. Luttmer, and Kelly Shue. 2016. "Screening Peers Softly: Inferring the Quality of Small Borrowers." *Management Science* 62 (6):1554–1577.

Jackson, Matthew O., Tomas Rodriguez-Barraquer, and Xu Tan. 2012. "Social Capital and Social Quilts: Network Patterns of Favor Exchange." *American Economic Review* 102 (5):1857–1897.

James, Christopher. 1987. "Some evidence on the uniqueness of bank loans." *Journal of financial economics* 19 (2):217–235.

Jiménez, Gabriel, Steven Ongena, José-Luis Peydró, and Jesús Saurina. 2014. "Hazardous times for monetary policy: What do twenty-three million bank loans say about the effects of monetary policy on credit risk-taking?" *Econometrica* 82 (2):463–505.

Jimenez, Gabriel, Vicente Salas, and Jesus Saurina. 2006. "Determinants of collateral." *Journal of financial economics* 81 (2):255–281.

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

Karlan, Dean, Beniamino Savonitto, Bram Thuysbaert, and Christopher Udry. 2017. "Impact of savings groups on the lives of the poor." *Proceedings of the National Academy of Sciences* 114 (12):3079–3084.

Khwaja, Asim Ijaz and Atif Mian. 2008. "Tracing the impact of bank liquidity shocks: Evidence from an emerging market." *American Economic Review* 98 (4):1413–1442.

Kleven, Henrik Jacobsen. 2021. "Sufficient Statistics Revisited." *Annual Review of Economics* 13:515–538.

Kundu, Shohini and Nishant Vats. 2024. "The Geography of Corporate Investment: Role of History Beyond Institutions." *Available at SSRN* .

Kysucky, Vlado and Lars Norden. 2016. "The benefits of relationship lending in a cross-country context: A meta-analysis." *Management Science* 62 (1):90–110.

Liberti, José María and Mitchell A. Petersen. 2019. "Information: Hard and Soft." *Review of Corporate Finance Studies* 8 (1):1–41.

Lin, Mingfeng, Nagpurnanand R. Prabhala, and Siva Viswanathan. 2013. "Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-Peer Lending." *Management Science* 59 (1):17–35.

List, John A. 2020. "Non est disputandum de generalizability? A glimpse into the external validity trial." Tech. rep., National Bureau of Economic Research.

Maitra, Pushkar, Sandip Mitra, Dilip Mookherjee, Alberto Motta, and Sujata Visaria. 2017. "Financing Smallholder Agriculture: An Experiment with Agent-Intermediated Microloans in India." *Journal of Development Economics* 127:306–337.

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

Oster, Emily. 2019. "Unobservable selection and coefficient stability: Theory and evidence." *Journal of Business & Economic Statistics* 37 (2):187–204.

Padilla, A. Jorge and Marco Pagano. 1997. "Endogenous Communication among Lenders and Entrepreneurial Incentives." *Review of Financial Studies* 10 (1):205–236.

Paravisini, Daniel, Veronica Rappoport, and Philipp Schnabl. 2023. "Specialization in bank lending: Evidence from exporting firms." *The Journal of Finance* 78 (4):2049–2085.

Petersen, Mitchell A. and Raghuram G. Rajan. 1994. "The Benefits of Lending Relationships: Evidence from Small Business Data." *Journal of Finance* 49 (1):3–37.

Rai, Ashok S. and Tomas Sjöström. 2004. "Is Grameen Lending Efficient? Repayment Incentives and Insurance in Village Economies." *Review of Economic Studies* 71 (1):217–234.

Rajan, Raghuram G. 1992. "Insiders and Outsiders: The Choice between Informed and Arm's-Length Debt." *Journal of Finance* 47 (4):1367–1400.

Ramakrishnan, Ram T. S. and Anjan V. Thakor. 1984. "Information Reliability and a Theory of Financial Intermediation." *Review of Economic Studies* 51 (3):415–432.

Saez, Emmanuel. 2001. "Using Elasticities to Derive Optimal Income Tax Rates." *Review of Economic Studies* 68 (1):205–229.

Sappington, David. 1983. "Limited Liability Contracts between Principal and Agent." *Journal of Economic Theory* 29 (1):1–21.

Schenone, Carola. 2010. "Lending relationships and information rents: Do banks exploit their information advantages?" *The Review of Financial Studies* 23 (3):1149–1199.

Stein, Jeremy C. 2002. "Information Production and Capital Allocation: Decentralized versus Hierarchical Firms." *Journal of Finance* 57 (5):1891–1921.

Stiglitz, Joseph E. 1990. "Peer Monitoring and Credit Markets." *World Bank Economic Review* 4 (3):351–366.

Stiglitz, Joseph E. and Andrew Weiss. 1981. "Credit Rationing in Markets with Imperfect Information." *American Economic Review* 71 (3):393–410.

Tedeschi, Gwendolyn Alexander. 2006. "Here Today, Gone Tomorrow: Can Dynamic Incentives Make Microfinance More Flexible?" *Journal of Development Economics* 80 (1):84–105.

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

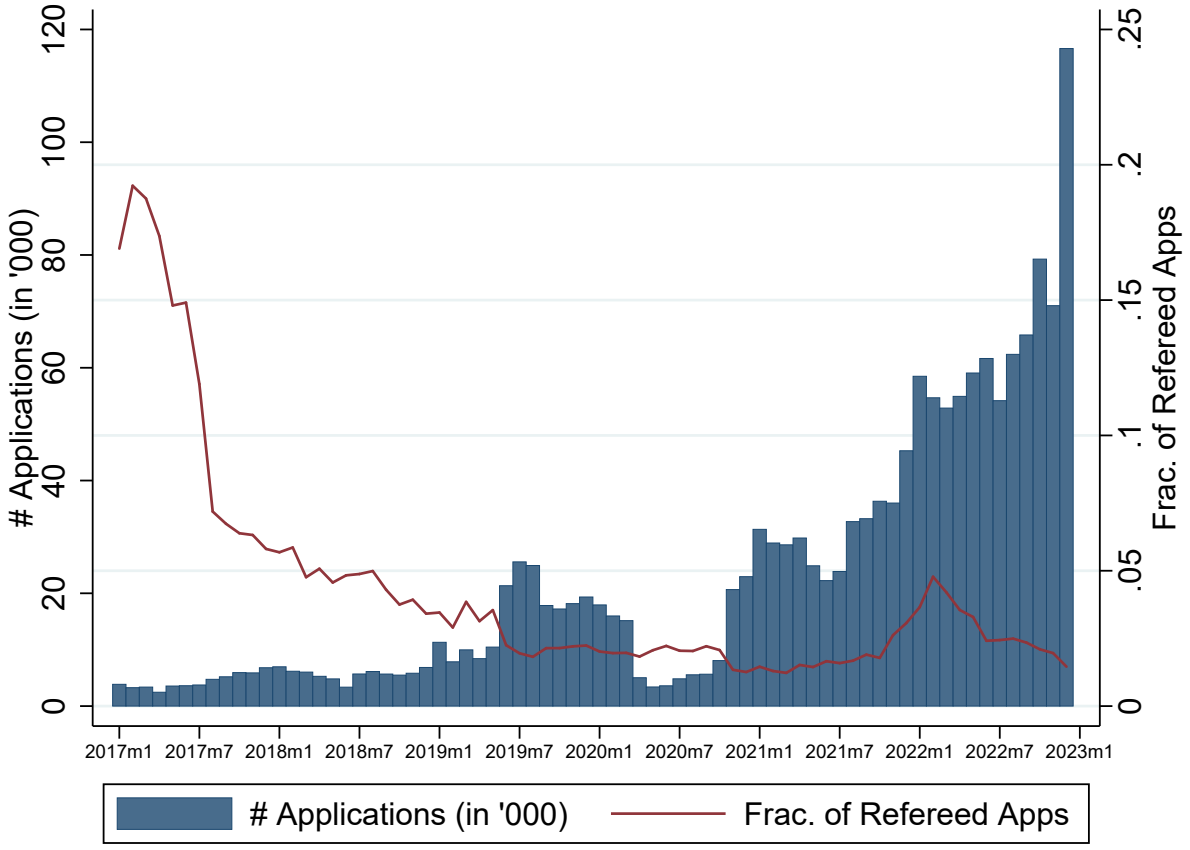
Tomy, Rimmy E. and Regina Wittenberg-Moerman. 2024. "A Review of the Effect of Financial Reporting on Debt Contracting and Related Topics." *Journal of Accounting and Economics* 78 (1). [VERIFY] title and page range.

Vallée, Boris and Yao Zeng. 2019. "Marketplace Lending: A New Banking Paradigm?" *Review of Financial Studies* 32 (5):1939–1982.

Vercammen, James A. 1995. "Credit Bureau Policy and Sustainable Reputation Effects in Credit Markets." *Economica* 62 (248):461–478.

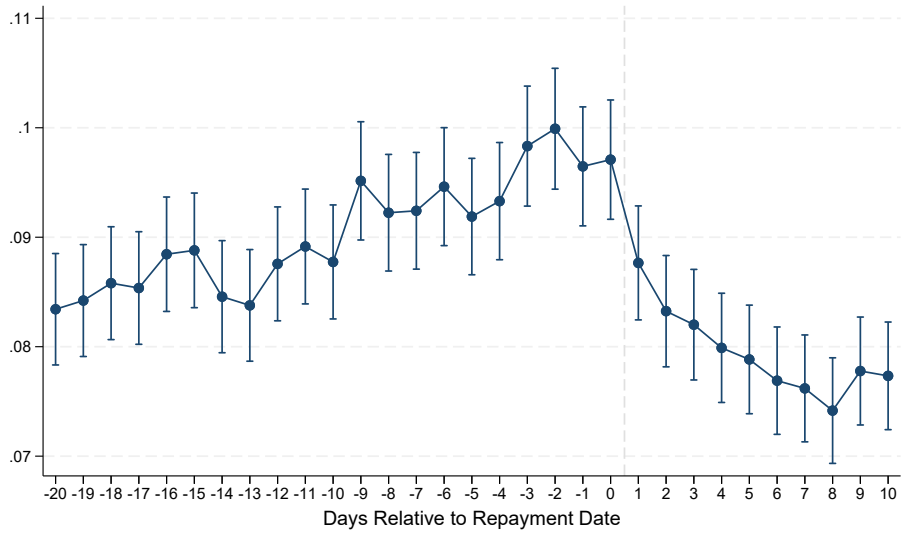
Wager, Stefan and Susan Athey. 2018. "Estimation and inference of heterogeneous treatment effects using random forests." *Journal of the American Statistical Association* 113 (523):1228–1242.

Figure 1: Aggregate Trend in Applications & Referrals

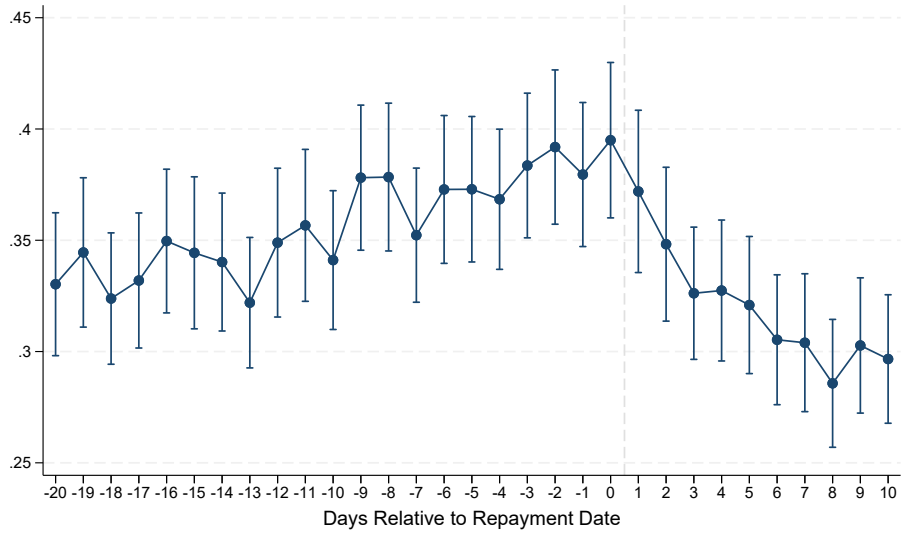


The figure plots the total number of applications received by the platform from January of 2017 until December of 2022. The number of applications is shown in blue bars and reported on the left Y-axis in '000 ₹. The solid red line presents the fraction of applications received in that month that were referred by the referrers with corresponding numbers reported on the right Y-axis.

Figure 2: Descriptive Evidence: Referrer & Referred Interactions around Repayment Dates



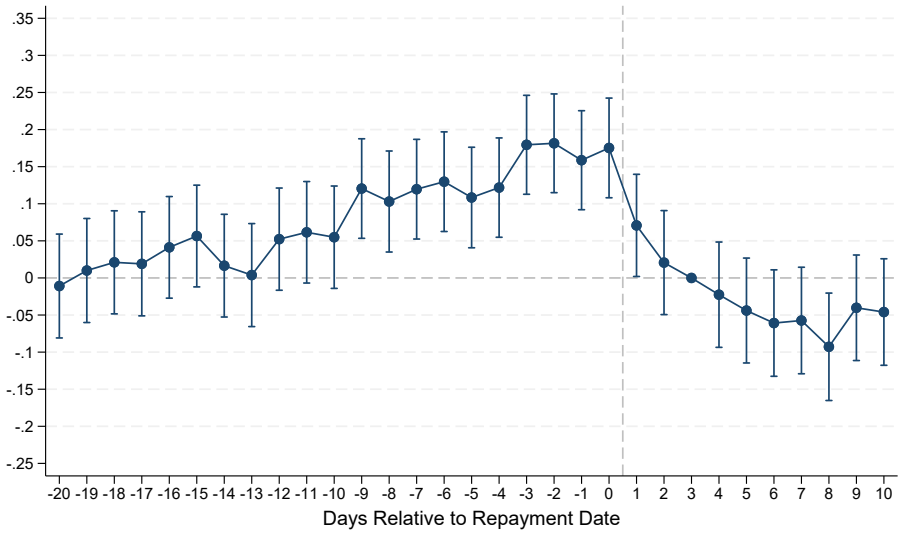
(a) Extensive Margin: Probability of a Phone Call between Referrer & Referred



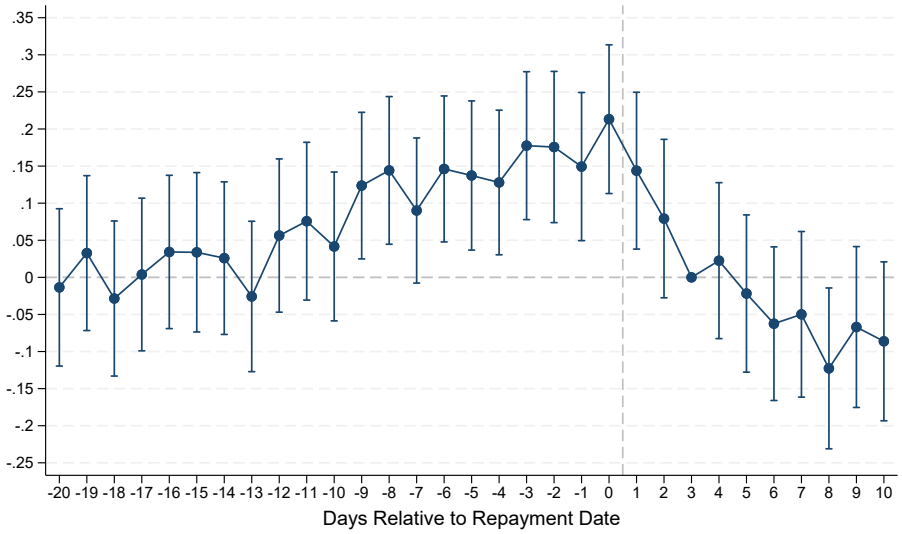
(b) Intensive Margin: Number of Phone Calls between Referrer & Referred

This figure illustrates the intensity of the interactions between the referrer and the referred around repayment dates using call logs data. Specifically, this figure plots the simple average likelihood of a phone call and the number of phone calls around the repayment date. We restrict the sample to customers who were referred and examine the daily phone call activity between the referrer and the referred around the date of repayment. $t = 0$ refers to the date when the loan installment or repayment is due. Figure 2a plots the average likelihood of a phone call between the referrer and the referred that day. Figure 2b plots the average number of phone calls between the referrer and the referred on that day. Both figures plot the point estimates along with the 95% confidence intervals.

Figure 3: Regression Evidence: Referrer & Referred Interactions around Repayment Dates



(a) Extensive Margin: Probability of a Phone Call between Referrer & Referred



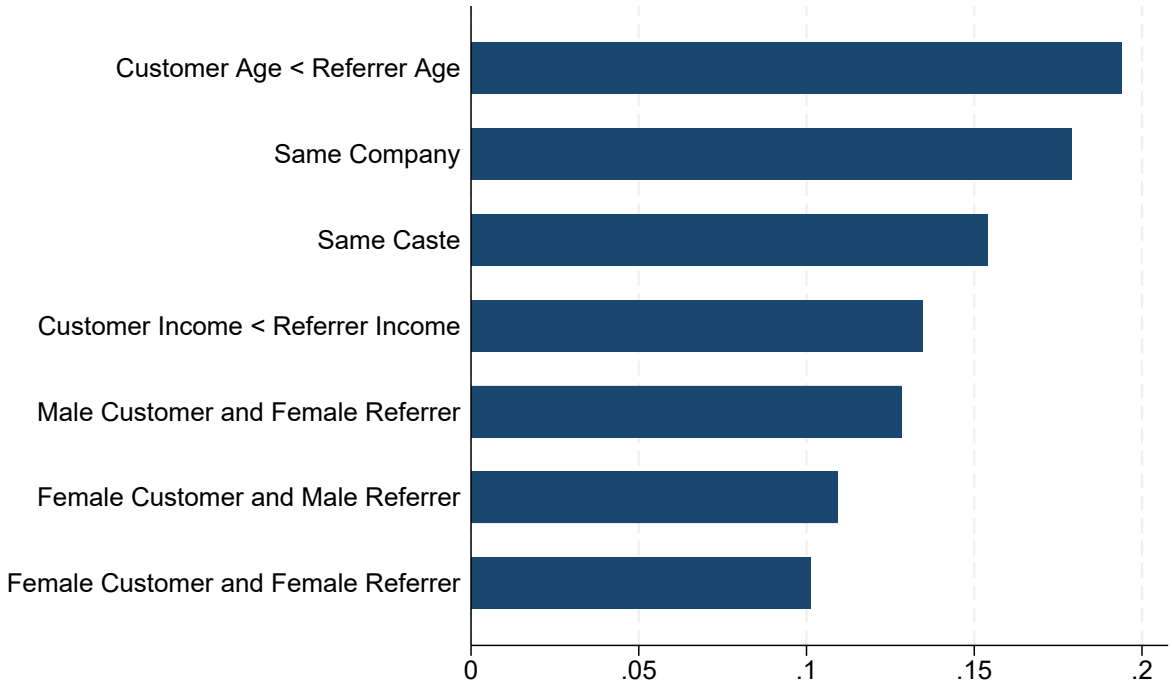
(b) Intensive Margin: Number of Phone Calls between Referrer & Referred

This figure illustrates the intensity of the interactions between the referrer and the referred around repayment dates using call logs data. Specifically, this figure plots the regression estimates from the following specification estimated using Poisson pseudo-likelihood regression with customer and ZIP X month fixed effects:

$$y_{i,t} = \sum_{k=-20, k\neq 3}^{k=10} \beta_k \times \text{Days until Repayment}_k + \theta_i + \theta_{z,m} + \theta_t + \varepsilon_{i,t}$$

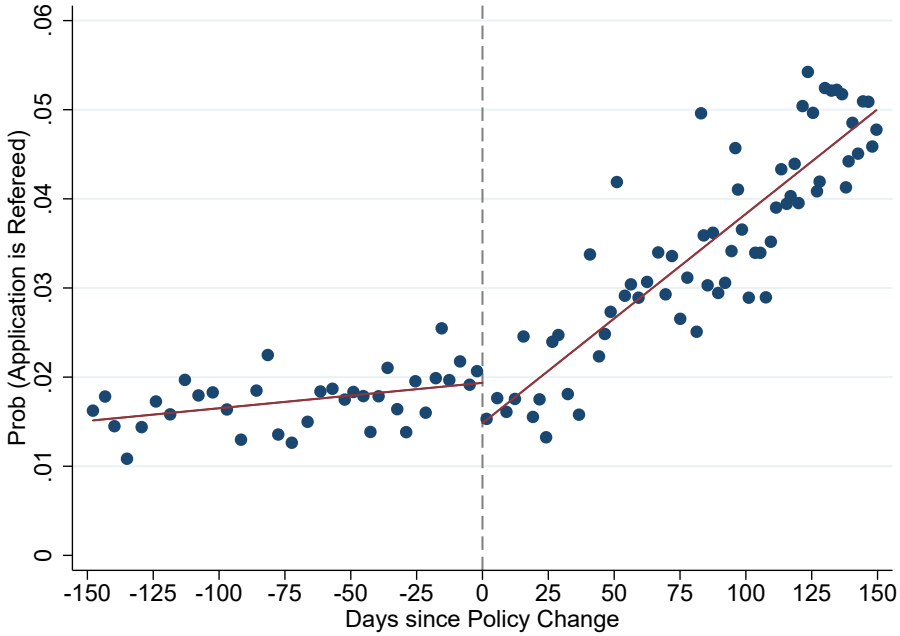
where θ_i denotes referred customer fixed effects and $\theta_{z,m}$ denotes Zip $\times$ month fixed effect, which month refers to the time (month-year) of the origination of the loan for which the customer i was referred. We restrict the sample to customers who were referred and examine the daily phone call activity between the referrer and the referred around the date of repayment. $t = 0$ refers to the date when the loan installment or repayment is due. Figure 3a plots the regression coefficients for β_k using a binary variable which takes a value of one if there was a phone call between the referrer and the referred that day and zero otherwise as the key dependent variable ($y_{i,t}$). θ_t denotes date fixed effect for the date on which $y_{i,t}$ is observed. Figure 3b plots the regression coefficients for β_k using the number of phone calls between the referrer and the referred on day t as the key dependent variable ($y_{i,t}$). Both figures plot the point estimates along with the 95% confidence intervals estimated using robust standard errors.

Figure 4: Feature Importance of Customer and Referrer Characteristics for Default



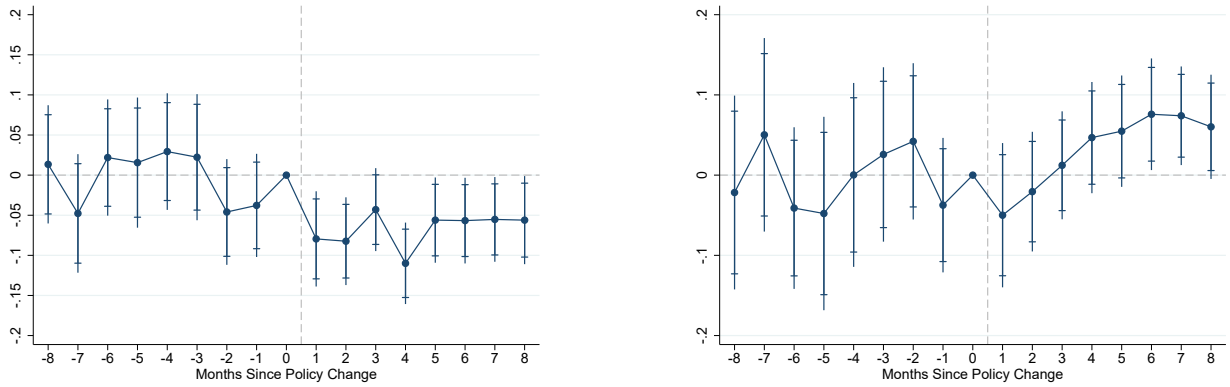
This figure plots the average feature importance from 11 linear and machine learning models. The model we consider include linear regression, logistic regression, decision tree, random forest, extra trees, gradient boosting, light gradient boosting machine (LightGBM), support vector machine (SVM), multilayer perception (MLP), TabNet, and wide & deep net. The feature importance for each of these model is reported in Panel A of Appendix Table A.1. We consider several features that define the realtionship between the referrer and the referred. *Customer Age < Referrer Age* is an indicator variable equal to one if the referred customer is younger than the referrer, and zero otherwise. *Same Company* is an indicator variable equal to one if the referred customer works in the same firm as the referrer, and zero otherwise. *Same Caste* is an indicator variable equal to one if a referred customer belongs to the same caste (social group) as the referrer, and zero otherwise. *Customer Income < Referrer Income* is an indicator variable equal to one if the referred customer has a lower income than the referrer, and zero otherwise. *Male Customer and Female Referrer* is an indicator variable equal to one if the referred customer is male and the referrer is female, and zero otherwise. *Female Customer and Male Referrer* is an indicator variable equal to one if the referred customer is female and the referrer is male, and zero otherwise. *Female Customer and Female Referrer* is an indicator variable equal to one if both the referred customer and the referrer are female, and zero otherwise. The omitted category is *Male Customer and Male Referrer*, where both the referred customer and the referrer are male. This figure plots the average of the feature importance associated with each feature from all these models. The ranking of the features is based on the average feature importance across these models.

Figure 5: Effect of Change in Monetary Bonus on Referral Activity



The figure displays a regression discontinuity plot illustrating the probability that an application is referred, before and after the change in the referral bonus policy. The zero point corresponds to October 2nd, 2021, the date when the bonus per successful referral was increased from ₹1,250 (approximately $15) to ₹3,500 (approximately $42). Days to the right of zero represent the period after the policy change, while days to the left represent the period before. The plot covers a 150-day window on each side of the policy change. The dots depict bin-level averages of the referral likelihood based on a binscatter plot, and the solid maroon line indicates the best-fitting trends before and after the policy change.

Figure 6: Assessment of Pre-Trends: Referral Bonus & Credit Market Outcomes

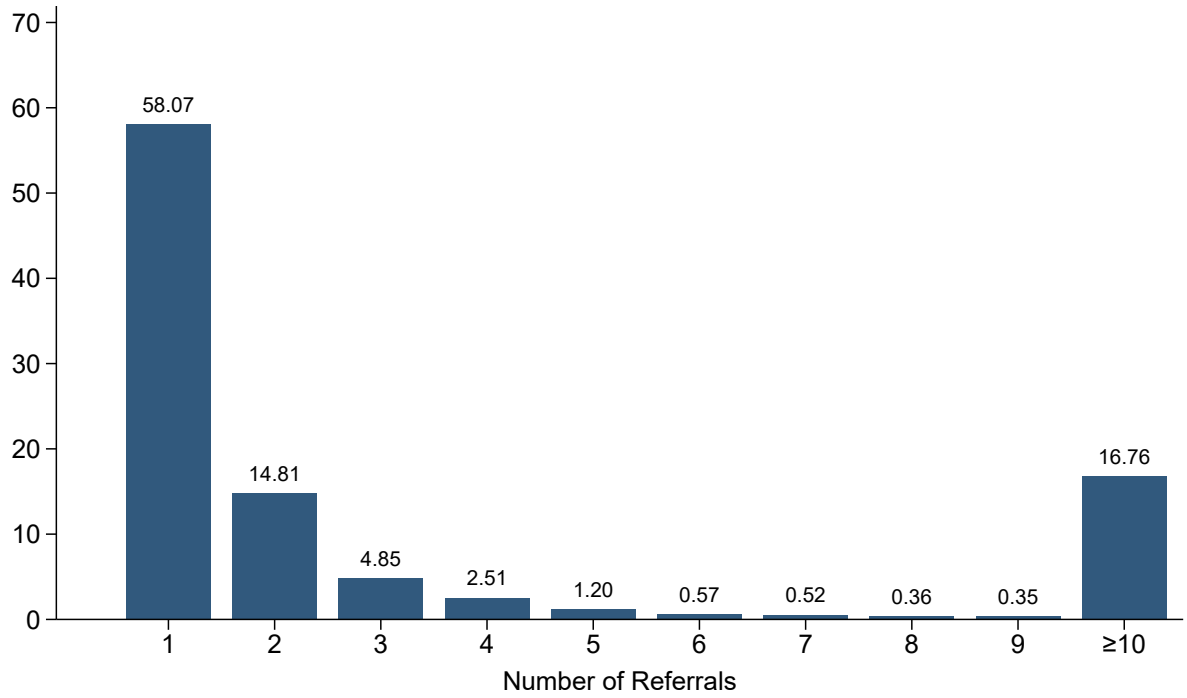


(a) Probability of Acceptance

(b) Probability of Default

This figure illustrates the parallel trend of acceptance rate and default rate around the policy change. Panel (a) plots the probability of acceptance, with the 90% and 95% confidence intervals, by month relative to the referral-bonus policy change; Panel (b) plots the probability of default, with the 90% and 95% confidence intervals, by month relative to the referral-bonus policy change. Eight months before and eight months after the policy change are displayed on the X-axis, where month 0 is the date of the policy change.

Figure 7: Distribution of Referred Applications by the Number of Unique Referrals per Referrer



This figure plots the proportion of applications referred by referrers, by total number of unique referrals. Bars show the fraction of applications referred by referrers with 1–9 referrals, and 10 or more referrals. For instance the bar labeled as 1 on the X-axis shows the fraction of referred applications that were referred by referrers that made exactly one referral throughout our sample.

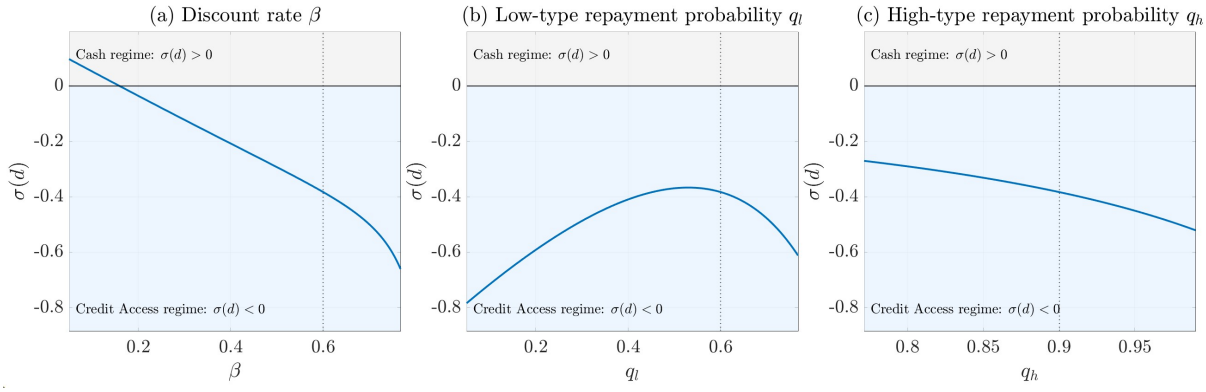


Figure 8: Comparative statics of the cost-minimizing referral contract. The figure plots $\sigma(d)$ against the referrer's discount factor β, the low-type repayment probability q_l, and the high-type repayment probability q_h. The shaded gray region with $\sigma(d) \geq 0$ corresponds to the cash regime, and the shaded blue region with $\sigma(d) < 0$ corresponds to the credit access regime. The parameter values are $\beta = 0.6, R = 1.3, q_l = 0.6, q_h = 0.9, q^* = 0.2$, unless stated otherwise.

Table 1: Summary Statistics

	# Obs	p25	p50	p75	Mean	SD
Panel A: Demographic Characteristics						
Female	1,216,524	0.0000	0.0000	0.0000	0.1235	0.3290
Age (in years)	1,216,524	24.8575	28.6658	33.6904	29.9494	6.8859
Completed High School	1,216,524	0.0000	0.0000	1.0000	0.2883	0.4530
Completed College/Undergraduate	1,216,524	0.0000	1.0000	1.0000	0.5760	0.4942
Completed Masters	1,216,524	0.0000	0.0000	0.0000	0.1357	0.3424
Panel B: Economic Characteristics						
Monthly Income (₹)	1,216,524	18,000.00	25,000.00	35,000.00	31,035.04	22,052.48
Has a Credit Score	1,216,524	0.0000	0.0000	1.0000	0.3428	0.4747
Credit Score	417,071	642.0000	694.0000	724.0000	675.9431	67.7755
Panel C: Application Characteristics						
Loan Amount (₹)	1,216,524	9,000.00	20,000.00	46,000.00	35,703.93	45,924.02
Debt-to-Income	1,216,524	0.3889	0.8434	1.7500	1.0965	0.9155
Maturity (in months)	1,216,524	3.0000	9.0000	13.0000	8.0320	4.8173
Prob(Acceptance)	1,216,524	0.0000	0.0000	1.0000	0.2604	0.4389
Prob(Referred)	1,216,524	0.0000	0.0000	0.0000	0.0285	0.1664
Prob(Default)	316,832	0.0000	0.0000	1.0000	0.2609	0.4391

This table presents summary statistics for the key variables in the application-level data. The sample comprises 1,216,524 loan applications from first-time customers to the fintech lender between January 2017 and December 2022. Panel A reports statistics for key demographic variables: gender (female), age in years, and indicator variables for whether the applicant has completed high school, a college undergraduate degree, or a master's degree. Panel B reports statistics on applicant-level economic characteristics: monthly income, presence of a credit score, and, conditional on having a score, the score's value. Panel C reports statistics for application-level characteristics: loan amount requested, debt-to-income ratio, loan maturity in months, and indicator variables for whether the application was approved, referred, and, if approved, whether the borrower defaulted.

Table 2: Comparison of Referred and Non-Referred Applicants

	(1)	(2)	(3)	(4)	(5)	(6)
	Mean (St Dev)		Diff. in Mean	p-value for	Diff. Relative	Diff. Relative
	Non Referred	Referred	(St Error)	difference	to the Mean	to St Dev
Female	0.1230	0.1413	-0.0183	0.3804	-0.148	0.055
	(0.3284)	(0.3483)	(0.0018)			
Age (in years)	29.9682	29.3088	0.6595	0.0928	0.022	0.096
	(6.8907)	(6.6883)	(0.0375)			
Completed High School	0.2893	0.2540	0.0354	0.1658	0.123	0.078
	(0.4534)	(0.4353)	(0.0025)			
Completed College	0.5752	0.6027	-0.0275	0.3336	-0.048	0.056
	(0.4943)	(0.4893)	(0.0027)			
Completed Masters	0.1354	0.1433	-0.0078	0.7068	-0.058	0.023
	(0.3422)	(0.3504)	(0.0019)			
Monthly Income (₹)	31,023.09	31,442.40	-419.31	0.7684	-0.014	0.019
	(22,047.77)	(22,209.04)	(120.17)			
Has a Credit Score	0.3404	0.4261	-0.0857	0.0040	-0.250	0.181
	(0.4738)	(0.4945)	(0.0026)			
Credit Score	676.06	672.71	3.3504	0.4310	0.005	0.049
	(67.6964)	(69.8213)	(0.5678)			
Loan Amount (₹)	35,823.34	31,632.68	4,190.66	0.0764	0.117	0.091
	(46,109.57)	(38,856.12)	(250.22)			
Debt-to-Income	1.0991	1.0073	0.0918	0.0756	0.084	0.100
	(0.9175)	(0.8401)	(0.0050)			
Maturity (in months)	8.0518	7.3579	0.6938	0.0140	0.086	0.144
	(4.8141)	(4.8774)	(0.0262)			

This table presents summary statistics for key variables in the application-level data, comparing referred and non-referred applicants. Column 1 presents the overall sample characteristics. Columns 2 and 3 present summary statistics for non-referred and referred applicants, respectively. Column 4 presents the difference in means between Columns 2 and 3. Column 5 presents the p-values from a t-test assessing the null hypothesis that the difference between the two groups is zero. Finally, Columns 6 and 7 express the difference presented in Column 4 relative to the sample mean and sample standard deviation, respectively. The table reports statistics for: (1) key demographic variables (gender (female), age in years, and indicator variables for educational attainment (high school, college degree, or master's degree)); (2) applicant-level economic characteristics (monthly income, presence of a credit score, and credit score value, conditional on having a score); and (3) application-level characteristics (loan amount requested, debt-to-income ratio, and loan maturity in months).

Table 3: Referral Program & Application Acceptance

Dep Var: *Acceptance*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Referred	0.0539***	0.0481***	0.0337***	0.0340***	0.0325***	0.0323***	0.0321***	0.0304***	0.0303***	0.0304***	0.0306***
	(0.0023)	(0.0025)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0015)	(0.0015)	(0.0015)
Month FE	Yes										
Month × Zipcode FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Credit Score Decile FE			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Decile FE				Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age Decile FE					Yes	Yes	Yes	Yes	Yes	Yes	Yes
Gender FE						Yes	Yes	Yes	Yes	Yes	Yes
Education FE							Yes	Yes	Yes	Yes	Yes
Loan Purpose FE								Yes	Yes	Yes	Yes
Product Type FE									Yes	Yes	Yes
Loan Amount Decile FE										Yes	Yes
DTI Decile FE											Yes
R^2	0.1928	0.3372	0.7697	0.7699	0.7707	0.7708	0.7710	0.7762	0.8006	0.8016	0.8021
Adjusted R^2	0.1928	0.2030	0.7231	0.7233	0.7243	0.7243	0.7246	0.7309	0.7602	0.7614	0.7620

This table reports the relationship between loan acceptance and referrals. *Acceptance* is an indicator variable equal to one if the loan application is approved, and zero otherwise. *Referred* is an indicator variable equal to one if a new customer is referred by an existing customer with the FinTech lender and zero otherwise. Month captures the application date, recorded in a month-year format. The zip code refers to the residential zip code reported by the applicant in their application. The credit score represents the applicant's Equifax credit score, with applicants grouped into deciles based on their scores; a separate category is included for new-to-credit applicants who do not have a credit score. Income decile and age decile are determined using verified income and age data, respectively. The variable gender is coded as one for female applicants and zero otherwise. The education fixed effect accounts for the highest level of education reported by the applicant. The loan purpose indicates the reason for the loan as specified by the applicant, while product type denotes the specific type of loan applied for. The loan amount decile is based on the requested amount, categorized into deciles, and the debt-to-income (DTI) ratio is constructed by dividing the requested loan amount by the applicant's verified monthly income. All regressions include an equal number of observations between 1.2-1.3 million. Standard errors, in parentheses, are clustered by zip code. $*p$ <.1; $**p$ <.05; $***p$ <.01.

Table 4: Referral Program & Loan Default

Dep Var: *Default*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Referred	-0.0588***	-0.0586***	-0.0602***	-0.0622***	-0.0621***	-0.0619***	-0.0619***	-0.0622***	-0.0608***	-0.0596***	-0.0609***
	(0.0042)	(0.0048)	(0.0048)	(0.0048)	(0.0048)	(0.0048)	(0.0048)	(0.0048)	(0.0046)	(0.0046)	(0.0046)
Month FE	Yes										
Month × Zipcode FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Credit Score Decile FE			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Decile FE				Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age Decile FE					Yes	Yes	Yes	Yes	Yes	Yes	Yes
Gender FE						Yes	Yes	Yes	Yes	Yes	Yes
Education FE							Yes	Yes	Yes	Yes	Yes
Loan Purpose FE								Yes	Yes	Yes	Yes
Product Type FE									Yes	Yes	Yes
Loan Amount Decile FE										Yes	Yes
DTI Decile FE											Yes
R^2	0.0495	0.2939	0.3027	0.3110	0.3111	0.3112	0.3115	0.3137	0.3261	0.3280	0.3302
Adjusted R^2	0.0492	0.0632	0.0748	0.0859	0.0860	0.0861	0.0864	0.0894	0.1058	0.1082	0.1111

This table reports the relationship between application default and referrals. *Default* is an indicator variable equal to one if a customer's loan is past-due for more than 90 days, and zero otherwise. *Referred* is an indicator variable equal to one if a new customer is referred by an existing customer with the FinTech lender and zero otherwise. Month captures the application date, recorded in a month-year format. The zip code refers to the residential zip code reported by the applicant in their application. The credit score represents the applicant's Equifax credit score, with applicants grouped into deciles based on their scores; a separate category is included for new-to-credit applicants who do not have a credit score. Income decile and age decile are determined using verified income and age data, respectively. The variable gender is coded as one for female applicants and zero otherwise. The education fixed effect accounts for the highest level of education reported by the applicant. The loan purpose indicates the reason for the loan as specified by the applicant, while product type denotes the specific type of loan applied for. The loan amount decile is based on the requested amount, categorized into deciles, and the debt-to-income (DTI) ratio is constructed by dividing the requested loan amount by the applicant's verified monthly income. All regressions include an equal number of observations between 255k-265K. Standard errors, in parentheses, are clustered by zip code. $*p$ <.1; $**p$ <.05; $***p$ <.01.

Table 5: Role of Selection & Disciplining in Referral Networks in Credit Markets

Panel A: Application Acceptance				
	(1)	(2)	(3)	(4)
FinTech × *Referred*	-0.0021	-0.0047	-0.0023	-0.0037
	(0.0030)	(0.0038)	(0.0034)	(0.0042)
Customer × Year FE	Yes	Yes	Yes	Yes
Lender × Year FE	Yes	Yes		
Month × Zipcode FE		Yes		Yes
Lender × Year × Zipcode FE			Yes	Yes
R^2	0.5455	0.6169	0.5571	0.6273
Adjusted R^2	0.4405	0.4317	0.4196	0.4036
Panel B: Default				
	(1)	(2)	(3)	(4)
FinTech × *Referred*	-0.0292***	-0.0366***	-0.0249**	-0.0373***
	(0.0105)	(0.0114)	(0.0114)	(0.0123)
Customer × Year FE	Yes	Yes	Yes	Yes
Lender × Year FE	Yes	Yes		
Month × Zipcode FE		Yes		Yes
Lender × Year × Zipcode FE			Yes	Yes
R^2	0.3698	0.4704	0.4017	0.4954
Adjusted R^2	0.2334	0.2235	0.2268	0.2038

This table reports the loan acceptance and default for referrals on their loans with the FinTech lender – which is the primary source of our data – and other lenders. *Acceptance* is an indicator variable equal to one if an applicant's loan application is accepted by the lender, and zero otherwise. *Default* is an indicator variable equal to one if a customer's loan is past-due for more than 90 days, and zero otherwise. *FinTech* is an indicator variable equal to one if a loan application is to the FinTech lender– which is the primary source of our data – and zero for other lenders. *Referred* is an indicator variable equal to one if the customer is referred by an existing customer with the FinTech lender and zero otherwise. The sample includes 2.15-2.17 million observations. Standard errors, in parentheses, are clustered by zipcode. *p <.1; **p <.05; ***p <.01.

Table 6: Effect of Referral Bonus on Application Acceptance

Dep Var: *Acceptance*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Referred × *Post*	-0.0419***	-0.0359***	-0.0246***	-0.0237***	-0.0242***	-0.0241***	-0.0241***	-0.0245***	-0.0206***	-0.0204***	-0.0206***
	(0.0047)	(0.0051)	(0.0038)	(0.0038)	(0.0038)	(0.0038)	(0.0038)	(0.0038)	(0.0035)	(0.0035)	(0.0035)
Referred	0.0808***	0.0720***	0.0494***	0.0491***	0.0479***	0.0476***	0.0474***	0.0458***	0.0433***	0.0433***	0.0429***
	(0.0038)	(0.0041)	(0.0035)	(0.0035)	(0.0035)	(0.0035)	(0.0035)	(0.0034)	(0.0033)	(0.0032)	(0.0033)
Month FE	Yes										
Month × Zipcode FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Credit Score Decile FE			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Decile FE				Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age Decile FE					Yes	Yes	Yes	Yes	Yes	Yes	Yes
Gender FE						Yes	Yes	Yes	Yes	Yes	Yes
Education FE							Yes	Yes	Yes	Yes	Yes
Loan Purpose FE								Yes	Yes	Yes	Yes
Product Type FE									Yes	Yes	Yes
Loan Amount Decile FE										Yes	Yes
DTI Decile FE											Yes
R^2	0.1877	0.3336	0.7713	0.7716	0.7724	0.7724	0.7726	0.7778	0.8019	0.8031	0.8065
Adjusted R^2	0.1877	0.1976	0.7247	0.7250	0.7259	0.7260	0.7263	0.7324	0.7615	0.7629	0.7670

This table reports the relationship between loan acceptance and referrals. *Acceptance* is an indicator variable equal to one if the loan application is approved, and zero otherwise. *Referred* is an indicator variable equal to one if a new customer is referred by an existing customer with the FinTech lender and zero otherwise. *Post* is an indicator variable equal to one if a loan application is submitted in or after October 2021, and zero otherwise. Month captures the application date, recorded in a month-year format. The zip code refers to the residential zip code reported by the applicant in their application. The credit score represents the applicant's Equifax credit score, with applicants grouped into deciles based on their scores; a separate category is included for new-to-credit applicants who do not have a credit score. Income decile and age decile are determined using verified income and age data, respectively. The variable gender is coded as one for female applicants and zero otherwise. The education fixed effect accounts for the highest level of education reported by the applicant. The loan purpose indicates the reason for the loan as specified by the applicant, while product type denotes the specific type of loan applied for. The loan amount decile is based on the requested amount, categorized into deciles, and the debt-to-income (DTI) ratio is constructed by dividing the requested loan amount by the applicant's verified monthly income. All regressions include an equal number of observations between 1.1-1.2 million. Standard errors, in parentheses, are clustered by zip code. $*p$ <.1; $**p$ <.05; $***p$ <.01.

Table 7: Effect of Referral Bonus on Loan Default

Dep Var: *Default*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Referred × *Post*	0.0569***	0.0609***	0.0511***	0.0473***	0.0472***	0.0471***	0.0471***	0.0461***	0.0419***	0.0416***	0.0423***
	(0.0075)	(0.0092)	(0.0090)	(0.0090)	(0.0090)	(0.0090)	(0.0090)	(0.0089)	(0.0089)	(0.0088)	(0.0088)
Referred	-0.0821***	-0.0850***	-0.0824***	-0.0827***	-0.0826***	-0.0823***	-0.0823***	-0.0822***	-0.0792***	-0.0778***	-0.0794***
	(0.0055)	(0.0068)	(0.0066)	(0.0067)	(0.0067)	(0.0067)	(0.0067)	(0.0066)	(0.0064)	(0.0063)	(0.0063)
Month FE	Yes										
Month × Zipcode FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Credit Score Decile FE			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Decile FE				Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age Decile FE					Yes	Yes	Yes	Yes	Yes	Yes	Yes
Gender FE						Yes	Yes	Yes	Yes	Yes	Yes
Education FE							Yes	Yes	Yes	Yes	Yes
Loan Purpose FE								Yes	Yes	Yes	Yes
Product Type FE									Yes	Yes	Yes
Loan Amount Decile FE										Yes	Yes
DTI Decile FE											Yes
R^2	0.0490	0.2952	0.3040	0.3126	0.3126	0.3128	0.3130	0.3152	0.3274	0.3294	0.3331
Adjusted R^2	0.0488	0.0640	0.0756	0.0870	0.0870	0.0872	0.0874	0.0904	0.1066	0.1091	0.1139

This table reports the relationship between application default and referrals. *Default* is an indicator variable equal to one if a customer's loan is past-due for more than 90 days, and zero otherwise. *Referred* is an indicator variable equal to one if a new customer is referred by an existing customer with the FinTech lender and zero otherwise. *Post* is an indicator variable equal to one if a loan application is submitted in or after October 2021, and zero otherwise. Month captures the application date, recorded in a month-year format. The zip code refers to the residential zip code reported by the applicant in their application. The credit score represents the applicant's Equifax credit score, with applicants grouped into deciles based on their scores; a separate category is included for new-to-credit applicants who do not have a credit score. Income decile and age decile are determined using verified income and age data, respectively. The variable gender is coded as one for female applicants and zero otherwise. The education fixed effect accounts for the highest level of education reported by the applicant. The loan purpose indicates the reason for the loan as specified by the applicant, while product type denotes the specific type of loan applied for. The loan amount decile is based on the requested amount, categorized into deciles, and the debt-to-income (DTI) ratio is constructed by dividing the requested loan amount by the applicant's verified monthly income. All regressions include an equal number of observations between 255k-265K. Standard errors, in parentheses, are clustered by zip code. *p <.1; **p <.05; ***p <.01.

Table 8: Effect of Referral Bonus: Role of Selection & Disciplining in Referral Networks

Panel A: Application Acceptance				
	(1)	(2)	(3)	(4)
FinTech × *Referred* × *Post*	-0.0073	0.0025	-0.0076	0.0012
	(0.0074)	(0.0088)	(0.0083)	(0.0097)
FinTech × *Referred*	0.0017	-0.006	0.0017	-0.0043
	(0.0067)	(0.0077)	(0.0074)	(0.0085)
FinTech × *Post*	0.0026	-0.0240***	0.0044	-0.0248***
	(0.0030)	(0.0037)	(0.0033)	(0.0039)
Referred × *Post*	0.0285***	-0.0052	0.0286***	-0.0055
	(0.0033)	(0.0046)	(0.0034)	(0.0047)
Customer × Year FE	Yes	Yes	Yes	Yes
Lender × Year FE	Yes	Yes		
Month × Zipcode FE		Yes		Yes
Lender × Year × Zipcode FE			Yes	Yes
R^2	0.5455	0.6169	0.5571	0.6273
Adjusted R^2	0.4405	0.4318	0.4196	0.4036

Panel B: Default				
	(1)	(2)	(3)	(4)
FinTech × *Referred* × *Post*	0.0649***	0.0485**	0.0676***	0.0451*
	(0.0215)	(0.0226)	(0.0239)	(0.0248)
FinTech × *Referred*	-0.0712***	-0.0657***	-0.0687***	-0.0648***
	(0.0184)	(0.0190)	(0.0208)	(0.0212)
FinTech × *Post*	0.0073	-0.0021	0.0065	-0.0024
	(0.0098)	(0.0101)	(0.0106)	(0.0109)
Referred × *Post*	0.0254***	0.0059	0.0256***	0.0061
	(0.0051)	(0.0050)	(0.0051)	(0.0051)
Customer × Year FE	Yes	Yes	Yes	Yes
Lender × Year FE	Yes	Yes		
Month × Zipcode FE		Yes		Yes
Lender × Year × Zipcode FE			Yes	Yes
R^2	0.3698	0.4704	0.4017	0.4954
Adjusted R^2	0.2334	0.2235	0.2268	0.2038

This table reports the loan acceptance and default for referrals on their loans with the FinTech lender – which is the primary source of our data – and other lenders. *Acceptance* is an indicator variable equal to one if an applicant's loan application is accepted by the lender, and zero otherwise. *Default* is an indicator variable equal to one if a customer's loan is past-due for more than 90 days, and zero otherwise. *FinTech* is an indicator variable equal to one if a loan application is to the FinTech lender and zero for other lenders. *Referred* is an indicator variable equal to one if the customer is referred by an existing customer with the FinTech lender and zero otherwise. *Post* is an indicator variable equal to one if a loan application is submitted in or after October 2021, and zero otherwise. The sample includes 2.15-2.17 million observations. Standard errors, in parentheses, are clustered by zipcode. *p <.1; **p <.05; ***p <.01.

Table 9: Frequent Referrer: Role of Selection & Disciplining in Referral Networks

Panel A: Application Acceptance				
	(1)	(2)	(3)	(4)
FinTech × *Referred*	-0.0015	-0.0053	-0.0023	-0.0038
	(0.0034)	(0.0043)	(0.0038)	(0.0046)
FinTech × *Referred by Frequent Referrer*	-0.0032	0.0043	0.0001	0.0009
	(0.0071)	(0.0089)	(0.0081)	(0.0104)
FinTech × *Referred* + *FinTech* × *Referred by Frequent Referrer* = 0				
F-Statistic	0.5875	0.0179	0.0929	0.1013
p-value	0.4434	0.8935	0.7605	0.7503
Customer × Year FE	Yes	Yes	Yes	Yes
Lender × Year FE	Yes	Yes		
Month × Zipcode FE		Yes		Yes
Lender × Year × Zipcode FE			Yes	Yes
R^2	0.5455	0.6169	0.5571	0.6273
Adjusted R^2	0.4405	0.4317	0.4196	0.4036

Panel B: Default				
	(1)	(2)	(3)	(4)
FinTech × *Referred*	-0.0431***	-0.0477***	-0.0384***	-0.0480***
	(0.0110)	(0.0121)	(0.0120)	(0.0131)
FinTech × *Referred by Frequent Referrer*	0.0839***	0.0728**	0.0850**	0.0727**
	(0.0302)	(0.0313)	(0.0332)	(0.0355)
FinTech × *Referred* + *FinTech* × *Referred by Frequent Referrer* = 0				
F-Statistic	2.0464	0.7362	2.2131	0.5487
p-value	0.1526	0.3909	0.1369	0.4589
Customer × Year FE	Yes	Yes	Yes	Yes
Lender × Year FE	Yes	Yes		
Month × Zipcode FE		Yes		Yes
Lender × Year × Zipcode FE			Yes	Yes
R^2	0.3698	0.4704	0.4017	0.4954
Adjusted R^2	0.2334	0.2235	0.2268	0.2038

This table reports the loan acceptance and default for referrals on their loans with the Fintech lender – which is the primary source of our data – and other lenders. *Acceptance* is an indicator variable equal to one if an applicant's loan application is accepted by the lender, and zero otherwise. *Default* is an indicator variable equal to one if a customer's loan is past-due for more than 90 days, and zero otherwise. *Fintech* is an indicator variable equal to one if a loan application is to the Fintech lender– which is the primary source of our data – and zero for other lenders. *Referred* is an indicator variable equal to one if the customer is referred by an existing customer with the Fintech lender and zero otherwise. *Referred by Frequent Referrer*, equal to one if a new customer is referred by an existing customer who refers five or more customers, and zero otherwise. The sample includes 2.15-2.17 million observations. Standard errors, in parentheses, are clustered by zipcode. *p <.1; **p <.05; ***p <.01.

Table 10: Role of Non-Monetary Incentive in Referral Networks

Panel A: One year since a customer's first acceptance			
	(1)	(2)	(3)
	Reapply	*# Reapply*	*# Acceptance*
Refer a New Customer	0.2221***	0.3347***	0.1441***
	(0.0071)	(0.0130)	(0.0121)
Pseudo R^2	0.0385	0.1969	0.2118
Panel B: Two years since a customer's first acceptance			
	(1)	(2)	(3)
	Reapply	*# Reapply*	*# Acceptance*
Refer a New Customer	0.1970***	0.3781***	0.2042***
	(0.0067)	(0.0140)	(0.0135)
Pseudo R^2	0.0336	0.1898	0.2274
Panel C: Three years since a customer's first acceptance			
	(1)	(2)	(3)
	Reapply	*# Reapply*	*# Acceptance*
Refer a New Customer	0.1861***	0.3888***	0.2237***
	(0.0062)	(0.0141)	(0.0142)
Pseudo R^2	0.0284	0.1853	0.2160
Cohort × Zipcode FE	Yes	Yes	Yes
Credit Score Decile FE	Yes	Yes	Yes
Income Decile FE	Yes	Yes	Yes
Age Decile FE	Yes	Yes	Yes
Gender FE	Yes	Yes	Yes
Education FE	Yes	Yes	Yes

This table reports the relationship between existing customers' decisions to refer new customers and to apply for new loans. *Refer a New Customer* is an indicator variable equal to one if a customer refers any new customers in the next one, two, or three years since their first loan acceptance with the FinTech lender, and zero otherwise. *Reapply* is an indicator variable equal to one if a customer applies for subsequent loans after their first loan acceptance with the FinTech lender, and zero otherwise. *# Reapply* is the number of subsequent loan applications after a customer's first loan acceptance with the FinTech lender. *# Acceptance* is the number of subsequent loan applications that are accepted after a customer's first loan acceptance with the FinTech lender. Cohort captures the date of a customer's first loan acceptance with the FinTech lender, recorded in a month-year format. We consider loan applications one, two, or three years after the cohort date. To avoid right-censoring bias as the sample period ends in 2022, we only include customers whose cohort date is before 2022 in Panel A, before 2021 in Panel B, and before 2020 in Panel C. The zip code refers to the residential zip code reported by the applicant in their application. The credit score represents the applicant's Equifax credit score, with applicants grouped into deciles based on their scores; a separate category is included for new-to-credit applicants who do not have a credit score. Income decile and age decile are determined using verified income and age data, respectively. The variable gender is coded as one for female applicants and zero otherwise. The education fixed effect accounts for the highest level of education reported by the applicant. The sample includes 200K-210K customer-level observations. Poisson Pseudo-Maximum Likelihood estimator is implemented in all regressions. Standard errors, in parentheses, are clustered by zip code. *p <.1; **p <.05; ***p <.01.

Table 11: Role of Non-Monetary Incentive in Referral Networks: Post Referral Outcomes for Referred Borrowers

Dep Var: *# Acceptance*	(1)	(2)	(3)
	Time since a customer's first acceptance:		
	One Year	Two Years	Three Years
Refer a New Customer × *Default*	-0.1564**	-0.1320**	-0.1449**
	(0.0767)	(0.0567)	(0.0584)
Refer a New Customer × *No Default*	0.0575	0.0860*	0.0803*
	(0.0535)	(0.0506)	(0.0474)
Omitted Category: *Refer a New Customer* × *Not Accepted*			
Refer a New Customer × *Default* = *Refer a New Customer* × *No Default*			
χ^2	12.99	13.30	14.42
p-value	0.0003	0.0003	0.0001
Cohort × Zipcode FE	Yes	Yes	Yes
Credit Score Decile FE	Yes	Yes	Yes
Income Decile FE	Yes	Yes	Yes
Age Decile FE	Yes	Yes	Yes
Gender FE	Yes	Yes	Yes
Education FE	Yes	Yes	Yes
Pseudo R^2	0.2864	0.2404	0.2385

This table reports the relationship between existing customers' loan outcomes and their referrals' loan performance. *Refer a New Customer* is an indicator variable equal to one if a customer refers any new customers in the next one, two, or three years since their first loan acceptance with the FinTech lender, and zero otherwise. *Default* is an indicator variable equal to one if at least one of the referred new customers defaults during the same period with *Refer a New Customer*, and zero otherwise. *No Default* is an indicator variable equal to one if none of the new customers is default, and zero otherwise. *# Acceptance* is the number of subsequent loan applications that are accepted after a customer's first loan acceptance with the FinTech lender. Cohort captures the date of a customer's first loan acceptance with the FinTech lender, recorded in a month-year format. We consider loan applications one, two, or three years after the cohort date, and the sample includes customers who refer new customers during the one, two, or three-year period. To avoid right-censoring bias as the sample period ends in 2022, we only include customers whose cohort date is before 2022 in column (1), before 2021 in column (2), and before 2020 in column (3). The zip code refers to the residential zip code reported by the applicant in their application. The credit score represents the applicant's Equifax credit score, with applicants grouped into deciles based on their scores; a separate category is included for new-to-credit applicants who do not have a credit score. Income decile and age decile are determined using verified income and age data, respectively. The variable gender is coded as one for female applicants and zero otherwise. The education fixed effect accounts for the highest level of education reported by the applicant. The sample includes 1.5K-1.7K customer-level observations. Poisson Pseudo-Maximum Likelihood estimator is implemented in all regressions. Standard errors, in parentheses, are clustered by zipcode. *p <.1; **p <.05; ***p <.01.

Table 12: Inputs to the sufficient statistic: values, category, and provenance

Symbol	Value	Category and source
Panel A. Calibrated or externally measured.		
β	0.75	Referrer's date-zero discount factor over the eight-month horizon.
Panel B. Estimated or observed.		
$q_l = q_B$	0.7391	One minus the sample mean of Prob(Default). Table 1, summary statistics, mean = 0.2609, n = 316,832.
$\Delta^{(0)}$	0.0794	Pre-policy disciplining effect. Table 7, column (11), coefficient on *Referred*, −0.0794 (s.e. 0.0063).
$\Delta^{(1)}$	0.0371	Post-policy disciplining effect. Table 7, column (11), *Referred* plus *Referred* × *Post*: $-0.0794 + 0.0423 = -0.0371$.
$\hat{a}_l/\hat{a}_h$	1.80	Observed penalty-to-reward ratio in approval outcomes over the three-year horizon. Table 11, column (3): *Refer a New Customer* × *Default* = −0.1449 and *Refer a New Customer* × *No Default* = 0.0803, giving $0.1449/0.0803 \simeq 1.80$.
π_0	₹1,250; ₹3,500	Upfront referral bonus before and after the October 2021 policy change. These amounts are approximately 3.5% and 9.8% of mean requested loan principal, respectively.
s_0, s_1	2.00%, 3.75–4.00%	Referred share of applications before and after the bonus increase, from Figure 5.
Panel C. Derived.		
κ_h	0.185744	$(1 - q_l R)/(1 - \beta q_l R)$, Equation (5).
$q_l/(1 - q_l)$	2.8329	Model-implied penalty-to-reward ratio in referrer-value units, implied by Equation (24).

Notes. Table and column references are to the empirical analysis. Panel A reports maintained or calibrated parameters used in Section 6. Panel B reports empirical estimates and program parameters used in Section 6. Panel C reports quantities derived from these inputs and used in Section 6. Digits are reported to make the arithmetic reproducible and do not indicate precision.

Internet Appendix for:

"Referrals in Lending Markets: Role of Social Networks in the Age of Technology"

.1 Proof of Lemma 1

Proof. Consider a candidate cost-minimizing contract with $T < 0$. We construct another contract that induces participation by the same set of referrers at a strictly lower cost.

Because $T < 0$, Equation (13) implies $a_l > 0$. Define

$$\delta := -\frac{T}{\beta(1 - q_B)} > 0, \tag{A.1}$$

and consider a new contract that leaves (π_0, π_r, a_h) unchanged and reduces the default penalty so that

$$\widetilde{a}_l W_l = a_l W_l - \delta. \tag{A.2}$$

Equation (13) implies $\delta < a_l W_l$, so $\widetilde{a}_l \geq 0$. By construction, the participation intercept under the new contract is

$$\widetilde{T} = T + \beta(1 - q_B)\delta = 0. \tag{A.3}$$

At the same time, Equation (12) implies

$$\widetilde{S} = S - \delta.$$

Consider now the participation constraint of any disciplinable referrer with type $\Delta \geq d$. Relative to the original contract,

$$\left(\widetilde{T} + \beta\Delta\widetilde{S}\right) - \left(T + \beta\Delta S\right) = \beta(1 - q_B - \Delta)\delta \geq 0, \tag{A.4}$$

where the inequality follows because $q_B + \Delta \leq 1$. Thus, every disciplinable referrer who participates under the original contract continues to participate under the new contract. Moreover, since $\widetilde{T} = 0$, the participation constraint in Equation (14) also implies the incentive-compatibility constraint.

The new contract is strictly less costly. Using the definition of $\bar{C}(d)$ in Equation (17), the reduction in the average contractual cost is

$$\bar{C}(d) - \bar{\widetilde{C}}(d) = \left(1 - q^*(d)\right)\kappa_l\delta > 0. \tag{A.5}$$

Hence, a contract with $T < 0$ cannot be cost-minimizing. Therefore, a cost-minimizing contract satisfies $T = 0$.

Given $T = 0$, the minimum feasible incentive slope is

$$S = \frac{e}{\beta d},$$

as stated in Equation (21). This proves the lemma. □

.2 Proof of Proposition 1

Proof. By Lemma 1, a cost-minimizing contract satisfies $T = 0$ and $S = S(d)$, where $S(d)$ is given by Equation (21). For convenience, define

$$H := a_h W_h, \qquad L := a_l W_l. \tag{A.6}$$

Equation (12) then gives

$$L = S - \pi_r - H.$$

Substituting this expression into Equation (13) and using $T = 0$ gives

$$\pi_0 = \beta \left[(1 - q_B)S - \pi_r - H \right]. \tag{A.7}$$

Because $\pi_0 \geq 0$, feasible contracts satisfy

$$\pi_r + H \leq (1 - q_B)S. \tag{A.8}$$

Assumption 2, together with $d \geq \Delta_m$, implies

$$S(d) = \frac{e}{\beta d} \leq \frac{e}{\beta \Delta_m} \leq \min\{W_h, W_l\}.$$

Thus, the bounds $a_h, a_l \leq 1$ do not bind for the contracts characterized below.

Using Equation (5), so that $c_h = \kappa_h W_h$ and $c_l = \kappa_l W_l$, Equation (17) can be written as

$$\bar{C}(d) = \pi_0 + q^*(d)\pi_r + q^*(d)\kappa_h H + \left(1 - q^*(d)\right)\kappa_l L.$$

Substituting $L = S - \pi_r - H$ and Equation (A.7) yields

$$\begin{aligned} \bar{C}(d) = S \left[\beta(1 - q_B) + \left(1 - q^*(d)\right)\kappa_l \right] + \sigma(d)H \\ + \left[\sigma(d) + q^*(d)(1 - \kappa_h) \right] \pi_r, \end{aligned} \tag{A.9}$$

where $\sigma(d)$ is defined in Equation (22).

From Equation (5) and $R < \beta^{-1}$,

$$\kappa_h < 1.$$

The coefficient on π_r in Equation (A.9) therefore exceeds the coefficient on H by $q^*(d)(1 - \kappa_h) > 0$. Thus, whenever repayment-contingent compensation is provided, the approval reward is less costly than the ex-post monetary bonus.

If $\sigma(d) > 0$, Equation (A.9) is minimized by setting $H = \pi_r = 0$. It follows that $L = S(d)$, and Equation (A.7) gives

$$\pi_0 = \beta(1 - q_B)S(d) = \frac{(1 - q_B)e}{d}.$$

This gives the Cash Regime in Equation (23).

If $\sigma(d) < 0$, the lender minimizes cost by setting $\pi_r = 0$ and increasing H until Equation (A.8) binds:

$$H = (1 - q_B)S(d).$$

Equation (A.7) then gives $\pi_0 = 0$, while

$$L = q_B S(d).$$

This gives the Credit Access Regime in Equation (24).

If $\sigma(d) = 0$, Equation (A.9) is independent of H, while the coefficient on π_r is strictly positive. Hence, $\pi_r = 0$ and any

$$H \in [0, (1 - q_B)S(d)] \tag{A.10}$$

is cost-minimizing. These contracts are precisely the convex combinations of the contracts in Equations (23) and (24).

Finally, substituting the cost-minimizing choices into Equation (A.9) gives

$$\bar{C}^*(d) = \left[\beta(1 - q_B) + \left(1 - q^*(d)\right)\kappa_l + (1 - q_B)\min\{\sigma(d), 0\}\right] S(d),$$

which is Equation (25). This proves the proposition. □

.3 Proof of Proposition 2

Proof. Let $C^*(\Delta; d)$ denote the cost in Equation (16), evaluated at the cost-minimizing contract that implements threshold d. The net value generated by the marginal participating referrer is

$$v(d) = \left[(q_B + d)R - 1\right]K_B - C^*(d; d). \tag{A.11}$$

Using Equation (15), the lender's value can be written as

$$\frac{V(d)}{\eta} = \int_d^{\bar{\Delta}} \left\{\left[(q_B + \Delta)R - 1\right]K_B - C^*(\Delta; d)\right\} g(\Delta)\, d\Delta. \tag{A.12}$$

Within either regime in Proposition 1, the contractual instruments are proportional to the incentive slope $S(d) = e/(\beta d)$. Therefore, holding Δ fixed,

$$\frac{\partial C^*(\Delta; d)}{\partial d} = -\frac{C^*(\Delta; d)}{d}. \tag{A.13}$$

Differentiating Equation (A.12) gives

$$\begin{aligned}\frac{V'(d)}{\eta} &= -g(d)v(d) + \frac{1}{d}\int_d^{\bar{\Delta}} C^*(\Delta; d)g(\Delta)\, d\Delta \\ &= -g(d)v(d) + \frac{1 - G(d)}{d}\bar{C}^*(d).\end{aligned} \tag{A.14}$$

The second equality uses the definition of the average contractual cost in Equation (17).

Using the definition of $\lambda(d)$ in Equation (28), Equation (A.14) becomes

$$\frac{V'(d)}{\eta} = g(d)\left[\frac{\bar{C}^*(d)}{\lambda(d)} - v(d)\right] = \frac{g(d)\bar{C}^*(d)}{\lambda(d)}\left[1 - \Lambda(d)\right], \tag{A.15}$$

where the second equality follows from Equation (27).

At an interior optimum, $V'(d^*) = 0$. Equation (A.15) therefore implies

$$\lambda(d^*)v(d^*) = \bar{C}^*(d^*),$$

or, equivalently,

$$\Lambda(d^*) = 1,$$

as stated in Equation (29).

Equation (A.15) also gives the final result. When $\Lambda(d) > 1$, $V'(d) < 0$, so locally lowering d increases the lender's value. When $\Lambda(d) < 1$, $V'(d) > 0$, so locally raising d increases the lender's value. This proves the proposition. □

Appendix A When are Referrers More Effective?

Table A.1: Characteristics between Customers and Referrers

Panel A: Feature Importance

Types of Model	Linear		Tree-based			Boosting		SVM	Neural Networks				
Variables	Linear	Logistic	Decision Tree	Random Forest	Extra Trees	Gradient Boosting	LightGBM	SVM	MLP	TabNet	Wide&Deep	Average	Rank
Customer Age < Referrer Age	0.2232	0.2203	0.0523	0.1846	0.1800	0.3140	0.1192	0.2640	0.1693	0.2652	0.1421	0.1940	1
Same Company	0.1346	0.1292	0.2139	0.1855	0.1859	0.1373	0.1476	0.2859	0.1535	0.2483	0.1465	0.1789	2
Same Caste	0.2066	0.2068	0.0644	0.1629	0.1465	0.0975	0.2367	0.1964	0.1385	0.0684	0.1690	0.1540	3
Customer Income < Referrer Income	0.0872	0.0857	0.2720	0.1512	0.1644	0.1392	0.1673	0.0453	0.1432	0.0798	0.1442	0.1345	4
Male Customer and Female Referrer	0.0371	0.0460	0.1597	0.1365	0.1448	0.2094	0.1082	0.1735	0.1423	0.1014	0.1518	0.1282	5
Female Customer and Male Referrer	0.1222	0.1170	0.1036	0.0966	0.0837	0.0602	0.1876	0.0349	0.1107	0.1492	0.1365	0.1093	6
Female Customer and Female Referrer	0.1891	0.1950	0.1341	0.0828	0.0947	0.0424	0.0334	0.0000	0.1426	0.0878	0.1099	0.1011	7

Panel B: Direction of the Effect

Types of Model	Linear		Tree-based			Boosting		SVM	Neural Networks					
Variables	Linear	Logistic	Decision Tree	Random Forest	Extra Trees	Gradient Boosting	LightGBM	SVM	MLP	TabNet	Wide&Deep	# Negative	# Positive	Sign
Customer Age < Referrer Age	–	–	–	–	–	–	–	–	–	–	–	11	0	Negative
Same Company	–	–	–	–	–	–	–	–	–	–	–	11	0	Negative
Same Caste	–	–	–	–	–	–	–	–	–	–	–	11	0	Negative
Customer Income < Referrer Income	–	–	–	–	–	–	–	+	–	–	–	10	1	Negative
Male Customer and Female Referrer	+	+	–	+	+	+	+	–	–	–	–	6	5	Uncertain
Female Customer and Male Referrer	–	–	–	–	–	–	–	–	–	–	–	11	0	Negative
Female Customer and Female Referrer	–	–	–	–	–	+	–	–	+	+	+	7	4	Uncertain

This table reports the relationship between default and customer characteristics. *Customer Age < Referrer Age* is an indicator variable equal to one if a customer is younger than the referrer, and zero otherwise. *Same Company* is an indicator variable equal to one if a new customer works in the same company with the referrer, and zero otherwise. *Same Caste* is an indicator variable equal to one if a customer is in the same caste with the referrer, and zero otherwise. *Customer Income < Referrer Income* is an indicator variable equal to one if a new customer has a lower income than the referrer, and zero otherwise. *Male Customer and Female Referrer* is an indicator variable equal to one if the new customer is male and the referrer is female, and zero otherwise. *Female Customer and Male Referrer* is an indicator variable equal to one if the new customer is female and the referrer is male, and zero otherwise. *Female Customer and Female Referrer* is an indicator variable equal to one if both the new customer and the referrer are female, and zero otherwise. The omitted category is *Male Customer and Male Referrer*, where both the new customer and the referrer are male. Panel A reports the feature importance in predicting default with linear and machine learning models, including Linear Regression, Logistic Regression, Decision Tree, Decision Tree, Random Forest, Extra Trees, Gradient Boosting, Light Gradient Boosting Machine (LightGBM), Support Vector Machine (SVM), Multilayer perception (MLP), TabNet, and Wide & Deep Net. Panel B reports the direction of the effect by each variable, where – indicates that a variable reduces the default rate while + indicates it increases the default rate.

Appendix B Robustness

Table B.1: Robustness: Role of Selection & Disciplining in Referral Networks in Credit Markets

Panel A: Removing other lenders with a referral program		
	(1)	(2)
Dep Var	*Acceptance*	*Default*
Fintech × *Referred*	-0.0037	-0.0373***
	(0.0042)	(0.0123)
Customer × Year FE	Yes	Yes
Month × Zipcode FE	Yes	Yes
Lender × Year × Zipcode	Yes	Yes
R^2	0.6273	0.4954
Adjusted R^2	0.4037	0.2038

Panel B: Comparing with other non-banks		
	(1)	(2)
Dep Var	*Acceptance*	*Default*
Fintech × *Referred*	-0.0015	-0.0262*
	(0.0048)	(0.0135)
Customer × Year FE	Yes	Yes
Month × Zipcode FE	Yes	Yes
Lender × Year × Zipcode	Yes	Yes
R^2	0.6361	0.4905
Adjusted R^2	0.4249	0.2096

This table reports the loan acceptance and default for referrals on their loans with the Fintech lender – which is the primary source of our data – and other lenders, where we remove loan applications with other lenders that have similar referral programs (Panel A) and compare the Fintech lender with non-banks (Panel B). *Acceptance* is an indicator variable equal to one if an applicant's loan application is accepted by the lender, and zero otherwise. *Default* is an indicator variable equal to one if a customer's loan is past-due for more than 90 days, and zero otherwise. *Fintech* is an indicator variable equal to one if a loan application is to the Fintech lender– which is the primary source of our data – and zero for other lenders. *Referred* is an indicator variable equal to one if the customer is referred by an existing customer with the Fintech lender and zero otherwise. The sample includes 2.15-2.17 million observations. Standard errors, in parentheses, are clustered by zipcode. $*p$ <.1; $**p$ <.05; $***p$ <.01.

Table B.2: Robustness: Internal Validity of the Khwaja-Mian Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full Sample		Khwaja-Mian Sample		Difference	Difference	p-value for
	Mean	St Dev	Mean	St Dev	(1)-(3)	relative to (1)	Difference
Female	0.1235	0.3290	0.1169	0.3213	0.007	0.053	0.7824
Age (in years)	29.9494	6.8859	30.3816	6.5298	-0.432	-0.014	0.4390
Completed High School	0.2883	0.4530	0.2415	0.4280	0.047	0.162	0.1844
Completed College	0.5760	0.4942	0.6041	0.4890	-0.028	-0.049	0.4844
Completed Masters	0.1357	0.3424	0.1544	0.3614	-0.019	-0.138	0.5384
Monthly Income (₹)	31,035.04	22,052.48	33,301.05	23,414.29	-2,266.01	-0.073	0.2290
Credit Score	675.94	67.78	686.40	60.70	-10.457	-0.015	0.0392
Loan Amount (₹)	35,703.93	45,924.02	32,003.18	40,852.64	3,700.75	0.104	
Debt-to-Income	1.0965	0.9155	0.8922	0.7491	0.204	0.186	
Maturity (in months)	8.0320	4.8173	5.8073	3.8076	2.225	0.277	
Referred	0.0285	0.1664	0.0303	0.1713	-0.002	-0.062	0.9982

This table presents summary statistics of key variables in the customer-level data, comparing the full sample and customers in Khwaja-Mian tests. The table reports statistics for: (1) key demographic variables (gender (female), age in years, and indicator variables for educational attainment (high school, college degree, or master's degree)); (2) applicant-level economic characteristics (monthly income, presence of a credit score, and credit score value, conditional on having a score); and (3) application-level characteristics (loan amount requested, debt-to-income ratio, loan maturity in months, and whether the application is referred, accepted, and default).

Table B.3: Robustness: Referral Program & Credit Market Outcomes

Panel A: Application Acceptance					
	(1)	(2)	(3)	(4)	(5)
Referred	0.0306***	0.0315***	0.0317***	0.0307***	0.0318***
	(0.0015)	(0.0016)	(0.0016)	(0.0015)	(0.0016)
R^2	0.8021	0.7947	0.7882	0.8095	0.7964
Adjusted R^2	0.7620	0.7535	0.7462	0.7715	0.7566
Panel B: Loan Default					
	(1)	(2)	(3)	(4)	(5)
Referred	-0.0609***	-0.0605***	-0.0574***	-0.0609***	-0.0573***
	(0.0046)	(0.0051)	(0.0056)	(0.0046)	(0.0057)
R^2	0.3302	0.3284	0.3284	0.3289	0.3267
Adjusted R^2	0.1111	0.1063	0.1080	0.1109	0.1077
Sample Restriction	Baseline	After Jan, 2018	After Jan, 2019	Excluding COVID	(3)-(4)
Month × Zipcode FE	Yes	Yes	Yes	Yes	Yes
Credit Score Decile FE	Yes	Yes	Yes	Yes	Yes
Income Decile FE	Yes	Yes	Yes	Yes	Yes
Age Decile FE	Yes	Yes	Yes	Yes	Yes
Gender FE	Yes	Yes	Yes	Yes	Yes
Education FE	Yes	Yes	Yes	Yes	Yes
Loan Purpose FE	Yes	Yes	Yes	Yes	Yes
Product Type FE	Yes	Yes	Yes	Yes	Yes
Loan Amount Decile FE	Yes	Yes	Yes	Yes	Yes
DTI Decile FE	Yes	Yes	Yes	Yes	Yes

This table reports the relationship between loan acceptance, loan default and referrals. *Acceptance* is an indicator variable equal to one if the loan application is approved, and zero otherwise. *Default* is an indicator variable equal to one if a customer's loan is past-due for more than 90 days, and zero otherwise. *Referred* is an indicator variable equal to one if a new customer is referred by an existing customer with the Fintech lender and zero otherwise. Month captures the application date, recorded in a month-year format. The zip code refers to the residential zip code reported by the applicant in their application. The credit score represents the applicant's Equifax credit score, with applicants grouped into deciles based on their scores; a separate category is included for new-to-credit applicants who do not have a credit score. Income decile and age decile are determined using verified income and age data, respectively. The variable gender is coded as one for female applicants and zero otherwise. The education fixed effect accounts for the highest level of education reported by the applicant. The loan purpose indicates the reason for the loan as specified by the applicant, while product type denotes the specific type of loan applied for. The loan amount decile is based on the requested amount, categorized into deciles, and the debt-to-income (DTI) ratio is constructed by dividing the requested loan amount by the applicant's verified monthly income. The sample includes 1.15-1.22 million observations. Column 1 presents the results with the baseline sample. Columns 2 and 3 drop all applications before January 2018 and January 2019, respectively. Column 4 removes the sample of applications during the COVID-19 period between April 2020 and October 2020. Column 5 drops the sample of applications before January 2019 and during COVID-19. Standard errors, in parentheses, are clustered by zip code. $^{*}p$ <.1; $^{**}p$ <.05; $^{***}p$ <.01.

Table B.4: Frequent Referrer: Referral Program & Application Acceptance

Dep Var: *Acceptance*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Referred	0.0726***	0.0637***	0.0381***	0.0383***	0.0376***	0.0373***	0.0369***	0.0351***	0.0344***	0.0350***	0.0352***
	(0.0026)	(0.0029)	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0017)	(0.0017)	(0.0017)
Referred by Frequent Referrer	-0.0983***	-0.0852***	-0.0240***	-0.0234***	-0.0278***	-0.0272***	-0.0266***	-0.0257***	-0.0228***	-0.0248***	-0.0250***
	(0.0053)	(0.0058)	(0.0041)	(0.0041)	(0.0041)	(0.0041)	(0.0041)	(0.0041)	(0.0037)	(0.0037)	(0.0037)
Referred + *Referred by Frequent Referrer* = 0											
F-Statistics	30.6142	17.3068	14.4844	16.3092	7.0226	7.5032	7.7771	6.4696	12.3301	9.4236	9.5488
p-value	<0.0001	<0.0001	0.0001	0.0001	0.0081	0.0062	0.0053	0.0110	0.0004	0.0021	0.0020
Month FE	Yes										
Month × Zipcode FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Credit Score Decile FE			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Decile FE				Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age Decile FE					Yes	Yes	Yes	Yes	Yes	Yes	Yes
Gender FE						Yes	Yes	Yes	Yes	Yes	Yes
Education FE							Yes	Yes	Yes	Yes	Yes
Loan Purpose FE								Yes	Yes	Yes	Yes
Product Type FE									Yes	Yes	Yes
Loan Amount Decile FE										Yes	Yes
DTI Decile FE											Yes
R^2	0.1930	0.3374	0.7697	0.7699	0.7707	0.7708	0.7710	0.7762	0.8006	0.8016	0.8021
Adjusted R^2	0.1930	0.2032	0.7231	0.7234	0.7243	0.7244	0.7246	0.7309	0.7602	0.7614	0.7620

This table reports the relationship between loan acceptance and referrals. *Acceptance* is an indicator variable equal to one if the loan application is approved, and zero otherwise. *Referred* is an indicator variable equal to one if a new customer is referred by an existing customer with the Fintech lender and zero otherwise. *Referred by Frequent Referrer*, equal to one if a new customer is referred by an existing customer who refers five or more customers, and zero otherwise. Month captures the application date, recorded in a month-year format. The zip code refers to the residential zip code reported by the applicant in their application. The credit score represents the applicant's Equifax credit score, with applicants grouped into deciles based on their scores; a separate category is included for new-to-credit applicants who do not have a credit score. Income decile and age decile are determined using verified income and age data, respectively. The variable gender is coded as one for female applicants and zero otherwise. The education fixed effect accounts for the highest level of education reported by the applicant. The loan purpose indicates the reason for the loan as specified by the applicant, while product type denotes the specific type of loan applied for. The loan amount decile is based on the requested amount, categorized into deciles, and the debt-to-income (DTI) ratio is constructed by dividing the requested loan amount by the applicant's verified monthly income. All regressions include an equal number of observations between 1.2-1.3 million. Standard errors, in parentheses, are clustered by zip code. $^{*}p$ <.1; $^{**}p$ <.05; $^{***}p$ <.01.

Table B.5: Frequent Referrer: Referral Program & Loan Default

Dep Var: *Default*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Referred	-0.0690***	-0.0667***	-0.0674***	-0.0688***	-0.0686***	-0.0682***	-0.0682***	-0.0686***	-0.0668***	-0.0655***	-0.0669***
	(0.0042)	(0.0049)	(0.0048)	(0.0049)	(0.0049)	(0.0049)	(0.0049)	(0.0049)	(0.0047)	(0.0047)	(0.0047)
Referred by Frequent Referrer	0.0746***	0.0608***	0.0546***	0.0491***	0.0487***	0.0477***	0.0474***	0.0483***	0.0456***	0.0448***	0.0449***
	(0.0111)	(0.0128)	(0.0126)	(0.0125)	(0.0125)	(0.0125)	(0.0125)	(0.0125)	(0.0123)	(0.0122)	(0.0122)
Referred + *Referred by Frequent Referrer* =0											
F-Statistics	0.2515	0.2112	1.0508	2.5128	2.5656	2.7444	2.8067	2.7029	3.1600	3.0162	3.4108
p-value	0.6160	0.6458	0.3053	0.1130	0.1092	0.0976	0.0939	0.1002	0.0755	0.0825	0.0648
Month FE	Yes										
Month × Zipcode FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Credit Score Decile FE			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Decile FE				Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age Decile FE					Yes	Yes	Yes	Yes	Yes	Yes	Yes
Gender FE						Yes	Yes	Yes	Yes	Yes	Yes
Education FE							Yes	Yes	Yes	Yes	Yes
Loan Purpose FE								Yes	Yes	Yes	Yes
Product Type FE									Yes	Yes	Yes
Loan Amount Decile FE										Yes	Yes
DTI Decile FE											Yes
R^2	0.0497	0.2939	0.3027	0.3111	0.3112	0.3113	0.3115	0.3138	0.3262	0.3281	0.3302
Adjusted R^2	0.0494	0.0633	0.0749	0.0860	0.0860	0.0861	0.0864	0.0894	0.1059	0.1083	0.1111

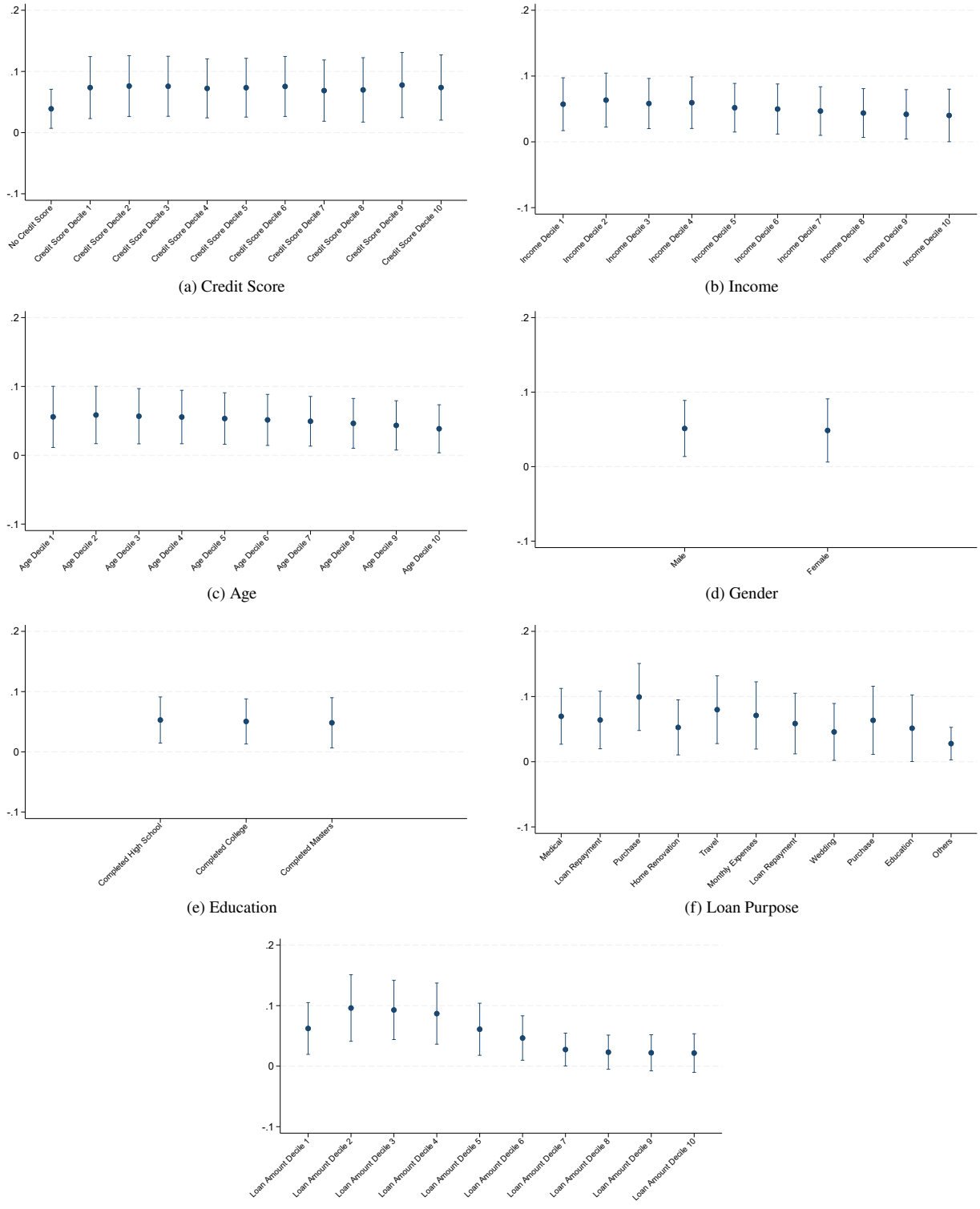
This table reports the relationship between application default and referrals. *Default* is an indicator variable equal to one if a customer's loan is past-due for more than 90 days, and zero otherwise. *Referred* is an indicator variable equal to one if a new customer is referred by an existing customer with the Fintech lender and zero otherwise. *Referred by Frequent Referrer*, equal to one if a new customer is referred by an existing customer who refers five or more customers, and zero otherwise. Month captures the application date, recorded in a month-year format. The zip code refers to the residential zip code reported by the applicant in their application. The credit score represents the applicant's Equifax credit score, with applicants grouped into deciles based on their scores; a separate category is included for new-to-credit applicants who do not have a credit score. Income decile and age decile are determined using verified income and age data, respectively. The variable gender is coded as one for female applicants and zero otherwise. The education fixed effect accounts for the highest level of education reported by the applicant. The loan purpose indicates the reason for the loan as specified by the applicant, while product type denotes the specific type of loan applied for. The loan amount decile is based on the requested amount, categorized into deciles, and the debt-to-income (DTI) ratio is constructed by dividing the requested loan amount by the applicant's verified monthly income. All regressions include an equal number of observations between 255k-265K. Standard errors, in parentheses, are clustered by zip code. *p <.1; **p <.05; ***p <.01.

Table B.6: Robustness: Role of Selection & Disciplining in Referral Networks by Frequent Referrers

Panel A: Application Acceptance				
	(1)	(2)	(3)	(4)
FinTech × *Referred*	-0.0041	-0.0038	-0.0038	-0.0033
	(0.0047)	(0.0046)	(0.0046)	(0.0046)
FinTech × *Referred by Frequent Referrer*	0.0019	0.0009	0.0007	-0.0034
	(0.0102)	(0.0104)	(0.0110)	(0.0105)
Fintech × *Referred* + *Fintech* × *Referred by Frequent Referrer* = 0				
F-Statistic	0.0564	0.1013	0.0997	0.4976
p-value	0.8123	0.7503	0.7522	0.4806
R^2	0.6273	0.6273	0.6273	0.6273
Adjusted R^2	0.4036	0.4036	0.4036	0.4036
Panel B: Default				
	(1)	(2)	(3)	(4)
FinTech × *Referred*	-0.0523***	-0.0480***	-0.0501***	-0.0517***
	(0.0133)	(0.0131)	(0.0128)	(0.0128)
FinTech × *Referred by Frequent Referrer*	0.0764**	0.0727**	0.0950**	0.1134***
	(0.0329)	(0.0355)	(0.0372)	(0.0384)
Fintech × *Referred* + *Fintech* × *Referred by Frequent Referrer* = 0				
F-Statistic	0.6335	0.5487	1.5902	2.8164
p-value	0.4261	0.4589	0.2073	0.0933
R^2	0.4954	0.4954	0.4954	0.4954
Adjusted R^2	0.2038	0.2038	0.2038	0.2038
Threshold of Frequent Referrer	≥ 3	≥ 5	≥ 7	≥ 10
Customer × Year FE	Yes	Yes	Yes	Yes
Month × Zipcode FE	Yes	Yes	Yes	Yes
Lender × Year × Zipcode FE	Yes	Yes	Yes	Yes

This table reports the loan acceptance and default for referrals on their loans with the Fintech lender – which is the primary source of our data – and other lenders. *Acceptance* is an indicator variable equal to one if an applicant's loan application is accepted by the lender, and zero otherwise. *Default* is an indicator variable equal to one if a customer's loan is past-due for more than 90 days, and zero otherwise. *Fintech* is an indicator variable equal to one if a loan application is to the Fintech lender– which is the primary source of our data – and zero for other lenders. *Referred* is an indicator variable equal to one if the customer is referred by an existing customer with the Fintech lender and zero otherwise. *Referred by Frequent Referrer*, equal to one if a new customer is referred by an existing customer who refers X or more customers, and zero otherwise, where X is 3 in column (1), 5 in column (2), 7 in column (3) and 10 in column (4). The sample includes 2.15-2.17 million observations. Standard errors, in parentheses, are clustered by zipcode. *p <.1; **p <.05; ***p <.01.

Figure B.1: Heterogeneous Effect of Referrals: Application Acceptance



(a) Credit Score

(b) Income

(c) Age

(d) Gender

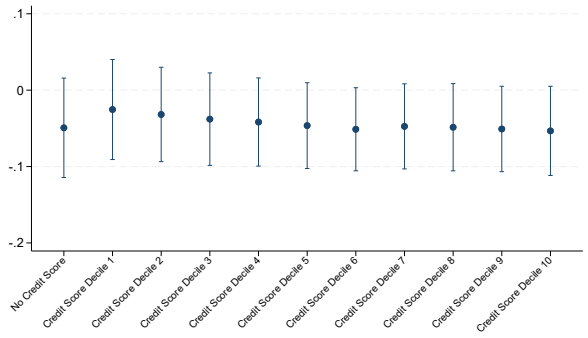
(e) Education

(f) Loan Purpose

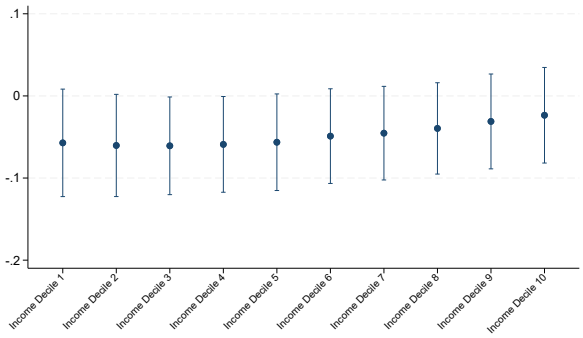
(g) Loan Amount

This figure reports the heterogeneous treatment effect estimates of referral on acceptance probability across characteristics of new customers: credit score, income, age, gender, education, loan purpose, and loan amount. Each panel shows the average effect associated with referral for each subgroup of a specific characteristic and its corresponding 90% confidence interval. These estimates are generated using causal forest, a tree-based machine learning algorithm that flexibly uncovers non-linear relationships and high-order interactions among customer characteristics à la Wager and Athey (2018). The methodology (causal forest) learns directly from the data to identify which customer subgroups benefit most or least from referrals. It is built to estimate how the effect of referrals on application acceptance varies across subgroups.

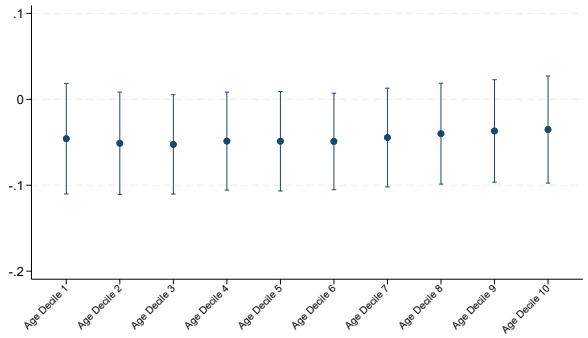
Figure B.2: Heterogeneous Effect of Referrals: Likelihood of Default



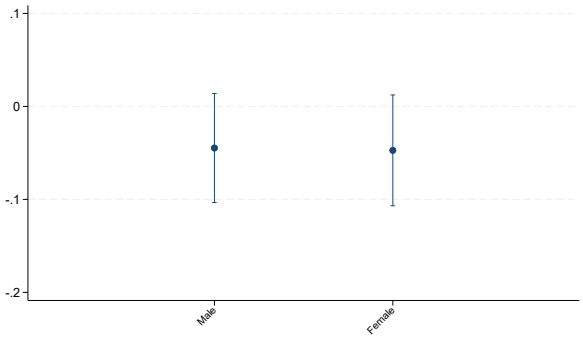
(a) Credit Score



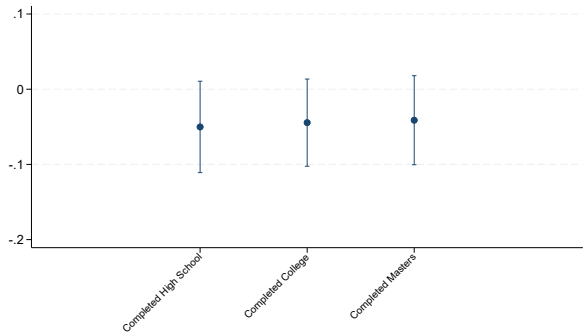
(b) Income



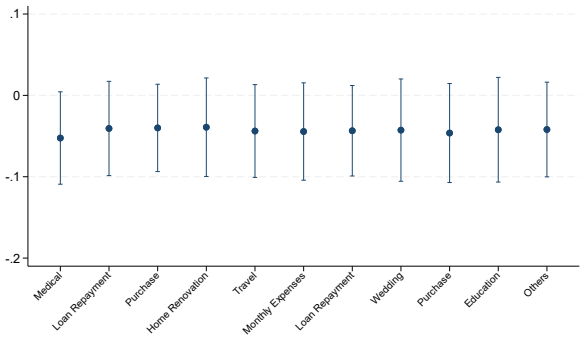
(c) Age



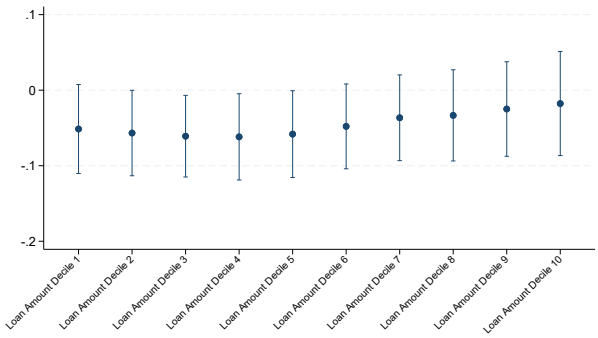
(d) Gender



(e) Education



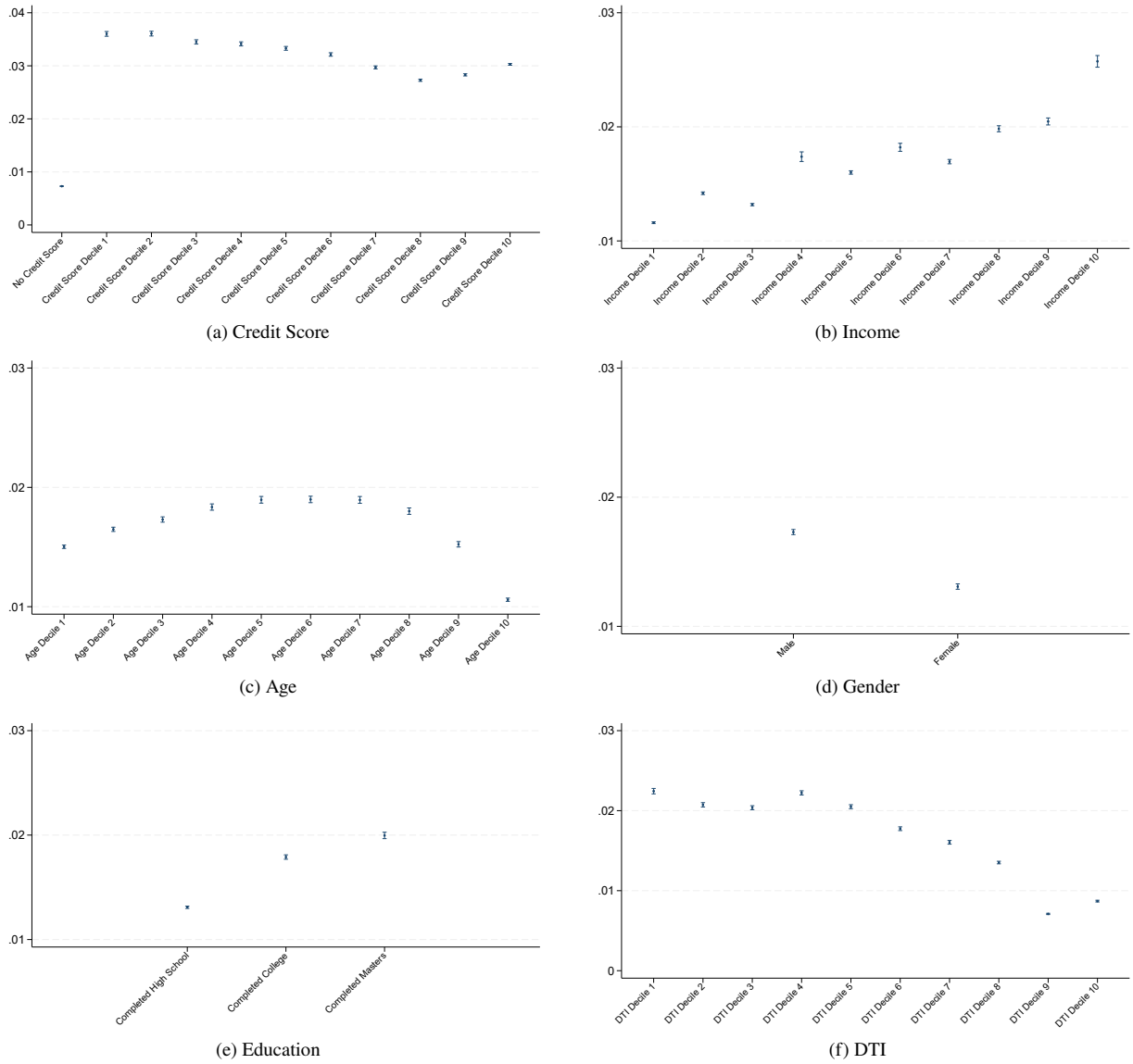
(f) Loan Purpose



(g) Loan Amount

This figure reports the heterogeneous treatment effect estimates of referral on the likelihood of default across characteristics of new customers: credit score, income, age, gender, education, loan purpose, and loan amount. Each panel shows the average effect associated with referral for each subgroup of a specific characteristic and its corresponding 90% confidence interval. These estimates are generated using causal forest, a tree-based machine learning algorithm that flexibly uncovers non-linear relationships and high-order interactions among customer characteristics à la Wager and Athey (2018). The methodology (causal forest) learns directly from the data to identify which customer subgroups benefit most or least from referrals. It is built to estimate how the effect of referrals on default varies across subgroups.

Figure B.3: The Probability of Referrer



(a) Credit Score

(b) Income

(c) Age

(d) Gender

(e) Education

(f) DTI

This figure reports the likelihood of referral across referrers' characteristics: credit score, income, age, gender, education, and DTI. Each panel shows the average effect associated with referral for each subgroup of a specific referrer characteristic and its corresponding 90% confidence interval. These estimates are generated with random forest, which builds and averages 500 decision trees, to capture non-linearities and interactions among referrer characteristics.