FinTech Lending and Cashless Payments*

Pulak Ghosh
Indian Institute of Management

Boris Vallee
Harvard Business School

Yao Zeng
Wharton

August 30, 2021

Click **here** for latest version

Abstract

This study provides a new perspective on the rise of FinTech lending by uncovering an informational synergy with cashless payments. Theoretically, FinTech lenders screen borrowers more efficiently when borrowers use cashless payments that produce transferable and verifiable information. In turn, because borrowers expect lenders to rely on such payment information to screen them, a strategic consideration to stand out from non-adopting borrowers pushes more borrowers to adopt cashless payments. Using novel loan application data from a leading Indian FinTech lender targeting small businesses and an instrumental variable based on the 2016 Indian Demonetization, we identify that a larger use of cashless payments predicts a higher likelihood of loan approval, a lower interest rate, and lower default. These relationships are stronger for firms of higher observable risk, and for firms of higher quality that can be only inferred from payment records. This synergy supports data sharing and open banking, and more broadly the development of an alternative banking model without a balance sheet or traditional banking relationships.

Keywords: FinTech, lending, payments, verifiability, data sharing, open banking.

*We are grateful to Jipeng Liu and Dolly Yu for outstanding research assistance. We thank Paul Beaumont (discussant), Tobias Berg, Sylvain Catherine, Shawn Cole, Itamar Drechesler, Andreas Fuster, Sasha Indarte, Filippo Mezzanotti, Thomas Philippon, Huan Tang (discussant), Laura Veldkamp, Ansgar Walther, Aluna Wang (discussant), and seminar and conference participants at the University of Zurich, Wharton; CICF, EFA, and the Entrepreneurial Finance Association Conference for helpful comments. We thank Indifi for generously providing us access to their data. All errors are ours only.

1 Introduction

Banks' informational advantage in lending stems from their relationships with borrowers resulting from monitoring and repeated lending (Diamond, 1991, Rajan, 1992) and from deposit-taking (Black, 1975, Fama, 1985, Berlin and Mester, 1999, Puri, Rocholl, and Steffen, 2017), both allowing to produce borrower information *inside* the bank. However, the past decade has witnessed a dramatic rise of lending by FinTech companies, which do not enjoy such relationships.[1] How can FinTech lenders compete with banks and even dominate certain lending markets, and what does it mean for the future of banking?

Our paper provides a new perspective, both theoretically and empirically, on the rise of FinTech lending by linking it to another important financial service: payments.[2] We uncover a synergy between FinTech lending and cashless payments, the latter producing verifiable information *outside* the lending relationship. We provide novel and causal evidence on how the use of outside information of varying verifiability affects lending outcomes, which is scarce in the literature. The uncovered synergy in producing and using outside information implies a joint rise of both FinTech lending and cashless payments, and suggests an alternative banking model without relationships in the traditional sense.

Our study builds on two simple observations. First, compared to traditional banks, FinTech lenders often have a technological advantage in accessing and using outside verifiable information beyond the usual credit bureau ratings. Second, cashless payment service providers collect abundant verifiable data that are economically relevant to borrower creditworthiness and are hard to manipulate. Building on these two observations, we first develop a theoretical framework showing that the interaction between FinTech lenders and cashless payments fosters the development of both. In one direction, FinTech lenders become more efficient in screening borrowers when borrowers adopt cashless payments that produce more verifiable information. In the other direction, because would-be borrowers expect lenders

[1]Financial Stability Board (FSB) and Basel Committee define FinTech as "technologically enabled financial innovation that could result in new business models, applications, process, or products." In the US, FinTech lending has been becoming dominating in some of the most important lending markets including the mortgage markets (e.g., Buchak, Matvos, Piskorski, and Seru, 2018, Fuster, Plosser, Schnabl, and Vickery, 2019) and has also developed dramatically in small-business lending markets (e.g., Gopal and Schnabl, 2020, Erel and Liebersohn, 2021). The rise of FinTech lending has also been particularly pronounced in developing economics (Claessens, Frost, Turner, and Zhu, 2018). However, few FinTech lenders engage in monitoring or take deposits.

[2]Payments are of first-order importance to financial markets and the real economy (Bianchi and Bigio, 2021, Piazzesi and Schneider, 2021). Payments have also been experiencing drastic digitalization and disintermediation (Brunnermeier, James and Landau, 2019, Duffie, 2019). The rise of cashless payments has sped up since the global financial crisis and has coincided with the rise of FinTech lending (Vives, 2019).

to rely on outside verifiable payment information to screen them, a strategic consideration for a borrower to stand out from worse borrowers emerges, which ultimately pushes more borrowers to adopt cashless payments.

Specifically, we build a simple model of FinTech lending and borrowers' choice of payment methods, which generate unique predictions. In the model, there is a risk-neutral firm and a risk-averse financier. A higher-type firm has a better technology and is more likely to produce a better product. The firm is privately informed about its type and needs financing. Thus, the financier relies on its prior and any outside information available in making a financing decision. Before the financing stage, each firm uses its endowment to produce, which we call the production stage. Importantly, the firm chooses to use either cash or cashless payments to process its production. While using cash does not leave any records, the cashless payment service records verifiable information about all production outcomes.[3] If the firm adopts cashless payments, it commits to provide the generated verifiable records to the financier in the financing stage, consistent with FinTech lenders' advantage in accessing such outside information in reality.

We establish the existence of synergy between lending and cashless payments in two steps, each of which highlights one direction of the synergy. We first show how cashless payments improve lending outcomes by highlighting two complementary informational effects of verifiable payment records. First, when a firm adopts the cashless payment service, the payment records help reveal the quality of the firm's technology. This information-revealing effect allows the financier to better screen the firm and to achieve more efficient financing outcomes. It particularly benefits firms of better types, and is stronger when the firm can establish more payment records or the payment records are more verifiable. Second, verifiable payment records also directly reduce the financing risk that the financier bears. This risk-reducing effect benefits both the financier and in turn all firm types, and is stronger when the lender's prior about firm types is less precise.

We then show how the reliance of lending on outside verifiable information fosters the adoption of cashless payments. Under a benchmark in which adoption is costless, all firm types optimally adopt the verifiable cashless payment service. This force stems from a strategic consideration among firm types. When high-type firms adopt cashless payments, the financier will rationally update its belief and expect any firm using cash to have a

[3] To focus on the informational role of cashless payments, we abstract away from their other benefits and costs (i.e., convenience yields or privacy concerns).

low type. Thus, a moderately low-type firm would find it optimal to stand out from even lower-type firms by adopting cashless payments, and this force unravels to push more firm types to adopt cashless payments.[4]

We move on to provide causal empirical evidence on how the use of outside information of varying verifiability affects lending outcomes for borrowers with different characteristics, which is consistent with our model predictions. In doing so, we use novel loan application-level data from one of the leading FinTech lenders in India, which focuses on small business loans. The data is unique and suited to our analysis for several reasons. First, the data includes the whole information set that the FinTech lender observes and uses, including applicants' business characteristics, traditional credit ratings from credit bureau inquiries, and most importantly, detailed payment records of different levels of informational verifiability. Second, the lender gets access to those outside payment records from bank statements, which the borrowers agree to share. Thus, our tests shed light on the implications of data sharing and open banking at the same time. Finally, as we elaborate on below, our data also uniquely allows for a causal identification thanks to the exogenous effect of the 2016 Indian Demonetization on payments.

We develop a methodology to classify each payment appearing on bank statements into cash and cashless payments. Within cashless payments, we can further break down between information-intensive and information-light methods of payments, depending on whether payments are partly aggregated, and whether the payment counter-party is identifiable. Being able to access such granular payment records not only allows Indifi to potentially screen the borrowers more efficiently, but also uniquely allows us as econometricians to test how payment information with varying levels of verifiability affects lending outcomes.

Equipped with this data and measures of use of cashless payments, we study whether such use correlates with loan screening outcomes on both the extensive and intensive margins, controlling non-linearly for a wealth of applicant characteristics such as business size, age, 3-digit zip code, and owner credit score. We find that a higher share of cashless payments is associated with improved lending outcomes: applicants using and submitting more cashless payments are more likely to obtain a loan, and when doing so obtain a lower interest rate. The economic magnitude of this relationship is large. For instance, an interquartile of cashless payments use is associated with an increase in the likelihood of obtaining a loan

[4] This logic resembles the "unraveling argument" analyzed in information disclosure (Milgrom, 1981).

by 3 percentage points, or one-tenth of the baseline approval rate. These results are more pronounced for more verifiable cashless payments, such as individual internet transfers, as opposed to less verifiable ones, such as mobile payments that aggregate payments and do not provide information about the transacting counter-party.

To gain causal identification on the impact of cashless payments on lending outcomes, we use a unique institutional setting in India, the 2016 Indian Demonetization, and the granularity of our data to implement an instrumental variable analysis. We instrument applicant reliance on cashless payments with an indicator variable for the borrower banking at a reserve chest bank branch, that is, a branch that distributes new banknotes and collects damaged old ones. Banking at such an establishment after the demonetization is indeed predictive of lower use of cashless payments, as chest bank clients have better access to the new banknotes during the shortage of cash caused by the Demonetization, and therefore switch less to other means of cashless payment. One innovation of our identification is that it rests on within-district variation but not on the cross-section of districts, as our data uniquely allows us to map borrowers to specific bank branches within a zip code. When instrumenting the use of cashless payments with this plausibly exogenous variation, our previous result is strengthened: a higher share of cashless payments is associated with a significantly higher likelihood of obtaining a loan.

In addition, we find heterogeneous benefits from using cashless payments across firm characteristics. On the one hand, firms that are observably riskier, that is, those who have previous publicly-observed delinquencies or written-off bank loans, benefit more from using cashless payments. This result is consistent with the risk-reducing effect we flesh out in the model. On the other hand, firms that have less volatile payment revenues, that is, those of better quality in terms of residual firm risks that are not publicly observable but can be inferred from outside payment records, also benefit more from using cashless payments. This finding supports the information-revealing effect.

Turning to loan default, we find that, within loans charging the same interest rate and everything else equal, firms that use more cashless payments are less likely to default on the FinTech loan. We interpret this empirical fact as consistent with our model prediction that the use of cashless payments helps the FinTech lender price loans more efficiently, leading to more efficient capital allocations among different borrowers.

Finally, although we focus on FinTech lending and cashless payments due to their increasing economic importance, our message is general. On the one hand, we find that the use of sales data on online marketplaces, which are verifiable and accessible to lenders, also improves lending outcomes. Such information is different from payment records but is conceptually the same in our framework because it is also outside the lending relationship, verifiable, and directly generated from borrowers' production process. Our results are thus externally valid to different types of outside information. On the other hand, our emphasis on FinTech lending does not exclude the possibility of traditional banks using outside verifiable information (other than credit bureau ratings) in lending.[5] Our key message is that FinTech lenders have a comparative advantage in using outside verifiable information, which helps them level the playing field with traditional banks, and we provide direct evidence of such information's effects on lending outcomes, which is scarce in the literature due to the lack of data on how traditional banks may have used outside information.

Taken together, our results provide a new perspective to understand the interaction between FinTech lending and cashless payments. In one direction, our findings of cashless payments improving lending outcomes support policies that promote data sharing and open banking. In the other direction, the prediction of adoption of verifiable payment services is consistent with the trends of many economies increasingly switching to cashless payments hand-in-hand with the rise of FinTech lending, particularly in developing economies such as China, India, or Kenya. The synergy we document also provides a rationale for the trends of FinTech lenders directly offering payment service, and payment service firms and marketplaces offering lending.[6] These developments suggest the emergence of an alternative banking model without a balance sheet or relationships in the traditional sense.

At the conceptual level, our paper contributes to the large relationship banking literature (see Diamond (1991) and Rajan (1992) for pioneer work and Liberti and Petersen (2019) for a survey). Specifically, a branch of this literature suggests that relationship-

[5]Anecdotal evidence suggests that traditional banks may use outside data and require direct access to firms' order books or revenue reporting systems in commercial loan markets (Liberti and Petersen, 2019). Traditional banks also reportedly rely on outside data such as magazine subscriptions and utility bills in consumer loan markets; see https://www.wsj.com/articles/need-cash-companies-are-considering-magazine-subscriptions-and-phone-bills-when-making-loans-11568280601?st=1.

[6]In the US, OnDeck provides payment service to its borrowers, and PayPal, Square, and Stripe all offer loans. In India, Paytm partners with banks to offer lending by providing both payment data access and screening technology to banks. Amazon and Alibaba are prominent examples of marketplaces offering loans to their partners. Relatedly, Frost, Gambacorta, Huang, Shin, and Zbinden (2020) examine a context in Argentina and find that large technology companies (i.e., "BigTech" firms) by acquiring non-traditional information have an advantage in lending relative to traditional credit ratings. Despite expanding their scope, these institutions remain fundamentally different from traditional banks as they do not take deposits and are not regulated as banks.

specific payment processing by a bank for its borrowers as depositors helps to ease information asymmetry between the two parties and facilitate lending by the same bank to those borrowers (e.g., Black, 1975, Fama, 1985, Berlin and Mester, 1999, Mester, Nakamura, and Renault, 2007, Norden and Weber, 2010, Puri, Rocholl, and Steffen, 2017), highlighting the informational spillover between the two sides of bank balance sheet. Our paper shows that the existence of a shared bank balance sheet or relationships inside the same bank is not a necessary condition for such informational spillover between lending and payments. As long as payment information is verifiable, despite being outside, a stand-alone lender may rely on it to better infer about borrowers' creditworthiness, and this reliance in turn encourages borrowers to use verifiable payment services in the first place.

Our paper contributes to the literature on both FinTech lending and cashless payments (see Philippon (2019), Vives (2019) and Allen, Gu, and Jagtiani (2021) for surveys). On the lending side, several studies explore the forces behind the boom of FinTech lending in consumer loans (Buchak, Matvos, Piskorski, and Seru, 2018, Fuster, Plosser, Schnabl, and Vickery, 2019) and small-business loans (Gopal and Schnabl, 2020, Beaumont, Tang, and Vansteenbergh, 2021, Erel and Liebersohn, 2021), highlighting regulatory arbitrage, higher convenience, and faster lending technologies.[7] Our paper highlights an informational channel for FinTech lenders to level the playing field with banks by using outside but verifiable payment information. In this aspect, our paper is closely related to Berg, Burg, Gombovic, and Puri (2020), who empirically show that "digital footprints" such as website registration inputs significantly predict consumer defaults on top of credit scores. Complementing theirs, we focus on payments, a type of footprint that is directly generated by the borrowers' production process and is therefore a direct and hard-to-manipulate signal. This focus allows us to economically pinpoint why and how lenders incorporate outside information in their screening, and to provide a rationale for the voluntary provision of such information by borrowers. Our paper also closely complements Fuster, Goldsmith-Pinkham, Ramadorai, and Walther (2020), who study to what extent better inference technologies such as machine learning can affect lending outcomes including approvals and interest rates.

On the payment side, the existing literature mainly focuses on the direct convenience benefits of cashless payments and their adoptions (e.g. Jack and Suri, 2014, Muralidharan,

[7] A fast-growing and related literature considers peer-to-peer or marketplace lending powered by FinTech platforms, which rely on end-investor screening and investing. Recent studies include Iyer, Khwaja, Luttmer and Shue (2015), Balyuk (2017), Tang (2019) and Vallee and Zeng (2019), for example. We do not focus on peer-to-peer lending but rather the broader screening and lending by general FinTech lenders.

Niehaus, and Sukhtankar, 2016, Higgins, 2019). Among them, our paper closely complements Chodorow-Reich, Gopinath, Mishra, and Narayanan (2020) and Crouzet, Gupta, and Mezzanotti (2020), which both provide comprehensive studies of the 2016 Indian Demonetization on the adoption of cashless payments, focusing on testing money neutrality and network externality, respectively. Parlour, Rajan, and Zhu (2020) theoretically study the competition between stand-alone FinTech payment service providers and traditional banks providing both lending and payment service, sharing with ours a focus on the informational role of payments. We contribute to this literature by showing to what extent using cashless payments of different informational verifiability can affect access to capital and improve lending outcomes. The synergy between lending and outside cashless payments provides an additional rationale for the surging adoption of cashless payments.

Our paper also contributes to a burgeoning literature on the data market and data sharing. Jones and Tonetti (2020) argue that the non-rivalry of data implies increasing returns to the sharing of data and better resource allocations. Ichihashi (2020) shows that sharing consumer data leads to more targeted product offering but also more accurate price discrimination. He, Huang, and Zhou (2020) theoretically examine the lending competition between traditional banks and FinTech lenders when borrowers share their bank data with FinTech lenders and explore the implications of open banking. Our paper complements theirs by explicitly identifying that the access to traditional bank statements that contain borrowers' payment records leads to better lending outcomes by FinTech lenders, and by considering the strategic interaction in adopting cashless payments that is critical in the synergy between FinTech lending and cashless payments. Our unraveling mechanism in the spirit of Milgrom (1981) complements the idea of data externality that an agent's data sharing creates an externality on other agents because one's data is revelatory of others' (Bergemann, Bonatti, and Gan, 2020), because data sharing increases platforms' market power (Kirpalani and Philippon, 2020), or because of agents' behavioral biases (Liu, Sockin, and Xiong, 2020). Our micro-level theory and evidence about the nature and consequences of data accumulation also complements Farboodi and Veldkamp (2020a,b) who examine the interaction between production, adoption of information technologies, and data accumulation at the macroeconomic level.

In what follows: Section 2 presents the theory framework, Section 3 describes the data and institutional details, Section 4 presents the empirical analysis, and Section 5 concludes.

2 Theoretical Framework

2.1 Setting

We build a simple model to capture the interaction of payment technology and lending outcomes. The model has two representative agents: a competitive risk-neutral firm and a competitive risk-averse financier. Time is discrete: $t = 0, 1, 2, ..., n, n + 1$ with $n \geq 1$. We call $\{0, ..., n - 1\}$ the production stage and n the lending stage. The firm has a risky technology, the quality of which is characterized by $z \in \overline{\mathbb{R}}$, the extended real set. Only the firm is privately informed about z, and thus we call z the firm's type, and the financier's prior follows a normal distribution $z \sim N(\mu, \tau_z^{-1})$. If the technology is operated at t, it can produce a product of quality y_t to be delivered at $t + 1$, and the product quality is also i.i.d. normally distributed given the quality of technology: $y_t \sim N(z, \tau_y^{-1})$. Intuitively, a better technology is more likely to produce a better product. The firm has enough capital to operate the technology in the production stage, that is, during $t \in \{0, ..., n-1\}$, yielding a series of realized production outcomes $Y = \{y_t | 0 \leq t \leq n - 1\}$.

At the beginning of the production stage $t = 0$, the firm chooses how to accept payments for the products produced. It can either accept payments in cash, which renders the production outcomes Y non-verifiable, or it can commit to using an outside cashless payment service that allows the realized production outcomes to be documented as a file of payment records, which is verifiable and can be accessed by the financier in the lending stage. For each production outcome y_t, the cashless payment service can generate a record $x_t \sim N(y_t, \tau_x^{-1})$, where the precision τ_x can be naturally interpreted as the level of verifiability. The higher the precision τ_x is, the more verifiable the payment record is. The file of all verifiable payment records can be then denoted by $X = \{x_t | 0 \leq t \leq n - 1\}$.

At the financing stage $t = n$, the firm does not have capital anymore to operate the technology, and thus has to finance the technology through the financier. We assume that the financier has a CARA utility function with absolute risk aversion ρ. Modeling FinTech lenders as being risk-averse allows us to parsimoniously capture the various financial constraints that they face, which effectively make them risk-averse. Importantly, the financier does not know the quality of the firm's technology, and thus must infer it based on the prior and the payment records X that the firm submits, if any. Under the CARA-normal framework, the financier effectively has a mean-variance utility, and thus the effective financing price bid by the competitive financier is

$$p = E[y_n|\mathcal{I}_n] - \frac{\rho}{2} Var[y_n|\mathcal{I}_n] \, ,$$

which is equivalent to

$$p = E[z|\mathcal{I}_n] - \frac{\rho}{2} Var[z|\mathcal{I}_n] \, , \tag{2.1}$$

where $\mathcal{I}_n$ is the financier's information set at $t = n$. Following the literature (e.g., Fishman and Parker, 2015, Vallee and Zeng, 2019), we can also interpret the reciprocal of the price p as an interest rate to map it to a lending context.

The equilibrium concept we consider is a standard sequential equilibrium. Specifically, the equilibrium profile consists of the firm's payment choice policy $x(z) : \overline{\mathbb{R}} \to \{\varnothing, X\}$, that is, whether a firm uses cash or verifiable cashless payments, and the financier's pricing policy $p : \{\mathcal{I}_n\} \to \overline{\mathbb{R}}$, and both agents maximize their expected utilities. According to sequential rationality, the financier makes an inference from both 1) the actual information content of $\mathcal{I}_n$, and 2) the firm's decision of using cashless payments or not, that is, the strategy $x(\cdot)$ itself. The proofs are given in Appendix A.

Before proceeding, we discuss several modeling choices and their roles in our analysis. These modeling choices are also important in tightly bridging our model predictions and the design of empirical tests. First, to focus on the informational role of payments on lending outcomes, we absorb common public signals about borrower creditworthiness such as past credit events all in the prior. In this sense, our notion of firm types captures residual firm risks that can be only potentially inferred from payment records, as opposed to observed risks, which capture the risks that can be inferred from other publicly available signals such as past credit events or firm characteristics. Second, to compare cashless payments to cash in the dimension of verifiability, we do not consider either the convenience yields or the physical costs of using different types of payment methods. The binary choice between cash and one single verifiable cashless payment service over the entire production period is a parsimonious way to capture the essence of payment technology choice.[8] Accordingly, the length of the production period does not correspond to the firm's life cycle but rather to how many verifiable payment records the firm may potentially establish.[9] Finally, we use

[8]In reality, a firm may mix a spectrum of payment methods with different degrees of verifiability. For example, checks are more verifiable than cash but less than online banking transfers.

[9]In testing our model, we will separately control for firms' age, credit history, and the number of total payment records submitted to the lender.

a general framework of financing without specifically modeling the type of securities used. This implies that our theoretical framework can be applied more broadly and may inform future research on the informational role of payments in both debt and equity financing contexts. Since our empirical analysis is based on debt financing, in Internet Appendix IA we also provide an alternative model with explicit debt financing featuring face value, interest rate, and default, yielding the same predictions qualitatively.

2.2 Impact of Payment Records on Optimal Financing

In this subsection, we study how the information content of verifiable cashless payments affects financing outcomes – one direction of the synergy between cashless payments and lending. We thus first consider a sub-game equilibrium in which the firm commits to using cashless payments, that is, $x(z) = X$ for all $z \in \overline{\mathbb{R}}$. In this sub-game equilibrium, observing the actual information context of X, the financier updates its belief about z directly from X only. The following result characterizes the expected informed price by any firm type z from the perspective of $t = 0$:

Proposition 1. *For a firm of type z, the expected informed financing price it gets by choosing cashless payments at $t = 0$ is*

$$p(z) \doteq \frac{\tau_z}{\tau_z + n\tau_s}\mu + \underbrace{\frac{n\tau_s}{\tau_z + n\tau_s}z}_{\textit{Information-revealing}} - \underbrace{\frac{\rho}{2}\frac{1}{\tau_z + n\tau_s}}_{\textit{Risk-reducing}}, \tag{2.2}$$

where $\tau_s = (\tau_x^{-1} + \tau_y^{-1})^{-1}$ captures the overall informational verifiability of cashless payments, which increases in τ_x.

The intuition behind Proposition 1 can be easily seen when comparing the informed price (2.2) to the uninformed price

$$p_\mu \doteq \mu - \frac{\rho}{2}\frac{1}{\tau_z}, \tag{2.3}$$

which is the counterfactual price that the financier offers if no firm establishes verifiable payment records at all. In this uninformed case, the financier prices the firm's technology completely based on its prior. For convenience and use below, we also define the expected price improvement from choosing cashless payments for firm type z as

$$\Delta p(z) = p(z) - p_\mu \,,$$

where $p(z)$ and p_μ are given in (2.2) and (2.3) respectively.

Compared to the uninformed price (2.3), the last two terms of the informed price (2.2) show two effects by establishing verifiable payment records. We elaborate on the two effects below by performing a number of straightforward comparative statics.

The first effect, which we call the *information-revealing effect*, comes from that the information context in the verifiable cashless payments, which is informative about the firm's true type, allows the financier's posterior belief to move closer to the firm's type x. Indeed, the first two terms in (2.2) represent a weighted average of the financier's prior μ and the firm's true type z, compared to the simple prior μ in (2.3). When the firm uses cashless payments for longer in the sense that n is larger, or the cashless payment service is more verifiable in the sense that τ_x is larger, the weight on the firm type z becomes larger, and consequently the informed price $p(z)$ is more reflective of z. In this sense, cashless payments are information-revealing, and the information-revealing effect is stronger when a firm establishes more cashless payment records and when cashless payments are more verifiable. We highlight that whether this information-revealing effect improves the informed price (compared to the uninformed price) depends on the firm type: only higher-than-average firm types enjoy a price improvement through this information-revealing effect.

The second effect, which we call the *risk-reducing effect*, stems from that verifiable cashless payments, regardless of the information context itself (i.e., independent of firm type), also directly reduce the financing risk that the financier has to bear. As shown in the third term in (2.2), more payment records or payment records of higher verifiability, represented by a higher n or a higher τ_x, can directly reduce the variance when the financier makes an inference about the firm type, compared to that in (2.3). The reduced risks thus encourage the financier to bid a higher financing price to the firm. As a result, the overall net effect on the firm's financing price depends on the sum of the information-revealing and the risk-reducing effects.

The results in Proposition 1 and the decomposition of the information-revealing and risk-reducing effects lead to several empirical predictions, which we elaborate on below using a series of corollaries.

First, we are interested in the average effect of using cashless payments on financing outcomes. To capture the average effect, we apply Proposition 1 to the average expected price improvement $E[\Delta p(z)]$:

Corollary **1.** *The average expected price improvement from choosing cashless payments is*

$$E[\Delta p(z)] = \frac{\rho}{2}\left(\frac{1}{\tau_z} - \frac{1}{\tau_z + n\tau_s}\right) > 0$$

which is increasing in n and τ_x.

The intuition behind Corollary 1 centers on the risk-reducing effect. Although the information-revealing effect is on average zero across all firm types, the risk-reducing effect is always positive. Thus, when the firm has more cashless payment records or the records are of higher verifiability, the risk-reducing effect is higher, and so is the overall price improvement. Mapped into our lending context and assuming a lender only grants a loan when the expected price is higher enough to justify the loan origination cost, we have the following prediction:

Prediction **1.** *With more verifiable cashless payment records or payment records of higher verifiability, a firm is more likely to be granted a loan and to enjoy a lower interest rate.*

Second, we are interested in how the effects of using cashless payments vary across firms with different levels of *observed* risks. Recall that in our framework, observed risks are captured by the lender's prior, and we have the following result:

Corollary **2.** *The contribution of using cashless payments to the expected price improvement is higher for firms with a lower prior precision τ_z in the sense that*

$$\frac{\partial^2 E[\Delta p(z)]}{\partial n \partial \tau_z} < 0 \text{ and } \frac{\partial^2 E[\Delta p(z)]}{\partial \tau_x \partial \tau_z} < 0\,.$$

Corollary 2 shows that the contribution of using cashless payments to the expected price improvement is higher for observably riskier firms, for whom the lender's prior is more uncertain. For example, the lender may be more skeptical about firms that have previous credit delinquency or written-off events, and thus the use of cashless payments may better help those firms, on average, to improve their lending outcomes. The intuition of this result again stems from the risk-reducing effect: the marginal informational benefit of a signal is higher when the lender's prior is less precise. In other words, the risk-reducing effect is stronger for observably riskier firms based on any publicly observable signals other than payments. This leads to the following prediction:

PREDICTION **2.** *The effects of more verifiable cashless payment records or payment records of higher verifiability in improving loan approvals and reducing interest rates are stronger for observably riskier firms (e.g., those with higher risks observable through past public credit history), all else equal.*

Having examined the differential effects of cashless payments on firms with different observed risks, we are interested in how the effects of using cashless payments vary when firm type varies. Recall that firm types in our framework capture *residual* risks that are unobservable through public signals such as past credit events, but only potentially inferable through verifiable payment records. We have the following result:

COROLLARY **3.** *A higher firm type enjoys a higher expected informed price from the financier, and the increase is higher when n or τ_x increases in the sense that*

$$\frac{\partial^2 p(z)}{\partial z \partial n} > 0 \text{ and } \frac{\partial^2 p(z)}{\partial z \partial \tau_x} > 0\,.$$

The intuition behind Corollary 3 instead centers on the information-revealing effect. Although the risk-reducing effect is type-independent, the information-revealing effect is stronger for more extreme firm types because it allows those firms to reveal their types more clearly from the average. Particularly, when the firm has more cashless payment records or the records are of higher verifiability, this type-dependent information-revealing effect becomes even stronger. Mapped into our lending context, we have the following prediction:

PREDICTION **3.** *The effects of more verifiable cashless payment records or payment records of higher verifiability in improving loan approvals and reducing interest rates are stronger for firms with lower residual risks that can be only inferable through payment information, all else equal.*

Finally, we explore the effect of cashless payments on efficient capital allocation, which can be proxied by the efficiency of the financier's inference problem. As standard in the statistical literature (e.g., DeGroot, 2005), we capture it by the mean squared error $E[(z - E[z|X])^2]$. We have the following straightforward result:

Corollary 4. *The mean squared error $E[(z - E[z|X])^2]$ of the financier's inference is decreasing in n and τ_x.*

Intuitively, a lower mean squared error as suggested by Corollary 4 means that the lender in our lending context bears lower residual risks in financing the loan at the offered interest rate. It then implies a lower probability of default conditional on a given interest rate, suggesting a more efficient capital allocation. Mapped into our lending context, we have the following prediction.

Prediction 4. *More cashless payment records or payment records of higher verifiability lead to less default, all else equal.*

We highlight that this prediction fundamentally comes from the fact of the financier being risk-averse in our framework. A lower mean squared error of the financier's inference does not imply higher lending profitability in our framework, because competition among financiers always competes any positive profits away in our theoretical framework. Looking forward, we will further clarify this point as we empirically test this prediction.

2.3 Optimal Payment Method Choice

When the financier is risk-averse, the risk-reducing effect of cashless payments creates a direct incentive for borrowers to rely on cashless payments. However, as FinTech lending develops, it is natural to think that the risk-aversion of such lenders will decrease. One implication from Proposition 1 is that as the financier become less risk-averse (potentially because the financier grows larger and more diversified as the development of FinTech lenders in reality), the information-revealing effect will dominate the risk-reducing effect, and hence only high firm types can enjoy an expected price improvement by committing to cashless payments. Does it imply that low firm types would deviate to use cash instead? The answer is no. When firms make the payment method choice decision $x(\cdot)$ endogenously, the financier will rationally update its belief about the firm type based on both the decision and the information content of the realized payment records X, if any. If some low-type firms were to use cash in equilibrium and no payment records were established, the financier would update its belief to reflect that, which would then lead some of those firms to consider establishing payment records instead. This force will unravel and may eventually lead all firm types to adopt cashless payments when the payment records become more verifiable. We formalize this idea in this subsection by fully solving the equilibrium. This equilibrium outcome illustrates the other direction of the synergy between lending and cashless payments.

Formally, we focus on an economic scenario in which the financier's risk-aversion ρ is arbitrarily small, and consider a monotone equilibrium in which there exists a cutoff firm type z^* such that higher firm types $z \geq z^*$ commit to using verifiable cashless payments whereas lower types $z \leq z^*$ use cash, where $z^* = -\infty$ means that all firm types adopt cashless payments.[10] By sequential rationality, when firm z^* chooses cashless payments at $t = 0$, the financier knows that the firm must be of type $z \geq z^*$ at $t = n - 1$ by obtaining its submitted payment records, and by (2.1) the firm's expected financing price from the perspective of $t = 0$ is now given by

$$p(z^*; z \geq z^*) = E[E[z|X, z \geq z^*|z^*] , \tag{2.4}$$

where $p(z; z \geq z^*)$ is a function of z for any non-lower firm type $z \geq z^*$ defined as

$$p(z; z \geq z^*) \doteq E[E[z|X, z \geq z^*|z] , \tag{2.5}$$

Rather, if it chooses cash at $t = 0$ and submits no payment records at $t = n$, the financier then knows that the firm must be of type $z \leq z^*$ at $t = n - 1$. In this case, again by (2.1), the firm's expected financing price at $t = 0$ is given by

$$p(z \leq z^*) \doteq E[z|z \leq z^*] , \tag{2.6}$$

and any lower firm type will get the same expected financing price. Analyzing the firm's expected payoff gain at $t = 0$ by choosing cashless payments over cash, we have the following formal result:

Proposition 2. *In a monotone equilibrium such that firm types $z \geq z^*$ adopt cashless payments while firm types $z \leq z^*$ uses cash, it must be that $z^* = -\infty$, meaning that all firm types will optimally adopt cashless payments.*

[10]Note that the cutoff type z^* is indifferent from the two payment method choices, and we assume it chooses cashless payments in any equilibrium path (although it may deviate in an off-equilibrium path). Also recall that the firm type is defined over the extended real set $\overline{\mathbb{R}}$. Mathematically, this ensures that the type set is compact and thus the lowest type exists.

The intuition of Proposition 2 can be seen from a strategic consideration among firms in the financing market that pushes all firms to adopt cashless payments. Suppose no firm uses cashless payments, in which case the financier's prior belief about the firm type is μ. Then all the better-than-average firms, that is, types $z > \mu$ will deviate to verifiable cashless payments, because the resulting payment information will allow them to be differentiated from the lower-than-average firms and consequently to enjoy a higher financing price. However, as they do so, the financier will rationally update its belief. The financier will now perceive the average of firm types without verifiable payment records to be lower than μ, say, $\nu < \mu$. Thus, better-than-ν firm types will deviate to cashless payments to stand themselves out from even lower firm types. This process unravels until all firm types have adopted cashless payments. Taken together, the idea behind Proposition 2 is reminiscent of the seminal "unraveling argument" of Milgrom (1981) in an information disclosure context.[11,12]

Although Proposition 2 may seem stark as other unmodeled frictions might prevent all firms from adopting cashless payments in reality, it provides a new theoretical perspective for us to understand the increasing popularity of cashless payments and their informational efficiency in the era of big data. Cashless payment service providers offer a convenience service to their users. However, as low-creditworthiness agents understand that the payment records they create are verifiable and potentially assessed by future lenders when screening loan applications, these agents might refuse to use cashless payments despite their convenience and low pecuniary cost. The potential lack of usage by low-creditworthiness agents would limit the market size and informational efficiency associated with these payment services. For this reason, regulators are also introducing policies to directly promote those

[11]In more detail, in a context of firms truthfully and voluntarily disclosing their product quality and disclosing being costless, the firm of the best quality will voluntarily disclose, and thus consumers will interpret no disclosure as indicating that the firm does not have the best quality. But given this, the second-best firm will disclose, followed by the third-best, and so on. This process unravels and all the firms thus disclose in the end. In our context, committing to using cashless payments can be indeed interpreted as committing to disclosing a series of unbiased but noisy signals about the creditworthiness of the firm to a potential financier, and the unraveling mechanism applies. Technically, our contribution here is to extend the Milgrom (1981) mechanism to a context in which the firm cannot truthfully disclose its type directly but only through a series of unbiased but noisy signals.

[12]We note that the literature of information closure is large and there are many other circumstances in which the unraveling mechanism may fail (see Okuno-Fujiwara, Postlewaite, and Suzumura (1990) for a systematic treatment and Ali, Lewis, and Vasserman (2020), Bond and Zeng (2021) for recent applications). We draw the analogy between adopting cashless payments and information disclosure to highlight the informational role of the adoption of cashless payments, rather than suggesting that information must be fully revealed in any lending context. To study the general effect of information disclosure in lending is beyond the scope of this paper.

services and improve their efficiency, such as the PSD2 in Europe. However, Proposition 2 suggests that those policies miss an important general equilibrium force that may ultimately push more prospective borrowers to voluntarily use cashless payments.

Broadly speaking, Propositions 1 and 2 also jointly shed light on the debates of data sharing and open banking. In our framework, firms effectively own the data of their creditworthiness before choosing their payment methods. Sharing their data by committing to using cashless payments that generate transferrable and verifiable information improves lending efficiency, consistent with the benefits of data sharing and open banking. Importantly, the synergy between FinTech lending and cashless payments we uncover suggests that the achievement of wide data sharing and open banking can be self-enforcing. More data sharing improves lending efficiency and leads to better capital allocation, which in turn encourages more data sharing due to the strategic consideration of data owners.

3 Data and Institutional Details

3.1 Data

The empirical analysis of this study relies on novel data provided by one of the largest Indian FinTech lenders, Indifi. The dataset represents the whole information set available to this FinTech lender when screening applications, the screening outcomes, and the loan performance as of September 2019. Indifi is an online lending platform that grants unsecured loans to micro businesses in India. To screen applications, Indifi collects information on loan applicants from several sources: from the application form on their website, from the Indian credit bureau, and from industry partners (e.g. online marketplace) for a subset of applications. Indifi requires applicants to submit the last six monthly statements for the bank account that will be used for receiving and repaying the loan. These bank statements contain transaction-level data. Our data combines all of those types of information for all loan applications received by Indifi from September 2015 to September 2019.

For all applications, we thus have information on the industry, location, the number of years of operations of the business, and the age of the business owner.

As shown in Figure 1 below, providing six months of bank statements is a necessary condition for obtaining a loan from Indifi, and we therefore only keep complete applications, which always include such data. This requirement is important in that it is in line with our theoretical framework, in which entrepreneurs commit to submitting whatever information recorded at the production stage to the lender at the lending stage. In practice, Indifi typically obtains this data by asking loan applicants to link their bank deposit accounts to the Indifi loan application protocol or to directly uploading bank statements in pdf forms. The bank statement data is harmonized by a specialized third party, and is structured at the payment level, and therefore provides the comprehensive payment record for the applicant for the six months before their application. Altogether, we thus observe more than 35 million transactions.

For the majority of applications, our dataset also includes the credit bureau data associated with each application. The data contains the Cibil score, that is, the Indian credit score of the business owner, as well as the start date of their credit history, the number of previous loans and their associated amount and interest rate. This data also includes overdue amounts on existing loans at the time at which Indifi pulled out the credit report of each of the applicants.

Last, we observe Indifi's decision to offer a loan or not, and for which amount and interest rate, and whether it is actually disbursed. We also observe whether the loan is delinquent as of September 2019.

Table 1 provides summary statistics on the application, loan, and borrower characteristics that we use in our empirical analysis. This table documents that applicants are micro to small firms, which rely primarily on cashless payments.[13]

[Insert Table 1 and Figure 1 here]

3.2 Exploiting Bank Statement Data to Classify Payment Types

By conducting text analysis on the payment labels from this dis-aggregated data, we can identify the technology used for more than three quarters of the payments appearing on bank statements. We summarize the key steps of the procedure below, and provide a more detailed elaboration of the procedure in Internet Appendix IB.

We first group identifiable payment records into two broad categories: cash payments, i.e. cash deposit or withdrawal, either at a branch or an ATM, and cashless payments,

[13] The share of transactions done in cash is most likely underestimated, and in turn the share of cashless payment is over-estimated because applicants might receive cash from their clients, which they spend without depositing it on their bank account. The lender faces the same empirical issue. This risk of a measurement error brings an additional motivation for instrumenting the use of cash in our empirical analysis.

which contain any payment for which we can identify the payment technology and that does not belong to the previous category.

We then further classify different types of cashless payments with different levels of verifiability. Specifically, within cashless payments, we further distinguish between information-intensive payments and information-light payments as follows. We classify internet banking transfers and certified check payments as information-intensive and highly verifiable, as there is no aggregation and it is relatively easy to identify the name and type of the payment counter-party. We classify payments through third-party mobile applications, mobile banking and POS machine as information-light payments. Indeed, most popular mobile payment methods in India aggregate payments over a day, which prevents them from being able to identify the counter-party or the motive of a given payment. Moreover, recipient identities are often disguised by specific mobile payment IDs even in singular mobile payments. This means that most mobile payments often have lower verifiability than expected with other verifiable cashless payments.

For each borrower, we aggregate this payment level data by calculating separately for revenues and spending the share of payments conducted in cashless technology, and averaging these two shares, over the six-month period for which we have bank statements.[14]

Table 1 provides summary statistics on the payment records we access, and on the informational proxies we build out of this data. Table A1 in the Appendix further provides the breakdown of cashless share by industries.

3.3 Currency Chest Banks and the 2016 Indian Demonetization

On November 8, 2016, the Government of India announced the demonetization of large banknotes, which needed to be immediately exchanged with banknotes of a new denomination. The action was intended to curtail the shadow economy and reduce the use of illicit and counterfeit cash to fund illegal activity and terrorism. The demonetization created an immediate and prolonged cash shortage in the months that followed, with a potentially lasting impact on cash use (see Chodorow-Reich, Gopinath, Mishra, and Narayanan (2020) and Crouzet, Gupta, and Mezzanotti (2020) for comprehensive analysis on the broad implications of the demonetization). Related to the 2016 Demonetization, a second cash shortage happened in 2018 as the lack of large denomination banknotes put strain on the

[14] We ignore payments whose technology we cannot identify when calculating these shares, meaning that cashless share of payments and cash share of payments add to one.

logistics of banknote distribution.[15]

The cash shortages triggered by demonetization were however less pronounced at the more than 3,000 bank branches that have the status of currency chest bank. Currency chests are branches of selected banks where banknotes and rupee coins are stored on behalf of the Reserve Bank of India for further distribution of these notes and coins in their area of operations. Firms with deposit accounts at a chest bank are therefore more likely to have been able to access the new banknotes, and in turn to conduct relatively more cash payments in the wake of the unexpected demonetization shock compared to those not having a chest bank account, all else being equal. The 2016 Indian demonetization was largely unexpected and thus a plausibly exogenous shock to the use of cashless payments in India, which differently affected bank depositors depending on whether they banked at a chest bank branch or not. Having a deposit account at a chest bank is unlikely to be directly correlated with a firm's (unobserved) creditworthiness given the wide distribution of chest banks in India and the absence of a particular benefit to bank at such a branch absent the unexpected cash shortage. To confirm this point, we provide a comparison of firm characteristics between those with and without a chest bank account in Table A2 in the Appendix, which shows that the two groups of firms are observably similar.

4 Empirical Results

Equipped with the loan application-level data and the proxies for payment information in borrowers' bank statements, we empirically test Predictions 1, 2, 3, and 4 derived from Proposition 1 in Section 2, focusing on the one direction of the synergy from using cashless payments to lending outcomes. We find supportive evidence of the economic mechanism described by our theoretical framework.[16]

[15] See, for instance, https://www.nytimes.com/2018/04/20/business/india-bank-cash-atm.html.

[16] Although Proposition 2 speaks to the long-term trend of cashless payment adoption and is therefore challenging to be empirically tested in our setting, that proposition is also important for our empirical analysis because it suggests that firms' choice of payment methods should be mostly independent to firm creditworthiness in equilibrium. Together with the instrumental variable analysis that we introduce later on, this prediction lends support to a causal interpretation for the correlations between payment technology and lending outcomes we first document. We leave the tests of the other direction of the synergy, that is, the effects of lending on the long-term adoption of cashless payments to future research.

4.1 Baseline Result: Cashless Payments Use and Lending Outcomes

We first empirically test Prediction 1, which results from the risk-reducing effect of using cashless payments. Specifically, we estimate the relationship between the use of cashless payments and lending outcomes and find that a higher share of cashless payments is associated with improved lending outcomes. Throughout our empirical analysis, we consider two dependent variables: an indicator variable for obtaining a loan, and the interest rate paid on the loan conditional on acceptance, which capture the extensive and intensive margin of the lender screening. We use the share of cashless payments out of all payment records, which serves as a baseline proxy for the level of verifiability of the entire profile of payment records (i.e., parameter τ_x in the model), as the main explanatory variable.

For illustration purposes, Figure 2 presents a binned scatterplot of the offered interest rate (conditional on a loan approval) against the share of cashless payments, controlling for application months. It shows a significant negative relationship between the two: when a firm exhibits a larger share of cashless payments in the records it submits when applying for a loan, the lender charges a lower interest rate. This correlation is consistent with the risk reduction channel fleshed out in the model: as the lender can observe a more precise signal of the firm quality, it faces lower uncertainty and in turn can offer a lower interest rate. Indifi maps well into the risk-averse financier we model, as it operates at a relatively low scale and therefore benefits less from diversification of idiosyncratic risks, and faces important balance sheet constraints due to its start-up nature.

[Insert Figure 2 here]

To make the test more precise, we then run regressions of both lending outcomes on the share of cashless payments, with a comprehensive set of granular controls and fixed effects allowing for non-linear relationships as in the following baseline specification:

$$\text{LendingOutcome}_i = \beta \text{CashlessShare}_i + \gamma X_i + \sum_k \theta_{F_k(i)} + \varepsilon_i \,.$$

Table 2 provides the regression results. We include borrower controls, namely the log number of payments, the credit history length, the business vintage, and the log of owner's age fixed effects for industry and application month in all specifications, to address potential composition effects along these dimensions. In columns 2 and 4, we add fixed effects for deciles of revenues to non-linearly control for business size, 3-digit zip-code fixed effects to mitigate potential geographical composition effects, and Cibil score 10-points range fixed effects to non-linearly control for credit score, thereby further mitigating concerns that borrower characteristics potentially correlated with their use of cashless payments are driving the relationship we document. Importantly, because we observe the same information set as the FinTech lender, and because our dependent variable is a decision from that same lender, unobserved variable bias is significantly less likely in our setting than in a standard banking empirical setting. The results robustly show that a higher share of cashless payments corresponds to a significantly higher likelihood of obtaining a loan, and a significantly reduced interest rate compared to the cost of the previous loan.

[Insert Table 2 here]

As shown in Table 2, the economic magnitude of the relationship between cashless payments and lending outcomes is large. Specifically, the estimate regarding loan approval in column 2 with the full set of controls suggests that, everything else equal, an interquartile increase in the share of cashless payments corresponds to a likelihood that is 0.087 * 31 = 2.7 percentage points lower to get a loan approved, which represents more than 10 percent the baseline likelihood of getting approved. In terms of interest rates conditional on loan approval, the estimate in column 4 suggests that an interquartile increase in cashless payments further corresponds to an interest rate that is lower by 1.556 * 27 = 42 basis points.

The estimated results of cashless payments on lending outcomes are robust to alternative proxies for payment verifiability. Recall that Corollary 1 shows that informational verifiability τ_x and the number of verifiable payment records n both contribute to the risk-reducing effect of using cashless payments. Hence, we expect a simply higher number of cashless payments, which proxies for n, to lead to better lending outcomes as well. Thus, we regress the two lending outcomes on the log of the number of cashless payments, with the same full set of controls. Note that we particularly control for the number of total payment records to mitigate the concern that the number of cashless payments may reflect unobserved firm characteristics that are related to firm size.

$$\text{LendingOutcome}_i = \beta \log \text{Cashless}_i + \gamma X_i + \sum_k \theta_{F_k(i)} + \varepsilon_i \,.$$

Table 3 presents the results, which show that a higher number of cashless payments is associated with a significantly higher likelihood of getting a loan, and a significantly lower interest rate conditional on loan approval. These results are consistent with the pattern in Table 2, further supporting Prediction 1.

[Insert Table 3 here]

The granularity of our data allows us to further test Prediction 1 by studying whether different types of cashless payments, varying in their levels of informational verifiability, affect lending outcomes differently. According to Prediction 1, we expect the use of cashless payment technologies with a higher level of verifiability to generate a larger risk-reducing effect, leading to better lending outcomes.

To implement this test, we further break down the share of cashless payments into the share of information-intensive technologies and that of information-light technologies. Information-intensive technologies include for instance online banking and certified checks: each payment record corresponds to a single payment, and the counter-party can be identified. These information-intensive payment technologies thus generate payment records of higher verifiability (i.e., they correspond to a higher τ_x in the model). On the other hand, information-light technologies, such as mobile payment apps, typically aggregate or net several payments within a day, and prevent the counter-party from being identified. Accordingly, we interpret these information-light payment technologies to generate payment records of lower verifiability (i.e., of a lower τ_x in the model).

We repeat the specifications of Table 2 based on this breakdown of cashless payments. Table 4 reports the results. They show that higher use of information-intensive payments leads to a higher likelihood of loan approval and a lower interest rate conditional on a loan approval, while in contrast, higher use of information-light payments leads to a lower likelihood of loan approval and a higher interest rate. The results are both statistically and economically significant in all the specifications.

[Insert Table 4 here]

Overall, our empirical findings are consistent with Prediction 1 and suggest a significant positive impact of using cashless payments on lending outcomes on both the extensive and intensive margins. The relationship between payment technology and lending outcomes appears not only from the use of cashless payments over cash, but also from the use of information-intensive payment technologies over information-light ones.

4.2 Instrumental Variable Analysis

Proposition 2 suggests that firms' payment method choices should be independent of their creditworthiness in the long term. In other words, it suggests that any observed variation in the use of cashless payments is not driven by borrower creditworthiness or a reverse causality from lending outcomes. However, one might still worry that borrowers in reality may not fully internalize the unraveling argument as the model predicts, or that the specification from the previous subsection does not properly control for relationships between cashless payment use and certain firm characteristics that the FinTech lender rely on when screening. To mitigate these concerns, as well as the previously mentioned potential measurement error in the cashless payment share, we implement an instrumental variable analysis based on a unique feature of our data as well as the 2016 Indian Demonetization.

Specifically, we instrument the share of cashless payments with an indicator variable for the borrower banking at a currency chest bank branch after the demonetization. The rationale of the instrument is that because shortages of cash following the demonetization were less pronounced at chest banks, their clients should rely relatively more on cash, and in turn less on cashless payments compared to those who do not bank with a chest bank, everything else equal. Notably, our data allows us to identify the bank branch by combining the bank name with the borrower zip code, which enables identification from within-zip-code variation in treatment.

The construction of our instrumental variable is inspired by and adapted from Chodorow-Reich, Gopinath, Mishra, and Narayanan (2020) and Crouzet, Gupta, and Mezzanotti (2020), which both use cross-district variation to identify the persistent effects of the 2016 Indian Demonetization on payment choices and other broad implications.[17] In contrast to their identifying approaches, a key contribution and difference of our instrumental variable is its identification from within-district variation in treatment, which is enabled by the unique feature of our data. This construction particularly suits our research question of studying the use of cashless payments on lending outcomes at the loan application level, because across-district variation could be correlated with local economic conditions unrelated to payment technology, which would likely violate the exclusion restriction of our instrument.

[17] Generally, this approach relates to a fast-growing macroeconomic literature using cross-regional variation to study macroeconomic topics, as comprehensively discussed in Nakamura and Steinsson (2018) and Chodorow-Reich (2019).

Given our construction, the main identifying assumption is therefore that the matching between small firms and chest bank branches is orthogonal to firm quality that is not adequately controlled for in our regression, but is used by the FinTech lender in its screening process. This assumption is likely to be satisfied due to the unexpected and exogenous nature of the demonetization shock, which makes it unlikely that some small firms of higher quality picked their bank branch in anticipation of this shock. In addition, the quality underlying this selection effect should be orthogonal to the set of controls we include in our specification, which virtually spans the whole information set that the FinTech lender possesses.

We present the regression coefficients for both stages of the two-stage least-square regression in Table 5. In columns 1 and 3, we use an indicator variable for banking at a chest bank branch as the instrument.[18] In columns 2 and 4, we allow the coefficient on our instrument to vary over time by using instead the interaction between the indicator variable for banking at a chest bank branch, and indicator variables for the year of application. This latter specification allows us to capture the time-varying nature of the cash shortages, and to keep applications prior to the demonetization in the sample. Media coverage suggests that cash shortages were particularly acute in 2017 and 2018. This specification also offers a form of placebo test by evidencing that banking at a chest bank is not associated with a lower cashless share in 2016.

The first-stage regression coefficient in column 1 is large and statistically significant: it indicates that having a deposit account at a chest bank increases the use of cash payments by 3 percentage points on average. The F-stat for the first stage is larger than 100, significantly above the threshold for strong instruments. The second-stage result is consistent with the OLS analysis: based on this plausibly exogenous variation in the use of cash payments at the borrower level, we find that more use of cash payments leads to a significantly lower likelihood of loan approval. The economic magnitude is also large. An interquartile increase of the share of cash payments, which represents 31 percentage points in our sample, is associated with approximately an increase of $0.8 * 31 = 24.8$ percentage points in the chance of getting a loan approved, and it is statistically significant as well. Turning to the event-study type of instrumentation, we observe that the coefficient on the interaction

[18] Because Indifi was created in 2015, the number of applications prior to the demonetization is limited, which prevents us from running a classic differences-in-difference specification. We therefore exclude these applications from the sample when running these regressions.

between banking at a chest bank branch and the year 2016 is not significant, which provides a sort of placebo test on the sorting. In turn, the coefficients post demonetization are all significant, and particularly so for 2018 that represented the second cash shortage. The second stage results are also consistent with the previous specification. Such analysis thus lends support to a causal interpretation of the test we previously run for Prediction 1.

[Insert Table 5 here]

A related concern for our instrumental variable analysis is that the local average treatment effect differs from the true causal effect, for instance because only a specific type of firms keep using cash due to their access to chest banks during the demonetization. This concern is mitigated by the relative homogeneity of the firms we are studying: micro to small firms which have a bank account. Our sample thus does not cover informal activity, nor sophisticated start-ups, where the relation between payment technology and creditworthiness might be particularly strong for reasons outside of the informational mechanism we study.

4.3 Cashless Payments, Firm Risks, and Lending Outcomes

We next test Predictions 2 and 3 in regards to how firm risks interact with the effects of cashless payments on lending outcomes. We ask whether riskier firms may benefit more from the use of cashless payments. Informed by the theoretical framework, we are able to explicitly distinguish between the two types of risks: publicly observable risk, such as previous credit events that are observable in the credit bureau data, and residual risks that can only be inferred from verifiable payment records. This distinction is important because Predictions 2 and 3 speak to effects in opposite directions: the effects of using cashless payments on lending outcomes are stronger for firms with higher observed risks due to the risk-reducing effect, but stronger for firms with lower residual risks due to the information-revealing effect. To empirically distinguish the two types of risks is thus important for us to more clearly pinpoint the role of cashless payments in affecting lending outcomes.

Accordingly, we construct proxies for both observable and residual risks. We first use whether an applicant firm has any previous delinquencies, or loans being written-off to proxy for its observable risk, which correspond to τ_z in the model, i.e. the precision of the lender's prior before seeing any payment records. Specifically, having previous delinquencies or write-offs corresponds to a low τ_z.

We then use the volatility of payment inflows of an applicant firm, measured at a weekly frequency, to proxy for its residual risk.[19] By construction, the volatility of payment revenues directly measures the risk of a firm production process, and thus serves as a proxy for the actual firm type z in the model. Specifically, lower volatility corresponds to a higher z. We highlight that the Indian credit bureau does not observe nor share firms' payment revenues, nor its volatility. Thus, this proxy indeed captures risk that is not publicly observable.

We first illustrate the relationship between the share of cashless payments and the offered interest rate by the two types of risk in Figures 3 and 4, simply controlling for application month fixed effects. Figure 3 shows that conditional on a loan approval, a firm that has previous loan write-offs (i.e., with higher observable risk) benefits more from using cashless payments: a higher share of cashless payments leads to a larger reduction in interest rate than for firms that do not have such events. In contrast, Figure 4 shows that a firm that has below-median volatility of payment inflows (i.e., with lower residual risk) benefits more from using more cashless payments than firms with above-median volatility of payment inflows. Those two trends are consistent with Predictions 2 and 3 and are supportive of the risk-reducing effect and the information-revealing effect of cashless payments, respectively.

[Insert Figures 3 and 4 here]

We turn to formal regressions focusing on the roles of the two types of risk with our full set of comprehensive controls. First, we interact the share of cashless payments with our two proxies for observable firm risk, using the same specification as in Table 2. Table 6 reports the results. Panel A shows that the marginal effect of a higher cashless share on the likelihood of a loan being approved is higher for firms with previous credit events, and Panel B shows that, conditional on a loan approval, the marginal effect of a higher cashless share on reducing the offered interest rate is also higher for those observably riskier firms. Notably, all the results are both economically and statistically significant, consistent with Prediction 2 and the risk-reducing effect being stronger for observably riskier firms.

[Insert Table 6 here]

We then interact the share of cashless payments with residual firm risk, proxied by the

[19] We obtain comparable results if we use monthly frequency.

volatility of payment inflows, also using the full set of controls and fixed effects:

$$\begin{aligned}\text{LendingOutcome}_i &= \beta_1 \text{CashlessShare}_i + \beta_2 \text{RevenueVol}_i \\ &+ \beta_3 \text{CashlessShare}_i \times \text{RevenueVol}_i + \gamma X_i + \sum_k \theta_{F_k(i)} + \varepsilon_i \,.\end{aligned}$$

Results are displayed in Table 7. The findings are consistent with Prediction 3 and the information-revealing effect: using cashless payments is particularly beneficial to firms with lower residual risks. Specifically, for applicant firms with lower revenue volatility, the estimate of the interaction term in column 1 suggests that a higher share of cashless payments leads to an even higher likelihood of loan approval than for those with higher revenue volatility. This relationship is unlikely to be driven by other firm-level characteristics given the comprehensive set of fixed effects that we include. Similarly, column 2 suggests that this differential effect is also present at the intensive margin. Specifically, a higher share of cashless payments leads to an even lower interest rate for firms with lower revenue volatility, and the result is even more statistically significant.

[Insert Table 7 here]

To further show the robustness of the information-revealing effect, we perform a similar test following the specifications in Table 4, taking advantage of our unique data of different payment technologies with different levels of verifiability intensity. Table 8 displays the regression coefficients. For higher-type applicant firms, i.e. with lower inflow volatility, higher use of information-intensive cashless payments increases the likelihood of loan approval and decreases the offered interest rate even more.

[Insert Table 8 here]

Taken together, the results above provide comprehensive evidence on how cashless payment use, by documenting transferable and verifiable information about borrower creditworthiness that is otherwise not available from publicly available sources, and that is costly to manipulate, significantly affects the lending outcomes at an outside lender that has access to this data. Such impact appears to take effect through two channels: the risk-reducing and information-revealing effects of accessing and exploiting such data. These findings thus directly speak to the motives behind the recent policy initiatives that aim to promote data sharing and open banking, such as the PSD2. Notably, our findings show that what matters for lending outcomes is not only bank statement aggregates that can be potentially opened. Rather, the detailed components of these statements matter in terms of how the quantity and the verifiability of the information such statements contain. Thus, observing the share of cashless payments ensures that the records are verifiable, and allows for building trustworthy proxies for risk from these payment records.

4.4 Cashless Payments and Defaults

We move on to empirically test Prediction 4 by relating the use of cashless payments to loan defaults. If cashless payments lead to more efficient loan screening and pricing, we should expect that a borrower's credit risk is better priced in at loan issuance, which bears implications for the efficiency of capital allocation. Because the lender makes fewer false positive errors when lending to borrowers that rely more on cashless payments, when controlling for interest rates, our model predicts that the loan default rate should be lower for such borrowers.[20]

To test this prediction, we regress an indicator variable of a loan having defaulted as of September 2019 in our sample on the share of cashless payments with our standard set of controls, which includes fixed effects for loan starting month, augmented with fixed effects for the interest rate level.

$$\text{Default}_i = \beta \text{CashlessShare}_i + \gamma X_i + \sum_k \theta_{F_k(i)} + \varepsilon_i \,.$$

Table 9 reports the results. Consistent with Prediction 4, we find that a higher share of cashless payments leads to a lower likelihood of loan default, conditional on an interest rate. This finding suggests that the use of cashless payments leads to more efficient screening and lending decisions.

[Insert Table 9 here]

The results in Table 9 also suggest that the lender may not fully exploit the information context of cashless payments in pricing loans. Because the share of cashless payments still

[20] In the model, the financier is competitive and therefore makes zero profit. The gains resulting from the lower number of false positive, i.e. including entrepreneurs from the left tail of the distribution of quality into a given price bucket, is indeed offset by reducing the gains it is making from the entrepreneurs from the right tail of the distribution it includes into this price bucket.

negatively predicts loan default after controlling for the interest rate, the lender is making a larger profit with these borrowers, and could have further reduced the offered interest rate it offers them. In reality, lenders not fully exploiting verifiable payment records is likely to happen for various reasons. First, consistent with our model, if the lender is effectively risk-averse (e.g., due to capital constraints), the optimal financing price always includes a discount as shown in equation (2.2) in Proposition 1, which corresponds to a premium in the offered interest rate in reality. Although the risk-reducing effect may lead to a higher financing price (i.e., a lower interest rate) as suggested by (2.2), it will never push the premium to zero in equilibrium. Second, although our model features a competitive financier, the lending market is unlikely to be perfectly competitive in reality. If the lender has market power, it may keep charging a higher interest rate even if it fully processes the information context it possesses. We highlight that although we view both reasons above to be plausible in reality, neither of them would weaken the key message underlying our findings that verifiable payment information leads to more efficient lending and capital allocation outcomes.

4.5 Alternative Verifiable Information and Lending Outcomes

Finally, we show that our findings can be generalized to other easily accessible verifiable data beyond cashless payments. To this end, we take advantage of another unique feature of our data. For a subset of borrowers, Indifi also has access to applicant firms' verifiable sales data on a set of partner online marketplaces. These data are different from the payment records that the applicant firms submit through their bank statements, but are conceptually similar as both are directly generated by the economic activity of the applicant firms. They are also similar in terms of their informational verifiability and informativeness about firm creditworthiness. Such data from marketplaces is also comparable to the data that BigTech firms such as Amazon and Alibaba rely on to make their lending decisions.

Using these third-party merchant sales data, we run regressions of lending and capital allocation outcomes on the share of merchant sales. Table 10 reports the results, which repeat our previous specifications with the full set of controls, as well as controlling for the share of cashless payments.[21] The results show that on top of the informational role of cashless payments, a higher share of verifiable merchant sales further and significantly

[21] In India, transactions on online marketplaces are not necessarily paid through cashless payments, as a significant share of transactions are paid in cash at delivery.

increases the likelihood of a loan being approved, and significantly predicts a lower chance of loan default.

[Insert Table 10 here]

Table 10 thus lends support to the external validity of our results. It echoes the earlier message we convey that cashless payments, although being economically important, are not the sole source of outside information in facilitating lending decisions. Rather, the key is that such outside information is verifiable and transferrable.

5 Conclusion

We provide a new perspective to understand the joint rise of FinTech lending and cashless payments. We uncover both theoretically and empirically a synergy between FinTech lenders and cashless payment providers, the latter producing borrower information outside the lender. In one direction, FinTech lenders become more efficient in screening high-versus low-quality borrowers when borrowers adopt cashless payments that produce more verifiable information. In the other direction, because would-be borrowers expect lenders to rely on outside verifiable payment information to screen them, a strategic consideration for a borrower to stand out of worse borrowers emerges, which ultimately pushes all borrowers to adopt cashless payments.

We use novel data from a large FinTech lender in India to test our predictions. We find that higher use of cashless payments is associated with improved borrowing outcomes: applicants relying heavily on cashless payments are more likely to obtain a loan, and when doing so obtain a lower interest rate. Consequently, such borrowers get a significantly lower rate from the FinTech lender than from previous loans with traditional institutions. This benefit is particularly pronounced for cashless payment users that present a low level of risk that is not publicly observed, due to the combination of both risk-reducing and information-revealing effects of verifiable payment records. These relationships appear to be more pronounced when focusing on more verifiable cashless payments. We also find that within loans charging the same interest rate, borrowers that use more cashless payments are less likely to default.

The mechanism we uncover provides a new perspective for the hand-in-hand rise of both FinTech lending and cashless payments that appear to be self-reinforcing, and also suggests an alternative banking model without a balance sheet and without relationships in the traditional sense. This mechanism further sheds light on the recent policy developments in data sharing and open banking. It first provides direct support to data sharing and open banking in term of improving lending efficiency. Notably, our findings show that what matters for lending outcomes is not bank statements themselves. Rather, the detailed components matter in terms of how much information is verifiable, as we capture by the relative share of cash versus cashless payment records. It also implies that even in the absence of policies to promote data sharing or open banking such as the Second Payment Services Directive (PSD2) in Europe, borrowers may voluntarily commit to data sharing in order to improve their outcome on the lending markets.

References

Agarwal, S., S. Alok, P. Ghosh, and S. Gupta. 2021. Financial inclusion and alternate credit scoring: role of big data and machine learning in Fintech. Working paper.

Ali, N., G. Lewis, and S. Vasserman. 2020. Voluntary disclosure and personalized pricing. *Review of Economic Studies*, forthcoming.

Allen, F., X. Gu, and J. Jagtiani. 2021. A survey of Fintech research and policy discussion. *Review of Corporate Finance* 1:259-339.

Balyuk, T. 2017. Financial innovation and borrowers: Evidence from peer-to-peer lending. Working paper.

Beaumont, P., H. Tang, and E. Vansteenbergh. 2021. The role of FinTech in small business lending. Working paper.

Berlin, M. and L. J. Mester. 1999. Deposits and relationship lending. *Review of Financial Studies* 12:579-607.

Berg, T., Burg, V., Gombovic, A. and Puri, M. 2020. On the rise of FinTechs: credit scoring using digital footprints. *Review of Financial Studies* 33:2845-2897.

Bergemann, D., A. Bonatti, and T. Gan. 2020. The economics of social data. Working paper.

Bianchi, J. and S. Bigio. 2021. Banks, liquidity management, and monetary policy. *Econometrica*, forthcoming.

Black, F. 1975. Bank funds management in an efficient market. *Journal of Financial Economics* 2:323-339.

Bond, P. and Y. Zeng. 2021. Silence is safest: information disclosure when the audience's preferences are uncertain. *Journal of Financial Economics*, forthcoming.

Brunnermeier, M., H. James, and J. Landau. 2019. The digitalization of money. Working Paper.

Buchak, G., G. Matvos, T. Piskorski, and A. Seru. 2018. FinTech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics* 130:453-483.

Chodorow-Reich, G. 2019. Geographic cross-sectional fiscal multipliers: what have we Llearned? *American Economic Journal: Economic Policy* 11:1-36.

Chodorow-Reich, G., Gopinath, G., Mishra, P. and Narayanan, A. 2020. Cash and the economy: Evidence from India's demonetization. *Quarterly Journal of Economics* 135:57-103.

Claessens, S., J. Frost, G. Turner and F. Zhu. 2018. Fintech credit markets around the world: size, drivers and policy issues. *BIS Quarterly Review* September 2018.

Crouzet, N., A. Gupta, and F. Mezzanotti. 2020. Shocks and technology adoption: Evidence from electronic payment systems. Working paper.

Diamond, D. 1991. Monitoring and reputation: the choice between bank loans and directly placed debt. *Journal of Political Economy* 99:698-721.

DeGroot, D. 2005. *Optimal statistical decisions.* John Wiley & Sons.

Darrell, D. 2019. Digital currencies and fast payment systems: disruption is coming. Asian Monetary Policy Forum.

Erel, I. and J. Liebersohn. 2020. Does FinTech substitute for banks? Evidence from the paycheck protection program. Working paper.

Fama, E. 1985. What's different about banks? *Journal of Monetary Economics* 15:29-39.

Farboodi, M., and L. Veldkamp. 2020a. Long run growth of financial technology. *American Economic Review* 110:2485-2523.

Farboodi, M., and L. Veldkamp. 2020b. A growth model of the data economy. Working paper.

Frost, J., L. Gambacorta, Y. Huang, H. Shin, and P. Zbinden. 2020. BigTech and the changing structure of financial intermediation. *Economic Policy* 34:761-799.

Fishman, M., and J. Parker. 2015. Valuation, adverse selection, and market collapses. *Review of Financial Studies* 28:2575–607.

Fuster, A., P. Goldsmith-Pinkham, T. Ramadorai, and A. Walther. 2020. Predictably unequal? The effects of machine learning on credit markets. *Journal of Finance*, forthcoming.

Fuster, A., M. Plosser, P. Schnabl, and J. Vickery. 2019. The role of technology in mortgage lending. *Review of Financial Studies* 32:1854-1899.

Gopal, M. and P. Schnabl. 2020. The rise of finance companies and FinTech lenders in small business lending. Working paper.

Gordon, R. 1941. Values of Mills' ratio of area to bounding ordinate and of the normal probability integral for large values of the argument. *Annals of Mathematical Statistics* 12:364-366.

He, Z., J. Huang, and J. Zhou. 2020. Open banking: credit market competition when borrowers own the data. Working paper.

Higgins, S. 2019. Financial technology adoption. Working paper.

Ichihashi, S. 2020. Online privacy and information disclosure by consumers. *American Economic Review* 110:569-595.

Jack, W. and T. Suri. 2014. Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution. *American Economic Review* 104:183-223.

Jones, C. and C. Tonetti. 2020. Nonrivalry and the economics of data. *American Economic Review* 110:2819-2858.

Kirpalani, R. and T. Philippon. 2020. Data sharing and market power with two-sided platforms. Working paper.

Iyer, R., A. Khwaja, E. Luttmer, and K. Shue. 2015. Screening peers softly: Inferring the quality of small borrowers. *Management Science* 62:1554-1577.

Liberti, J. and M. Petersen. 2019. Information: hard and soft. *Review of Corporate Finance Studies* 8:1-41.

Liu, X., M. Sockin, and W. Xiong. 2020. Data privacy and temptation. Working paper.

Mester, L., L. Nakamura, and M. Renault. 2007. Payments accounts and loan monitoring. *Review of Financial Studies* 20: 529-556.

Milgrom, P. 1981. Good news and bad news: representation theorems and applications. *Bell Journal of Economics* 12:380-391.

Muralidharan, K., P. Niehaus, and S. Sukhtankar. 2016. Building state capacity: Evidence from biometric smartcards in India. *American Economic Review* 106: 2895-2929.

Nakamura, E. and J. Steinsson. 2018. Identification in macroeconomics. *Journal of Economic Perspectives* 32:59-86.

Norden, L. and M. Weber. 2010. Credit line usage, checking account activity, and default risk of bank borrowers. *Review of Financial Studies* 23:3665-3699.

Okuno-Fujiwara, M., A. Postlewaite, and K. Suzumura. 1990. Strategic information revelation. *Review of Economic Studies* 57:25-47.

Parlour, C., U. Rajan, and H. Zhu. 2020. When FinTech competes for payment flows. Working paper.

Piazzesi, M. and M. Schneider. 2021. Payments, credit and asset prices. Working Paper.

Philippon, T. 2019. On Fintech and financial inclusion. Working paper.

Puri, M., J. Rocholl, and S. Steffen. 2017. What do a million observations have to say about loan defaults? Opening the black box of relationships. *Journal of Financial Intermediation* 31:1-15.

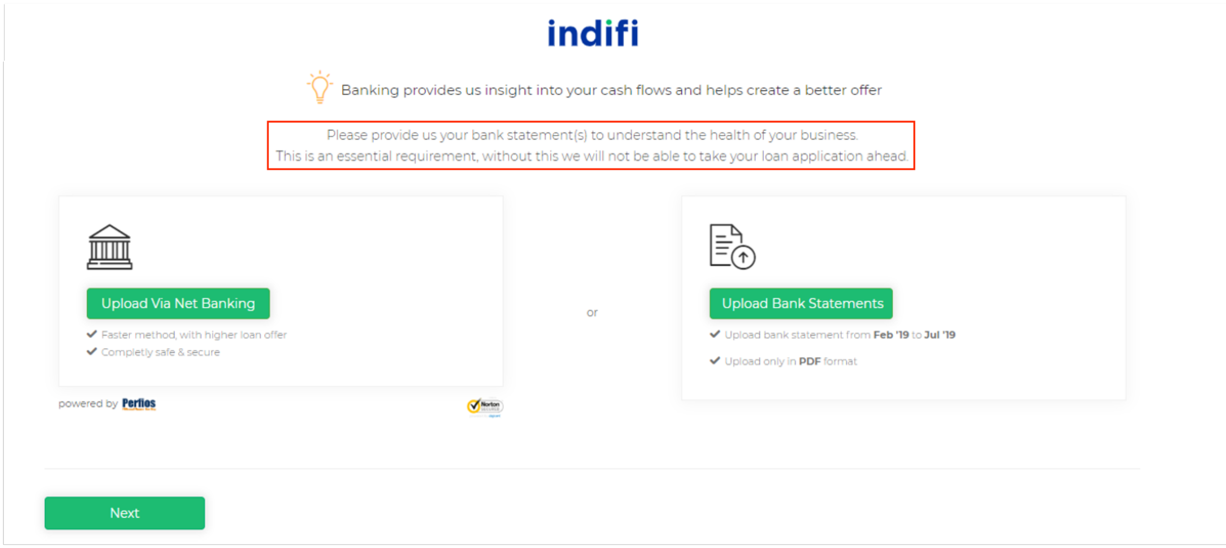
Rajan, R. 1992. Insiders and outsiders: the choice between informed and arm's-length debt. *Journal of Finance* 47:1367-1400.

Tang, H. 2019. Peer-to-peer lenders versus banks: substitutes or complements? *Review of Financial Studies* 32:1900-1938.

Vallee, B., and Y. Zeng. 2019. Marketplace lending: a new banking paradigm? *Review of Financial Studies* 32:1939-1982.

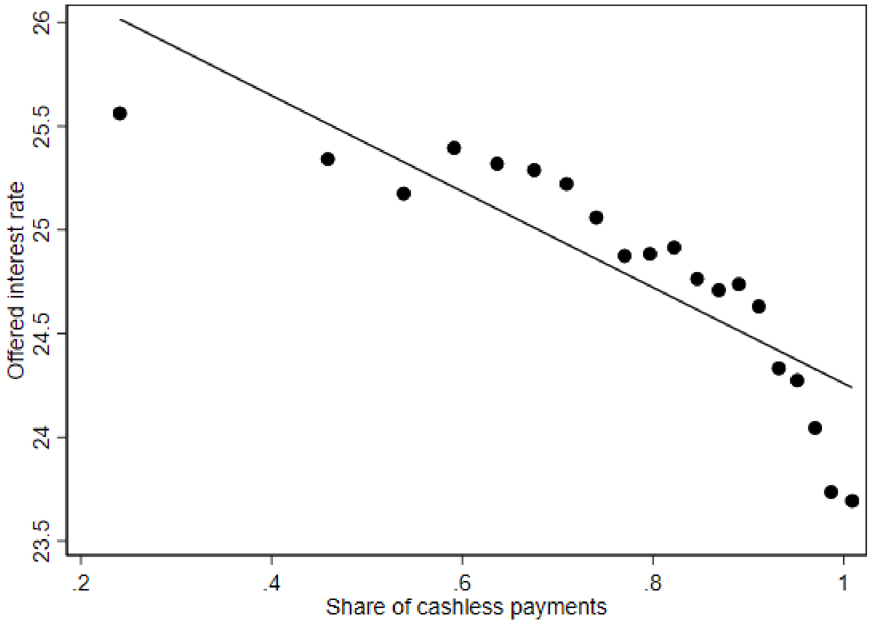
Vives, X. 2019. Digital disruption in banking. *Annual Review of Financial Economics* 11:243-272.

Figure 1: **Indifi: the application platform**



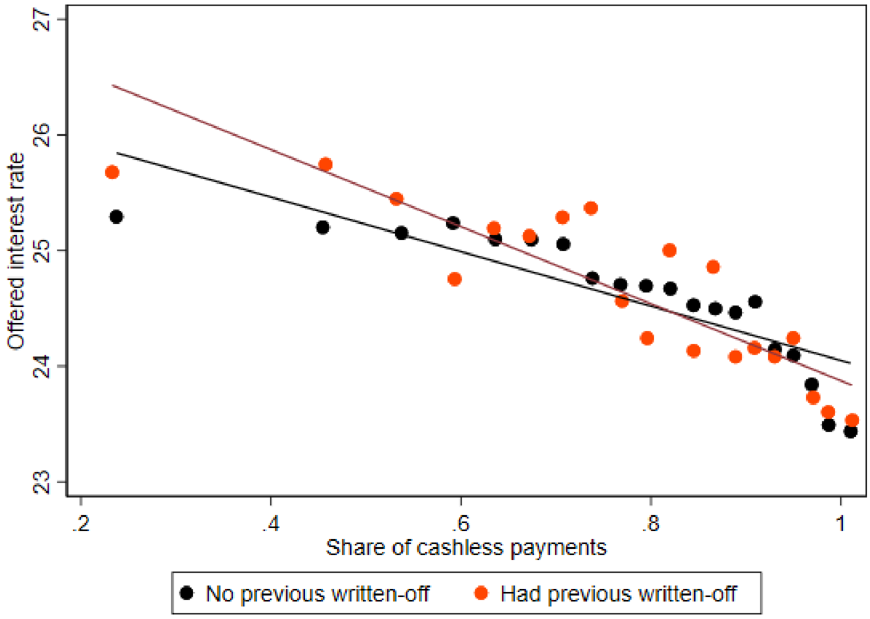
Notes: This figure shows the webpage on which Indifi mandatorily requires loan applicants to submit bank statements, which consist of payment records of different levels of verifiability. It also shows that loan applicants can submit bank statements by either linking the bank account or directly uploading bank statements in pdf forms.

Figure 2: **Payment verifiability and lending outcomes**



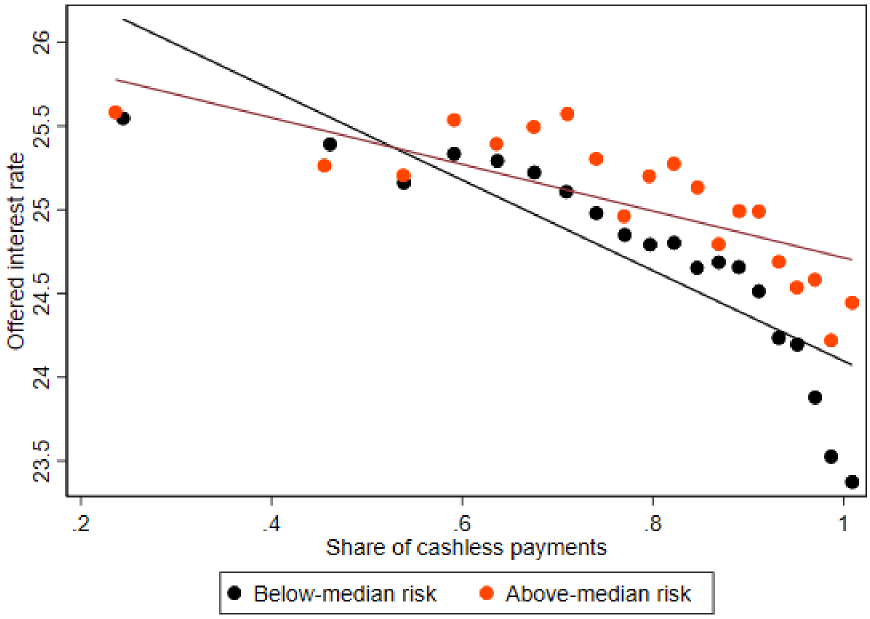
Notes: This figure reports the 20-binned scatterplot of the offered interest rate conditional on a loan approval against the share of cashless payments, controlling for application-month fixed effects.

Figure 3: **Payment verifiability, observed risks, and lending outcomes**



Notes: This figure reports the 20-binned scatterplot of the offered interest rate conditional on a loan approval against the share of cashless payments by whether the applicant firms have any previous bank loan written-off events, controlling for application-month fixed effects.

Figure 4: **Payment verifiability, residual risks, and lending outcomes**



Notes: This figure reports the 20-binned scatterplot of the offered interest rate conditional on a loan approval against the share of cashless payments by the weekly volatility of payment revenues, controlling for application-month fixed effects.

Table 1: **Summary statistics**

	Obs	**Mean**	**p25**	**p75**
	Panel A. Loan Applicant Characteristics			
Year of Application	86,447	2018	2018	2019
Applicant Age	86,443	35.3	29.1	40.4
Business Age	85,916	4.8	1.4	5.9
Cibil Score	86,447	534	562	715
Credit History Length (in Years)	63,433	7.0	2.4	11.1
	Panel B. Banking Statement Data			
Share of Cashless Payments	86,349	0.717	0.585	0.894
of which: Information-intensive Payments	86,349	0.541	0.341	0.761
Share of Cash Payments	86,349	0.283	0.106	0.415
Borrower Banks at Chest Banks (0/1)	86,447	0.039	0	0
Avg. Monthly Revenue (INR)	86,447	1,615,696	114,154	823,622
Avg. # of Transactions	86,447	554	155	600
	Panel C. Loan Application Outcomes			
Approved Loan (0/1)	86,447	0.252	0	1
Offered Interest Rate	21,746	24.8	23.4	27
Default (0/1)	9,138	0.073	0	0

Notes: This table reports summary statistics for main variables used in the regressions. Panel A reports summary statsitics of loan applicants by years of application, age, credit score and credit history. Panel B displays summary statistics of banking payments of applicants by payment technologies, bank type and monthly average revenue. Panel C reports loan consequences of the applications with available banking records.

Table 2: **Payment verifiability and lending outcomes**

	Approved Loan (1/0)		Offered Interest Rate	
	(1)	(2)	(3)	(4)
Share of cashless payments	0.109***	0.087***	-1.718***	-1.556***
	(0.010)	(0.010)	(0.142)	(0.142)
Borrower controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Application month FE	Yes	Yes	Yes	Yes
Cibil score group FE	No	Yes	No	Yes
3-digit zip code FE	No	Yes	No	Yes
Revenue deciles FE	No	Yes	No	Yes
Observations	53,042	52,696	17,606	17,548
R^2	0.259	0.310	0.446	0.489

Notes: This table presents OLS regressions that use the share of cashless payments to predict loan consequences. The set of controls includes log # of payments, credit history length, business vintage, Log of owner's age, missing credit score indicator and top-up loan indicator. Standard errors are clustered at application-month level.

Table 3: **Number of verifiable payments and lending outcomes**

	Approved Loan (1/0) (1)	Offered Interest Rate (2)
log(# of cashless payments)	0.015***	-0.159***
	(0.004)	(0.044)
Borrower controls	Yes	Yes
Industry FE	Yes	Yes
Application month FE	Yes	Yes
Cibil score group FE	Yes	Yes
3-digit zip code FE	Yes	Yes
Revenue deciles FE	Yes	Yes
Observations	52,726	17,550
R^2	0.309	0.483

Notes: This table presents OLS regressions that use the number of payments and number of cashless payments to predict loan consequences. The set of controls includes log # of payments, credit history length, business vintage, Log of owner's age, missing credit score indicator and top-up loan indicator. Standard errors are clustered at application-month level.

Table 4: **Intensity of payment verifiability and lending outcomes**

	Approved Loan (1/0)		Offered Interest Rate	
	(1)	(2)	(3)	(4)
Share of information-intensive payments	0.154***	0.125***	-1.449***	-1.693***
	(0.007)	(0.009)	(0.106)	(0.141)
Share of information-light payments		-0.064***		-0.583***
		(0.012)		(0.179)
Borrower controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Application month FE	Yes	Yes	Yes	Yes
Cibil score group FE	Yes	Yes	Yes	Yes
3-digit zip code FE	Yes	Yes	Yes	Yes
Revenue deciles FE	Yes	Yes	Yes	Yes
Observations	52,696	52,696	17,548	17,548
R^2	0.315	0.316	0.492	0.492

Notes: This table presents OLS regressions that use the share of payments of different level of information intensiveness and its interaction with borrower's risk measure to predict loan consequences. The set of controls includes log # of payments, credit history length, business vintage, Log of owner's age, missing credit score indicator and top-up loan indicator. Standard errors are clustered at application-month level.

Table 5: **IV: Cashless payment use and loan approval**

	Share of Cashless Payments		Approved Loan (1/0)	
	(1)	(2)	(3)	(4)
Borrower banks at Chest Bank (1/0)	-0.028***			
	(0.004)			
Borrower banks at Chest Bank (1/0)× 2016		-0.005		
		(0.031)		
Borrower banks at Chest Bank (1/0)× 2017		-0.023*		
		(0.012)		
Borrower banks at Chest Bank (1/0)× 2018		-0.030***		
		(0.006)		
Borrower banks at Chest Bank (1/0)× 2019		-0.023***		
		(0.007)		
Share of cashless payments			0.827**	0.762**
			(0.305)	(0.290)
Borrower controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Application month FE	Yes	Yes	Yes	Yes
Cibil score group FE	Yes	Yes	Yes	Yes
Revenue deciles FE	Yes	Yes	Yes	Yes
Observations	51,396	53,038	51,396	53,038
F-statistic	107.030	91.309		

Notes: This table presents regression coefficients for a 2SLS specification. Column 1 and 2 presents the first stage, where we regress the share of cashless payments on an indicator variable for the borrower banking at a chest bank. Column 3 and 4 displays the second stage, where we regress the indicator variable for getting the loan application approved on the instrumented *Share of Cashless Payments*.

Table 6: **Payment verifiability, observed risks, and lending outcomes**

	(1)	(2)	(3)	(4)
	Previous delinquency		Previous written-off	
	No	Yes	No	Yes
	Panel A. Approved Loan (1/0)			
Share of cashless payments	0.065***	0.106***	0.085***	0.149***
	(0.013)	(0.012)	(0.011)	(0.039)
Observations	19,200	31,782	47,400	3,560
R^2	0.353	0.298	0.315	0.337
	Panel B. Offered Interest rate			
Share of cashless payments	-1.318***	-1.862***	-1.678***	-2.218***
	(0.237)	(0.167)	(0.156)	(0.402)
Borrower controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Application month FE	Yes	Yes	Yes	Yes
Cibil score group FE	Yes	Yes	Yes	Yes
3-digit zip code FE	Yes	Yes	Yes	Yes
Revenue deciles FE	Yes	Yes	Yes	Yes
Observations	5,813	11,233	15,916	1,123
R^2	0.459	0.521	0.481	0.691

Notes: The set of controls includes log # of payments, credit history length, business vintage, Log of owner's age, missing credit score indicator and top-up loan indicator. Standard errors are clustered at application-month level.

Table 7: **Payment verifiability, residual risks, and lending outcomes**

	Approved Loan (1/0) (1)	Offered Interest Rate (2)
Share of cashless payments	0.095***	-2.111***
	(0.013)	(0.222)
Share of cashless payments × Weekly revenue volatility	-0.152*	12.983***
	(0.088)	(3.165)
Weekly revenue volatility	-0.505***	-11.428***
	(0.069)	(2.724)
Borrower controls	Yes	Yes
Industry FE	Yes	Yes
Application month FE	Yes	Yes
Cibil score group FE	Yes	Yes
3-digit zip code FE	Yes	Yes
Revenue deciles FE	Yes	Yes
Observations	52,695	17,548
R^2	0.312	0.490

Notes: This table presents OLS regressions that use the share of cashless payments interacted with borrower's risk measure to predict loan consequences. The set of controls includes log # of payments, credit history length, business vintage, Log of owner's age, missing credit score indicator and top-up loan indicator. Standard errors are clustered at application-month level.

Table 8: **Intensity of payment verifiability, residual risks, and lending outcomes**

	Approved Loan (1/0) (1)	Offered Interest Rate (2)
Share of information-intensive payments	0.152***	-2.220***
	(0.012)	(0.212)
Share of information-light payments	-0.066***	-0.669***
	(0.011)	(0.186)
Share of information-intensive payments × Weekly revenue volatility	-0.411***	12.250***
	(0.111)	(2.808)
Weekly revenue volatility	-0.455***	-9.668***
	(0.072)	(2.333)
Borrower controls	Yes	Yes
Industry FE	Yes	Yes
Application month FE	Yes	Yes
Cibil score group FE	Yes	Yes
3-digit zip code FE	Yes	Yes
Revenue deciles FE	Yes	Yes
Observations	52,695	17,548
R^2	0.318	0.493

Notes: This table presents OLS regressions that use the share of payments of different level of information intensiveness and its interaction with borrower's risk measure to predict loan consequences. The set of controls includes log # of payments, credit history length, business vintage, Log of owner's age, missing credit score indicator and top-up loan indicator. Standard errors are clustered at application-month level.

Table 9: **Payment verifiability and loan default**

	Default (1/0)	
	(1)	(2)
Share of cashless payments	-0.076***	-0.057***
	(0.013)	(0.015)
Borrower controls	Yes	Yes
Industry FE	Yes	Yes
Application month FE	Yes	Yes
Cibil score group FE	No	Yes
3-digit zip code FE	No	Yes
Revenue deciles FE	No	Yes
Interest Rate FE	Yes	Yes
Observations	9,084	8,139
R^2	0.055	0.115

Notes: This table presents OLS regressions that use the share of cashless payments to predict loan default rate. The set of controls includes log # of payments, credit history length, business vintage, Log of owner's age, missing credit score indicator and top-up loan indicator. Standard errors are clustered at application-month level.

Table 10: **Alternative verifiable information and lending outcomes**

	Approved Loan (1/0) (1)	Offered Interest Rate (2)	Default (1/0) (3)
Share of merchant sales	0.327***	0.033	-0.040***
	(0.012)	(0.152)	(0.011)
Share of cashless payments	0.097***	-1.585***	-0.059***
	(0.009)	(0.144)	(0.015)
Borrower controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Application month FE	Yes	Yes	Yes
Cibil score group FE	Yes	Yes	Yes
3-digit zip code FE	Yes	Yes	Yes
Revenue deciles FE	Yes	Yes	Yes
Interest rate FE	No	No	Yes
Observations	52,696	17,548	8,114
R^2	0.307	0.472	0.116

Notes: The set of controls includes log # of payments, credit history length, business vintage, Log of owner's age, missing credit score indicator and top-up loan indicator. Standard errors are clustered at application-month level.

Appendix

A Proofs

PROOF OF PROPOSITION 1. When the firm adopts verifiable cashless payments, the financier makes inference about the firm type based on the established and submitted payment records X. We first note that

$$x_t|z \sim N\left(z, \tau_s^{-1}\right) ,$$

where $\tau_s \doteq (\tau_x^{-1} + \tau_y^{-1})^{-1}$ captures the effective overall informational verifiability of each payment record. By Bayesian updating, we then have

$$z|X \sim N\left(\frac{\tau_z \mu + n\tau_s \bar{x}}{\tau_z + n\tau_s}, \frac{1}{\tau_z + n\tau_s}\right) ,$$

where $\bar{x} = \frac{1}{n}\sum_{t=0}^{n-1} x_t$ is the sample mean of X. On the hand, note that $E[\bar{x}|z] = E[x_t|z] = z$. We then immediately have

$$E[E[z|X]|z] = \frac{\tau_z \mu + n\tau_s z}{\tau_z + n\tau_s} ,$$

and

$$E[Var[z|X]|z] = \frac{1}{\tau_z + n\tau_s} .$$

yielding the result. All the three corollaries then follow by direct calculation. □

PROOF OF PROPOSITION 2. We consider $\rho = 0$ to focus on the economic environment where the information-revealing effect dominates. When the cutoff firm type z^* adopts cashless payments, (2.4) suggests that the expected financing price from the perspective of $t = 0$ is

$$p(z^*; z \geq z^*) = E[E[z|X, z \geq z^*|z^*] \geq p(z^*) ,$$

where $p(z^*)$ is given by (2.2) in Proposition 1 under $\rho = 0$. On the other hand, if the cutoff firm type z^* uses cash, (2.6) suggests that the expected financing price from the perspective of $t = 0$ is

$$p(z \leq z^*) = E[z|z \leq z^*] = \mu - \frac{\phi(\zeta^*)}{\Phi(\zeta^*)}\sigma ,$$

where $\sigma = \sqrt{\tau_z^{-1}}$ is the prior standard deviation, $\zeta^* = \frac{z^* - \mu}{\sigma}$ is the standardized cutoff firm type given the prior distribution, and

$$\phi(\zeta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\zeta^2\right),$$

and

$$\Phi(\zeta) = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{\zeta}{\sqrt{2}}\right)\right),$$

are the probability density function and the cumulative distribution function of a standard normal distribution.

Define

$$\delta(z^*) \doteq p(z^*; z \geq z^*) - p(z \leq z^*)$$

as the expected payoff gain at $t = 0$ for the cutoff firm type z^* by choosing cashless payments over cash. Direct calculation yields:

$$\begin{aligned} \delta(z^*) &\geq p(z^*) - p(z \leq z^*) \\ &= \left(\frac{n\tau_s}{\tau_z + n\tau_s}\zeta^* + \frac{\phi(\zeta^*)}{\Phi(\zeta^*)}\right)\sigma, \end{aligned} \tag{A.1}$$

where again $p(z^*)$ is given by (2.2) in Proposition 1 under $\rho = 0$. Further define

$$M(\zeta^*) \doteq \frac{\phi(\zeta^*)}{\Phi(\zeta^*)} > 0.$$

By standard statistical result (e.g., Gordon, 1941), we know that $M(\zeta^*)$ has the following properties:

i). $\lim_{\zeta^* \to -\infty}(\zeta^* + M(\zeta^*)) = 0$,

ii). $-1 < M'(\zeta^*) < 0$, and

iii). $M''(\zeta^*) > 0$,

the three of which jointly imply that

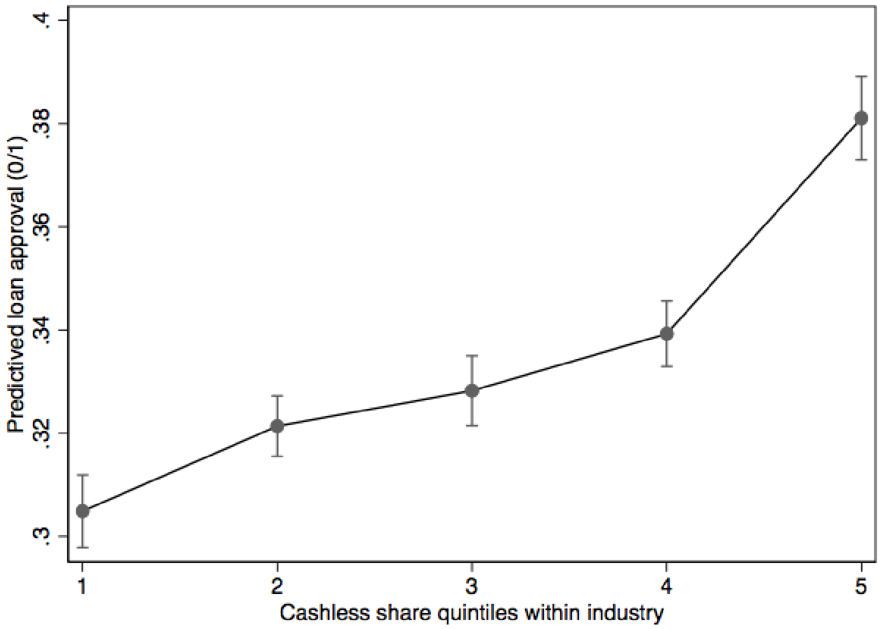
$$\zeta^* + M(\zeta^*) > 0 \tag{A.2}$$

for any $\zeta^* > -\infty$. Because $\frac{n\tau_s}{\tau_z + n\tau_s} < 1$, (A.1), (A.2) and property i) above then jointly imply that $\delta(z^*) > 0$ for all $-\infty \leq \zeta^* < 0$. On the other hand, (A.1) also directly means that $\delta(z^*) > 0$ holds for all $\zeta^* \geq 0$. Thus, $\delta(z^*) > 0$ for all $z \in \overline{\mathbb{R}}$. Finally, note that

$p(z; z \geq z^*)$ increases in z by construction (2.5), confirming that a monotone equilibrium exists only if $z^* = -\infty$, concluding the proof. $\square$

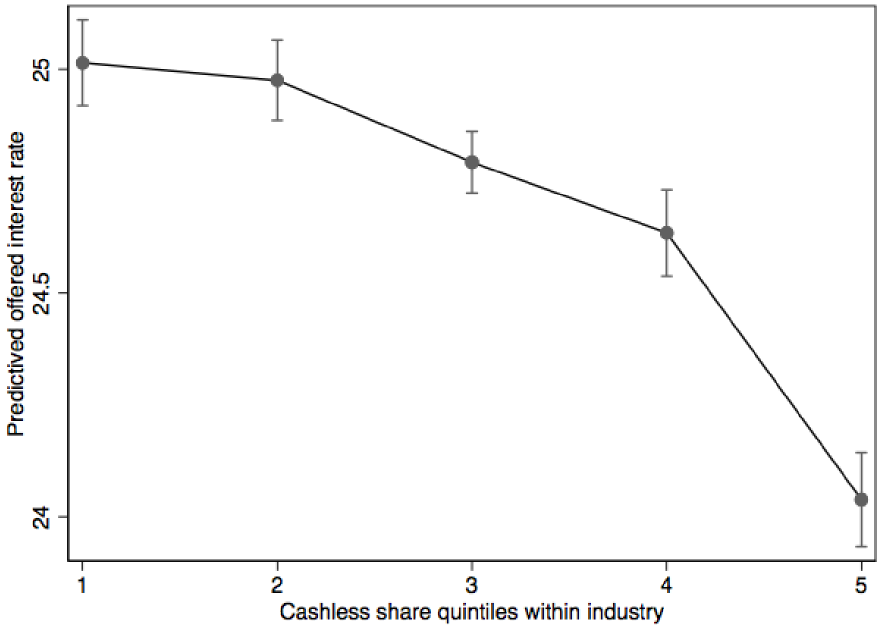
B Additional Empirical Results

Figure A1: **Cashless share quintiles and loan approval**



Notes: This figure reports the marginal effect of cashless share on the loan approval outcome (0/1). The underlying estimation is the OLS regression that uses the quintile of cashless share within a given industry to predict loan approval outcome (0/1) with fixed effects and controls same as Table A3.

Figure A2: **Cashless share quintiles and offered interest rate**



Notes: This figure reports the marginal effect of cashless share on the offered interest rate. The underlying estimation is the OLS regression that uses the quintile of cashless share within a given industry to predict the offered interest rate with fixed effects and controls same as Table A3.

Table A1: **Cashless payments use by industry**

	Obs	**Mean**	**p10**	**p25**	**p75**	**p90**
E-commerce	1164	0.790	0.534	0.685	0.949	0.994
ITeS / call centers	695	0.784	0.504	0.672	0.945	0.995
Transportation & logistics	2278	0.783	0.517	0.690	0.926	0.981
Restaurant	9996	0.771	0.527	0.665	0.910	0.977
Metal & mining	544	0.755	0.409	0.633	0.944	0.991
IT consulting	710	0.752	0.455	0.637	0.920	0.979
Hotels & restaurants	2027	0.748	0.458	0.631	0.919	0.979
Travel services	7457	0.744	0.481	0.631	0.898	0.963
Contractors	1346	0.743	0.447	0.604	0.913	0.976
Construction	442	0.734	0.472	0.623	0.903	0.961
Advertising	973	0.731	0.424	0.599	0.915	0.984
Services	1594	0.731	0.435	0.592	0.911	0.979
Hardware equipment	1658	0.728	0.419	0.598	0.912	0.976
Healthcare	604	0.727	0.425	0.595	0.909	0.976
Computers(hardware) & electronics	1857	0.724	0.456	0.601	0.892	0.965
Electrical store	554	0.723	0.423	0.590	0.895	0.973
Home furniture & furnishing	696	0.723	0.429	0.610	0.887	0.963
Automobiles	1240	0.718	0.422	0.597	0.883	0.967
Cement & construction	549	0.712	0.405	0.579	0.884	0.966
Education consulting	840	0.694	0.394	0.550	0.872	0.963
Retail / shop	1412	0.688	0.398	0.561	0.858	0.956
Wood & wood products	375	0.684	0.329	0.517	0.863	0.955
Agriculture	1489	0.682	0.352	0.524	0.884	0.973
Books, office supplies & stationery	629	0.675	0.368	0.533	0.861	0.958
Jewellery	513	0.673	0.349	0.528	0.855	0.968
FMCG & household products	2197	0.671	0.357	0.528	0.850	0.958
Textiles	7077	0.669	0.339	0.529	0.855	0.953
Trading	2747	0.666	0.348	0.525	0.853	0.956
Chemist / druggist	874	0.659	0.351	0.526	0.838	0.930
Grocery store	2304	0.635	0.314	0.492	0.822	0.928

Notes: This table reports the share of cashless payment use by industry sorted by mean. The selected industries represent the most frequent ones in our bank statement data as they constitute 90% of the observations excluding unclassified data (i.e. industries labelled as “NA” or “others”).

Table A2: **Characteristics comparison for borrowers with or without chest bank accounts**

	Without chest bank			**With chest bank**		
	Obs	**Mean**	**SD**	**Obs**	**Mean**	**SD**
			Panel A. Loan Applicant Characteristics			
Year of Application	83,089	2018	0.7	3,358	2018	0.8
Applicant Age	83,085	35.3	9.4	3,358	34.9	9.4
Business Age	82,585	4.8	13.0	3,331	5.5	6.5
Cibil Score	83,089	534	295.7	3,358	532	301.1
Credit History Length (in Years)	60,955	7.0	5.5	2,478	6.7	5.4
			Panel B. Banking Statement Data			
Share of Cashless Payments	82,994	0.718	0.2	3,355	0.689	0.2
of which: Information-intensive Payments	82,994	0.541	0.3	3,355	0.540	0.3
Share of Cash Payments	82,994	0.282	0.2	3,355	0.311	0.2
Avg. Monthly Revenue (INR in thousand)	83,089	1,621	12,977.2	3,358	1,483	10,191.0
Avg. # of Transactions	83,089	558	1,350.0	3,358	458	839.4

Notes: This table reports a comparison on characteristics for borrowers that bank at chest banks vs. the ones that do not. Panel A compares applicants characteristics by years of application, age, credit score and credit history. Panel B compares banking payments of applicants by payment technologies, bank type and monthly average revenue.

Table A3: **Cashless share as a categorical variable and lending outcomes**

	(1) Approved Loan (1/0)	(2) Offered Interest Rate
Cashless share quintiles by industry=2	0.016***	-0.039
	(0.005)	(0.062)
Cashless share quintiles by industry=3	0.023***	-0.222***
	(0.006)	(0.056)
Cashless share quintiles by industry=4	0.034***	-0.380***
	(0.005)	(0.080)
Cashless share quintiles by industry=5	0.076***	-0.976***
	(0.006)	(0.087)
Borrower controls	Yes	Yes
Industry FE	Yes	Yes
Application month FE	Yes	Yes
Cibil score group FE	Yes	Yes
3-digit zipcode FE	Yes	Yes
Revenue deciles FE	Yes	Yes
Observations	52,696	17,548
R^2	0.288	0.476

Notes: This table presents OLS regressions that use the share of cashless payments as a categorical variable to predict loan consequences. Base level ("Cashless share quintiles by industry=1") is omitted in the report. The regressor "Cashless share quintiles by industry" indicates the quintile of cashless share within a given industry. E.g., The variable "Cashless share quintiles by industry=2" indicates the applicant's cashless share use is between 20% to 40% quintile in the industry that the applicant belongs to. The set of controls includes log # of payments, credit history length, business vintage, log of owner's age, missing credit score indicator and top-up loan indicator. Standard errors are clustered at application-month level.

Table A4: **Cashless payments, Cibil score and lending outcomes**

	Approved Loan (1/0)		Offered Interest Rate	
	(1)	(2)	(3)	(4)
Share of cashless payments	0.117***	0.114***	-1.713***	-1.704***
	(0.010)	(0.011)	(0.147)	(0.143)
	[7.1e-15]	[6.2e-14]	[4.9e-15]	[2.3e-15]
Cibil score	0.003***	0.003***	-0.024***	-0.024***
	(0.000)	(0.000)	(0.002)	(0.002)
	[1.1e-19]	[6.5e-19]	[6.1e-16]	[6.9e-16]
Borrower controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Application month FE	Yes	Yes	Yes	Yes
3-digit zipcode FE	No	Yes	No	Yes
Revenue deciles FE	No	Yes	No	Yes
Observations	53,042	52,700	17,606	17,548
R^2	0.248	0.264	0.432	0.459

Notes: Standard errors are in parentheses and p-values are in sqaure brackets. This table presents OLS regressions that use the share of cashless payments and Cibil score to predict loan consequences. The set of controls includes log # of payments, credit history length, business vintage, log of owner's age, missing credit score indicator and top-up loan indicator. Standard errors are clustered at application-month level.

Table A5: **Cashless share measured over different time lengths and lending outcomes**

	Approved loan (1/0)		Offered interest rate	
	(1)	(2)	(3)	(4)
Panel A. Measured over 3 months before application				
Share of cashless payments	0.108***	0.088***	-1.647***	-1.501***
	(0.010)	(0.010)	(0.133)	(0.130)
Observations	52,160	51,820	17,514	17,455
R^2	0.237	0.288	0.423	0.468
Panel B. Measured over the last month before application				
Share of cashless payments	0.097***	0.078***	-1.479***	-1.319***
	(0.009)	(0.009)	(0.108)	(0.106)
Observations	48,126	47,820	17,062	17,004
R^2	0.222	0.271	0.423	0.465
Borrower controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Application month FE	Yes	Yes	Yes	Yes
Cibil score group FE	No	Yes	No	Yes
3-digit zipcode FE	No	Yes	No	Yes
Revenue deciles FE	No	Yes	No	Yes

Notes: This table presents OLS regressions that use the share of cashless payments over different measurement lengths to predict loan consequences. The set of controls includes log # of payments, credit history length, business vintage, Log of owner's age, missing credit score indicator and top-up loan indicator. Standard errors are clustered at application-month level.

Table A6: **Direction of payments and lending outcomes**

	Approved Loan (1/0) (1)	Offered Interest Rate (2)	Default (1/0) (3)
Share of inbound cashless payments (revenues)	-0.003 (0.006)	-0.923*** (0.117)	-0.014 (0.011)
Share of outbound cashless payments (expenditures)	0.123*** (0.008)	-1.007*** (0.077)	-0.060*** (0.016)
Borrower controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Application month FE	Yes	Yes	Yes
Cibil score group FE	Yes	Yes	Yes
3-digit zip code FE	Yes	Yes	Yes
Revenue deciles FE	Yes	Yes	Yes
Interest rate FE	No	No	Yes
Observations	52,300	17,456	8,082
R^2	0.291	0.474	0.116

Notes: The set of controls includes log # of payments, credit history length, business vintage, Log of owner's age, missing credit score indicator and top-up loan indicator. Standard errors are clustered at application-month level.

Internet Appendix for

FinTech Lending and Cashless Payments

Pulak Ghosh Boris Vallee Yao Zeng

IA Alternative Model of Debt Financing

In this appendix, we build an alternative model with otherwise a similar setting as the baseline model, but with specific debt financing features. We show that this alternative modeling approach generates qualitatively the same predictions as our baseline model.

This alternative model has two agents: an entrepreneur and a lender. Time is discrete: $t = 0, 1, 2, ..., T, T+1$ with $T \geq 1$. We call $\{0, ..., T-1\}$ the production stage and T the lending stage. The entrepreneur has a risky technology whose output is characterized by an i.i.d. random variable R. The technology requires one unit of capital good as investment at any date $t \leq T$. If an investment is made at t, then that unit of capital good gets deployed and fully depreciated, while at $t+1$, the technology produces $R_H > 1$ consumption goods (in a good state) with probability π, and $0 < R_L < 1$ consumption goods (in a bad state) with probability $1-\pi$. Each unit of consumption goods is valued at 1 by (unmodeled) consumers. Consumers are competitive so that they always bid up the sale price to 1. Economically, π captures the quality of the entrepreneur's technology, and as we elaborate on later, its resulting creditworthiness. We call π the entrepreneur's type. The entrepreneur knows her own type but the lender does not. The entrepreneur has T units of capital goods during $t \in \{0, ..., T-1\}$, and her reservation value of the capital goods is 0. Thus, she always invests at the production stage, yielding a series of realized production outcomes $\{R_t | R_t \in \{R_H, R_L\}, 0 \leq t \leq T-1\}$.

At the beginning of the production stage $t = 0$, the entrepreneur chooses how to accept payments for the consumption goods produced. She can either accept payments in cash, which is costless but renders the production outcomes non-verifiable, or she can commit to using an outside cashless payment service. The cashless payment service requires a fixed up-front fee of C and the realized production outcomes $\{R_t\}$ will be documented as a file of payment records, which is verifiable and can be accessed by the lender in the lending stage. Under this setting, a file of payment records at T can be sufficiently summarized by a couple (n, x), where $n \in \{0, T\}$ denotes the total number of payment records, and $x \leq n$ denotes the realized number of "good" records that indicate realized good production outcomes R_H. Note that $(0, 0)$ denotes a degenerate file for an entrepreneur who uses cash.

Similar to our baseline model, the binary choice between cash and verifiable cashless payments over the entire production period is a parsimonious way to capture the essence of payment technology choice. In reality, an entrepreneur may mix a spectrum of payment methods with different degrees of verifiability. We also note that the length of the production period does not correspond to the entrepreneur's life cycle but rather to how many verifiable payment records she may potentially establish, and therefore proxies for the maximum quantity of information contained in a given payment record.

At the lending stage $t = T$, the entrepreneur does not have any more capital goods, and thus has to borrow from the lender, whose values of both capital and consumption goods are 1. Also at $t = T$, the lender can costlessly access the entrepreneur's established payment records (n, x), update his belief, and optimally offer the entrepreneur a standard debt contract with interest rate r to maximize his expected lending profit. The lender will reject the entrepreneur if his expected lending profit is negative. The entrepreneur also has a private reservation interest rate $\bar{r}$ and will only accept the lender's offered interest rate if the offered interest rate is lower than her reservation interest rate is higher.[22]

The equilibrium concept we consider is a standard sequential equilibrium. Specifically, the equilibrium profile consists of the entrepreneur's payment choice policy $n(\pi) : [0, 1] \to \{0, T\}$, that is, whether an entrepreneur uses cash or verifiable cashless payments, and the lender's interest rate policy $r(n, x) : \{\mathcal{I}_T\} \to \mathbb{R}$, where $\mathcal{I}_T$ is the lender's information set at $t = T$, and both agents maximize their expected profits. According to sequential rationality, the lender makes inference about the entrepreneur's type from both 1) the actual information content of the entrepreneur's payment records (n, x), and 2) the entrepreneur's decision of using cashless payments or not, that is, the strategy $n(\cdot)$ itself. To help deliver closed-form solutions and explicit comparative statics, we assume that the lender's prior of π follows a standard uniform distribution. We also assume that the lender's prior about the entrepreneur's reservation interest rate $\bar{r}$ follows a uniform distribution supported on $[0, R_H - 1]$.[23]

[22]The existence of reservation interest rate is necessary to ensure an interior solution of equilibrium interest rate because the lender is a monopoly; see Parlour, Rajan, and Zhu (2020) for similar assumptions. It can be motivated by the competition from another unmodeled lender.

[23]The reservation interest rate cannot be higher than $R_H - 1$ because if so, the entrepreneur will get a

IA.1 Optimal Loan Pricing and Capital Allocation

We first consider how the lender determines the optimal offered interest rate r. Denote the lender's profit conditional on the entrepreneur accepting the offered loan by

$$V(r) = \pi(1+r) + (1-\pi)R_L - 1\,.$$

The lender's problem is:

$$\max_r E_T\left[V(r)\left(1 - \frac{r}{R_H - 1}\right)|\mathcal{I}_T\right]\,,$$

in which the expectation is taken at date T and the second term is the probability of the entrepreneur accepting the offered loan.

Let $\tilde{\pi} \doteq E[\pi|\mathcal{I}_T]$ be the lender's posterior belief about the entrepreneur type at T given an information set $\mathcal{I}_T$. Direct calculation yields:

Lemma **1.** *The lender's optimal interest rate is given by:*

$$r^* = \frac{1-R_L}{2\tilde{\pi}} + \frac{R_H + R_L - 2}{2}\,, \tag{IA.1}$$

which is decreasing in $\tilde{\pi}$. Moreover, the lender's optimal expected profit conditional on extending a loan is given by

$$E[V(r^*|\mathcal{I}_T)] = \tilde{\pi}\left(1 + \frac{1-R_L}{2\tilde{\pi}} + \frac{R_H + R_L - 2}{2}\right) + (1-\tilde{\pi})R_L - 1\,,$$

which is increasing in $\tilde{\pi}$ and is positive if and only if

$$\tilde{\pi} > \frac{1-R_L}{R_H - R_L}\,. \tag{IA.2}$$

Lemma 1 explicitly characterizes a bridge between 1) lending outcomes, that is, loan pricing and capital allocation, and 2) the lender's posterior belief $\tilde{\pi}$ about the entrepreneur's type. When the lender believes the entrepreneur is of a higher type, he is more likely to extends a loan (as suggested by condition (IA.2)) and offers a lower interest rate (as suggested by condition (IA.1)). We can therefore focus on how the entrepreneur' payments record affects the lender's posterior beliefs.

negative surplus even in the good state.

IA.2 Impact of Payment Records on Optimal Loan Pricing and Lending

In this subsection, we study how the information content of verifiable cashless payments affects loan pricing and lending outcomes in expectation – one direction of the synergy between cashless payments and lending. We thus consider a class of sub-game equilibria in which all entrepreneur types commit to cashless payments, that is, $n(\pi) = T$ for all $\pi \in [0, 1]$.[24]

Under this class of sub-game equilibria, the lender updates his belief about π only from the actual information context of (T, x), and not from $n(\cdot)$. More good payments in his record, i.e., a higher x, naturally implies a better posterior belief by the lender, and by Lemma 1 immediately leads to a higher chance of a loan being granted and a lower offered interest rate. At the same time, x is itself a random variable at $t = 0$ and its realization depends on the entrepreneur type: a better entrepreneur expects to establish more good payment records when she commits to using verifiable cashless payments. Moreover, x should better reflect the entrepreneur's true type rather than luck when the entrepreneur uses cashless payments longer and more independent payment records can be established, i.e., when T is larger. We formalize the above intuition in the next proposition by solving the expected interest rate in closed-form and perform comparative statics.

Proposition **3.** *For an entrepreneur of type π, the expected interest rate she gets by choosing cashless payments is*

$$r^*(\pi, T) = \frac{(1 - R_L)(2 + T)}{2(1 + \pi T)} + \frac{R_H + R_L - 2}{2},$$

which is decreasing in T if $\pi \geq \frac{1}{2}$, and

$$\frac{\partial^2 r^*(\pi, T)}{\partial \pi \partial T} < 0. \tag{IA.3}$$

Furthermore, when $T \to \infty$, the expected interest rate is set as if the lender knows the true type of the entrepreneur:

$$r^*(\pi, \infty) = \frac{1 - R_L}{2\pi} + \frac{R_H + R_L - 2}{2}. \tag{IA.4}$$

[24] We solve for the full equilibrium in the next subsection and explain then why this specific class of sub-game equilibria are particularly relevant.

Proof of Proposition 3. Because all entrepreneur types use cashless payments and establish verifiable payment records, the lender's prior of π (before seeing the actual payment records) is $\frac{1}{2}$. By standard Bayesian updating under uniform distributions, the lender's posterior about π after seeing a realized file of payment records (T, x) is given by

$$\tilde{\pi} = E_T[\pi | T, x] = \frac{1+x}{2+T} . \tag{IA.5}$$

Intuitively, the more good payment records are established at the production stage (i.e., the higher x is), the higher the lender's posterior is about the entrepreneur's type.

Note that x in (IA.5) is itself a random variable: better entrepreneurs generate more good payment records in expectation. Conditional on any true type π and total number of payment records to be established T, we have $E_0[x|\pi, T] = \pi T$ where the expectation is taken at $t = 0$. Thus,

$$E[\tilde{\pi}|\pi, T] = \frac{1+\pi T}{2+T} , \tag{IA.6}$$

and

$$\lim_{T\to\infty} E[\tilde{\pi}|\pi, T] = \pi ,$$

implying that the lender's posterior converges to the true type when there are sufficiently many payment records.

Thus, in view of Lemma 1, we have

$$r^*(\pi, T) = \frac{1-R_L}{2E_0[\tilde{\pi}|\pi, T]} + \frac{R_H + R_L - 2}{2} ,$$

where $E_0[\tilde{\pi}|\pi, T]$ is given by condition (IA.6). Direct calculation yields the desired results. □

The results in Proposition 3 lead to two empirical predictions, which are in line with the predictions of our baseline model. First, the comparative statics (IA.3) suggests that when more verifiable payment records are established, a better-than-average entrepreneur expects a lower interest rate, and the marginal effect in interest rate reduction is stronger for better entrepreneurs. In contrast, worse-than-average entrepreneurs expect higher interest rates with more verifiable payment records. Combined with the calculation of lender expected profit as in Lemma 1, it further implies that better entrepreneurs are more likely to be granted a loan when establishing more verifiable payment records.

Furthermore, the limiting result (IA.4) suggests that more verifiable payment records allow the lender to detect the entrepreneur's true type more precisely. This allows the lender to offer a more efficient interest rate, yielding a more efficient capital allocation. Empirically, this should be reflected by a lower default rate conditional on an interest rate, since the firm's risk has been more efficiently priced in in the lending decision.

IA.3 Optimal Payment Method Choice

One implication from Proposition 3 is that when all entrepreneur types use cashless payments, only better-than-average entrepreneurs can enjoy a lower interest rate by committing to cashless payments and establishing more payment records. Does it imply that lower-than-average entrepreneurs would deviate to use cash instead? The answer is no. When entrepreneurs make the payment processing decision $n(\cdot)$ endogenously, the lender will update his belief about the entrepreneur type based on that decision before evaluating the information content of the realized payment records (T, x). If lower-than-average entrepreneurs used cash in equilibrium, the lender would update his belief to reflect that, which would then lead some of those entrepreneurs to consider establishing payment records instead. If the cost of the cashless payment service C is low enough, that force will eventually lead all entrepreneur types to adopt cashless payments. We formalize this idea in this subsection by fully solving the equilibrium. This equilibrium outcome illustrates the other direction of the synergy between lending and cashless payments and also confirms the relevance of the sub-game equilibria we focus on in Section IA.2.

We consider a monotone equilibrium in which there exists a cutoff entrepreneur type π^* such that higher entrepreneur types $\pi > \pi^*$ commit to using verifiable cashless payments whereas lower types $\pi < \pi^*$ use cash. By definition this cutoff type π^* is indifferent between those two payment choices. We consider the two choices below, respectively.

First, if the cutoff type π^* commits to cashless payments and establishes payment records, the sender will first know that her type must be no lower than π^* before evaluating the payment records, and will further update his belief about π based on the realized payment records (T, x) at $t = T$. The lender's posterior belief after evaluating the realized payment records (T, x) is:

$$\tilde{\pi}(T, x) = E[\pi | \pi \geq \pi^*, T, x] \,. \tag{IA.7}$$

Then by Lemma 1, the expected offered interest rate to type π^* who commits to establish T payment records at $t = 0$ is

$$r^*(\pi^*, T) = \frac{1 - R_L}{2E[\tilde{\pi}(T, x)|\pi^*]} + \frac{R_H + R_L - 2}{2}, \tag{IA.8}$$

where $\tilde{\pi}(T, x)$ is given by (IA.7) and the expectation is taken at $t = 0$ by the entrepreneur with respect to x, the potential number of good payment records. Accordingly, the total interest she expects to pay back in a good state is

$$I(\pi^*, T) = E_{\bar{r}}[\min(r^*(\pi^*, T), \bar{r})], \tag{IA.9}$$

where $r^*(\pi^*, T)$ is given by (IA.8) and the outer expectation is taken with respect to the reservation interest rate $\bar{r}$.[25]

Second, if the cutoff type π^* uses cash instead, the lender will anyway know that her type is no higher than π^* despite not seeing any payment records, that is, $\tilde{\pi} = E[\pi|\pi \leq \pi^*]$. Thus, by Lemma 1 again, the expected offered interest rate to type π^* who chooses cash at $t = 0$ (and thus establishes none payment record) is

$$r^*(\pi^*, 0) = \frac{1 - R_L}{2E[\pi|\pi \leq \pi^*]} + \frac{R_H + R_L - 2}{2}, \tag{IA.10}$$

and similarly, the total interest she expects to pay back in a good state is

$$I(\pi^*, 0) = E_{\bar{r}}[\min(r^*(\pi^*, 0), \bar{r})], \tag{IA.11}$$

where $r^*(\pi^*, 0)$ is given by (IA.10). Now using expressions (IA.9) and (IA.11) and applying the cutoff entrepreneur type's indifference condition, we have the following formal result:

Proposition 4. *There exists a cutoff entrepreneur $\pi^* \in [0, 1]$ who is indifferent between using cashless payments or cash:*

$$\pi^* I(\pi^*, T) + C = \pi^* I(\pi^*, 0), \tag{IA.12}$$

[25] This is because the entrepreneur will take that reservation interest rate if the lender's offered interest rate is higher or the lender rejects her. The same applies to the expectation in (IA.11).

when C is sufficiently small. Notably,

$$\lim_{C \to 0} \pi^* = 0\,,$$

meaning that all entrepreneurs commit to establishing payment records if the verifiable cashless payment service is sufficiently cheap.

PROOF OF PROPOSITION 4. We repeatedly use the expressions of the two expected interest rates, $r^*(\pi, T)$ for entrepreneur-type π who commits to cashless payments (see condition (IA.8)) and $r^*(\pi, 0)$ for entrepreneur-type π who uses cash (see condition (IA.10)) in this proof. Consider an arbitrary cutoff entrepreneur type $\hat{\pi}$'s date-0 expected interest payment in a candidate cutoff equilibrium. If she commits to using cashless payments, it is given by $\hat{\pi} E_{\bar{r}}[\min(r^*(\hat{\pi}, T), \bar{r})]$, in which the minimum operator captures the fact that the entrepreneur chooses her reservation interest rate if the offered interest rate is higher. Instead, if, she uses cash, it is given by $\hat{\pi} E_{\bar{r}}[\min(r^*(\pi, 0), \bar{r})]$. Thus, the ex-ante net benefit of committing to cashless payments is given by

$$V_{cashless}(\hat{\pi}) = -\hat{\pi} E_{\bar{r}}[\min(r^*(\hat{\pi}, T), \bar{r})] - C\,,$$

which is the opposite of the expected interest payment, minus the physical cost of using cashless payments, and the ex-ante net benefit of using cash is given by

$$V_{cash}(\hat{\pi}) = -\hat{\pi} E_{\bar{r}}[\min(r^*(\pi, 0), \bar{r})\,.$$

Define

$$\Delta V(\hat{\pi}) = V_{cashless}(\hat{\pi}) - V_{cash}(\hat{\pi}) \tag{IA.13}$$

as the ex-ante net benefit gain for entrepreneur type $\hat{\pi}$ from choosing cashless payments over cash. First, notice that for the worst type 0 as the cutoff type:

$$\Delta V(0) = -C < 0\,,$$

implying the worst type is strictly worse-off by establishing payment records.

On the other hand, for the best type 1 as the cutoff type:

$$\Delta V(1) = E_{\bar{r}}[\min(r^*(E[\pi], 0), \bar{r})] - E_{\bar{r}}[\min(r^*(1, T), \bar{r})] - C \tag{IA.14}$$

$$\geq E_{\bar{r}}[\min(r^*(E[\pi], 0), \bar{r})] - E_{\bar{r}}[\min(r^*(1, 1), \bar{r})] - C\,, \tag{IA.15}$$

where equation (IA.14) follows from that the uninformed interest rate is set as if every entrepreneur is of the average type, that is, $E[\pi|\pi \leq 1] = E[\pi]$, and inequality (IA.15) follows from that $r^*(1,T) \leq r^*(1,1)$ for any $T \geq 1$ and the fact that function $E_{\bar{r}}[\min(z, \bar{r})]$ is strictly increasing in z by construction. Combine the first two terms in the right hand side of (IA.15) and define

$$A = E_{\bar{r}}[\min(r^*(E[\pi], 0), \bar{r})] - E_{\bar{r}}[\min(r^*(1, 1), \bar{r})]\,.$$

Notice that $E[\pi] < 1$, it must be that $r^*(E[\pi], 0) > r^*(1,1)$ by Proposition 3. Again because function $E_{\bar{r}}[\min(z, \bar{r})]$ is strictly increasing in z, A must be strictly positive. It immediately follows that the right hand side of (IA.15) must be strictly positive if $C < A$, that is, if C is sufficiently small. This implies that the best type is strictly better-off by committing to cashless payments if they are cheap enough.

Consequently, as defined in (IA.13), when $C < A$, $\Delta V(0)$ is strictly negative while $\Delta V(1)$ is strictly positive. By the continuity of $\Delta V(\hat{\pi})$ in $\hat{\pi}$, there must exists a $\pi^* \in (0,1)$ at which the indifference condition (IA.12) holds. The limiting result immediately follows from $\lim_{C\to 0} \Delta V(0) = 0$ and $\lim_{C \to 0} \Delta V(1) = A > 0$, concluding the proof. $\square$

Economically, the left hand side of the indifference condition (IA.12) denotes the total cost of using cashless payments, which is the expected interest payment to the informed lender plus the physical cost of using cashless payments. The right hand side denotes the total cost of using cash, that is, the expected interest payment to the uninformed lender.

IB Classification of Payment Records by Technologies

In this appendix, we provide a detailed elaboration of the identification and classification methods, by which we classify identifiable payment records into either cash or cashless payments.

To start, we identify and classify cash payments. Payment records with strings "cash withdrawal" or "cash deposit" are identified and classified as cash payments.

We then identify and classify cashless payments by different technologies. First, we identify information-intensive payments, including Internet banking payments and certified checks. In the context of India, those two broad payment technologies are information-intensive in the sense that each payment records corresponds to a single transaction (i.e., it is not aggregated), and the name and type of the payment counter-party is identifiable in the record.

For Internet banking payments, we can identify and classify six sub-categories, as below. First, payments made by Immediate Payment Service (IMPS), which is the instant payment inter-bank electronic funds transfer system in India, are identified by a description including strings "INB IMPS" or "ONL IMPS". Second, Internet banking transactions on bank's websites are identified by the relevant strings, including large corporate transfers and any transfers between saving/checking accounts. Third, direct deposits are are identified by the relevant strings, including loan disbursals, salary, investment income, etc. Fourth, online payments made by debit cards are identified by the string "purchase by card." Fifth, Mastercard Money Transfer is identified by a description starts with "MMT". Finally, automatic fee deduction (excluding bank charges) and automatic payments (loan repayments, utilities, credit card repayment etc.) are identified by the relevant strings.

For certified check payments, they are identified by a description including "CLG-CHQ", "ECS/ACH", "Cheque" or starts with "CLG".

Second, we identify information-light payments, including payments through third-party mobile applications and mobile banking, as well as POS machine payments. In India, those payments records are either aggregated in bank statements, or lacking the name or typeof the payment counter-party.

For mobile payments, we can identify and classify five subcategories, as below. First, payments made through third-party mobile gateways, including Paytm, Razorpay, PayU, Phonepe, Paypal, are identified by the relevant strings. Second, mobile banking payments made through mobile banking apps are identified by a description starting with "MB", "MOB/TPFT", "MOBFT" and including 'PayZapp'. Third, RuPay, which are payments made over texts, are identified using the relevant string. Fourth, payments made through Unified Payments Interface (UPI), which is a mobile payment system developed by India's National Payment System for interactions between two bank accounts over mobile platform, are identified by a description with "UPI". Finally, payments made through IMPS Mobile Transactions, another mobile system developed by India's National Payment system, are identified by a unique 7-digit "Mobile Money Identifier" (MMID).

For POS machine payments, they are identified by a description including "POS", "ECOM", or "PUR".

We finally note three adjustments and exceptions in the classification procedure. First, the payment counter-party is not identifiable for a small number of Internet banking payments and certified checks in our sample, which we will instead classify as information-light payments. Second, any cash transfer to electronic wallets on third-party mobile apps are classified as information-light cashless payments. Third, rejected payments and penalty charges due to rejected payments, whenever identified, are not classified as either cash or cashless payments because they are indicative of technical errors in the payment process and do not fall into our economic framework.

Following the procedure described above, we have been able to identify and classify 75% of the total payment records observed in the data.