FinTech Lending and Cashless Payments*

Pulak Ghosh
Indian Institute of Management

Boris Vallee
Harvard Business School

Yao Zeng
Wharton

December 29, 2021

Click **here** for the latest version

Abstract

We uncover an informational synergy between FinTech lending and cashless payments. Theoretically, FinTech lenders screen borrowers more efficiently when borrowers use cashless payments that produce transferable and verifiable information. In turn, a strategic consideration to stand out from non-adopting borrowers pushes borrowers to adopt cashless payments. Empirically, a larger use of cashless payments predicts a higher likelihood of loan approval, a lower interest rate, and a higher loan amount, especially for firms of higher credit quality. This synergy provides an economic rationale for open banking, and more broadly for data sharing and a lending model without 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 (discussant), Daniel Bjorkegren, Sylvain Catherine, Shawn Cole, William Cong (discussant), Itamar Drechesler, Cesare Fracassi (discussant), Andreas Fuster, Zhiguo He, Sasha Indarte, Amit Khandelwal (discussant), Filippo Mezzanotti, Thomas Philippon, Raghu Rajan, Uday Rajan (discussant), Huan Tang (discussant), Laura Veldkamp, Ansgar Walther (discussant), Aluna Wang (discussant), Wei Xiong, Haoxiang Zhu, and seminar and conference participants at Bank for International Settlements, Federal Reserve Board, Harvard University, Hong Kong Baptist University, Hong Kong University of Science and Technology, Luohan Academy, Monash University, National Taiwan University, Office of the Comptroller of the Currency, Toulouse School of Economics, University of Pennsylvania, University of Technology Sydney, University of Toronto, University of Zurich; American Finance Association (AFA) Annual Meeting, China International Conference in Finance (CICF), European Finance Association (EFA) Annual Meeting, Entrepreneurial Finance Association Conference, HEC Paris Entrepreneurship Conference, NBER Asset Pricing Meeting, Norges Bank-CEPR Banking Workshop, Texas Finance Festival, and World Bank Asia Economic Policy Network Conference for helpful comments. We thank Indifi for providing us with access to their data. All errors are ours.

1 Introduction

Borrowing and lending in credit markets, a central theme in economics, has been traditionally dominated by banks. A fundamental economic rationale for such dominance is the existence of an informational advantage in lending for banks resulting from the scope of their activities. Banks produce borrower information *inside* the bank through monitoring and repeated lending (Diamond, 1991, Rajan, 1992) and from deposit-taking (Black, 1975, Fama, 1985). However, the past decade has witnessed a dramatic rise of lending by financial technology (FinTech) companies, which do not enjoy such relationships with borrowers.[1] How can FinTech lenders level the playing field with banks and even dominate certain lending markets, and what does it mean for the future of banking?[2]

Our paper provides a simple answer to this question by uncovering an informational synergy between FinTech lending and cashless payments. Payments are of first-order importance in financial markets and have been experiencing drastic digitalization and disintermediation in recent decades (Brunnermeier, James and Landau, 2019, Duffie, 2019), which increases their informational verifiability and transferability. Using cashless payments produces verifiable information about would-be borrowers *outside* lending relationships, which can be however easily accessed and used by technology-savvy lenders even if those lenders do not process payments themselves. The uncovered synergy in using and producing outside information implies a joint rise of FinTech lending and cashless payments, and supports the development of an alternative banking model without relationships in the traditional sense. The synergy also provides a rationale for the emergence of FinTech platforms in the sense of stand-alone FinTech lenders starting to offer payment services, and stand-alone payment firms and marketplaces starting to offer lending products.[3]

[1]The Financial Stability Board (FSB) defines FinTech as "technologically enabled financial innovation." Compared to traditional banks, FinTech lenders typically rely on electronic technologies and alternative data sources to screen loans but do not engage in relationship-developing or deposit-taking.

[2]Combining new credits extended by both stand-alone FinTech firms and large technology companies, the Bank for International Settlements estimates the global annual volume of FinTech lending to have grown from \$17 billion to \$769 billion from 2013 to 2019, at a 87.3% annual growth rate. In the US, FinTech lending dominates 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 is particularly pronounced in developing economics (Claessens, Frost, Turner, and Zhu, 2018).

[3]In the US, OnDeck, a once stand-alone FinTech lender, has started to provide payment service to its borrowers, and PayPal, Square, and Stripe, once stand-alone payment service providers, have all started to offer loans. Globally, Amazon and Alibaba are prominent examples of marketplaces starting to offer loans to their partners. Despite realizing the synergy, these institutions remain fundamentally different from traditional banks.

We first build a simple model showing that the interaction between FinTech lenders and cashless payments fosters the development of both. In the model, there is a firm and a financier. A higher-type firm is more likely to produce a better product, but the financier does not know the type. The firm first uses its endowment to produce, after which it seeks finance from the financier. Importantly, the firm chooses to use either cash or cashless payments to process its production. Compared to cash, the cashless payment service records verifiable information about production outcomes, which the firm commits to providing to the financier in the financing stage.

We establish the existence of a 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. First, thanks to the precision of the signal provided by payment records, the financier faces a reduced level of risk and therefore provides improved financing terms for firms of all types – a risk-reducing effect. Second, the level of the signal helps reveal the quality of the firm to the financier, meaning that better quality firms will get even better financing terms – an information-revealing effect.

We then show how the reliance of lending on outside verifiable information fosters the adoption of cashless payments, highlighting 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 be of a low type. Thus, a moderately low-type firm would find it optimal to stand out from even lower-type firms by adopting cashless payments. This force pushes more firm types to adopt cashless payments, with an unraveling towards full adoption if adoption costs are sufficiently small.

In transparently uncovering the synergy and providing an economic rationale for the growth potential of FinTech lending, we purposefully abstract away from competition between banks and FinTech lenders, or market power of FinTech lenders. We show in extensions of the model that the existence of the synergy is robust to these considerations.

We move on to provide empirical evidence on how the use of outside information of varying verifiability affects lending outcomes. In doing so, we use novel loan application-level data from one of the leading FinTech lenders in India, which focuses on unsecured lending to small businesses. This empirical setting uniquely fits our analysis for several reasons. First, India has been experiencing a pronounced joint rise in FinTech lending – India's volume of FinTech lending grew from $9 bn in 2012 to $150 bn in 2020 – and cashless payments, following both public and private initiatives.[4,5] Second, unsecured loans to small businesses are a segment where information asymmetry is large, making it a compelling setting to study the informational mechanism we flesh out in the model. Third, the data includes the whole information set that the FinTech lender observes and uses, including applicants' business characteristics, traditional credit scores, and most importantly, detailed payment records of different levels of verifiability. This feature allows us as econometricians to test how payment information with varying levels of verifiability affects lending outcomes. Fourth, the lender gets access to those outside payment records from bank statements, which the borrowers agree to share when applying. Thus, our empirical tests shed light on the implications of data sharing and open banking at the same time. Finally, as we elaborate on below, our setting allows for an identification strategy based on the effects of the 2016 Indian Demonetization on payments, which helps us rule out a spurious relationship between borrowing outcomes and cashless payment use.

We develop a parsimonious text-analysis methodology to classify each payment into cash and cashless payments based on their transaction label. 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. At the application level, we build measures of cashless payment use intensity by calculating the value-weighted share of transactions executed with this technology. Equipped with these measures and our comprehensive data, we turn to studying whether cashless payment use correlates with loan screening outcomes on both the extensive and intensive margins, controlling non-linearly for a comprehensive set of applicant characteristics such as firm size, age, 3-digit zip code, and credit score.

We find that a higher share of cashless payments is associated with improved lending outcomes: regardless of their credit quality, applicants using more cashless payments are significantly more likely to obtain a loan, and when doing so obtain a lower interest rate and a higher loan amount. These results are consistent with the risk-reducing effect that benefits all firms. The results are robust to a host of alternative specifications and measures of cashless payments, suggesting that the findings are unlikely to be driven by measurement errors or selection biases. The economic magnitude is also large: controlling for other observable characteristics, including the borrower credit score and industry, a one standard

[4]See https://www.statista.com/statistics/1202533/india-digital-lending-volume.

[5]See, for instance, https://techwireasia.com/2021/03/indias-road-to-becoming-a-cashless-economy.

deviation increase in cashless payment use is associated with a 2 percentage points higher likelihood of obtaining a loan, or 8% of the unconditional likelihood of obtaining a loan in our sample. Conditional on obtaining a loan, a one standard deviation increase in cashless payment use corresponds to a 36 bps lower interest rate, and 19% larger amount granted.

What is crucial for cashless payment use to affect lending outcomes in our framework is that the information it produces is verifiable. To this end, we further show that the positive effects of cashless payments on lending outcomes are more pronounced for more verifiable payment records, such as individual internet transfers, as opposed to less verifiable ones, such as aggregated or anonymized payments. Moreover, we find that borrowers conducting a large share of their sales on online marketplaces that share their data with the FinTech lender also benefit from improved lending outcomes.[6] Online marketplace transaction data is conceptually comparable to the one provided by cashless payment records 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 verifiable information.

To further gain causal identification, 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 applicant banking at a reserve chest bank branch, that is, a branch that distributes new banknotes and coins. Banking at such a branch after the Demonetization is indeed predictive of lower use of cashless payments, due to the better access to the new banknotes during cash shortage, whereas it does not appear to be the case prior to the demonetization. One innovation of our identification is that it rests on within-district variation and not on the cross-section of districts, as our data uniquely allows us to map borrowers to specific bank branches within a zip code. Since the Demonetization is unexpected, and the matching between firms and chest bank branches is plausibly orthogonal to firm quality once we control for observable characteristics, we use this exogenous variation in the use of cashless payments to strengthen our findings in a causal sense: a higher share of cashless payments leads to a significantly higher likelihood of obtaining a loan. Note that we do not aim at completely ruling out reverse causality due to the two-way synergy we highlighted, but rather mitigating concerns over unobservable variable biases.

[6] Transactions on online marketplace in India are often paid in cash, at delivery of the good ordered, and therefore exhibits a correlation of only 0.07 with our measure of cashless payment use.

Furthermore, we find that the positive effects of using cashless payments are stronger for firms of better quality, consistent with the information-revealing effect. We capture firm quality with several proxies: the applicant's credit score, as well as the level and volatility of payment revenues, which we can uniquely construct from the content of payment records, and find robust results. These findings further support the notion that payment records provide incremental and verifiable information beyond the usual credit scores.

Finally, we find that firms that use more cashless payments are less likely to default, all else equal. We interpret this empirical fact as consistent with our model prediction that the use of cashless payments helps the FinTech lender screen loans more efficiently, leading to more efficient capital allocations among different borrowers.

Although we focus on FinTech lending and cashless payments due to their increasing economic importance, the implications of our paper are broader and contribute to the large relationship banking literature (see Diamond (1991) and Rajan (1992) for pioneer work and Liberti and Petersen (2019) for a survey). First, our framework and results do not exclude the possibility of traditional banks using outside verifiable information (other than credit scores) in lending and exploiting the synergy we describe.[7] Indeed, it is well documented in the literature that a bank typically learns from relationship-specific deposit flows to facilitate lending by the same bank to its depositors (e.g., 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.[8] Our paper complement and differ from those findings by showing that the existence of an integrated 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, any lender may rely on it in screening loans, and we provide causal evidence which is rare in the existing literature. In this regard, our study also speaks to the recent and future evolution of loan screening by traditional banks given the development of open banking. Second, our results are externally valid to any type of outside verifiable information. Although cashless payment records in our empirical context come from bank statements, what matters economically for our study

[7]Anecdotal evidence suggests that traditional banks also increasingly 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.

[8]The idea of synergy across the two sides of a bank balance sheet dates back to Moulton (1918). In addition to an informational spillover, such synergy may also happen due to bank liquidity management (e.g., Calomiris and Kahn, 1991, Diamond and Rajan, 2001, Kashyap, Rajan, and Stein, 2002) or bank market power in the funding markets (e.g., Drechsler, Savov, and Schnabl, 2007).

is that these records are produced outside of the lender in question, as is the marketplace data we also explore. Our key message is therefore that FinTech lenders have a higher incentive and a potential comparative advantage in using outside verifiable information to compensate for the lack of inside information production, which helps them level the playing field with traditional banks and suggests an alternative banking model without relationships in the traditional sense.

Our paper then contributes to the literature on both FinTech lending and cashless payments (see Vives (2019), Allen, Gu, and Jagtiani (2021), and Berg, Fuster, and Puri (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.[9] Our paper highlights an informational channel for FinTech lenders to level the playing field with banks by using outside but verifiable information. In this aspect, our paper is closely related to Björkegren and Grissen (2018, 2020), who find that borrowers' mobile phone call records predict repayment behaviors and defaults, and to Berg, Burg, Gombovic, and Puri (2020), who document that simple "digital footprints" such as website registration inputs predict consumer defaults on top of credit scores.[10] Complementing these studies, we focus on the verifiability of payments, a type of information that is directly generated by the borrowers' production process and is therefore hard to manipulate, and examine the impact on lending outcomes such as loan approval and interest rate directly. Strengthened by causal evidence, our focus allows us to economically pinpoint why and how lenders incorporate outside information, and to provide a rationale for the voluntary provision of such information by borrowers. In this lending aspect, our study also complements Fuster, Goldsmith-Pinkham, Ramadorai, and Walther (2020) and Bartlett, Morse, Stanton, and Wallace (2021), who study to what extent new inference technologies such as algorithmic and machine learning

[9]A related literature considers peer-to-peer or marketplace lending specifically, which relies on end-investor screening and direct investing. Recent studies include Iyer, Khwaja, Luttmer and Shue (2015), Tang (2019) and Vallee and Zeng (2019), for example. Despite the overall growth of FinTech lending, this peer-to-peer aspect has been losing momentum recently because such platforms have been increasingly catering to large, institutional investors, as Vallee and Zeng (2019) predict. We do not focus on the peer-to-peer aspect but rather the broader FinTech lending model that relies on outside verifiable information.

[10]Relatedly, Frost, Gambacorta, Huang, Shin, and Zbinden (2020) examine a large technology company (i.e., "BigTech") in Argentina and find that it has an information advantage in credit assessment relative to a traditional credit bureau.

can affect lending outcomes, focusing on consumer discrimination.

On the payment side, the existing literature mainly focuses on the direct benefits on consumptions and productions of cashless payments, and their adoptions (e.g. Jack and Suri, 2014, Muralidharan, Niehaus, and Sukhtankar, 2016, Higgins, 2020). 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. Theoretically, Parlour, Rajan, and Zhu (2020) study the competition between stand-alone FinTech payment service providers and traditional banks providing both lending and payment service, showing that such competition may indeed disrupt the information spillover within a bank, and Bianchi and Bigio (2021) and Piazzesi and Schneider (2021) incorporate payments into quantitative models and deliver implications on asset prices and monetary policies. We complement them by showing to what extent using cashless payments of different informational verifiability can affect access to capital and improve lending outcomes, and the synergy we highlight provides a new rationale for the surging adoption of cashless payments.

Finally, our paper contributes to a burgeoning literature on the economics of data (see Admati and Pfleiderer (1990) for early work and Bergemann and Bonatti (2019) for a recent survey). Jones and Tonetti (2020) argue that the non-rivalry of data implies increasing returns to the sharing of data and leads to better resource allocations. Bonatti and Cisternas (2020) and Ichihashi (2020) show that aggregating past purchase histories into a transferrable score or sharing consumer data in general leads to more targeted price discrimination. He, Huang, and Zhou (2021) examine the lending competition between traditional banks and FinTech lenders when borrowers share their bank data with FinTech lenders and explore the rich welfare implications of open banking. This literature is so far mostly theoretical, and thus our paper complements it by empirically identifying that the access to traditional bank statements that contain borrowers' payment records leads to better lending outcomes for FinTech lenders, and by considering the strategic interaction in adopting cashless payments that is critical in the synergy between FinTech lending and cashless payments. This strategic consideration 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' (Acemoglu, Makhdoumi, Malekian, and Ozdaglar, 2020, Bergemann, Bonatti, and Gan, 2020), because data sharing increases two-sided platforms' market power

(Kirpalani and Philippon, 2020), or because of agents' behavioral biases such as self-control issues to temptation goods (Liu, Sockin, and Xiong, 2020). Our micro-level analysis also complements Farboodi and Veldkamp (2020a,b), Cong, Xie, and Zhang (2021), and Cong, Wei, Xie, and Zhang (2021) who examine the interaction between growth, 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

We build a simple model to qualitatively uncover the informational synergy between lending and payments. The key of the model is to clarify the different economic forces underlying this synergy in a transparent way, which helps inform our empirical analysis.

2.1 Setting

The model has two representative agents: a competitive risk-neutral firm and a competitive risk-averse financier who has a CARA utility function with absolute risk aversion ρ. Modeling the financier as being risk-averse allows us to parsimoniously capture how information helps reduce uncertainty, as in the literature focusing on information and data (e.g., Farboodi and Veldkamp, 2020a). In reality, financial institutions face various financial constraints, effectively making them risk-averse in investment decisions (e.g., Rampini and Viswanathan, 2010, Bolton, Chen and Wang, 2011) despite diversification.

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 $\tilde{z} \sim N(\mu, \tau_z^{-1})$. 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. To focus on the role of information on financing outcomes without complicating the analysis, we follow the literature (e.g., Fishman and Parker, 2015) to model this process by considering the firm effectively selling the product of uncertain quality y_n to the financier and getting a price of p, where p can be interpreted as loan amount or interest rate to map to a credit market context. 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. Since the financier is competitive, it breaks even in equilibrium, and under the CARA-normal framework, 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 $t = n$.

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. These modeling choices help simplify the analysis while highlighting the essence of the economic forces. We show in the Internet Appendix that our qualitative predictions are robust to these choices.

First, to compare cashless payments to cash in the dimension of verifiability, we abstract away from either the convenience benefits or the physical or implicit costs (e.g., privacy concerns) of using different types of payment methods. We show in Internet Appendix IA.1 that incorporating a net cost of adopting cashless payments would not affect the empirical predictions of the model. In addition, 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 varying informational verifiability.[11] 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.[12]

Second, to highlight the general lending model based on outside information, we abstract away from potential competition between banks and FinTech lenders, or lender market power. In reality, FinTech lenders typically compete for a different market segment than banks do (e.g., Tang, 2019), and our framework indeed aims to capture the early stage of FinTech development, in which most FinTech lenders are small and competitive, to help explain why they can level the playing field and grow. We use a simple extension in Internet Appendix IA.2 to show that our main empirical predictions still hold even if borrowers self-select into FinTech lending from banks based on their types or realized payment records. In Internet Appendix IA.3, we provide another alternative model featuring a monopolistic lender, following Pagano and Jappelli (1993), and show that the empirical predictions are preserved as well.

Finally, to focus on the informational role of payments on financing outcomes, we model financing as a product sale without specifically modeling the type of securities used, and interpret the reciprocal of the sale price as the cost of capital. This implies that our theoretical framework can be applied broadly in considering both debt and equity financing contexts. In the alternative model in Internet Appendix IA.3 with a monopolistic lender, we indeed model explicit debt financing featuring face value, interest rate, and default, yielding the same predictions qualitatively.

[11] 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.

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

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 consider a sub-game equilibrium in all firm types commit 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. In our baseline model, we later verify that this is indeed the only equilibrium in Section 2.3. We also illustrate in Internet Appendices IA.1 and IA.2 that all the model predictions are robust to the potential of only higher-type firms adopting cashless payments and applying for FinTech credit (i.e., $x(z) = X$ for $z > \underline{\mu}$ only) due to either adoption costs or self-selection from banks into FinTech lenders.

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*

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

which is increasing in τ_x and n.

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 and independent of firm type. 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 in which a higher financing price p can be translated into a higher likelihood of loan approval, a lower interest rate, or a larger loan amount, we have the following prediction:

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

Second, we are interested in how the effects of using cashless payments vary across firms with different types, i.e., firm quality in reality. We have the following result:

COROLLARY **2.** *A higher firm type enjoys a higher expected price improvement from choosing cashless payments, and the increase is higher when τ_x or n increases in the sense that*

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

The intuition behind Corollary 2 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. When the records are of higher verifiability or the firm has more cashless payment records, this type-dependent information-revealing effect becomes even stronger. Mapped into our lending context, we have the following prediction:

PREDICTION **2.** *The effects of payment records of higher verifiability or more verifiable cashless payment records in improving loan approvals, reducing interest rates, and increasing loan amounts are stronger for firms of higher quality, all else equal.*

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

Corollary 3. *The mean squared error $E[(z - E[z|X])^2]$ of the financier's inference is decreasing in τ_x and n.*

Intuitively, a lower mean squared error as suggested by Corollary 3 means that the lender bears lower risks in financing the loan. It then implies a lower probability of default, suggesting a more efficient capital allocation. Thus, we have the following prediction.

Prediction 3. *Payment records of higher verifiability or more cashless payment records lead to less default, all else equal.*

We highlight that Prediction 3 fundamentally comes from the risk-reducing effect. A lower mean squared error of the financier's inference does not necessarily imply higher lending profitability in our framework, because competition among financiers always competes away any positive profits under our framework. Looking forward, we will further clarify this point as we empirically test this prediction in Section 4.6.

2.3 Optimal Payment Method Choice

In this subsection, we study how the firm's expectation of its payment records being screened in turn affects its optimal payment method choice – the other direction of the synergy between cashless payments and lending. 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 adoption is costless. We formalize this idea in this subsection by fully solving the equilibrium. This equilibrium outcome also verifies that $x(z) = X$ for all $z \in \overline{\mathbb{R}}$, as we conjected in Section 2.2.

Formally, we 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.[13] By sequential rationality, when firm z^* chooses cashless payments at $t = 0$,

[13]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 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 $z \leq z^*$ use cash, it must be that $z^* = -\infty$, meaning that all firm types will optimally adopt cashless payments.*

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.[14]

We highlight that the main goal of Proposition 2 is to provide a new perspective to help understand the increasing popularity of cashless payments in the long run, rather than literally predicting a universal adoption of cashless payments. In reality, the presence of physical or implicit adoption costs (i.e., privacy concerns) may prevent lower-type firms from adopting cashless payments, as we show in Internet Appendix IA.1.[15] However, as shown there, the empirical predictions we derive in Section 2.2 still hold as long as some high-type firms adopt cashless payments, and the strategic consideration of adopting always persists in any equilibrium.

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

[14]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. 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.

[15]This is analogous to the information closure literature that in many circumstances the unraveling mechanism applies but full unraveling may not be achieved (see Okuno-Fujiwara, Postlewaite, and Suzumura (1990) for a systematic treatment and Acharya, DeMarzo, and Kremer (2011), Ali, Lewis, and Vasserman (2020), Bond and Zeng (2021) for recent applications in pricing contexts).

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 upload 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.[16]

[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, 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.[17]

[16] 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.

[17] 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.

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 a strain on the logistics of banknote distribution.[18]

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 creditworthiness given the wide and relatively random distribution of chest bank branches in India with the idea of distributing cash as efficiently as possible to the whole population, and the absence of a particular benefit to bank at such a bank branch absent the unexpected cash shortage. Indeed, according to our conversation with Indian borrowers, most of them do not know themselves whether they ever bank with a chest bank branch or not. To confirm this point, we provide a comparison of firm

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

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 along various dimensions regarding firm fundamentals.

4 Empirical Findings

Equipped with the loan application-level data and the proxies for payment information in borrowers' bank statements, we empirically test Predictions 1, 2, and 3 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.[19]

4.1 Cashless Payments and Lending Outcomes: Baseline Result

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 three dependent variables: 1) an indicator variable for obtaining a loan, 2) the interest rate offered on the loan conditional on acceptance, and 3) the loan amount ultimately extended, which jointly capture the extensive and intensive margin of lending outcomes. 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

[19]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.

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.

[Insert Figure 2 here]

To make the test more precise, we run formal regressions of all the three lending outcomes on the standardized share of cashless payments, with a comprehensive set of granular controls and fixed effects allowing for non-linear relationships. We use the following baseline specification:

$$\text{LendingOutcome}_i = \beta \text{CashlessShare}_i + \gamma X_i + \sum_k \theta_{F_k(i)} + \varepsilon_i \,. \tag{4.1}$$

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, 4, and 6, 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 in all columns robustly show that a higher share of cashless payments is associated with a significantly higher likelihood of obtaining a loan, a significantly reduced interest rate, and a significantly higher loan amount.

[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, a one standard-deviation increase in the share of cashless payments corresponds to a likelihood that is 2 percentage points higher to get a loan approved, which represents 8 percent the baseline likelihood of getting approved in our sample. In terms of interest rates conditional on loan approval, the estimate in column 4 suggests that a one standard-deviation increase in cashless payments further corresponds to an interest rate that is lower by 36 basis points. Such an increase in cashless payment use is also associated with a 19% larger granted loan amount.

Without claiming causality for now, we provide a number of additional tests to help establish the robustness of the baseline results, which taken together show that the baseline results are unlikely to be driven by measurement errors or selection biases. These additional tests are all informed by our model and highlight the economic channel of information verifiability.

First, we repeat the baseline specification (4.1) by using the quintile of cashless shares with a given industry to predict lending outcomes, allowing for a potential non-linear relationship between cashless payment use and lending outcomes. We present the results in Figure A1, with the underlying regression shown in Table A3 in the Appendix. The results show that the positive effects of cashless payments on lending outcomes are generally present for all levels of cashless use, but most economically significant for the highest quintile which has the largest information verifiability τ_x. This pattern helps to mitigate potential measurement concerns because the highest quintile is least likely to be subject to measurement errors.

Second, we explore how the cashless share measured over different time horizons may affect lending outcomes. Recall that the lender requires all applicants to submit six months' payment records and use this full information set in screening loans. If we as econometricians were to use a cashless share measured in a shorter time horizon, the prediction power on lending outcomes would be lower due to a potentially lower information variability or a higher measurement error. Based on this idea, we repeat the baseline regression (4.1) using the share of cashless payments measured over three months, one month, and two weeks prior to the loan application. Table A4 presents the results, in which the three panels correspond to the three different time horizons discussed above. As expected, the effects of cashless payments on lending outcomes gradually decline as the measured time horizon becomes shorter, corresponding to a lower information verifiability τ_x. Importantly, this pattern helps to mitigate the concern that our baseline results are driven by applicants timing their application date given realized payment records, because all the applicants are always required to submit six months' payment records regardless of how we measure them.

Third, we show that the estimated results of cashless payments on lending outcomes are robust to alternative proxies for the use of cashless payments. 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 lending outcomes on the log of the number of cashless payments, with the same full set of controls as in the baseline regression (4.1). 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. Table A5 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. Overall, the results based on the additional empirical specifications are all consistent with the pattern in Table 2. They taken together help to further mitigate measurement errors and selection biases, further supporting Prediction 1.

4.2 Cashless Payments and Lending Outcomes: Information Verifiability

Economically, the key of cashless payments facilitating lending stems from the resulted payment records being potentially more verifiable, but not necessarily from the payment technology being physically cashless itself. 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 baseline specification (4.1) based on this breakdown of cashless payments. Table 3 reports the results. The coefficient in row 1 shows that higher use of information-intensive payments leads to a higher likelihood of loan approval, a lower interest rate conditional on a loan approval, and a higher loan amount. At the same time, higher use of information-light payments still leads to a lower interest rate compared to cash use conditional on a loan approval, but the economic magnitude is smaller than that for the use of information-intensive payments (row 2, column 4). Interestingly, higher use of information-light payments leads to a lower likelihood of loan approval compared to cash use, although the economic magnitude is small (row 2, column 2). This pattern can be potentially reconciled by the fact that informative-light payments, despite being cashless, do not provide an economically significant amount of verifiable information compared to cash, due to the loss of information in aggregation and anonymization. These results suggest that, economically, what matters for lending outcomes is indeed how verifiable payment records are, but not necessarily whether the payment technology itself is cashless or not.

[Insert Table 3 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.

4.3 Cashless Payments and Lending Outcomes: External Validity

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, built upon the baseline specification (4.1). Table 4 reports the results, which repeat our previous specifications with the full set of controls, as well as controlling for the share of cashless payments.[20] The results show that on top of the informational role of cashless payments, a higher share of verifiable merchant sales further and significantly increases the likelihood of a loan being approved, and significantly predicts a lower chance of loan default.

[Insert Table 4 here]

Table 4 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.

4.4 Cashless Payments and Lending Outcomes: 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 relies 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,

[20] 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.

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's 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.[21] 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.

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. As shown in Table A2 in the Appendix, firms that bank at a chest bank branch are indeed comparable to those not along various dimensions regarding firm fundamentals, consistent with the idea that most borrowers do not know if they bank at a chest bank branch. 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.[22] In columns 2 and 4, we allow the coefficient on our

[21] 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).

[22] 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

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 Kleibergen-Paap F-statistic for the first stage is around 28, above the threshold for weak instruments. The second-stage result is directionally consistent with the OLS analysis: based on this plausibly exogenous variation in the use of cashless payments at the borrower level, we find again that more use of cashless payments leads to a significantly higher likelihood of loan approval. In terms of magnitude, the instrumental variable provides a significantly larger coefficient than the OLS: a one-standard deviation in cashless payment use results in a 22 percentage points increase in the likelihood of getting a loan. Such estimates suggests that potential sources of endogeneity are biasing against the positive relationship between cashless payment use and loan application outcomes. This magnitude should however be taken interpreted cautiously, as the local average treatment might be inflated compared to the true causal effect, if access to banknotes within the chest bank branches was for firms of specific characteristics. Turning to the event-study type of instrumentation, we observe that the coefficient on the interaction between banking at a chest bank branch and the year 2016 is not significant, which provides a reassuring placebo test. In turn, the coefficients post demonetization are all significant, and particularly so for 2018 that represented the combination of the first and second cash shortage events after the initial Demonetization. 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]

applications from the sample when running these regressions.

4.5 Cashless Payments, Firm Quality, and Lending Outcomes

We move on to test Prediction 2 in regards to how firm quality (i.e., firm type) interacts with the effects of cashless payments on lending outcomes, which speaks to the information-revealing effect of using cashless payments. Built upon the baseline specification (4.1), we interact the share of cashless payments with firm quality, also using the full set of controls and fixed effects:

$$\begin{aligned}\text{LendingOutcome}_i &= \beta_1 \text{CashlessShare}_i + \beta_2 \text{FirmQuality}_i \\ &+ \beta_3 \text{CashlessShare}_i \times \text{FirmQuality}_i + \gamma X_i + \sum_k \theta_{F_k(i)} + \varepsilon_i \,. \end{aligned} \tag{4.2}$$

The main challenge in implementing specification (4.2) is to find a proxy for firm quality, which is not observable ex-ante. To this end, we construct two sets of empirical proxies using different information and show that the results are robust across those different proxies.

First, we use credit scores as an intuitive proxy for firm quality, based on the simple idea that credit scores as a public signal correlate with the true firm type z. Intuitively, a firm with a higher credit score is more likely to have a higher credit quality going forward, which may be more strongly revealed by a higher use of cashless payments. We thus interact the standardized share of cashless payments with credit scores, using specification (4.2). Table 6 reports the results. The regression coefficients in row 3 show that for firms with a higher credit score, a higher cashless share leads to an even higher likelihood of loan approval, a lower offered interest rate conditional on loan approval, and a higher loan amount, compared to firms with a lower credit score. These results jointly suggest a stronger information-revealing effect for firms with higher credit quality, consistent with Prediction 2.

[Insert Table 6 here]

Although intuitive, credit scores may not fully capture firm quality because a firm's credit-relevant economic activities may not be fully incorporated by past credit histories. In this aspect, the uniqueness of our payment data allows us to construct more granular proxies for firm quality by directly using the content of the payment records. Intuitively, the detailed content in payment records provides verifiable and incremental information about firm quality on top of usual credit scores, which is in turn consistent with the use of payment records by FinTech lenders in the first place.

Specifically, we construct two complementary proxies for firm quality based on the content of payment records: the monthly average balance level (scaled by revenues) and the weekly volatility of payment revenues.[23] Intuitively, a higher average level and a lower volatility of payment revenues are suggestive of a healthier business model, implying a higher firm quality. In this sense, they are economically valid proxies for firm type z in the model. We highlight that those two proxies based on payment content are economically different from the measure of cashless share we constructure earlier, which speaks to the information verifiability of payment records and proxies for τ_x.

We interact the standardized share of cashless payments with these two payment-content-based proxies for firm quality, using specification (4.2), with a full set of comprehensive controls and fixed effects including the credit score group fixed effect. Results are presented in Table 7. The findings are consistent with Prediction 2 and the information-revealing effect: using cashless payments is particularly beneficial to higher-quality firms as revealed by their payment records. Specifically, for applicant firms with a higher average payment revenue and lower revenue volatility, the estimates of the interaction term in columns 1 and 2 suggest that a higher share of cashless payments leads to an even higher likelihood of loan approval. This relationship is unlikely to be driven by other firm-level characteristics given the comprehensive set of fixed effects that we include. Similarly, columns 3 and 4 suggest that this differential effect is also present at the intensive margin of lending outcomes, that is, offered interest rates. Specifically, a higher share of cashless payments leads to an even lower interest rate for firms with a higher average payment revent and lower revenue volatility.

[Insert Table 7 here]

To further show the robustness of the information-revealing effect, we perform a similar test following the specifications in Table 3, taking advantage of our unique data of different payment technologies with different levels of verifiability intensity. Table 8 presents the regression coefficients. For higher-type applicant firms, i.e. those with a higher average payment revenue and lower revenue volatility, higher use of information-intensive cashless payments increases the likelihood of loan approval and decreases the offered interest rate even more, further supporting Prediction 2.

[Insert Table 8 here]

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

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.

4.6 Cashless Payments and Defaults

Finally, we test Prediction 3 by relating the use of cashless payments to loan defaults. If cashless payments indeed lead to more efficient loan screening, we should expect a borrower to be less likely to default if a loan is granted, which bears implications for the efficiency of capital allocation.

To this end, we consider the following specification 4.3 by regressing an indicator variable of a loan having defaulted as of September 2019 in our sample on the standardized share of cashless payments with our standard set of controls:

$$\text{Default}_i = \beta \text{CashlessShare}_i + \gamma X_i + \sum_k \theta_{F_k(i)} + \varepsilon_i \,. \tag{4.3}$$

Table 9 reports the results. Consistent with Prediction 3, we find that a robust pattern of a higher share of cashless payments leading to a lower likelihood of loan default. Particularly, we further include the share of revenues from online marketplace, as used in Table 4, in columns 3 and 6. The coefficients suggest that a higher use of verifiable outside information other than cashless payments also significantly negatively predicts lower defaults, further supporting the external validity of our framework. Notably, the coefficients of cashless share in columns 3 and 6 also preserve the same statistical and economical significance, suggesting that different sources of outside information, as long as being verifiable, may independently and incrementally help improve lending outcomes without weakening each other's informational role. Taken together, these findings suggest that the use of outside verifiable information leads to more efficient screening and lending decisions.

[Insert Table 9 here]

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 and a higher loan amount. 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 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

Acemoglu, D., A. Makhdoumi, A. Malekian and A. Ozdaglar. 2020. Too much data: prices and inefficiencies in data markets. *American Economic Journal: Microeconomics*, forthcoming.

Acharya, V., P. DeMarzo, and I. Kremer. 2011. Endogenous information flows and the clustering of announcements. *American Economic Review* 101:2955-2979.

Admati, A. and P. Pfleiderer. 1990. Direct and indirect sale of information. *Econometrica* 58:901-928.

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.

Bartlett, R., A. Morse, R. Stanton, and N. Wallace. 2021. Consumer-lending discrimination in the FinTech era. *Journal of Financial Economics*, forthcoming.

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.

Berg, T., Fuster, A. and Puri, M. 2021. FinTech lending. *Annual Review of Financial Economics*, forthcoming.

Bergemann, D and A. Bonatti. 2019. Markets for information: an introduction. *Annual Review of Economics* 11:85-107.

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.

Björkegren, D. and D. Grissen. 2018. The potential of digital credit to bank the poor. *American Economic Review* 108:68-71.

Björkegren, D. and D. Grissen. 2020. Behavior revealed in mobile phone usage predicts credit repayment. *World Bank Economic Review* 34:618-634.

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

Bolton, P., H. Chen, and N. Wang. 2011. A unified theory of Tobin's q, corporate investment, financing, and risk management. *Journal of Finance* 66:1545-1578.

Bonatti, A. and G. Cisternas. 2020. Consumer scores and price discrimination. *Review of Economic Studies*, 87:750-791.

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.

Calomiris, C., and C. Kahn. 1991. The role of demandable debt in structuring optimal banking arrangements. *American Economic Review* 81:497-513.

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.

Cong, L. W., D. Xie, and L. Zhang. 2021. Knowledge accumulation, privacy, and growth in a data economy. *Management Science*, forthcoming.

Cong, L. W., W. Wei, D. Xie, and L. Zhang. 2021. Endogenous growth under multiple uses of data. Working paper.

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.

Diamond, D. and R. Rajan. 2001. Liquidity risk, liquidity creation, and financial fragility: A theory of banking. *Journal of Political Economy* 109:287-327.

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

Drechsler, I., A. Savov, and P. Schnabl. 2017. The deposits channel of monetary policy. *Quarterly Journal of Economics* 132: 1819-1876.

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. 2021. Open banking: credit market competition when borrowers own the data. Working paper.

Higgins, S. 2020. Financial technology adoption. *American Economic Review*, forthcoming.

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.

Kashyap, A., R. Rajan, and J. Stein. 2002. Banks as liquidity providers: An explanation for the coexistence of lending and deposit-taking. *Journal of Finance* 57: 33-73.

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.

Moulton, H. 1918. Commercial banking and capital formation. *Journal of Political Economy* 26:484-508, 638-663, 705-631, 849-881.

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.

Pagano, M. and T. Jappelli. 2019. Information sharing in credit markets. *Journal of Finance* 48:1693-1718.

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.

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.

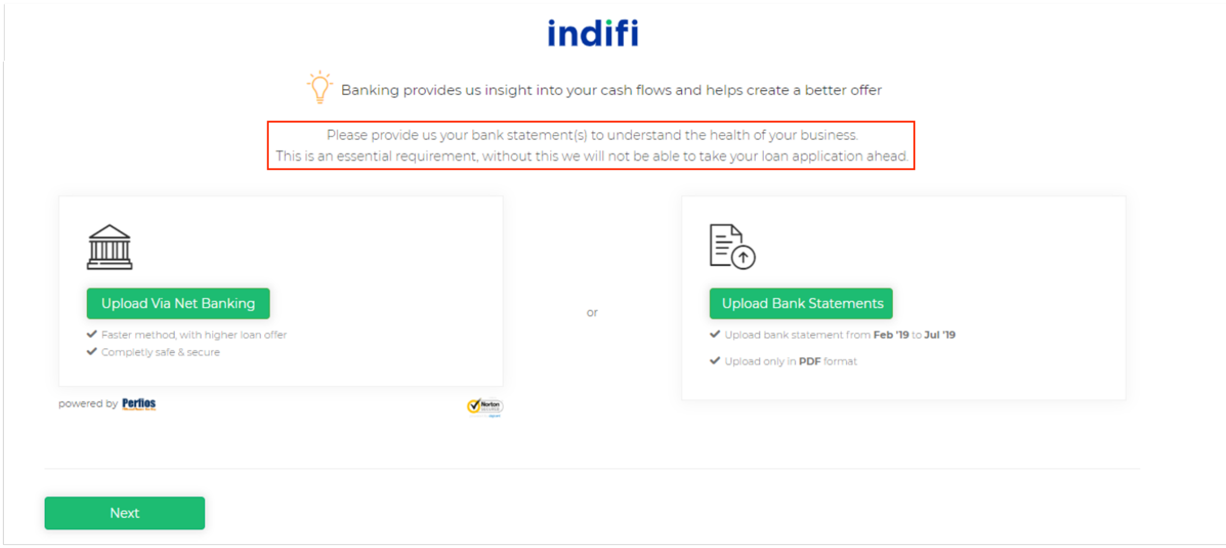
Rampini, A., and S. Viswanathan. 2010. Collateral, risk management, and the distribution of debt capacity. *Journal of Finance* 65:2293-2322.

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.

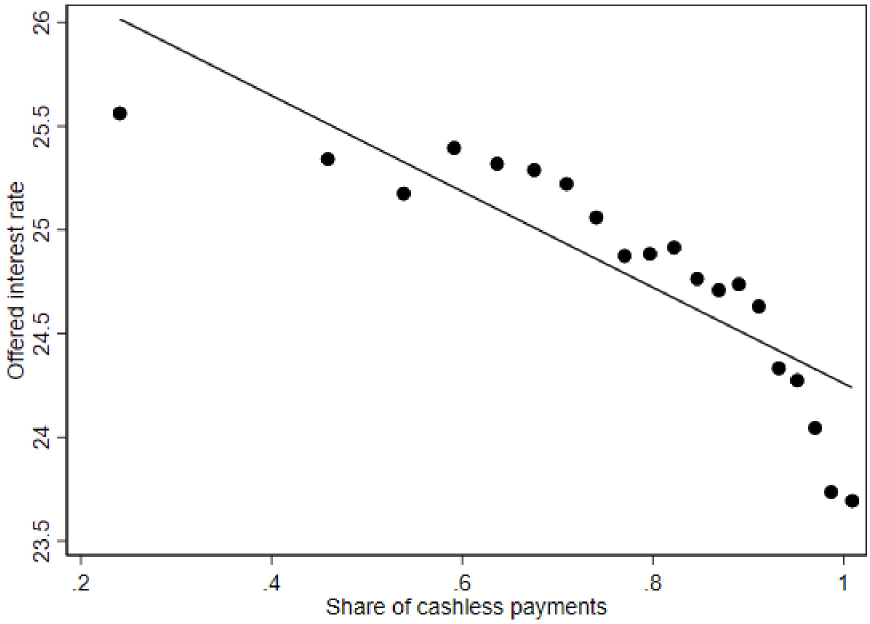
Vives, X. 2019. Digital disruption in banking. *Annual Review of Financial Economics* 11:243-272.

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.

Table 1: **Summary statistics**

	Obs	**Mean**	**Std Dev**	**p25**	**p75**
		Panel A. Loan Applicant Characteristics			
Year of Application	71,293	2018	0.7	2018	2019
Applicant Age	71,293	35.5	9.1	29.1	40.5
Business Age	71,293	4.7	13.8	1.4	5.9
Credit History Length (in Years)	63,433	7.0	5.5	2.4	11.1
Cibil Score	57,762	670	135	655	724
Missing Cibil Score (0/1)	71,293	0.188	0.391	0	0
		Panel B. Transaction Data			
Share of Cashless Payments	71,227	0.717	0.221	0.585	0.894
of which: Information-intensive Payments	71,227	0.541	0.273	0.341	0.761
Share of Cash Payments	71,227	0.283	0.221	0.106	0.415
Share of Revenues from Online Marketplace	20,141	0.383	0.320	0.116	0.586
Borrower Banks at Chest Banks (0/1)	71,227	0.039	0.193	0	0
Avg. Monthly Revenue (INR)	71,227	1,615,696	12,880,247	114,154	823,622
Avg. Monthly Balance (INR)	71,227	38,141	108,915	9,613	66,655
Avg. # of Transactions	71,227	554	1334	155	600
Weekly Revenue Volatility	71,227	30,003	89,924	6,097	27,789
		Panel C. Loan Application Outcomes			
Approved Loan (0/1)	71,293	0.306	0.461	0	1
Offered Interest Rate	21,746	24.8	3.3	23.4	27
Amount (INR)	14,227	421,223	642,058	100,000	500,000
Default (0/1)	13,499	0.049	0.216	0	0

Notes: This table reports summary statistics for the main variables used in the regressions. The sample consists of all complete applications to Indifi between september 2015 and september 2019. Panel A reports summary statistics of loan applicants by years of application, age, credit score and credit history. Panel B displays summary statistics of transactions data, obtained through disaggregation of bank statements, and covers payment technologies, bank type, monthly average revenue, monthly average balance and weekly revenue volatility. Panel C reports lending outcomes of the applications.

Table 2: **Cashless payments and lending outcomes**

	Approved Loan (1/0)		Offered Interest Rate		log(Amount)	
	(1)	(2)	(3)	(4)	(5)	(6)
Share of cashless payments	0.023***	0.019***	-0.391***	-0.360***	0.197***	0.186***
	(0.002)	(0.002)	(0.030)	(0.029)	(0.013)	(0.013)
Borrower controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Application month FE	Yes	Yes	Yes	Yes	Yes	Yes
Cibil score group FE	No	Yes	No	Yes	No	Yes
3-digit zipcode FE	No	Yes	No	Yes	No	Yes
Revenue deciles FE	No	Yes	No	Yes	No	Yes
Observations	70,394	69,853	21,663	21,610	14,064	14,003
R^2	0.265	0.309	0.430	0.473	0.281	0.330

Notes: This table presents coefficients from OLS regressions that regress an indicator for having its loan application approved (column 1 and 2), the interest rate offered (column 3 and 4), and the loan amount offered (column 5 and 6), on the share of cashless payments. The share of cashless payments is calculated as per section 3.2, and is then standardized. The set of borrower controls includes log # of payments, credit history length, business vintage, Log of owner's age, missing credit score indicator and top-up loan indicator. There are 67 industries. Cibil score groups represent 10 points bands of Cibil Score, the Indian equivalent of FICO score. 3-digit zipcode are the first 3 digit of the Indian Pincode. Revenues are calculated as the sum of inflows on the bank account over the last six months. Standard errors are clustered at application-month level and displayed below their coefficient of interest. *, **, and *** represent statistical significance at the 10%, 5%, and 1% confidence levels, respectively.

Table 3: **Intensity of payment verifiability and lending outcomes**

	Approved Loan (1/0)		Offered Interest Rate		log(Amount)	
	(1)	(2)	(3)	(4)	(5)	(6)
Share of information-intensive payments	0.041***	0.033***	-0.432***	-0.491***	0.273***	0.267***
	(0.001)	(0.002)	(0.027)	(0.036)	(0.011)	(0.014)
Share of information-light payments		-0.012***		-0.104***		-0.011
		(0.002)		(0.031)		(0.015)
Borrower controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Application month FE	Yes	Yes	Yes	Yes	Yes	Yes
Cibil score group FE	Yes	Yes	Yes	Yes	Yes	Yes
3-digit zipcode FE	Yes	Yes	Yes	Yes	Yes	Yes
Revenue deciles FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	69,853	69,853	21,610	21,610	14,003	14,003
R^2	0.314	0.314	0.477	0.477	0.359	0.359

Notes: This table presents coefficients from OLS regressions that regress an indicator for having its loan application approved (column 1 and 2), the interest rate offered (column 3 and 4), and the loan amount offered (column 5 and 6), on the share of cashless payments. The share of information-intensive payments and the share of information-light payments are calculated as per section 3.2, and are then standardized. 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. There are 67 industries. Cibil score groups represent 10 points bands of Cibil Score, the Indian equivalent of FICO score. 3-digit zipcode are the first 3 digit of the Indian Pincode. Standard errors are clustered at application-month level and displayed below their coefficient of interest. *, **, and *** represent statistical significance at the 10%, 5%, and 1% confidence levels, respectively.

Table 4: **Alternative verifiable information and lending outcomes**

	Approved Loan (1/0) (1)	Offered Interest Rate (2)	log(Amount) (3)
Share of revenues from online marketplace	0.077***	0.024	0.076***
	(0.003)	(0.028)	(0.010)
Share of cashless payments	0.021***	-0.365***	0.210***
	(0.002)	(0.029)	(0.014)
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 zipcode FE	Yes	Yes	Yes
Revenue deciles FE	Yes	Yes	Yes
Interest rate FE	No	No	Yes
Observations	69,853	21,610	9,040
R^2	0.304	0.457	0.374

Notes: This table presents coefficients from OLS regressions that regress an indicator for having its loan application approved (column 1 and 2), the interest rate offered (column 3 and 4), and the loan amount offered (column 5 and 6), on the share of revenues originating from an online marketplace that has a data-sharing agreement with Indifi. The share of revenues from online marketplace is calculated as the average monthly revenue on the partner online marketplace divided by the average total monthly revenue as per the banking statements, and is then standardized. The share of cashless payments is standardized as well. 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. There are 67 industries. Cibil score groups represent 10 points bands of Cibil Score, the Indian equivalent of FICO score. 3-digit zipcode are the first 3 digit of the Indian Pincode. Standard errors are clustered at application-month level and displayed below their coefficient of interest. *, **, and *** represent statistical significance at the 10%, 5%, and 1% confidence levels, respectively.

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

	Share of Cashless Payments			Approved Loan (1/0)	
	(1)	(2)	(3)	(4)	(5)
Borrower banks at Chest Bank Branch (1/0)	-0.095***	0.055			
	(0.018)	(0.124)			
Chest Bank Branch (1/0) × 2016			0.057		
			(0.113)		
Chest Bank Branch (1/0) × 2017			-0.078**		
			(0.036)		
Chest Bank Branch (1/0) × 2018			-0.125***		
			(0.022)		
Chest Bank Branch (1/0) × 2019			-0.069**		
			(0.027)		
Share of cashless payments				0.249***	0.221***
				(0.094)	(0.077)
Borrower controls	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Application month FE	Yes	Yes	Yes	Yes	Yes
Cibil score group FE	Yes	Yes	Yes	Yes	Yes
Revenue deciles FE	Yes	Yes	Yes	Yes	Yes
Observations	68,400	1,978	70,390	68,400	70,390
F-statistic	28.59	-	14.26	-	-

Notes: This table presents regression coefficients for a 2SLS specification. Column 1 presents the first stage, where we regress the share of cashless payments on an indicator variable for the borrower banking at a chest bank branch. The sample is restricted to applications happening after the demonetization (december 2016). Column 2 conducts the same analysis on applications prior to the demonetization. Column 3 presents another first stage specification, where the share of cashless payments is regressed on the interaction between the indicator variable for banking at chest bank branch and year fixed effects, using the whole sample. Column 4 and 5 displays the second stage of the 2SLS corresponding to columns 1 and 3, where we regress the indicator variable for getting the loan application approved on the instrumented *Share of Cashless Payments*. The share of cashless payments is calculated as per section 3.2, and is then standardized. 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. There are 67 industries. Cibil score groups represent 10 points bands of Cibil Score, the Indian equivalent of FICO score. Standard errors are clustered at the application-month level and displayed below their coefficient of interest. *, **, and *** represent statistical significance at the 10%, 5%, and 1% confidence levels, respectively.

Table 6: **Cashless payments, credit score, and lending outcomes**

	Approved Loan (1/0) (1)	Offered Interest Rate (2)	log(Amount) (3)
Share of cashless payments	0.023***	-0.322***	0.196***
	(0.002)	(0.034)	(0.014)
Cibil Score	0.057***	-0.402***	0.144***
	(0.004)	(0.030)	(0.012)
Share of cashless payments × Cibil Score	0.012***	-0.160***	0.041**
	(0.002)	(0.034)	(0.017)
Borrower controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Application month FE	Yes	Yes	Yes
3-digit zipcode FE	Yes	Yes	Yes
Revenue deciles FE	Yes	Yes	Yes
Observations	57,239	18,953	8,153
R^2	0.258	0.449	0.367

Notes: This table presents coefficients from OLS regressions that regress an indicator for having its loan application approved (column 1), the interest rate offered (column 2), and the loan amount offered (column 3), on the share of revenues originating from an online marketplace that has a data-sharing agreement with Indifi. The share of cashless payments and the Cibil Score (Indian equivalent to FICO score) are standardized. 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. There are 67 industries. 3-digit zipcode are the first 3 digit of the Indian Pincode. Standard errors are clustered at application-month level and displayed below their coefficient of interest. *, **, and *** represent statistical significance at the 10%, 5%, and 1% confidence levels, respectively.

Table 7: **Cashless payments, payment revenue, and lending outcomes**

	Approved Loan (1/0)		Offered Interest Rate	
	(1)	(2)	(3)	(4)
Share of cashless payments	0.019***	0.019***	-0.347***	-0.294***
	(0.002)	(0.002)	(0.031)	(0.031)
Average monthly balance / revenue	0.005**		0.043	
	(0.002)		(0.030)	
Share of cashless payments × Average monthly balance / revenue	0.005***		-0.035*	
	(0.001)		(0.018)	
Weekly revenue volatility		-0.035***		-0.121*
		(0.004)		(0.063)
Share of cashless payments × Weekly revenue volatility		-0.002*		0.164***
		(0.001)		(0.040)
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 zipcode FE	Yes	Yes	Yes	Yes
Revenue deciles FE	Yes	Yes	Yes	Yes
Observations	52,681	52,695	17,541	17,548
R^2	0.311	0.312	0.490	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 variable share of cashless payments is standardized. 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 and displayed below their coefficient of interest. *, **, and *** represent statistical significance at the 10%, 5%, and 1% confidence levels, respectively.

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

	Approved Loan (1/0)		Offered Interest Rate	
	(1)	(2)	(3)	(4)
Share of information-intensive payments	0.034***	0.035***	-0.461***	-0.405***
	(0.003)	(0.003)	(0.038)	(0.037)
Share of information-light payments	-0.014***	-0.014***	-0.123***	-0.140***
	(0.002)	(0.002)	(0.038)	(0.039)
Share of information-intensive payments × Average monthly balance / revenue	0.004***		0.006	
	(0.001)		(0.026)	
Average monthly balance / revenue	0.006***		0.024	
	(0.002)		(0.035)	
Share of information-intensive payments × Weekly revenue volatility		-0.006***		0.192***
		(0.002)		(0.044)
Weekly revenue volatility		-0.039***		-0.174**
		(0.004)		(0.068)
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 zipcode FE	Yes	Yes	Yes	Yes
Revenue deciles FE	Yes	Yes	Yes	Yes
Observations	52,681	52,695	17,541	17,548
R^2	0.316	0.318	0.492	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. Both the share of information-intensive payments and the share of information-light payments are standardized. 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 and displayed below their coefficient of interest. *, **, and *** represent statistical significance at the 10%, 5%, and 1% confidence levels, respectively.

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

	Default (1/0)					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of cashless payments	-0.019***	-0.018***	-0.018***	-0.015***	-0.015***	-0.016***
	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)
Share of revenues from online marketplace			-0.008***			-0.009***
			(0.003)			(0.003)
Borrower controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Application month FE	Yes	Yes	Yes	Yes	Yes	Yes
Cibil score group FE	No	Yes	Yes	No	Yes	Yes
3-digit zipcode FE	No	Yes	Yes	No	Yes	Yes
Revenue deciles FE	No	Yes	Yes	No	Yes	Yes
Observations	9,138	9,069	9,069	9,084	9,014	9,014
R^2	0.060	0.100	0.101	0.071	0.109	0.110

Notes: This table presents OLS regressions that use the share of cashless payments to predict loan default rate. The variable share of cashless payments is standardized. 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 and displayed below their coefficient of interest. *, **, and *** represent statistical significance at the 10%, 5%, and 1% confidence levels, respectively.

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 $\rho = 0$ in the proof to focus on the economic environment where the information-revealing effect dominates. When $\rho > 0$, the following two key inequalities (A.1) and (A.2) hold strictly, yielding the same equilibrium. Intuitively, if all firm types find it optimal to adopt cashless payments without the risk-reducing effect, adding the risk-reducing effect will make all firm types even more willing to adopt.

Formally, 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

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

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

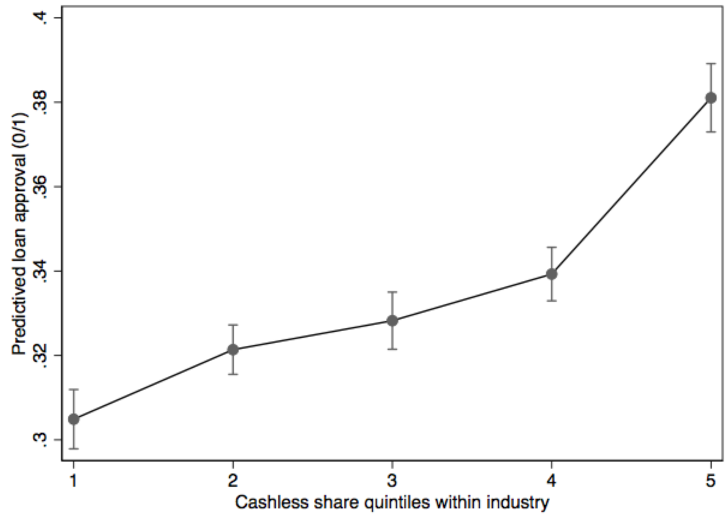
the three of which jointly imply that

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

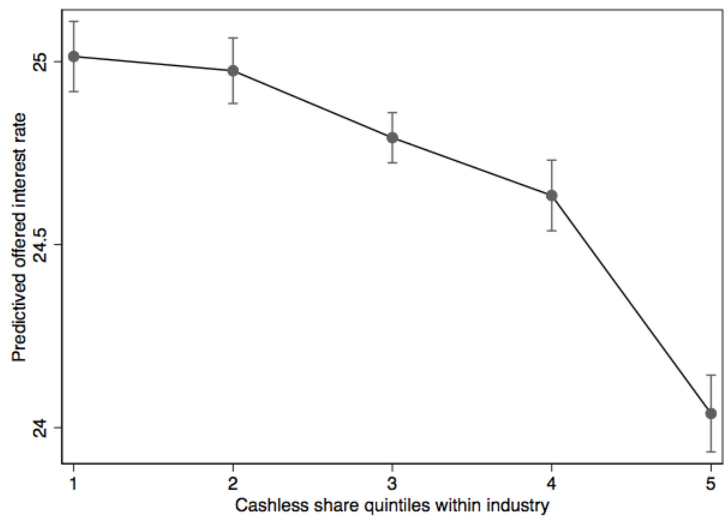
for any $\zeta^* > -\infty$. Because $\frac{n\tau_s}{\tau_z+n\tau_s} < 1$, (A.1), (A.2) and property i) above then jointly imply that $\delta(z^*) > 0$ for all $-\infty \leq \zeta^* < 0$. On the other hand, (A.1) also directly means that $\delta(z^*) > 0$ holds for all $\zeta^* \geq 0$. Thus, $\delta(z^*)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



Cashless share and offered interest rate



Notes: The figure in Panel A 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. The figure in Panel B 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 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
Panel C. Measured over the 2 weeks before application				
Share of cashless payments	0.084***	0.064***	-1.281***	-1.129***
	(0.011)	(0.011)	(0.119)	(0.118)
Observations	30,981	30,755	12,585	12,528
R^2	0.202	0.255	0.433	0.477
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 A5: **Number of cashless payments and lending outcomes**

	Approved Loan (1/0)	Offered Interest Rate
	(1)	(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 A6: **Sanctioned amount / revenue and cashless payment**

	Sanctioned amount / revenue	
	(1)	(2)
Share of cashless payments	456.231**	413.364**
	(175.022)	(156.945)
Average monthly balance / revenue		546.940**
		(251.260)
Share of cashless payments × Average monthly balance / revenue		650.494**
		(270.259)
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	52,681
R^2	0.010	0.013

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 Model Extensions and Alternatives

IA.1 Costly Adoption of Cashless Payments

In this appendix, we extend the baseline model by incorporating a net cost of adopting cashless payments, due to physical (e.g., the cost of purchasing a mobile phone) or implicit (e.g., privacy concerns) reasons. We note that cashless payments also typically provide convenience benefits to adopters. Thus, the mere existence of those physical or implicit costs may not necessarily suggest that the net cost of adopting is positive. The combination of all the benefits and costs, in reality, may well imply that the net cost of adoption is zero or even negative, falling into our baseline model.

Under this extended model, we can characterize the optimal adoption decision as follows:

PROPOSITION **IA1.** *Suppose the net cost of adopting cashless payment is finite and positive: $0 < c \leq \infty$. In a monotone equilibrium such that firm types $z \geq z^*$ adopt cashless payments while firm types $z \leq z^*$ use cash, there must exist $z^* \geq -\infty$, meaning that at least some high-type firms will optimally adopt cashless payments.*

PROOF OF PROPOSITION IA1. Consider equation

$$p(z^*; z \geq z^*) - c = p(z \leq z^*) \tag{IA.1}$$

where $p(z^*; z \geq z^*)$ and $p(z \leq z^*)$ are given in (2.4) and (2.6). Notice that both sides of (IA.1) are continuous in z^*. When z^* approaches positive infinity, the left-hand side of (IA.1) approaches positive infinity, but the right-hand side remains bounded, implying that the left-hand side is greater than the righ-hand side when z^* is sufficiently large. Thus, if equation (IA.1) does not have a solution, it implies that $z^* = -\infty$, that is, all firm types optimally adopt cashless payments despite the net adoption cost. Otherwise if equation (IA.1) has a solution of $z^* > -\infty$, because $p(z^*; z \geq z^*)$ as given in (2.5) increases in z, it implies that higher firm types $z \geq z^*$ adopt cashless payments while lower firm types $z \leq z^*$ use cash. $\square$

The key idea underlying Proposition IA1 is that the strategic consideration of standing out of non-adopting firm types still pushes more firm types to adopt, although some low-type firms may find the net cost to be too large to be justified by the informational benefits on the lending market. Intuitively, after paying the net adoption cost, the cutoff firm type just breaks even in expectation, and all the higher types find it profitable to adopt cashless payments.

Despite some low-type firms may not end up adopting cashless payments with the net adoption cost, we nonetheless show that as long as some high-type firms adopt cashless payments, our qualitative predictions of cashless payments improving lending outcomes would not be affected.

PROPOSITION **IA2.** *In a monotone equilibrium such that firm types $z \geq z^*$ adopt cashless payments while firm types $z \leq z^*$ use cash with $z^* > -\infty$, the average expected price improvement $E[\Delta p(z)]$ from choosing cashless payments is increasing in τ_x and n, and a higher firm type enjoys a higher expected price improvement from choosing cashless payments, and the increase is higher when τ_x or n increases in the sense that*

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

PROOF OF PROPOSITION IA2. As $z^* > -\infty$, the equilibrium path features the separation of firm types who adopt cashless payments and those who do not. Thus, the expected price improvement from adopting cashless payments can be re-defined as

$$\Delta p(z) = p(z; z \geq z^*) - p(z \leq z^*) + \frac{\rho}{2}\left(E[Var[z|z \leq z^*]|z] - E[Var[z|X, z \geq z^*]|z]\right) \quad \text{(IA.2)}$$

where $p(z; z \geq z^*)$ and $p(z \leq z^*)$ are given in (2.5) and (2.6).

Although the truncated normal distribution does not have a conjugate prior distribution and thus unfortunately (IA.2) does not admit a closed-form solution, we can re-write the prior probability density of the corresponding truncated normal distribution $N(\mu, \tau_z^1, z \geq z^*)$:

$$f(z) = \lambda(z^*)\phi\left((z - \mu(z^*))\sqrt{\tau(z^*)}\right)\mathbf{1}_{z \geq z^*}\,,$$

where ϕ is the density function of a standard normal distribution, and the normalizing factor λ and the effective mean $\mu(z^*)$ and prevision $\tau(z^*)$ of the truncated distribution only depend on z^* but not on τ_x or n. By Bayes's rule, the posterior probability density satisfies:

$$f(z|X) \propto \exp\left(-\frac{1}{2}\left(\tau(z^*)+n\tau_s\right)\left(z-\frac{\tau(z^*)\mu(z^*)+n\tau_s\bar{x}}{\tau(z^*)+n\tau_s}\right)^2\right)\mathbf{1}_{z\geq z^*}\,,$$

where $\bar{x}=\frac{1}{n}\sum_{t=0}^{n-1}x_t$ is the sample mean of X.

Note that the second and third terms in (IA.2) do not depend on either τ_x or n, and the fourth term does not depend on z. By Fatou-Lebesgue Theorem, we therefore have

$$\begin{aligned}
sign\left(\frac{\partial E[\Delta p(z)]}{\partial(n\tau_s)}\right) &= -sign\left(\frac{\partial E[E[Var[z|X,z\geq z^*]|z]]}{\partial(n\tau_s)}\right)\\
&= -sign\left(\frac{\partial\frac{1}{\tau(z^*)+n\tau_s}}{\partial(n\tau_s)}\right)\\
&> 0\,.
\end{aligned}$$

Similarly,

$$\begin{aligned}
sign\left(\frac{\partial^2\Delta p(z)}{\partial z\partial(n\tau_s)}\right) &= sign\left(\frac{\partial^2 p(z;z\geq z^*)}{\partial z\partial(n\tau_s)}\right)\\
&= sign\left(\frac{\partial^2\frac{\tau(z^*)\mu(z^*)+n\tau_s z}{\tau(z^*)+n\tau_s}}{\partial z\partial(n\tau_s)}\right)\\
&> 0\,,
\end{aligned}$$

yielding the desirable results. □

The key idea is that as long as a pool of high-type firms uses cashless payments, higher informational verifiability leads to better lending outcomes, and such positive effects are stronger for higher firm types. Proposition IA2 shows that this simple intuition is robust, despite that the expected price improvement $\Delta p(z)$ does not anymore admit a closed-form solution as in the baseline model. Thus, the baseline model is a preferable benchmark to explicitly illustrate the underlying economic forces.

IA.2 Self-Selection into FinTech Lenders from Banks

In this appendix, we provide another stripped-down extension of the baseline model by incorporating competition between a FinTech lender and a bank. Specifically, we allow for an endogenous choice of the firm between a traditional bank loan and FinTech credit. We show that the presence of bank competition effectively constructs an opportunity cost for adopting cashless payments and applying for FinTech credit, but the resulted self-selection would not affect the qualitative predictions of our baseline model.

To fix ideas and highlight the comparative advantage of FinTech lenders in using outside verifiable information, we consider a benchmark case of the bank not using outside information but having a potentially more precise prior $\tau_b > \tau_z$ due to its production of inside information. To keep other dimensions comparable, we assume that the bank is also risk-averse and has the same absolute risk aversion ρ as the FinTech lender. The bank and the FinTech lender together face the same pool of firms types as in the baseline model, and the firm at $t = 0$ also chooses whether it plans to apply for a traditional bank loan or FinTech credit in the financing stage.

Under this extended model, we show that the firm's optimal lender choice can be characterized as follows:

Proposition **IA3.** *There exists a monotone equilibrium such that firm types $z \geq z_b^*$ choose the FinTech lender and adopt cashless payments while firm types $z \leq z_b^*$ choose the bank and use cash, in which $z_b^* \geq -\infty$.*

The proof of Proposition IA3 is similar to that of Proposition IA1, and we illustrate the intuition here. Since the bank does not use payment records to screen firms, it always offers an uninformed price on any equilibrium path. If no firm type adopts cashless payments, the uninformed price from the bank is higher than that from the FinTech lender:

$$\mu - \frac{\rho}{2}\frac{1}{\tau_b} > \mu - \frac{\rho}{2}\frac{1}{\tau_z},$$

and this difference effectively constructs a net cost of adopting cashless payments and subsequently applying for the FinTech credit. If such a difference is large enough, it may discourage some low-type firms from using cashless payments, and they consequently apply for a traditional bank loan in the financing stage, getting the uninformed price.

An important implication of Proposition IA3 is that at least some high-type firms will adopt cashless payments and apply for FinTech credit, despite the competition from the bank that may attract some low-type firms. We highlight that the hidden type in our model captures a form of firm quality that can be efficiently conveyed through payment records but may not be efficiently recognized by traditional banks. Therefore, this prediction is indeed consistent with the empirical pattern that FinTech borrowers typically do not have a good enough prior from banks, potentially due to a lack of past bank relationships, but may benefit from FinTech presence thanks to the use of outside verifiable information.

Another important implication of Proposition IA3 is that despite the competition from banks, the equilibrium still admits a cutoff structure. Consequently, Proposition IA2 applies, which implies that all the qualitative predictions of our baseline model regarding the effects of cashless payments on lending outcomes are unchanged.

IA.3 Explicit Debt Financing

In this appendix, we build an alternative model with otherwise a similar setting as the baseline model, but with more explicit debt financing features such as interest rates and defaults. We show that this alternative modeling approach generates qualitatively the same predictions about the synergy between FinTech lending and cashless payments.

There are two reasons why we only consider this model version as an alternative rather than as our baseline model. First, given the focus on the informational effects of payments on financing outcomes, we believe that our current baseline model is more general in the sense that the financing price p can be easily translated into a notion of cost of capital, and thus allows our model to speak to both debt and equity financing. Second, modeling defaults explicitly prevents modeling the financier as being risk-averse, and thus we are unable to illustrate the risk-reducing and information-revealing effects as transparent as we do in the 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 enough 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.[24]

The equilibrium concept we consider is a standard sequential equilibrium. Specifically, the equilibrium profile consists of the entrepreneur's payment choice policy $n(\pi) : [0, 1] \rightarrow$

[24]The existence of reservation interest rate is necessary to ensure an interior solution of equilibrium interest rate because the lender is a monopoly; see Pagano and Jappelli (1993) and Parlour, Rajan, and Zhu (2020) for similar assumptions. It can be motivated by the competition from another unmodeled lender.

$\{0, T\}$, 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 $[0, R_H - 1]$.[25]

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 **IA1.** *The lender's optimal interest rate is given by:*

$$r^* = \frac{1 - R_L}{2\tilde{\pi}} + \frac{R_H + R_L - 2}{2} \,, \tag{IA.3}$$

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

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

[25] The reservation interest rate cannot be higher than $R_H - 1$ because if so, the entrepreneur will get a negative surplus even in the good state.

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

$$\tilde{\pi} > \frac{1 - R_L}{R_H - R_L}\,. \tag{IA.4}$$

Lemma IA1 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.4)) and offers a lower interest rate (as suggested by condition (IA.3)). We can therefore focus on how the entrepreneur' payments record affects the lender's posterior beliefs.

Now we can 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 $\pi \in [0, 1]$.[26]

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 IA1 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 **IA4.** *For an entrepreneur of type π, the expected interest rate she gets by choosing cashless payments is*

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

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

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

$$\frac{\partial^2 r^*(\pi, T)}{\partial \pi \partial T} < 0\,. \tag{IA.5}$$

Proof of Proposition IA4. 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

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

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.6) 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,

$$E[\tilde{\pi}|\pi, T] = \frac{1+\pi T}{2+T}\,, \tag{IA.7}$$

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 IA1, 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.7). Direct calculation yields the desired results. $\square$

The results in Proposition IA4 lead to two empirical predictions, which are in line with the predictions of our baseline model. First, the comparative statics (IA.5) 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 IA1, it further implies that better entrepreneurs are more likely to be granted a loan when establishing more verifiable payment records.

We move on to illustrate the other direction of the synergy between lending and cashless payments. Similarly, it simultaneously verifies the sub-game equilibria we focus on above.

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:

$$\tilde{\pi}(T, x) = E[\pi | \pi \geq \pi^*, T, x] \,. \tag{IA.8}$$

Then by Lemma IA1, the expected offered interest rate to type π^* who commits to establish T payment records at $t = 0$ is

$$r^*(\pi^*, T) = \frac{1 - R_L}{2E[\tilde{\pi}(T, x)|\pi^*]} + \frac{R_H + R_L - 2}{2} \,, \tag{IA.9}$$

where $\tilde{\pi}(T, x)$ is given by (IA.8) 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

$$I(\pi^*, T) = E_{\bar{r}}[\min(r^*(\pi^*, T), \bar{r})] \,, \tag{IA.10}$$

where $r^*(\pi^*, T)$ is given by (IA.9) and the outer expectation is taken with respect to the reservation interest rate $\bar{r}$.[27]

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 IA1 again, the expected offered interest rate to type π^* who chooses cash

[27] 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.12).

at $t = 0$ (and thus establishes none payment record) is

$$r^*(\pi^*, 0) = \frac{1 - R_L}{2E[\pi|\pi \leq \pi^*]} + \frac{R_H + R_L - 2}{2}\,, \tag{IA.11}$$

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

$$I(\pi^*, 0) = E_{\bar{r}}[\min(r^*(\pi^*, 0), \bar{r})]\,, \tag{IA.12}$$

where $r^*(\pi^*, 0)$ is given by (IA.11). Now using expressions (IA.10) and (IA.12) and applying the cutoff entrepreneur type's indifference condition, we have the following formal result:

Proposition **IA5.** *There exists a cutoff entrepreneur* $\pi^* \in [0, 1]$ *who is indifferent between using cashless payments or cash:*

$$\pi^* I(\pi^*, T) + C = \pi^* I(\pi^*, 0)\,, \tag{IA.13}$$

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 IA5. 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.9)) and $r^*(\pi, 0)$ for entrepreneur-type π who uses cash (see condition (IA.11)) 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

$$\Delta V(\hat{\pi}) = V_{cashless}(\hat{\pi}) - V_{cash}(\hat{\pi}) \tag{IA.14}$$

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:

$$\begin{aligned} \Delta V(1) &= E_{\bar{r}}[\min(r^*(E[\pi], 0), \bar{r})] - E_{\bar{r}}[\min(r^*(1, T), \bar{r})] - C && \text{(IA.15)} \\ &\geq E_{\bar{r}}[\min(r^*(E[\pi], 0), \bar{r})] - E_{\bar{r}}[\min(r^*(1, 1), \bar{r})] - C\,, && \text{(IA.16)} \end{aligned}$$

where equation (IA.15) 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.16) 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.16) 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 IA4. 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.16) 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.14), 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.13) 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.13) 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 banks' websites are identified by the relevant strings, including large corporate transfers and any transfers between saving/checking accounts. Third, direct deposits 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 type of 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" or 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.